

Örebro Universitet  
Institutionen för Ekonomi, Statistik och Informatik  
Statistik C, C-uppsats 10 poäng  
Handledare, Örebro Universitet: Thomas Laitila  
Handledare, Statistiska Centralbyrån: Lennart Nordberg  
HT-06, 070108

**Utvärdering av granskningsystem för SCB:s  
undersökningar Kortperiodisk  
Sysselsättningsstatistik och  
Konjunkturstatistik över Vakanser**

Författare:  
Chandra Adolfsson 801124

## Sammanfattning

I denna studie har undersökningarna Kortperiodisk Sysselsättningsstatistik (KS) och Konjunkturstatistik över Vakansers (KV) befintliga granskningssystem utvärderats med avseende på hur effektivt det är. Processdata har framställts och analyserats. Resultaten tyder på att många av de inkomna blanketterna med misstänkt felaktiga uppgifter inte rättas upp, utan tvingas igenom trots att granskningssystemet ej accepterade uppgifterna. Det befintliga granskningssystemet har en högre träffsäkerhet avseende KS-undersökningen, men både KS och KV skulle kunna granskas mer effektivt.

För att utvärdera det befintliga granskningssystemet ytterligare användes en poängfunktion. Till studien fanns tillgång till både helt ogranskat material och helt granskat material och dessa material användes i poängfunktionen. Det uppräknade ogranskade värdet för varje objekt jämfördes med det uppräknade granskade värdet och ställdes i relation till respektive skattade branschtotale. De poängsatta blanketterna rangordnades sedan. Därefter analyserades materialet för att försöka finna var det skulle vara lämpligt att sätta det tröskelvärde som skulle skilja det material som "egentligen" skulle ha behövts granskas från det som kunde ha lämnats orört. Att sätta tröskelvärdet är svårt. Här gjordes det godtyckligt utifrån kriterierna att det fel som införs i skattningarna för att allt material inte granskas skulle hållas så lågt som möjligt samt att antalet blanketter som skulle behöva granskas manuellt av produktionsgruppen också skulle hållas så lågt som möjligt. Även här visade det sig att det befintliga granskningssystemet inte är så effektivt som önskas. När resultaten från denna del av utvärderingen analyserades upptäcktes problem som beror på blankettutformningen. Skulle blanketterna ses över och åtgärdas skulle det fel som införs för att allt material inte granskas kunna minskas avsevärt. Genom att minska det införda felet kan tröskelvärdet förmodligen sättas på en ny nivå vilket medför att omfattningen av granskningen skulle minska ytterligare.

Hur skulle då ett mer effektivt granskningssystem kunna se ut? I den här studien har valet fallit på att testa "*significance editing*" på KS-undersökningen, det som på svenska kallas för *effektgranskning*. En poängfunktion användes även här, denna tilldelar de inkomna blanketterna varsin poäng och dessa poäng rangordnas därefter. Efter att poängen rangordnats bestäms en gräns, ett tröskelvärde, och de blanketter med en poäng som överstiger tröskelvärdet granskas och rättas upp av produktionsgruppen. De blanketter med en poäng som understiger det satta tröskelvärdet rättas inte upp, utan behåller sina originalvärden. Poängfunktionen jämför det inkomna ogranskade, uppräknade, värdet med ett uppräknat "förväntat" värde och ställer denna differens i relation till den skattade branschtotalen. Svårigheten ligger ofta i att hitta ett bra förväntat värde och detta problem uppstår ideligen i urvalsundersökningar. Tanken med effektgranskning är att omfattningen av granskningen ska minska och den granskning som utförs ska ha effekt på slutresultatet.

Det var inte lätt att hitta ett bra förväntat värde på den tid som stod till förfogande. Två problem som snabbt upptäcktes var dels att i KS-undersökningen finns inte uträknade säsong- eller trendfaktorer per variabel. Dessutom byttes en mycket stor del av urvalet ut till kvartal 2 (som denna studie har avgränsats till att behandla). Detta har fått till följd att cirka hälften av objekten i urvalet inte går att följa bakåt i tiden eftersom de inte ingått i urvalet tidigare. I studien har respektive stratumets medelvärde använts som förväntat värde. Resultaten visar att det valda förväntade värdet inte skulle ha använts i praktiken, men det fungerar bra i syfte att illustrera hur det i praktiken skulle kunna gå till att införa en mer effektiv granskning.

# Innehållsförteckning

<b>1</b>	<b>INLEDNING .....</b>	<b>1</b>
1.1	SYFTE.....	2
1.2	AVGRÄNSNING.....	2
<b>2</b>	<b>BESKRIVNING AV UNDERSÖKNINGARNA KS &amp; KV, PRODUKTIONSPROCESSEN &amp; DET BEFINTLIGA GRANSKNINGSSYSTEMET.....</b>	<b>2</b>
2.1	GEMENSAMT FÖR KS OCH KV AVSEENDE PRIVAT SEKTOR .....	2
2.2	KORTPERIODISK SYSSELSÄTTNINGSTATISTIK FÖR PRIVAT SEKTOR .....	4
2.2.1	<i>Företagare</i> .....	4
2.2.2	<i>Anställda</i> .....	4
2.2.3	<i>Frånvarande</i> .....	4
2.2.4	<i>Personalomsättning</i> .....	5
2.3	KONJUNKTURSTATISTIK ÖVER VAKANSER I PRIVAT SEKTOR .....	5
2.3.1	<i>Lediga jobb</i> .....	5
2.3.2	<i>Vakanser</i> .....	6
2.3.3	<i>Rekryteringsgrad, vakansgrad samt genomsnittlig rekryteringstid [7]</i> .....	6
2.4	PRODUKTIONSPROCESSEN .....	6
2.5	KS/KV-UNDERSÖKNINGARNAS GRANSKNINGSSYSTEM .....	9
2.5.1	<i>Processdata</i> .....	10
<b>3</b>	<b>GRANSKNING – NÅGRA METODASPEKTER.....</b>	<b>15</b>
3.1	GRANSKNINGENS SYFTE OCH ROLL .....	15
3.2	OLIKA KATEGORIER AV FEL .....	15
3.3	HIDIROGLOU-BERTHELOTS METOD .....	15
3.4	EFFEKTGRANSKNING MED HJÄLP AV POÅNGFUNKTIONER .....	16
3.4.1	<i>Pseudo bias (PB)</i> .....	18
3.4.2	<i>Förslag till granskningsstrategi</i> .....	18
<b>4</b>	<b>RESULTAT AV UTVÄRDERINGEN AV KS/KV:S BEFINTLIGA GRANSKNINGSSYSTEM.....</b>	<b>20</b>
4.1	UTVÄRDERING AV HB-METODEN.....	20
4.2	JÄMFÖRELSE MELLAN KS-OGRANSKADE DATA OCH KS-GRANSKADE DATA.....	20
4.2.1	<i>Analys av variabeln tillsvidareanställda män</i> .....	27
4.3	JÄMFÖRELSE MELLAN KV-OGRANSKADE DATA OCH KV-GRANSKADE DATA .....	28
<b>5</b>	<b>RESULTAT AV TESTET DÄR RESPEKTIVE STRATUMMEDELVÄRDE ANVÄNTS SOM FÖRVÄNTAT VÄRDE .....</b>	<b>31</b>
<b>6</b>	<b>DISKUSSION OCH SLUTSATS .....</b>	<b>35</b>
6.1	DELSYFTE 1 - UTVÄRDERINGEN.....	36
6.2	DELSYFTE 2 – DET ALTERNATIVA GRANSKNINGSSYSTEMET OCH PROBLEMATIKEN KRING DET FÖRVÄNTADE VÄRDET .....	38
	<b>REFERENSER .....</b>	<b>40</b>
	<b>APPENDIX 1 - BLANKETTUTSEENDEN.....</b>	<b>42</b>
	<b>APPENDIX 2 – TABELLER ÖVER DE KODER SOM ANVÄNDS I KS/KV .....</b>	<b>45</b>
	<b>APPENDIX 3 – FLÖDESSCHEMA ÖVER PRODUKTIONSPROCESSEN .....</b>	<b>47</b>
	<b>APPENDIX 4 – YTTERLIGARE PROCESSDATARESULTAT .....</b>	<b>48</b>
	<b>APPENDIX 5 - YTTERLIGARE UTVÄRDERINGSRESULTAT AV KS .....</b>	<b>49</b>
	<b>APPENDIX 6 – YTTERLIGARE UTVÄRDERINGSRESULTAT AV KV .....</b>	<b>64</b>
	<b>APPENDIX 7 – YTTERLIGARE RESULTAT AV DET GENOMFÖRDA TESTET .....</b>	<b>65</b>
	<b>APPENDIX 8 – NYCKLAR TILL NÄRINGSRENSINDELNING I KS OCH KV .....</b>	<b>70</b>

## 1 Inledning

Sedan år 1858 har Statistiska Centralbyrån (SCB) försett samhället med statistik som underlag för beslutsfattande, debatt och forskning. SCB producerar en mångfald av statistik, drygt 200 statistikprodukter tas fram. För att kunna framställa dessa statistikprodukter krävs insamling och bearbetning av data. De insamlade data lagras i databaser och för att kunna bearbeta och granska data krävs IT-system. De allra flesta statistikprodukter har sin egen databas och sitt egna IT-system. Detta innebär att SCB för närvarande har cirka 200 olika IT-system som behöver underhållas vilket är en orsak till produkternas höga framställningskostnader.

Under år 2004 genomförde SCB en intern omorganisation som ett första steg för att komma till rätta med detta. Omorganisationen bestod i att försöka samla så mycket av processerna insamling och granskning för företag på en enhet på SCB. Istället för att som tidigare ha produkterna och deras insamlings- och granskningsprocesser utspridda på respektive ämnesenhet. Avdelningen för Datainsamling från Företag och Organisationer och enheten för Företagens Uppgiftslämnande (DFO/FU) skapades. På DFO/FU finns i dagsläget ett fyrtiotal olika produkter och nästan lika många olika IT-system.

För att höja statistikens kvalitet och minska produktionskostnaderna togs ett andra steg i augusti år 2006. Ett arbete med att standardisera processer, verktyg och metoder påbörjades. Arbetet utförs i projektform under namnet Lotta-projektet och målet är att projektet ska vara klart i september 2007, med vissa justeringar fram till årsskiftet 2007/2008.

Lotta-projektet består av tre paraplyprojekt, Standardiserade processer och gemensamma verktyg, System för kvalitetssäkring, kvalitetskontroll och utvärdering samt Kompetensförsörjning och styrningsfrågor. Under paraplyprojekten kommer ett stort antal delprojekt att bedrivas. En del av arbetet som ligger under paraplyprojekt 2, det vill säga under System för kvalitetssäkring, kvalitetskontroll och utvärdering, är att se över produktionsprocesserna på SCB. Detta kommer att göras bland annat genom att utföra fallstudier på utvalda produkter. Fallstudierna kommer att fokusera på den produktionsgranskning som utförs. Respektive fallstudie ska bland annat leverera en slutsats om det är lämpligt att införa selektiv granskning. Förslag till vilken/vilka poängfunktioner, baserat på misstanke och potentiell effekt, ska också ges.

Två tätt sammanlänkade produkter som samlas in och granskas på DFO/FU på uppdrag av avdelningen för Näringsliv & Arbetsmarknad (NA) är den Kortperiodiska Sysselsättningsstatistiken (KS) och Konjunkturstatistik över Vakanser (KV). Ansvarig ämnesenhet för produkterna KS och KV på avdelningen NA är enheten för Företags- och Registerbaserad Sysselsättning (NA/FRS).

En av de fallstudier som ska utföras kommer att omfatta dessa två statistikprodukter och denna uppsats kommer att utgöra delar av fallstudien.

I effektmålet för fallstudierna ingår att de efter genomförande bör kunna ge omedelbara effektiviseringar av den granskning som utförs på någon eller flera av de utvalda produkterna. Detta bör på sikt kunna realiseras i kostnadsbesparingar. Ett mer effektivt granskningssystem för KS/KV-undersökningarna som är enkelt att uppdatera och inte alltför svårt att lära sig hantera är önskvärt. Att minska arbetstiden som idag läggs på produktionsgranskning liksom att minska framställningskostnaderna är också ett mål.

## 1.1 Syfte

Det föreligger en misstanke om att en väsentlig del av den produktionsgranskning som utförs i undersökningarna Kortperiodisk Sysselsättningsstatistik (KS) och Konjunkturstatistik över Vakanser (KV) inte har så stor effekt på slutresultatet. Syftet med denna uppsats är dels att utvärdera det befintliga granskningssystemet med avseende på hur effektivt det är. Utvärderingen kommer att ske genom att ta fram processdata och jämföra ogranskade data med granskade data med hjälp av en poängfunktion.

Uppsatsen syftar också till att illustrera hur ett alternativt effektivt granskningssystem där man använder sig av effektgranskning skulle kunna se ut samt testa en variant av en poängfunktion på det datamaterial som står till förfogande.

## 1.2 Avgränsning

Denna uppsats avgränsas till att omfatta den del av granskningen som avser privat sektor. Till privat sektor förs enskilda firmor, enkla bolag, handelsbolag, kommanditbolag, samtliga aktiebolag, ekonomiska och ideella föreningar samt offentligt ägda aktiebolag, affärsverken och Riksbanken. Branschen transport, magasinering och kommunikationer (SNI I) redovisas exklusive ombordsanställda i utrikes- och kustsjöfart registrerade i Sverige.

Med granskning menas här identifiering och åtgärdande av fel och outliers i individuella data som används för framställningen av undersökningarna Kortperiodisk Sysselsättningsstatistik (KS) och Konjunkturstatistik över Vakanser (KV). Granskningen som utförs inom privat sektor omfattar tre olika delar; granskning av sysselsättning, frånvaro, personalomsättning, granskning av sjuklöneavviseringar och granskning av lediga jobb. Denna uppsats kommer inte att behandla granskningen av sjuklöneavviseringarna. Alla data som ligger till grund för uppsatsen avser inkomna uppgifter avseende kvartal 2 (april, maj och juni) år 2006.

Utvärderingen av det befintliga granskningssystemet kommer att utföras på både undersökningen Kortperiodisk Sysselsättningsstatistik och på Konjunkturstatistik över Vakanser. Testet där respektive stratummedelvärde används som förväntat värde kommer endast att utföras på Kortperiodisk Sysselsättningsstatistik.

## 2 Beskrivning av undersökningarna KS & KV, produktionsprocessen & det befintliga granskningssystemet

### 2.1 Gemensamt för KS och KV avseende privat sektor

Målobjekt för KS/KV avseende privat sektor är arbetsställe. Arbetsställe definieras som varje adress, fastighet eller grupp av närliggande fastigheter där företaget bedriver verksamhet. Alla företag har minst ett arbetsställe. Målpopulationen utgörs av de arbetsställen i privat sektor som har minst en anställd enligt SCB:s Företagsdatabas (FDB) [1].

Urvalsramen skapas ur FDB och består av arbetsställen/enheter som där är registrerade som aktiva och har minst en anställd. Undersökningarna ingår i SAMU-systemet [2] (SAMU står för "Systemet för Samordnade Urval") och den urvalsmetod som används ger ett obundet slumpmässigt urval (OSU) av arbetsställen inom strata. Urvalsramen stratifieras efter variablerna storleksklass och branschtillhörighet. Beroende på

strukturen inom de olika strata totalundersöks vissa strata medan ett urval av arbetsställen dras från de resterande strata.

Utvalda arbetsställen byts successivt ut enligt en given metod (JALES-metoden [2]) som nyttjas för att minska företagens uppgiftslämnarbörda. Nytt urval dras två gånger per år. Dels används mars-SAMU vilket innebär ett nytt urval från och med mät månaden april och dels används augusti-SAMU som innebär ett nytt urval från och med mät månaden oktober. Vid nytt urval byts cirka 25 procent av samtliga arbetsställen ut.

Undersökningarna Kortperiodisk Sysselsättningsstatistik [3] och Konjunkturstatistik över Vakanser [4] nyttjar samma urval. Det samordnade urvalet medför att SCB har uppgifter om exempelvis antalet lediga jobb, anställda och nyanställda på varje arbetsställe i urvalet. Detta ger SCB möjlighet att publicera rekrytringsgrad, vakansgrad och rekryteringstid. Samordningen gör också att SCB kan uppnå konsistens mellan sysselsättning, personalomsättning och vakanser. Ytterligare en samordningsvinst är att det oftast är samma kontaktperson på båda undersökningarna ute på företagen och på SCB har samma person i produktionsgruppen hand om både KS och KV avseende samma arbetsställe. Detta bidrar till att minska bördan för uppgiftslämnarna då de kan vända sig till en och samma person om fler undersökningar. Antalet kuvert från SCB minskar då båda undersökningarna skickas i samma kuvert om de ska besvaras av samma uppgiftslämnare.

Skattningarna redovisas för olika gruppindelningar. Dessa definieras av arbetsställets näringsgren (för närvarande används SNI 2002 [5]), storleksklass och region (NUTS 3 [6] på KS, NUTS 2 [6] på KV) samt av företagets sektortillhörighet. Redovisningen av Kortperiodisk Sysselsättningsstatistik sker även efter kön.

För de utvalda arbetsställena samlas uppgifter in om de variabler som ingår i respektive undersökning under en angiven arbetsdag i mät månaden. Den angivna arbetsdagen är alltid den mittersta onsdagen i mät månaden. Anledningarna till valet av mät dag är flera. En av dessa anledningar är jämförbarheten, en annan är att SCB vill underlätta för uppgiftslämnarna genom att låta dem veta i god tid vilka de aktuella mät dagarna är så att de kan lägga in dessa i sina lönesystem (från vilka de tar ut de efterfrågade uppgifterna). Ytterligare en anledning är att vissa variabler påverkas av veckodag. Exempelvis är det känt från statistik över den korta sjukfrånvaron som delvis också samlas in via KS-undersökningen att sjukfrånvaron är högst på måndagar för att sedan falla under resten av veckodagarna. Det medför att onsdagen förmodligen speglar den genomsnittliga sjukfrånvaron bäst.

Målsättningen är att svarsbortfallet ska ligga på max 8 procent för respektive undersökning vid publicering. För att minska bortfallsfelet på KS görs vissa imputeringar. Exempelvis imputeras frånvarovariablerna om ett arbetsställe inte kan redovisa dessa, men har redovisat övriga variabler. Mer om frånvaroimputeringen står att läsa i avsnitt 2.5.1. För att minska bortfallet ytterligare används utskick av så kallade tack- och påminnelsekort, påminnelsebrev och telefonpåminnelser. Rak uppräknings används för bortfallskompensation. Rak uppräknings bygger på antagandet att bortfallet kan betraktas som helt slumpmässigt. Om detta inte stämmer finns det risk för att skattningarna innehåller skevheter vilket i sin tur bland annat medför att intervallskattningarnas konfidensgrad minskar. Publicering av undersökningarna sker cirka sex veckor efter mät kvartalets utgång.

Användare av Kortperiodisk Sysselsättningsstatistik och Konjunkturstatistik över Vakanser är Närings- och Finansdepartementet, Arbetsmarknadsstyrelsen (AMS), Medlingsinstitutet, Riksbanken, Konjunkturinstitutet och Nationalräkenskaperna (MP/NR, SCB). Även universitet och fristående forskare använder statistiken.

## 2.2 Kortperiodisk Sysselsättningsstatistik för privat sektor

Undersökningen Kortperiodisk Sysselsättningsstatistik har samlats in och publicerats för privat sektor sedan år 1989. Det föreligger uppgiftslämnarplikt på undersökningen. Syftet med undersökningen är främst att med hög precision snabbt indikera förändringar av antalet anställda på detaljerad näringsgrensnivå. Ett annat syfte är att presentera sysselsättningsuppgifter över hela arbetsmarknaden fördelat på län.

Urvalsstorleken kvartal 2 år 2006 var cirka 18 600 arbetsställen i privat sektor varav cirka 2 400 totalundersöktes.

Variablerna som undersöks är de som följer tillsammans med definitioner och förklaringar:

### 2.2.1 Företagare

- *Antalet verksamma företagare* vid mättillfället avser verksamma företagare/delägare i enskild firma, enkelt bolag, handelsbolag och kommanditbolag samt eventuella medhjälpande familjemedlemmar som arbetar vid arbetsstället utan att vara anställda.

Uppgifter om antal verksamma företagare samlas in per kön.

### 2.2.2 Anställda

- *Antalet anställda totalt* vid mättillfället avser personer (inklusive deltids- och visstidsanställda) som var anställda på mät dagen. Som anställda räknas även personer som vid mättillfället inte var i tjänst på grund av arbetstidens förläggning, sjukdom, semester, tjänstledighet eller dylikt. Verksamma aktiebolagsägare är inkluderade. En anställd som har anställning vid flera olika arbetsställen räknas flera gånger. Om en anställd däremot har flera tjänster inom samma arbetsställe räknas denne bara en gång.
- *Antalet tillsvidareanställda* som vid mättillfället hade en formell tillsvidareanställning. Verksamma aktiebolagsägare och delägare i ekonomiska föreningar förs hit.
- *Antalet visstidsanställda* vid mättillfället avser personer (såväl heltids- som deltidsanställda) som anställts för en viss period eller ett visst uppdrag. Timavlönade räknas som visstidsanställda om de vid lönekörningen omfattas av någon lönetransaktion. Undantagna är exempelvis personer i arbetspraktik.

Uppgifter om antal tillsvidare- och visstidsanställda samlas in per kön.

### 2.2.3 Frånvarande

Frånvaron redovisas efter tre frånvaroorsaker:

- a) *Totalt antal heldagsfrånvarande* vid mättillfället dock ska inte personer med schemalagd ledighet ingå
- b) *Sjukdom*
- c) *Semester*
- d) *Övrig frånvaro*, denna kan vara både semesterlönegrundande och inte semesterlönegrundande. Hit räknas ledighet med havandeskaps- eller föräldrapenning, ledighet för smittbärare, ledighet för viss utbildning enligt studiestödslagen, facklig utbildning, militär grundutbildning, militär repetitionsutbildning, civilförsvarsutbildning, permittering, semesterpermittering, kompensationsledighet och arbetskonflikt.

Uppgifter om frånvaron samlas in per kön.

#### 2.2.4 Personalomsättning

Personalomsättningsmättet utgör medeltalet av uppgifterna för de månader som ingår i kvartalet. Personer som har haft flera tillfälliga anställningar under månaden räknas endast en gång.

- *Nyanställda*, personer som har anställts vid arbetsstället under mätmånaden. Här samlas *totalt antal nyanställda* in, liksom de personer som *nyanställts* som *tillsvidare-* eller *visstidsanställda* uppdelat per kön.
- *Avgångna*, personer som har lämnat arbetsstället efter uppsägning eller på egen begäran under mätmånaden. Här samlas *totalt antal avgångna* in, liksom de personer som *avgått* och var *tillsvidare-* eller *visstidsanställda* uppdelat per kön.

KS-blanketten finns i flera olika blankettvarianter, i appendix 1 visas de två varianter som är aktuella för denna studie.

Kortperiodisk Sysselsättningsstatistik används till att beräkna Bruttonationalprodukten (BNP), göra konjunkturbedömningar såväl i Sverige som i Europa, utföra analyser av sjukfrånvaron i samhället och av arbetsmarknadens funktionssätt avseende personalomsättning samt för att göra landsjämförelser på de nivåer som ofta efterfrågas. Statistiken används också av EU för att beskriva och analysera regioners ekonomiska utveckling.

### 2.3 Konjunkturstatistik över Vakanser i privat sektor

Konjunkturstatistik över Vakanser introducerades i privat sektor under tredje kvartalet år 2000. Uppgiftslämnarplikt infördes som en följd av högt bortfall den första juli år 2003. Syftet med undersökningen är att med hög precision beskriva efterfrågan på arbetskraft. Uppgifter presenteras om lediga jobb och vakanser för hela arbetsmarknaden fördelat på region och storleksklass. Urvalsstorleken kvartal 2 år 2006 var cirka 16 600 arbetsställen i privat sektor varav cirka 2 400 totalundersöktes.

Variablerna som undersöks är de som följer tillsammans med definitioner och förklaringar:

#### 2.3.1 Lediga jobb

- *Antalet Lediga jobb totalt* vid mättillfället avser tjänster till vilka arbetsgivaren har påbörjat extern rekrytering av en medarbetare men ännu inte tillsatt någon. Med extern rekrytering avses rekrytering som inte är begränsad till företaget, det lediga jobbet ska alltså vara tillgängligt för personer utanför organisationen. De lediga jobben delas i två grupper dels bemannade lediga jobb och dels obemannade lediga jobb.



De bemannade lediga jobben kan fördelas på följande kategorier:

- a) ledigt jobb där avgående medarbetare bemannar det lediga jobbet, detta kan ske vid exempelvis ersättningsrekrytering. Här ska också lediga jobb som kommer att vara bemannade av semestervikarier och av ersättare för anställda som beviljats friår.
- b) ledigt jobb som tillfälligt bemannas av vikarie, behovsanställd eller annan tillfälligt anställd.
- c) ledigt jobb som tillfälligt bemannas av konsult eller personal från bemanningsföretag. Här avses ej personal från bemanningsföretag som permanent sköter vissa arbetsuppgifter, exempelvis vid outsourcing.

De obemannade lediga jobben kan fördelas på följande kategorier:

- a) ledigt jobb som är obemannat och där bemanning önskades på mätdagen.
- b) ledigt jobb som är obemannat och där bemanning önskades efter mätdagen. Under denna kategori lediga jobb faller rekryteringar till exempelvis framtida projekt eller för vilka arbetsplats inte är klar ännu.

### **2.3.2 Vakanser**

Vakanser avser obemannade lediga jobb som kan tillträdas omedelbart, detta motsvaras av den ovanstående kategorin ”ledigt jobb som är obemannat och där bemanning önskades på mätdagen”.

### **2.3.3 Rekryteringsgrad, vakansgrad samt genomsnittlig rekryteringstid [7]**

Rekryteringsgrad definieras som andelen lediga jobb av antalet anställda i respektive redovisningsgrupp. Detta är ett relativt mått där antalet lediga jobb normaliserats för antalet anställda inom redovisningsgruppen. Ju högre rekryteringsgrad desto större är rekryteringsaktiviteten. Skälet till denna redovisning är att antalet lediga jobb tenderar att öka med antalet anställda inom redovisningsgruppen.

Vakansgraden definieras som andelen obemannade lediga jobb av antalet anställda i respektive redovisningsgrupp. Vakansgraden visar relativ brist på arbetskraft. Ju högre vakansgrad desto högre är den relativa bristen på arbetskraft.

Genomsnittlig rekryteringstid definieras som antalet lediga jobb relaterade till antalet nyanställningar. Detta mått kan tas fram då KS mäter antalet nyanställningar på de arbetsställen som ingår i KV. Antal lediga jobb utgörs av antalet lediga jobb en specifik mätdag i månaden. Antal nyanställningar utgörs av antalet personer som anställts under hela månaden. Kvartalsresultat estimeras som ett kvartalsgenomsnitt för de månader som ingår i kvartalet. Genomsnittlig rekryteringstid kan tolkas som ett mått på matchningseffektivitet på arbetsmarknaden.

KV-blankettens utseende visas i appendix 1.

## **2.4 Produktionsprocessen**

Produktionsprocessen startar med att ramar skapas och allokering utförs så att urvalet till Kortperiodisk Sysselsättningsstatistik och Konjunkturstatistik över Vakanser för privat sektor sedan kan dras. En del av urvalet totalundersöks, detta gäller främst stora arbetsställen (med > 99 anställda) som är viktiga för slutresultatet. Den totalundersökta delen undersöks varje månad. Den del av urvalet som inte totalundersöks fördelas slumpvis i tre ungefär lika stora delar inom varje stratum där varje tredjedel svarar för en

månad vardera i kvartalet. Genom att tredela urvalet speglas kvartalet bättre än om endast en månad i kvartalet undersöks.

När urvalet väl är draget bearbetas det av produktionsgruppen. Produktionsgruppen består i dagsläget av åtta personer som tillhör enheten DFO/FU på SCB. Bearbetningen av urvalet består till stor del av att hitta kontaktpersoner till undersökningarna, utföra adressändringar samt uppdatera koder i applikationen som produktionsgruppen använder sig av i det dagliga arbetet med undersökningarna. En annan del av bearbetningen består i att fördela vissa delar av urvalet per person som ingår i produktionsgruppen. Samtliga totalundersökta arbetsställen fördelas samt arbetsställen som tillhör organisationsnummer med minst ett totalundersökt arbetsställe. Detta kommer sig av önskemål både från uppgiftslämnarna, ämnesenheten, produktionsgruppen och från Storföretagsgruppen på SCB. Storföretagsgruppen på SCB tillhör också DFO/FU och arbetar med de 50 största koncernerna/företagen i Sverige. De agerar kontaktyta gentemot företagen och SCB och utför löpande registervård samt viss samgranskning. Samtliga företag som Storföretagsgruppen arbetar med ingår i KS:s och KV:s urval. Arbetsställen som inte är totalundersökta och inte tillhör ett organisationsnummer med totalundersökta arbetsställen fördelas ej. Produktionsgruppen plockar åt sig av dessa arbetsställen om de inkommer med felaktiga uppgifter och/eller behöver påminnas.

När urvalet har bearbetats skickas de blanketter ut som tillhör det nya urvalets första mätmånad. Utskick av blanketter sker i slutet av mät månaden och sedan har uppgiftslämnaren cirka en månad på sig att fylla i och skicka in blanketten. Som ett alternativ till den pappersblankett som skickas ut finns det en elektronisk blankett för redovisning via Internet. Uppgiftslämnaren har också möjlighet att faxa, maila, skicka in fil, använda FTP-överföring eller lämna in uppgifterna via telefon.

Uppgifter som inkommer via post tas varje veckodag om hand av enkätenheten, DIH/ENK, på SCB. Personalen där öppnar kuverten, förgranskar de inkomna blanketterna och lägger de blanketter som innehåller en ändring av adress- och/eller kontaktpersonsuppgifter och/eller text utanför läsområdet överst i bunten. Därefter skannar de blanketterna, verifierar och överför uppgifterna till KS/KV:s databas. Att verifiera innebär att de siffror som skanningsprogrammet inte kan tolka får tolkas manuellt. Alla tecken massverifieras också och detta innebär att man får se exempelvis alla skannade ettor ihop för att se så att alla tecken som tolkats som ettor verkligen är ettor (snedstreck råkar ofta tolkas som ettor). När enkätenheten har skannat, tolkat och verifierat materialet överförs det till KS/KV:s databas. Där läggs materialet i en tabell i väntan på att någon ur produktionsgruppen ska ladda in uppgifterna i applikationens ordinarie tabeller. Applikationen används främst av produktionsgruppen och benämns härnäst KS/KV-systemet.

När uppgifterna kommer in i KS/KV-systemet granskas de maskinellt. De uppgifter som av KS/KV-systemet anses orimliga, felaktiga eller åtminstone tveksamma faller ut som fel och dessa tas om hand och rättas upp av produktionsgruppen. Denna typ av granskning kallas här för mikrogranskning. Ibland är det möjligt att rätta upp en felaktig uppgift genom att titta på den skannade bilden (skanningsfel förekommer) eller genom att titta på föregående inlämnade uppgifter, i annat fall kontaktas uppgiftslämnaren. Hur det befintliga granskningssystemet är uppbyggt beskrivs i avsnitt 2.5.

Uppgiftslämnarna har, som nämnts ovan, möjlighet att lämna in de efterfrågade uppgifterna via fler media än via pappersblankett. De kan välja att lämna in KS-uppgifterna via fil och i så fall lämnar de in uppgifter för flera arbetsställen och ofta även för flera organisationsnummer på samma fil. De kan även inkludera uppgifter i filen som avser arbetsställen som inte ingår i urvalet, dessa sorteras automatiskt bort och

berör aldrig produktionsgruppen. Uppgifter som lämnas via fil lämnas antingen via e-post eller FTP-överföring. Väljer uppgiftslämnaren att faxa in pappersblanketten går den inte att skanna och detta medför att produktionsgruppen får registrera uppgifterna manuellt. Uppgifter som lämnas in via telefon registreras också manuellt i KS/KV-systemet av den person i produktionsgruppen som talar med uppgiftslämnaren.

Enkätenheten fyller boxar med de pappersblanketter de skannar och dessa tar produktionsgruppen hand om. Då enkätenheten lägger alla blanketter som innehåller en ändring av adress- och/eller kontaktpersonsuppgifter och/eller text utanför läsområdet överst i bunten är det enkelt för produktionsgruppen att hitta de blanketter som behöver uppdateras.

Produktionsgruppens huvudsakliga arbete består av att rätta upp inkomna felaktiga uppgifter, utföra adress- och kontaktpersonsuppdateringar och att utföra påminnelsearbete. Påminnelsearbetet utförs främst via telefon, men även via fax och e-post.

Cirka tio kalenderdagar innan sista insändningsdag skickas ett tack- och påminnelsekort ut till en del av urvalet och ett par dagar efter sista insändningsdag mottar de uppgiftslämnare som inte fått något tack- och påminnelsekort och som ännu inte skickat in de efterfrågade uppgifterna ett påminnelsekort eller ett påminnelseutskick. Påminnelseutskicket innehåller ett påminnelsebrev samt ny blankett.

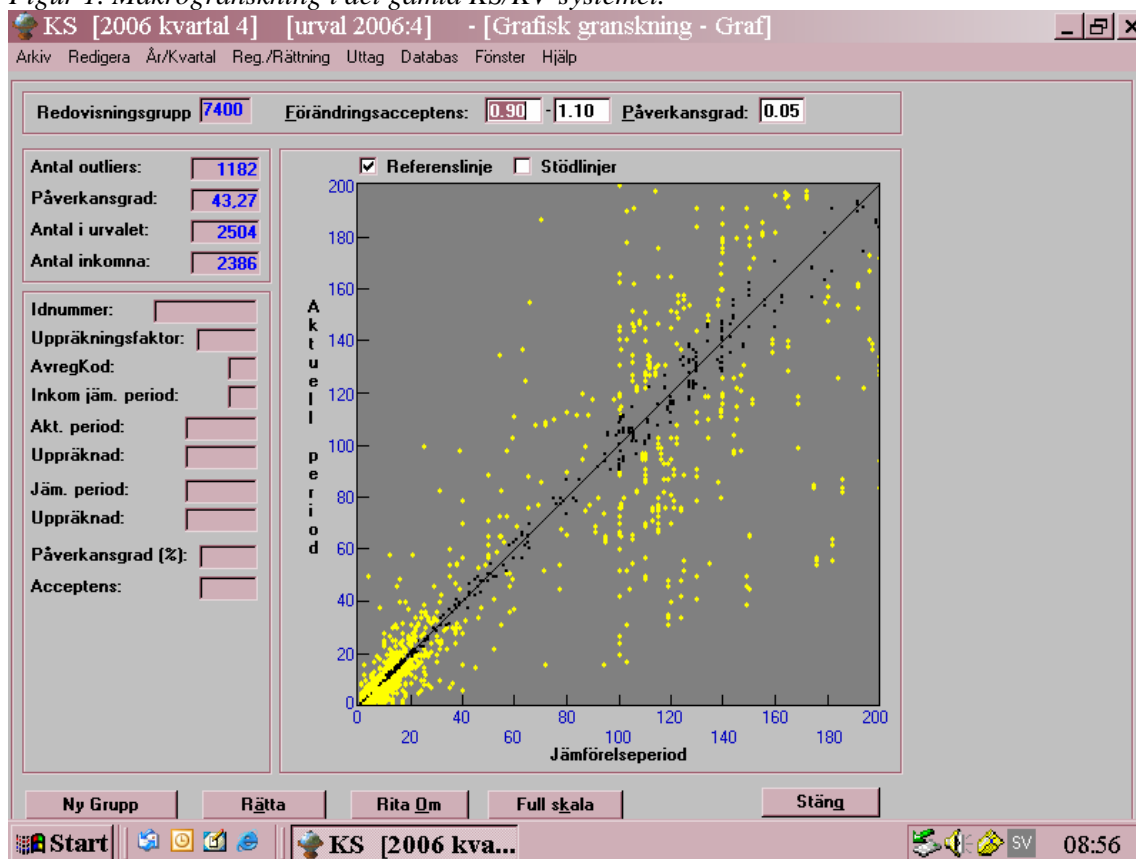
Arbetsställen som trots tack- och påminnelsekort, påminnelsekort eller påminnelseutskick inte lämnar in uppgifterna blir uppringda i mån av tid av antingen produktionsgruppen eller av en intervjuare från intervjuenheten, DIH/CATI, på SCB. I samtalet blir uppgiftslämnaren upplyst om vikten av att delta i undersökningarna, då båda undersökningarna är urvalsundersökningar försämrar varje uteblivet svar statistikens kvalitet. Uppgiftslämnaren ges också möjlighet att lämna uppgifterna direkt via telefon och slipper då fylla i pappersblanketten.

Varje månadsomgång avslutas med en så kallad månadsslutkörning. Under denna försöker produktionsgruppen få ned bortfallet genom att rätta upp inkomna uppgifter som av KS/KV-systemet bedömts som felaktiga och ta in så mycket uppgifter som möjligt via telefon och fax. Vid en månadsslutkörning är målet att nå ett bortfall på max 10 procent på respektive undersökning.

Efter att en månadsomgång är avslutad makrogranskas materialet. Tidigare makrogranskade produktionsgruppen månad 1 och 2 i kvartalet och de produktansvariga på ämnesenheten månad 3 samt hela kvartalet. Anledningen till att hela kvartalet ses över trots att respektive månad redan blivit granskad efter avslutad månadsomgång är att uppgifter kommer in och tas emot även efter avslutad månadsomgång. Det nya KS/KV-systemet infördes i februari 2006 och i detta finns inte makrogranskningen inbyggd ännu. Därför makrogranskar de produktansvariga samtliga omgångar åtminstone under år 2006. Makrogranskningen utförs endast på variabeln *totalt antal anställda*. Makrogranskningen är en så kallad grafisk granskning där varje branschgrupp (43st) granskas för sig. Målsättningen är att hitta de avvikelser som har betydelse för slutresultatet.

Figur 1 visar hur makrogranskningen såg ut i det gamla KS/KV-systemet.

Figur 1. Makrogranskning i det gamla KS/KV-systemet.



Varje arbetsställe representeras av en ”prick” i den grafiska granskningen. De arbetsställen som ligger längst från linjen granskas närmare genom att produktionsgruppen får reda ut vad avvikelserna beror på. Oftast behöver uppgiftslämnaren kontaktas och nya uppgifter måste begäras. Ett vanligt misstag som görs av uppgiftslämnaren är att lämna uppgifter för hela företaget trots att undersökningen är en så kallad arbetsställeundersökning vilket leder till att antalet anställda överskattas. Andra fel som upptäcks i makrogranskningen kan vara skanningsfel eller att uppgiftslämnaren missat att medräkna en viss personalgrupp, exempelvis tjänstemännen.

I appendix 3 finns ett flödesschema där man kan följa varje process från första steget i urvalsprocessen fram till publicering och uppdragsprocessen.

## 2.5 KS/KV-undersökningarnas granskningssystem

Inkomna uppgifter granskas kontinuerligt av produktionsgruppen. När en rad läggs till eller ändras i de tre tabeller som innehåller blankettuppgifter körs ett granskningsprogram som ligger i databasen igång automatiskt. En fördel med detta är att det inte spelar någon roll vilket media som används för att registrera uppgifter för så fort en uppgift ändras så granskas den automatiskt. Detta granskningsprogram kontrollerar uppgifterna utifrån redan satta granskningsvillkor. På de blanketter där någon granskningskontroll faller ut sätts styrkoden 02 som indikerar felaktiga/tveksamma uppgifter. På övriga blanketter sätts styrkoden 03 som betyder godkända uppgifter. Felkoden på de granskningskontroller som har fallit ut sparas i egna tabeller, en för respektive undersökning. En blankett kan således få flera olika felkoder om flera variabelvärden är felaktiga/tveksamma.

I KS-undersökningen är ett exempel på ett uppenbart fel när uppgiftslämnaren har fyllt i fler frånvarande än antal anställda personer. Ett misstänkt fel är när antalet anställda skiljer sig markant från föregående inlämnad uppgift eller från antalsuppgiften i FDB för det aktuella arbetsstället. Det senare nämnda felet granskas med en speciell metod, Hidiroglou-Berthelots metod (HB-metoden [8]), vilken beskrivs i kapitel 3.

I KV-undersökningen är ett exempel på ett uppenbart fel när totalt antal lediga jobb skiljer sig från summan av de olika kategorierna av lediga jobb. Ett misstänkt fel på en KV-blankett är när ett arbetsställe enligt antalsuppgiften i FDB har fler än 199 anställda men inte redovisar något ledigt jobb. Se appendix 2 för komplett lista över vilka styr- och felkoder som undersökningarna använder sig av. Av felkodslistan går att utläsa vilka kriterier och/eller gränser som ligger till grund för de felkoder som produktionsgruppen ansvarar för att ta om hand.

Medlemmarna i produktionsgruppen laddar varje dag uppgifter som inkommit via pappersblankett, fil, FTP-överföring eller via webben. Fax och uppgifter som inkommer via telefon registreras manuellt. Produktionsgruppen har möjlighet att varje dag välja menyvalet "Rätta dagens bunt" i KS/KV-systemet och kan på så vis enkelt komma åt och ta om hand dagens felaktiga blanketter. Produktionsgruppen har också möjlighet att ta ut listor med de arbetsställen som har minst en felkod att rätta upp.

Oftast behöver uppgiftslämnaren kontaktas om felkoder på den inlämnade blanketten har fallit ut, detta görs främst via telefon och e-post. I KS/KV-systemet har granskaren möjlighet att skriva en kommentar om vad orsaken till felet/felen var och sätta en markör som signalerar att granskaren tagit kontakt med uppgiftslämnaren via telefon. Markören möjliggör framtagning av processdata över hur många återkontakter via telefon som görs. Erfarenheten säger dock att denna markör inte alltid används av granskaren vilket medför att resultat av processdata på denna variabel ska tolkas med försiktighet. Antalet återkontakter via telefon underskattas rejält vid uttag av processdata baserat på denna markör. I dagsläget kan granskaren inte sätta någon markör för exempelvis kontakt via e-post utan får istället skriva det i kommentarrutan. Det befintliga systemet saknar även orsakskodning. Orsakskodning innebär att granskaren kodar de blanketter som fallit ut som fel, men som rättats upp, efter orsaken till felet. Orsakskodning ger möjlighet att sammanställa de orsaker till varför blanketter inte blir godkända av granskningssystemet. Vid analys av sammanställningen kan exempelvis strukturfel i blanketten upptäckas och åtgärdas.

### **2.5.1 Processdata**

Processdata innebär att ta fram statistik över den del i produktionsprocessen som avser granskning. Utan data om processen går det inte att ständigt förbättra respektive undersökning vilket är något SCB försöker att leva upp till. Tre av syftena med att ta fram processdata är att få underlag till förbättring av undersökningen, mäta effekter av processförändringar och att få underlag till analys och kvalitetsredovisningar. För att till fullo beskriva och utvärdera det befintliga granskningssystemet behövde processdata framställas. Arbetet började med en analys av datamaterialet. Då detta hade gjorts upptäcktes behovet av att göra vissa bearbetningar av materialet innan processdata kunde tas fram. Bearbetningar av datamaterialet, script och tabeller har gjorts i programvaran SAS (version 9.1). Programvaran Query Analyzer har använts till att göra vissa datakörningar mot databaserna som ligger i Microsoft SQL 2000, även Excel (Microsoft Office Excel 2003) har nyttjats. Alla data som utvärderingen grundas på är uppgifter avseende kvartal 2 (april, maj och juni) 2006. Materialet bestod från början av både privat- och offentlig sektor, men då utvärderingen endast ska utföras för privat sektor så har materialet rensats på blanketter tillhörande offentlig sektor. Även blanketter som inkom efter publiceringsdatumet har rensats bort.

Ett problem vid framtagningen av processdata och med att jämföra de ogranskade värdena med de granskade värdena var olika koder som används i produktionen för att signalera att något inte är som det ska, främst rörde det sig om en signaleringskod och en imputeringskod. Signaleringskoden ('8888') registreras av enkätenheten då de i skanningen stöter på svårtolkade värden, bokstäver istället för siffror eller decimaltal. Om en uppgiftslämnare har problem i sitt system och inte kan redovisa någon sjukfrånvaro har produktionsgruppen möjlighet att lägga en imputeringskod på variabeln *heldagsfrånvarande man p.g.a. sjukdom* ('9999') som signalerar att när skattningen tas fram ska frånvaron imputeras. Samtliga frånvarovariabler imputeras då.

Värdet efter imputering räknas fram för respektive frånvarovariabel av följande formel enligt samtal med Annika Norlén (NA/FRS, SCB, 2006):

$$Y_k^* = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n Z_i} * Z_k$$

där  $\sum_{i=1}^n Y_i$  = Antalet frånvarande i det aktuella stratumet,

$\sum_{i=1}^n Z_i$  = Antalet anställda i det aktuella stratumet,

$Z_k$  = Antalet anställda på det aktuella arbetsstället

Dessa koder ligger i det ogranskade materialet och gör att differenserna mellan ogranskade- och granskade värden blir orimligt stora ('8888' och '9999' räknas då som värden). Först behövde alla blanketter med sådana här koder tas om hand, i de fält där koderna fanns sattes värdena till noll istället.

När det ogranskade materialet skulle matchas ihop med det granskade stämde inte materialen helt överens. En del blanketter sänds in och kommer med i ogranskat-tabellen, men tas bort exempelvis på grund av övertäckning och finns därför inte med i granskat-tabellen. Dessa blanketter har tagits bort från materialet. Materialet har även rensats på data avseende offentlig sektor (kommuner, statliga myndigheter, landsting).

I KS-undersökningen följs personalen på de utvalda arbetsställena. Detta skapar problem i de fall då ett arbetsställe förändras under en urvalsomgång eller då uppgiftslämnaren inte har den information i sitt lönesystem som krävs för att kunna lämna de efterfrågade uppgifterna. Ett arbetsställe som ingår i urvalet, men som inte har möjlighet att redovisa de uppgifter som efterfrågas kontaktas angående problemet.

Ett återkommande problem är att KS/KV-undersökningarna efterfrågar uppgifter på arbetsställenivå och en del företag skulle hellre vilja lämna uppgifter på företagsnivå. I dessa specialfall har produktionsgruppen möjlighet att tilldela arbetsstället särskilda styrkoder som gör att blanketten hanteras annorlunda när den kommer in. Ett exempel är när ett företag inte kan särredovisa sina arbetsställen. Låt säga att ett företag består av fem arbetsställen totalt varav ett ingår i undersökningarna. Företaget uppger att de inte har möjlighet att lämna in uppgifter endast avseende det efterfrågade arbetsstället, men däremot kan de lämna uppgifter som omfattar hela organisationsnumret, det vill säga för

alla fem arbetsställena. Produktionsgruppen lägger då styrkoden 73 som tillåter uppgiftslämnaren att redovisa hela organisationsnumret på blanketten. När blanketten kommer in räknas uppgifterna ned med en kvot. Kvoten som används är antalsuppgiften från FDB avseende arbetsstället dividerat med antalsuppgiften från FDB avseende hela organisationsnumret. Efter omräkning är uppgifterna i KS/KV-systemet inte exakta, men de är betydligt bättre än alternativet att inte få in några uppgifter alls.

Ett liknande problem som också tas om hand med hjälp av en annan styrkod och omräkning av de inkomna uppgifterna är när ett arbetsställe delas eller om uppgiftslämnaren inte har möjlighet att redovisa både tjänstemän och kollektivanställda på samma blankett. I det senare fallet är det vanligt att det krävs två uppgiftslämnare för att få uppgifter som omfattar hela arbetsstället, en som redovisar uppgifter avseende tjänstemännen och en som redovisar uppgifter avseende de kollektivanställda. I dessa fall skapar produktionsgruppen redovisningsenheter. I praktiken innebär det att de båda uppgiftslämnarna får varsin blankett och när de har fyllt i dessa och skickat in dem så slås deras uppgifter ihop. På så sätt får SCB exakta uppgifter som avser hela det utvalda arbetsstället. När dessa redovisningsenhetssuppgifter inkommer till SCB granskas de och kan falla ut som fel. Uppgifterna summeras automatiskt av KS/KV-systemet om samtliga uppgifter på samtliga redovisningsenheter är korrekta. Om det finns felkoder på minst en redovisningsenhetssuppgift summeras inte uppgifterna utan felkoderna måste rättas upp först. Rättas inte samtliga felkoder upp räknas arbetsstället som bortfall. Materialet till denna uppsats har rensats på redovisningsenheter och endast "hela" arbetsställen är således medräknade.

För att få en uppfattning om hur mycket av produktionsgruppens tid som läggs på mikrogranskning och av vilken vikt den är för slutresultatet har processdata tagits fram. I diagram 1 och diagram 2 visas fördelningen mellan olika styrkoder avseende blanketterna tillhörande kvartal 2 år 2006 för KS respektive KV.

Diagram 1. Fördelning av styrkoder under kvartal 2 år 2006 på KS-undersökningen.

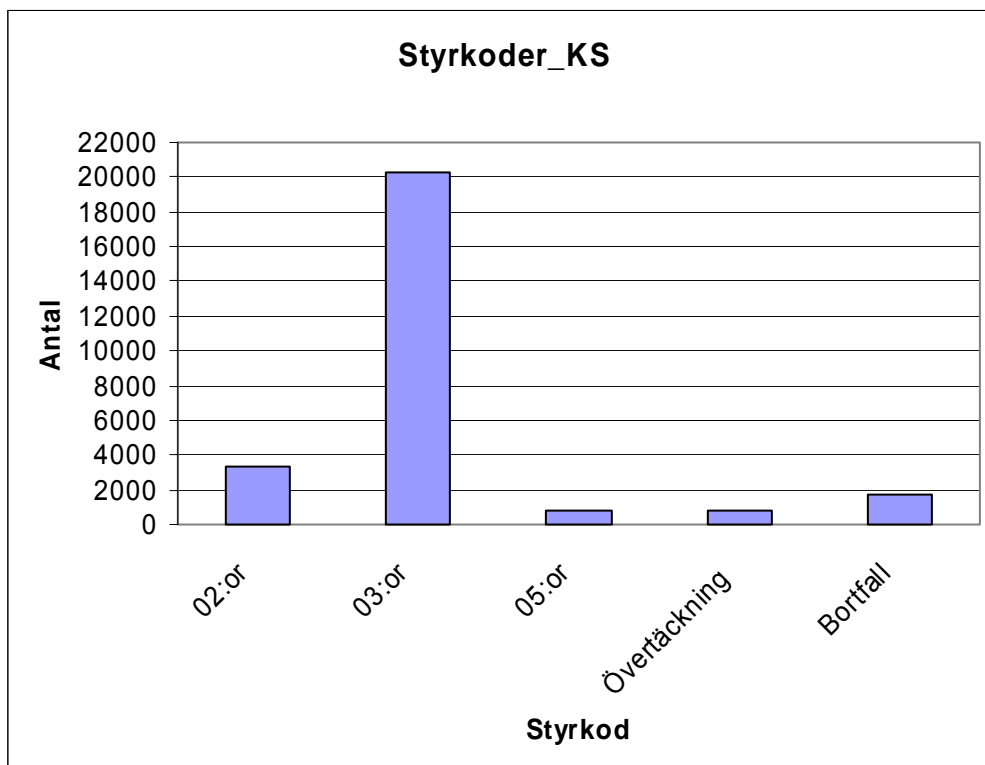
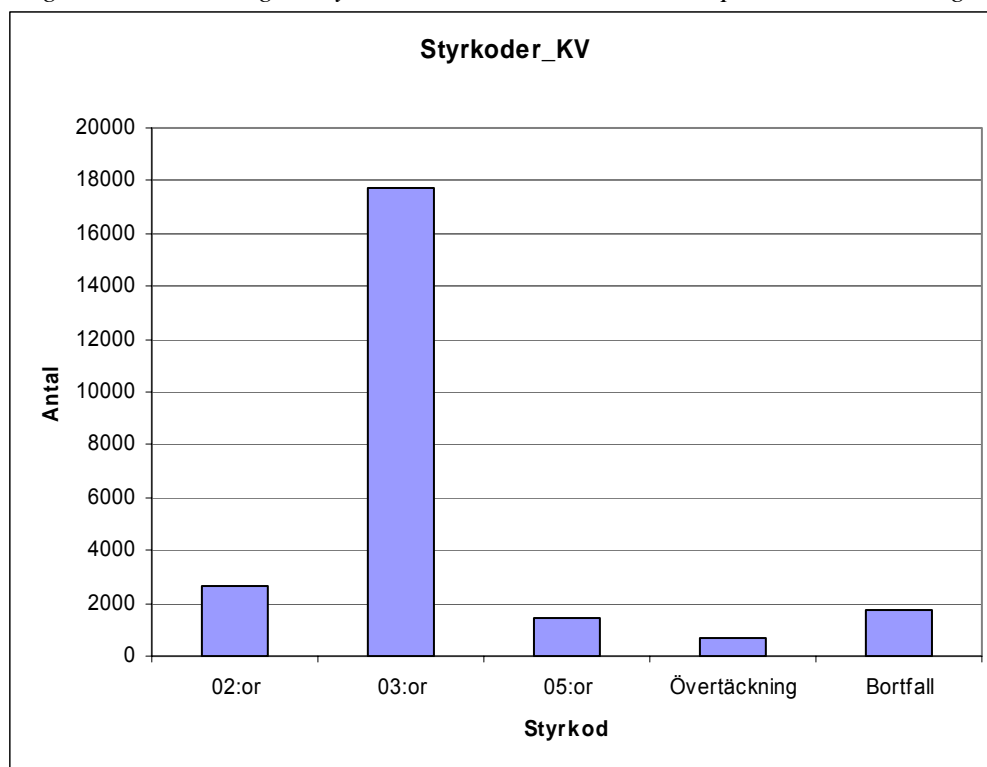


Diagram 2. Fördelning av styrkoder under kvartal 2 år 2006 på KV-undersökningen.



Som nämnts tidigare anger styrkoden 02 att en blankett är inkommen med icke godkända uppgifter. En 03:a är en godkänd blankett. När en 02:a rättas upp får den automatiskt styrkoden 03. 05:or är blanketter som inkom med enligt KS/KV-systemet icke godkända uppgifter, men som vid återkontakt med uppgiftslämnaren visade sig vara korrekta enligt denna. Produktionsgruppen har då tvingat igenom blanketten trots ej accepterade uppgifter. Övertäckningen består av arbetsställen som vid urvalsdragningen var aktiva och hade minst en anställd enligt FDB, men som under kvartalets gång försatts i konkurs, bytt sektor, inte hade några sysselsatta på mätdagen, blivit vilande eller så har verksamheten lagts ned. Bortfallet avser objektsbortfall, det vill säga de efterfrågade uppgifterna från arbetsstället saknas. 02:orna som visas i diagram 1 och 2 är antalet inkomna 02:or tillhörande kvartal 2, merparten av dessa har innan publiceringen av kvartal 2 rättats upp och blivit 03:or. 02:or som inte rättats upp i tid till publiceringen tvingas alltså igenom med styrkod 05, räknas som övertäckning eller bortfall. På grund av detta innehåller stapeln över 03:or både blanketter som inkom och fick styrkoden 02, men som rättades upp i tid och blev godkända och blanketter som blev godkända från början. Stapeln över 05:or består helt och hållet av blanketter som inkom och fick styrkoden 02, men som sedan har tvingats igenom med hjälp av styrkoden 05. Även stapeln över övertäckning innehåller en mindre andel blanketter som inkom som 02:or och detsamma gäller för stapeln som visar bortfallet. För att ge en rättvis bild ska det dock nämnas att antalet 02:or som räknas som bortfall vid publicering av kvartal 2 2006 var färre än 20st.



Tabell 1 visar en tabell som sammanfattar resultaten av de processdata som framställts.

Tabell 1. Processdata avseende kvartal 2 år 2006 KS och KV.

	KS	KV
<b>Antal inkomna blanketter</b>	21 239	19 310
<b>Antal felsignalerade blanketter (02:or)</b>	3 348	2 617
<b>Antal ändrade blanketter</b>	2 746	473
<b>Antalet tvingade blanketter (05:or)</b>	789	1 457
<b>Därav antalet tvingade blanketter (05:or) på den totalundersökta delen</b>	212	1 418

I appendix 4 finns en tabell över hur många och vilka felkoder som föll ut avseende kvartal 2.

Dagens granskningssystem fungerar inte effektivt för KV-undersökningen då det utläses ur tabell 1 att de flesta av de inkomna felaktiga blanketterna tvingas trots ej accepterade värden. Det utläses också att granskningen verkar fungera väsentligt sämre för den totalundersökta delen.

Den totalundersökta delen karaktäriseras av erfarna uppgiftslämnare. De mottar KV-blanketten varje månad och faller endast i undantagsfall ur urvalet, de flesta har således besvarat enkäten sedan undersökningen startade år 2000. Problemet här är inte att uppgiftslämnarna fyller i blanketten fel, har missförstått frågorna eller liknande utan problemet är granskningskontrollen av variabeln *lediga jobb*. De flesta blanketter tvingas med kod 05 för att uppgiftslämnaren har fyllt i att arbetsstället inte hade några lediga jobb på mätdagen och vid återkontakt med uppgiftslämnaren har det visat sig att uppgiften är korrekt. Cirka 80 procent svarar att de inte hade något ledigt jobb på mätdagen. Ett antagande som SCB gör är att arbetsställen med många anställda, för närvarande går gränsen vid > 199st, borde söka åtminstone en medarbetare varje månad, men detta stämmer inte med verkligheten. Vissa företag anställer nya medarbetare, men deras rekryteringsprocesser passar inte in på SCB:s definition av rekrytering. Exempel på detta är företag som är försiktiga med att anställa och rekryterar i första hand inom företaget innan de vänder sig externt, dessa lediga jobb skall inte medräknas i statistiken eftersom de inte står till arbetsmarknadens förfogande. Andra företag ingår i branscher som inte expanderar eller är i behov av att ersättningsrekrytera. Det kan gå månader eller till och med år mellan annonstillfällena.

Ett annat problem är att ett arbetsställe kan söka tio nya medarbetare i exempelvis månad 1, men i månad 2 månad söker de inte någon ny och då blir månad 1:s blankett godkänd, men månad 2:s blankett faller ut som fel. Den nuvarande granskningen tar således inte hänsyn till tidigare inlämnade uppgifter. KV-undersökningen, liksom KS-undersökningen, använder sig av mätdag vilket innebär att om uppgiftslämnaren exempelvis har en annons ute dagen innan mätdagen och efter mätdagen, men inte på mätdagen så faller blanketten ut som fel på grund av utebliven extern rekrytering. Variabeln är därför svår att granska.

Tabell 1 kan bland annat användas till att beräkna andelen blanketter som inkommit och fått minst ett värde ändrat innan publicering. Om denna beräkning ger ett litet värde jämfört med andelen felsignalerade blanketter antyder det att granskningsprocessen är ineffektiv. Andelen ändrade blanketter i kombination med andelen felsignalerade blanketter ger här ytterligare bevis för att granskningen avseende KV behöver effektiviseras. Cirka 82 procent av de KS-blanketter som inkommit med felaktiga värden har ändrats innan publicering medan motsvarande siffra för KV endast är cirka 18 procent.

### 3 Granskning – några metodaspekter

Statistiken är behäftad med många olika sorters fel, exempelvis ramfel, urvalsfel, mätfel, bortfallsfel och bearbetningsfel. När en produktionsprocess för att ta fram statistik ska utformas gäller det att hitta en rimlig balans mellan olika felkällor. Därför fanns det goda skäl att försöka utvärdera effektiviteten i KS/KV:s granskningssystem.

#### 3.1 Granskningens syfte och roll

Granskningens syfte kan beskrivas av nedanstående punkter [9]:

- Förbättra respektive undersökning, försöka att förhindra att fel uppkommer
- Öka kvaliteten i ingående och utgående data på ett effektivt sätt
- Bidra till kvalitetsbedömningar av statistiken

Granskningens roll kan beskrivas av nedanstående punkter [9]:

- Identifiera felkällor i respektive undersökning, speciellt problem för uppgiftslämnaren att besvara frågorna
- Identifiera och åtgärda betydelsefulla fel

Begreppet betydelsefulla fel kräver en förklaring. Granskningen ska hitta de fel i indata som, om de förblir oupptäckta, skulle medföra större systematiska eller slumpmässiga fel i statistiken.

#### 3.2 Olika kategorier av fel

I granskning utförs kontroller av data, en del kontroller signalerar fel i materialet medan andra kontroller signalerar så kallade misstänkta fel. Avseende de misstänkta felen så kan inte kontrollen med säkerhet säga att det ifrågasatta värdet är fel utan bara att det ser misstänkt ut. Värdet som kontroller med säkerhet kan konstatera att de är fel kallas här *uppenbara fel* och de värden som ser misstänkta ut och visar sig vara fel kallas här för *misstänkta fel*. De misstänkta felen kan klassificeras i *avvikelsefel*, värdet på variabeln är för stort eller för litet, och i *definitionsfel*. Ett exempel definitionsfel är felaktigt svarsbeteende, det vill säga många uppgiftslämnare uppfattar en fråga eller underliggande definitioner på ett likartat men felaktigt sätt.

#### 3.3 Hidioglou-Berthelots metod

Den metod som för närvarande används för att kontrollera rimligheten i inkomna uppgifter om antalet anställda i KS är Hidioglou-Berthelots metod (HB-metoden) [8]. Metoden används i granskningskontrollen bakom felkod 09. Felkod 09 ger en indikation om att det inrapporterade värdet för *totalt antal anställda* inte verkar vara korrekt. Om arbetsstället inkommit med uppgiften *totalt antal anställda* föregående månad (då arbetsstället tillhör den totalundersökta delen) eller föregående kvartal (då arbetsstället ingår i den kvartalsundersökta delen) och den uppgiften är granskad och vid behov rättad jämförs den nu inkomna uppgiften mot denna uppgift. Om arbetsstället inte ingått i urvalet tidigare eller inte inkom med uppgiften om *totalt antal anställda* föregående omgång jämförs istället den nu inkomna uppgiften mot arbetsställets antalsuppgift i FDB. HB-metoden beskrivs ingående av Hidioglou och Berthelot [8] Här följer en kort beskrivning av metoden;

Låt  $y_k(t)$  och  $y_k(t-1)$  vara värden för arbetsställe  $k$  vid tidpunkt  $t$  respektive  $t-1$  för variabeln  $y$ .

Låt  $R_k = \frac{y_k(t)}{y_k(t-1)}$  vara förändringen i  $y$  för arbetsställe  $k$ .

Det kan vara svårt att sätta upp gränser för vad som ska vara ett godkänt värde på  $R_k$ . Företagsdata är ofta mycket skevt fördelade. På ett litet arbetsställe med säg fem anställda kan ett  $R_k$ -värde motsvarande 50 procents ökning vara helt betydelselöst medan 50 procents ökning på ett stort arbetsställe kan behöva kontrolleras. HB-metoden har utvecklats för att komma tillrätta med den här typen av problem. Metoden syftar till att identifiera betydelsefulla förändringar, både relativa och absoluta.

HB-metoden går ut på att transformera data i två steg. I första steget bildas variabeln  $S_k$ .

$$S_k = \begin{cases} (R_k - R_{median}) / R_k & \text{om } 0 < R_k < R_{median} \\ (R_k - R_{median}) / R_{median} & \text{om } R_k \geq R_{median} \end{cases}$$

Där  $R_{median}$  är medianvärdet för  $R_k$ -kvoterna. I KS-undersökningen är  $R_{median}$  beräknad på hela det inkomna datamaterialet avseende variabeln *totalt antal anställda*.

Transformationen innebär att hälften av alla  $S_k$  är mindre än noll och den andra hälften är större än noll. Hidioglou och Berthelot hävdar att denna transformation gör att avvikande värden upptäcks lika enkelt i båda "svansarna" (stora respektive små värden på  $R_k$ ) [8].

Data transformeras i ytterligare ett steg för att anpassa granskningen till den skeva storleksfördelningen i  $y$ :

$$E_k = S_k \cdot \{ \text{Max}(y_k(t), y_k(t-1)) \}^U \text{ där parametern } U \text{ väljs i intervallet } [0,1].$$

Valet av  $U$  ger möjlighet att anpassa kontrollen till observationens betydelse.  $U$  väljs alltså i KS-undersökningens fall utifrån hur mycket hänsyn som ska tas till det totala antalet anställda på arbetsstället.  $U = 0$  medför att kontrollen inte alls beror på storleken av  $y$  medan  $U = 1$  ger full betydelse åt storleken av  $y$ . Om  $U = 0$  får det till följd att om ett arbetsställe ökar från exempelvis två till fyra personer och ett annat arbetsställe ökar från 100 till 200 personer ska båda dessa arbetsställen granskas. Båda arbetsställets ökning antas då vara lika viktiga att reda ut. Om  $U = 1$  får det till följd att om föregående exempel skulle inträffa skulle endast arbetsstället som uppvisade en ökning från 100 till 200 personer granskas.

Värdena för  $E_k$  fördelas kring noll. De  $E_k$  som ligger långt från noll betraktas som misstänkta värden och bör kontrolleras exempelvis genom en återkontakt med uppgiftslämnaren.

### 3.4 Effektgranskning med hjälp av poängfunktioner

Ett problem med HB-metoden är att den inte ger någon direkt koppling mellan granskningskontroll och effekten på statistiken. Australiens statistikbyrå (Australian Bureau of Statistics, ABS) är framstående inom det område som kallas "significance editing". Denna metodik ger en tydligare sådan koppling och har beskrivits och tillämpats på några statistikbyråer, däribland på ABS. Lawrence och McDavitt [10] och Lawrence och McKenzie [11] är två viktiga referenser med anknytning till arbetet vid ABS. En lämplig benämning på svenska för "significance editing" är *effektgranskning*

eftersom metoden syftar till att spåra upp och kontrollera de data som har eller kan förväntas ha störst effekt på statistiken. Effektgranskning är en form av selektiv granskning där baskonceptet är användandet av poängar baserade på förväntad effekt av granskningen.

Lawrence och McDavitt redovisar resultat från ansatsen med effektgranskning som tyder på att granskningens omfattning skulle kunna halveras jämfört med ett mera konventionellt förfarande [10], som exempelvis det som för närvarande görs i KS/KV.

För att beskriva principen för effektgranskning behövs några beteckningar och definitioner, dessa är baserade på Lennart Nordbergs PowerPoint-presentation [12].

Låt

i)  $y_{jk}^{(OG)}$  beteckna ett inkommet och ogranskat värde på variabel  $y_j$  för arbetsställe  $k$ .

ii)  $y_{jk}^{(G)}$  vara inkommet, granskat och vid behov rättat värde på variabel  $y_j$  för arbetsställe  $k$ .

iii)  $w_k$  vara uppräkningsfaktorn för arbetsställe  $k$ .

Uppräkningsfaktorn, är en vikt (beräknas i CLAN [13]) och behövs då KS och KV är urvalsundersökningar som skattar totaler, det vill säga data samlas in på arbetsställenivå, men det är aggregerade data som publiceras (riksnivå, branschgruppsnivå etc.).

iv)  $\hat{T}_j^{(G)}(b) = \sum_{k \in b} w_k y_{jk}^{(G)}$  vara en skattning av totalen för variabel  $y_j$  i bransch  $b$ ,

baserad på granskade data. Vidare är  $\hat{T}_j^{(G)} = \sum_b \hat{T}_j^{(G)}(b)$ , det vill säga motsvarande punktskattning summerad över alla förekommande branscher.

v)  $\hat{T}_j^{(OG)}(b)$  och  $\hat{T}_j^{(OG)}$  är motsvarande storheter som i (iv) men baserade på ogranskade data  $\{y_{jk}^{(OG)}\}$ .

Avvikelsen mellan  $\hat{T}_j^{(G)}(b)$  och  $\hat{T}_j^{(OG)}(b)$ , det vill säga mellan punktskattningarna på granskat och ogranskat material är vanligtvis stor, men problemet skulle lösas om det gick att identifiera de arbetsställen som bidrar till de största felen. Det kan dock vara svårt att göra under löpande produktion (återkommer till den frågan senare), men kan enkelt utföras i efterhand.

Inför nu följande funktioner:

$$LS_{jk} = \frac{w_k |y_{jk}^{(OG)} - y_{jk}^{(G)}|}{\hat{T}_j^{(G)}(b)} \quad \text{för } k \in b.$$

$LS_{jk}$ , där  $LS$  står för *Local Score*, *poängfunktion* på svenska, är ett mått på hur stor den uppräknade avvikelsen är mellan granskat och ogranskat värde på variabel  $y_j$  för arbetsställe  $k$  i förhållande till branschtalen.

Låt vidare  $GS_k = \max_j \{LS_{jk}\}$  (där  $GS$  står för *Global Score*).

Antag att en gräns (tröskelvärde)  $\alpha$  bestäms och att endast för de arbetsställen där  $GS_k > \alpha$  granskas och rättas data, medan ogranskade data behålls för övriga arbetsställen.

### 3.4.1 Pseudo bias (PB)

Följande funktion  $PB_j(b, \alpha)$  är ett mått på hur stort fel man inför i skattningen av branschtotalen för  $y_j$  genom att granska och rätta endast de arbetsställen för vilka  $GS_k > \alpha$ . Även här står  $\alpha$  för tröskelvärdet,  $b$  för bransch och formlerna är baserade på Lennart Nordbergs PowerPoint-presentation [12].

$$PB_j(b, \alpha) = \frac{\hat{T}_j^{(PG, \alpha)}(b) - \hat{T}_j^{(G)}(b)}{\hat{T}_j^{(G)}(b)}$$

där

$$\hat{T}_j^{(PG, \alpha)} = \sum_{GS_k > \alpha} w_k y_{jk}^{(G)} + \sum_{GS_k \leq \alpha} w_k y_{jk}^{(OG)}$$

Ett annat sätt att skriva funktionen är:

$$PB_j(b, \alpha) = \frac{\sum_{GS_k \leq \alpha} w_k (y_{jk}^{(OG)} - y_{jk}^{(G)})}{\hat{T}_j^{(G)}(b)}$$

Om  $y_{jk}^{(G)}$  kunde ersättas i funktionen  $PB_j(b, \alpha)$  med de sanna värdena  $y_{jk}$  skulle en skattning av den bias (relativt punktskattningen) som de ogranskade data orsakar fås. I brist på kunskap om de sanna värdena används istället  $y_{jk}^{(G)}$  vilket är de bästa data SCB har, de används som ”proxies” för de sanna värdena. Detta medför att måttet på biasen endast är approximativt. Av det skälet benämns  $PB_j(b, \alpha)$  *pseudo bias*.

Ett alternativt mått på pseudo bias fås om  $\hat{T}_j^{(G)}(b)$  ersätts med medelfelet (eng. standard error)  $SE\left(\hat{T}_j^{(G)}(b)\right)$ . I denna studie har valet gjorts att använda  $\hat{T}_j^{(G)}(b)$ .

### 3.4.2 Förslag till granskningsstrategi

I kapitel 4 används funktionerna  $LS_{jk} = \frac{w_k |y_{jk}^{(OG)} - y_{jk}^{(G)}|}{\hat{T}_j^{(G)}(b)}$  för  $k \in b$

och måttet pseudo bias

$$PB_j(b, \alpha) = \frac{\sum_{GS_k \leq \alpha} w_k (y_{jk}^{(OG)} - y_{jk}^{(G)})}{\hat{T}_j^{(G)}(b)}$$

till att utvärdera det befintliga granskningsystemet. Dessa kan inte användas i löpande produktion eftersom tillgång då inte finns till samtliga  $y_{jk}^{(G)}$ -värden. Däremot kan en

ersättare ofta hittas i formlerna för respektive  $y_{jk}^{(G)}$ , ett ”förväntat” värde. I en kortperiodisk undersökning där undersökningsobjekten är med i urvalet under ett antal omgångar skulle värdet från föregående månad/kvartal/år kunna tjäna som förväntat värde, eventuellt efter justering med en säsongsfaktor och/eller trendfaktor.

$y_{jk}^{(G)}$  skulle således behöva ersättas med något slags förväntat värde  $\tilde{y}_{jk}$ :

$$L\tilde{S}_{jk} = \frac{w_k |y_{jk}^{(OG)} - \tilde{y}_{jk}|}{\hat{T}_j^{(G)}(b)}$$

$$P\tilde{B}_j(b, \alpha) = \frac{\sum_{G\tilde{S}_k \leq \alpha} w_k (y_{jk}^{(OG)} - \tilde{y}_{jk})}{\hat{T}_j^{(G)}(b)}$$

Värdet i nämnaren,  $\hat{T}_j^{(G)}(b)$  kan vara den publicerade punktskattningen från föregående undersökningsomgång.

Tröskelvärdet,  $\alpha$ , är inte enkelt att komma fram till. Det som föreslås är att pröva sig fram empiriskt genom att under några undersökningsomgångar ”övergranska” materialet.

En strategi för granskning skulle kunna se ut på det sätt som beskrivs härnäst. Förslaget är löst baserat på Lawrences och McDavitts artikel [10] och Lawrences och McKenzies [11].

- 1) Genomför en omfattande granskning av inkommande data under några produktionsomgångar. Spara både ogranskade och granskade data, det vill säga  $y_{jk}^{(OG)}$  och  $y_{jk}^{(G)}$ .

- 2) Välj förväntade värden  $\tilde{y}_{jk}$ .

- 3) Beräkna poängfunktionerna  $L\tilde{S}_{jk} = \frac{w_k |y_{jk}^{(OG)} - \tilde{y}_{jk}|}{\hat{T}_j^{(G)}(b)}$  och  $G\tilde{S}_k = \max_j \{L\tilde{S}_{jk}\}$ .

- 4) Prova olika tröskelvärden  $\alpha$  och välj ett värde så att

$$P\tilde{B}_j(b, \alpha) = \frac{\sum_{G\tilde{S}_k \leq \alpha} w_k (y_{jk}^{(OG)} - \tilde{y}_{jk})}{\hat{T}_j^{(G)}(b)} \text{ är tillräckligt liten.}$$

- 5) Utvärdera det valda tröskelvärdet genom att beräkna

$$PB_j(b, \alpha) = \frac{\sum_{G\tilde{S}_k \leq \alpha} w_k (y_{jk}^{(OG)} - y_{jk}^{(G)})}{\hat{T}_j^{(G)}(b)} \text{ och kontrollera att även } PB_j(b, \alpha) \text{ är liten.}$$

- 6) Använd det valda tröskelvärdet  $\alpha$  i kommande produktionsomgångar. Beräkna poängfunktionerna  $L\tilde{S}_{jk} = \frac{w_k |y_{jk}^{(OG)} - \tilde{y}_{jk}|}{\hat{T}_j^{(G)}(b)}$  och  $G\tilde{S}_k = \max_j \{L\tilde{S}_{jk}\}$ . Kontrollera

endast de arbetsställen för vilka  $G\tilde{S}_k > \alpha$ .

Att välja det förväntade värdet kan, som nämnts tidigare, vara svårt. I kortperiodiska undersökningar kan det alltså eventuellt vara möjligt att använda föregående månads/kvartals/års värde. Det hänger på att undersökningsobjekten har lämnat in

uppgiften som avses att användas som förväntat värde och objekten måste alltså ingå i urvalet minst två gånger. Ett problem kan också vara att den tidigare lämnade uppgiften kan behöva justeras exempelvis med en säsons- och/eller trendfaktor och dessa är inte alltid enkla att ta fram. I urvalsundersökningar uppstår problematiken att objekten byts ut över tid och i vissa undersökningar dras ett helt nytt urval varje gång och då kan objekten inte följas bakåt i tiden alls.

## **4 Resultat av utvärderingen av KS/KV:s befintliga granskningssystem**

### **4.1 Utvärdering av HB-metoden**

Då denna studie bland annat syftar till att utvärdera det befintliga granskningssystemet är det också nödvändigt att utvärdera hur effektivt HB-metoden fungerar i KS-undersökningen. Om det befintliga granskningssystemet inte fungerar effektivt, men HB-metoden gör det vore det lämpligt att fortsätta använda den. I så fall bör ett nytt sätt att granska endast tillämpas på övriga variabler. Utvärderingen av HB-metoden har inte tillägnats mycket plats i denna studie.

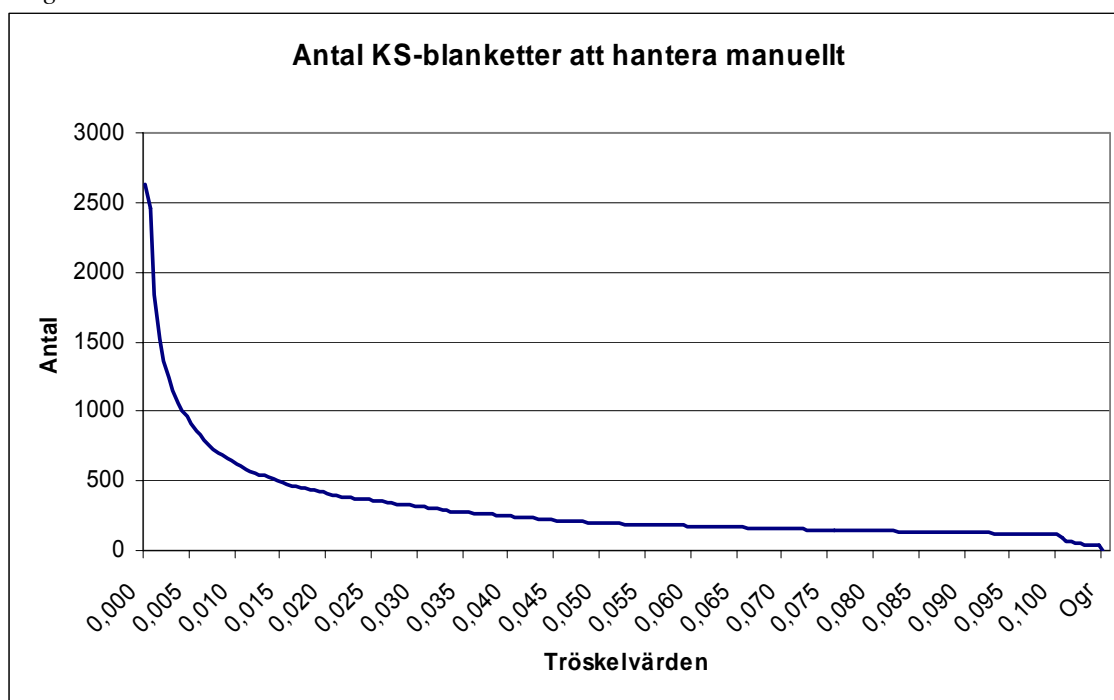
I denna del av utvärderingen har antalet icke godkända blanketter på grund av felkod 09 (antalet anställda skiljer sig mot antalsuppgiften för arbetsstället i FDB) tagits fram. Vidare har det undersökts hur många av dessa felaktiga blanketter som fått det felaktiga värdet ändrat. Antalet ändrade blanketter i relation till antalet inkomna och granskade blanketter ger ett mått på träffsäkerheten i dagens granskning. Dock bör två orsaker till att resultatet inte är exakt nämnas. Felkod 09 kan falla ut och vid återkontakt får produktionsgruppen en ny uppgift som registreras, men inte heller denna accepteras av granskningssystemet. Detta gör att blanketten bidrar till att höja andelen blanketter som har ändrats efter att denna felkod fallit ut, men blanketten har ändå felkoden kvar och är alltså inte upprättad. Den andra orsaken är att om en blankett inkommer utan anställda faller inte felkod 09 ut. Men då nya uppgifter fås vid återkontakt granskas dessa och då kan felkod 09 falla ut. Dessa fall kommer inte med i följande resultat. De nämnda fallen inträffar dock inte ofta och påverkar därför inte resultatet av utvärderingen nämvärt.

Avseende kvartal 2 föll 515st 02:or (ej godkända blanketter) med felkod 09 ut. Av dessa fick 230st ett ändrat värde innan publicering. De blanketter som inkom med felkod 09, men som efter återkontakt med uppgiftslämnaren fick en övertäckningskod är inte medräknade här. Det är inte enkelt att svara på om HB-metoden fungerar effektivt eller inte i KS-undersökningen. Det går inte att avgöra hur många eller vilka av de blanketter som föll ut med felkod 09 som påverkade slutresultatet och heller inte om samtliga ändringar som utfördes var av nytta.

### **4.2 Jämförelse mellan KS-ogranskade data och KS-granskade data**

Totalt finns det 2 746st inkomna KS-blanketter som innan publicering fått minst ett värde ändrat av totalt 21 239st inkomna KS-blanketter. Alla ändringar som görs påverkar inte slutresultatet så mycket att det är värt den tid som läggs ned på att granska och rätta upp materialet. En illustration av hur många färre blanketter som skulle ha kunnat granskas utan att slutresultatet påverkades i alltför stor utsträckning visas av diagram 3. På Y-axeln avläses antalet blanketter som skulle ha behövts hanteras manuellt vid de olika gränser för poängfunktionen som avläses på X-axeln. Poängfunktionen som avses finns beskriven i avsnitt 3.4.

Diagram 3. Antal KS-blanketter att hantera manuellt.



I diagram 3 kan avläsas hur många blanketter som skulle behöva hanteras manuellt för olika tröskelvärden. Diagram 3 skall avläsas i kombination med diagram 4-10 som följer. Diagram 4-10 visar, för variabeln *tillsvidareanställda män* och *tillsvidareanställda kvinnor* samt för varje total, den uppnådda pseudo biasen för olika gränser. Pseudo biasen avläses på Y-axeln och de olika gränserna på X-axeln. Motsvarande diagram över resterande variabler som ingår i KS finns i appendix 5. Pseudo bias är det fel som uppstår för att allt material inte granskas och är således ett viktigt mått, och det mått som valts att användas i denna utvärdering, för att se hur effektivt det befintliga granskningssystemet är. Felet som införs får givetvis inte bli alltför stort. Lawrence och McKenzie nämner att pseudo biasen helst skall ligga mellan 10-20 procent av medelfelet [11], men det går inte att tillämpa detta på alla undersökningar. I vissa undersökningar kan det tolereras högre bias och i andra krävs det att biasen är lägre. Således är det mycket svårt att bestämma tröskelvärdet,  $\alpha$ . Då gränsen väl är satt kan det kontrolleras att pseudo biasen ligger på en godkänd nivå genom att avläsa diagram 4-10 vid den valda gränsen. För att se hur många blanketter som skulle ha behövts hanteras manuellt om den valda gränsen använts avläses diagram 3. Diagram 4-10 är konstruerade genom tillämpning av den metod som presenterades i kapitel 3. I diagrammen nedan används inte den absoluta pseudo biasen, detta för att visa att ändringar av inkomna värden utförs i båda riktningarna.



Diagram 4. Pseudo bias på totalnivå för variabeln totalt antal anställda.

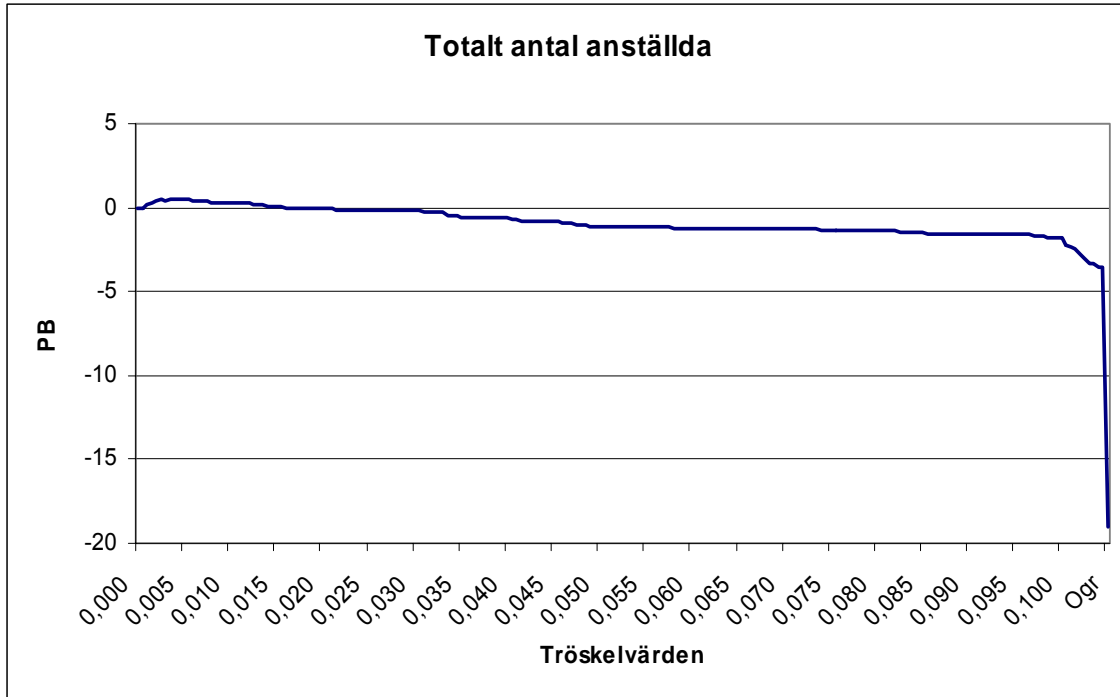


Diagram 5. Pseudo bias på totalnivå för variabeln tillsvidareanställda män.

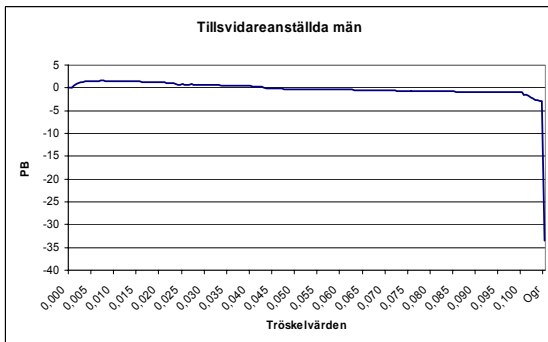


Diagram 6. Pseudo bias på totalnivå för variabeln totalt antal verksamma företagare

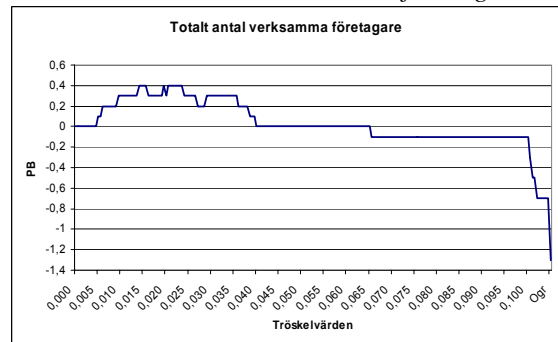


Diagram 7. Pseudo bias på totalnivå för variabeln tillsvidareanställda kvinnor.

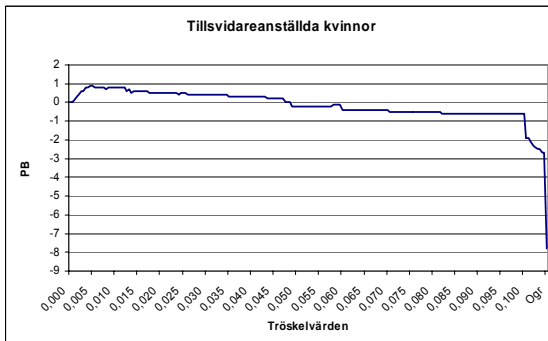


Diagram 8. Pseudo bias på totalnivå för variabeln totalt antal frånvarande.

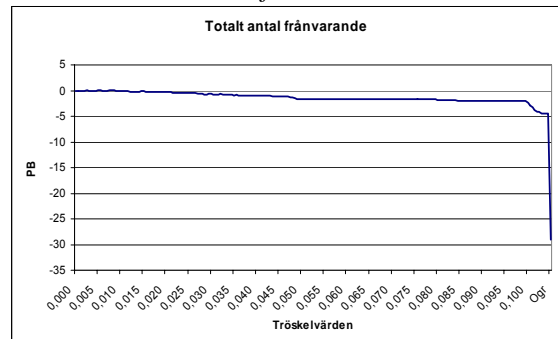


Diagram 9. Pseudo bias på totalnivå för variabeln totalt antal nyanställda.

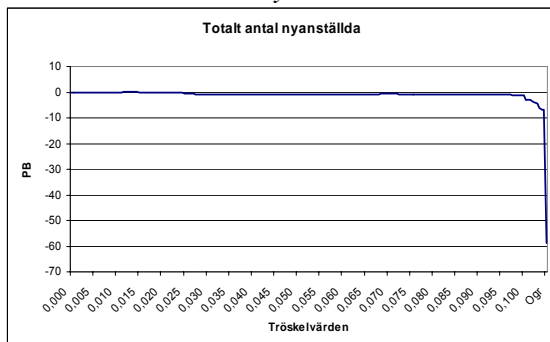


Diagram 10. Pseudo bias på totalnivå för variabeln totalt antal avgångna.

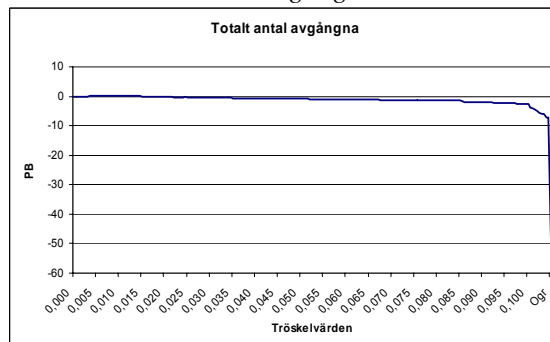


Diagram 4-10 visar att produktionsgranskning av materialet är nödvändigt, men också att det i dagsläget övergranskas. Det är också tydligt att om produktionsgranskning inte utfördes alls skulle samtliga variabler underskattas.

Diagram 3-10 baseras på över 200 olika gränser på X-axeln och att inkludera gränsen Ogranskat (Ogr i diagrammen) påverkar skalan vilket i sin tur medför en missvisande jämn linje i respektive diagram. Av den anledningen följer diagram som visar detsamma som ovan, men för ett kortare intervall av gränser. Detta ger en mer rättvis bild. Intervallet 0,002-0,02 har valts att visas. Diagram över de variabler som inte visas här återfinns i appendix 5.

Diagram 11. Antal KS-blanketter att hantera manuellt.

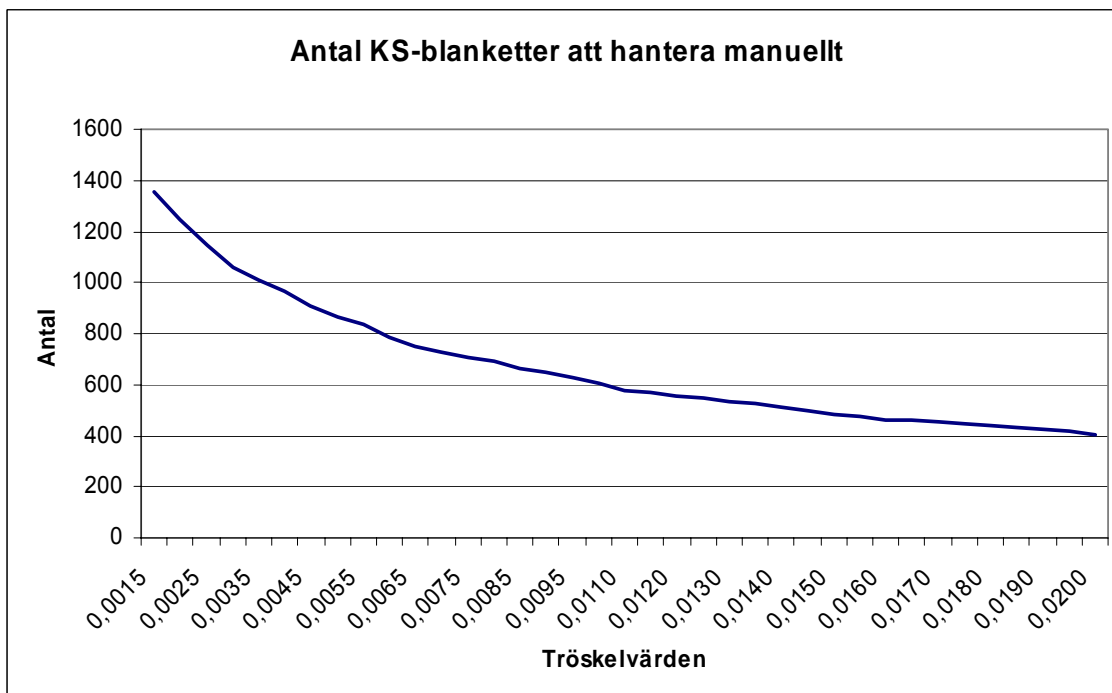


Diagram 12. Pseudo bias på totalnivå för variabeln totalt antal anställda.

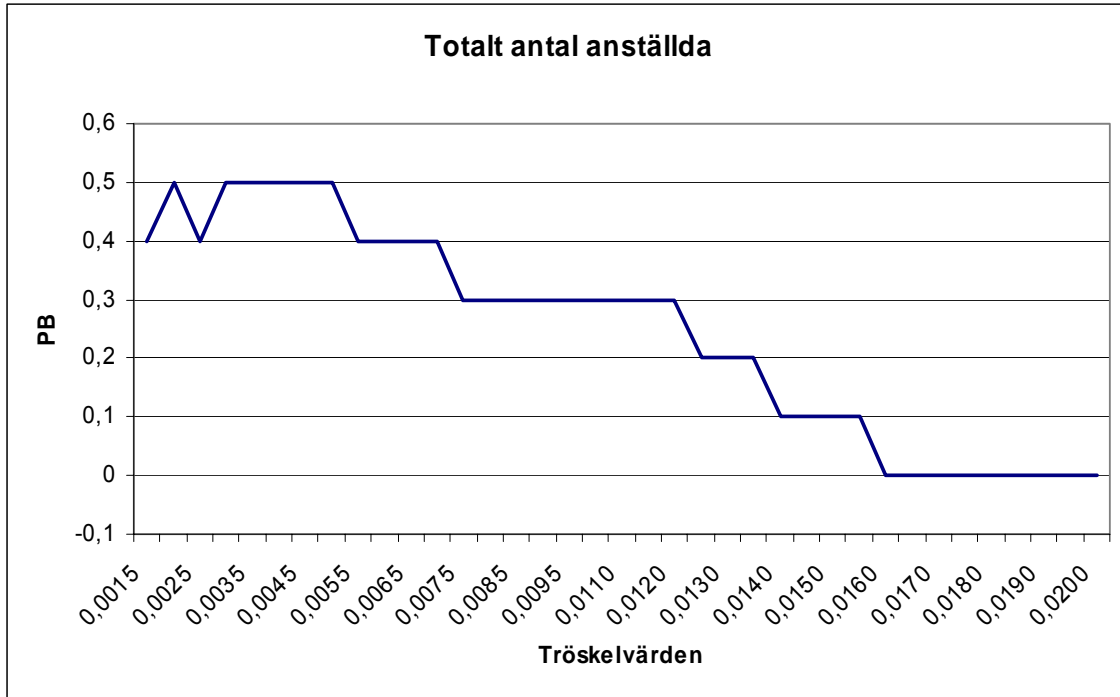


Diagram 13. Pseudo bias på totalnivå för variabeln tillsvidareanställda män.

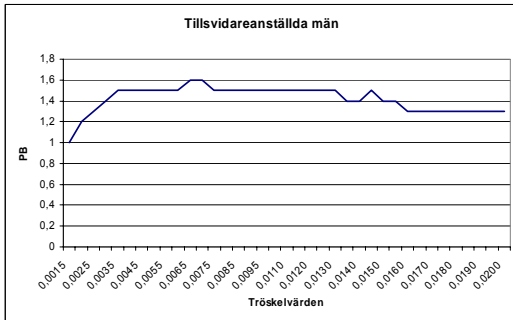


Diagram 14. Pseudo bias på totalnivå för variabeln tot. antal verksamma företagare.

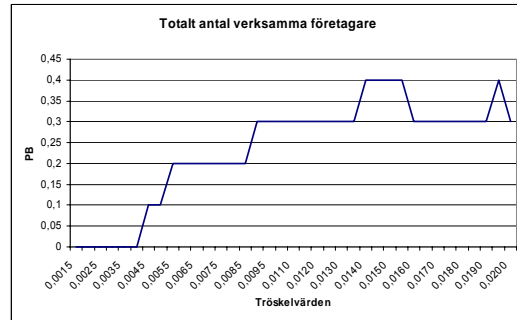


Diagram 15. Pseudo bias på totalnivå för variabeln tillsvidareanställda kvinnor.



Diagram 16. Pseudo bias på totalnivå för variabeln totalt antal frånvarande.

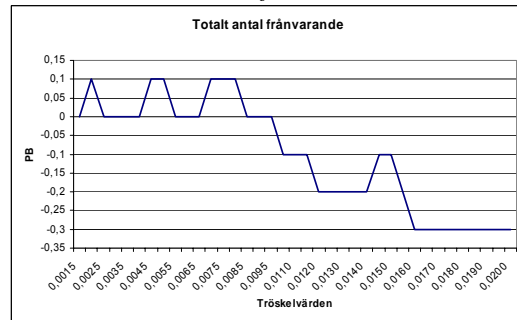


Diagram 17. Pseudo bias på totalnivå för variabeln totalt antal nyanställda.

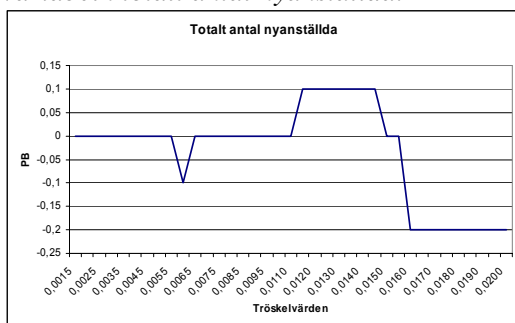
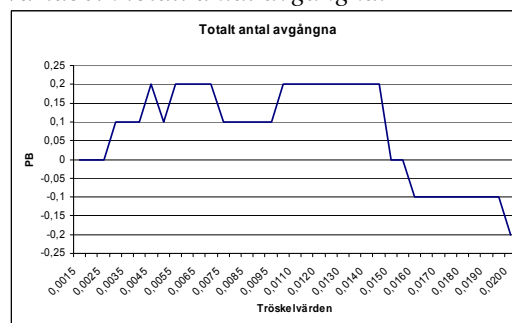


Diagram 18. Pseudo bias på totalnivå för variabeln totalt antal avgångna.



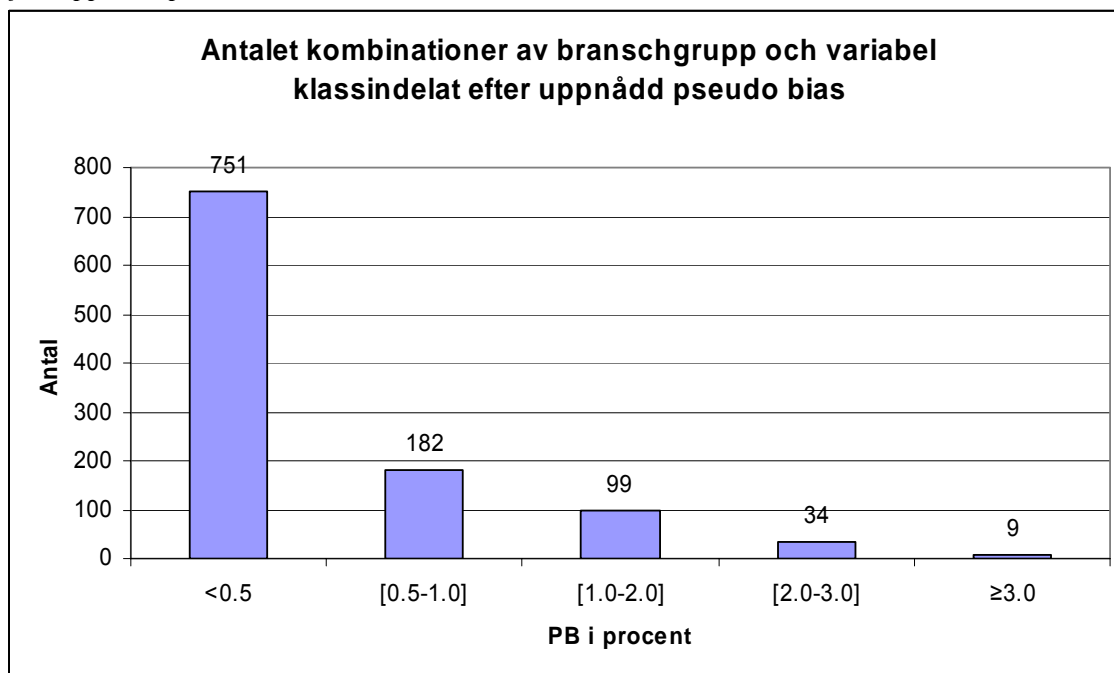
Som visas av diagram 4-10 och vilket blir ännu tydligare i diagram 12-18 är det inte enkelt att avgöra vilken gräns som ska väljas. I denna studie har det inte funnits tillräcklig tid att lägga ned på att välja "rätt" gräns. Den valda gränsen är således godtycklig och vald utifrån kriterierna att pseudo biasen hålls så låg som möjligt och att antalet blanketter som produktionsgruppen måste hantera manuellt också hålls så lågt som möjligt. Vald gräns i denna studie är 0,01. Att en gräns, trots tidsbristen, valts är för att illustrera hur arbetet med att införa ett granskningssystem som baseras på poängfunktioner skulle genomföras i praktiken steg för steg.

Diagram 4-10 och 12-18 visar hur stor brist på totalnivå, det vill säga riksnivå, som uppstår för att allt material inte granskas. Då den högst prioriterade redovisningen är branschgruppsnivå är det viktigt att undersöka storleken på pseudo biasen på den nivån för den valda gränsen, detta utfördes härnäst i utvärderingen.

KS innehåller 25 variabler och redovisas på 43 olika branschgrupper (exklusive punktskattningen för hela riket), detta ger sammanlagt 1 075 kombinationer mellan branschgrupp och variabel. Efter att ha klassindelade materialet med avseende på den uppnådda pseudo biasen för den valda gränsen, 0,01, visade resultatet att cirka 70 procent av kombinationerna har en pseudo bias som understiger 0,5 procent. Endast nio kombinationer har en pseudo bias på 3 procent eller mer och av dessa återfinns tre stycken i branschgrupp 4510<sup>1</sup>. Diagram 19 visar en sammanställning av resultatet för det klassindelade materialet, absolutbeloppet av pseudo biasen har använts i diagrammet.

<sup>1</sup> Nyckel finns i appendix 8.

Diagram 19. Fördelningen av antalet kombinationer av branschgrupp och variabel klassindelad efter uppnådd pseudo bias.



I appendix 5 visas en fullständig tabell över pseudo biasen som uppstår då den valda gränsen är satt till 0,01 per branschgrupp. Vid en analys av samtliga variabler sticker pseudo biasen på variabeln *tillsvidareanställda män* ut och även till viss del på variabeln *tillsvidareanställda kvinnor* samt på de variabler där verksamma företagare undersöks. KS-undersökningen publicerar inte företagari variablerna ännu och därför undersöks orsaken till den högre pseudo biasen endast för variabeln *tillsvidareanställda män* i denna studie. Resultatet av analysen av pseudo biasen inom vissa branscher avseende variabeln *tillsvidareanställda män* redovisas i avsnitt 4.2.1.

I tabell 2 visas hur många KS-blanketter produktionsgruppen skulle behöva granska per branschgrupp om gränsen  $\alpha$  sätts till 0,01.

Tabell 2. Antal blanketter att granska manuellt då gränsen är satt till 0,01.

<b>NGGr2<sup>2</sup></b>	<b>Manuell_KS</b>	<b>NGGr2</b>	<b>Manuell_KS</b>
<b>Totalt</b>	606	<b>4530</b>	20
<b>0100</b>	17	<b>5000</b>	18
<b>1000</b>	9	<b>5100</b>	35
<b>1500</b>	9	<b>5200</b>	13
<b>1800</b>	11	<b>5210</b>	9
<b>2000</b>	6	<b>5220</b>	14
<b>2100</b>	3	<b>5500</b>	22
<b>2200</b>	9	<b>6000</b>	20
<b>2300</b>	12	<b>6100</b>	17
<b>2500</b>	6	<b>6400</b>	9
<b>2600</b>	9	<b>6500</b>	14
<b>2700</b>	5	<b>6600</b>	8
<b>2800</b>	15	<b>7000</b>	21
<b>2900</b>	12	<b>7100</b>	18
<b>31XX</b>	15	<b>7200</b>	14
<b>3200</b>	1	<b>7300</b>	7
<b>3300</b>	9	<b>7400</b>	36
<b>3400</b>	10	<b>8000</b>	13
<b>3500</b>	2	<b>8500</b>	24
<b>3600</b>	3	<b>9000</b>	18
<b>4000</b>	8	<b>9100</b>	24
<b>4510</b>	34	<b>9200</b>	27

#### 4.2.1 Analys av variabeln *tillsvidareanställda män*

För att kunna svara på varför pseudo biasen på variabeln *tillsvidareanställda män* blev betydligt högre jämfört med andra variabler beslutades att analysera denna variabel närmare. Som nämnts tidigare finns det en fullständig tabell i appendix 5 som visar storleken på pseudo biasen på branschgruppsnivå då den valda gränsen är satt till 0,01. Vissa branscher visar i denna tabell upp en högre pseudo bias än övriga och det är dessa som har analyserats vidare. De utvalda branscherna är 2000, 4510, 4530, 5000, 5100, 5200, 6000 och 7400. Pseudo biasen beror på de blanketter som innehåller felaktiga värden, men vars globala poäng inte överstiger den valda gränsen. En analys av ett stickprov av de blanketter som hör till de utvalda branscherna utfördes. Då den globala poängen på dessa blanketter understiger den valda gränsen har inte dessa blanketters ogranskade värden ersatts med de granskade. Analysen visade att det finns två orsaker till problemet. Dels att det görs många små ändringar på variabeln *tillsvidareanställda män*, främst från noll tillsvidareanställda män till en tillsvidareanställd man. Den andra orsaken är att en del uppgiftslämnare fyller i totalen på frågan om antalet anställda, men fördelar därefter inte den ifyllda uppgiften på kön och anställningsform som också efterfrågas. Pseudo biasen på totalnivå för den aktuella variabeln, då gränsen sattes till 0,01, uppgick till cirka 1,5 procent. Detta motsvarar cirka 22 000 tillsvidareanställda män. Cirka 19 000 av dessa 22 000 tillsvidareanställda män förklaras av dessa två orsaker, vilket motsvarar en pseudo bias på 1,3 procent. Om dessa problem åtgärdas skulle pseudo biasen på variabeln *tillsvidareanställda män* sänkas till cirka 0,2 procent vilket vore en acceptabel nivå.

<sup>2</sup> Nyckel finns i appendix 8.

Förklaring till de små ändringar som utförs:

- Aktiebolag med aktiebolagsägaren/ägarna som enda verksamma personer inom bolaget medräknar inte sig själva i statistiken och blanketten faller då ut som fel eftersom SCB antar att det finns minst en verksam person på varje utvalt arbetsställe i KS/KV-undersökningarna. På KS-blanketten som skickas ut efterfrågar SCB antalet anställda på arbetsstället inklusive verksam/verksamma aktiebolagsägare, men av erfarenhet vet produktionsgruppen att uppgiftslämnaren ofta inte ser ägaren/ägarna som anställda. Därför medräknas de inte i redovisningen, analysen visade på att detta stämmer. De arbetsställen som inkommer med noll anställda återkontakts alltid av produktionsgruppen och rättas då oftast upp genom korrigerings av antalet anställda.

Förklaring till att variabeln *totalt antal anställda* fylls i, men inte följdfrågan:

- Analysen visade att ett problem som ger upphov till en väsentlig del av den granskning som behöver effektiviseras är partiellt bortfall på frågan om antalet anställda. Uppgiftslämnare fyller i blanketten fel på så vis att de fyller i *totalt antal anställda*, men de fördelar inte den uppgiften på kön och anställningsform vilket efterfrågas i samma frågeblock. Av erfarenhet vet produktionsgruppen att en del uppgiftslämnare har svårt att göra en exakt fördelning och därför fylls endast totalen i. Vid återkontakt med uppgiftslämnaren går det dock alltid att hitta en lösning, exempelvis kan fördelningen uppskattas. En annan orsak kan vara att uppgiftslämnaren tror att det räcker att fylla i totalen om det inte har skett någon förändring av antalet anställda sedan föregående mätning. De tror att SCB då kan göra fördelningen åt dem. Även detta löses vid återkontakt då uppgiftslämnaren får det förklarat för sig varför SCB inte kan ta emot endast totalen och lämnar då oftast fördelningen via telefon. Den tredje orsaken som analysen visade är att fördelning ibland görs i form av kryss på följdfrågan om fördelning på kön och anställningsform istället för siffror. Endast numeriska värden godkänns i granskningen vilket medför att även då detta inträffar tas en återkontakt med uppgiftslämnaren för att få den korrekta fördelningen.

#### **4.3 Jämförelse mellan KV-ogranskade data och KV-granskade data**

Totalt finns det 473st inkomna KV-blanketter som innan publicering fått minst ett värde ändrat av totalt 19 310st inkomna KV-blanketter. En illustration av hur många färre blanketter som skulle ha kunnat granskas utan att slutresultatet påverkades i alltför stor utsträckning visas av diagram 20 i kombination med diagram 21-23. På Y-axeln i diagram 20 avläses antalet blanketter som skulle ha behövts hanteras manuellt vid de olika gränser för poängfunktionen som avläses på X-axeln. Poängfunktionen som avses finns beskriven i avsnitt 3.4.

Diagram 20. Antal KV-blanketter att hantera manuellt.

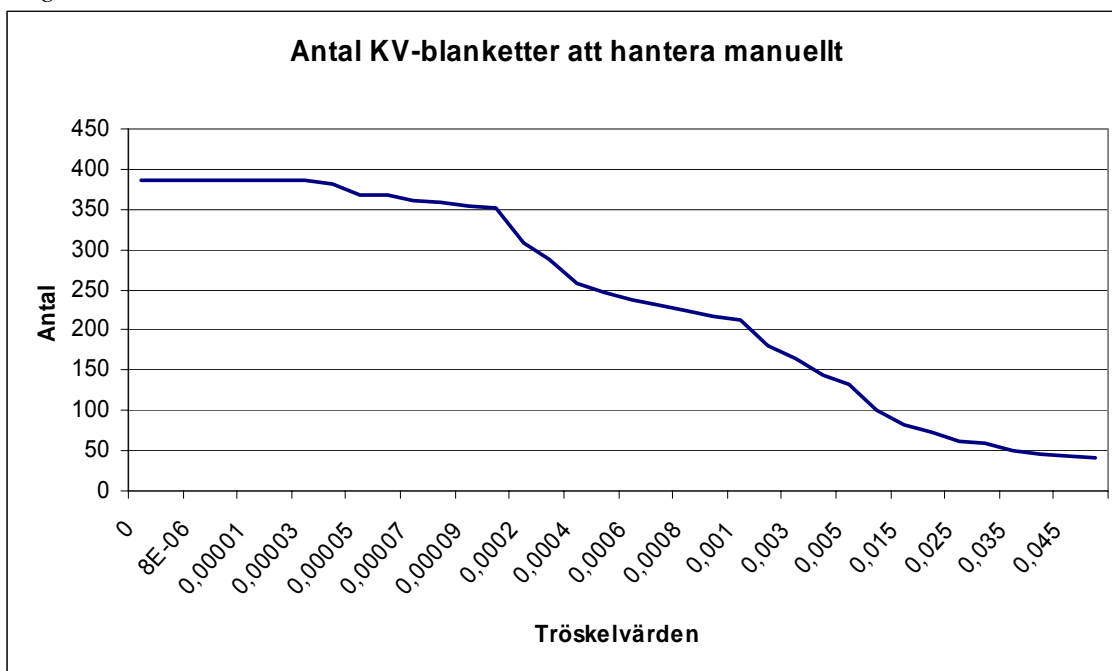


Diagram 21-23 visar den uppnådda pseudo biasen per variabel, vilken avläses på Y-axeln, för de olika gränser som avläses på X-axeln. Diagrammen visar tydligt att produktionsgranskning av materialet är nödvändigt, men också att det i dagsläget övergranskas. Diagram 21, 22 och 23 är konstruerade genom tillämpning av den metod som presenterades i kapitel 3. I diagrammen används inte den absoluta pseudo biasen, detta för att visa att ändringar av inkomna värden utförs i båda riktningarna. Diagram över följande variabler återfinns i appendix 6: *lediga jobb bemannade av avgående medarbetare, lediga jobb bemannade av vikarier/tillfälligt anställda och lediga jobb bemannade av bemanningsföretag/konsulter.*

Diagram 21. Pseudo bias på totalnivå för variabeln totalt antal lediga jobb.





Diagram 22. Pseudo bias på totalnivå för variabeln obemannade lediga jobb.

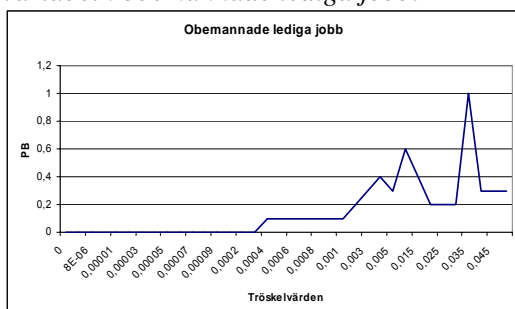
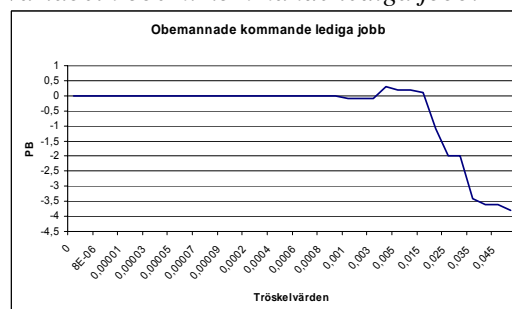


Diagram 23. Pseudo bias på totalnivå för variabeln obem. kommande lediga jobb.



Som visas av diagram 20-23 är det inte enkelt att avgöra vilken gräns som ska väljas. I denna studie har det inte funnits tillräcklig tid att lägga ned på att välja ”rätt” gräns. Den valda gränsen är således godtycklig och vald utifrån kriterierna att pseudo biasen hålls så låg som möjligt och att antalet blanketter som produktionsgruppen måste hantera manuellt också hålls så lågt som möjligt. Vald gräns i denna studie är 0,002.

Diagrammen visar hur stor brist på totalnivå, det vill säga riksnivå, som uppstår för att allt material inte granskas. Då den högst prioriterade redovisningen är branschgruppsnivå är det viktigt att visa hur pseudo biasen ser ut på branschgruppsnivå för den valda gränsen, detta visas i tabell 3.

Tabell 3 visar pseudo biasen som uppstår då den valda gränsen är satt till 0,002. Därefter visas en tabell där antalet blanketter att hantera manuellt per branschgrupp utläses, då gränsen är satt till 0,002.

Tabell 3. Pseudo biasen per branschgrupp då gränsen är satt till 0,002.

NGrp <sup>3</sup>	Tot. antal lediga jobb	Led. jobb bem. av avg. medarb.	Led. jobb bem. av vikarie/t.anst.	Led. jobb bem. av fgt/kons.	Obem. lediga jobb	Kommande obem. lediga jobb
Alla	0,3	-0,1	-0,2	-0,2	0,2	-0,1
1	0	0	0	0	0	0
2	0,4	-0,3	-0,4	-0,3	0,2	-0,6
3	0,2	0	0	0	-0,1	0
4	0,3	0	0	0	0,4	0,1
5	0,2	0	0	0	0,2	-0,1
6	0,3	-0,1	-0,2	-0,3	0,3	0,1
7	0	0	-0,1	0	-0,2	0
8	-0,1	0	0	0	0	0
9	0,1	-0,2	0	0	0	0,1

Tabell 4. Antal blanketter att granska manuellt per branschgrupp då gränsen är satt till 0,002.

NGrp	Manuell_KV
Alla	181
1	2
2	32
3	15
4	39
5	18
6	42
7	14
8	8
9	11

<sup>3</sup> Nyckel finns i appendix 8.

## 5 Resultat av testet där respektive stratummedelvärde använts som förväntat värde

Effektgranskningen i denna studie är tänkt att fungera som en illustration över hur införandet och tillämpningen av effektgranskning skulle ha kunnat utföras i praktiken. Som nämnts tidigare är KV:s variabler svårare att granska effektivt än KS:s då drygt 80 procent av uppgiftslämnarna svarar att de inte hade något ledigt jobb på mätdagen. Resonemang kring huruvida samma förväntade värde skulle kunna användas för både KS och KV har förts och det har konstaterats att en poängfunktion för KV skulle behöva använda ett annat förväntat värde än KS för att fungera effektivt. För att ge en så bra illustration över hur effektgranskning fungerar som möjligt har valet gjorts att i denna studie endast testa effektgranskningen på KS-undersökningen.

KS och KV är båda urvalsundersökningar och normalt vid urvalsbyte byts cirka 25 procent av urvalet ut. I april 2006 förändrades urvalsdesignen och detta fick bland annat till följd att mer än 40 procent av urvalet byttes ut vid detta urvalstillfälle. I denna studie hade det varit intressant att använda föregående kvartals inkomna KS-uppgifter som förväntat värde, men då en mycket hög andel av arbetsställena i kvartal 2 inte har varit med i undersökningarna tidigare finns det inte tillräckligt med material att använda. Dessutom skulle det ta för mycket tid att göra alla bearbetningar som har gjorts på kvartal 2:s data även på kvartal 1:s och det skulle vara nödvändigt att göra det för att kunna matcha ihop materialen. Ytterligare en faktor som har bidragit till att valet av förväntat värde i denna studie inte blev föregående omgångs värde är att kvartal 1:s data skulle ha behövt justeras med en säsongsfaktor och denna finns i dagsläget inte framräknad. Istället har valet gjorts att använda ett annat förväntat värde. Kvartal 2:s inkomna och granskade uppgifter har använts för att beräkna samtliga stratummedelvärden och dessa stratummedelvärden har sedan använts som förväntade värden i denna studie.

Poängfunktionen som beskrivs i kapitel 3 har tillämpats på samma material som utvärderingen av det befintliga granskningssystemet grundade sig på, men endast på KS-undersökningens data. Diagram 25-31 och diagram 33-39 är uppbyggda på samma sätt som diagram 4-10 och diagram 12-18 i kapitel 4. Diagram 24 skall avläsas i kombination med diagram 25-31. Detta diagram används nämligen till att utläsa hur många blanketter som skulle behöva hanteras manuellt av produktionsgruppen vid ett visst val av tröskelvärde  $\alpha$ . Diagram över de variabler som inte visas här återfinns i appendix 7.

Diagram 24. Antal KS-blanketter att hantera manuellt.

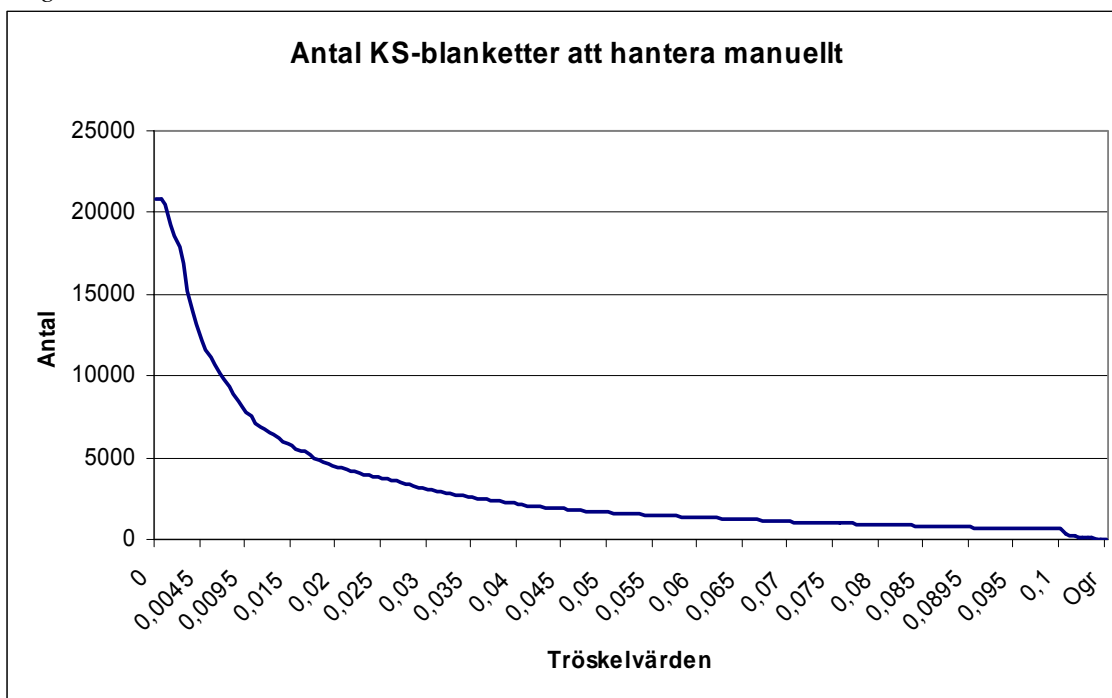


Diagram 25. Pseudo bias på totalnivå för variabeln totalt antal anställda.

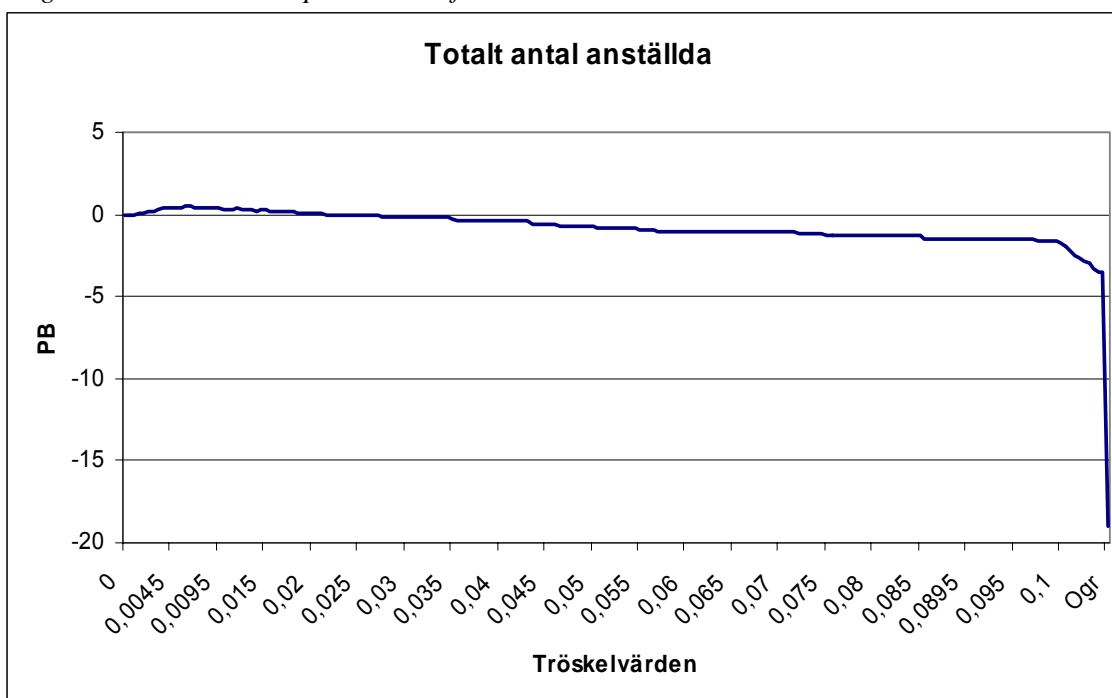


Diagram 26. Pseudo bias på totalnivå för variabeln tillsvidareanställda män.

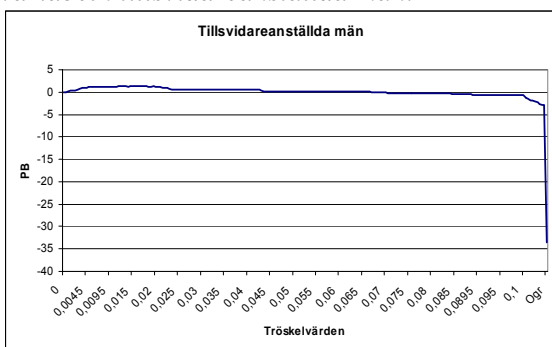


Diagram 27. Pseudo bias på totalnivå för variabeln totalt antal frånvarande.

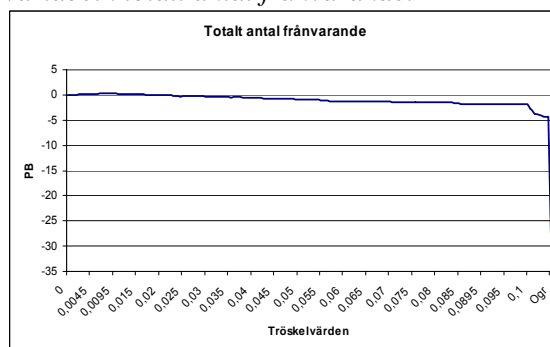


Diagram 28. Pseudo bias på totalnivå för variabeln tillsvidareanställda kvinnor.

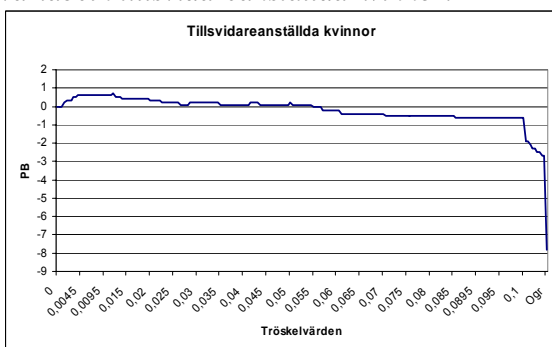


Diagram 29. Pseudo bias på totalnivå för variabeln totalt antal nyanställda.

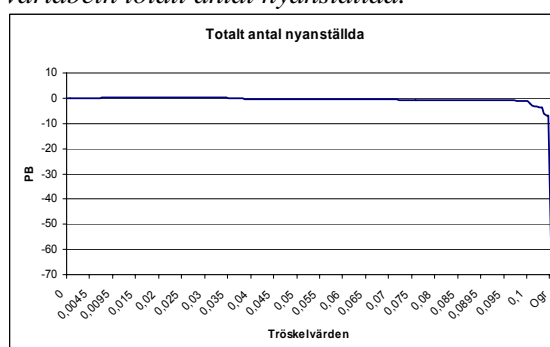


Diagram 30. Pseudo bias på totalnivå för variabeln totalt antal verksamma företagare.

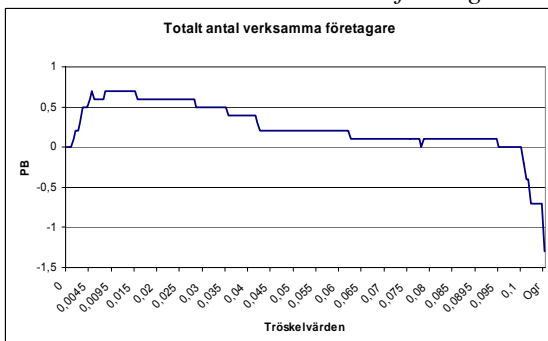


Diagram 31. Pseudo bias på totalnivå för variabeln totalt antal avgångna.

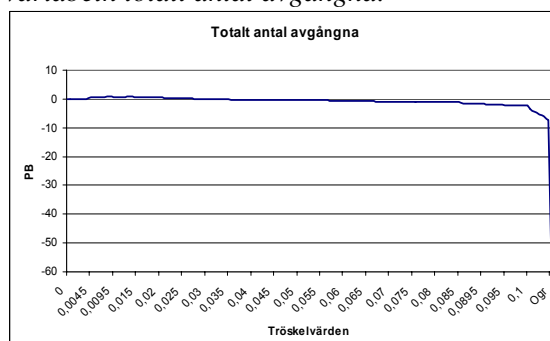


Diagram 24-31 baseras på över 200 olika gränser på X-axeln och att inkludera gränsen Ogranskat (Ogr i diagrammen) påverkar skalan vilket i sin tur medför en missvisande jämn linje i respektive diagram. Av den anledningen följer diagram 32-39 som visar detsamma som diagram 24-31, men för ett kortare intervall av gränser. Detta ger en mer rättvis bild. Intervallet 0,002-0,02 har valts att visas. Diagram 32 skall avläsas i kombination med diagram 33-39. Detta diagram används nämligen till att utläsa hur många blanketter som skulle behöva hanteras manuellt av produktionsgruppen vid ett visst val av tröskelvärdet  $\alpha$ . Diagram över de variabler som inte visas här återfinns i appendix 7.

Diagram 32. Antal KS-blanketter att hantera manuellt.

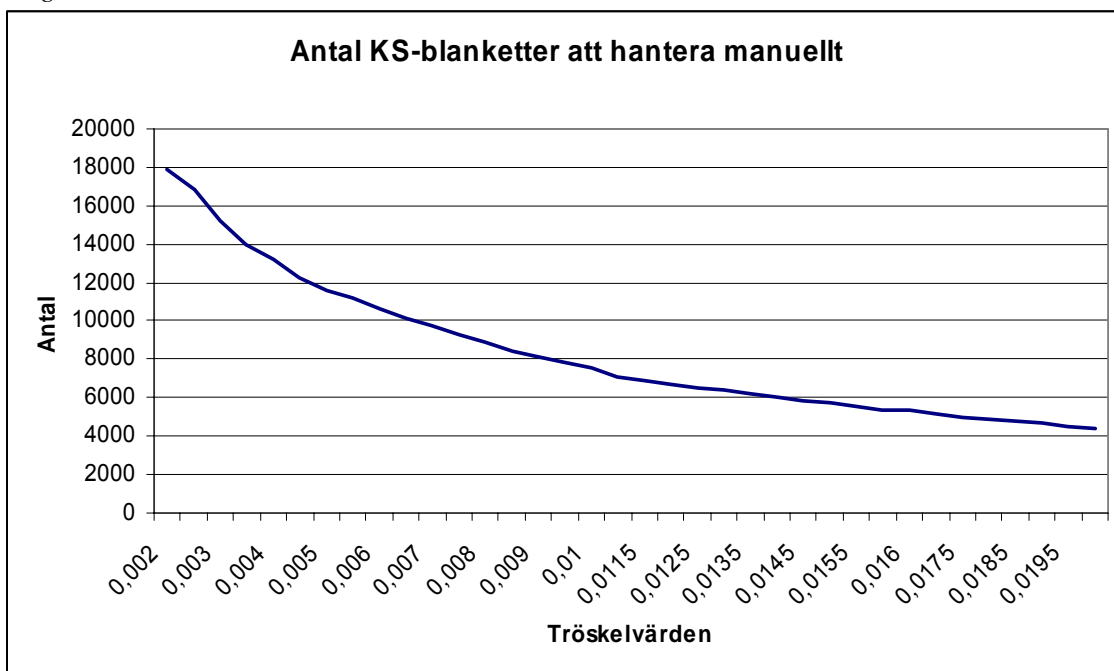


Diagram 33. Pseudo bias på totalnivå för variabeln totalt antal anställda.

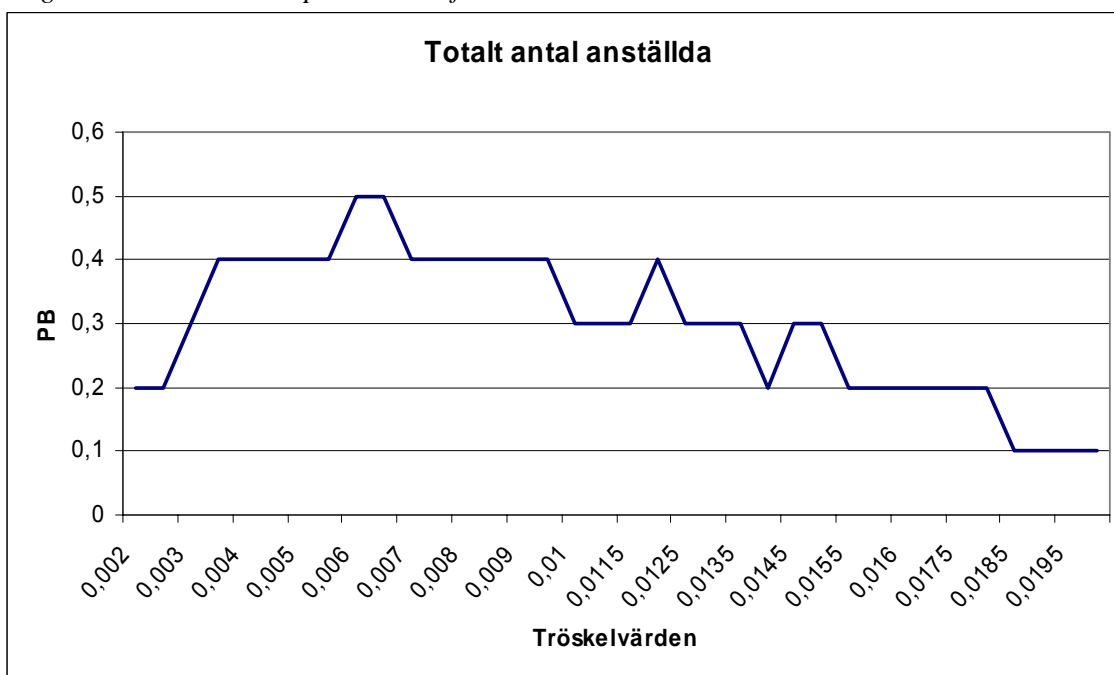


Diagram 34. Pseudo bias på totalnivå för variabeln tillsvidareanställda män.

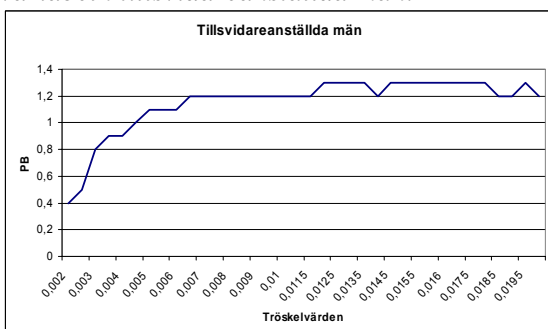


Diagram 35. Pseudo bias på totalnivå för variabeln tillsvidareanställda kvinnor.

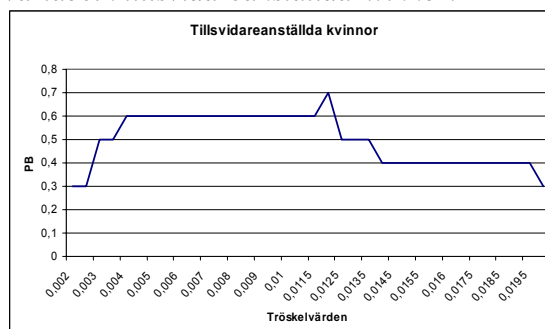


Diagram 36. Pseudo bias på totalnivå för variabeln tot. antal verksamma företagare.

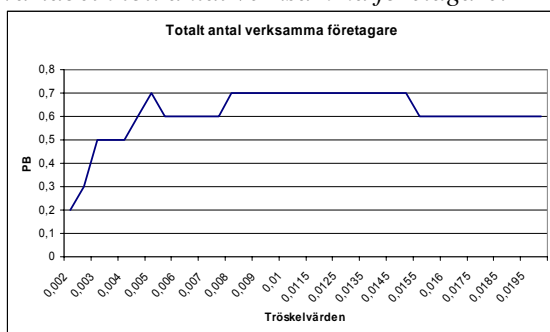


Diagram 37. Pseudo bias på totalnivå för variabeln totalt antal nyanställda.

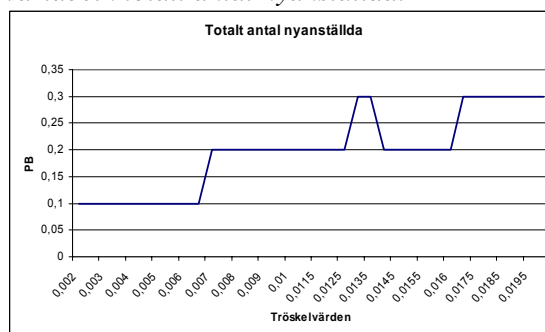


Diagram 38. Pseudo bias på totalnivå för variabeln totalt antal frånvarande.

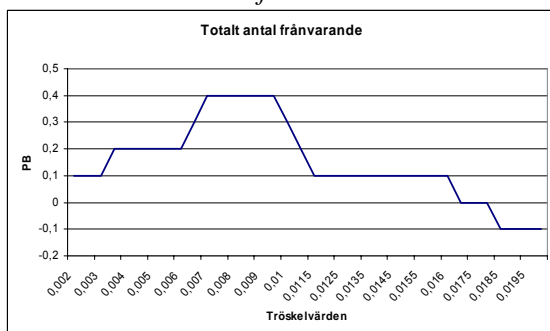
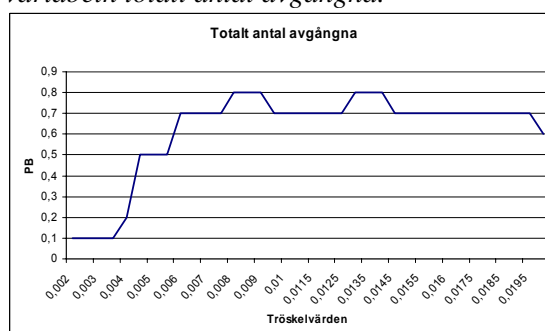


Diagram 39. Pseudo bias på totalnivå för variabeln totalt antal avgångna.



Diagrammen visar att om respektive stratummedelvärde används som förväntat värde resulterar det i en acceptabel storlek på pseudo biasen på totalnivå för de flesta variablerna. Linjen i diagram 33-39 kan tyckas variera mycket, men i de flesta fall varierar den i promille, inte i procent.

## 6 Diskussion och slutsats

Det tog tid innan det blev helt bestämt vilka delar av det totala materialet som denna studie skulle baseras på. Skulle beräkningarna baseras på samtliga uppgifter tillhörande en månad i kvartal 2 eller skulle beräkningarna göras på samtliga uppgifter tillhörande hela kvartal 2? Skulle endast den totalundersökta delen ingå i studien eller skulle hela urvalet ingå? Skulle samtliga inkomna blanketter ingå i beräkningarna eller skulle redovisningsenhetsblanketter och blanketter som innan publicering blivit övertäckning plockas bort? Vilka variabler skulle ingå i studien och skulle offentlig sektors blanketter inkluderas eller inte? Till slut beslutades, vilket framgick tidigare, att studien skulle omfatta privat sektor, hela det inkomna materialet avseende kvartal 2, både den totalundersökta- och den urvalsbaserade delen samt alla variabler förutom de som tillhör sjuklöneaviseringsdelen på KS-blanketten. Samtidigt beslutades också att blanketter som fått övertäckningskoder innan publicering och redovisningsenhetsblanketter skulle plockas bort ur materialet. Alla dessa frågor som ställdes innan arbetet kunde påbörjas gav insikt om att en väsentlig del av den tid som läggs på en studie består av planering. Alla frågor som dyker upp i planeringsstadiet behöver besvaras om möjligt för att minimera risken att upptäcka/göra fel i studien.

Som verktyg i studien användes SAS och Query Analyzer och Excel. Det tog tid att lära sig grunderna i SAS-programmering, då detta var ett nytt verktyg att arbeta i, men det var väl spenderad tid. Utbildningen har nämligen bland annat resulterat i att ett 20-tal script som använts till att ta fram resultaten i denna studie har kunnat skrivas självständigt. När problem avseende programmeringen uppstod fanns ett stöd i Susann

Gustavsson (DFO/IT, SCB), hon agerade bollplank och problemen kunde lösas gemensamt. Lennart Nordberg (DFO/MET, SCB) har också varit behjälplig i denna studie som diskussionspartner avseende frågor om SAS-programmeringen och upplägget av denna studie.

## 6.1 Delsyfte 1 - Utvärderingen

Att utvärdera granskningen i KS- och KV-undersökningarna kunde ha gjorts endast genom att ta fram processdata och analysera denna. Därför var det särskilt intressant att kombinera framställning av processdata med att också använda en poängfunktion som utvärderingsmått.

I denna studie har störst del av tiden lagts på att bearbeta data. Ämnesenheten tillhandahöll rådata, vilket innebär data som är helt obearbetad. Det var tidskrävande, men absolut nödvändigt, att sätta sig in i materialet och att lära sig vilken information som finns i vilka tabeller och hur den informationen kunde nyttjas på bästa sätt. Vid analys av materialet som stod till förfogande upptäcktes att en mängd olika bearbetningar skulle bli tvungna att genomföras innan processdata skulle kunna framställas. Detta för att få fram ett material som ger en mer sann bild av den produktionsgranskning som utförs. De bearbetningar som behövde göras bestod bland annat i att rensa materialet på objekt tillhörande offentlig sektor, redovisningsenhetsblanketter och på blanketter som inkom efter publiceringsdatumet.

I materialet fanns även skanningsfel och vissa koder som används av enkätenheten och produktionsgruppen. Skanningsfelen och koderna skulle ha snedvridit resultaten av jämförelserna mellan det ogranskade och det granskade materialet om de inte hade tagits om hand innan processdata och jämförelserna gjordes. Om dessa bearbetningar inte skulle ha genomförts skulle det alltså inte ha varit möjligt att dra några slutsatser om den befintliga granskningen. Exempelvis koden '8888' som registreras av enkätenheten, för att signalera till produktionsgruppen att de i skanningen stött på ett svårtolkat värde, bokstäver istället för siffror eller decimaltal, skulle orsaka oerhörda differenser om de hade lämnats orörda och jämförts med motsvarande värde i granskat-tabellen. Under studiens gång upptäcktes fler brister vilket gjorde att många datakörningar fick göras om flera gånger efter att scripten uppdaterats. Resultatet av alla bearbetningar som gjorts är ett datamaterial som kommer att kunna ligga till grund för fler än denna studie och det kommer definitivt att användas i fallstudien som ingår i Lotta-projektet.

I utvärderingen upptäcktes att variabeln *tillsvidareanställda män* inte uppträdde som förväntat. Pseudo biasen för denna variabel blev i varje beräkning betydligt högre jämfört med övriga variabler och om problemet med variabeln kunde upptäckas och åtgärdas skulle det innebära en väsentlig sänkning av pseudo biasen. En analys på branschgruppsnivå genomfördes därför och de branscher som uppvisade en högre bias än de övriga analyserades ytterligare. Tillfredsställande nog kunde de två orsaker som ger den förhöjda biasen upptäckas och förhoppningsvis kommer de att kunna åtgärdas inom kort. Studien syftade bland annat till att utvärdera dagens granskning i KS- och KV-undersökningarna. Ett av granskningens huvudsyften är att identifiera felkällor för senare åtgärder i undersöknings- och produktionsprocessen [9]. Detta för att den ständigt ska förbättras och inte stagnera och för att ge svar på om frågeformulären är effektiva och ger svar på det som efterfrågas. Denna studie har i utvärderingsfasen funnit just problem i KS-frågeformulär. Efter åtgärd kommer denna upptäckt att ha bidragit till minskad granskning och färre återkontakter med uppgiftslämnarna och detta får till följd att inte bara ett av studiens delsyften är uppfyllt utan även ett av granskningens huvudsyften.

De två orsaker som ger upphov till en högre pseudo bias på framför allt variabeln *tillsvidareanställda män* är att många små ändringar görs på variabeln samt att frågan om antalet anställda många gånger fylls i på ett felaktigt sätt vilket ger partiellt bortfall på frågan. Om dessa två orsaker eliminerades skulle den återstående biasen endast bero på att allt material inte granskas. Följande stycke förklarar problemen mer ingående och även rekommendationer för att komma till rätta med problemen ges.

Det var alltså när processdata togs fram och analyserades som det upptäcktes att många ändringar på KS-undersökningen gjordes på variablerna *tillsvidareanställda män* och *tillsvidareanställda kvinnor*. Aktiebolag med aktiebolagsägaren/ägarna som enda verksamma personer inom bolaget medräknar inte sig själva i statistiken. Blanketten faller då ut som fel eftersom SCB antar att det finns minst en verksam person på varje utvalt arbetsställe i undersökningarna. På KS-blanketten som skickas ut efterfrågar SCB antalet anställda på arbetsstället inklusive verksam/verksamma aktiebolagsägare. Av erfarenhet vet dock produktionsgruppen att uppgiftslämnaren ofta inte ser ägaren/ägarna som anställda och därför medräknas de inte i redovisningen. Utvärderingen visade på att detta stämmer.

På de blanketter som skickas till enskilda firmor, enkla bolag, handelsbolag och kommanditbolag finns en fråga där SCB ber uppgiftslämnaren att endast fylla i verksamma företagare och att på den efterföljande frågan fylla i antalet anställda. Då både produktionsgruppens erfarenhet och utvärderingens resultat påvisar underskattning av antalet anställda avseende den del av urvalet som består av aktiebolag föreslås en förändring av frågeformuläret som motverkande åtgärd. Ändringen skulle bestå i att lägga till en fråga på frågeformuläret där SCB ber uppgiftslämnaren att fylla i antalet verksamma aktiebolagsägare och att sedan på fråga 2 endast fylla i antalet anställda. Denna ändring medför inte bara en förändring av blanketten utan även av webbformuläret, vägledningar och tabeller. Även IT-systemet kommer att behöva uppdateras. Detta genererar kostnader i första läget, men är mödan värd i långa loppet. Granskningen och antalet återkontakter med uppgiftslämnarna kommer förhoppningsvis att minska vilket i sin tur minskar SCB:s kostnader och uppgiftslämnarbördan.

Ytterligare ett skäl till att se över frågekonstruktionen på KS-blanketten är att utvärderingen visade att ett problem som ger upphov till en väsentlig del av den granskning som behöver effektiviseras är partiellt bortfall på frågan om antalet anställda. Uppgiftslämnare fyller i blanketten fel på så vis att de fyller i *totalt antal anställda*, men de fördelar inte den uppgiften på kön och anställningsform vilket efterfrågas i samma frågeblock. I utvärderingen framkom också att uppgiftslämnare ibland fördelar det totala antalet anställda, men i form av kryss på blanketten istället för med siffror. Detta medför extra arbete för enkätenheten, som bland annat har till uppgift att verifiera de skannade uppgifterna, samt för produktionsgruppen och uppgiftslämnarna.

Avseende KV-undersökningen behöver granskningskontrollen för felkod 53 förändras och effektiviseras. I dagsläget om ett arbetsställe har > 199 anställda enligt FDB, men inte redovisar något ledigt jobb på KV-blanketten faller blanketten ut som fel (02:a). Den idag statiska gränsen fungerar inte då utvärderingen tydligt visade att de blanketter som faller ut på grund av felkod 53 inte rättas upp utan tvingas igenom med styrkod 05. Detta är extra tydligt på den totalundersökta delen av urvalet där majoriteten av de blanketter som faller ut som fel tvingas igenom av produktionsgruppen. De uppgiftslämnare som hör till den totalundersökta delen är redan hårt belastade eftersom de mottar KV-blanketten en gång per månad och inte en gång per kvartal som resterande uppgiftslämnare. Detta medför att de genererar fler inkomna blanketter som behöver granskas vilket gör denna grupp av uppgiftslämnare till de stora vinnarna om en mer effektiv granskning införs.



Även för KV-undersökningens del rekommenderas att frågeformuläret och vägledningen ses över. Idag fyller många uppgiftslämnare i att de har ett eller flera lediga jobb men sedan på fråga 3, där de lediga jobben ska fördelas på olika kategorier, sätter de kryss på de olika kategorierna istället för att fylla i uppgifterna med siffror. Detta resulterar i onödigt kodningsarbete ('8888') för enkätenheten, i onödigt granskningsarbete för produktionsgruppen och i onödiga återkontakter med uppgiftslämnarna.

Som nämnts tidigare utförs många små ändringar av de inkomna värdena, exempelvis från noll anställda till en eller två anställda. Skulle dessa uppgifter kunna imputeras maskinellt? Låt säga att ett arbetsställe inkommit med en tillsvidareanställd man i två omgångar. Då skulle det kunna tänkas att systemet imputerar en tillsvidareanställd man den tredje gången blanketten inkommer med noll anställda, om arbetsstället fortfarande har en anställd enligt FDB. Istället för att uppgiftslämnaren ska behöva återkontakts angående detta. Det finns fler exempel där imputering skulle kunna vara en effektiv åtgärd. Detta behöver dock utredas ytterligare innan implementering. Genomtänkta kriterier och villkor behöver sättas upp och det bör undersökas vilka konsekvenser imputering kan ge.

I utvärderingen framkom det att både KS- och KV-undersökningen skulle kunna granskas mer effektivt. Problem har också upptäckts och förslag på lösningar har presenterats.

Det var intressant att produktionsgruppens erfarenhet av granskningsarbetet sammanföll med utvärderingens resultat. Detta ger tyngd till ytterligare satsningar på fortsatt utredningsarbete om hur granskningen i KS/KV-undersökningarna kan effektiviseras på bästa sätt både för uppgiftslämnarna och för SCB.

## **6.2 Delsyfte 2 – Det alternativa granskningsystemet och problematiken kring det förväntade värdet**

Vissa variabler är av sådant slag att de är svåra att hitta bra förväntade värden till, ett exempel på det är KV:s variabel *totalt antal lediga jobb*. Variabeln är svår att granska då den kan anta ett stort antal olika värden och då det är svårt att förutse vilka värden uppgiftslämnaren "borde" redovisa. Ett arbetsställe kan i månad efter månad redovisa noll lediga jobb och sedan plötsligt söka, i extremfallet, ett par hundra nya medarbetare. Frågan som uppstår är om användandet av poängfunktion är det mest effektiva sättet att granska variabler av detta slag på och i så fall hur denna poängfunktion skulle se ut? En tanke är att använda föregående omgångs inlämnade värde om det finns, men att också inkludera fler faktorer i granskningsvillkoren, exempelvis läns-, bransch- och storleksklassstillhörighet, tidpunkt, trend och säsong. Finns det fler faktorer som skulle kunna bidra till ett väl fungerade jämförelsemått och finns det någon bra hjälpinformation att använda? Skulle det vara möjligt i praktiken för produktionsgruppen att kunna granska kontinuerligt om alla dessa faktorer och kanske ännu fler ska ingå i granskningsvillkoren? För att besvara detta behöver frågan studeras ytterligare och det fanns det inte utrymme till i denna studie. Förhoppningsvis kan detta utredas ytterligare av SCB vid ett annat tillfälle. I denna studie testades en variant av en poängfunktion därför endast på KS-undersökningen.

Poängfunktionen med respektive stratummedelvärde som förväntat värde fungerade inte optimalt, men det var ändå intressant att testa teorin i praktiken och det visade att effektgranskning har stor potential. Så stor att den definitivt förtjänar att utredas vidare och att på sikt, när ett väl fungerade förväntat värde har hittats, ersätta dagens "gammalmodiga" sätt att granska. Dagens granskning är tänkt att upptäcka alla fel, men det finns ingen granskning som hittar alla fel och alla fel är inte ens möjliga att

eliminera. Resultaten av utvärderingen och testet visade att produktionsgranskning är nödvändigt och nu finns också vetskapen att det är en omöjlighet att hitta alla fel, därför är effektgranskning väl värt att överväga. Med effektgranskning kan omfattningen av produktionsgranskningen minska och produktionsgruppen får då tid att koncentrera sig på att utreda och rätta upp den granskning som verkligen har effekt på slutresultatet. Effektgranskning underlättar också granskarnas arbete jämfört med dagens befintliga granskningssystem. Dagens granskningssystem tar ingen hänsyn till vilken blankett som borde granskas i första hand av produktionsgruppen med avseende på att det arbetsställe som blanketten avser har störst inverkan på slutresultatet. Med effektgranskning uppstår aldrig detta prioriteringsproblem eftersom blanketterna tilldelas poäng och rangordnas efter dessa.

I praktiken skulle inte valet ha gjorts att använda stratummedelvärdena som förväntade värden i KS-undersökningen. Resultaten av testet visade att om respektive stratummedelvärde används som förväntat värde i praktiken skulle en ohanterlig mängd blanketter falla ut till manuell granskning. Som utläses av diagram 24 skulle materialet övergranskas om detta val av förväntat värde gjordes vid en implementering av ett nytt granskningssystem. Med större resurser, mer tid och kompetens går det att ta fram förväntade värden som är bättre anpassade för undersökningen och som ger bättre resultat än de som visas i kapitel 5. Vid implementering av effektgranskning i den australiensiska undersökningen Australian Survey of Average Weekly Earnings (ASAW) [10] minskade omfattningen av produktionsgranskningen med drygt 50 procent. Det är möjligt att det inte går att åstadkomma en lika stor minskning av produktionsgranskningen i KS/KV-undersökningarna. Resultaten från denna studie visar dock att produktionsgranskningen i KS/KV undersökningarna definitivt går att minska i omfattning, men det är svårt att sja om hur mycket den går att minska.

Det finns inte mycket litteratur att tillgå på området effektgranskning och inte många tidigare utförda studier. Teorin i sig är inte svår, men att tillämpa den i praktiken och uppnå ett acceptabelt resultat visade sig vara desto mer problematiskt. Det visade sig att det inte är enkelt att hitta ett bra förväntat värde och att det inte är lätt att bestämma tröskelvärdet. Trots att KS och KV har samlats in under flera år och att mycket data finns tillgängligt gavs här prov på hur urvalsproblematiken kan försvåra tillämpning av effektgranskning i praktiken. Det faktum att säsong- och trendkomponenter inte används i dagsläget per variabel försvårade också möjligheten att kunna använda ett "bra" förväntat värde i denna studie. Det hade varit intressant att även använda kvartal 1:s värden som förväntade värden avseende de objekt som lämnat in uppgifter både avseende kvartal 1 och 2 för att se om misstanken att de skulle fungera bättre som förväntade värden än vad respektive stratummedelvärde gjorde stämmer. Detta rymdes dock inte inom ramen för denna studie, men är förhoppningsvis något som rymms inom den fallstudie som ingår i Lotta-projektet. Testet som utfördes gav trots svårigheter och problem en bild av hur mycket mindre granskning som skulle kunna utföras och hur mycket mer effektiv den granskning som skulle utföras skulle kunna vara. Detta motiverar verkligen till att fortsätta med detta arbete!

## Referenser

Förutom de i studien nämnda referenserna, ingår nedan några som är av mer allmän karaktär för det aktuella ämnesområdet.

- [1] SCB (2006). *Företagsdatabasen (FDB)*.  
<[http://www.scb.se/templates/Uppgiftslamnare\\_99134.asp](http://www.scb.se/templates/Uppgiftslamnare_99134.asp)>
- [2] SCB (2003). *Systemet för Samordnade Urval*. Bakgrundsfakta till ekonomisk statistik 2003:3, SCB, Örebro.
- [3] SCB (2006). *Kortperiodisk Sysselsättningsstatistik*.  
<[http://www.scb.se/templates/Product\\_7820.asp](http://www.scb.se/templates/Product_7820.asp)>
- [4] SCB (2006). *Konjunkturstatistik över Vakanser*.  
<[http://www.scb.se/templates/Product\\_8132.asp](http://www.scb.se/templates/Product_8132.asp)>
- [5] SCB (2003). *Standard för svensk näringsgrensindelning (SNI 2002)*. Meddelanden i samordningsfrågor 2003:2, SCB, Örebro.
- [6] SCB (2003). *Regionala indelningar i Sverige*. Meddelanden i samordningsfrågor 2003:1, SCB, Örebro.
- [7] SCB (2005). *Beskrivning av statistiken*.  
<[http://www.scb.se/statistik/AM/AM0701/\\_dokument/AM0701\\_BS\\_2005.doc](http://www.scb.se/statistik/AM/AM0701/_dokument/AM0701_BS_2005.doc)>
- [8] Hidiroglou, M.A., and Berthelot, J.-M. (1986). Statistical editing and imputation for periodic business surveys. *Survey Methodology*, No. 12, 73-83.
- [9] SCB (2002). *Guide till granskning*. CBM 2002:1, SCB, Örebro
- [10] Lawrence D. and McDavitt C. (1994). Significance Editing in the Australian Survey of Average Weekly Earnings. *Journal of Official Statistics*, Vol. 10. No. 4, 437-447.
- [11] Lawrence D. and McKenzie R. (2000). The General Application of Significance Editing. *Journal of Official Statistics*, Vol. 16, No. 3, 243-253.
- [13] Andersson C. and Nordberg L. (1998). *A User's Guide to CLAN 97 – a SAS-program for computation of point- and standard error estimates in sample surveys*. SCB, Örebro.
- Dahmström K. (2005). *Från datainsamling till rapport – att göra en statistisk undersökning*. Fjärde upplagan, Studentlitteratur, Lund.
- Farwell K. (2005). *The General Application of Significance Editing to Economic Collections*. Australian Bureau of Statistics.
- Hedlin D. (2003). Score Functions to Reduce Business Survey Editing at the U.K. Office for National Statistics. *Journal of Official Statistics*, Vol. 19, No. 2, 177-199.
- Höglund Davila E. (1989). Studier av Hidiroglou-Berthelots metod tillämpad på data, dels från order- och leveransstatistiken, dels från löne- och sysselsättningsstatistiken för industriarbetare. *F-metod*, Nr. 23, SCB.

## **Opublicerade källor i PowerPoint-format**

[12] Nordberg L. (2006). *Significance Editing*. SCB, Örebro.

Farwell K. (odaterad). *Significance Editing for a Variety of Survey Situations*. Australian Bureau of Statistics.

## Appendix 1 - Blankettutseenden

Här följer de blankettvarianter som är relevanta i denna studie avseende KS- och KV-undersökningen.

Blankett A1.1. KS-lång.

Kortperiodisk sysselsättningsstatistik för privat sektor

2 +

Inlämns till SCB senast 2006-05-19

Ni kan lämna in via Internet: [www.insamling.scb.se](http://www.insamling.scb.se)

Användarnamn:

Lösenord:

AM KSP 2006 04

Arbetsställe

Organisationsnr och cennr

1

<b>1. Hur många verksamma företagare fanns vid ovanstående arbetsställe 2006-04-12?</b> Här menas verksamma företagare/delägare i enskild firma, enkelt bolag, HB och KB. Räkna även med eventuella familjemedlemmar som arbetar i företaget utan formell anställning.		Totalt	Män	Kvinnor
		<input type="text"/>	<input type="text"/>	<input type="text"/>
<b>2. Hur många anställda hade ovanstående arbetsställe 2006-04-12?</b> Här inkluderas även verksamma aktiebolagsägare samt delägare i ekonomiska föreningar. Även frånvarande anställda skall ingå. Om arbetsstället saknar anställda, fyll i totalrutan med 0, skriv kommentar och återvänd blanketten till SCB.		Totalt	Tillsvidareanställda	Visstidsanställda
		<input type="text"/>	<input type="text"/>	<input type="text"/>
<b>3. Hur många var frånvarande hela dagen 2006-04-12?</b> på grund av ...		Totalt	Sjukdom	Semester
		<input type="text"/>	<input type="text"/>	<input type="text"/>
			Övrig orsak	<input type="text"/>
<b>4. Hur många personer nyanställdes på arbetsstället under april månad?</b>		Totalt	Tillsvidareanställda	Visstidsanställda
		<input type="text"/>	<input type="text"/>	<input type="text"/>
<b>5. Hur många personer slutade på arbetsstället under april månad?</b>		Totalt	Tillsvidareanställda	Visstidsanställda
		<input type="text"/>	<input type="text"/>	<input type="text"/>

Första kontaktperson (TEXTA)	Telefon (inkl riktnr)	Mobilnr	Telefax (inkl riktnr)
Andra kontaktperson (TEXTA)	Telefon (inkl riktnr)	Mobilnr	Telefax (inkl riktnr)
E-post till första kontaktperson	E-post till andra kontaktperson		
Kommentarer			SCB:s kod

UPPGIFTSPLIKT enligt SFS 2001:99 och SGB-FS 1992:28.

Här lämnade uppgifter är sekretesskyddade enligt 9 kap 4 § sekretesslagen (SFS 1980:100)

## Kortperiodisk sysselsättningsstatistik för privat sektor

1

Insänds till SCB senast 2006-05-19

Ni kan lämna in via Internet: [www.insamling.scb.se](http://www.insamling.scb.se)  
 Användarnamn: Lösenord:

AM KSP 2006 04

Arbetsställe

Organisationnr och övrigt

1

## 1. Hur många verksamma företagare fanns vid ovanstående arbetsställe 2006-04-12?

Här menas verksamma företagare/delägare i enskild firma, enkelt bolag, HB och KB. Räkna även med eventuella familjemedlemmar som arbetar i företaget utan formell anställning.

Totalt	Män	Kvinnor
<input type="text"/>	<input type="text"/>	<input type="text"/>

## 2. Hur många anställda hade ovanstående arbetsställe 2006-04-12?

Här inkluderas även verksamma aktiebolagsägare samt delägare i ekonomiska föreningar. Även frånvarande anställda skall ingå. Om arbetsstället saknar anställda, fyll i totalrutan med 0, skriv kommentar och återsänd blanketten till SCB.

Totalt	Tillsvidareanställda	Visstidsanställda
<input type="text"/>	<input type="text"/>	<input type="text"/>
	<input type="text"/>	<input type="text"/>

Första kontaktperson (TEXTA)	Telefon (inkl riktnr)	Mobilnr	Telefax (inkl riktnr)
Andra kontaktperson (TEXTA)	Telefon (inkl riktnr)	Mobilnr	Telefax (inkl riktnr)
E-post till första kontaktperson	E-post till andra kontaktperson		
Kommentarer			SCB:s kod

UPPGIFTSPLIKT enligt SFS 2001:99 och SGB-FS 1992:28

Här lämnade uppgifter är sekretesskyddade enligt 9 kap 4 § sekretesslagen (SFS 1980:100)



Blankettutgivare  
**Statistiska centralbyrån**  
 Statistics Sweden  
 Avdelningen för näringsliv och arbetsmarknad

Postadress SCB:s kontaktperson  
 701 89 Örebro

Gruppnr 018 - 17 88 10  
 Telefax 018 - 17 88 08

## Konjunkturstatistik över vakanser i privat sektor

4

+

Insänds till SCB senast 2006-05-19

Ni kan lämna in via Internet: [www.insamling.scb.se](http://www.insamling.scb.se)

Användarnamn:

Lösenord:

AM KVP 2006 04

Arbetsställe

Organisationer och adress

1

## 1. Hade arbetsstället några lediga jobb 2006-04-12?

Vi menar jobb där extern rekrytering påbörjats men ingen ännu hade anställts 2006-04-12. Jobbet räknas som ledigt oavsett om det under rekryteringsperioden är bemannat eller obemannat.

 Ja

 Nej  Tack för din medverkan!  
Skriv ev. kommentar och  
sänd blanketten till SCB.

## 2. Hur många?

lediga jobb

## 3. Hur fördelar sig de lediga jobben som noterats i fråga 2 på nedanstående kategorier?

a) Obemannat trots att bemanning önskades 2006-04-12

b) Obemannat och bemanning önskades inte förrän efter 2006-04-12

c) Bemannat med medarbetare som ska sluta

d) Bemannat med vikarier/tillfälligt anställda

e) Bemannat med konsulter/personal från bemanningsföretag

Kontrollera att fråga 2 = 3a + 3b + 3c + 3d + 3e

## Tack för din medverkan! - Återsänd till SCB!

Första kontaktperson (TEXTA)	Telefon (inkl riktnr)	Mobilnr	Telefax (inkl riktnr)
Andra kontaktperson (TEXTA)	Telefon (inkl riktnr)	Mobilnr	Telefax (inkl riktnr)
E-post till första kontaktperson	E-post till andra kontaktperson		
Kommentarer			SCB:s kod

UPPGIFTSPLIKT enligt SFS 2001:99 och SCS-FS 2003:15

Här lämnade uppgifter är sekretesskyddade enligt 9 kap 4 § sekretesslagen (SFS 1980:100)



Blankettgivare  
**Statistiska centralbyrån**  
Statistiska Swedens  
Avdelningen för näringsliv och arbetsmarknad

Postadress SCB:s kontaktperson  
701 88 Örebro

Gruppnr 018 - 17 88 10  
Telefax 018 - 17 88 03

## Appendix 2 – Tabeller över de koder som används i KS/KV

Tabell A2.1 visar de styrkoder som används i KS- och KV-undersökningarna.

*Tabell A2.1. Styrkoder i KS- och KV-undersökningen.*

<b>Styrkod</b>	<b>Klartext</b>
00	Ej inkommen blankett
02	Granskad men ej godkänd
03	Granskad och godkänd
05	Granskad och godkänd trots fel
06	Granskad, godkänd trots fel och partiellt imputerad för frånvaro
07	Granskad, godkänd och partiellt imputerad för frånvaro
08	Godkänd, men ej korrekt RFV-del
09	Inkommen, ej åtgärdad
10	Anstånd
11	Anstånd, ej skarp påminnelse
20	Vilande, tillfälligt ingen drift
21	Inga sysselsatta
22	Konkurs
23	Företag med < 10 anställda, nedlagda via systemet med automatik
24	Helt nedlagt arbetsställe
26	Bortslumpade för kvartal 1, 2006
60	Vägrare
61	Sjukdom, dataproblem o.dyl.
70	Kan inte redovisa på arbetsställenivå - flera arbetsställen ingår
71	Redovisas av kod 70-blankett
72	Arbetsstället delat, kan inte redovisa arbetare och tjänstemän på samma blankett
73	Kan ej redovisa på arbetsställenivå ett arbetsställe ingår
80	Ringkod
89	Dataliste- eller datafilföretag



Tabell A2.2 visar de felkoder som används i KS- och KV-undersökningarna.

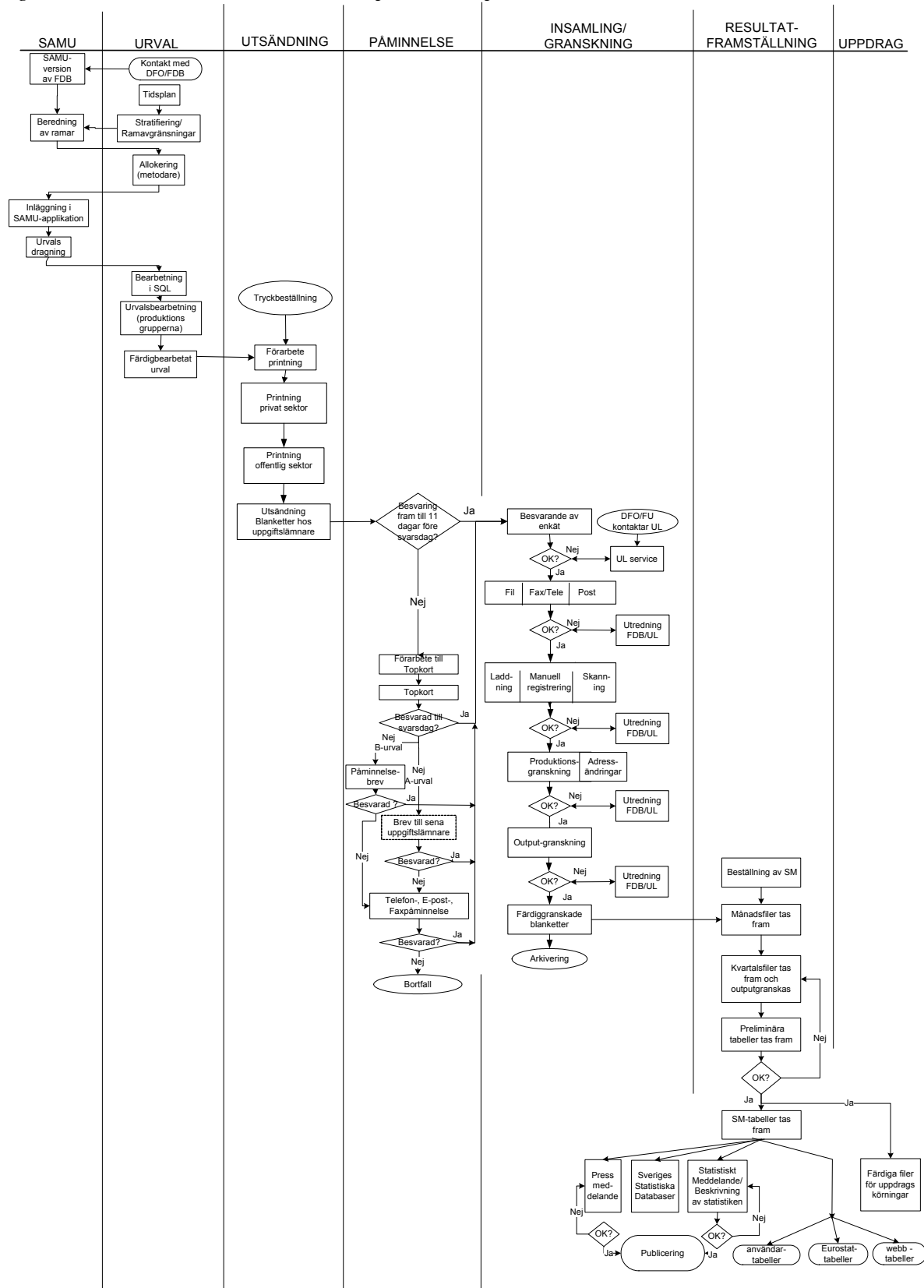
Tabell A2.2. Felkoder i KS- och KV-undersökningen.

<b>Felkod</b>	<b>Klartext</b>
01	Detta är ett aktiebolag, det ska inte ha företagare
02	Detta är ett HB/KV, det ska ha företagare
03	Angivet antal företagare stämmer inte överens med summan angiven i totalrutan
04	Tillsvidareanställda saknas
08	Arbetsstället saknar anställda
09	Totala antalet anställda skiljer sig jämfört med CFAR
10	Angivet antal anställda stämmer inte överens med summan angiven i totalrutan
11	Totala frånvaron för män överstiger antalet anställda män
12	Totala frånvaron för kvinnor överstiger antalet anställda kvinnor
13	Sjukfrånvaron överstiger 30% av antal tillsvidareanställda och antal tillsvidareanställda överstiger 20 stycken
14	Angivet antal frånvarande stämmer inte överens med summan angiven i totalrutan
15	Sjukfrånvaro saknas, antal tillsvidareanställda överstiger 150 stycken
16	Angivet antal nyanställda stämmer inte överens med summan angiven i totalrutan
17	Angivet antal avgångna stämmer inte överens med summan angiven i totalrutan
18	Antalet nyanställda är fler än 6 och mer än 25% av antalet anställda
19	Antalet avgångna är fler än 6 och mer än 25% av antalet anställda
20	Personalomsättning saknas och antalet tillsvidareanställda är fler än 200
21	Personnr felaktigt eller saknas
22	Sjukfrom har felaktigt datumformat, är efter dagens datum eller sjukfrom saknas när sjuktom finns.
23	Sjuktom har felaktigt datumformat, är efter dagens datum eller sjuktom saknas när sjukfrom finns.
24	Sjukfrom ligger efter sjuktom eller antalet dagar mellan sjukfrom och sjuktom är fler än 21
25	Antal hel- eller deldagar är fler än perioden mellan sjukfrom och sjuktom eller är fler än 21
26	Sjukperioderna för samma person går i varandra
27	Datum samt antal dagar saknas
28	Antal dagar saknas
29	Datum saknas
30	Sjuktomdatum ligger inte i mätmånaden
31	Sjukperioden överstiger 21 dagar
32	Sjukavisering saknas och det är fler än 19 anställda enl CFAR
33	Antalet angivna personnummer är fler än angivet antalet anställda
34	Inga sjuka angivna i KS-delen när det finns sjukaviseringsperiod på detta datum
35	Antalet dagar är 21 eller flera utan att försäkringskassan är för kryssad
36	När sjuka är angivna i KS-delen men inga personnummer är angivna i RFV-delen
37	Perioden mellan from och tom är mindre än 21 dagar men försäkringskassan är förkryssad
38	Statlig sektor har övriga anställda
50	Totalt är inte summan av fråga 3
52	Många lediga jobb jämfört med Syssa
53	Totalt antal lediga jobb är noll när Syssa är större än 199
54	Antal vakanser och vakanskrysset är inkonsekventa
55	Denna vakansblankett är oifylld
70	Dubblett finns

## Appendix 3 – Flödesschema över produktionsprocessen

Figur A3.1 visar ett flödesschema där varje process kan följas från första steget i urvalsprocessen fram till publicering och uppdragsprocessen avseende KS- och KV-undersökningarna.

Figur A3.1. Flödesschema över KS/KV:s produktionsprocess.



## Appendix 4 – Ytterligare processdataresultat

Tabell A4.1 och A4.2 anger antalet blanketter som föll ut med respektive felkod i KS- och KV-undersökningen under kvartal 2 2006.

*Tabell A4.1. Antal blanketter per felkod avseende kvartal 2 år 2006, KS.*

<b>Felkod</b>	<b>Antal</b>
01	33
02	181
03	55
04	1 493
08	806
09	569
10	1 155
11	134
12	92
13	31
14	418
15	192
16	145
17	132
18	28
19	20
20	201
34	1
36	1
<b>Totalt</b>	<b>5 687</b>

*Tabell A4.2. Antal blanketter per felkod avseende kvartal 2 år 2006, KV.*

<b>Felkod</b>	<b>Antal</b>
50	708
52	111
53	1 424
54	604
55	324
<b>Totalt</b>	<b>3 171</b>

## Appendix 5 - Ytterligare utvärderingsresultat av KS

Diagram A5.1-A5.18 visar storleken på pseudo biasen i procent på Y-axeln för de olika gränserna på X-axeln. Diagrammen är konstruerade genom tillämpning av den metod som presenteras i kapitel 3. I diagrammen används inte den absoluta pseudo biasen, detta för att visa att ändringar av inkomna värden utförs i båda riktningarna.

Diagram A5.1. Pseudo bias på totalnivå för variabeln visstidsanställda män.

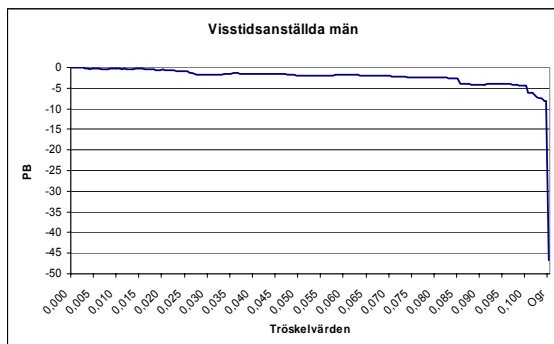


Diagram A5.2. Pseudo bias på totalnivå för variabeln frånvarande sjuka män.

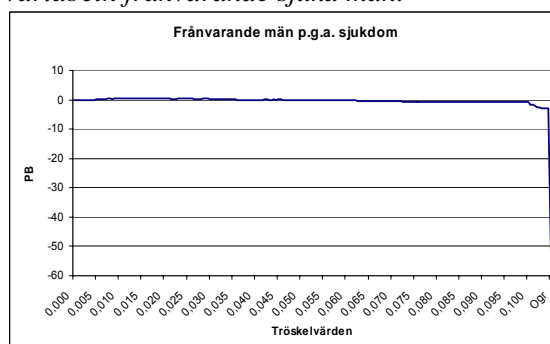


Diagram A5.3. Pseudo bias på totalnivå för variabeln visstidsanställda kvinnor.

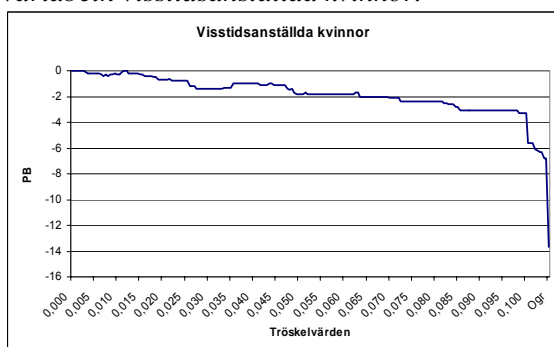


Diagram A5.4. Pseudo bias på totalnivå för variabeln frånvarande sjuka kvinnor.

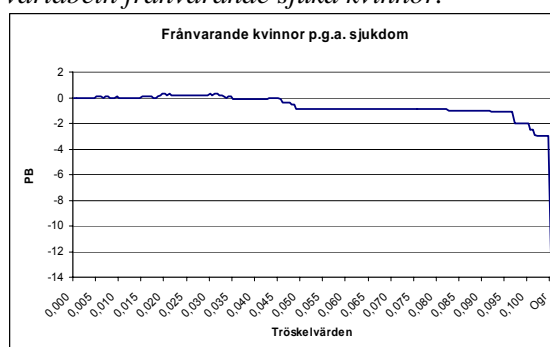


Diagram A5.5. Pseudo bias på totalnivå för variabeln verksamma företagare män.

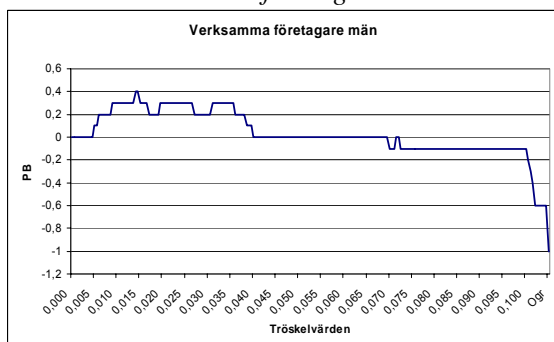


Diagram A5.6. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester män.

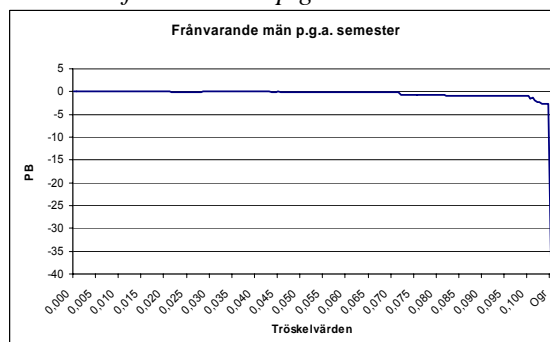


Diagram A5.7. Pseudo bias på totalnivå för variabeln verksamma företagare kvinnor.

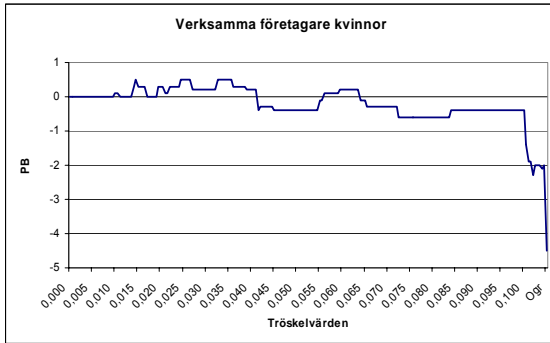


Diagram A5.8. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda män.

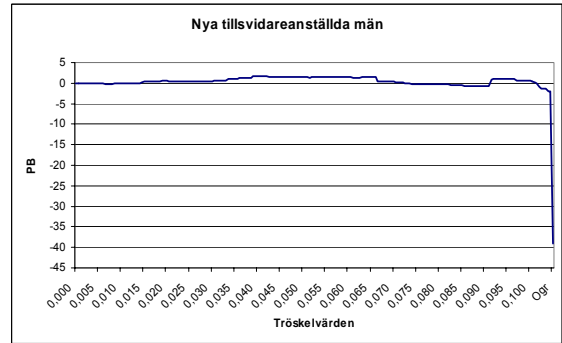


Diagram A5.9. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester kvinnor.

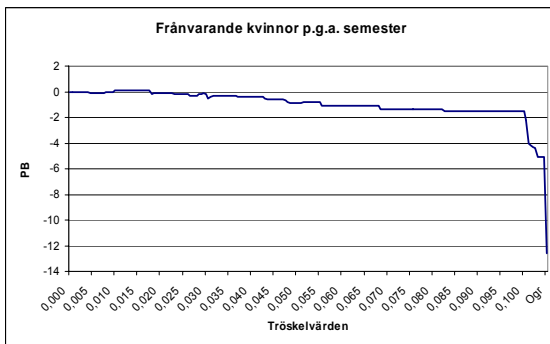


Diagram A5.10. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda kvinnor.

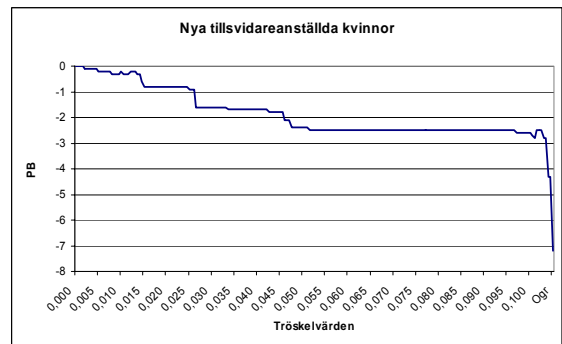


Diagram A5.11. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak män.

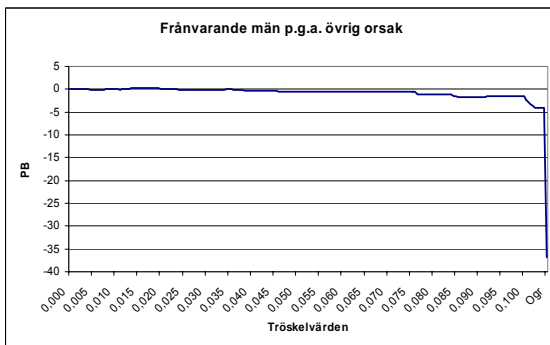


Diagram A5.12. Pseudo bias på totalnivå för variabeln nya visstidsanställda män.

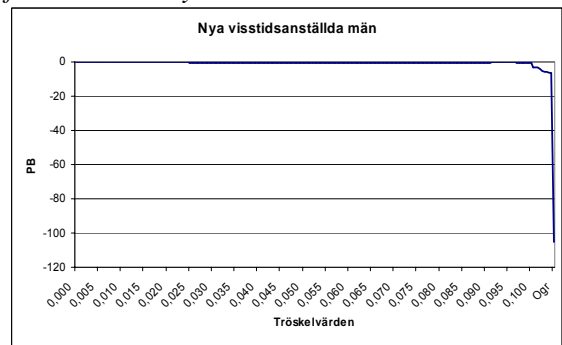


Diagram A5.13. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak kvinnor.

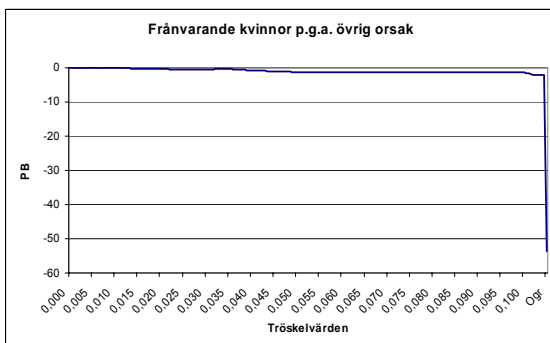


Diagram A5.14. Pseudo bias på totalnivå för variabeln nya visstidsanställda kvinnor.

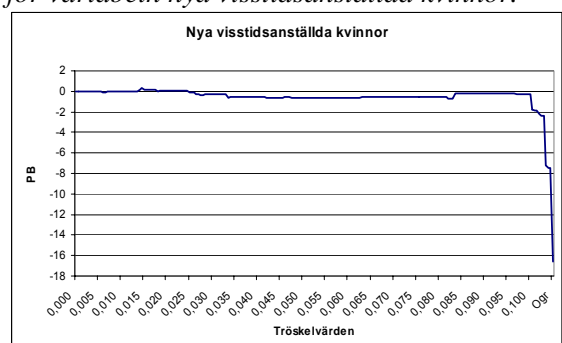


Diagram A5.15. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda män.

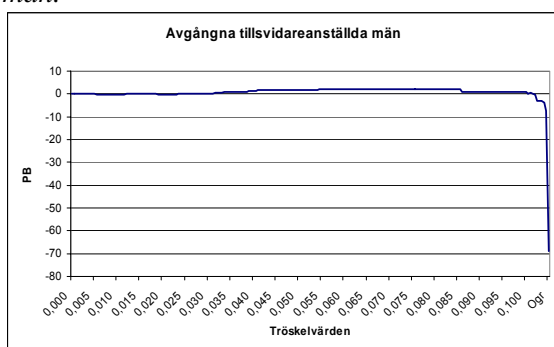


Diagram A5.16. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda män.

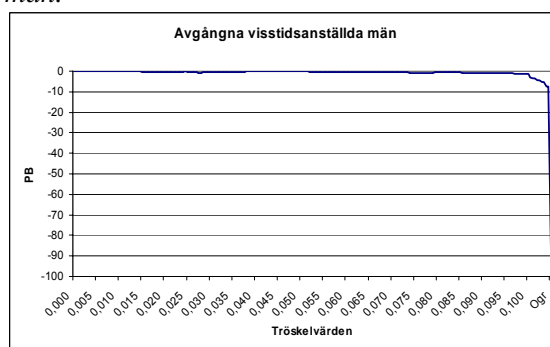


Diagram A5.17. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda kvinnor.



Diagram A5.18. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda kvinnor.

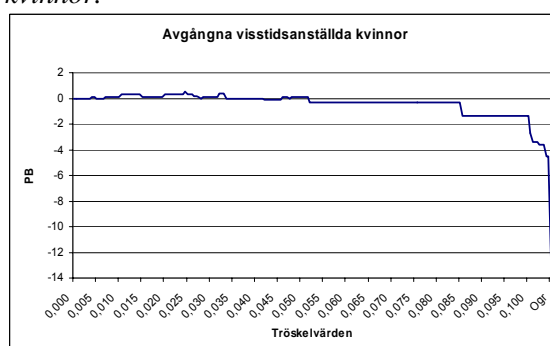


Diagram A5.19-A5.36 visar storleken på pseudo biasen i procent på Y-axeln för de olika gränserna på X-axeln, men endast för intervallet av gränser mellan 0,002-0,02.

Diagrammen är konstruerade genom tillämpning av den metod som presenteras i kapitel 3. I diagrammen används inte den absoluta pseudo biasen, detta för att visa att ändringar av inkomna värden utförs i båda riktningarna.

Diagram A5.19. Pseudo bias på totalnivå för variabeln visstidsanställda män.

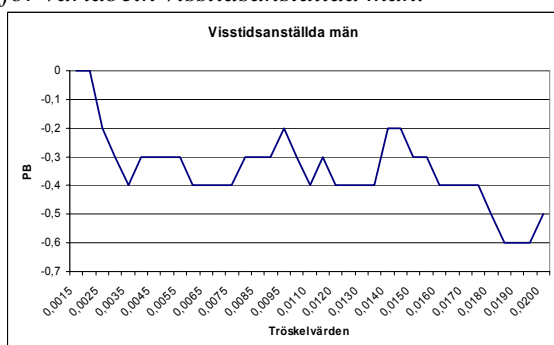


Diagram A5.20. Pseudo bias på totalnivå för variabeln visstidsanställda kvinnor.

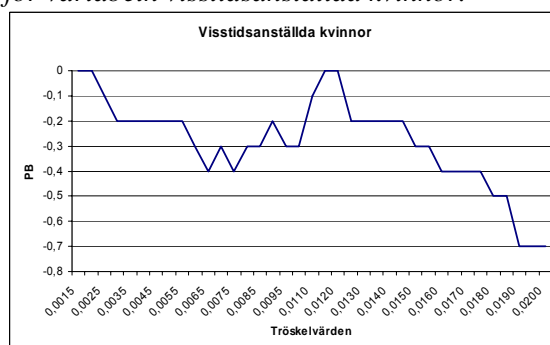


Diagram A5.21. Pseudo bias på totalnivå för variabeln verksamma företagare män.

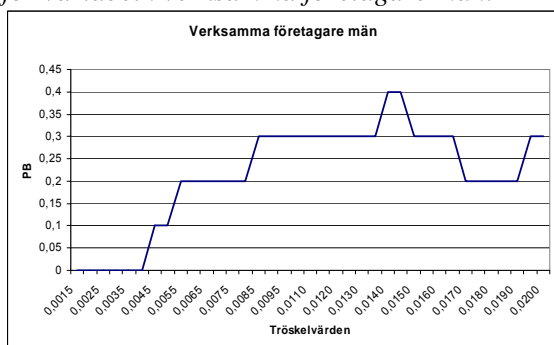


Diagram A5.22. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda män.

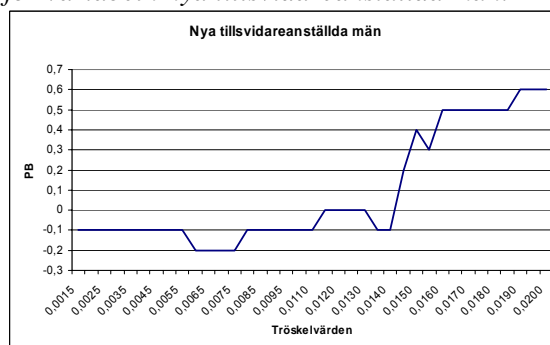


Diagram A5.23. Pseudo bias på totalnivå för variabeln verksamma företagare kvinnor.

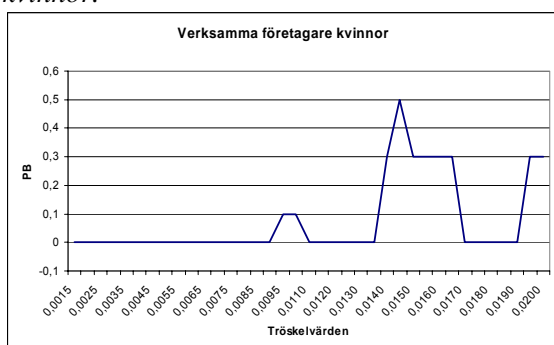


Diagram A5.24. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda kvinnor.



Diagram A5.25. Pseudo bias på totalnivå för variabeln frånvarande sjuka män.

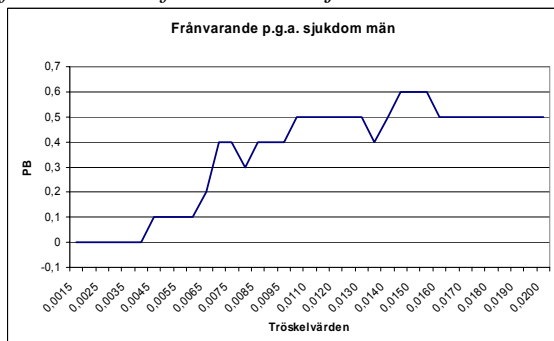


Diagram A5.26. Pseudo bias på totalnivå för variabeln nya visstidsanställda män.

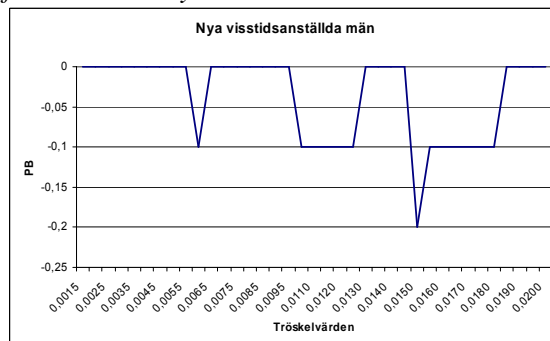


Diagram A5.27. Pseudo bias på totalnivå för variabeln frånvarande sjuka kvinnor.

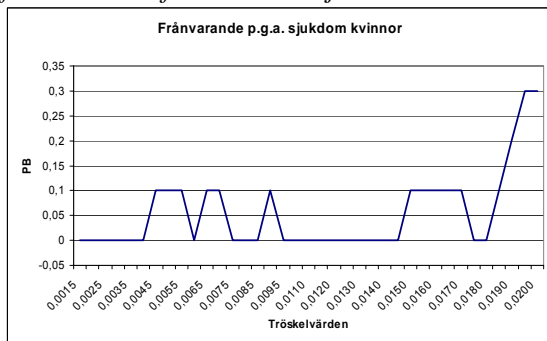


Diagram A5.28. Pseudo bias på totalnivå för variabeln nya visstidsanställda kvinnor.



Diagram A5.29. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester män.

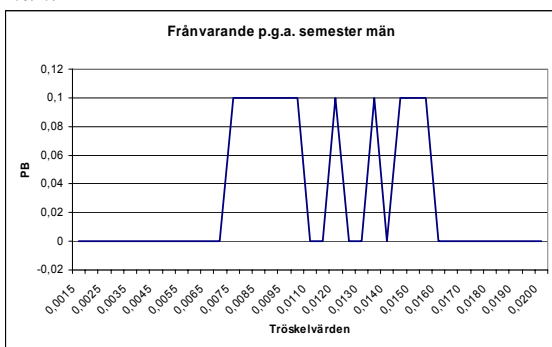


Diagram A5.30. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda män.

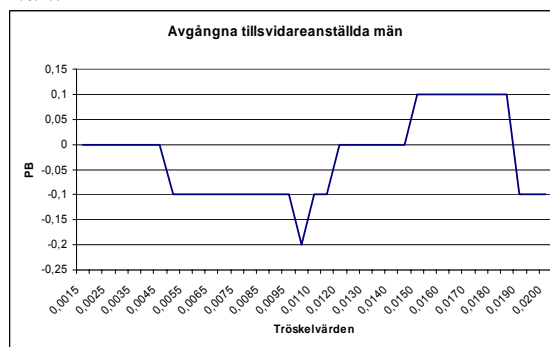


Diagram A5.31. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester kvinnor.

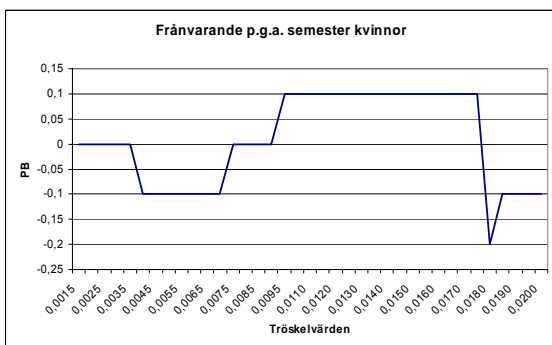


Diagram A5.32. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda kvinnor.

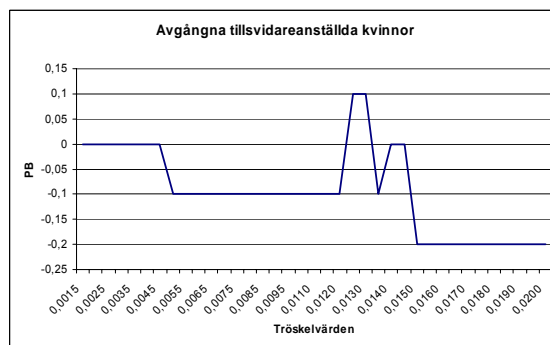


Diagram A5.33. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak män.

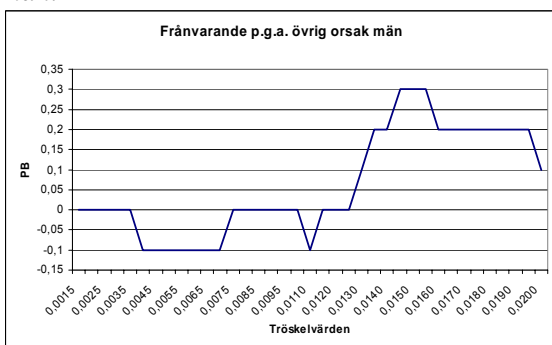


Diagram A5.34. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda män.

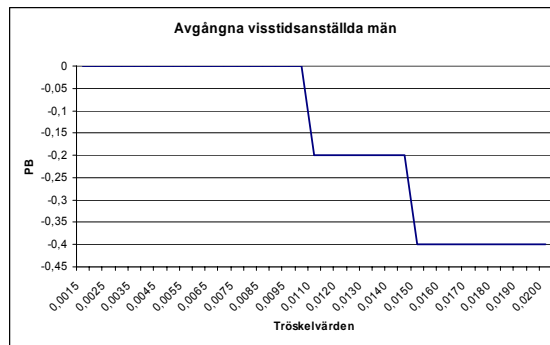


Diagram A5.35. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak kvinnor.

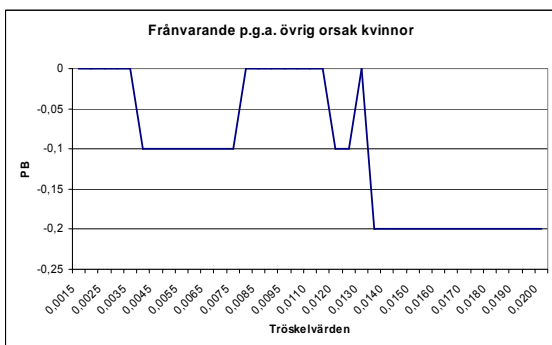


Diagram A5.36. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda kvinnor.





Tabell A5.1 visar punktskattningen för varje variabel per branschgrupp i KS-undersökningen avseende kvartal 2 år 2006.

Tabell A5.1. Punktskattning per branschgrupp för samtliga variabler i KS-undersökningen.

NGGr2	Verks. ftgare män	Verks. ftgare kvinnor	Tot. antal verks. ftgare	Tillsvidare- anställda män	Tillsvidare- anställda kvinnor	Visstids- anställda män	Visstids- anställda kvinnor
<b>Totalt</b>	191 242	14 382	205 624	1 464 633	796 048	185 465	180 682
<b>0100</b>	5 650	1 788	7 438	21 252	6 046	4 653	2 122
<b>1000</b>	21	40	60	6 447	876	682	174
<b>1500</b>	434	223	657	31 481	17 801	5 220	3 576
<b>1800</b>	127	125	252	4 588	3 653	419	326
<b>2000</b>	155 503	115	155 618	27 506	5 135	2 386	543
<b>2100</b>	1	1	1	27 317	7 504	1 950	872
<b>2200</b>	209	76	285	22 734	14 841	2 436	2 525
<b>2300</b>	23	24	47	22 995	15 971	1 551	1 214
<b>2500</b>	55	16	71	15 197	6 832	1 357	827
<b>2600</b>	53	36	89	11 713	2 795	1 115	337
<b>2700</b>	10	1	10	27 534	5 402	1 645	510
<b>2800</b>	344	155	498	57 442	11 296	4 046	1 162
<b>2900</b>	227	39	266	75 760	14 576	4 442	1 590
<b>31XX</b>	43	1	43	18 463	6 884	1 115	566
<b>3200</b>	1	1	1	16 897	6 817	553	277
<b>3300</b>	60	1	60	14 776	7 244	795	558
<b>3400</b>	31	1	31	56 440	14 640	3 302	1 088
<b>3500</b>	117	20	136	15 518	2 566	585	184
<b>3600</b>	173	38	211	26 957	14 062	1 437	708
<b>4000</b>	1	1	1	19 018	5 866	938	632
<b>4510</b>	2 760	345	3 105	86 110	7 861	10 324	850
<b>4530</b>	2 157	299	2 456	78 205	7 446	3 255	487
<b>5000</b>	1 383	237	1 620	54 019	12 237	4 352	2 820
<b>5100</b>	988	384	1 372	116 253	48 280	8 619	5 277
<b>5200</b>	2 324	2 124	4 448	44 397	66 259	7 262	16 591
<b>5210</b>	415	418	833	18 300	42 121	5 092	13 620
<b>5220</b>	604	616	1 220	3 799	6 283	618	1 661
<b>5500</b>	4 090	2 033	6 123	30 705	35 415	10 369	16 028
<b>6000</b>	4 334	593	4 927	80 128	13 215	15 053	2 310
<b>6100</b>	280	39	319	33 537	23 286	4 871	3 834
<b>6400</b>	1 636	303	1 939	38 043	21 614	6 479	3 872
<b>6500</b>	1	1	1	22 596	31 066	1 817	3 369
<b>6600</b>	113	1	113	10 776	11 419	508	838
<b>7000</b>	834	426	1 260	32 342	18 174	2 807	2 498
<b>7100</b>	137	100	236	6 484	1 697	747	463
<b>7200</b>	470	130	599	64 349	20 538	2 508	1 189
<b>7300</b>	21	1	21	5 788	4 109	1 561	1 101
<b>7400</b>	3 794	1 385	5 179	130 063	99 648	27 996	29 815
<b>8000</b>	165	148	313	13 437	33 226	3 955	7 120
<b>8500</b>	648	669	1 317	22 076	66 063	11 274	31 379
<b>9000</b>	660	1 068	1 728	10 194	10 702	2 027	2 034
<b>9100</b>	1	1	1	19 074	26 515	4 498	5 698
<b>9200</b>	351	372	724	23 924	18 071	8 848	8 039

Tabell A5.1. Fortsättning...

<b>NGGr2</b>	<b>Totalt antal anställda</b>	<b>Frånv. p.g.a. sjukdom män</b>	<b>Frånv. p.g.a. sjukdom kvinnor</b>	<b>Frånv. p.g.a. semester män</b>	<b>Frånv. p.g.a. semester kvinnor</b>
<b>Totalt</b>	2 624 950	40 696	34 443	36 371	23 487
<b>0100</b>	34 073	349	151	189	95
<b>1000</b>	8 178	168	21	153	25
<b>1500</b>	58 078	1 543	1 312	951	454
<b>1800</b>	8 985	152	135	73	61
<b>2000</b>	35 570	1 128	317	422	83
<b>2100</b>	37 642	1 109	473	1 096	301
<b>2200</b>	42 535	625	500	839	451
<b>2300</b>	41 731	646	700	858	668
<b>2500</b>	24 213	552	460	374	179
<b>2600</b>	15 960	518	166	265	56
<b>2700</b>	35 090	1 218	298	729	126
<b>2800</b>	73 946	1 591	639	814	200
<b>2900</b>	96 368	2 355	722	1 572	414
<b>31XX</b>	27 028	536	352	384	116
<b>3200</b>	24 544	413	447	664	257
<b>3300</b>	23 373	351	378	398	185
<b>3400</b>	75 471	2 511	1 008	992	261
<b>3500</b>	18 852	534	94	374	66
<b>3600</b>	43 163	2 050	1 982	527	315
<b>4000</b>	26 454	350	159	671	245
<b>4510</b>	105 144	2 084	97	981	90
<b>4530</b>	89 392	2 159	149	1 254	213
<b>5000</b>	73 428	1 256	340	1 102	358
<b>5100</b>	178 428	2 372	1 393	2 829	1 123
<b>5200</b>	134 489	765	3 056	1 154	2 238
<b>5210</b>	79 133	460	1 886	541	1 078
<b>5220</b>	12 361	54	242	10	113
<b>5500</b>	92 516	525	951	828	718
<b>6000</b>	110 706	2 426	652	1 769	352
<b>6100</b>	65 527	1 202	891	1 347	1 072
<b>6400</b>	70 009	1 642	1 593	1 930	1 148
<b>6500</b>	58 849	389	1 003	1 195	1 437
<b>6600</b>	23 541	205	379	473	449
<b>7000</b>	55 820	681	623	886	433
<b>7100</b>	9 391	86	58	104	57
<b>7200</b>	88 583	903	568	1 716	564
<b>7300</b>	12 559	66	123	255	171
<b>7400</b>	285 665	2 580	3 924	3 163	2 555
<b>8000</b>	57 738	359	1 196	247	908
<b>8500</b>	130 792	477	3 007	829	2 324
<b>9000</b>	24 956	301	254	239	109
<b>9100</b>	55 786	487	1 118	423	944
<b>9200</b>	58 881	519	629	752	473

Tabell A5.1. Fortsättning...

<b>NGGr2</b>	<b>Frånv. p.g.a. övrig orsak män</b>	<b>Frånv. p.g.a. övrig orsak kvinnor</b>	<b>Totalt antal frånvarande</b>	<b>Nya tillsvidareanst. män</b>	<b>Nya tillsvidareanst. kvinnor</b>
<b>Totalt</b>	40 101	48 945	224 089	12 342	6 658
<b>0100</b>	627	305	1 715	121	6
<b>1000</b>	287	57	710	21	5
<b>1500</b>	1 000	1 247	6 507	128	120
<b>1800</b>	77	163	661	11	23
<b>2000</b>	750	289	2 989	161	5
<b>2100</b>	886	486	4 351	67	18
<b>2200</b>	553	850	3 819	119	54
<b>2300</b>	995	1 670	5 537	177	115
<b>2500</b>	470	372	2 408	41	31
<b>2600</b>	307	106	1 418	61	15
<b>2700</b>	964	336	3 669	85	23
<b>2800</b>	1 133	508	4 885	436	46
<b>2900</b>	2 358	1 026	8 448	471	88
<b>31XX</b>	525	407	2 319	95	56
<b>3200</b>	726	597	3 104	132	49
<b>3300</b>	423	460	2 195	130	98
<b>3400</b>	2 884	1 386	9 041	242	79
<b>3500</b>	685	208	1 961	57	11
<b>3600</b>	1 149	849	6 873	238	47
<b>4000</b>	703	421	2 550	82	28
<b>4510</b>	1 014	147	4 414	1 546	91
<b>4530</b>	1 289	282	5 346	988	90
<b>5000</b>	1 154	630	4 840	465	92
<b>5100</b>	2 947	2 579	13 243	656	423
<b>5200</b>	1 651	5 058	13 923	377	505
<b>5210</b>	619	2 603	7 188	84	264
<b>5220</b>	148	505	1 072	56	100
<b>5500</b>	831	2 082	5 935	332	362
<b>6000</b>	1 939	539	7 678	709	104
<b>6100</b>	1 017	1 946	7 476	209	117
<b>6400</b>	702	911	7 926	238	208
<b>6500</b>	907	2 775	7 706	232	172
<b>6600</b>	273	778	2 558	98	98
<b>7000</b>	499	939	4 061	103	150
<b>7100</b>	85	113	504	104	25
<b>7200</b>	1 495	1 320	6 565	702	190
<b>7300</b>	183	200	998	42	31
<b>7400</b>	3 186	5 752	21 160	2 039	1 623
<b>8000</b>	281	2 405	5 396	46	66
<b>8500</b>	537	2 996	10 216	184	477
<b>9000</b>	205	580	1 688	33	145
<b>9100</b>	465	1 095	4 533	101	176
<b>9200</b>	1 168	966	4 506	123	232

Tabell A5.1. Fortsättning...

<b>NGGr2</b>	<b>Nya visstids- anställda män</b>	<b>Nya visstids- anställda kvinnor</b>	<b>Totalt antal nyanställda</b>	<b>Avgångna tillsvidare- anst. män</b>	<b>Avgångna tillsvidare- anst. kvinnor</b>
<b>Totalt</b>	39 605	29 456	88 060	12 578	7 203
<b>0100</b>	1 184	498	1 809	205	22
<b>1000</b>	312	90	428	31	5
<b>1500</b>	1 688	1 147	3 083	313	191
<b>1800</b>	80	61	175	19	28
<b>2000</b>	539	95	799	202	42
<b>2100</b>	801	360	1 246	169	44
<b>2200</b>	539	611	1 324	219	129
<b>2300</b>	529	330	1 151	229	104
<b>2500</b>	368	215	655	135	45
<b>2600</b>	308	98	482	98	39
<b>2700</b>	741	258	1 106	147	22
<b>2800</b>	950	286	1 718	358	53
<b>2900</b>	1 258	398	2 215	514	106
<b>31XX</b>	370	204	726	194	26
<b>3200</b>	210	132	522	169	53
<b>3300</b>	147	88	463	135	74
<b>3400</b>	933	373	1 627	377	107
<b>3500</b>	132	36	235	93	27
<b>3600</b>	357	196	838	181	113
<b>4000</b>	303	151	564	49	32
<b>4510</b>	3 339	179	5 155	587	85
<b>4530</b>	983	80	2 141	716	115
<b>5000</b>	728	230	1 514	519	101
<b>5100</b>	2 222	977	4 278	902	385
<b>5200</b>	873	2 397	4 153	518	610
<b>5210</b>	1 108	2 239	3 695	200	406
<b>5220</b>	162	386	704	101	76
<b>5500</b>	2 076	3 419	6 190	491	562
<b>6000</b>	1 803	519	3 136	639	79
<b>6100</b>	904	735	1 965	252	166
<b>6400</b>	3 303	2 150	5 899	575	292
<b>6500</b>	234	545	1 183	115	131
<b>6600</b>	57	143	396	59	63
<b>7000</b>	941	674	1 868	131	102
<b>7100</b>	217	16	363	47	2
<b>7200</b>	241	94	1 227	765	159
<b>7300</b>	119	59	250	58	48
<b>7400</b>	4 743	3 383	11 788	1 588	1 300
<b>8000</b>	204	453	770	59	124
<b>8500</b>	942	2 647	4 250	190	649
<b>9000</b>	578	254	1 010	72	37
<b>9100</b>	662	732	1 672	82	260
<b>9200</b>	1 417	1 516	3 289	75	192

Tabell A5.1. Fortsättning...

<b>NGGr2</b>	<b>Avgångna visstidsanst. män</b>	<b>Avgångna visstidsanst. kvinnor</b>	<b>Totalt antal avgångna</b>
<b>Totalt</b>	13 123	12 945	45 849
<b>0100</b>	288	379	895
<b>1000</b>	27	6	69
<b>1500</b>	285	196	985
<b>1800</b>	28	19	93
<b>2000</b>	215	28	486
<b>2100</b>	60	32	306
<b>2200</b>	234	279	860
<b>2300</b>	102	55	490
<b>2500</b>	47	41	268
<b>2600</b>	47	14	197
<b>2700</b>	56	28	254
<b>2800</b>	248	13	671
<b>2900</b>	338	88	1 047
<b>31XX</b>	67	46	334
<b>3200</b>	48	30	300
<b>3300</b>	32	31	273
<b>3400</b>	108	47	639
<b>3500</b>	48	5	174
<b>3600</b>	66	76	435
<b>4000</b>	79	37	197
<b>4510</b>	603	41	1 315
<b>4530</b>	217	61	1 108
<b>5000</b>	171	62	853
<b>5100</b>	649	425	2 361
<b>5200</b>	582	1 195	2 904
<b>5210</b>	255	320	1 182
<b>5220</b>	50	91	317
<b>5500</b>	1 405	1 420	3 879
<b>6000</b>	581	26	1 325
<b>6100</b>	317	223	958
<b>6400</b>	1 481	981	3 328
<b>6500</b>	95	188	528
<b>6600</b>	11	32	165
<b>7000</b>	206	120	559
<b>7100</b>	11	3	63
<b>7200</b>	142	50	1 116
<b>7300</b>	36	19	161
<b>7400</b>	1 698	1 638	6 223
<b>8000</b>	385	531	1 100
<b>8500</b>	368	2 717	3 924
<b>9000</b>	325	133	567
<b>9100</b>	137	219	698
<b>9200</b>	975	1 002	2 243

Redovisningen per branschgrupp är den redovisningsnivå som prioriteras högst av ämnesenheten NA/FRS. Det är därför inte tillräckligt kvalitetssäkrande att bestämma cut off-gränsen,  $\alpha$ , utifrån storleken på pseudo biasen på totalnivå. Den bör i det här fallet bestämmas utifrån hur stor pseudo bias ämnesenheten är villig att acceptera på branschgruppsnivå. Tabell A5.2 visar hur stor pseudo biasen skulle bli per branschgrupp i KS-undersökningen om den valda gränsen sätts till 0,01.

Tabell A5.2. Pseudo biasen per branschgrupp i KS då gränsen är satt till 0,01.

NGGr2	Verk- samma ftgare män	Verk- samma ftgare kvinnor	Tot. antal verk- samma ftgare	Tills- vidare- anställda män	Tills- vidare- anställda kvinnor
<b>Totalt</b>	0,3	0,1	0,3	1,5	0,8
<b>0100</b>	1,8	0	<b>2,3</b>	<b>3,3</b>	0,5
<b>1000</b>	0	0	0	0,3	0
<b>1500</b>	0,7	0	0,5	0,1	0,3
<b>1800</b>	0	0	0	0,5	0,2
<b>2000</b>	0	0	0	<b>4,1</b>	<b>2,2</b>
<b>2100</b>	0	0	0	-0,3	-0,9
<b>2200</b>	0,5	0	0,4	1,1	0,4
<b>2300</b>	0	0	0	0,3	-0,1
<b>2500</b>	0	0	0	-0,2	0,1
<b>2600</b>	0	0	0	-0	-0,2
<b>2700</b>	0	0	0	0	0
<b>2800</b>	0	0	0	1,6	0,6
<b>2900</b>	0	0	0	1,2	0,4
<b>31XX</b>	0	0	0	0,5	-0,6
<b>3200</b>	0	0	0	0,3	0,2
<b>3300</b>	0	0	0	1,5	1,8
<b>3400</b>	0	0	0	0	-0,3
<b>3500</b>	0	0	0	1,1	0
<b>3600</b>	0	0	0	1,2	1,8
<b>4000</b>	0	0	0	0,1	0,4
<b>4510</b>	<b>3,9</b>	0	<b>3,5</b>	<b>3</b>	0,9
<b>4530</b>	0	0	0	<b>2,2</b>	0,3
<b>5000</b>	0	0	0	<b>2,1</b>	0,9
<b>5100</b>	0	0	0	1,8	1,3
<b>5200</b>	<b>2,6</b>	0,9	1,4	<b>3,3</b>	<b>2,1</b>
<b>5210</b>	0	0	0	1,6	1,3
<b>5220</b>	0	0	0	1,8	0,7
<b>5500</b>	<b>3,3</b>	0	<b>2,2</b>	1,8	1,2
<b>6000</b>	<b>2,9</b>	0	<b>2,5</b>	1,6	<b>2,8</b>
<b>6100</b>	0	0	0	<b>2,2</b>	1,9
<b>6400</b>	0,5	0	0,4	0,9	0,1
<b>6500</b>	0	0	0	-0,7	-0,6
<b>6600</b>	0	0	0	1,7	2
<b>7000</b>	0	0	0	<b>2,7</b>	1,3
<b>7100</b>	0	0	0	1,3	0
<b>7200</b>	0	0	0	0,5	-0,1
<b>7300</b>	0	0	0	<b>2,5</b>	-0,1
<b>7400</b>	0,6	0	0,5	1,9	0,7
<b>8000</b>	0	0	0	0,4	-0,4
<b>8500</b>	0	0	0	0,5	1,1
<b>9000</b>	0	0	0	<b>2,5</b>	<b>2,4</b>
<b>9100</b>	0	0	0	-0,6	-1,3
<b>9200</b>	0	0	0	<b>3,8</b>	<b>2,3</b>

Tabell A5.2. Fortsättning...

NGGr2	Visstids- anställda män	Visstids- anställda kvinnor	Totalt antal anställda	Frånvarande p.g.a. sjukdom män	Frånvarande p.g.a. sjukdom kvinnor
<b>Totalt</b>	-0,3	-0,3	0,3	0,5	0
<b>0100</b>	0,2	0,5	0,6	0	0
<b>1000</b>	0	0	0,2	0,6	0
<b>1500</b>	-0,3	0	0,1	-0,9	-1,1
<b>1800</b>	0,2	0,3	-0	-0,7	0
<b>2000</b>	0,8	0,6	<b>2,6</b>	1	0
<b>2100</b>	-0,4	-0,1	0	-1,4	0
<b>2200</b>	-1,6	-0,7	0,2	0,8	0,2
<b>2300</b>	-0,2	-0,1	0,1	1,9	0,4
<b>2500</b>	-0,9	-0,5	-0,4	-1,1	-0,9
<b>2600</b>	-0,8	0	-0,1	0	0
<b>2700</b>	0	0	0	0,2	0
<b>2800</b>	1,3	0	0,5	0,9	0
<b>2900</b>	-0,9	-0,3	0,3	1,3	-0,1
<b>31XX</b>	-0,5	-1,6	0,1	1,5	0
<b>3200</b>	0	0	0,3	0,5	-0,4
<b>3300</b>	0,3	0	0,5	0,6	0,8
<b>3400</b>	-0,6	-1,4	-0,6	-1	-1,7
<b>3500</b>	0	0	0,2	0	0
<b>3600</b>	-1	0	0,5	0,7	0
<b>4000</b>	0,2	0,2	0,2	0,6	-0,6
<b>4510</b>	-0,1	-0,6	0,9	0,8	-1
<b>4530</b>	-1	0	-0,1	-0,3	-2
<b>5000</b>	-1,7	0,6	0,2	0,6	0
<b>5100</b>	-1,6	0,2	0,5	<b>3,4</b>	0,1
<b>5200</b>	0,1	-0,9	0	0,3	-0,5
<b>5210</b>	0	-0,1	0,2	0	0,1
<b>5220</b>	0	0	0,9	0	0
<b>5500</b>	0,5	-0,9	0,3	0	-0,1
<b>6000</b>	-1,9	-0,9	0	0,7	0,5
<b>6100</b>	0,8	0,4	0,6	0,4	0,8
<b>6400</b>	0,2	0	0,6	-0,7	0,9
<b>6500</b>	0,6	0,2	-0,2	0,5	0,4
<b>6600</b>	<b>2,8</b>	1,1	1,4	<b>2,9</b>	<b>2,1</b>
<b>7000</b>	-1,4	0,2	1,6	-0,3	-0,2
<b>7100</b>	0	0	0,5	0	0
<b>7200</b>	-1	0,4	-0,3	0,9	0,2
<b>7300</b>	-0,1	-0,1	0,2	1,5	0,8
<b>7400</b>	1,8	1,8	0,9	1,6	0,4
<b>8000</b>	-1,1	-0,5	-0,1	-0,3	-0,1
<b>8500</b>	-1,8	-1,4	-0	-0,8	0,3
<b>9000</b>	0,4	0,1	1,3	-1	2
<b>9100</b>	<b>-2,5</b>	<b>-2,4</b>	-1,3	-0,2	0,3
<b>9200</b>	<b>-2,3</b>	-2	0,7	-0,2	-0,3

Tabell A5.2. Fortsättning...

NGGr2	Frånvarande	Frånvarande	Frånvarande	Frånvarande	Totalt antal från- varande
	p.g.a. semester män	p.g.a. semester kvinnor	p.g.a. övrig orsak män	p.g.a. övrig orsak kvinnor	
<b>Totalt</b>	0,1	0,1	-0	0	-0,1
<b>0100</b>	0	0	0	0	-0,1
<b>1000</b>	0	0	0	0	0,6
<b>1500</b>	0,1	-0,4	-0,4	-0,2	-0,5
<b>1800</b>	0	0	0	-0,6	-0,3
<b>2000</b>	1,7	0	0,3	0,3	0,7
<b>2100</b>	-0,2	-0,3	-0,8	-0,8	-0,1
<b>2200</b>	0,7	0,7	0,4	0,1	-0,8
<b>2300</b>	0,6	0,3	0,3	0	0,4
<b>2500</b>	-0,8	0	-1,3	-0,8	-0,8
<b>2600</b>	0	0	0	0	-0,1
<b>2700</b>	0	0	0	0	0,2
<b>2800</b>	0	0	0	0	0,5
<b>2900</b>	0	1	-0,9	0,6	0,6
<b>31XX</b>	0,5	-0,9	0,4	-0,7	-0,2
<b>3200</b>	0,3	0	0,1	0	-0,2
<b>3300</b>	0,8	0,5	0,9	0,9	0,9
<b>3400</b>	-1	-1,1	-1,2	-1,1	-0,7
<b>3500</b>	0,5	0	0	0	0
<b>3600</b>	0	0	-0,4	-0,1	0,3
<b>4000</b>	0,4	0	-0,1	0,5	0,1
<b>4510</b>	-0,2	0	0	-0,7	0,2
<b>4530</b>	-0,7	0	-0,4	0	0,1
<b>5000</b>	0	0	0,4	-0,6	-1,2
<b>5100</b>	1,4	0,1	0,9	0,1	0,9
<b>5200</b>	0,3	1,2	0,1	0,2	-0,7
<b>5210</b>	0,7	0	0,8	0,9	0,7
<b>5220</b>	0	0	0	0	0
<b>5500</b>	0	-0,3	0,1	0,2	-0,2
<b>6000</b>	0,7	0,9	0,3	0	0,3
<b>6100</b>	0,2	-0,2	0,9	0,2	1,6
<b>6400</b>	0	0	0,1	0	0,2
<b>6500</b>	0,2	0,3	0	0,2	0,7
<b>6600</b>	1,1	<b>2</b>	<b>2,2</b>	1,2	<b>2</b>
<b>7000</b>	-0,3	-0,2	-0,6	-0,4	-0,4
<b>7100</b>	0	0	1,2	0	0,4
<b>7200</b>	0,1	0,2	-0,3	0,2	0,2
<b>7300</b>	-0,4	0	0	-0,5	0,1
<b>7400</b>	-0,8	-0,3	0,6	-0,5	<b>-2,1</b>
<b>8000</b>	0	-0,1	-1,1	-0,2	0,3
<b>8500</b>	-0,4	-0,8	-0,4	0,1	0,1
<b>9000</b>	-0,4	0,9	-0,5	0,5	0,2
<b>9100</b>	0,2	0,2	0,2	0,2	0,2
<b>9200</b>	-0,1	0	0,1	-0,1	0,1



Tabell A5.2. Fortsättning...

<b>NGGr2</b>	<b>Nya tillsvidareanställda män</b>	<b>Nya tillsvidareanställda kvinnor</b>	<b>Nya visstidsanställda män</b>	<b>Nya visstidsanställda kvinnor</b>	<b>Totalt antal nyanställda</b>
<b>Totalt</b>	-0,1	-0,3	-0,1	0	0
<b>0100</b>	0	0	0	0	-0,3
<b>1000</b>	0	0	0	0	0
<b>1500</b>	0	0	-0,6	-0,3	-0,5
<b>1800</b>	0	0	0	0	0
<b>2000</b>	0,6	0	0,9	0	0,3
<b>2100</b>	0	0	-0,1	0,3	0
<b>2200</b>	-0,8	0	0	0,7	0
<b>2300</b>	0	0	0,2	0,3	0,3
<b>2500</b>	0	0	-1,1	-0,9	-0,8
<b>2600</b>	0	0	0	0	0
<b>2700</b>	1,2	0	0,4	0	0,2
<b>2800</b>	0	0	0	0	0
<b>2900</b>	0	0	-0,6	-1,3	-0,5
<b>31XX</b>	0	0	-0,8	-1	-0,7
<b>3200</b>	0	0	0	0	0
<b>3300</b>	0	0	0	0	0
<b>3400</b>	-0,8	0	-0,1	-1,3	-0,7
<b>3500</b>	0	0	0	0	-0,4
<b>3600</b>	0	0	-0,3	0	-0,1
<b>4000</b>	1,2	0	0,7	0,7	0,7
<b>4510</b>	-0,4	0	-0,5	-1,7	-0,5
<b>4530</b>	0	0	-0,4	0	-0,2
<b>5000</b>	0	0	0	0	0
<b>5100</b>	0	0	-0	-0,1	0
<b>5200</b>	0	0	0,3	0,3	0,2
<b>5210</b>	0	0	0	0	0,9
<b>5220</b>	0	0	0	0	0
<b>5500</b>	0	0,3	-0,3	-0	-0
<b>6000</b>	-0,4	-1	0,4	0,4	0,8
<b>6100</b>	0	0	-0,1	-0,1	-0,2
<b>6400</b>	0	0	0	0	0,1
<b>6500</b>	0	0	0	-0,2	0,3
<b>6600</b>	2	0	1,8	-0,7	0,5
<b>7000</b>	-1	0	-0,1	-0,3	-0,2
<b>7100</b>	0	0	0	0	0
<b>7200</b>	-0,4	0	0	0	-0,5
<b>7300</b>	0	0	0	0	0
<b>7400</b>	-0,2	-0,6	0,2	0,1	0,1
<b>8000</b>	0	0	0	0	0,3
<b>8500</b>	-1,1	-1,5	0,1	0,9	0,4
<b>9000</b>	0	0	0	0,4	0,1
<b>9100</b>	0	0	-0,2	-0,3	-0,2
<b>9200</b>	0	0	0	0	0

Tabell A5.2. Fortsättning...

NGGr2	Avgångna tillsvidareanställda män	Avgångna tillsvidareanst. kvinnor	Avgångna visstidsanst. män	Avgångna visstidsanst. kvinnor	Totalt antal avgångna
<b>Totalt</b>	-0,2	-0,1	-0	0,3	0,2
<b>0100</b>	0	0	0	0	0,3
<b>1000</b>	0	0	0	0	0
<b>1500</b>	-0,3	-1	-0,4	0	0,1
<b>1800</b>	0	0	0	0	0
<b>2000</b>	0	0	0,5	0	0,4
<b>2100</b>	-0,6	0	1,7	0	-0,3
<b>2200</b>	0	-0,8	0	0	-0,1
<b>2300</b>	0	0	0	0	0,2
<b>2500</b>	<b>-2,2</b>	0	0	<b>-2,4</b>	-1,5
<b>2600</b>	0	0	0	0	0
<b>2700</b>	0	0	0	0	0
<b>2800</b>	0	0	0	0	0
<b>2900</b>	0,6	0	-0,3	0	0,4
<b>31XX</b>	0	0	-1,5	0	0
<b>3200</b>	0	0	0	0	0,3
<b>3300</b>	0	0	0	0	0
<b>3400</b>	-0,5	-0,9	0	0	0,5
<b>3500</b>	0	0	0	0	0,6
<b>3600</b>	0,6	0	0	0	0,2
<b>4000</b>	0	0	1,3	0	0,5
<b>4510</b>	-0,5	0	-1,2	0	-0,8
<b>4530</b>	-0,8	0	-0,9	0	0,3
<b>5000</b>	0	0	0	0	-0,1
<b>5100</b>	-0,3	0	0	0	-0,1
<b>5200</b>	0	0,2	0	0,3	-0,6
<b>5210</b>	0	0	0	0	0
<b>5220</b>	0	0	0	0	0
<b>5500</b>	0	0,2	-0,1	0	0,8
<b>6000</b>	0,5	1,3	0	0	0,3
<b>6100</b>	0	-0,6	0,3	0	-0,1
<b>6400</b>	0	0	0	0	0,1
<b>6500</b>	0	1,5	1,1	0	0,4
<b>6600</b>	1,7	0	0	0	0,6
<b>7000</b>	0	0	0,5	0	-0,2
<b>7100</b>	0	0	0	0	0
<b>7200</b>	0,3	0,6	0	0	-0,5
<b>7300</b>	0	0	0	0	0
<b>7400</b>	-0,6	-0,4	0,1	0,5	-0
<b>8000</b>	0	0	0	0	0,1
<b>8500</b>	-0,5	-0,6	0	1	1,7
<b>9000</b>	0	0	0	0	0,2
<b>9100</b>	-1,2	0	0	0	-0,1
<b>9200</b>	0	0	0	0	-0

De markerade uppgifterna i tabell A5.2 visar de branschgrupper i kombination med respektive variabel där pseudo biasen överstiger  $\pm 2$  procent.

## Appendix 6 – Ytterligare utvärderingsresultat av KV

Diagram A6.1-A6.3 visar storleken på pseudo biasen i procent på Y-axeln för de olika gränserna på X-axeln. Diagrammen är konstruerade genom tillämpning av den metod som presenteras i kapitel 3. I diagrammen används inte den absoluta pseudo biasen, detta för att visa att ändringar av inkomna värden utförs i båda riktningarna.

Diagram A6.1. Pseudo bias på totalnivå för variabeln lediga jobb bemannade av avgående medarbetare.



Diagram A6.2. Pseudo bias på totalnivå för variabeln lediga jobb bemannade av bem.företag/konsulter.

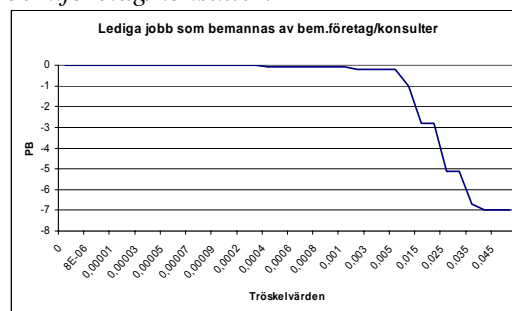


Diagram A6.3. Pseudo bias på totalnivå för variabeln lediga jobb bemannade av vikarier/tillfälligt anställda.



Tabell A6.1 visar punktskattningen för varje variabel per branschgrupp i KV-undersökningen.

Tabell A6.1. Punktskattning per branschgrupp för varje variabel i KV-undersökningen.

NGrp	Tot. antal lediga jobb	Led. jobb bem. av avg. medarb.	Led. jobb bem. av vikarie/t.anst.	Led. jobb bem. av bem.ftg/kons.	Obem. lediga jobb	Kommande obem. lediga jobb
Alla	39 049	5 953	3 747	1 909	13 623	13 860
1	352	23	59	0	68	203
2	7 815	1 009	912	896	2 546	2 496
3	3 569	268	272	268	1 891	871
4	5 824	1 158	712	125	1 812	2 017
5	3 060	416	326	40	1 213	1 062
6	11 077	1 124	431	380	4 857	4 285
7	3 082	1 153	712	52	519	646
8	1 822	195	72	0	252	1 303
9	2 447	608	250	147	465	977

## Appendix 7 – Ytterligare resultat av det genomförda testet

Diagram A7.1-A7.18 visar storleken på pseudo biasen i procent på Y-axeln för de olika gränserna på X-axeln. Här har poängfunktionen som beskrivs i kapitel 3 använts. Det förväntade värdet som använts i funktionen är respektive stratummåttvärde.

Diagram A7.1. Pseudo bias på totalnivå för variabeln visstidsanställda män.

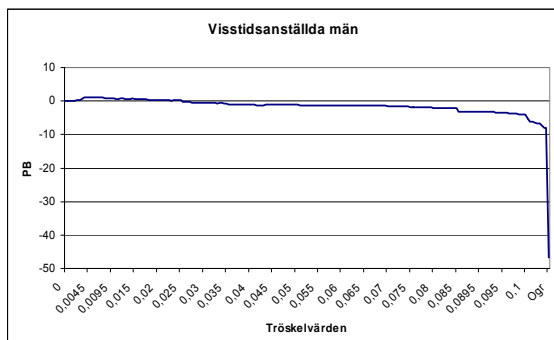


Diagram A7.2. Pseudo bias på totalnivå för variabeln verksamma företagare kvinnor.

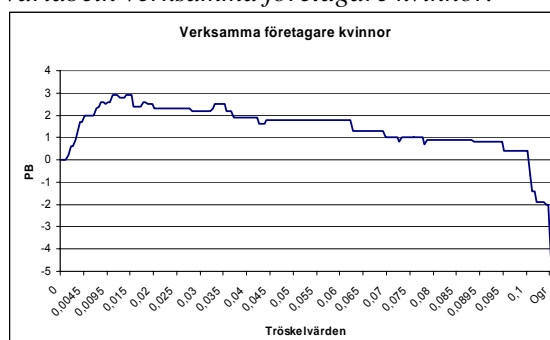


Diagram A7.3. Pseudo bias på totalnivå för variabeln visstidsanställda kvinnor.

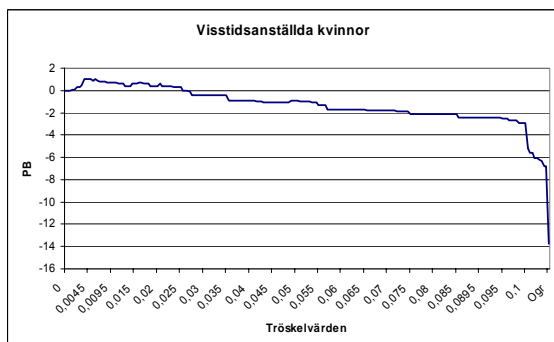


Diagram A7.4. Pseudo bias på totalnivå för variabeln frånvarande sjuka män.

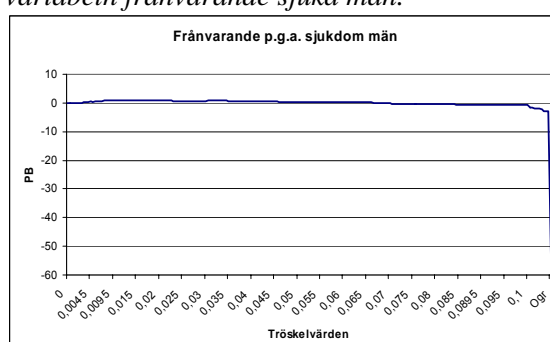


Diagram A7.5. Pseudo bias på totalnivå för variabeln verksamma företagare män.

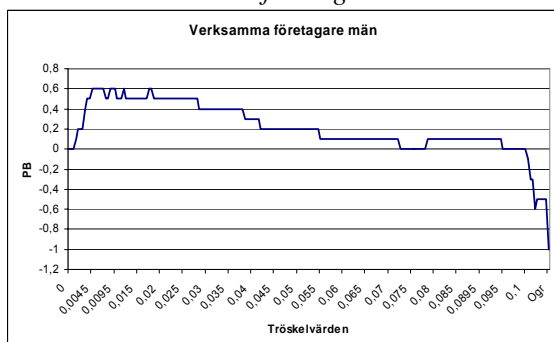


Diagram A7.6. Pseudo bias på totalnivå för variabeln frånvarande sjuka kvinnor.

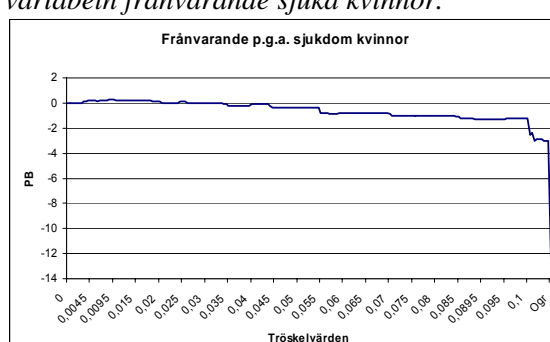


Diagram A7.7. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester män.

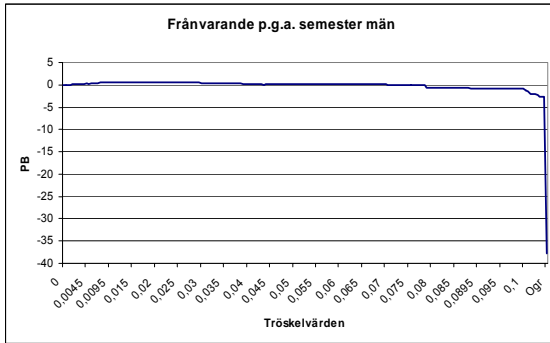


Diagram A7.8. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda män.

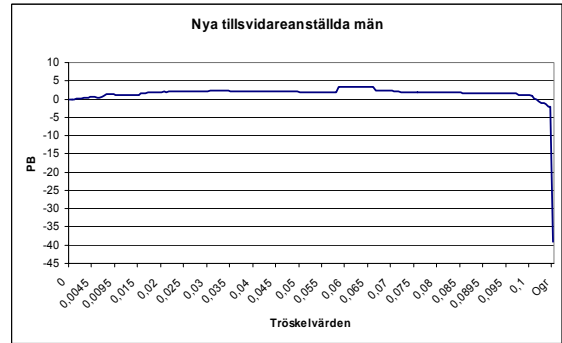


Diagram A7.9. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester kvinnor.

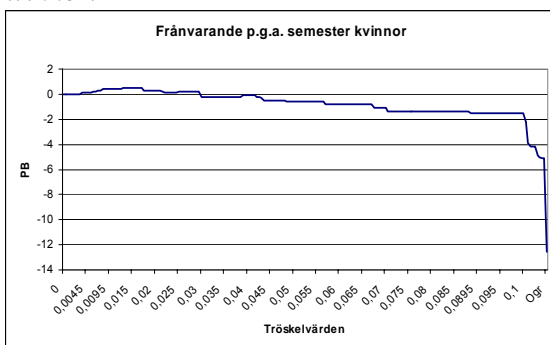


Diagram A7.10. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda kvinnor.

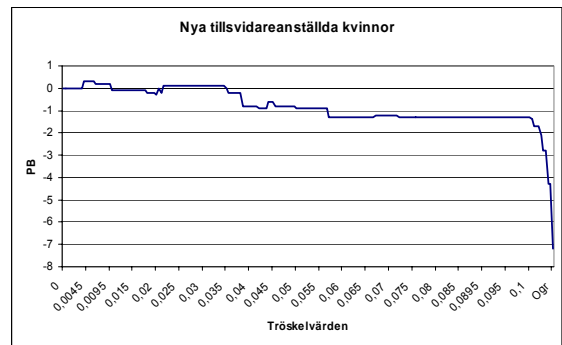


Diagram A7.11. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak män.

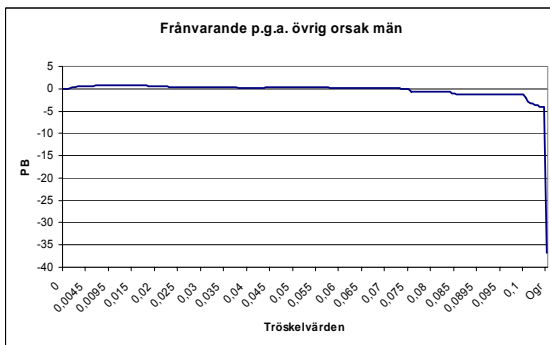


Diagram A7.12. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak kvinnor.

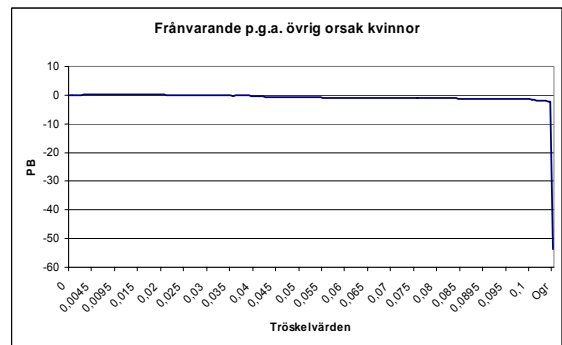


Diagram A7.13. Pseudo bias på totalnivå för variabeln nya visstidsanställda män.

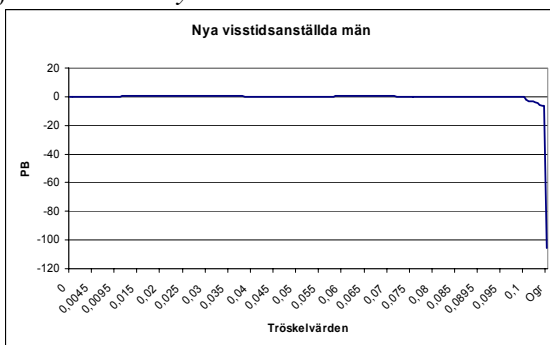


Diagram A7.14. Pseudo bias på totalnivå för variabeln nya visstidsanställda kvinnor.

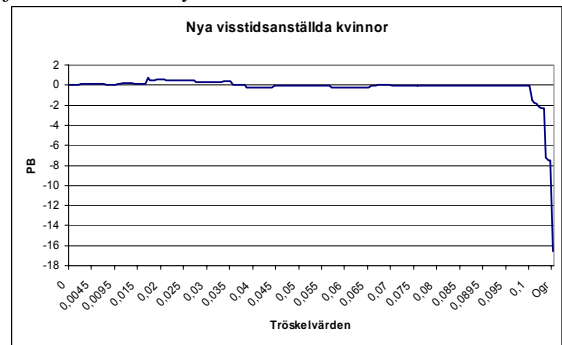


Diagram A7.15. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda män.

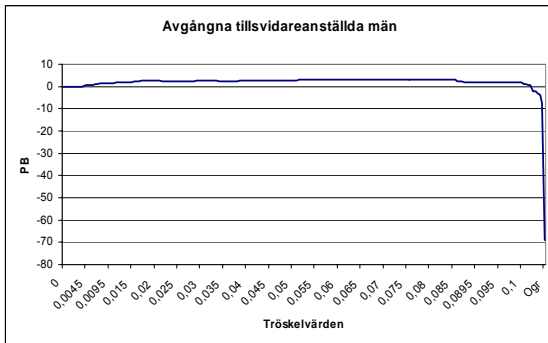


Diagram A7.16. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda män.

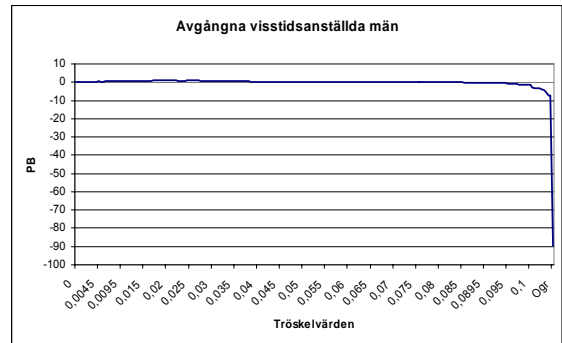


Diagram A7.17. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda kvinnor.

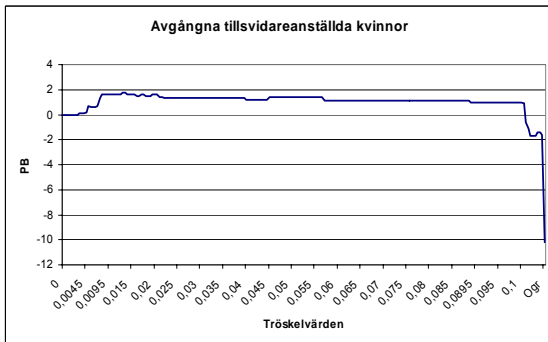


Diagram A7.18. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda kvinnor.

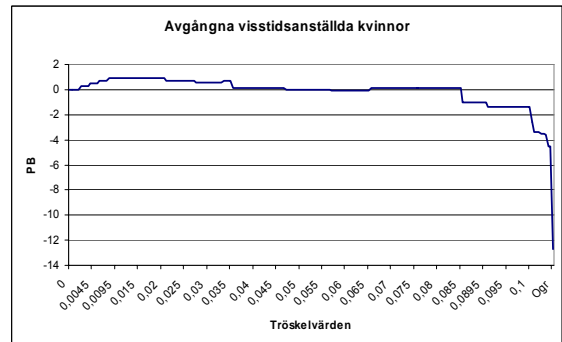


Diagram A7.19-A7.36 visar storleken på pseudo biasen i procent på Y-axeln för de olika gränserna på X-axeln, men endast för intervallet av gränser mellan 0,002-0,02. Även här har respektive stratummedelvärde använts som förväntat värde i poängfunktionen.

Diagram A7.19. Pseudo bias på totalnivå för variabeln visstidsanställda män.

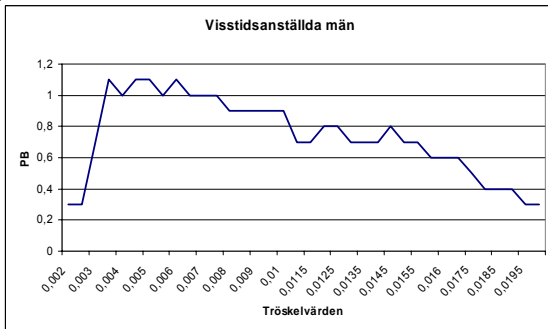


Diagram A7.20. Pseudo bias på totalnivå för variabeln visstidsanställda kvinnor.

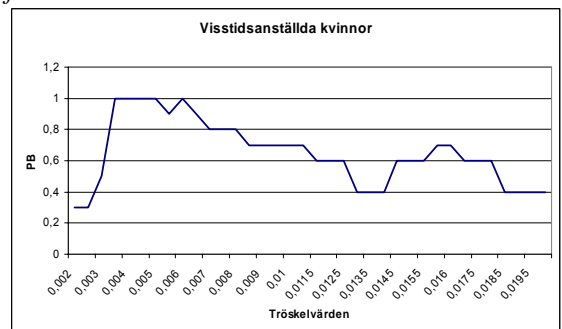


Diagram A7.21. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester män.

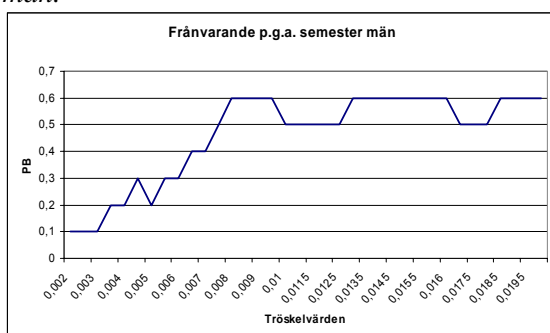


Diagram A7.22. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. semester kvinnor.

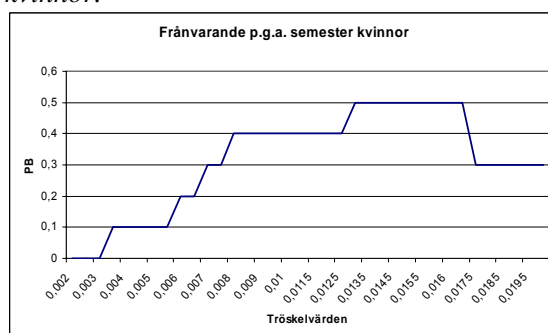


Diagram A7.23. Pseudo bias på totalnivå för variabeln verksamma företagare kvinnor.

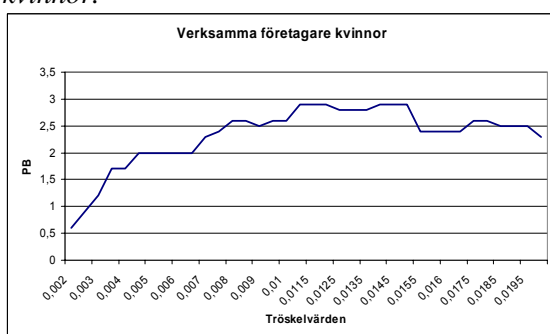


Diagram A7.24. Pseudo bias på totalnivå för variabeln verksamma företagare män.

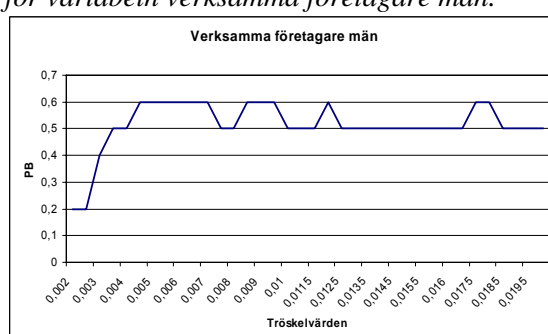


Diagram A7.25. Pseudo bias på totalnivå för variabeln frånvarande sjuka män.

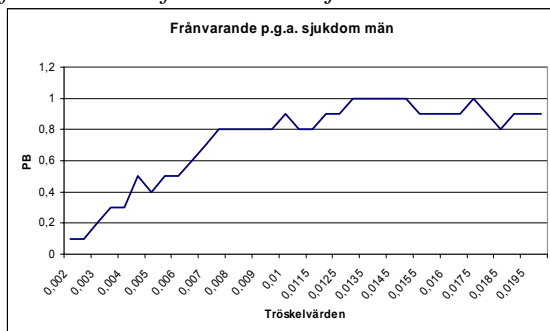


Diagram A7.26. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak män.

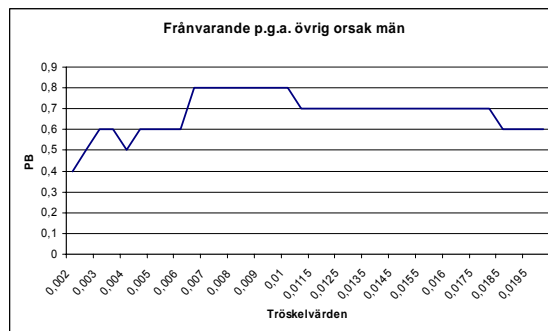


Diagram A7.27. Pseudo bias på totalnivå för variabeln frånvarande sjuka kvinnor.



Diagram A7.28. Pseudo bias på totalnivå för variabeln frånvarande p.g.a. övrig orsak kvinnor.

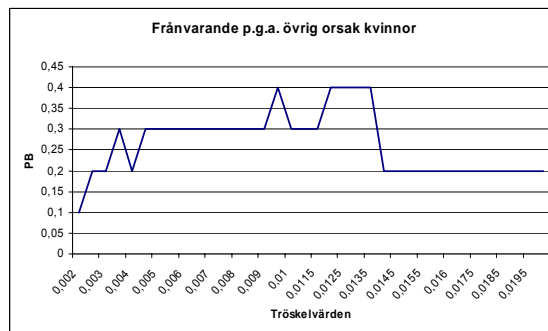


Diagram A7.29. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda män.



Diagram A7.30. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda män.

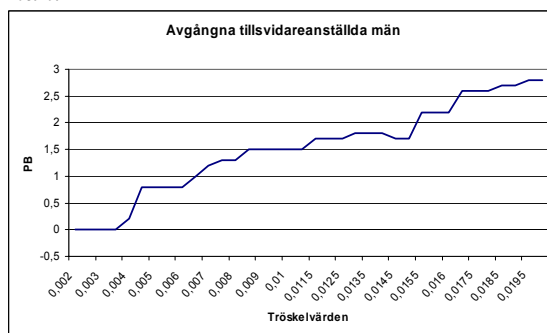


Diagram A7.31. Pseudo bias på totalnivå för variabeln nya tillsvidareanställda kvinnor.



Diagram A7.32. Pseudo bias på totalnivå för variabeln avgångna tillsvidareanställda kvinnor.



Diagram A7.33. Pseudo bias på totalnivå för variabeln nya visstidsanställda män.

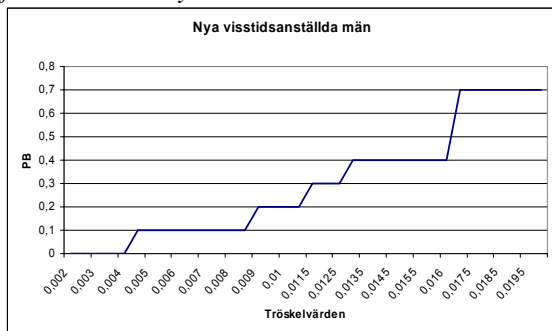


Diagram A7.34. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda män.

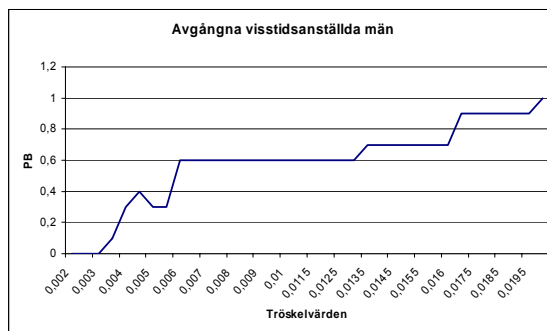


Diagram A7.35. Pseudo bias på totalnivå för variabeln nya visstidsanställda kvinnor.

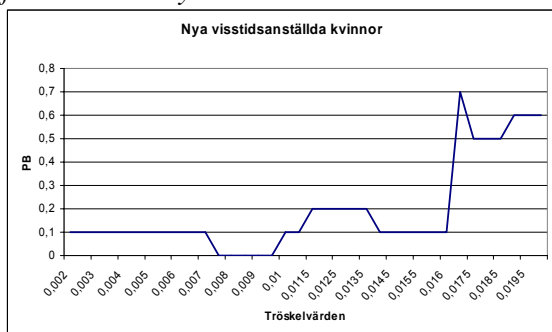


Diagram A7.36. Pseudo bias på totalnivå för variabeln avgångna visstidsanställda kvinnor.





## Appendix 8 – Nycklar till näringsgrensindelning i KS och KV

Tabell A8.1. Näringsgrensindelning i KS-undersökningen (SNI 2002 2006).

<b>NGGr2</b>	<b>Klartext</b>
<b>0100</b>	Jordbruk, jakt och skogsbruk
<b>1000</b>	Gruvor och mineralutvinning
<b>1500</b>	Livsmedel, dryckesvaror och tobaksindustri
<b>1800</b>	Tillverkning av textilier, beklädnad, läder och lädervaror
<b>2000</b>	Trävaruindustri
<b>2100</b>	Massa, pappers och pappersvarutillverkning
<b>2200</b>	Förlag, grafisk och annan reproduktionsindustri
<b>2300</b>	Kemisk och stenkolsindustri, petroleum och kärnbränsle
<b>2500</b>	Gummi och plastvaruindustri
<b>2600</b>	Jord och stenvaruindustri
<b>2700</b>	Stål och metallverk
<b>2800</b>	Metallvaruindustri utom maskiner och apparater
<b>2900</b>	Maskinindustri som ej ingår i annan avd.
<b>31XX</b>	Tillverkning av kontorsmaskiner, datorer och annan elektronik
<b>3200</b>	Teleproduktindustri
<b>3300</b>	Industri för optikprodukter
<b>3400</b>	Industri för motorfordon
<b>3500</b>	Annan transportmedelindustri
<b>3600</b>	Övrig tillverknings- och återvinningsindustri
<b>4000</b>	El, gas, värme- och vattenverk
<b>4510</b>	Mark, grund och byggnadsentreprenörer, uthyrningsfirmor för byggnadsmask.
<b>4530</b>	Byggnadsinstallations och slutbehandlingsfirmor
<b>5000</b>	Handel med och serviceverkstäder för motorfordon
<b>5100</b>	Parti och agenturhandel utom med motorfordon
<b>5200</b>	Detaljhandel ej motorfordon och reparationsverkstäder (exkl. 5210 och 5220)
<b>5210</b>	Detaljhandel med brett sortiment
<b>5220</b>	Specialiserade livsmedelsbutiker samt systembutiker och tobaksaffärer
<b>5500</b>	Hotell och restauranger
<b>6000</b>	Landtransportföretag
<b>6100</b>	Transport, magasinering och kommunikation (exkl. 6000 och 6400)
<b>6400</b>	Post och telekommunikationsföretag
<b>6500</b>	Banker och andra kreditinstitut
<b>6600</b>	Försäkringsbolag
<b>7000</b>	Fastighetsbolag och fastighetsförvaltare
<b>7100</b>	Uthyrningsfirmor
<b>7200</b>	Datakonsulter och dataservicebyråer
<b>7300</b>	Forsknings och utvecklingsinstitutioner
<b>7400</b>	Andra företagservicefirmor
<b>8000</b>	Utbildningsväsendet
<b>8500</b>	Hälso och sjukvård, socialtjänst och veterinärkliniker
<b>9000</b>	Andra samhälleliga och personliga tjänster (exkl. 9100 och 9200)
<b>9100</b>	Intresseorganisationer och samfund
<b>9200</b>	Enheter för rekreation och sport

*Tabell A8.2. Näringsgrensindelning i KV-undersökningen (SNI 2002) 2006.*

<b>NGrp</b>		<b>Klartext</b>
<b>1</b>	A+B	Jordbruk, jakt och skogsbruk
<b>2</b>	C+D+E	Industrin
<b>3</b>	F	Byggindustri
<b>4</b>	G	Parti- och detaljhandel, reparationsverkstäder
<b>5</b>	I	Transport, magasinering och kommunikation
<b>6</b>	J+K	Finansiell verksamhet och företagstjänster
<b>7</b>	M+N	Utbildning samt hälso- och sjukvård
<b>8</b>	H	Hotell och restauranger
<b>9</b>	O	Andra samhällliga och personliga tjänster