CHALMERS | GÖTEBORGS UNIVERSITET

Modellering av marknadspåverkan på aktiepriser

Kandidatarbete inom civilingenjörsprogrammen Teknisk fysik och Teknisk matematik vid Chalmers tekniska högskola

Maria Bergqvist
Sandra Jansson
Pia Lidman
Jacob Söderström

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborgs universitet
Göteborg 2016

Modellering av marknadspåverkan på aktiepriser

Kandidatarbete i matematik inom civilingenjörsprogrammet Teknisk fysik vid Chalmers tekniska högskola
Jacob Söderström

Kandidatarbete i matematik inom civilingenjörsprogrammet Teknisk matematik vid Chalmers tekniska högskola
Maria Bergqvist Sandra Jansson Pia Lidman

Handledare: Sören Christensen
Examinator: Marina Axelson-Fisk
Maria Roginskaya

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborgs universitet
Göteborg 2016

Marknadspåverkan: en möjlig pengabov

Ett sätt att tjäna snabba pengar kan vara att placera tillgångar i aktier, men att bli rik på aktiehandel är inte helt enkelt. Du ska inte bara välja rätt aktier vid rätt tidpunkt utan om du ska köpa eller sälja en stor mängd aktier finns det risk för att din handel pressar upp respektive ned priset. Det här leder till att din förtjänst inte blir så stor som du hade förväntat. Detta fenomen kallas marknadspåverkan och kan tydliggöras genom att titta på en fruktmarknad.

Tänk dig att du på en fruktmarknad ska köpa 100 äpplen till dina kollegor. På marknaden finns ett antal fruktstånd och när du anländer säljer alla äpplen för 5 kr styck. Första frukthandlaren har bara 35 äpplen, vilka du köper till det priset. Eftersom du inte har äpplen till alla kollegor går du vidare till nästa frukthandlare. Denna har sett att du har köpt alla äpplen från den första handlaren och höjer därför priset till 7 kr per äpple. Du behöver äpplena så du köper de 35 äpplen som den här frukthandlaren har, men för det något högre priset. Det saknas fortfarande 30 äpplen och därför går du vidare till nästa handlare. Även den tredje handlaren har märkt att du köper mycket äpplen så handlaren höjer priset till 10 kr styck och du handlar de resterande äpplena där. Därmed har du köpt 100 äpplen men för ett genomsnittspris på 7,2 kr per äpple. Din handel har därmed påverkat marknaden så att genomsnittspriset du betalat är 2,2 kr mer per äpple än vad du hade förväntat dig. När du går förbi marknaden senare under dagen märker du att priset återgått till 5 kr per äpple och du har därmed betalat 220 kr mer än vad som är marknadsvärdet för äpplena.

Detta kallas som sagt marknadspåverkan och innebär att genom att handla stora mängder kan priset på varan drivas upp eller ned beroende på om du köper eller säljer. Denna effekt gäller inte bara för fruktmarknader utan kan även uppstå vid stora aktieköp och kan då innebära stora oväntade förluster för köparen. Det kan även ske vid stora aktieförsäljningar men då med förluster för säljaren. Ett sätt att minska förlusterna är därför att försöka förutspå vilka aktier som leder till störst marknadspåverkan och därmed anpassa handeln efter det.

Idag finns inget allmänt vedertaget sätt att förutspå marknadspåverkan. Däremot finns det en del tidigare publicerade modeller som används, av vilka en ligger till grund för den här studien. En anledning till att det inte finns någon vedertagen modell kan vara att om en modell skulle fungera väl skulle den troligtvis inte publiceras. Om den publiceras skulle alla handla på detta sätt och modellen skulle inte längre fungera lika väl. Tänk bara tillbaka på fruktmarknaden. Där hade det inneburit att om du hade tänkt att äpplen kommer ha en stor marknadspåverkan medan päron har en liten väljer du att köpa päron istället för äpplen. Skulle alla andra också tänka på samma sätt kommer efterfrågan på päron istället att öka och efterfrågan på äpplen minska. Det leder till att marknadspåverkan blir stor för päron men liten för äpplen och förutsägelsen är inte längre aktuell.

För att kunna skapa ett sätt att förutspå marknadspåverkan behöver vi se bakåt i tiden och försöka avgöra om det finns något historiskt mönster i denna effekt. Eftersom det inte går att både handla och låta bli att handla en aktie samtidigt leder det till att marknadspåverkan inte kan mätas exakt. Därmed finns inte data att tillgå som beskriver värdet på marknadspåverkan och att uppskatta marknadspåverkan så nära det verkliga värdet som möjligt blir därför en utmaning.

I studien har fyra olika sätt att uppskatta marknadspåverkan undersökts. Dessa användes sedan i en tidigare framtagen modell, för att se hur väl modellen fungerar. De olika metoderna kommer även de att åskådliggöras med hjälp av fruktmarknaden.

Det första sättet jämför priset när du bestämmer dig att handla äpplen med det genomsnittliga priset du köper för. Antag att du på morgonen går förbi marknaden och ser att äpplena kostar 4 kr. Marknaden har inte öppnat så du bestämmer dig för köpa äpplena på lunchen. När du kommer tillbaka då visar det sig att det var gårdagens pris du sett och köpen går till som i det tidigare exemplet. Med denna metod blir marknadspåverkan då 3,2 kr per äpple eftersom när du bestämde dig för att handla kostade äpplena 4 kr styck och genomsnittspriset för köpen var 7,2 kr per äpple. Här har vi då antagit att marknadspriset på äpplen inte förändras under dagen utan prisökningarna är till följd av att handlarna ser en möjlighet att tjäna extra. Det andra sättet jämför priset du köpte det första äpplet för med det genomsnittliga priset. Alltså 5 kr mot 7,2 kr, vilket ger en marknadspåverkan på 2,2 kr per äpple.

I slutet av varje dag sker en auktion på fruktmarknaden där de säljare som har kvar frukt från dagen och de köpare som är villiga att köpa frukt deltar. Vid denna auktion sätts ett fast pris för varje frukt, vilket alla köper och säljer för. Den tredje metoden för uppskattning av marknadspåverkan jämför priset på äpplen vid auktionen med vad äpplena kostade på gårdagens auktion. Även det fjärde sättet att uppskatta marknadspåverkan använder sig av denna auktion men där jämförs istället auktionens pris med det genomsnittliga priset för marknaden när den öppnar dagen efter.

Inget av de fyra sätten att uppskatta marknadspåverkan gav några användbara resultat. Utifrån hur marknadspåverkan fungerar är de framtagna resultaten helt enkelt orimliga. På fruktmarknaden skulle resultaten motsvara att frukthandlarna sänker sina priser när de märker att någon är intresserad av att handla stora mängder. Detta innebär antingen att den ovan nämnda modellen eller att sätten att uppskatta marknadspåverkan inte fungerar. Ytterligare en anledning kan vara att den prisinformation som används inte är helt korrekt och att uppskattningarna inte tar hänsyn till andra händelser på marknaden. Till exempel kan ett ökat äppelpris bero på att en kändis har berättat om en ny diet som innebär att man bara ska äta äpplen. Därmed ökar efterfrågan på äpplen och till följd av det höjs även priset. Denna påverkan är inte en följd av din handel men du påverkas ändå av den.

Frågan kring hur detta fenomen förutspås är fortfarande öppen. Kanske går det inte ens att förutspå denna förlust som kallas marknadspåverkan. Men en sak är säker, den finns och det är många som är intresserade av att lösa den här gåtan. Skulle du göra det kan storkovan bli din, men endast så länge du inte berättar om din placeringstaktik för någon annan.

Sammandrag

En oväntad förlust vid aktiehandel med stora ordrar kan ske till följd av en effekt kallad marknadspåverkan. Marknadspåverkan definieras som den procentuella skillnaden mellan transaktionspriset och marknadspriset om ordern inte nått marknaden. Denna effekt blir tydligare vid handel med stora ordrar eftersom en order då kan behöva genomföras i omgångar. Därmed märks en förändring i efterfrågan av aktien och således ändras priset. Denna prisförändring hade inte ansetts vara en förlust om det inte empiriskt påvisats att den delvis är av temporär karaktär.

I dagsläget finns ingen allmänt vedertagen modell för marknadspåverkan. Syftet med projektet är därför att skapa en enkel modell för denna effekt. Modellen togs fram i samarbete med myndigheten Andra AP-fonden då de är intresserade av att använda en sådan modell vid optimering av aktieportföljer. Detta gjordes genom att utgå från en befintlig modell där marknadspåverkan, MI, modelleras enligt

$$\mathrm{MI} = c_1 \frac{\sigma}{V} \left(\frac{\phi}{V} \right)^{c_2} q.$$

I ekvationen är σ aktiens volatilitet, V är den genomsnittliga volymen som handlas per dag, ϕ är antal tillgängliga aktier på marknaden och q är antal aktier som ordern består av. För att skapa modellen behövdes konstanterna c_1 och c_2 bestämmas med icke-linjär regression utifrån given data för de nämnda parametrarna. För att göra det krävdes en skattning av marknadspåverkan eftersom denna parameter inte är direkt observerbar.

Tre uppsättningar data innehållande information om transaktioner, som är genomförda av Andra AP-fonden, analyserades för att kunna skatta marknadspåverkan på fyra olika sätt. Den första metoden jämförde priset när ordern träffade marknaden med det genomsnittliga transaktionspriset. Den andra metoden jämförde priset för första transaktionen med det genomsnittliga transaktionspriset. Den tredje metoden för skattning jämförde aktiens avkastning från stängningspris mellan två dagar med avkastningen för aktiens referensgrupp för samma tidsperiod. Den fjärde metoden skattade marknadspåverkan genom att jämföra aktiens avkastning med avkastningen för aktiens referensgrupp. Skillnaden var att i det senare fallet var det inte den dagliga avkastningen mellan två stängningspris utan avkastningen från stängningspris till öppningspris nästkommande dag som användes.

För varje skattning av marknadspåverkan delades ordrarna upp i fem geografiska regioner och c_1 samt c_2 beräknades. Inte för någon kombination av skattningsmetod och region lyckades dock ett rimligt resultat erhållas. Detta beror troligtvis på brister i metoderna för hur marknadspåverkan skattas samt brister i modellen. Resultatet kan även vara till följd av stora mängder brus i den erhållna datan.

Abstract

When trading with large stock orders an unexpected loss may be caused by an effect called market impact. Market impact is defined as the difference between the price of the transaction and the market price in absence of the transaction. This effect is more apparent when trading with large orders as the order then may be executed through several transactions. A partially executed order will be noticed by other traders who will adjust their price to the increased demand. These price changes would not have been considered a loss if it had not been empirically proven that the price changes were partially temporary.

Today there is no generally accepted model for estimating market impact. Therefore, the aim of this project is to develop a simple model for estimating market impact. The project was done in cooperation with the Second Swedish National Pension Fund, because of their interest in applying the model in their portfolio optimization. The model was developed from a preexisting model where market impact, MI, is modelled as

$$\mathrm{MI} = c_1 \frac{\sigma}{V} \left(\frac{\phi}{V} \right)^{c_2} q.$$

In the model, σ is the volatility of the stock, V is the average daily volume, ϕ is the amount of shares outstanding and q is the amount of stocks in the order. In order to develop the model the constants c_1 and c_2 had to be determined by nonlinear regression using existing data for the four mentioned parameters. To determine the constants, a good estimation of market impact was required since market impact is not directly observable.

In this project, market impact has been estimated using four different methods with data given by the Second Swedish National Pension Fund. This data contains information about several orders performed by the Fund. The first method was to compare the price when the order hit the market with the average price for the order. The second method was to compare the price of the first transaction with the average price for the order. The third method was to compare the daily return for the stock with the daily return for a peer group. The fourth method did also compare the return of the stock with the return of a peer group. The difference between the third and fourth method was that the fourth method used the overnight return, i.e. how much the price had changed from the market closing to the market opening the following day.

For all four methods of estimating market impact the orders were divided into five geographical regions. For each combination of method and region estimations of c_1 and c_2 were calculated. Although none of the calculations resulted in any reasonable values. This is most likely due to flaws in the methods of estimating market impact and flaws in the model, as well as noise in the data.

Förord

Följande kandidatarbete har utförts i samarbete med Andra AP-fonden. Det har varit mycket givande att arbeta tillsammans med en extern part och se vad våra utbildningar kan leda till efter examen. Under arbetets gång har vi fört loggbok över tiden som lagts ner och vad som har gjorts. Alla delar i rapporten har skrivits tillsammans av samtliga gruppmedlemmar.

Vi skulle vilja ge ett hjärtligt tack till Jan Lennartsson, Carl Lindberg och övriga medarbetare på avdelningen för kvantitativ förvaltning på Andra AP-fonden för ett gott och givande samarbete. Slutligen skulle vi även vilja ge ett tack till vår handledare Sören Christensen för den tid och engagemang han har lagt ned i vårt projekt och ett litet tack till Claes Ohlsson för korrekturläsning och fackspråklig vägledning.

Innehåll

1 Inledning **1**
- 1.1 Bakgrund . . . 1
- 1.2 Syfte och uppgift . . . 1
- 1.3 Generella avgränsningar . . . 2

2 Teori **3**
- 2.1 Ekonometri . . . 3
- 2.2 Regressionsanalys . . . 3
- 2.3 Marknadspåverkan . . . 4

3 Metod **6**
- 3.1 Litteraturstudiens utförande . . . 6
- 3.2 Skattning av marknadspåverkan . . . 6
  - 3.2.1 Skattning av marknadspåverkan med priset när ordern träffar marknaden och det genomsnittliga priset . . . 8
  - 3.2.2 Skattning av marknadspåverkan med första transaktionspriset och det genomsnittliga priset . . . 9
  - 3.2.3 Skattning av marknadspåverkan med referensgrupper och dagsavkastning 9
  - 3.2.4 Skattning av marknadspåverkan med referensgrupper och avkastning mellan stängnings- och öppningspris . . . 9
  - 3.2.5 Operationella avgränsningar . . . 10

4 Resultat **11**

5 Diskussion **15**
- 5.1 Analys av beräknade värden för konstanterna c_1 och c_2 . . . 15
- 5.2 Analys av konfidensintervall för c_1 och c_2 . . . 15
- 5.3 Analys av regionsuppdelning . . . 16
- 5.4 Analys av data . . . 16
- 5.5 Brister i modellen . . . 16
- 5.6 Diskussion kring uteslutande av extremvärden . . . 17
- 5.7 Felanalys av skattningsmetoderna . . . 17
  - 5.7.1 Skattning av marknadspåverkan när ordern träffar marknaden och genomsnittliga priset . . . 17
  - 5.7.2 Skattning av marknadspåverkan med första transaktionspriset och genomsnittliga priset . . . 18
  - 5.7.3 Skattning av marknadspåverkan med referensgrupp och daglig avkastning 18
  - 5.7.4 Skattning av marknadspåverkan med referensgrupper vid öppningspris 19
- 5.8 Utvecklingspotential . . . 19
  - 5.8.1 Utveckling av data . . . 19
  - 5.8.2 Utveckling av modell . . . 20
  - 5.8.3 Utveckling av skattningsmetoder . . . 20

6 Slutsats **21**

Referenser **22**

Bilaga A Figurer för skattning av marknadspåverkan **23**
- A.1 Figurer för skattning av marknadspåverkan med priset när ordern träffar marknaden och det genomsnittliga priset . . . 23
- A.2 Figurer för skattning av marknadspåverkan med första transaktionspriset och det genomsnittliga priset . . . 26
- A.3 Figurer för skattning av marknadspåverkan med referensgrupper och dagsavkastning . . . 29
- A.4 Figurer för skattning av marknadspåverkan med referensgrupper och avkastning mellan stängnings- och öppningspris . . . 32

1 Inledning

1.1 Bakgrund

På aktiemarknader köps och säljs stora mängder värdepapper. Ett värdepappers pris sätts till det värde som både en köpare och en säljare är villiga att handla för. Aktiens börspris är därmed aktiens pris vid den senaste transaktionen, det vill säga senaste priset som en köpare och en säljare kommit överens om. Det innebär att varje lagd order, köp eller sälj, kommer påverka aktiepriset. Prisförändringarna gör att aktiehandel anses vara en möjlighet för investerare att göra en ekonomisk vinst.

För att maximera det framtida värdet på sitt kapital kan en investerare placera kapitalet i en kombination av värdepapper, vilka tillsammans utgör en så kallad aktieportfölj. Genom att placera aktieportföljen enligt ett specifikt index följer portföljens värde marknadens upp- och nedgångar. Ett index är en samling vikter som anger hur stor andel av kapitalet som ska investeras i respektive värdepapper. Vikterna summerar till ett och avspeglar den genomsnittliga förändringen av aktiemarknaden.

Aktiva investerare har dock en portfölj som avviker från index för att förhoppningsvis generera en större ekonomisk vinst. Svårigheten för en sådan investerare är att avgöra hur aktierna ska fördelas i den avvikande portföljen. Portföljen kan optimeras utifrån de olika aktiernas förväntade avkastning och volatilitet, hur mycket en aktie varierar, samt beroendestrukturen mellan avkastningarna [1]. Här definieras avkastningen som den procentuella prisförändringen med avseende på ursprungspriset.

Myndigheten Andra AP-fonden är en investerare som använder optimeringsmetoder för att försöka maximera vinsten av det investerade kapitalet. De lägger stora ordrar vilket påverkar aktiepriserna och därmed även den förväntade avkastningen. Denna effekt kallas marknadspåverkan och definieras som den procentuella skillnaden mellan transaktionspriset och marknadspriset om ordern inte nått marknaden [2]. Effekten blir tydligare vid handel med stora ordrar eftersom ordern då kan behöva genomföras i omgångar. Därmed förändras efterfrågan av aktien i högre grad och således även priset. Detta hade inte ansetts som en förlust om det inte empiriskt påvisats att prispåverkan delvis är av temporär karaktär. Givet att ordern påverkar priset och att marknadspåverkan till viss del är temporär innebär det en direkt värdeminskning av portföljen att genomföra transaktionen.

Värdeminskningarna till följd av marknadspåverkan kan dock inte observeras direkt. Detta eftersom om ordern genomförs så är transaktionspriset känt medan priset om ordern inte nått marknaden är okänt. I motsatt fall, att ordern inte når marknaden, är priset om ordern inte nått marknaden känt men transaktionspriset finns inte då transaktionen inte skett. Båda priser kan således inte vara kända och marknadspåverkan kan som tidigare nämnts inte observeras direkt.

1.2 Syfte och uppgift

I dagsläget finns ingen allmänt vedertagen modell för marknadspåverkan på aktiepriser. Syftet med projektet är att skapa en enkel modell som Andra AP-fonden kan använda vid optimering av aktieportföljer.

Uppgiften inleds genom att med hjälp av tidigare framsteg inom området ta fram en enkel modell för marknadspåverkan. För att fastställa modellen behöver det undersökas vilka parametrar som är del av denna. Parametrarna kräver i sin tur relevant data vilken tillhandahålls av Andra AP-fonden. Modellen kräver även skattning av marknadspåverkan eftersom denna parameter inte är direkt observerbar. Skattningarna bygger antingen på att jämföra det genomsnittliga priset för en order med ett pris innan handeln påbörjas eller på att jämföra aktiens avkastning med den viktade avkastningen från en referensgrupp till aktien. Detta för att finna den mest optimala modell som kan användas vid optimering av aktieportföljer.

1.3 Generella avgränsningar

Andra AP-fonden efterfrågade en enkel modell då komplicerade modeller är svåra att tillämpa rent praktiskt tillsammans med den nuvarande optimeringsmetoden som används av Andra AP-fonden. Modellen skulle bestå av ett antal parametrar som tillsammans bildade en funktion, där en förmodad parameter var antalet aktier i ordern. För att fungera väl med den nuvarande optimeringsmetoden ville Andra AP-fonden att denna parameter skulle vara antingen linjär eller kvadratisk, vilket avgränsade modellens struktur. Däremot fanns det inga avgränsningar för vilka de andra parametrarna var eller funktionens utseende i övrigt.

En ytterligare avgränsning var att den tillhandahållna datan var begränsad till en viss tidsperiod och därmed begränsades även antalet aktieordrar. Dessutom innehöll datan enbart ett begränsat antal kända parametrar och därför kunde skattningen av marknadspåverkan endast bero på dessa.

Tidsbegränsning av projektet gjorde att enbart fyra olika metoder för skattning av marknadspåverkan hann analyseras och appliceras på en tidigare modell. Dessa applicerades endast på en publicerad modell eftersom det ansågs mer tidsödande att börja helt från grunden.

2 Teori

Nedan presenteras tre huvudområden för projektets bakomliggande teori. Dessa områden är ekonometri, regression samt marknadspåverkan. Ekonometri är ett brett område där matematiska modeller appliceras på ekonomisk data, och således faller projektet inom ekonometrins ramar. Ett frekvent använt verktyg inom ekonometri är regression vilket innebär att anpassa en modell efter given data. Ekonomisk data utgör även grunden för aktieportföljer vilka Andra AP-fonden strävar efter att optimera. Optimeringen kan dock bli bristfällande vid handel med stora ordrar eftersom en effekt kallad marknadspåverkan då påverkar aktiepriset.

2.1 Ekonometri

Ekonometri är ett brett område där matematiska och statistiska metoder appliceras på ekonomisk data. Syftet med detta är att bekräfta eller dementera ekonomiska teorier genom att testa ekonomiska modeller och se hur väl dessa stämmer med verkligheten [3, s. 1].

Det finns stora likheter mellan ekonometriska och andra statistiska modeller. En viktig skillnad är dock att inom ekonometri kan endast observerad data användas och inte experimentellt framställd data. Detta eftersom experimentellt framställd data inte kan garanteras generera en exakt bild av marknaden då ekonomisk data är stokastisk. Därmed kan samtliga modellparametrar variera vid samma tillfälle vilket gör att det inte går att utföra kontrollerade experiment för att undersöka en modell [4].

En gren av ekonometri är finansiell ekonometri. Den skiljer sig från andra former av ekonometri då den endast fokuserar på data från finansmarknader. Huvudsakligen syftar finansiell ekonometri till att analysera finansiella tillgångar, exempelvis aktier och obligationer. Detta medför att finansiell ekonometri kan användas inom allt från portföljförvaltning till riskanalys [5].

Inom ekonometri används modeller för att förenkla representationen av en verklig process. Det som är riskfyllt är att dessa modeller kan bli alltför förenklade eller att antagandena blir orimliga. En ekonometrisk modell består av en mängd beteendeekvationer som härrör från ekonomiska teorier. Dessa ekvationer består av observerade variabler samt störningar. Modellen redogör även huruvida det återfinns brus i de observerade variablerna eller inte. Dock är sådana modeller är oftast väldigt begränsade i sin applicerbarhet [3, s. 3-4].

Målet med ekonometri är alltså att formulera en rimlig modell som sedan skattas och testas för att slutligen kunna användas för att skapa prognoser. En av de vanligaste metoderna som används inom ekonometri är regression [3, s. 3-4].

2.2 Regressionsanalys

Regression är ett matematiskt verktyg för att analysera hur en variabel beror på flera prediktorer, oberoende variabler. Målet är att få fram ett samband för att förutspå hur variabeln påverkas av ändringar hos prediktorerna. Detta undersöks genom att ändra en av prediktorerna och hålla de andra fixa. På så sätt fås ett samband mellan den givna prediktorn och variabeln.

Utgångspunkten är n punkter på formen $\{y_i, \mathbf{x}_i\}_{i=1}^n$ där y_i är värdet på variabeln och $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^{\mathrm{T}}$ är värdena på prediktorerna. Därefter eftersträvas att hitta en funktion, $y = f(\mathbf{x}, \boldsymbol{\beta})$, som förutspår värdet på y givet $\mathbf{x} \in \mathbb{R}^p$ genom en skattning av $\boldsymbol{\beta}$ [6].

En form av regression är icke-linjär regression. Där anpassas data till en funktion i vilken parametrarna har ett icke-linjärt beroende. Icke-linjär regression använder vanligtvis iterativa metoder, som kräver stora databeräkningar. En svårighet är att det inte går att använda en bestämd lösningsmetod för samtliga problem eftersom de slutgiltiga funktionerna som uppnås är vitt skilda [7].

För att skatta $\boldsymbol{\beta}$ finns det flera tillvägagångssätt. Ett sätt är använda numeriska algoritmer. Applikationen *cftool* i MATLAB är en sådan plattform. Där utförs regressionen genom Gauss-Newtons iterativa algoritm, som bygger på minimering av den euklidiska normen i kvadrat av skillnaden mellan datan och modellens skattning [8],

$$\min_{\boldsymbol{\beta}} \|\hat{\mathbf{y}} - \mathbf{y}\|^2 \tag{1}$$

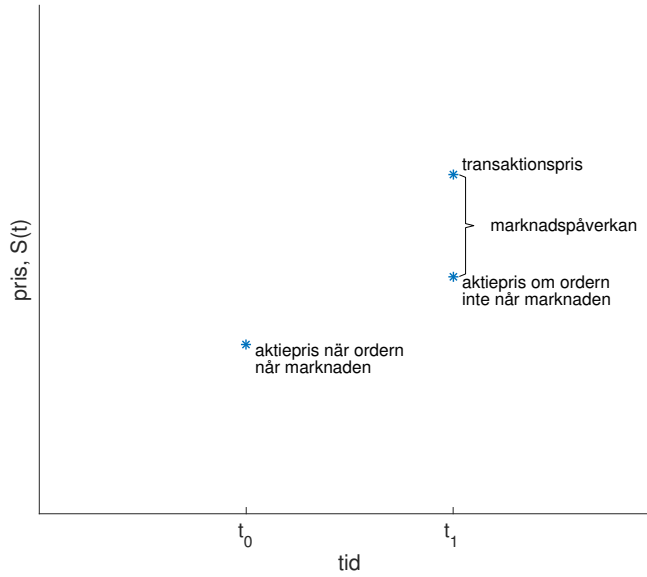
där $\hat{y}_i = f(\mathbf{x}_i, \boldsymbol{\beta})$ är skattningen av y_i.

För varje skattning beräknas även ett konfidensintervall, vilket är ett mått på hur tillförlitliga skattningarna är. Applikationen *cftool* beräknar konfidensintervallet med hjälp av Students t-fördelning[9]. En tillförlitlig skattning bör ha ett litet konfidensintervall som inte innehåller noll. Motiveringen till detta är att då intervallet innehåller noll går det inte att förkasta att skattningen kan vara noll och ett brett intervall implicerar en stor osäkerhet i det skattade värdet.

2.3 Marknadspåverkan

Vid handel på finansiella marknader tillkommer kostnader som är kopplade till transaktionerna, så kallade transaktionskostnader. Kostnaderna kan delas upp i direkta och indirekta kostnader. Direkta kostnader, exempelvis förvaltningsavgifter och courtage, är kända i förväg och kan därför tas i beaktande när det kommer till portföljoptimering. Indirekta kostnader är inte kända i förväg och är således betydligt mer komplicerade att skatta. Vid stora investeringar är den största indirekta kostnaden marknadspåverkan [10].

Marknadspåverkan definieras som den procentuella skillnaden mellan priset på aktien givet att ordern genomförs på marknaden och priset på aktien givet att ordern aldrig når marknaden, se figur 1 för ett tydliggörande exempel. Denna effekt är inte direkt mätbar då transaktionen omöjligen kan vara både genomförd och inte genomförd. Det finns olika sätt att skatta marknadspåverkan, exempelvis att jämföra det genomsnittliga transaktionspriset med marknadspriset precis innan transaktionen sker [2].



Figur 1: Figuren visar aktiepriset som en diskret funktion av tiden. Vid tidpunkt t_0 bestämmer sig en investerare för att köpa aktien. Ett tidssteg senare, t_1, återfinns två punkter i figuren. Den övre punkten representerar det pris som investeraren handlar aktien för. Den nedre punkten visar det pris som aktien hade haft om investeraren inte hade handlat och därmed om ordern inte hade nått marknaden. Prisskillnaden mellan t_0 och den nedre punkten för t_1 är till följd av eventuella marknadsfluktuationer som inte är kopplade till denna investerares handel. Den procentuella skillnaden mellan de två punkterna vid t_1 är det som definieras som marknadspåverkan, nämligen skillnaden mellan transaktionspriset för ordern och priset om ordern inte nått marknaden.

Vid handel med stora ordrar är marknadspåverkan som tydligast eftersom ordern då kan behöva genomföras i flera transaktioner. Detta sker om det inte finns en ensam handlare som kan tillgodose hela ordern och således behöver ordern ske med olika handlare och därmed genom olika transaktioner. När alla aktier inte kan handlas samtidigt förändras efterfrågan av

aktien. En förändrad efterfrågan kan även innebära förändring av priset och därmed handlas inte alla aktier till det ursprungliga marknadspriset.

En konsekvens av detta är att orderns genomsnittliga pris inte överensstämmer med det förväntade priset, det vill säga priset när ordern läggs. Att priserna inte längre stämmer överens innebär ett högre köppris eller ett lägre säljpris för investeraren beroende på om aktier ska köpas eller säljas. Detta kan enligt Jan Lennartsson och Carl Lindberg på Andra AP-fonden anses som en förlust för investeraren eftersom det empiriskt påvisats att marknadpåverkan ofta är delvis temporär. Vilket innebär att priset på aktien efter handeln kommer röra sig tillbaka mot priset innan första transaktionen. För ett köp innebär det här att priset sjunker efter att ordern genomförts, medan det för en försäljning innebär att priset stiger efter genomförandet. Denna prisförändring efter genomförandet av ordern medför därför att portföljens värde minskar om ordern genomförs.

Värdeminskningarna gör att valet av placeringar inte genererar det tänkta värdet. För att kompensera för detta inkluderas en modell för marknadspåverkan vid portföljoptimering. Idag finns ingen vedertagen sådan modell, däremot finns det ett fåtal modeller som försöker skatta marknadspåverkan. En som används frekvent, och används som bakomliggande modell för det här projektet, är en förenklad variant av Almgrens modell. Där modelleras effekten som

$$\mathrm{MI} = c_1 \frac{\sigma}{V} \left(\frac{\phi}{V} \right)^{c_2} q, \tag{2}$$

där c_i, $i = 1, 2$ är konstanter och de andra parametrarna finns beskrivna i tabell 1 [11]. Denna valdes som bakomliggande modell eftersom den används frekvent samt att Andra AP-fondens modell bygger på den.

Tabell 1: Namn och förklaring till de olika parametrarna i den förenklade versionen av Almgrens modell för marknadspåverkan.

Parameter	**Namn**	**Förklaring**
MI	Marknadspåverkan	Skillnaden mellan transaktionspriset och marknadspriset om transaktionen inte skett.
σ	Volatilitet	Beskriver variationen hos aktiepriser.
V	Daglig genomsnittlig volym	Genomsnittligt antal aktier som handlas per dag per specifik aktie.
ϕ	Antal tillgängliga aktier	Antalet aktier i företaget som finns tillgängliga för handel på marknaden.
q	Antal aktier	Antal aktier som önskas köpas/säljas.

3 Metod

Projektet inleddes med en litteraturstudie för att ge grundläggande kunskaper inom ämnet. Därefter behandlades och analyserades tre uppsättningar data för att kunna skatta marknadspåverkan på fyra olika sätt. Skattningarna behövdes sedan för att kunna beräkna de två obekanta konstanterna c_1 och c_2 i ekvation (2). Vilket gjordes med icke-linjär regression genom att minimera residualerna. Efter att konstanterna var funna utvärderades deras rimlighet genom att studera konfidensintervallen för dem.

3.1 Litteraturstudiens utförande

Litteraturstudien utfördes för att få kunskap kring hur marknadspåverkan uppkommer och varför. Därmed erhölls även en uppfattning om strukturen för en potentiell modell av marknadspåverkan. I litteraturen fanns även tidigare publicerade modeller som gav inblick i vilka parametrar som kan anses rimliga vid modellering av marknadspåverkan.

Litteraturstudien gjordes även för att ge grundläggande kunskaper inom de ämnen som skulle komma att beröras och användas i studien. Dessa ämnen var utöver marknadspåverkan: ekonometri och regression vilka kan läsas om i avsnitt 2. Ekonometri var intressant eftersom projektet går under det ämnesområdet. Marknadspåverkan valdes då det var vad som åsyftades att modellera. Regression var av intresse då det är fundamentalt för att kunna bestämma de två obekanta konstanterna i ekvation (2).

Arbetet med litteraturstudier inleddes med att alla i gruppen studerade var sitt teoriavsnitt och sedan diskuterades dessa gemensamt i gruppen. Slutligen skrevs en sammanfattande text för varje ämne för att ge gruppens medlemmar inblick i samtliga ämnen.

3.2 Skattning av marknadspåverkan

Andra AP-fonden tillhandahöll tre olika uppsättningar data. Den första uppsättningen innehöll ordrar lagda mellan 2015-07-28 och 2015-09-15. Den andra innehöll ordrar lagda mellan 2015-07-29 och 2015-09-30. Den tredje innehöll ordrar lagda mellan 2015-07-28 och 2015-11-01.

Den första uppsättningen data innehöll 360 kolumner med olika typer utav information för 6923 ordrar. Först skapades en förståelse för vad informationen i de olika kolumnerna innebar och sedan valdes de kolumner som var av relevans för modelleringen ut. Den andra uppsättningen data bestod av 967 ordrar. Den här datan innehöll relevanta kolumner från den första datan men även, efter önskemål, det första transaktionspriset för en order. Den tredje uppsättningen data innehöll den första uppsättningen data samt ytterligare data med information för ordrar mellan 2015-09-16 till 2015-11-01, totalt erhölls då 9220 ordrar. För att generera data som skulle kunna appliceras på modellen krävdes det att en del av ordrarna togs bort då de av olika anledningar inte var kompatibla med modellen, för specifikationer se avsnitt 3.2.5. Av sekretesskäl redovisas dock inga specifika uppgifter från de tre uppsättningarna data.

Den kvarstående datan delades upp i geografiska regioner, efter önskemål från Andra AP-fonden, baserade på aktiens ursprung. De fem olika regionerna var Europa och Israel (ME), Nordamerika (MN), Nya Zeeland, Singapore, Hongkong och Australien (MP), Japan (MJ) samt tillväxtländer (EW). Uppdelningen gjordes då aktiemarknaderna är likartade inom dessa regioner, enligt Jan Lennartsson och Carl Lindberg på Andra AP-fonden, vilket kan innebära att även marknadspåverkan är likartad där. Därmed skulle modellen bli bättre anpassad efter de specifika marknaderna och således generera ett bättre resultat än en generell modell för hela världen. I modellen är c_1 en skalfaktor och därför borde den till stor del återspegla känsligheten för de olika marknaderna.

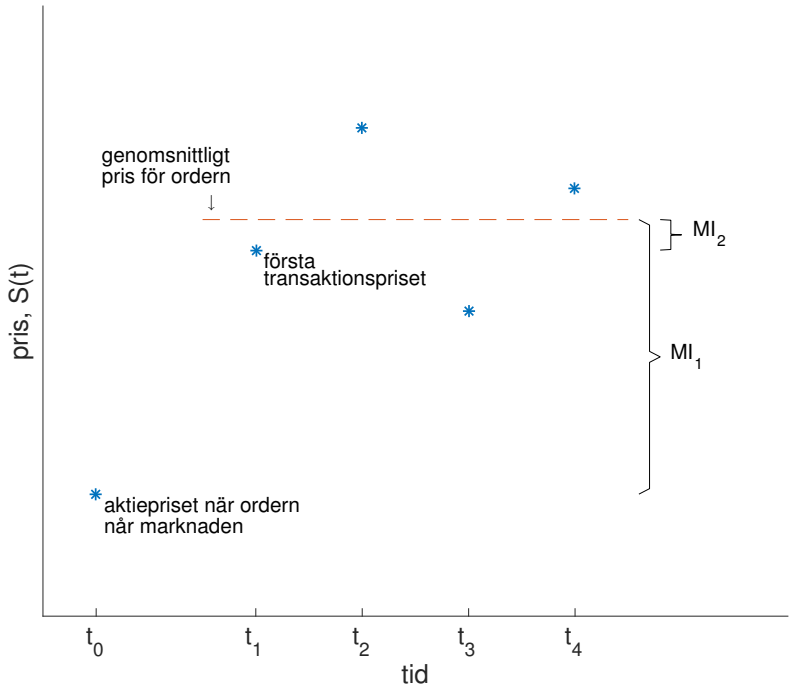
För att skapa en fungerade modell krävdes en skattning av marknadspåverkan för respektive order. Detta för att kunna uppskatta parametrarna c_1 och c_2 i ekvation (2) genom icke-linjär regression. Marknadspåverkan kan, som tidigare nämnts, inte mätas exakt då det skulle kräva att en order både genomfördes och inte genomfördes samtidigt. För att ändå erhålla ett värde på marknadspåverkan analyserades fyra olika metoder för skattning.

Den första metoden för skattning av marknadspåverkan använde den första uppsättningen data och byggde på skillnaden mellan det genomsnittliga transaktionspriset för ordern och aktiepriset då ordern träffade marknaden, se avsnitt 3.2.1. Där det genomsnittliga transaktionspriset beräknas genom

$$\bar{S} = \frac{1}{q} \sum_i q_i S_i$$

där q är antalet aktier i hela ordern, q_i är antalet aktier som handlas vid transaktion i och S_i är priset på aktierna som handlas vid transaktion i.

Den andra metoden för skattning använde den andra uppsättningen data och byggde på skillnaden mellan det genomsnittliga priset för ordern och priset för den första transaktionen i ordern, se avsnitt 3.2.2. Det två metoderna att skatta marknadspåverkan tydliggörs även i figur 2.



Figur 2: Figuren visar aktiepriset som en diskret funktion av tiden för att tydliggöra de två första metoderna för skattning av marknadspåverkan. Vid t_0 återfinns priset för aktien när ordern når marknaden och t_1-t_4 är fyra tidpunkter för de transaktioner som tillsammans utgör hela orderns handel. Dessa fyra punkter ger ett genomsnittligt transaktionspris för ordern vilket representeras med den streckade linjen i figuren. Det första sättet att skatta marknadspåverkan, MI_1, är då den procentuella skillnaden mellan det genomsnittliga priset för transaktionerna och priset när ordern når marknaden. Det andra sättet att skatta marknadspåverkan, MI_2, är den procentuella skillnaden mellan det genomsnittliga priset för transaktionen och det första transaktionspriset, priset vid t_1.

Den tredje metoden använde den tredje uppsättningen data men utgick från skillnaden mellan avkastningen för respektive aktie, med avseende på aktiens stängningspris en dag och aktiens stängningspris nästkommande dag, och den marknadsviktade avkastningen för en referensgrupp till aktien för samma tidsperiod, se avsnitt 3.2.3. Även den fjärde metoden för skattning använde den tredje uppsättningen data och beräknades med hjälp av avkastning för aktiers referensgrupper. Skillnaden var att avkastningen nu beräknades från stängningspris ena dagen till öppningspris nästa dag, se avsnitt 3.2.4

De ordrar som hanterade försäljning av aktier förväntades ge negativa värden på marknadspåverkan för samtliga skattningsmetoder. För att kunna modellera sälj- och köpordrar tillsammans behöver dessa dock ha samma tecken, därför teckenkorrigerades marknadspåverkan för säljordrarna med ett minustecken. Korrigeringen medförde att skattningarna alltid förväntades generera en positiv marknadspåverkan och därmed en förlust för portföljinnehavaren.

Eftersom alla variabler i ekvation (2) utom konstanterna c_1 och c_2 var givna i den tillhandahållna datan ansågs det lämpligt att skriva om modellen enligt följande

$$\begin{aligned} \widehat{\mathrm{MI}} &= c_1 \frac{\sigma}{V} \left(\frac{\phi}{V} \right)^{c_2} q \Leftrightarrow y = c_1 x^{c_2} \\ \text{där } y &= \frac{\widehat{\mathrm{MI}} V}{\sigma q} \text{ och } x = \frac{\phi}{V}, \end{aligned} \tag{3}$$

där $\widehat{\mathrm{MI}}$ är den skattade marknadspåverkan. Från datan skapades variablerna x och y enligt ekvation (3) och dessa användes för att bestämma c_1 och c_2 genom icke-linjär regression. Regressionen utfördes med hjälp av applikationen *cftool*, vilket är ett program i MATLAB som utför kurvanpassning. Applikationen hittar den bäst anpassade kurvan till uppmätta datapunkter genom att använda sig av minsta kvadratmetoden som minimerar residualen, skillnaden mellan det uppmätta värdet och det anpassade funktionsvärdet, i kvadrat.

För att värdena på c_1 och c_2 skulle bli så representativa som möjligt uteslöts extremvärden för de uppmätta datapunkterna (x, y). Detta gjordes genom att i *cftool* begränsa det tillåtna intervallet för x respektive y. Med hjälp av *cftool* beräknades även ett 95 % konfidensintervall för c_1 och c_2. Detta gjordes för att analysera rimligheten för de framtagna värdena på konstanterna. Ett stort konfidensintervall innebär att de uppmätta datapunkterna har stor spridning medan ett litet konfidensintervall antyder att punkterna är väl samlade. Desto mindre konfidensintervall, desto bättre är därför skattningen av c_1 och c_2.

3.2.1 Skattning av marknadspåverkan med priset när ordern träffar marknaden och det genomsnittliga priset

För modellering av marknadspåverkan krävs en skattning av denna parameter eftersom den inte är direkt mätbar. Det första sättet som analyserades var en metod som beräknade skillnaden mellan det genomsnittliga priset, $\bar{S}$, och priset på aktien då ordern nådde marknaden, $S(t_0)$, se figur 2 för exempel. Marknadspåverkan kunde då beräknas enligt

$$\widetilde{\mathrm{MI}} = \frac{\bar{S} - S(t_0)}{S(t_0)}. \tag{4}$$

Skattningen ansågs bristande eftersom priset är en stokastisk process och därför ses plötsliga prisändringar, vilka inte är en följd av marknadspåverkan, som en del av marknadspåverkan. Alltså blir den modellerade marknadspåverkan summan av den verkliga marknadspåverkan och marknadens upp- eller nedgångar.

För att kompensera för detta standardiserades medelvärdet på handelspriserna. Detta innebar att medelvärdet justerades med dagsavkastningen för det land där aktien handlades. Det medförde att marknadspåverkan skattades med

$$\widehat{\mathrm{MI}} = \frac{\bar{S}/(1 + \alpha(d)) - S(t_0)}{S(t_0)} \tag{5}$$

där korrektionstermen $\alpha(d)$ är den dagliga avkastningen på landsnivå. Denna avkastning beräknades genom följande formel

$$\alpha(d_j) = \frac{\sum_{i \in j} w_i R_i}{\sum_{i \in j} w_i}, \text{ där } R_i = \frac{S_i(d) - S_i(d-1)}{S_i(d-1)}. \tag{6}$$

R_i är dagsavkastningen för aktie i från land j, w_i är vikten för ett marknadsvärdesviktat index för respektive aktie i vid öppning på dag d och $S_i(d)$ är stängningspriset för aktie i på dag d.

3.2.2 Skattning av marknadspåverkan med första transaktionspriset och det genomsnittliga priset

Den första metoden ansågs potentiellt innehålla stora mängder brus, vilket innebär stora variationer i data som inte kan förklaras. I ett försök att undvika detta brus byggde den andra metoden på det första transaktionspriset istället för aktiepriset när ordern nådde marknaden. Den använde även den andra uppsättningen data som skiljde sig innehållsmässigt från den tidigare uppsättningen eftersom första transaktionspriset inte fanns i den första uppsättningen data.

Marknadspåverkan kunde då istället skattas enligt

$$\widehat{\mathrm{MI}} = \frac{\bar{S} - S(t_1)}{S(t_1)}. \tag{7}$$

där $\bar{S}$ var det genomsnittliga priset för ordern och $S(t_1)$ var priset för den första transaktionen i ordern, se figur 2 för exempel. Syftet med att använda denna metod var att minimera brus eftersom denna skattning endast bygger på transaktionspriser för ordern och inte marknadspriser. Den här metoden kräver att ordern ska handlas sekventiellt under dagen, det vill säga att ordern genomförs i flera transaktioner, och efter undersökning insågs att större delen av Andra AP-fondens transaktioner sker vid en auktion i stängningset av varje dag. Därför är metoden endast applicerbar på en liten del av Andra AP-fondens handel och därmed undersöktes istället andra möjligheter att skatta marknadspåverkan.

3.2.3 Skattning av marknadspåverkan med referensgrupper och dagsavkastning

Den del av Andra AP-fondens handel som inte sker sekventiellt handlas på en auktion i stängningset av dagen. Stängningsauktionen är till för de köp- och säljordrar som inte genomförts under dagen och där sätts ett stängningspris för varje aktie, vilket resterande handel går igenom för. För att kunna hantera dessa ordrar behövdes en annan skattningsmetod. Den tredje metoden baserades därför på skillnad i den dagliga avkastningen mellan en aktie och en referensgrupp för aktien.

En referensgrupp innehåller olika företag som liknar varandra på ett eller flera sätt. Det kan vara att de tillhör samma sektor, exempelvis transport, bank eller fastigheter, är ungefär lika stora eller att de ligger i samma region. De referensgrupper som används i denna metod tillhör samma region och samma industriklass.

Den tredje metoden använde den tredje uppsättningen data. Datan innehöll dock inte tidskontinuerlig prisdata för dagen. Därför beräknades avkastningen för referensgruppen med hjälp av stängningspris för dag d, $S_C(d)$, och stängningspris för dag $d-1$, $S_C(d-1)$. Skattningen av marknadspåverkan kunde då modelleras enligt

$$\widehat{MI} = R - R_r, \text{ där } R = \frac{S_C(d) - S_C(d-1)}{S_C(d-1)}. \tag{8}$$

R är avkastningen från handlingsdagen d för den eftersökta aktien och R_r är den marknadsviktade avkastningen för aktiens referensgrupp under handelsdagen.

3.2.4 Skattning av marknadspåverkan med referensgrupper och avkastning mellan stängnings- och öppningspris

För att förbättra den tredje metoden ansågs det rimligt att försöka minimera eventuellt brus. Detta gjordes genom att minska tidsintervallet för avkastningen, vilket borde minimera prisfluktuationer som inte var till följd av den sökta marknadspåverkan.

Den fjärde metoden för att skatta marknadspåverkan använde därför likt den tredje sig av referensgrupper men beräknade avkastningen som skillnaden mellan stängningspriset för dag d, $S_C(d)$, och öppningspriset för nästa dag, $S_O(d+1)$. Detta var möjligt eftersom likt en stängningsauktion för varje dag finns det en öppningsauktion varje morgon. Vid öppningsauktionen lägger flertalet aktörer in budpriser på köp- och säljordrar innnan marknaden öppnar. Priserna jämförs sedan och ett öppningspris sätts för varje aktie, vilket de lagda ordrarna går igenom för.

Skattningen av marknadspåverkan kunde då modelleras enligt

$$\widehat{MI} = R_r - R, \text{ där } R = \frac{S_O(d+1) - S_C(d)}{S_C(d)}. \tag{9}$$

R är avkastningen från stängningspris, för dag d, till öppningspriset dagen efter för en aktie och R_r är den motsvarande marknadsviktade avkastningen för referensgruppen som aktien tillhör.

3.2.5 Operationella avgränsningar

Av de tre uppsättningar data innehållande information om flera köp- och säljordrar, vilken tillhandahölls av Andra AP-fonden, krävdes att ett flertal ordrar sållades bort. Nedan presenteras en lista med urvalsregler som användes på respektive skattning av marknadspåverkan samt hur dessa urval påverkade modellen.

För alla fyra sätt att skatta marknadspåverkan användes följande urvalsregler:

- Ordrar med ofullständig data togs bort eftersom de inte var applicerbara på modellen. Exempelvis återfanns felaktiga uppgifter vilka generade orimligt stor marknadspåverkan så därför raderades ordrar där marknadspåervkan var större än 75 %.
- Eftersom marknadspåverkan inte är lika framträdande då ordern utgör en liten del av den genomsnittliga handeln raderades de ordrar där antalet aktier som hade handlats utgjorde mindre än 0,1 % av den dagliga handeln. Denna gräns ansågs rimlig då en högre gräns hade uteslutit för stor del av den redan begränsade datan.
- Aktier från Qatar och Förenade Arabemiraten är borttagna på grund av bristfällig data.
- Ordrar som inte var 100 % genomförda togs bort då det inte var känt varför de inte gått igenom. Det kan exempelvis bero på handelsstopp vilket inte är relevant för modellen.
- Aktier handlade på den svenska marknaden sorterades bort eftersom de har begränsningar på priset, det vill säga aktien handlas inte om priset blir för högt eller lågt. Därmed begränsas marknadspåverkan och aktierna är således inte relevanta för modellen.
- Då marknadspåverkan endast ska behandlas för aktier raderades den data som innehöll information om andra typer av värdepapper.

Vid skattning av marknadspåverkan då det genomsnittliga transaktionspriset jämförs med priset när ordern träffar marknaden samt då det jämförs med första transaktionspriset:

- För att marknadspåverkan ska kunna uppstå krävs att hela ordern inte går igenom vid endast en transaktion eftersom priset då inte påverkas av ordern. Av denna anledning raderades ordrar som hade gått igenom på mindre än två transaktioner.

Vid skattning av marknadspåverkan med första transaktionspriset och det genomsnittliga priset användes även följande urvalsregler:

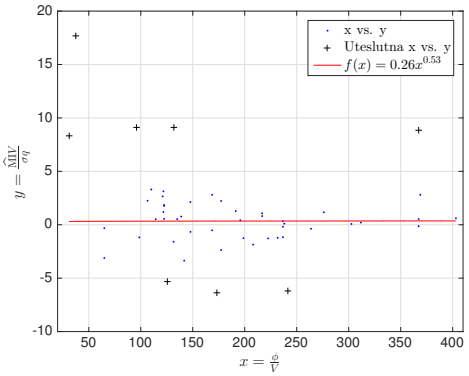
- Upprepningar av transaktioner raderades. Detta eftersom samma transaktion fanns med flera gånger i datan till följd av manuella fel från källan.

Efter avgränsningar för respektive region återstod 2080 ordrar för metoden då marknadspåverkan skattas med priset när ordern träffar marknaden och det genomsnittliga priset. För metoden då marknadspåverkan skattas med första transaktionspriset och det genomsnittliga priset återstod 886 ordrar. Metoden som använde referensgrupper och dagsavkastning för beräkning av marknadspåverkan hade 6298 ordrar kvar efter att avgränsningarna gjordes. Den sista metoden som använde referensgrupper och avkastning från stängnings- till öppningspris hade 6836 ordrar kvar efter avgränsningar.

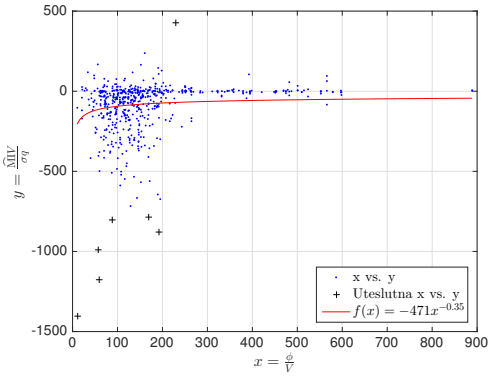
4 Resultat

För de fyra metoderna att skatta marknadspåverkan har $y = \frac{\widehat{MIV}}{\sigma q}$ analyserats som funktion av $x = \frac{\phi}{V}$ för respektive region, se avsnitt 3.2. Figur 3-6 visar uppmätta datapunkter (x, y) för utvalda regioner för samtliga skattningsmetoder. De visar även regressionkurvan efter borttagande av extremvärden. I figurerna redovisas även de två skattade konstanterna c_1 och c_2 från ekvation (2). Dessa anges med två decimalers noggrannhet för regressionskurvan $y = c_1 x^{c_2}$. Uteslutna extremvärden för de uppmätta datapunkterna är i figurerna markerade med $+$.

Figur 3 visar regressionskurvan och uppmätta datapunkter för regionerna MJ, figur 3a, och MN, figur 3b, då marknadspåverkan skattas med aktiens pris när ordern träffar marknaden och det genomsnittliga priset. I figur 4 redovisas plottar för två regioner då marknadspåverkan skattas med hjälp av första transaktionspriset och det genomsnittliga priset för ordern. De två regionerna är MP, figur 4a, och EW, figur 4b. För metoden att skatta marknadspåverkan med referensgrupper och dagsavkastning visas resultat för regionerna ME och MN i figur 5a respektive figur 5b. Figur 6a och 6b visar resultat för regionerna MN respektive EW när marknadspåverkan skattas med referensgrupper och avkastning mellan stängnings- och öppningspris. Samtliga regioner och skattningsmetoder återfinns i bilaga A.



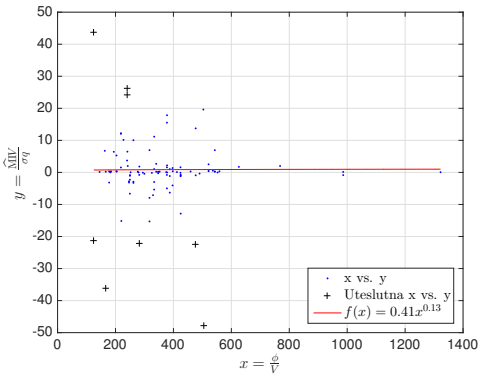
(a) I region MJ ges approximation av $y = 0{,}26x^{0{,}53}$



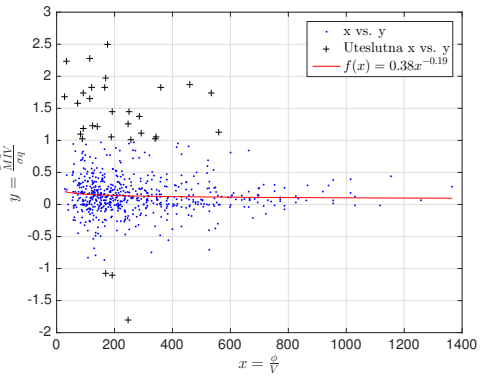
(b) I region MN ges approximation av $y = -471x^{-0{,}35}$

Figur 3: I figurerna plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ då marknadspåverkan skattas med hjälp av aktiens pris när ordern träffar marknaden och det genomsnittliga priset för ordern. Datapunkter som anses vara extremvärden är uteslutna och markerade som $+$ i graferna. Bästa $y = c_1 x^{c_2}$ approximation ges av den heldragna linjen i graferna.

Samtliga värden på konstanterna c_1 och c_2 och dess 95 % konfidensintervall för de olika regionerna redovisas i tabell 2-5 tillsammans med antal ordrar som analyseras. Tabellerna redovisar både det ursprungliga antalet ordrar som analyserades efter de avgränsningar som gjorts och antal ordrar efter exkluderande av extrempunkter. Tabell 2 visar värden på konstanterna, dess konfidensintervall och antal återstående ordrar då marknadspåverkan skattas genom priset då ordern träffar marknaden och det genomsnittliga priset för ordern. Tabell 3 visar värden på konstanterna, dess konfidensintervall samt antal återstående ordrar då marknadspåverkan skattas med hjälp av det första transaktionspriset och det genomsnittliga priset på ordern. Tabell 4 visar dessa uppgifter för metoden då marknadspåverkan skattas med hjälp av referensgrupper och dagsavkastning. Då marknadspåverkan skattas med referensgrupper men avkastningen beräknas som skillnaden mellan stängnings- och öppningspris redovisas resultat för samtliga värden i tabell 5.

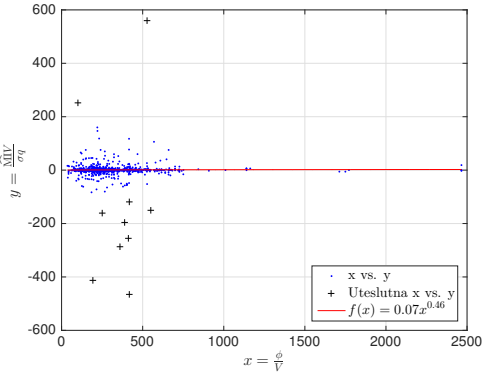


(a) I region MP ges approximation av $y = 0{,}41x^{0{,}13}$

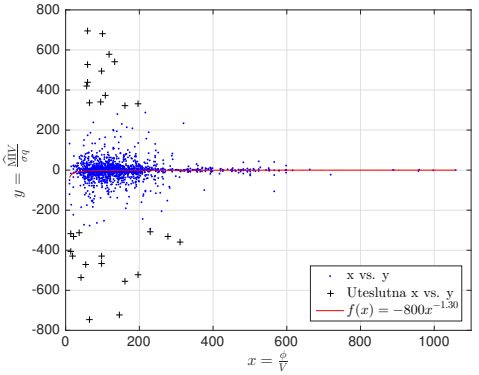


(b) I region EW ges approximation av $y = 0{,}38x^{-0{,}19}$

Figur 4: I figurerna plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ då marknadspåverkan skattas med första transaktionspriset och det genomsnittliga priset på ordern. Datapunkter som anses vara extremvärden är uteslutna och markerade som $+$ i graferna. Bästa $y = c_1 x^{c_2}$ approximation ges av den heldragna linjen i graferna.

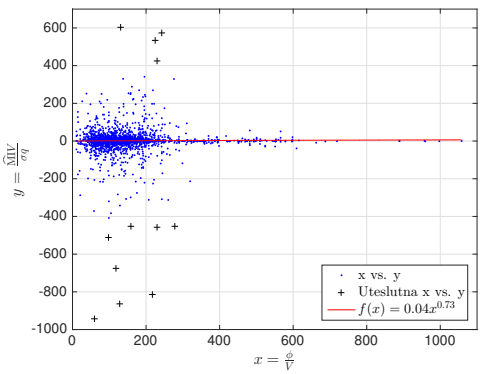


(a) I region ME ges approximation av $y = 0{,}07x^{0{,}46}$

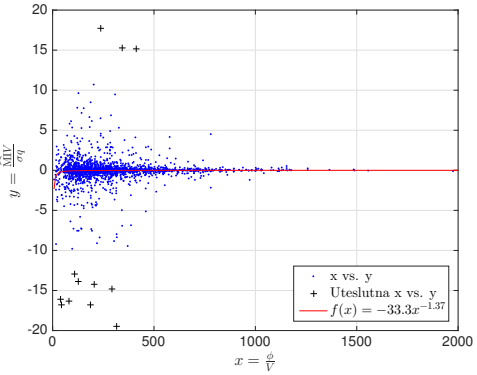


(b) I region MN ges approximation av $y = -800x^{-1{,}30}$

Figur 5: I figurerna plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ då marknadspåverkan skattas med referensgrupper och avkastningen räknas från stängningspris ena dagen till stängningspris nästa dag. Datapunkter som anses vara extremvärden är uteslutna och markerade som $+$ i graferna. Bästa $y = c_1 x^{c_2}$ approximation ges av den heldragna linjen i graferna.



(a) I region MN ges approximation av $y = 0{,}04x^{0{,}73}$



(b) I region EW ges approximation av $y = -33{,}3x^{-1{,}37}$

Figur 6: I figurerna plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ då marknadspåverkan skattas med referensgrupper och avkastningen räknas från stängningspris ena dagen till öppningspris nästkommande dag. Datapunkter som anses vara extremvärden är uteslutna och markerade som + i graferna. Bästa $y = c_1 x^{c_2}$ approximation ges av den heldragna linjen i graferna.

Tabell 2: Skattade värden på konstanterna c_1 och c_2 i ekvation (2), dess 95 % konfidensintervall samt antal ordrar både exklusive och inklusive extrempunkter för varje region. Marknadspåverkan skattas här som skillnaden mellan det genomsnittliga priset för ordern och aktiens pris när ordern träffar marknaden.

Region	$\mathbf{c_1}$ **(Konfidensintervall)**	$\mathbf{c_2}$ **(Konfidensintervall)**	**Antal ordrar (Inkl. extremvärden)**
ME	0,005764 $(-0{,}03175; 0{,}04328)$	1,375 $(0{,}3028; 2{,}448)$	348 (395)
MN	-471 $(-895{,}4; -46{,}51)$	$-0{,}3506$ $(-0{,}5428; -0{,}1585)$	568 (575)
MP	$-5{,}016 \cdot 10^5$ $(-2{,}296 \cdot 10^7; 2{,}196 \cdot 10^7)$	$-2{,}917$ $(-15{,}12; 9{,}287)$	199 (219)
MJ	0,2621 $(-4{,}421; 4{,}945)$	0,5263 $(-3{,}378; 3{,}483)$	40 (49)
EW	$-15{,}98$ $(-89{,}24; 57{,}28)$	$-0{,}5971$ $(-1{,}586; 0{,}3914)$	830 (842)

Tabell 3: Skattade värden på konstanterna c_1 och c_2 i ekvation (2), dess 95 % konfidensintervall samt antal ordrar både exklusive och inklusive extrempunkter för varje region. Marknadspåverkan skattas här som skillnaden mellan det genomsnittliga priset för ordern och första transaktionspriset.

Region	$\mathbf{c_1}$ **(Konfidensintervall)**	$\mathbf{c_2}$ **(Konfidensintervall)**	**Antal ordrar (Inkl. extremvärden)**
ME	32,25 $(-194{,}7; 259{,}2)$	$-1{,}174$ $(-2{,}518; 0{,}1711)$	90 (99)
MN	$-3{,}937 \cdot 10^6$ $(-8{,}148 \cdot 10^7; 8{,}935 \cdot 10^7)$	$-3{,}161$ $(-7{,}206; 0{,}8854)$	15 (17)
MP	0,4105 $(-7{,}625; 8{,}446)$	0,1252 $(-3{,}206; 3{,}456)$	89 (97)
MJ	0,01303 $(-0{,}1514; 0{,}1775)$	0,7356 $(-1{,}624; 3{,}095)$	37 (37)
EW	0,3797 $(-0{,}083; 0{,}8423)$	$-0{,}1883$ $(-0{,}4199; 0{,}04341)$	603 (636)

Tabell 4: Skattade värden på konstanterna c_1 och c_2 i ekvation (2), dess 95 % konfidensintervall samt antal ordrar både exklusive och inklusive extrempunkter för varje region. Marknadspåverkan skattas här med hjälp av referensgrupper och avkastningarna beräknas från aktiens stängningspris till aktiens stängningspris nästföljande dag.

Region	$\mathbf{c_1}$ **(Konfidensintervall)**	$\mathbf{c_2}$ **(Konfidensintervall)**	**Antal ordrar (Inkl. extremvärden)**
ME	0,06647 (−0,6436; 0,7765)	0,4587 (−1,346; 2,263)	1019 (1030)
MN	−800 (−2931; 1331)	−1,297 (−2,068; −0,5267)	1637 (1670)
MP	98,08 (−380; 576,2)	−0,7482 (−1,698; 0,2017)	476 (482)
MJ	6,95 (−125,5; 139,4)	−0,4161 (−4,383; 3,551)	1007 (1020)
EW	0,02742 (−1,034; 1,089)	−0,05411 (−7,228; 7,12)	2075 (2096)

Tabell 5: Skattade värden på konstanterna c_1 och c_2 i ekvation (2), dess 95 % konfidensintervall samt antal ordrar både exklusive och inklusive extrempunkter för varje region. Marknadspåverkan skattas här med hjälp av referensgrupper och avkastningarna beräknas från aktiens stängningspris ena dagen till aktiens öppningspris nästkommande dag.

Region	$\mathbf{c_1}$ **(Konfidensintervall)**	$\mathbf{c_2}$ **(Konfidensintervall)**	**Antal ordrar (Inkl. extremvärden)**
ME	−0,2389 (−3,981; 3,503)	0,1513 (−2,584; 2,886)	1155 (1161)
MN	0,03743 (−0,4634; 0,5382)	0,7347 (−1,703; 3,173)	1873 (1888)
MP	335,2 (−1940; 2610)	−1,231 (−2,717; 0,2555)	543 (554)
MJ	1,082 (−36,45; 38,61)	−0,07248 (−7,023; 6,878)	1134 (1139)
EW	−33,3 (−106,6; 40,04)	−1,367 (−2,076; −0,6585)	2079 (2094)

5 Diskussion

I tabell 2-5 redovisas värden för de beräknade konstanterna till ekvation (2) för de olika regionerna och skattningsmetoderna. Där ses att c_1 varierar kraftigt mellan väldigt stora negativa och positiva tal, men flertalet ligger dock inom intervallet $[-1; 1]$. I tabellerna ses även värden på exponenten c_2 vilka också de varierar en aning. Den största andelen ligger dock inom intervallet $[-1{,}3; 1{,}3]$. Ingen kombination av metod och region kan anses ge rimliga värden på dessa konstanter, vilket motiveras nedan. Sedan diskuteras även brister i modellen samt förbättringsmöjligheter vid uppföljning av projektet.

5.1 Analys av beräknade värden för konstanterna c_1 och c_2

Konstanten c_1 är som tidigare nämnts en skalfaktor för modellen. Denna bör vara positiv eftersom modellen korrigerades vid säljordrar med ett minustecken, vilket innebär att marknadspåverkan alltid definieras som en förlust då den är positiv för både köp- och säljordrar. Får c_1 ett negativt värde medför det att marknadspåverkan är negativ och därmed en vinst. Negativa värden på $y = \frac{\widehat{MIV}}{\sigma q}$ innebär att den skattade marknadspåverkan är negativ eftersom de resterande parametrarna alltid är positiva. I figur 3-6 kan vi se att det trots teckenkorrigering finns skattningar av marknadspåverkan som är negativa. Det kan bero på att andra aktörer har påverkat priset i motsatt riktning mot marknadspåverkan.

Då c_1 är en skalfaktor beror värdet på denna till stor del på marknaden. För en marknad som är oviss, exempelvis Japan, anses det ur ett teoretiskt perspektiv att värdet på c_1 är större än för andra, mer stabila marknader. Detta eftersom marknadspåverkan där kan bli större än i en mer stabil marknad. De beräknade konstanterna visar inte upp den förväntade strukturen utan de beror på hur marknadspåverkan har skattats. Det är svårt att dra en slutsats om rimliga värden för c_1 men ett värde över storleksordning 100 bör anses orimligt eftersom det skulle generera en väldigt stor marknadspåverkan med den antagna modellen.

Konstanten c_2 är exponent till kvoten mellan antal tillgängliga aktier för handel och den genomsnittliga volymen som handlas varje dag, $\frac{\phi}{V}$. Om kvoten är stor, det vill säga att av antalet tillgängliga aktier är det endast en liten del som handlas varje dag, blir aktien mer känslig för handel. Detta eftersom ordrarna då utgör en större del av dagshandeln och marknadspåverkan bör därmed bli större. Det leder till att c_2 bör vara positiv. Skulle exponenten vara negativ vänds kvoten till $\frac{V}{\phi}$ och ett litet värde på V i förhållande till ϕ innebär ett litet värde på kvoten och därmed en liten marknadspåverkan, vilket är motsägelsefullt.

Värdet på exponenten bör även vara mindre än ett, annars skulle marknadspåverkan växa väldigt snabbt när kvoten $\frac{\phi}{V}$ växer. I tabell 2 återfinns trots detta ett fall där c_2 är 1,375, det vill säga c_2 är större än ett. Detta genererade att c_1 är 0,005764 vilket är ungefär en tiondel mindre än de näst minsta värdena på c_1. I sex fall är c_2 mindre än minus ett. Dessa kan förkastas enligt motiveringen ovan men värt att notera är att det inte finns någon direkt mönster för värden på c_1 då c_2 är mindre än minus ett.

5.2 Analys av konfidensintervall för c_1 och c_2

Gemensamt för alla metoder för skattning, utom för fyra fall vilka diskuteras separat, är att konstanternas konfidensintervall innehåller talet noll. Det innebär att det inte går att förkasta att båda konstanterna också har värdet noll. Skulle konstanten c_1 vara noll fås en trivial modell eftersom modellen i sådant fall alltid skulle hävda att marknadspåverkan är obefintlig. Är däremot c_2 noll blir kvoten $\frac{\phi}{V}$ lika med ett och marknadspåverkan skulle vara oberoende av antal tillgängliga aktier. För de fyra fallen då konfidensintervallet inte innehåller noll är antingen någon av konstanterna negativ eller så är c_2 större än ett, vilket som tidigare motiverats är orimligt. Allt detta leder till att ingen slutsats om hur en rimlig modell skulle se ut kan göras.

Anledningen att noll finns med i konfidensintervallen kan bero på att en stor del av värdena för $y = \frac{\widehat{MIV}}{\sigma q}$ är negativa. Som tidigare nämnts innebär negativa y att den skattade marknadspåverkan är negativ. Det är därför rimligt att anta att metoderna för att skatta marknadspåverkan innehåller flertalet brister. Dessa brister tillsammans med eventuella fördelar för respektive metod kommer diskuteras närmare i avsnitt 5.7.

5.3 Analys av regionsuppdelning

För varje metod har konstanterna c_1 och c_2 beräknats utifrån respektive region. Detta gjordes eftersom marknaderna anses vara olika känsliga för den marknadspåverkan som uppstår då Andra AP-fonden placerar sina tillgångar där. Den varierande känsligheten leder således till att olika regioner bör ha olika marknadspåverkan. Det är dock problematiskt att dra en slutsats huruvida detta är sant eller inte eftersom det inte framkommit något användbart resultat för de olika regionerna. Det finns inte heller någon trend i att ovanstående skattning eller att en specifik region skulle följa ett visst mönster. Resultatet påvisar inte heller om det existerar en marknadspåverkan eller inte när Andra AP-fonden handlar i olika regioner.

Då Andra AP-fonden efterfrågade olika modeller för olika regioner har det inte undersökts om en gemensam modell för marknadspåverkan för samtliga regioner skulle stämma bättre överens med verkligheten. Detta bör ge ett sämre resultat då marknaderna skiljer sig i stabilitet och därmed bör generera olika marknadspåverkan. Huruvida den regionsuppdelningen som gjordes här är den mest optimala har inte heller analyserats utan den efterfrågade regionsuppdelningen användes konsekvent genom analysen.

5.4 Analys av data

Med tanke på att en stor del av de uppmätta datapunkterna är negativa är det troligast att det är metoderna för skattning av marknadspåverkan som brister. Det behöver dock inte vara den enda faktorn som gör att en slutsats inte kan dras. En annan anledning kan vara att det var en för liten mängd data som användes vid skattningarna. De två första metoderna analyserade betydligt färre ordrar än de två sista. För dessa metoder är det för alla regioner utom EW omöjligt att styrka något resultat och även för EW är detta problematiskt då det är högst troligt att det var för lite ordrar som analyserades. De andra två metoderna analyserade visserligen mer data men eftersom de inte genererade rimliga resultat kan det även här vara för få ordrar som analyserats. Eftersom finansmarknaden är stokastisk är det möjligt att färre datapunkter, i det här fallet färre ordrar, ger en sämre bild av verkligheten och att ett resultat givet endast ett fåtal ordrar inte är representativt för marknaden.

Huruvida en ökad datamängd hade genererat ett bättre resultat är svårt att dra några slutsatser kring. Värt att notera är att i den bakomliggande rapporten [11] fanns ursprungligen närmare 700 000 ordrar för en region under en period på 19 månader. Av de 700 000 ordrarna kvarstod ungefär 30 000 stycken efter ett flertal avgränsningar. Det är ungefär 14 gånger så mycket data som analyserades som mest för någon region i detta projekt. Det finns inget annat än den bakomliggande rapporten som styrker att mer data skulle ge ett bättre resultat. Mer data skulle dock ge en bättre inblick i hur sanningsenliga skattningsmetoderna är.

Längden på tidsperioden som den analyserade datan härrör från skiljer sig också mot den bakomliggande rapportens. Den data som tillhandahölls av Andra AP-fonden innehöll handel från två till tre månader medan den bakomliggande rapporten använde data från 19 månader. Eftersom det är rimligt att det finns vissa konjukturcykler över året samt att marknadspåverkan kan vara olika vid olika tider på året är det troligt att även detta är en bidragande faktor till de orimliga resultaten.

5.5 Brister i modellen

Låt oss anta att de fundamentala problemen inte ligger i hur marknadspåverkan skattas. Det skulle innebära att det är den bakomliggande modellen [11] som är bristfällig. Hur väl den bakomliggande modellen verkligen fungerar och även processen bakom den är vagt formulerad i den publicerade rapporten. Det finns, som tidigare nämnts, inte mycket publicerat inom området. Anledningen till det kan vara att när framsteg görs publiceras inte dessa eftersom ett placeringssätt fungerar bäst då ingen annan placerar på samma sätt.

En viktig skillnad mellan Almgrens fullständiga modell och den modell som används i denna rapport är att den fullständiga modellen har en term för den permanenta marknadspåverkan som uppstår. Hur termen skulle påverka resultatet är svårt att dra några slutsatser kring men det kan antas att termen borde gett rimligare resultat. I den bakomliggande rap-

porten har det även gjorts en avgränsning som innebar exkluderande av alla ordrar där det efterfrågades att ordern skulle handlas på öppnings- eller stängningsauktion. Motiveringen var att modellen inte var anpassad för sådana transaktioner. Detta leder till att modellen som användes i denna rapport kan vara dåligt lämpad för Andra AP-fondens handel då större delen av deras handel sker på just öppnings- eller stängningsauktion. Att modellen ändå valdes motiveras med att det är svårt att hitta andra publicerade modeller att utgå ifrån, samt att fondens nuvarande modell bygger på den bakomliggande modellen.

Då den bakomliggande modellen endast använder data från Nordamerika och marknaderna skiljer sig åt en aning var det mest troligt att modellen skulle ge ett bra resultat för handel i Nordamerika. Tabell 2-5 visar att så var inte fallet. Att modellen inte ger ett bra resultat för de skattningsmetoder av marknadspåverkan med referensgrupper och skillnad i avkastning är rimligt. Detta då i dessa fall sker den största delen av handeln på stängningsauktionen och de ordrarna hade sorterats bort i den bakomliggande rapporten. Däremot borde modellen ha gett ett bättre resultat i de fall när marknadspåverkan skattades med det genomsnittliga priset och antingen första transaktionspriset eller priset när ordern träffar marknaden. Detta eftersom då behandlas endast de ordrar som gått igenom vid flera transaktioner, vilket innebär att de inte är handlade vid stängningsauktionen utan sekventiellt under dagen. Anledningen till att modellen ändå inte funkar för dessa två fall kan som tidigare nämnts bero på att det finns för lite data.

5.6 Diskussion kring uteslutande av extremvärden

Vid beräkningen av konstanterna med *cftool* togs extremvärden bort om de ansågs orimliga för modellen. Extremvärdena skiljde sig markant från resten av de observerade datavärdena och detta kan exempelvis bero på fel i datan eller ovanliga värden. I tabell 2-5 redovisas hur många extremvärden som tagits bort för varje kombination av region och metod för skattning. Slutsatsen som kan dras från dessa tabeller är att ju fler datapunkter, ordrar, som analyseras desto mindre del av punkterna anses vara extremvärden. Det vill säga att procentuellt tas färre antal punkter bort. Detta skulle innebära att en ökad mängd data skulle leda till en mer sanningsenlig bild av aktiemarknaden och de skattade konstanterna skulle kunna ge ett rimligare resultat.

Ytterligare en anledning som styrker att resultatet kan ifrågasättas är att konstanternas värden är ytterst beroende av vilka punkter som definieras som extremvärden. Exkluderas eller inkluderas en viss punkt kan detta generera helt skilda resultat vilket inte är önskvärt för en stabil modell. I det framtagna resultatet har endast de punkter som ligger på ett avsevärt avstånd från de andra ansetts som extremvärden. Det är svårt att dra en exakt gräns och det är därför fullt rimligt att för många eller för få punkter har hanterats som extremvärden.

Idealiskt hade varit att analysera varje extrempunkt individuellt. Det genom att undersöka hur datan samlats in samt vilka förutsättningar och händelser som påverkade ordern. Tyvärr fungerar inte finansmarknaden idealiskt vilket innebär att detta inte är genomförbart i praktiken. Modellen kommer alltså behöva utesluta vissa typer av handel. Detta är inte något unikt utan större delen av det som publiceras inom ekonometri är modeller som endast är applicerbara under specifika omständigheter.

5.7 Felanalys av skattningsmetoderna

En trolig anledning till det orimliga resultatet är att det kan finnas stora brister i hur de olika metoderna skattar marknadspåverkan. Detta gäller framför allt om det antas att den bakomliggande modellen är korrekt. Nedan följer därför en diskussion kring de olika metodernas för- och nackdelar.

5.7.1 Skattning av marknadspåverkan när ordern träffar marknaden och genomsnittliga priset

Skattningen bygger på att jämföra ett pris innan handeln startar, priset på aktien då ordern träffar marknaden, med det genomsnittliga priset för ordern. Ur ett teoretiskt perspektiv är

det ett rimligt sätt att skatta marknadspåverkan. Detta eftersom innan ordern har börjat genomföras bör marknadspåverkan vara noll, och när den är genomförd bör det genomsnittliga priset vara startpriset tillsammans med den genomsnittliga marknadspåverkan. Problemet är att priset innan handeln startar inte är direkt kopplat till Andra AP-fondens transaktioner, och därmed inte är det pris som de börjar handla för. Det skapas därför en potentiell källa för brus eftersom priset kan förändras från det att ordern träffar marknaden till det att den första transaktionen har gått igenom.

För att minska det potentiella bruset hade det var mer lämpligt att jämföra det genomsnittliga priset för ordern med det första transaktionspriset. Detta eftersom i det fallet är det garanterat att båda priserna är kopplade till Andra AP-fondens handel och inte till marknadens. I denna metod gjordes ett försök att korrigera för bruset genom att skala om det genomsnittliga priset utifrån hur marknaden har gått. Det ansågs vara en rimlig korrektion att göra för att kunna särskilja vad som är marknadspåverkan och vad som är marknadens fluktuationer. Dock leder detta fortfarande till ett orimligt resultat eftersom denna skattning ger för mycket negativa värden på y. Detta beror kanske på att korrektionen också skapar brus då den endast avspeglar marknadens rörelser. Det kan hända något med marknaden som inte påverkar Andra AP-fondens handel och således blir korrektionen felaktig.

Den här metoden för skattning kan bara användas för handel som sker sekventiellt och därmed exkluderas all handel som sker på stängningsauktionen. Då större delen av Andra AP-fondens handel sker på just stängningsauktionen är denna metod inte optimal ur deras perspektiv. Det var anledningen till att den här samt metoden med första transaktionspriset inte analyserades på en större mängd data, vilket gjordes med de två skattningsmetoderna med referensgrupper.

5.7.2 Skattning av marknadspåverkan med första transaktionspriset och genomsnittliga priset

Likt den ovan nämnda metoden för att skatta marknadspåverkan jämförs i denna metod ett pris vid orderns start med det genomsnittliga priset för ordern. Men denna metod bygger endast på priser för ordern vilket bör innebära att den är bättre ur ett brusminimerande perspektiv än den förra metoden. Huruvida detta är sant eller inte är svårt att analysera eftersom det finns för lite data för att kunna dra några slutsatser om skattningen. Denna svårighet återspeglas även på konfidensintervallen, där bredden på intervallen tyder på att det är för få punkter för att inte få en spridning på mellan punkterna.

Denna metod anses vara den mest rimliga skattningen av de två hittills nämnda metoderna ur ett teoretiskt perspektiv. Fonden handlar endast en liten del av sina ordrar på detta vis vilket gör att skattningen inte är så applicerbar för dem och därför undersöktes denna inte vidare. Alltså är den troligtvis den mest korrekta metoden men också den minst givande.

5.7.3 Skattning av marknadspåverkan med referensgrupp och daglig avkastning

Den största delen av Andra AP-fondens handel sker som tidigare nämnts på stängningsauktion. Den tredje skattningsmetoden byggde på sådan handel och beräknades som skillnaden i daglig avkastning mellan aktien och en referensgrupp för aktien. Det gör att den här metoden är betydligt mer applicerbar för Andra AP-fonden än de tidigare diskuterade metoderna.

Referensgrupperna fungerar som en korrigering för bruset som uppstår men trots det fås inga rimliga resultat. En möjlig anledning att denna metod inte ger rimliga resultat kan vara att aktien inte har följt referensgruppens prisförändring och därmed uppstår en felaktig korrektion. Resultaten kan även bero på att i skattningsmetoden användes priser på aktier från stängningsauktion ena dagen till stängningsauktion nästkommande dag. Det innebär att metoden inte tar hänsyn till händelser under dagen som påverkar marknaden och att den utgår från att det endast är fondens handel som påverkar marknaden. Då detta inte sker i verkligheten är det som egentligen söks, för att kunna skatta marknadspåverkan, priset precis innan stängningsauktionen och priset från stängningsauktionen. Detta för att undvika de fluktuationer som sker på marknaden under dagen efter transaktionen. Även val av referensgrupp kan vara avgörande för att generera rimliga resultat. Här har endast givna

referensgrupper från Andra AP-fonden använts och därför har de inte analyserats vidare ur ett felanalysperspektiv.

Problemet med att skatta marknadspåverkan med priser vid stängningsauktion är att det är många ordrar utöver Andra AP-fondens som påverkar priset vid stängningsauktionen. Vid handel under dagen kan däremot andra aktörer märka om en aktör handlar stora kvantiteter och därmed kan de anpassa sin handel utefter marknadsförändringarna. Vid handel på stängningsauktion fås inte denna förvarning utan flera stora aktörer kan försöka handla samtidigt och på så sätt tillsammans skapa en större marknadspåverkan.

5.7.4 Skattning av marknadspåverkan med referensgrupper vid öppningspris

Istället för att skatta marknadspåverkan genom att se till avkastningen från stängningsauktion till stängningsauktion använder denna metod avkastningen från stängningsauktion till öppningsauktion. Detta borde ge en bättre skattning då en del av bruset försvinner till följd av att tidsperioden förkortas och det inte hinner hända lika mycket på marknaden. Tabell 5 visar att resultaten inte är rimliga för denna metod heller. En anledning till det kan vara att det är för kort tid mellan de två auktionerna för att marknadspåverkan ska hinna klinga av. Därmed skulle även öppningspriset ha påverkats av marknadspåverkan från stängningsauktionen. Även val av referensgrupper kan påverka utgången, precis som nämndes i föregående skattningsmetod, detta är dock något som inte analyserats vidare av samma anledning som ovan.

Ytterligare en källa till brus kan vara att öppningspriset beror på vilka ordrar som lagts, vilka i sig kan leda till marknadspåverkan, och de händelser som skett under natten. Dessa ordrar och händelser påverkar marknaden och det pris som sätts på aktien vid öppningsauktionen men har egentligen inget att göra med marknadspåverkan från Andra AP-fondens ordrar. Det finns inga data på priset innan öppningsauktionen gått igenom, eftersom börsen är stängd, och därmed går det inte att förbättra denna skattningsmetod på det sättet. Ett möjligt alternativ att förbättra metoden kan vara, precis som nämndes i den tidigare skattningen med referensgrupper och avkastning, att analysera priset innan stängningsauktionen och priset vid stängningsauktionen.

5.8 Utvecklingspotential

I styckena ovan påvisas att det finns flertalet brister vid framställning av den tänkta modellen, vilket medför en stor utvecklingspotential för projektet. Nedan diskuteras flertalet områden vilka har identifierats med möjlig förbättringskapacitet. Tidsbegränsning av projektet har lett till att dessa förbättringsmöjligheter inte har genomförts.

5.8.1 Utveckling av data

En potentiell förbättring för projektet är att analysera mer data över en längre tidsperiod. Det för att kunna skatta marknadspåverkan mer tillförlitligt. Som nämnts tidigare hade den bakomliggande rapporten ursprungligen närmare 700 000 ordrar och efter avgränsningar närmare 30 000 ordrar. De ordrarna härrör dessutom enbart från Nordamerika, MN, vilket endast är en geografisk region. Alltså analyseras där 30 000 ordrar för en region medan i denna rapport analyseras 1873 ordrar som mest för en region. Hur många ordrar som krävs för att skapa tillförlitliga resultat är svårt att avgöra, men desto fler ordrar som finns att tillgå desto mer sanningsenliga bör skattningarna bli.

Finansiella marknader är dock inte statiska, så en längre tidsperiod av data ger ingen garanti för att skapa en mer tillförlitlig modell. Används för gammal data, som inte längre är representativ, kan det leda till att modellen inte längre är aktuell. Därmed behöver modellen kontinuerligt utvärderas och uppdateras samt i vissa fall förkastas. Eftersom den bakomliggande modellen framtogs 2005 [11] kan det innebära att den modellen som används i projektet inte längre är aktuell och en förbättring av projektet skulle kunna vara att utgå från en annan modell.

När det gäller att skatta marknadspåverkan hade det varit önskvärt med andra prisuppgifter. En möjlighet är att skatta marknadspåverkan utifrån priset på stängningsauktionen

och ett pris direkt efter stängningsauktionen. Detta för att minska påverkan från andra händelser och ordrar från andra aktörer på den marknadspåverkan som är skapad av Andra AP-fonden. Det finns inga prisuppgifter mellan stängnings- och öppningsauktionen vilket innebär att detta inte går att genomföra eftersom börsen då är stängd.

En prisuppgift som finns att tillgå är priset precis innan stängningsauktionen. Marknadspåverkan kan då skattas som skillnaden mellan priset precis innan och priset på stängningsauktionen. När många aktörer är med och påverkar priset kan det även leda till en permanent marknadspåverkan och denna metod skulle skatta summan av den permanenta och den temporära marknadspåverkan. Detta är något som kan korrigeras genom att ta hänsyn till prisförändringen mellan priset innan auktionen och priset på stängningsauktionen för en referensgrupp.

5.8.2 Utveckling av modell

Något som inte har undersökts närmare är hur stor del av den dagliga genomsnittliga handeln som ordern utgör. Vid en kort undersökning märktes att när andelen var stor fanns det ordrar där marknadspåverkan blev stor. För att utveckla strukturen i den befintliga modellen är detta därför något som kan undersökas närmare i ett framtida projekt. Optimalt hade kanske varit att dela upp datan efter hur stor andel av den dagliga genomsnittliga handeln som ordern utgör. Detta för att om möjligt resultera i olika värden på konstanten c_2 beroende på storleken på kvoten.

Andra AP-fonden krävde att beroendet på antalet aktier som ska handlas är antingen linjärt eller kvadratiskt i modellen. Tidsbegränsning har gjort att endast det linjära fallet har undersökts. Det skulle därför vara rimligt att också undersöka hur en modell med ett kvadratiskt beroende på antal aktier som ska handlas skulle fungera för att skatta marknadspåverkan.

Den bakomliggande rapporten har valt att utesluta de ordrar som handlas på öppnings- och stängningsauktion. Detta eftersom de anser att dessa transaktioner har icke-linjära egenskaper [11]. Strukturen av modellen kan därför behöva modifieras för att vara applicerbar även på stängnings- och öppningsauktion.

5.8.3 Utveckling av skattningsmetoder

För att få en mer korrekt jämförelse mellan två skattningsmetoder bör samma data analyseras för båda metoderna. Detta har inte gjorts utan olika avgränsningar har applicerats för de olika skattningsmetoderna. En annan anledningen till att det blev olika ordrar som analyserades är att vissa ordrar saknade den information som krävdes för en viss skattningsmetod. Därmed uteslöts den ordern för just den skattningsmetoden. Dessa ordrar kunde dock ha tillräckligt med information för att användas med en annan skattningsmetod och har därför fått vara kvar efter avgränsningarna för den skattningsmetoden.

Att kunna analysera samma data är däremot inte genomförbart vid jämförelse mellan metoder som tar hänsyn till sekventiell handel och metoder där ordrar handlas på stängningsauktion. De ordrar som handlas sekventiellt under dagen kräver fler än en transaktionen för att marknadspåverkan ska kunna uppstå. Detta skiljer sig mot de ordrar som handlas på stängningsauktion då de endast kräver en transaktion för att genomföras. Kriterierna är alltså motsägelsefulla vilket innebär att det inte finns något nämnvärt överlapp av ordrar mellan dessa metoder. Därför är det inte möjligt att undersöka samma ordrar för dessa kategorier av skattningsmetoder.

De två skattningsmetoderna baseras på två olika typer av handel. Därför kan det vara av intresse att undersöka möjligheten att skatta marknadspåverkan på två olika sätt. Ett sätt för de sekventiella ordrarna och ett sätt för de ordrar som handlats på stängningsauktion. Alla skattade värden kan då sammanställas och appliceras på samma modell. Detta för att modellen ska vara tillämpbar på all typ av handel som Andra AP-fonden hanterar. På så sätt fås parametrar c_1 och c_2 som kan användas för både sekventiella ordrar såväl som ordrar som genomförs på stängningsauktion.

6 Slutsats

Att modellera marknadspåverkan är likt många ekonometriska problem en utmaning. Datan som analyseras härrör från stokastiska marknader och innehåller därmed mycket brus. Marknadspåverkan som effekt är inte heller direkt mätbar och det finns lite publicerat inom området. De publicerade modellerna har dock strukturer som liknar den bakomliggande modell [11] som har använts i detta projekt.

I projektet analyserades fyra sätt att skatta marknadspåverkan utan att finna ett användbart resultat. Denna slutsats kan styrkas genom att konfidensintervallen på de två skattade konstanterna är delvis stora och delvis innehåller flertalet av intervallen noll. Därför går det inte att dra någon slutsatsen att modellen är statistiskt signifikant.

De bristfälliga resultaten kan vara till följd av den stora mängd brus som fanns i den givna datan. De kan även vara till följd av hur väl utformade de analyserade skattningsmetoderna var. Antas dessa faktorer inte vara dominerande kan även det fundamentala felet vara kopplat till modellens utformning. Alla dessa orsaker leder dock till att det finns stor utvecklingspotential för projektet.

Referenser

[1] Granit Fonder. Portföljteori: Avkastning och risk. http://www.granitfonder.se/upload/2015/fondskolan/Fondskolan_Portföljteori%20Avkastning%20och%20Risk.pdf, 1 juni 2016.

[2] N.G. Torre & M.J. Ferrari. The market impact model. Technical report, Barra, 1999.

[3] G.S Maddala. *Introduction to Econometrics.* Macmillan Publishing Company, 2nd ed. edition, 1992.

[4] J.M. Wooldridge. *Introductory econometrics: a modern approach.* Thomson/South-Western, cop., 2006.

[5] S. Hurn, V. Martin, P. Phillips & J. Yu. Financial econometric modelling. http://www.mysmu.edu/faculty/yujun/MSFE_FEc/PartI.PDF, 2015.

[6] X. Yan. *Linear Regression Analysis: Theory and Computing.* World Scientific Publishing Company, 2009.

[7] J. Frost. What is the difference between linear and nonlinear equations in regression analysis? http://blog.minitab.com/blog/adventures-in-statistics/what-is-the-difference-between-linear-and-nonlinear-equations-in-regression-analysis, 9 februari 2016.

[8] Wolfram. Gauss–newton methods. https://reference.wolfram.com/language/tutorial/UnconstrainedOptimizationGaussNewtonMethods.html, 1 Juni 2016.

[9] MathWorks. Describe nonlinear relationships and make predictions from experimental data. http://se.mathworks.com/discovery/nonlinear-regression.html, 1 juni 2016.

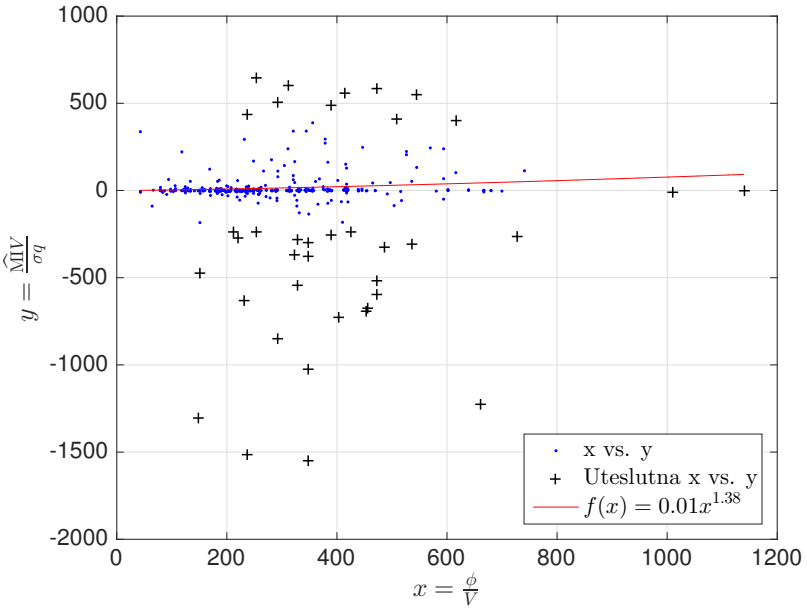
[10] Captor. Kostnaden för ”market impact”. http://captor.se/kostnaden-for-market-impact/, 7 februari 2015.

[11] E. Hauptmann & Hong Li R. Almgren, C. Thum. Direct estimation of equity market impact. Technical report, University of Toronto Departments of Mathematics and Computer Science, and Citigroup Global Quantitative Research, 2005.

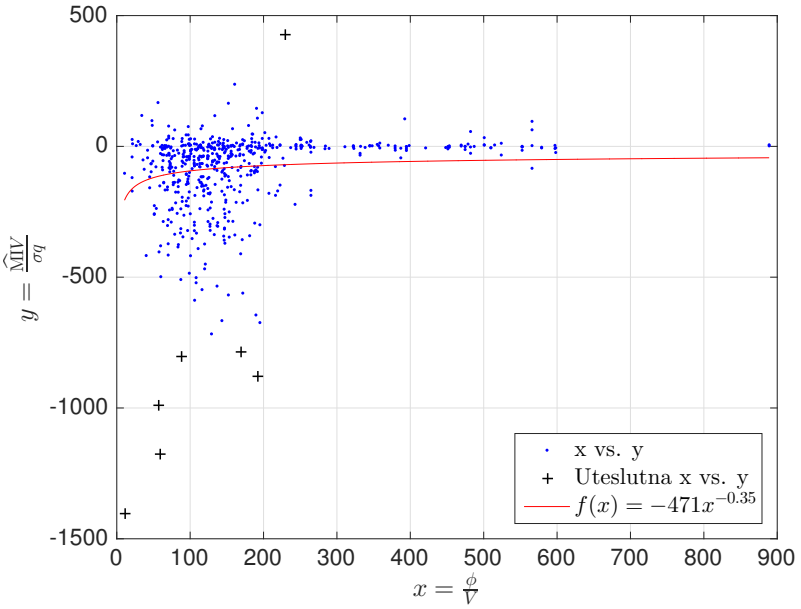
A Figurer för skattning av marknadspåverkan

I A.1-A.4 återfinns resultatet för respektive skattningsmetod. För varje metod visas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för respektive region. De visar även regressionkurvan efter borttagande av extremvärden. I figurerna redovisas även de två skattade konstanterna c_1 och c_2 från ekvation (2). Dessa anges med två decimalers noggrannhet för regressionskurvan $y = c_1 x^{c_2}$. Uteslutna extremvärden för de uppmätta datapunkterna är i figurerna markerade med $+$.

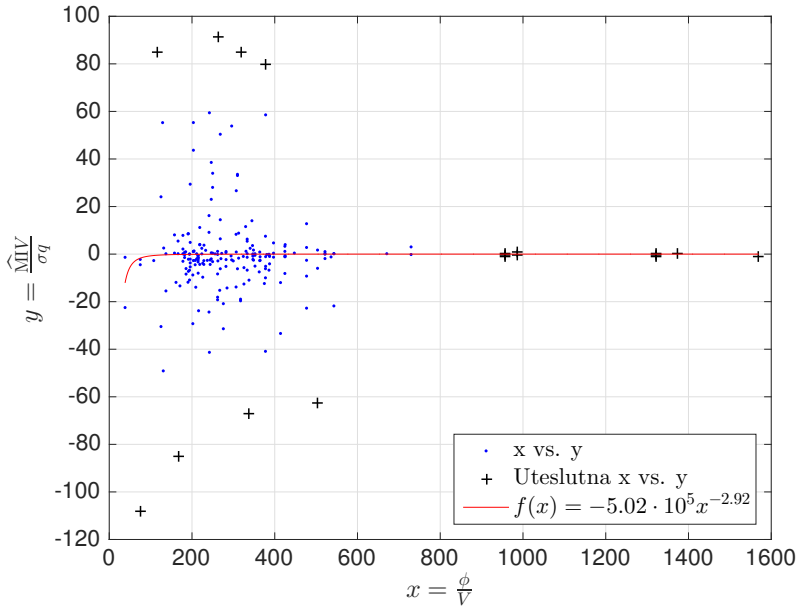
A.1 Figurer för skattning av marknadspåverkan med priset när ordern träffar marknaden och det genomsnittliga priset



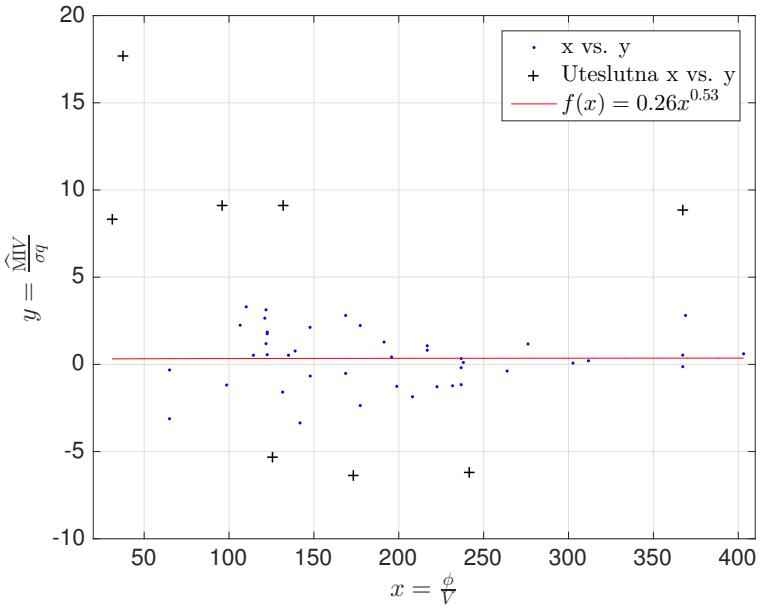
Figur 7: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region ME med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som $+$ i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}01x^{1{,}38}$ som är den heldragna linjen i grafen.



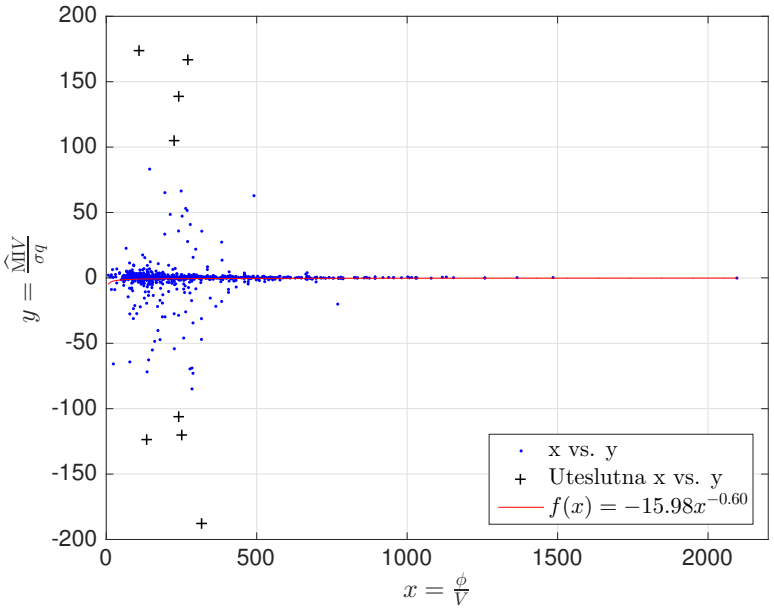
Figur 8: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MN med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -471x^{-0,35}$ som är den heldragna linjen i grafen.



Figur 9: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MP med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -5{,}02x^{-2,92}$ som är den heldragna linjen i grafen.

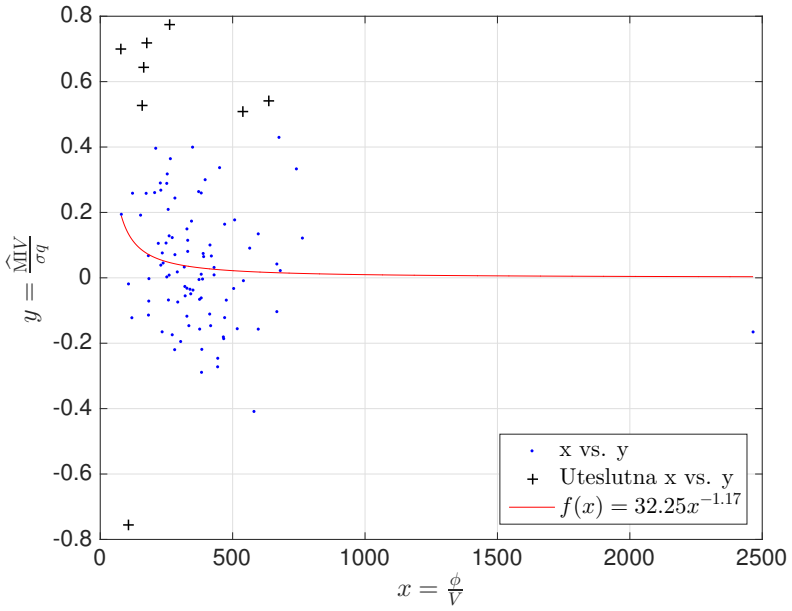


Figur 10: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MJ med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}26x^{0{,}53}$ som är den heldragna linjen i grafen.

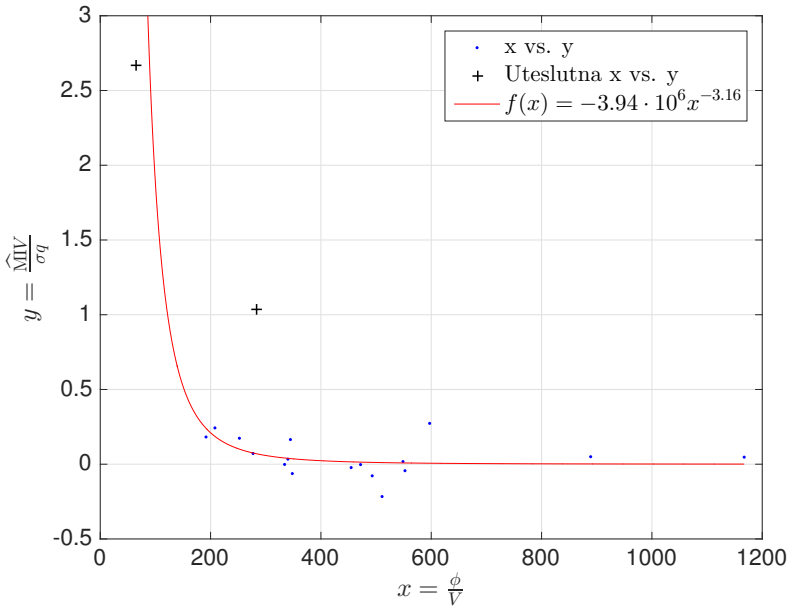


Figur 11: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region EW med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -15{,}98x^{-0{,}60}$ som är den heldragna linjen i grafen.

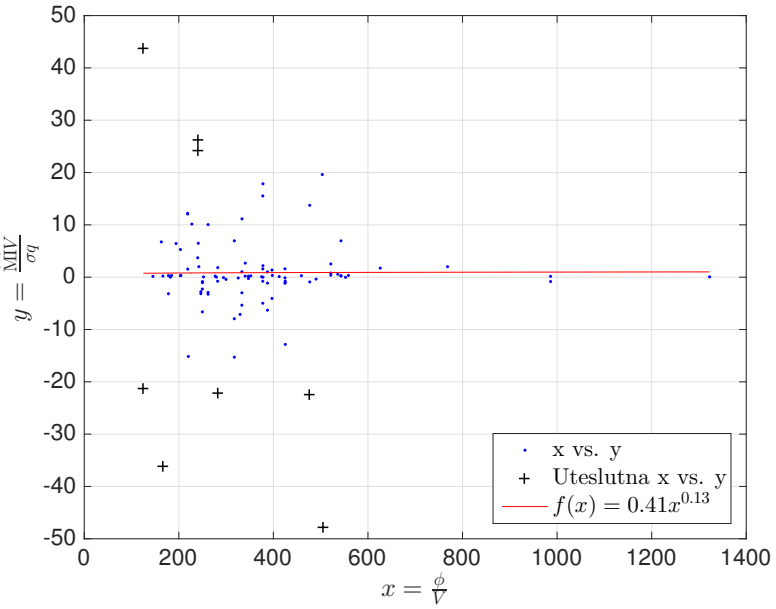
A.2 Figurer för skattning av marknadspåverkan med första transaktionspriset och det genomsnittliga priset



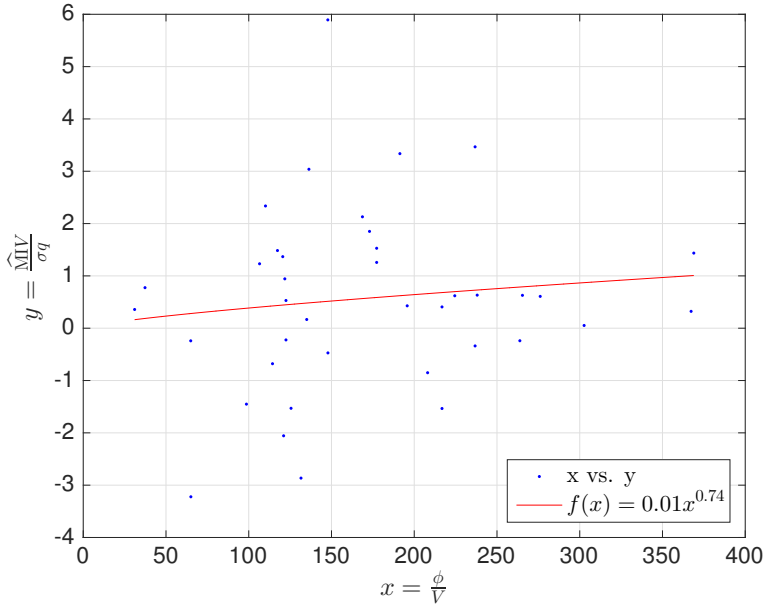
Figur 12: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region ME med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 32{,}25x^{-1{,}17}$ som är den heldragna linjen i grafen.



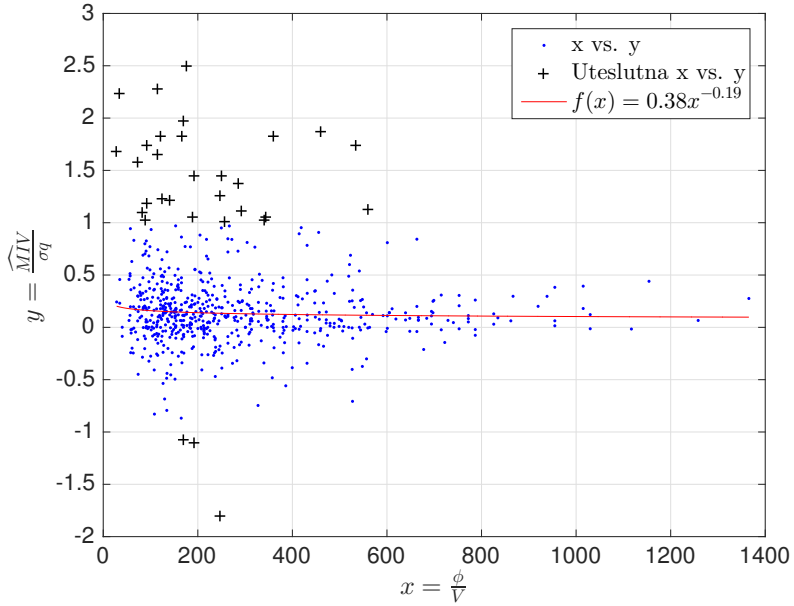
Figur 13: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MN med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -3{,}94 \cdot 10^6 x^{-3{,}16}$ som är den heldragna linjen i grafen.



Figur 14: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MP med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}41x^{0{,}13}$ som är den heldragna linjen i grafen.

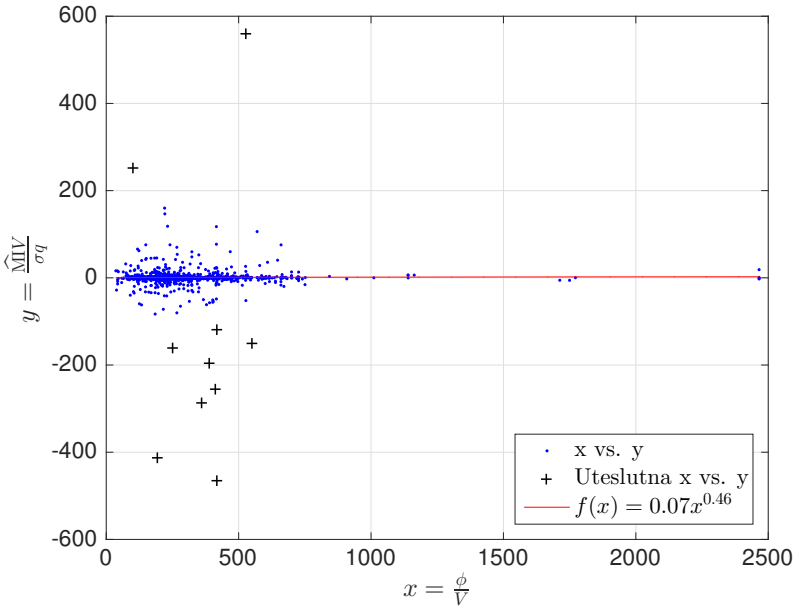


Figur 15: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MJ med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}01x^{0{,}74}$ som är den heldragna linjen i grafen.

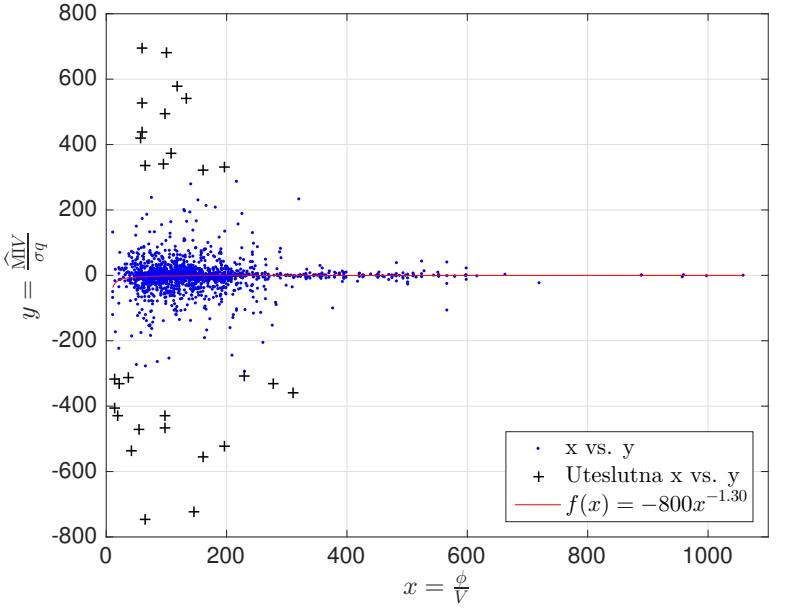


Figur 16: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region EW med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}38x^{-0{,}19}$ som är den heldragna linjen i grafen.

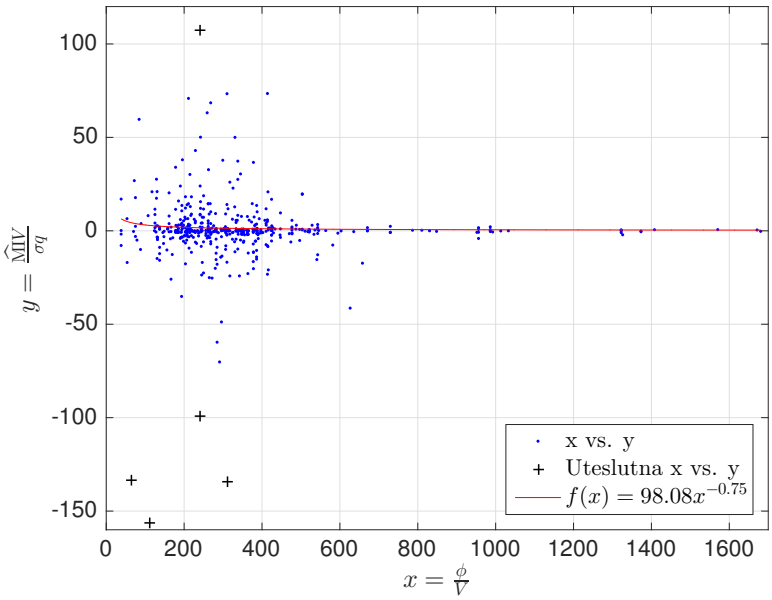
A.3 Figurer för skattning av marknadspåverkan med referensgrupper och dagsavkastning



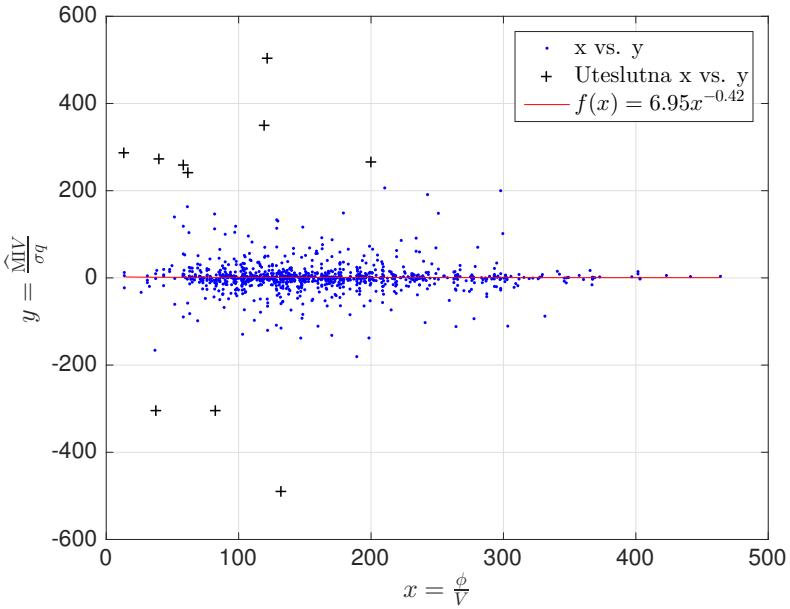
Figur 17: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region ME med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}07x^{0{,}46}$ som är den heldragna linjen i grafen.



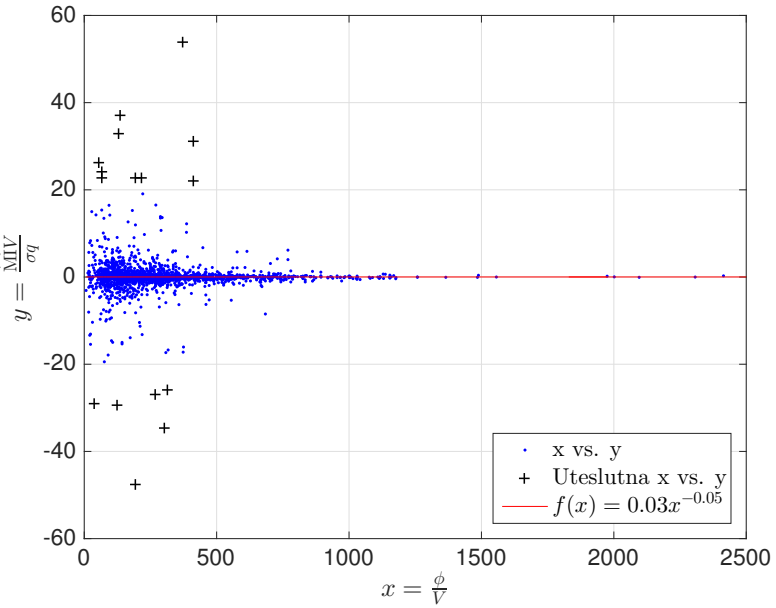
Figur 18: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MN med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -800x^{-1{,}30}$ som är den heldragna linjen i grafen.



Figur 19: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MP med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 98{,}08x^{-0{,}75}$ som är den heldragna linjen i grafen.

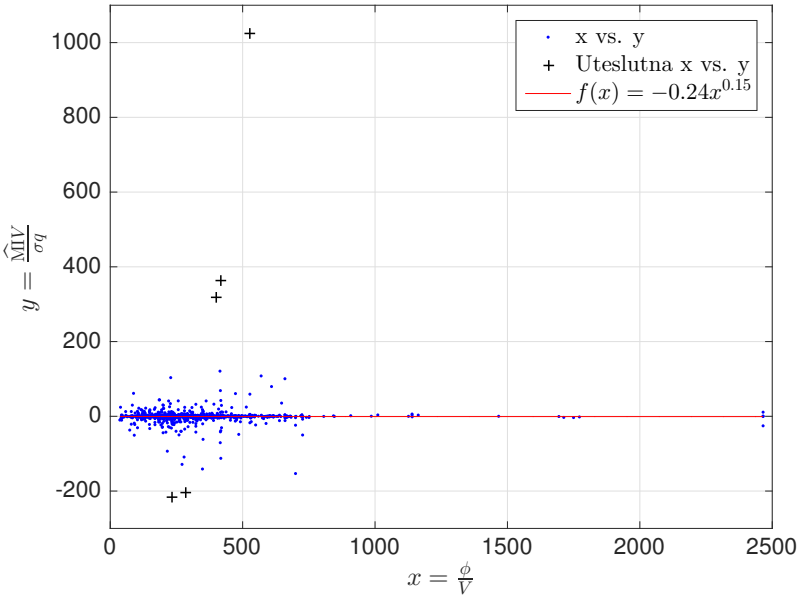


Figur 20: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MJ med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 6{,}95x^{-0{,}42}$ som är den heldragna linjen i grafen.

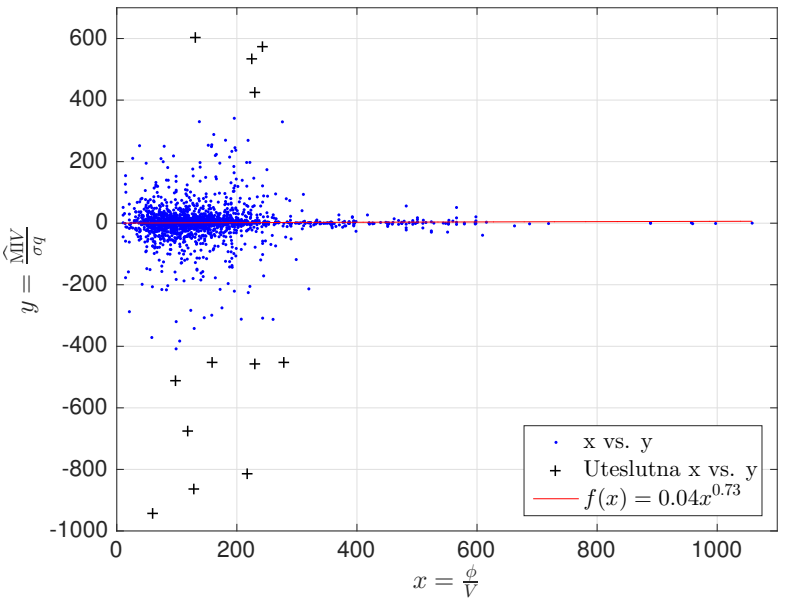


Figur 21: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region EW med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}03x^{-0{,}05}$ som är den heldragna linjen i grafen.

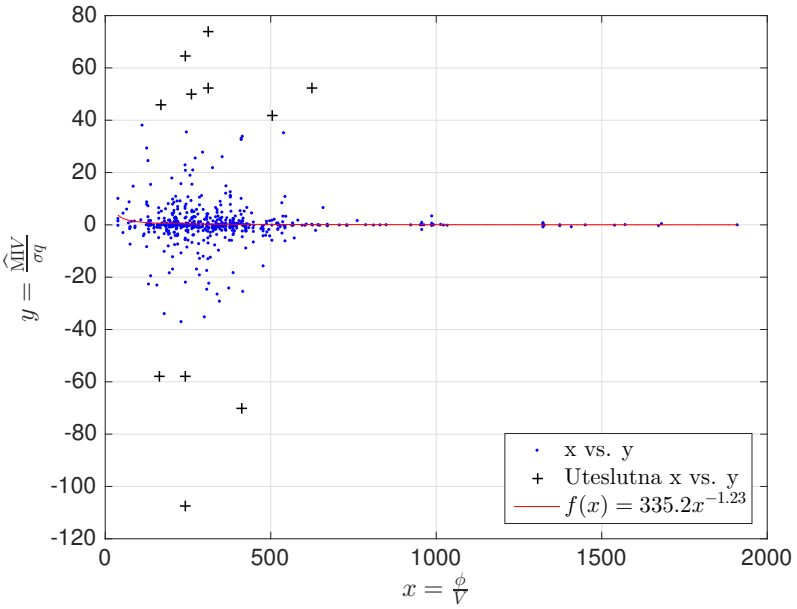
A.4 Figurer för skattning av marknadspåverkan med referensgrupper och avkastning mellan stängnings- och öppningspris



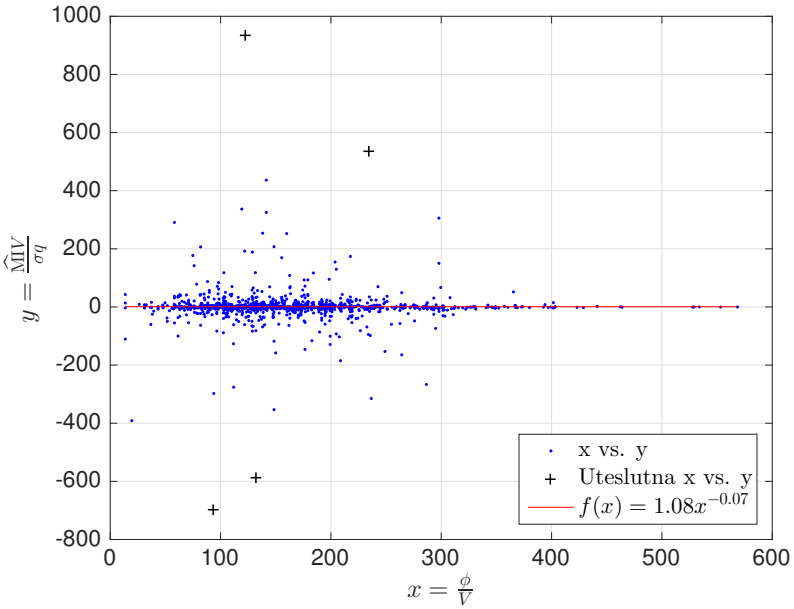
Figur 22: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region ME med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -0{,}24x^{0{,}15}$ som är den heldragna linjen i grafen.



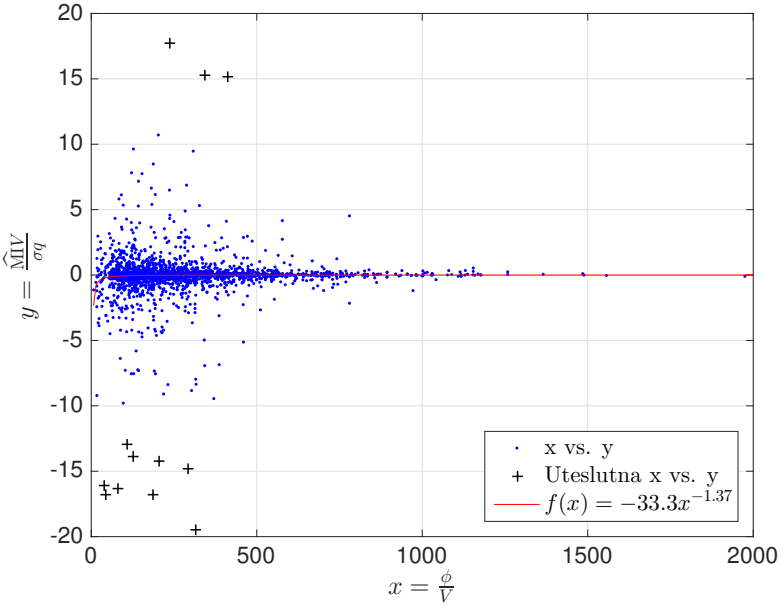
Figur 23: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MN med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 0{,}04x^{0{,}73}$ som är den heldragna linjen i grafen.



Figur 24: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MP med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 335{,}2x^{-1{,}23}$ som är den heldragna linjen i grafen.



Figur 25: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region MJ med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = 1{,}08x^{-0{,}07}$ som är den heldragna linjen i grafen.



Figur 26: I figuren plottas $y = \frac{\widehat{MIV}}{\sigma q}$ som funktion av $x = \frac{\phi}{V}$ för region EW med punkter. Datapunkter som ansågs vara extremvärden uteslöts därför och är markerade som + i grafen. Bästa potensapproximationen till datan gavs av $y = -33{,}3x^{-1{,}37}$ som är den heldragna linjen i grafen.