Estimating discretionary accruals using a grouping genetic algorithm

برآورد اقلام تعهدی اختیاری با استفاده از یک الگوریتم ژنتیک گروه بندی:

a b s t r a c t :

چکیده:

A number of different models have been suggested for detecting earnings management but the linear regression-based model presented by Jones (1991) is the most frequently used. The underlying assumption with the Jones model is that earnings are managed through accounting accruals. Typically, the companies for which earnings management is studied are grouped based on their industries. It is thus assumed that the accrual generating process for companies within a specific industry is similar. However ,some studies have recently shown that this assumption does not necessarily hold. An alternative approach which returns a grouping which is, if not optimal, at least very close to optimal is the use of genetic algorithms. The purpose of this study is to assess the performance of the cross-sectional Jones accrual model when the data set firms are grouped using a grouping genetic algorithm. The results provide strong evidence that the grouping genetic algorithm method outperforms the various alternative grouping methods.

تعداد مدل های متفاوتی برای تشخیص مدیریت عایدات پیشنهاد میشود، اما مدل رگرسیون خطی که به وسیله جونز ارائه شده مکررا استفاده میشود.فرض اساسی مدل جونز این است که عایدات از طریق حسابداری تعهدی اداره میشود.معمولا،شرکت هایی که مدیریت عایدات آنها مورد مطالعه قرار گرفته برمبنای صنعتشان گروه بندی شده اند.به این ترتیب فرض میشود که فرآیند تولید اقلام تعهدی برای شرکت های داخل یک صنعت خاص مشابه است.با این وجود برخی از مطالعات اخیرا نشان داده اند که چنین فرضی لزوما در اختیارنیست.رویکرد جایگزینی که یک گروه بندی را باز میگرداند این است که،اگر بهینه نبود،حداقل خیلی نزدیک به بهینه باشد که از روش الگوریتم ژنتیک استفاده شود.هدف این مطالعه ارزیابی عملکرد مقطعی مدل اقلام تعهدی جونز است، وقتی که مجموعه داده های شرکت هابا استفاده از الگوریتم ژنتیک گروه بندی، گروه بندی میشوند.نتایج، شواهد قوی را ارائه میدهد که روش گروه بندی الگوریتم ژنتیک عملکرد بهتری ازدیگر روش های گروه بندی جایگزین مختلف دارد.

1. Introduction

مقدمه:

The occurrence of earnings management has been a widely studiedsubject for the past 30 years. One of the major challenges whenexamining possible earnings management is that the magnitude ofit is difficult to assess. A number of different models have beensuggested for detecting earnings management but the linear regression-based model presented by Jones (1991) is the most frequentlyused. The underlying assumption with the Jones model is thatearnings are managed through accountingaccruals. Typically, thecompanies for which earnings management is studied are groupedbased on their industries. It is thus assumed that the accrual generatingprocess for companies within a specific industry is similar.Recently, however, some studies have shown that this assumptiondoes not necessarily hold. Dopuch, Mashruwala, Seethamraju, andZach (2012), for example, showed that a violation of the homogenousaccrual generating process within an industry causesmeasurement errors. In another study Ecker, Francis, Olsson, andSchipper (2011)showed that the performance of the Jones modelis improved when lagged totalassets are used as a grouping variableinstead of the industry membership. Eventhough alternativemethods have been used for grouping companies when usingtheJones-model, none of them has clearly outperformed thegroupingbased onindustry membership.

وقوع مدیریت عایدات موضوعی است که به طور گسترده ای برای ۳۰ سال گذشته مورد مطالعه قرار گرفته .یکی از چالش های عمده هنگام بررسی امکان مدیریت سود(عایدات)آن است که مقدار آن برای ارزیابی دشوار است.تعداد مدل های متفاوتی برای تشخیص(آشکار،نمایان ساختن) مدیریت عایدات پیشنهاد شده است، اما مدل رگرسیون خطی جونز به طور مکرر استفاده میشود.فرض اساسی مدل جونز این است که عایدات از طریق اقلام تعهدی حسابداری اداره میشود.معمولا،شرکت هایی که برای مدیریت سود مورد مطالعه قرار میگیرند بر مبنای صنعت شان گروه بندی میشوند.بنابراین چنین فرض میشود که روند تولید اقلام تعهدی برای شرکت های داخل یک صنعت خاص، مشابه است.با این وجود، اخیرا بعضی از مطالعات نشان میدهد که این فرض لزوما در اختیار نیست. برای مثال:

Dopuch, Mashruwala, Seethamraju, andZach (2012)

نشاندادند که نقص در فرآیند تولید اقلام تعهدی همگن در یک صنعت باعث خطاهای اندازگیری است.در دیگر مطلعات انجام شده توسط ایکر، فرانسیس،آلسون و اسچیپر نشان داده شد که کارایی مدل جونز زمانی بهبود

میابد که کل دارایی هایغیر جاریبه عنوان متغیر گروهبندی به جای عضو صنعتی استفاده میشود. اگر چه در زمان استفاده از مدل جونز شیوه های جایگزین دیگری برای گروه بندی شرکت ها استفاده شده است، هیچ یک از آن ها به وضوح، عملکرد بهتری از گروه بندی انجام شده بر مبنای عضویت صنعتی ،نداشته اند.

An exhaustive search for the best possible grouping is in mostcases impossible considering the large number of possible combinationseven with moderate size data sets. An alternative approachthat returns a grouping which is, if not optimal, at least very closeto optimal is the use of genetic algorithms. Genetic algorithmshave proven efficient in solving difficult problems such as the travellingsalesman and the equal piles problems.

پژوهش های وسیعی برای یافتن بهترین گروه بندی ممکن، در بسیاری از موارد با توجه به تعداد زیادی از ترکیبات احتمالی، حتی با متوسط اندازه مجموعه داده ها ،غیر ممکن است.رویکرد جایگزینی که گروه بندی را برمیگرداند این است که، اگر بهینه نبود حداقل خیلی نزدیک به بهینه در استفاده از الگوریتم ژنتیک باشد.الگوریتم های ژنتیک کارایی در حل مشکلات سخت مانند فروشنده سیار و مشکلات شمع برابر را به اثبات رساند.

Genetic algorithms have been used in a number of accountingapplications. Back,Laitinen, and Sere (1996) used a genetic algorithmto determine the optimal predictorsfor a neural networkbasedbankruptcy prediction model. A similar study was carriedout by Shin and Lee (2002) when they used a genetic algorithm togenerate bankruptcyprediction rules. Hoogs, Kiehl, Lacomb, andSenturk (2007) presented a geneticalgorithm approach for detectingfinancial statement fraud. Their model successfullyclassified63% of the companies that had been accused by the Securities andExchangeCommission (SEC) for improperly recognizing revenue.

روش الگوریتم ژنتیک در تعدادی از کاربرد های حسابداری استفاده میشود. Back,،Laitinenو Sereاز الگوریتم ژنتیک به منظور تعیین پیش بینی های بهینه برای شبکه عصبی بر مبنای مدل پیش بینی ورشکستگی استفاده شد. مطالعه مشابهی نیز به وسیله Shin and Lee (2002)هنگامی که آن ها از روش الگوریتم ژنتیک برای تولید(ایجاد) قوانین مربوط به ورشکستگی استفاده میکردند، انجام شد بود(Hoogs, Kiehl, Lacomb, andSenturk (2007 یکروش الگوریتم ژنتیک را برای کشف تقلب صورت های مالی ارائه کردند. مدل آن هابه طور موفقیت آمیزی ٦٣% از شرکت هایی را که به وسیله کمیسیون بورس اوراق بهادار به خاطر شناسایی درآمد غیر احتمالی متهم شده بودندطبقه بندی کرد.

The purpose of this study is to assess the performance of thecross-sectional Jones accrual model when the data set firms aregrouped using a grouping genetic algorithm. The performance ofthe grouping genetic algorithm approach is compared with theperformance of a number of other grouping techniques.

هدف این مطالعه ارزیابی کارایی یا عملکرد مدل مقطعی اقلام تعهدی جونز در مواقعی که مجموعه داده های شرکت با استفاده گروه بندی الگوریتم ژنتیک دسته بندی شده است، میباشد. کارایی رویکرد گروه بندی الگوریتم ژنتیک با کارایی تعداد دیگری از تکنیک های گروه بندی مقایسه شده است.

The remainder of this study is organized as follows. The basicoperating principle of the linear regression-based accrual modelsis covered in Section 2. In Section 3 an overview of both classicand grouping genetic algorithms is given. The research design ispresented in Section 4 and the results from the empirical study are presented in Section 5. Section 6 concludes the study.

باقی مانده این مطالعه یه صورت زیر سازمان دهی شده است.اصل عامل اساسی مدل های اقلام تعهدی برمبنای رگرسیون خطی در بخش دوم پوشش داده شده. در بخش ۳ مروری کلی بر هر دوی الگوریتم های ژنتیک گروهبندی و کلاسیک داده شده است.دربخش ۴ طرح تحقیق ، و نتایج ناشی از مطالعه تجربی در بخش ۵ ارائه شده است. بخش ۶ نتایج مطالعه است.

2. Discretionary accrual estimation models

مدل های برآورد اقلام تعهدی(اختیاری)

widest acceptance among researchers is the linear regression-based model presented by Jones (1991). The general assumption with the Jones-model is that earnings are managed through accounting accruals. The accruals are defined as the difference between net earnings before extraordinary items and cash flows from operations.The purpose of the Jones-model is to split the total accruals of a company into non-discretionary (expected) and discretionary (unexpected) accruals.

مقبول ترین روش در میان محققین مدل رگرسیون خطی ارائه شده به وسیله جونز است.فرض عمومی مدل جونز این است که عایدات از طریق اقلام تعهدی حسابداری مدیریت میشوند.اقلام تعهدی به عنوان تفاوت بین عایدات خالص قبل از اقلام غیرمترقبه و وجه نقد حاصل از عملیات، شناسایی میشوند.هدف مدل جونز جدا کردن(تجزیه کردن) کل اقلام تعهدی(تعهدات) یک شرکت به اقلام تعهدی غیر اختیاری و اقلام تعهدی اختیاری است.

The non-discretionary accruals are accruals that the company management has no or little control over ,whereas the discretionary accruals are a proxy for earnings management .In the model the reciprocal of total assets, change in revenues(DREV) and gross property, plant and equipment (PPE) are regressed on total accruals (TACC). The variable for change in revenues controls for current accruals such as receivables, payables and inventory whereas the variable for property, plant and equipment mainly controls for depreciation and amortization accruals. In the linear regression-based Jones-model the regression error term (e) equals the discretionary accruals

تعهدات غیر اختیاری تعهداتی هستند که مدیریت شرکت یا کنترل روی آن ندارد یا کنترل کمی روی آن دارد.درحالی که تعهدات اختیاری نماینده ای از مدیریت عایدات است.در مدل متقابل کل دارایی ها،تغییر در عایدات و اموال ناخالص،کارخانه و تجهیزات در کل تعهدات قرار گرفته میشوند. متغیر تغییر در کنترل های درآمد برای اقلام تعهدی جاری مواردی مانند حساب های دریافتنی، حسابهای پرداختنی و موجودی هستند، درحالی که متغیر اموال، کارخانه و تجهیزات عمدتا کنترل هایی

برای استهلاک دارایی های ثابت و استهلاک دارایی های نامشهود اقلام تعهدی است.
در مدل رگرسیون خطی جونز اصطلاح خطای رگرسیون برابر اقلام تعهدی اختیاری است

In the original Jones-model the regression was run separately foreach company using a time-series of at least 10 observations. Thetime-series approach has, however, several drawbacks. First, therequirement of at least 10 years of financial statement data mightlead to both a survivorship bias and a selection bias (Jeter &Shivakumar, 1999). Second, the assumption that the accrual generatingprocess of a company is stable over longer periods of timedoes not necessarily hold (Dopuch et al., 2012). Due to these drawbacksthe time-series approach has largely been replaced by thecross-sectional approach first suggested by Dechow (1994).

مدل اصلی رگرسیون جونز به طور جداگانه برای هر شرکت با استفاده از سری های زمانی حداقل ۱۰ داده از مشاهدات اجرا شده بود.با این وجود رویکرد سری زمانی دارای اشکالاتی است.اول،نیاز به حاقل ۱۰ سال از اطلاعات صورت های مالی دارد که ممکن است منجر به هم گرایش به بقا و هم گرایش به انتخاب شود.دوم،این فرض که روند تولید اقلام تعهدی یک شرکت طی مدت زمان زیادی پایدار است، لزوما در اختیار نیست.به علت اشتباهاتی که رویکرد سری زمانی داشت تا حد زیادی با رویکرد مقطعی که اولین بار توسط دکو(۱۹۹۴) پیشنهاد شده بود جایگزین شد.

the cross-sectional approach it is assumed that companies withina specific industry have a similar accrual generating process. Usually the industry is defined at a two-digit level SIC (e.g. Bartov,Gul, &Tsui, 2000; Jeter &Shivakumar, 1999). A number of recent studies have questioned whether the assumption of similar accounting generating processes within the industries is valid. Dopuch et al. (2012) showed that the assumption of a homogenous accruals generating process does not apply for several industries. Furthermore, they also showed that the violation of this assumption caused some measurement error of discretionary accruals. They do not, however, suggest any alternative method of grouping the companies when using the Jones-model. Ecker et al. (2011) used a number of different variables for grouping companies when using the Jones-model

and their findings showed that the best result was achieved when lagged total assets were used.

رویکرد مدل مقطعی فرض میکند که شرکت های داخل یک صنعت خاص روند تولید اقلام تعهدی مشابهی دارند.معمولا یک صنعت در سطح sic دو رقمی تعریف شده است.تعدادی از مطالعات اخیر این پرسش را طرح کردند که آیا فرض روندهای(فرایند های) تولید حسابداری مشابه داخل صنعت ها معتبر است.دوپاچ و همکاران،نشان دادند که فرض روند تولید اقلام تعهدی همگن برای صنایع مختلف صدق نمیکند(کاربرد ندارد).علاوه بر این ،آن ها نشان دادند که تخطی از این فرض باعث اندازه گیری اشتباهاقلام تعهدی اختیاری میشود.با این وجود آن ها مدل جایگزینی برای گروهبندی شرکت ها را هنگام استفاده از مدل جونز ارائه نکردند.ایکر و همکاران،

از تعدادی متغیر های مختلف برای گروه بندی شرکت ها هنگام به کارگیری مدل جونزاستفاده کردندویافته های آن ها نشان داد که بهترین نتیجه زمانی که کل دارایی های..... استفاده شده،بدست می آید.

3. Genetic algorithms

3.الگوریتم های ژنتیک

3.1. Operating principle

اصول عملیاتی

Genetic algorithms are an optimization technique based onmodels of natural selection and evolution. The fundamentalprinciples of the genetic algorithm were first presented by Holland(1975). The starting point when using genetic algorithms is apopulation consisting of a certain number of chromosomes(individuals), where the chromosomes represent valid solutionsto the problem. Once the population size has been determined, the initial population is usually randomly generated. The size ofthe initial population depends on the complexity of the problem.A genetic algorithm with a smaller population is faster but at thesame time the risk of premature convergence increases (Koljonen,Mannila, &Wanne, 2007).

الگوریتم های ژنتیک یک روش بهینه سازی مبتنی بر مدل های انتخاب طبیعی و تکامل هستند. اصول بنیادی الگوریتم ژنتیک اولین بار توسط هولاند ارائه شده بود.نقطه شروع هنگام استفاده از الگوریتم ژنتیک،جمعیتی متشکل از تعدادی معینی از کروموزم(افراد) است،که در آن کروموزم ها نشان دهنده راه حل های معتبر برای مشکلات هستند.هنگامی که اندازه جمعیت تخمین زده میشود،جمعیت

اولیه است که معمولا به طور تصادفی ایجاد شده.اندازه اولیه جمعیت بستگی به پیچیدگی مسئله دارد.یک الگوریتم ژنتیک با یک جمعیت کوچکتر سریع تر است اما در همان زمان مخاطره همگراییزود رس را افزایش میدهد.

Once the initial population has been generated,the fitness of each chromosome is evaluated. Based on thefitness values, parent chromosomes are selected from the initialpopulation to form new chromosomes to the next generationthrough a breeding process. The most commonly used method forselecting the parent chromosomes is the roulette wheel selection(Butun, Erfidan, &Urgun, 2006) in which a proportion of the wheelis assigned to each chromosome based on their fitness values. Thelarger the proportion of the wheel, the higher the probability ofgetting selected. A central part of the breeding process is thecross-over function.

هنگامی که جمعیت اولیه ایجاد شد،تناسب (ارزش تناسب، fitness نوعی تابع برای ارزیابی تناسب است)هر کروموزم ارزیابی میشود.براساس ارزش های تناسب،کروموزم های والدین از جمعیت اولیه برایتشکیل کروموزم های جدیدی برای نسل بعدی از طریق فرآیند پرورش (اصلاح نژاد)انتخاب میشوند. متداول ترین روش استفاده شده برای انتخاب کروموزم های والدین ،انتخاب چرخه منگنه دار رولت است.که در آن بخش ازچرخه، به هر کروموزم برمبنای ارزیابی های تناسب شان عددی اختصاص یافته.به موازاتبزرگترشدن بخشی از چرخه (افزایش متوسط عدد برازش جامعه)، سنگینی انتخاب هم بیشتر و جزئی‌تر می‌شود.بخش مرکزی فرآیند پرورش،cross-over(چسبیدن دو کروموزوم از طول به یکدیگر و تبادل برخی قطعات بین دو کروموزوم است). می باشد.

The parent chromosomes are combined usinga cross-over function to form new chromosomes. A commonly used cross-over technique is the single-point cross-over where a single cross-over point is selected in the chromosome. The part beyondthe cross-over point is then swapped between the two parentchromosomes. Once the next generation has been formed, thenew chromosomes are subjected to random mutation. The purposeof the mutation is to prevent the premature convergence of the genetic algorithm. After the mutation process has been completed,the fitness of the chromosomes in the new generation is assessedand the selection and cross-over procedures start all over. The evolution process is stopped when a satisfactory solution has beenreached or when some other predetermined condition has beenmet. To prevent from losing the best chromosomes during thecross-over and mutation operations, elitism can be employed.Elitism means that the best or a few of the best chromosomes aredirectly copied from the previous to the next generation. Elitismcan improve the performance of the genetic algorithm but thereis also a risk that it leads to premature convergence.

کروموزم های والدین با استفاده از عملcrossover برای تشکیل کروموزم های جدید ترکیب میشوند. روش crossover معمولا..... .سپس آن سوی نقطه crossover یا همگذری بین کروموزم های والدین معاوضه میشود.هنگامی که نسل بعدی شکل گرفت،کروموزم های جدید در معرض جهش تصادفی قرار میگیرند. هدف جهش جلوگیری از همگرایی زودرس الگوریتم ژنتیک است.بعد از اینکه فرآیند جهش کامل شد، میزان تناسب یا برازش کروموزم ها در نسل جدید مورد ارزیابی قرار میگیرد و شیوه های انتخاب و متقاطع در سراسر فرآیند شروع میشود.فرایند تکامل زمانی که راه حل رضایت بخشی بدست آید یا زمانی که بعضی دیگر از شرایط از پیش تعیین شده ای ملاقات شود،متوقف میشود. به منظور جلوگیری از، ازدست دادن بهترین کروموزم ها در خلال عملیات متقاطع و جهش،elitismمیتواند به کارگرفته شود.elitism به این معنی است که بهترین کروموزم ها یا کروموزم های بهتری به طور مستقیم از گذشته برای تولید نسل بعدی روبرداری شود(مناسب‌ترین عضو هر اجتماع انتخاب می‌شودElitist Selection) .elitism

میتواند عملکرد الگوریتم ژنتیک را بهبود بخشد، اما خطری که همچنان وجود دارد این است که elitism منجر به همگرایی زودرس(پیش از موعد، نارس) میشود.

3.2. Grouping genetic algorithms

گروه بندی الگوریتم های ژنتیک

Classic genetic algorithms do generally not perform well ongrouping problems. The main problems are a high redundancyamong the population chromosomes and a context insensitivityof the cross-over function (Falkenauer, 1996). To deal with theseshortcomings of the classic genetic algorithm, Falkenauer (1992)suggested a modified genetic algorithm suited for grouping problems.The general difference between the two types of geneticalgorithms is that in classic genetic algorithms the focus is on individualitems whereas in grouping genetic algorithms the focus is on groups of ithems.

الگوریتم های ژنتیک کلاسیک معمولا روی مشکلات گروه بندی به خوبی عمل نمی کنند. مشکلات اصلی شامل فراوانی یا افزونگی بالا میان کروموزم های جمعیت و بافت غیر حساس عملکرد crossoverمیباشد. به منظور مقابله با کمبود های ناشی از الگوریتم ژنتیک کلاسیکFalkenauerیک الگوریتم ژنتیک مناسب اصلاح شده را برای گروه بندی مشکلات پیشنهاد کرده. تفاوت کلی بین دو نوع الگوریتم های ژنتیک آن است که،الگوریتم های ژنتیک کلاسیک روی موارد یا آیتم های فردی تمرکز میکنند،در حالی که تمرکز در الگوریتم های ژنتیک گروه بندی بر روی گروه هایی از اقلام است .

In a grouping genetic algorithm the chromosome is divided intoa number of groups which in turn contain a number of individualitems. The values of the individual items as such are not important.Instead, it is the group membership of the items that is of value. Inthe grouping genetic algorithmsuggested by Falkenauer, the chromosomelength is variable. The groupinggenetic algorithm can,however, also be used with fixed lengthchromosomes when requiredby the nature of the problem (e.g. the equal piles problem,Falkenauer, 1996).

در یک الگوریتم ژنتیک گروهبندی کروموزم به تعدادی از گروه هایی که به نوبه خود شامل تعدادی از آیتم های فردی است تقسیم شده است. مقادیری از اقلام فردی به این ترتیب مهم نیستند.به جای آن،آن عضو گروهی از اقلام است که دارای ارزش میباشد در الگوریتم ژنتیک گروهبندی پیشنهاد شده به وسیله Falkenauer,طول کروموزم متغیر است.با این وجود الگوریتم ژنتیک گرو بندی نیز میتواند برای کرومزم های با طول ثابت در مواقعی که ماهیت مشکل مورد نیاز است، استفاده شود.(به عنوان مثال: مشکل توده های(شمع) برابر)

The cross-over function for the grouping genetic algorithm isillustrated inFig. 1. First, two cross-point are selected randomlyin both chromosomes involved in the cross-over function. To generatethe first new chromosome,the genes between the crosspointsin the second old chromosome areinjected after the firstcross-point in the first old chromosome. At this point some items occur in two groups. The groups coming from the first old chromosome and that have items also occurring in the groups coming from the second old chromosome are deleted. After removing these groups,some items might not be present in the remaining groups.These items are allocated to new groups using various methods.The missing items can, for example, form a new group or they can be distributed over a certain number of new groups. Once the first new chromosome has been created,a second new chromosome is formed using the same two oldchromosomes but in a reverse order.

شیوه crossover برای الگوریتم ژنتیک گروه بندی در تصویر شماره ۱ نشان داده شده است. در ابتدا، دو نقطه متقابل به طور تصادفی در هر دو کروموزم های درگیر در شیوه crossover انتخاب میشود.برای تولید اولین کروموزم جدید، ژن های بین نقاط متقابل در دومین کروموزم قدیمی بعد از اولین نقطه متقابل در اولین کروموزم قدیمی تزریق میشود.در این مرحله بعضی از آیتم ها در گروه ها رخ میدهد.گروه های از اولین کروموزم قدیمی و همچنین آن اقلامی که در گروه هایی از دومین کروموزم قدیمی رخ داده اند، حذف میشوند.پس از خارج کردن این گروه ها،بعضی از موارد ممکن است در گروه های باقی مانده حاضر نباشند. این اقلام به گروه های جدیدی با استفاده از روش های مختلف تخصیص داده شده است. اقلام از دست رفته،برای مثال، می تواند گروه جدیدی را شکل دهد یا اینکه میتواند در طول تعداد معینی از گروه های جدید توزیع شوند.هنگامی که اولین کروموزم جدید خلق

شد،یه کروموزم جدید دومی نیز با استفاده از همان دو کروموزم قدیمی اما در جهت معکوس تشکیل میشود.

As with the cross-over function, the mutation operator for grouping genetic algorithms works on groups rather than items.
مطابق با شیوه cross-over، اپراتور جهش برای الگوریتم های ژنتیک گروه بندی روی گروه ها به جای اقلام کار میکند.

Falkenauer (1996) describes three general strategies for mutation for the grouping genetic algorithm: creating new groups, eliminating existing groups or shuffling a small number of randomly selected items among the groups. The exact implementation details depend on what type of a problem is to be solved.

Falkenauer سه استراتژی عمومی یا کلی را برای جهش در الگوریتم های ژنتیک گروه بندی توصیف میکند: خلق گروه های جدید،حذف گروه های موجود یا در هم آمیختن یک تعداد اندکی از موارد انتخاب شده به طور تصادفی در میان گروه ها.
اجرای دقیق جزئیات به نوع مشکلی که باید حل شود بستگی دارد.

4. Research methodology

روش تحقیق

4.1. Research task

کار پژوهشی

The purpose of this study is to assess the performance of the cross-sectional Jones accrual model when the data set companiesare grouped using a grouping genetic algorithm. The result sobtained with the grouping genetic algorithm approach are compared with results obtained with various other grouping techniques,such as grouping the companies by their two-digit SIC.The performance of the cross-sectional Jones model with the different grouping techniques is measured both by the standard deviation of the discretionary accruals as well as by the ability to detect simulated earnings management of various levels.

هدف این مطالعه ارزیابی عملکرد مقطعی مدل تعهدی جونز هنگامی که مجموع داده ها ی شرکت ها با استفاده از الگوریتم ژنتیک گروه بندی، دسته بندی میشوند،است.نتایج بدست آمده با رویکرد گروه بندی الگوریتم ژنتیک با نتایج بدست آمده در روش های مختلف دیگر تکنیک های گروه بندی قابل مقایسه

هستند. مانند گره بندی شرکت ها توسط sic دو رقمی آن ها.عملکرد مدل مقطعی جونزباروش های گروه بندی مختلف،که هر دو باانحراف معیارازاقلام تعهدی اختیاری و نیزتوانایی تشخیص مدیریت سودشبیه سازی شده ازسطوح مختلف،اندازه گیری شد.

4.2. Data set description

توصیف مجموعه داده ها

The data set comprises financial statement data for public UScompanies with complete data for years 2009 and 2010 retrievedfrom Burea van DijksOrbis database. In accordance with previousstudies, utilities (SIC 4900-4999) and financial companies (SIC6000-6999) are excluded from the dataset. Furthermore, companieswith lagged total assets less than 1 million USD, zero revenues for atleast one of the two years, absolute total accruals equal to or abovelagged total assets or market capitalization equal to zero are removed.

مجموعه داده ها ،شاملداده هایصورت های مالی برای شرکت های دولتی ایالات متحده، همراه با اطلاعات کامل برای سال های ۲۰۰۹ و ۲۰۱۰ بازیابی شده از بانک اطلاعاتی دفترvan DijksOrbisمیباشد.بر اساس مطالعات قبلی، شرکت های آب و برق (SIC 4900-4999) وشرکت های مالیsic(6000-6999)از مجموعه داده ها حذف شدند. بنابراین شرکت هایی با کل دارایی های غیر جاری کمتر از ۱میلیون دولار ، نداشتن درآمد برای حداقل یکی از ۲ سال،کل اقلام تعهدی مطلق برابربا، یا بیش تر از کل دارایی های غیر جاری یا ارزش بازار برابر صفر، رد میشوند

To avoid the impact of possible outliers, the independentvariables in the Jones-model (change in sales and property, plant& equipment) are truncated at the 1st and 99th percentiles. Finally,all companies belonging to two-digit SIC groups with less than 20companies are excluded. The criteria above results in a data set comprising2590 companies in 30 two-

digit SIC groups. The final data setcontains 10 groups with 20 companies with the same two-digit SICin each group (see Table 1). These 10 groups are selected randomlyout of the available 30 two-digit SIC groups. If a selected group hasmore than 20 companies, 20 companies are selected randomly fromall companies with the same two-digit SIC. The following 10 two-digitSIC groups have been randomly selected.

برای جلو گیری از تاثیر احتمالی اقلام دارای نوسان، متغیر های مستقل در مدل جونز(تغییر در فروش ودارایی، کارخانه و تجهیزات) در ۱ و ۹۹..... کوتاه میشوند. سرانجام همه شرکت های متعلق به گروه های با کد دورقمی sic با کم تر از ۲۰ شرکت در حال مطالعه حذف میشوند.معیار های نتایج فوقدر یکمجموعه داده هاشامل2590شرکتدر گروه های با ۳۰ کد دورقمی می باشد.مجموعه داده های نهایی شامل ۱۰ گروه با ۲۰ شرکت با همان کد دورقمی sic در هر گروه میباشد. این ۱۰ گروه به طور تصادفی از خارج از گروه هایی با ۳۰ کد دورقمی sic موجود انتخاب شدند.اگر یک گروه انتخاب شده بیش از ۲۰ شرکت را در برگیرد، ۲۰ شرکت به طور تصادفی از همه شرکت ها، با همان کد دورقمی sic انتخاب میشوند. در زیر گروه هایی با ۱۰ کد دو رقمی به طور تصادفی انتخاب شده است.

4.3. Grouping genetic algorithm discretionary accrual model

٤,٣ گروه بندی الگوریتم ژنتیک مدل اقلام تعهدی اختیاری

The discretionary accrual model used in this study is a linearregression-based cross-sectional version of the Jones-model (e.g.Dechow, 1994). In the model the reciprocal of total assets, changein revenues (DREV) and gross property, plant and equipment (PPE)are regressed on total accruals (TACC). Total accruals are calculatedby subtracting operating cash flows from net earnings beforeextraordinary items. All variables in the regression equation aredeflated by lagged total assets (TAt_1).

مدل اقلام تعهدی اختیاری استفاده شده در این مطالعه یک رگرسیون خطی مبتنی بر نسخه ای از مدل جونز است.در مدل متقابل کل دارایی ها،تغییر در عایدات و اموال ناخالص،کارخانه و تجهیزات در کل اقلام تعهدی نشان داده شده .مجموع اقلام تعهدی به وسیله کم کردن جریان وجوه نقد عملیاتی از سود خالص قبل از اقلام غیر مترقبهمحاسبه میشود.همه متغییر های در معادله توسط کل دارایی های غیر جاریی

The grouping genetic algorithm is used for dividing the data setcompanies into 10 groups with 20 companies in each. For each ofthe 10 groups the Jones-model regression model is run separately.The residuals from the regressions (e) equal the discretionaryaccruals and the fitness of a specific grouping is assessed basedon the standard deviation of the discretionary accruals for all 10groups combined. The lower the standard deviation of the discretionaryaccruals, the better the grouping of the companies.

گروه بندی الگوریتم ژنتیک برای تقسیم مجموعه داده های شرکت ها به ۱۰ گروه با ۲۰ شرکت در هر کدام، استفاده شده است.برای هر کدام از ۱۰ گروه مدل رگرسون جونز به طور جداگانه به اجرا درآمد.مانده های از رگرسیون ها، برابر اقلام تعهدی اختیاری است ،و برازش یک گروه بندی خاص برمبنای انحراف معیار اقلام تعهدی اختیاری برای همه ۱۰ گروه ترکیب شده، ارزیابی شده است. کم ترین انحراف معیار اقلام تعهدی اختیاری،بهترین گروه بندی از شرکت هاست.

The parameters for the genetic algorithm are presented inTable 2. Thereis no exact method for determining the parametersand thus someheuristics are required. Both the number of groupsand the group size(number of items per group) are fixed and donot change during theevolution process. The population comprises100 chromosomes which areinitially generated at random.The maximum number of generations is setto 25,000 but theevolution process is stopped if the best fitness valuereaches zero.In order to prevent the loss of the chromosomes with thebest fitness,elitism is employed. In this study, the chromosomes withthe10 best fitness values are copied directly to the next generation.To avoidgetting stuck in local minima and to prevent earlyconvergence, a mutationmechanism is used. The probability thata chromosome is mutated is set to5%. The chromosomes selectedby elitism are excluded from possiblemutation.

پارامتر ها برای الگوریتم ژنتیک در جدول ۲ ارائه شده است.شیوه دقیقی برای تعیین پارامتر ها وجود ندارد،و به این ترتیب برخی از روش های ابتکاری مورد نیاز است. هم تعداد گروه ها و هم اندازه گروه، هر دو(تعداد اقلا بر حسب گروه) ثابت هستند، و درخلال فرآیند تکامل تغییر نمیابند.جمعیت شامل ۱۰۰ کروموزم که در ابتدا به طور تصادفی ایجاد شدند، است. بیشترین تعداد تولیدات مجموعه ای بالغ بر ۲۵۰۰۰ میباشد. اما فرایند تکامل زمانی که بهترین ارزش تناسب به صفر برسد، متوقف میشود. به منظور ممانعت از ، از دست دادن کروموزم ها با بهترین تناسب، elitism به کار گرفته میشود. در این مطالعه،کروموزم ها با ۱۰ مورد از بهترین ارزش های تناسب به طور مستقیم برای نسل بعدی کپی میشوند. برای پرهیز از گرفتار

شدن در حداقل محل،و جلوگیری از همگرایی زود هنگام ،مکانیسم جهش استفاده میشود.این احتمال که یک کروموزم جهش یابد ۵% است.کروموزم های انتخاب شده به وسیله elitism از امکان جهشحذف میشوند.

When generating a new generation with the genetic algorithm,the chromosomes with the 10 best fitness values are first copiedfrom the old generation. The remaining chromosomes are generatedfrom chromosomes in the old generation using the roulettewheel selection and a cross-over mechanism so that two chromosomesfrom the old generation results in two chromosomes in thenew population. The cross-points in both chromosomes areselected randomly.Once the new generation is completed،the chromosomes are subjected to possible mutation. If a chromosomeis mutated, two genes are randomly selected from the chromosomeand the items in these genes are randomly shuffled

هنگام تولید یک نسل جدید با استفاده از الگوریتم ژنتیک، کروموزم هایی با ۱۰ مورد از بهترین ارزش های تناسببرای اولین بار از نسل قدیمی کپی میشوند.کروموزم های باقیمانده ،از کروموزم های در نسل قدیمی با استفاده ازانتخاب چرخه رولت و سازو کارcrossover ایجاد شده، به طوری که دو کروموزم از نتایج نسل های قدیمی در دو کروموزم در جمعیت جدید .نقاط تقاطع در هر دو کروموزم به طور تصادفی انتخاب میشوند.هنگامی که نسل جدید کامل شد.کروموزم ها در معرض جهش احتمالی قرار میگیرند.اگر یک کروموزم جهش یافته باشد، دو ژن از کروموزم به طور تصادفی اتنخاب میشوند، و آیتم های موجود در این ژن ها به طور تصادفی در هم می آمیزند.

4.4. Strategy of analysis

استراتژی تجزیه و تحلیل

The performance of the Jones-model when grouping the companiesusing a grouping genetic algorithm is compared with theperformance of a number of alternative grouping methods. First,the performance is assessed when the companies are randomlygrouped and when one regression is run for all 200 companiespooled. Second, the companies are grouped based on their two-digitSIC. This is a common method of grouping, used frequentlyin previous studies. Last, six other grouping variables suggestedby Ecker et al. (2011) are used. These six variables are currentand lagged total assets, current and lagged revenues, market capitalizationand company age in years. When forming the groupsusing these six variables, the companies are first sorted accordingto the value of a variable. The first group then consists of companies1–20, the second group of companies 21–40 and so forth.

عملکرد مدل جونز هنگام گروه بندی شرکت ها با استفاده از گروه بندی الگوریتم ژنتیک با عملکرد تعدادی از روش هایگروه بندیجایگزین مقایسه میشود.در ابتدا، عملکرد، هنگامی که شرکت ها به طور تصادفی گروهبندی میشوند و زمانی که یک رگرسیون برای جمع همه ۲۰۰ شرکت به اجرا در میآید،ارزیابی میشود.دوم،شرکت هایی که براساس کد دو رقمی sic شان گروه بندی میشوند.این روشی معمول در گروه بندی است که به طور مکرر در مطالعات قبلی استفاده شده.در آخر،۶ گروه بندی دیگر از متغیر ها که به وسیله ایکر پیشنهاد شده مورد استفاده قرار گرفت.این ۶ متغیر کل دارایی های جاری و غیر جاری ،درآمد های جاری و دارای تاخیر زمانی ،بازار سرمایه و سن شرکت ها در سال میباشد. هنگام تشکیل گروه ها با این ۶ متغیر،شرکت هابرای اولین باربا توجه بهارزش یک متغیرطبقه بندی میشوند. بنابراین گروه اول شامل شرکت های ۱ تا ۲۰ ، گروه دوم شرکت های ۲۱ تا ۴۰ و غیره

The performance of the different grouping methods is first assessed by comparing the standard deviation of the discretionaryaccruals. A lower standard deviation of the discretionary accruals equals a better performance. Furthermore, the mean and median values of the discretionary accruals are examined. The closer to zero these values are, the better the performance of a method. To assess the earnings

management detection ability of the different grouping methods, various levels of simulated earnings managementare inserted to the data set. Two types of earnings management are simulated: expense manipulation and revenue/baddebt-manipulation. In expense manipulation a certain level of earnings management is added to total accruals, whereas in revenue/bad-debt manipulation earnings management is added both to the total accruals and to the change in revenues. The levels of the simulated earnings management ranges between _5% and 5% of lagged total assets with 1% increments.

عملکرد شیوه های مختلف گروه بندی، برای اولین بار با مقایسه انحراف معیار اقلام تعهدی اختیاری ارزیابی شده است.انحراف معیار پایین تر اقلام تعهدی اختیاری، برابر عملکرد بهتر است.علاوه بر این، مقادیر میانه و میانگین اقلام تعهدی اختیاری مورد بررسی قرار گرفت. به صفر نزدیکتر بودن این مقادیر نشانه عملکرد بهتر یک روش است.برای ارزیابی توانایی تشخیص مدیریت عایدات شیوه های مختلف گروه بندی ،سطوح مختلف شبیه سازی شده مدیریت سود به مجموعه داده ها وارد شده. دو نوع از مدیریت عایدات شبیه سازی شده اند، دستکاری هزینه و دستکاری درآمد یامطالبات غیر قابل وصول . در دستکاری هزینه یک سطح معینی از مدیریت عایدات به کل اقلام تعهدی اضافه شد، در حالی که در دستکاری درآمد یا مطالبات غیر قابل وصول ،مدیریت سود، هم به کل اقلام تعهدی و هم به تغییر در درآمد ها، اضافه شده است.سطوح شبیه سازی شده مدیریت عایدات دامنه ای بین ۵% و ۵% کل دارایی های غیر جاریی با ۱% افزایش میباشد.
11111111111111

To test whether a grouping method is able to detect a certain level of earnings management,simulated earnings management is added to 20 companies which are randomly selected from the total data set of 200 companies.The discretionary accruals for the companies with the simulated earnings management are calculated using the regression coefficients for the group to which they belong. Furthermore, for the20 companies that have simulated earnings management a partitioning variable (PART) is set to 1. For the remaining 180 companies the partitioning variable is set to 0. Next, the partitioning variable is regressed on the discretionary accruals (DACC).The ability of the Jones-model to detect certain level of simulated earnings management when using a specific grouping method is determined based on the regression coefficient of the partitioning variable (b1). If the b1 coefficient is significant at a 5% level, the simulated earnings management is considered detected. The earnings management

simulation process is repeated 500 times. The performance of a specific grouping method is assessed based on the number of significant coefficients for the partitioning variable at different levels of simulated earnings management.

برای آزمون اینکه آیا یک روش گروه بندی قادر هست تا یک سطح معینی از مدیریت عایدات را کشف کند یا نه، مدیریت عایدات شبیه سازی شده، به ۲۰ شرکتی که به طور تصادفی از مجموعه کل اطلاعات ۲۰۰ شرکت انتخاب شده بود،اضافه شد. اقلام تعهدی اختیاری برای شرکت هایی با مدیریت سود شبیه سازی شده با استفاده از ضرایب رگرسیون، برای گروهی که به آن تعلق داشتند محاسبه شد. علاوه براین، برای ۲۰ شرکتی که مدیریت عایدات شبیه سازی شده داشتندمتغیر پارتیشن بندی برابر ۱ است. برای ۱۸۰ شرکت باقی مانده متغیر گروه بندی برابر ۰ میباشد،سپس متغیر پارتیشن بندی روی اقلام تعهدی اختیاری ..... . توانایی مدل جونز برای تعیین سطح معینی از مدیریت عایدات شبیه سازی شده،درهنگام استفاده ازیک روش گروه بندی خاص،بر اساس ضریب رگرسیون متغیرپارتیشن بندی(B1)،تعیین شده است. اگر ضریب b1 در سطح معنی داری ۵% باشد،مدیریت عایدات شبیه سازی شده، تشخیص داده شده در نظر گرفته میشود.فرآیند شبیه سازی مدیریت عایدات ۵۰۰ مرتبه تکرار شده است.عملکرد یک روش گروه بندی خاص، براساس تعداد ضرایب قابل توجه برای متغیر پارتیشن بندی، در سطوح مختلف مدیریت عایدات شبیه سازی شده، ارزیابی شده است.

5. Results and discussion

بحث و نتایج

5.1. Descriptive statistics

آمار توصیفی

The total data set consists of 200 public US companies from 10different industries. The descriptive statistics presented in Table 3show that there are considerable differences between the data set companies.

مجموع کل داده ها شامل ۲۰۰ شرکت دولتی از ۱۰ صنعت مختلف هستند.آمار های توصیفی ارائه شده در جدول ۳ نشان میدهد که تفاوت قابل توجهی بین مجموعه داده های شرکت ها وجود دارد.

The largest company has total assets of more than 39billion USD, whereas the smallest company shows total assets ofjust above 1 million USD.

Similarly, the revenues range between more than 25 billion USD and about 15000 USD. The average total accruals are _0.069. This is expected as accruals typically are negative due to depreciation. The age of the companies ranges between1 and 155 years, with an average age of 35 years.

بزرگترین شرکت بیشتر از ۳۹ میلیارد دولار دارایی داشت، در حالی که کوچکترین شرکت تنها بالای ۱ میلیون دولار دارایی را نشان میداد، به طور مشابه، محدوده درآمد ها بین بیش از ۲۵ میلیارد دولار و حدود ۱۵۰۰۰ دولار بود.میانگین کل اقلام تعهدی۰,۰۶۹ هستند.این انتظار میرود که اقلام تعهدی، معمولا به علت استهلاک، منفی هستند. سن شرکت ها محدوده ای بین ۱ تا ۱۵۵ سال، با یک میانگین سنی ۳۵ سال در برداشت.

The evolution process is illustrated in Fig. 2. The dotted line shows the average standard deviation of the discretionary accruals and the solid line shows the lowest standard deviation of the discretionary accruals for the population in each generation. At around the 6000nd generation the population starts to converge and from the 10,000nd generation and onwards only minor improvements to the lowest standard deviation of the discretionaryaccruals occur. When the evolution process is stopped at the25,000nd generation the lowest value of the standard deviation of the discretionary accruals is 0.060. The evolution process was also run using various parameter values for population size, elitism and mutation probability. The evolution processes with the alternative parameters did, however, not improve the results.

فرایند تکامل در تصویر شماره ۳ نشان داده شده است. خط نقطه چین نشان دهنده متوسط انحراف معیار اقلام تعهدی اختیاری و خط ممتد نشان دهنده کمترین انحراف معیار اقلام تعهدی اختیاری برای جمعیت در هر نسل می باشد. در ..... ٦٠٠٠ ..... و بعد تنها بهبود جزئی تا پایین ترین اقلام تعهدی اختیاری رخ میدهد.زمانی که فرایند تکامل در ۲۵۰۰۰ نسل متوقف شد، کمترین مقدار انحراف معیار اقلام تعهدی اختیاری ۰,۰٦۰ است.فرآیند تکامل همچنین با استفاده از مقادیر پارامترهای مختلف برای اندازه جمعیت ، مانندelitism وجهش احتمالی اجرا شد. پروسه های تکامل با پارامتر های جایگزین انجام شد، بااین وجود ،نتایج را بهبود نبخشید.

5.2. Assessing the performance of the model

ارزیابی عملکرد مدل

In this study the performance of the cross-sectional Jonesmodelis assessed with the companies grouped using a groupinggenetic algorithm. The performance of the genetic algorithmgrouping method is compared with the performance of nine othergrouping methods. The first test of performance is based on thestandard deviation of the discretionary accruals. The results in Table4 clearly show that the genetic algorithm grouping method hasthe lowest standard deviation at 0.060. The second best groupingmethod is the two-digit SIC grouping with a standard deviationof 0.115 whereas the worst performing grouping method is thepooled regression with a standard deviation of 0.146. The remainingseven grouping methods show similar standard deviationsranging between 0.130 and 0.13 .

در این مطالعه عملکرد مدل مقطعی جونز، با شرکت های گروه بندی شده با استفاده از الگوریتم ژنتیک، ارزیابی شد. عملکرد روش گروه بندی الگوریتم ژنتیک با عملکرد ۹ شیوه گروه بندی دیگر مقایسه شد. اولین آزمون عملکرد مبتنی است بر انحراف معیار اقلام تعهدی اختیاری.نتایج در جدول ٤ به وضوح نشان میدهد که روش گروه بندی الگوریتم ژنتیک دارای کمترین انحراف معیار برابر ۰,۰٦۰ است. دومین و بهترین شیوه گروه بندی ، گروه بندی بااستفاده از کد دو رقمی sic با یک انحراف معیار ۰,۱۱۵ است، در حالی که بدترین شیوه اجرای گروه بندی رگرسیون..... با یک انحراف معیار ۰,۱۴۶ میباشد. هفت شیوه باقی مانده گروه بندی دیگرنشان دهنده انحراف معیار مشابه در محدوده ای بین ۰,۱۳۰ و ۰,۱۳ هستند.

None of the 10 grouping methods show discretionary accrualssignificantly different from zero. The grouping method with themean discretionary accruals closest to zero is the genetic algorithmgrouping method, whereas the grouping method based on the marketcapitalization shows the median discretionary accruals closestto zero. Finally, the genetic algorithm grouping method clearlyshows the smallest values for both minimum and maximum discretionaryaccruals. Overall, these results show that the geneticalgorithm grouping method outperforms the other nine groupingmethods.

هیچ کدام از ۱۰ شیوه گروه بندی ، به طور قابل ملاحضه ای اقلام تعهدی اختیاری را متفاوت از صفر نشان نمیدهد.شیوه گروه بندی اقلام تعهدی اختیاری با میانگین نزدیک به صفر روش گروه بندی الگوریتم ژنتیک است، در حالی که روش گروه بندی مبتنی بر بازار سرمایه نشان دهنده اقلام تعهدی اختیاری با میانگین نزدیک به صفر است.سرانجام، روش گروه بندی الگوریتم ژنتیک به وضوح کمترین مقدار برای هر دو حداقل و حداکثر اقلام تعهدی اختیاری را نشان میدهد.به طور کلی، این نتایج نشان میدهد که شیوه گروه بندی الگوریتم ژنتیک، بهتر از ۹ روش دیگر گروه بندی عمل میکند.

In the second test of performance, known levels of earningsmanagement is added to 20 companies which are randomly selectedfrom the total data set of 200 companies. The levels of thesimulated earnings management ranges between _5% and 5% oflagged total assets with 1% increments. For each level of simulatedearnings management, 500 data sets with 20 randomly selectedcompanies are created and evaluated. The higher the number ofdata sets where earnings management is detected, the better

theperformance. The results from the simulated expense manipulationare presented in Table 5.

در دومین آزمون عملکرد،سطوح شناخته شده مدیریت سود، به ۲۰ شرکتی که به طور تصادفی از مجموع ۲۰۰ شرکت انتخاب شدند، اضافه شد است.سطوح شبیه سازی شده مدیریت عایدات، محدوده ای بین ۵% و ۵% کل دارایی های..... با ۱% افزایش میباشد. برای هر سطح از مدیریت عایدات شبیه سازی شده،۵۰۰ مجموعه داده ها با ۲۰ شرکتی که به طور تصادفی انتخاب شدند، ایجاد و مورد ارزیابی قرار گرفت. بالاترین مجموعه داده ها که در آن مدیریت عایدات شناسایی شده، عملکرد بهتر است.نتایج ناشی از دستکاری هزینه شبیه سازی شده در جدول ۵ ارائه شده است.

With no simulated earnings management,a well specified model is not expected to detect any earnings management in a randomly selected data set. When no simulatedearnings management has been added to the data set, the detectionrate varies between 4.4% and 7.8%. As expected, the higherthe level of the simulated earnings management, the higher thedetection rate is for all grouping methods. At a level of ±1% the geneticalgorithm grouping method shows the highest detectionrates but the difference to the other grouping methods is not distinct.However, already at a level of ±2% the difference between the genetic algorithm grouping method and the other groupingmethods becomes much clearer.

بدون مدیریت سود شبیه سازی شده، انتظار نمیرود که یک مدل مطلوبی، هر مدیریت عایدات در مجموعه داده به طور اتنخاب تصادفی را آشکار کند. هنگامی که مدیریت عایدات شبیه سازی نشده به مجموع داده ها اضافه شود،نرخ کشف متغیر ها بین ۴,۴% و ۷,۸% خواهد بود، همان طور که انتظار میرود، بالاترین سطح شبیه سازی شده مدیریت عایدات، بالاترین نرخ کشف برای همه روش های گروه بندی است. سطح1٪ ±از روش گروه بندی الگوریتم ژنتیکبالاترین نرخ تخیص را نشان میدهد، اما این اختلاف برای دیگر روش های گروه بندی مشخص نشده.با این حال،در حال حاضر
در سطح ±2% اختلاف بین شیوه گروه بندی الگوریتم ژنتیک و دیگر شیوه های گروه بندی بسیار واضح تر میشود

With simulated earnings managementequaling ±2% of lagged total assets, the detection rate for thegenetic algorithm grouping method is 29.0% whereas the groupingmethod with the second best detection rate is much

lower at12.0%. At the highest level of simulated earnings management,the detection rate for the genetic algorithm grouping method isover 90%. At this level of simulated earnings management thedetection rate for the other grouping methods ranges between29.4% and 44.0%.

با مدیریت سود شبیه سازی شده برابر ±2% کل دارایی های غیر جاریی،نرخ کشف برای روش گروه بندی الگوریتم ژنتیک ۲۹% است.در حالی که روش گروه بندی با دومین و بهترین نرخ کشف بسیار پایین تر از ۱۲% است.در بالاترین سطح مدیریت عایدات شبیه سازی شده، نرخ کشف برای روش گروه بندی الگوریتم ژنتیک بلای ۹۰% است. در این سطح از مدیریت عایدات شبیه سازی شده ،نرخ کشف برای دیگر شیوه های گروه بندی محدوده ای بین ۲۹,٤% و ٤٤,۰۰% است.

To assess the overall detection rate for the grouping methods,an effectiveness score is calculated by first dividing the detectionrate for the grouping method at a specific level of earnings managementwith the maximum detection rate at that level of earningsmanagement. This results in 10 effectiveness scores for eachgrouping method. The overall effectiveness score is then calculatedby averaging these 10 effectiveness scores. As the genetic algorithmgrouping method has the highest detection rate at all levelsof simulated earnings management, the effectiveness score for thismethod equals 100%. The method with the second highest effectivenessscore is the two-digit SIC grouping method whereas therandom grouping method shows the lowest effectiveness score.

با هدف ارزیابی نرخ بدست آمده کلی برای شیوه های گروه بندی،یک رقم تاثیر گذار به وسیله نخستین نرخ یافت شده تقسیم بندی برای شیوه گروهبندی در سطحی خاص از مدیریت منافع، با بیشترین نرخ یافت شده در آن سطح از مدیریت عایدات، محاسبه شده است. این محاسبه در ۱۰ رقم تاثیر گذار نتیجه میدهد. کل امتیاز موثر سپس از طریق میانگین این ۱۰ امتیاز موثر محاسبه میشود. چون که شیوه گروه بندی الگوریتم ژنتیک دارای بالاترین نرخ تشخیص، در همه سطوح مدیریت عایدات شبیه سازی شده است،امتیاز موثر برای این شیوه برابر ۱۰۰% میباشد. شیوه دوم با بالاترین امتیاز موثر روش گروهبندی با کد sic دو رقمی است، در حالی که شیوه گروه بندی تصادفی پایین ترین امتیاز موثر را نشان میدهد.

The detection rates for the revenue/bad-debt manipulation arepresented in Table 6. In general, the detection rates for the simulatedrevenue/bad-debt manipulation are similar to the detectionrates for the simulated expense manipulation. It is mainly atthe highest level of the simulated earnings management that thedetection rates are somewhat higher for the simulated expensemanipulation. As with the expense manipulation, the genetic algorithmgrouping method clearly shows the highest detection ratesfor revenue/bad-debt manipulation. The second best groupingmethod is again the two-digit SIC grouping and the lowestperforming grouping method is the random grouping.

نرخ های تشخیص برای دستکاری درآمد/بدهی وصول نشده در جدول ٦ ارائه شده است.به طور کلی،نرخ های تشخیص برای درآمد/ بدهی وصول نشده مشابه با نرخ های تشخیص، برای دستکاری هزینه شبیه سازی شده میباشد.عمدتا در بالاترین سطح از مدیریت منافع شبیه سازی شده ، نرخ های تشخیص، تا حدودی برای دستکاری هزینه های شبیه سازی شده بالاتر هستند.چون با دستکاری هزینه، شیوه گروه بندی الگوریتم ژنتیک به وضوح بالاترین نرخ تشخیص را برای دستکاری درآمد/بدهی وصول نشده نشان میدهد.بهترین و دومین شیوه گروه بندی ، گروه بندی دوباره کد دو رقمی sicو پایین ترین عملکرد شیوه گروه بندی، گروه بندی تصادفی است.

5.3. Discussion

بحث

The results in this study provide strong evidence that thecross-sectionalJones-model with the companies grouped using agrouping geneticalgorithm outperforms the cross-sectional Jonesmodelwith the companies grouped with alternative methods. First,the standard deviation of the discretionary accruals using the geneticalgorithm grouping method is less than half of the standarddeviation obtained with the other grouping methods. Second, thegenetic algorithm method shows significantly higher earningsmanagement detection rates on various levels of simulatedexpenseand revenue/bad-debt manipulation compared with the othergrouping methods. At a 5% level of simulated earnings managementthegenetic algorithm method has a detection rate above 90%,whereas therest of the methods have detection rates around 30%.

نتایج این مطالعه شواهد قوی را فراهم کرد که مدل مقطعی جونز با شرکت های گروه بندی شده با استفاده از الگوریتم ژنتیک عملکرد بهتری از مدل مقطعی جونز با شرکت های گروه بندی شده با روش های جایگزین دارند.اولا، انحراف معیار اقلام تعهدی اختیاری با استفاده از شیوه گروه بندی الگوریتم ژنتیک، کمتر از نیمی از انحراف معیار بدست آمده با دیگر شیوه های گروه بندی است.ثانیا،شیوه الگوریتم ژنتیک به طور قابل ملاحضه ای بالاترین نرخ تشخیص مدیریت عایدات را روی سطوح مختلف هزینه شبیه سازی شده و دستکاری درآمد/..... در مقایسه با دیگر شیوه های گروه بندی نشان میدهد.در سطح ۵% از مدیریت عایدات شبیه سازی شده، شیوه الگوریتم ژنتیک نرخ تشخیص بالای ۹۰% داشت، در حالی که باقی مانده روش ها نرخ تشخیص در حدود ۳۰% داشتند.

6. Conclusion

نتایج

The purpose of this study was to assess the performance of the cross-sectional Jones accrual model when the data set firms are grouped using a grouping genetic algorithm. The results providestrong evidence that the grouping genetic algorithm method outperformsthe various alternativegrouping methods. This is especiallyclear when evaluating the earningsmanagement detectionpower at different levels of simulated earningsmanagement. Theresults also show that alternative grouping methods, such as thefrequently used two-digit SIC grouping, perform only marginallybetter compared with a random grouping approach. This studycould be extended by including various modifications of theJones-model as some studies have shown that there is a differencein performancebetween different versions of the Jones-model. Furthermore,a wider range of settings for the genetic algorithm could be assessed. Especially, lifting the constraint of a fixed number of groups could improve the performance of the genetic groupingalgorithm approach.

هدف این مطالعه ارزیابی عملکرد مدل مقطعی اقلام تعهدی جونز هنگامی که مجموعه داده های شرکت ها با استفاده ازالگوریتم ژنتیک گروه بندی، گروه بندی شده اند میباشد.نتایج شواهد قویرا فراهم میکند که شیوه گروه بندی الگوریتم ژنتیک عملکرد بهتری از دیگر شیوه های گروه بندی دارد.این امر به ویژه در زمان ارزیابی قدرت تشخیص مدیریت عایدات، درسطوح مختلفشبیه سازی شده مدیریت عایدات، روشن است.نتایج همچنین نشان میدهد که شیوه های گروه بندی جایگزین،مانند گروه بندی کد دو رقمیsic که به طور مکرر استفاده شده، تنها انکی بهتر در مقایسه با رویکرد گروه بندی تصادفی عمل کرد. این مطالعه میتواند به وسیله تغییرات مختلف مدل جونز، همانطور که بعضی از مطالعات نشان دادند که تفاوت در عملکرد بین شیوه های مختلف مدل جونز وجود دارد، میتواند توسعه یابدعلاوه بر این، محدوده گسترده از تنضیمات برای الگوریتم ژنتیک میتواند ارزیابی شود.به ویژه،..... میتواند به وسیله رویکرد الگوریتم گروه بندی ژنتیکی بهبود یابد.