

UKLJUČIVANJE SEZONSKIH UTJECAJA U BODOVANJE KREDITA BASEL II

dr. sc. Goran KLEPAC
Raiffeisen bank Austria d.d., Zagreb

Izvorni znanstveni članak*
UDK:657.92(035)
JEL: G21

Sažetak

Cilj je rada predstaviti novu metodologiju mjerenja sezonskih utjecaja (sezonskih oscilacija, vremenskih uzoraka) na ciljnu varijablu modela bodovanja kredita koji uzima u obzir ponašanje dužnika, i uključivanje tih utjecaja, ovisno o stupnju pouzdanosti, u spomenute modele kako bi oni dali što bolje rezultate. Takvi modeli bodovanja kredita imaju važnu ulogu unutar standarda Basel II prema internom sustavu rangiranja koji, za razliku od standardnog pristupa, "preciznije odražava individualni rizični profil banke". Taj se pristup oslanja na činjenicu kako banka najbolje poznaje svoje dužnike.

Rad također daje rješenje kako u modele bodovanja koji uzimaju u obzir ponašanje dužnika kredita uključiti i makroekonomske pokazatelje u obliku vremenskih serija (npr. tečaj, stopu likvidnosti, stope nezaposlenosti) radi poboljšanja mogućnosti predviđanja.

Zbog naravi takvih pojava koje su predočene vremenskim serijama, pojavljuju se određene teškoće, kako prilikom mjerenja, tako i prilikom uključivanja tih utjecaja u tradicionalne modele bodovanja. U radu se prikazuju rješenja tih problema primjenom REF II modela transformacije i analize vremenskih serija.

Ključne riječi: bodovanje kredita, REF II, analiza vremenskih serija, rudarenje podataka, temporalni utjecaji, sezonske oscilacije, Basel II

* Primljeno (Received): 1.6.2007.
Prihvaćeno (Accepted): 4.9.2007.

1. Uvod

Temeljni modeli bodovanja unutar standarda Basel II prema internom sustavu rangiranja jesu pristupni modeli bodovanja (engl. *application scorecards*). Njihova je temeljna zadaća inicijalno prepoznavanje potencijalno rizičnog klijenta prilikom razmatranja plasmana rizičnog proizvoda tom klijentu.

Modeli bodovanja koji uzimaju u obzir ponašanje dužnika primjenjuju se za predviđanje potencijalnoga rizičnog ponašanja (neplaćanja dospjelih obaveza) korisnika rizičnih bankarskih proizvoda i usluga prema standardima Basel II. Za izgradnju takvih vrsta modela služe povijesni podaci o ponašanju tih korisnika u prošlosti te se na temelju njih pokušava predvidjeti vjerojatnosti rizičnog ponašanja u budućnosti, u nastavku korištenja spomenutih proizvoda i usluga na temelju otkrivenih pravilnosti u uzorcima podataka (Thomas, 2002).

Korištenjem takvih vrsta modela financijske institucije mogu uspješnije upravljati svojim portfeljima rizičnih plasmana. Kako portfelj *sazrijeva* i s vremenom mijenja svoja svojstva bilo zbog ulazanja novih klijenata u portfelj, bilo zbog djelovanja mikroekonomskih i makroekonomskih utjecaja, potrebno je kontinuirano periodički mjeriti utjecaje i jačinu tih utjecaja na rizičnost portfelja. Za to se koristimo modelima bodovanja koji uzimaju u obzir ponašanje dužnika.

U razvoju modelâ bodovanja koji uzimaju u obzir ponašanje dužnika, a integralni su dio standarda Basel II, za predviđanje se koristimo uglavnom izvedenim atributima koji pokazuju određene oblike ponašanja tijekom nekoga vremenskog razdoblja, a predočeni su oblicima vrijednosti varijabli. Takva je metodologija pogodna za dizajniranje modela bodovanja primjenom uobičajenih metoda poput binominalne logističke regresije, neuronskih mreža, stabala odlučivanja i sl.

Tradicionalni pristup razvoju bodovanja kredita pri kojemu se uzima u obzir ponašanje dužnika uglavnom se ne bavi problematikom utjecaja sezonskih oscilacija na određene oblike rizičnog ponašanja (kašnjenja u otplatama). Autori su u literaturi uglavnom fokusirani na stvaranje uzorka podataka za izgradnju modela pomoću kojih će se minorizirati sezonski utjecaji na rizično ponašanje (Siddiqi, 2005; Thomas, 2002). Želimo li stabilan uzorak, takav pristup ima opravdanja, no on otvara i neka dodatna pitanja.

Ako sezonske oscilacije imaju jak utjecaj na portfelj korisnika rizičnih usluga, za zanemarivanje sezonskih utjecaja možemo razviti netransparentni model bodovanja izlaznog razdoblja (engl. *outcome period*) kraćega ili duljega od godine dana. Postavlja se pitanje zašto ne bismo mjerili te utjecaje, i ako se oni pokažu signifikantnima, zašto ih ne bismo uveli u model u obliku predikcijskih varijabli jer će one zasigurno pridonijeti pouzdanosti i većoj stabilnosti modela.

Kako ciljnu varijablu *stanje prijestupa* zbog kašnjenja možemo promatrati i kao vremensku seriju, moguće je izmjeriti sezonske utjecaje te ih, razmjerno otkrivenom stupnju utjecaja, uključiti u scoring model.¹

¹ Prema standardima Basel II stanje prijestupa (engl. *default*) jest kašnjenje u podmirenju novčanih obveza od 90 ili više dana.

Nadalje, makroekonomske pojave poput fluktuacije tečaja, povećanja stope nezaposlenosti ili pada likvidnosti također mogu imati velik utjecaj na varijacije stupnja kreditnog rizika.

U slučaju procijenjenog utjecaja makroekonomskih faktora predočenih u obliku vremenskih serija na rizično ponašanje korisnika, njihovim uključivanjem kao predikcijskih varijabli, također možemo povećati stupanj transparentnosti i pouzdanosti modelâ bodovanja koji uzimaju u obzir ponašanje dužnika (npr. slabljenje kune s obzirom na euro utječe na povećani rizik otplate kredita s valutnom klauzulom u eurima).

Jedan od mogućih razloga tradicionalno nedovoljnog uzimanja u obzir sezonskih utjecaja i makroekonomskih pokazatelja predočenih u obliku vremenskih serija uglavnom je rezultat činjenice da postoje problemi uključivanja vremenskih serija u klasične scoring modele. Ti se problemi ne očituju samo pri bodovanju, već su karakteristični i često spominjani pri rudarenju podataka (engl. *data mining*) (Berry, 1997; Han, 2001; Pyle, 2001; Williams, 2002).

Kako bi se uspješno riješio problem integracije vremenskih serija u klasične scoring modele i povećao stupanj pouzdanosti spomenutih modela, može se primijeniti model transformacije vremenske serije REF II (Klepac, 2005; 2006)².

Uvođenje sezonskih varijabli u modele bodovanja koji uzimaju u obzir ponašanje dužnika zahtijeva nov pristup kreiranju uzorka za provođenje analize jer je u izlaznom razdoblju dodatnu pozornost potrebno pridati razdoblju ulaska u stanje prijestupa. To znači da je potrebno ne samo evidentirati je li nastalo stanje prijestupa u izlaznom razdoblju, već i kada.

2. Preobrazba vremenske serije u model REF II

Preduvjet uključivanja sezonskih varijabli u scoring modele jest njihova preobrazba u model REF II, koja se ostvaruje u nekoliko koraka (Klepac, 2005; 2006).

Vremensku seriju možemo deklarirati kao niz vrijednosti $S(s_1, \dots, s_n)$, pri čemu S označava vremensku seriju, a (s_1, \dots, s_n) elemente serije S . Pogledajmo kako to izgleda u osam osnovnih koraka.

- Vremenska interpolacija (1)

Formiranje samostalnog vremenskog niza $V_i(v_{i1}, \dots, v_{in})$ na intervalu $\langle 1..n \rangle$ (dani, tjedni, mjeseci, kvartali, godine) s vrijednostima 0. Na temelju tako formiranog niza potrebno je interpolirati nedostajuće vrijednosti u $S(s_1, \dots, s_n)$ s 0 na temelju formiranog niza V_i . Rezultat takve obrade jest niz $S(s_1, \dots, s_n)$ s interpoliranim nedostajućim vrijednostima iz niza $V_i(v_{i1}, \dots, v_{in})$.

- Vremenska granulacija (2)

U tom koraku definiramo stupanj sažimanja vremenske serije $S(s_1, \dots, s_n)$ koja se nalazi u elementarnoj vremenskoj jedinici (dan, tjedan, mjesec...). Elemente postojeće vremenske serije sažimamo korištenjem statističkih funkcija poput sume, prosječne vrijednosti i

² REF je akronim engleskih riječi Raise, Equal Fall, dok II označava drugu generaciju modela.

moda na razini granuliranog odsječka. Na taj način vremensku seriju možemo svesti na veći stupanj granulacije (dani u tjedne, tjedni u mjesecima...), te dobivamo vremensku seriju $S(s_1, \dots, s_n)$ s većim stupnjem granulacije.

Ovisno o ciljevima, na taj se korak možemo vraćati tijekom procesa analize, što podrazumijeva obvezno ponovno provođenje procesa opisanih u idućim koracima.

• Normiranje (3)

Postupak normiranja podrazumijeva transformaciju vremenske serije $S(s_1, \dots, s_n)$ u vremensku seriju $T(t_1, \dots, t_n)$, pri čemu je svaki element niza podvrgnut postupku min-max normalizacije na intervalu $\langle 0, 1 \rangle$, i to:

- vremenska serija T sastoji se od elemenata (t_1, \dots, t_n) , pri čemu se t_i izračunava kao $t_i = [s_i - \min(S)] / [\max(S) - \min(S)]$, gdje su $\min(S)$ i $\max(S)$ minimalna i maksimalna vrijednost vremenske serije S ;
- vremenski pomak između elementarnih uzoraka (mjerilo vremenske kompleksnosti) odsječka osi X određen je s $d(t_i, t_{i+1}) = a$.

• Transformacija u REF notaciju prema formuli $T_r = t_{i+1} - t_i$ $T_r > 0 \Rightarrow R$; $T_r < 0 \Rightarrow F$; $T_r = 0 \Rightarrow E$, gdje su Y_i elementi niza N_s (4)

• Proračun nagiba pravca na osnovi kuta (5)

Koeficijent kutnog odklona \Rightarrow

$T_r > 0$ (R) koeficijent = $t_{i+1} - t_i$

$T_r < 0$ (F) koeficijent = $t_i - t_{i+1}$

$T_r = 0$ (E) koeficijent = 0

• Proračun površine ispod krivulje (6)

Numerička integracija metodom pravokutnika

$$p = [(t_i * a) + (t_{i+1} * a)] / 2$$

• Kreiranje vremenskih indeksa (7)

Riječ je o građenju hijerarhijskog stabla indeksa ovisno o karakteru analize, pri čemu element strukturiranog indeksa može biti i atribut poput šifre klijenta.

• Kreiranje razreda (8)

Obuhvaća kreiranje izvedenih vrijednosti atributa na temelju površine ispod krivulje i odklona kutova.

Moguće je kreirati razrede primjenom izrazite (engl. *crisp logic*) ili neizrazite logike (engl. *fuzzy logic*).

Navedenih osam osnovnih koraka temelj su algoritimiziranog postupka na kojemu se temelji model REF II, čiji je krajnji rezultat formiranje matrice transformacije. Matrica transformacije temelj je za provođenje daljnjih analitičkih postupaka u analizi vremenske serije.

Model REF II u osnovi je koncepcija sastavljena od tri podcjeline (REF-a, površine ispod krivulje te koeficijenta kutnog odklona), kojoj je glavna zadaća transformacija vre-

menske serije u niz pokazatelja koji jednoznačno definiraju vremensku seriju. REF, površina ispod krivulje te koeficijent kutnog otklona pokazatelji su koji jednoznačno opisuju odsječak vremenske serije, a niz takvih odsječaka zajedno čini transformiranu vremensku seriju. Takav niz pokazatelja poredan je redosljedom kako se pojavljuje u jedinici vremena i objedinjen je u zajedničku koncepcijsku strukturu transformirane vremenske serije koju nazivamo *matricom transformacije*.

Grafički prikaz te strukture dan je u tablici 1.

Tablica 1. Matrica transformacije

Indeks vremenskog odsječka	I1	I2	...	In
REF oznaka	REF 1	REF 2	...	REF n
Koeficijent kutnog otklona	KKO 1	KKO 2	...	KKO n
Površina vremenskog odsječka	P1	P2	...	Pn

Izvor: izračun autora

Pokazatelji vremenskog odsječka izračunani su na temelju koordinata dviju susjednih vrijednosti u vremenskoj seriji. Tako je odsječak s indeksom I1 formiran na osnovi vrijednosti koordinata vremenske serije t_0 i t_1 . Indeks vremenskog odsječka služi za jednoznačnu identifikaciju vremenskog odsječka radi njegove analize, odnosno kreiranje bazičnog uzorka.

Indeksi mogu biti složeno strukturirani te mogu sadržavati hijerarhijske elemente, elemente pripadnosti i vezne elemente prema ostalim izvorima podataka.

Elementi obuhvaćeni u prethodnoj tablici temelj su modela REF II i njima je moguće jednoznačno opisati krivulju te provesti sve analize za koje je taj model i razvijen. Osim opisanih, moguće je obuhvatiti i izvedene pokazatelje, ali to ovisi o prirodi analize. Nakon transformacije vremenske serije putem modela REF II dobijemo transformiranu vremensku seriju prikazanu u tablici 1.

3. Mjerenje mikroekonomskih i makroekonomskih sezonskih utjecaja na podmirivanje obveza

Uredno podmirenje dospjelih obveza rizičnih bankarskih usluga može ovisiti o profilu korisnika i dijelu godine u kojemu korisnik mora podmiriti svoje obveze. Zsigurno postoji profil korisnika čije ponašanje određuje, primjerice, sezona božićnih i novogodišnjih blagdana, ili pak sezona godišnjih odmora.

Ti su utjecaji na rizik podmirjenja dospjelih obveza manje ili više izraženi, ovisno o tržišnom segmentu koji čini portfelj korisnika. Pomoću modela REF II možemo izmjeriti koliko na segment za koji se razvija model bodovanja što uzima u obzir ponašanje dužnika na urednost otplate kredita utječu sezonski faktori (doba godine). Takvu vrstu utjecaja možemo ubrojiti u mikroekonomske temporalne faktore jer su karakteristični za određeni segment na razini promatrane financijske institucije.

Oscilacije tečaja također mogu uvelike utjecati na urednost podmirivanja dospjelih obveza po rizičnim bankarskim uslugama vezanima za stranu valutu. Te oscilacije možemo promatrati u obliku vremenskih serija te njihove utjecaje na rizik nepodmirenja dospjelih obveza mjeriti REF II modelom. Takvu vrstu utjecaja možemo ubrojiti u makroekonomske temporalne faktore. Osim oscilacija tečaja, možemo mjeriti kretanje stope nezaposlenosti ili pak likvidnost sektora u kojemu rade korisnici rizičnih usluga (na temelju segmentacije) te tako mjeriti urednost podmirivanja dospjelih obveza rizičnih bankarskih usluga. Metodologija Basel II dopušta razvoj bodovanja na razini određenih segmenata (engl. *asset class segments*), u kojemu nam kriteriji za provođenje segmentacije, među ostalima, mogu biti i gospodarski sektori.

Iz tablice je vidljivo ($IV = 1,04$) koliko je varijabla trenda stanja prijestupa na mjesečnoj razini značajna za stvaranje modela bodovanja koji uzimaju u obzir ponašanje dužnika. Također je vidljivo kako tijekom travnja, svibnja, lipnja i srpnja (mjere WE) imamo očekivano manji broj neurednih otplata kredita, dok krajem i početkom godine možemo očekivati veću vjerojatnost neurednih otplata kredita.

S mikrorazine te trendove krajem godine možemo hipotetički objasniti novogodišnjim praznicima i profilom korisnika kreditnih usluga unutar promatranog portfelja.

Postavlja se pitanje kako protumačiti uredne otplate tijekom ljetnog razdoblja.

Pretpostavimo da je ova analiza rađena na temelju korisnika kredita s deviznom klauzulom u eurima. Kad bismo željeli provesti dublju analizu, mogli bismo analizirati trend tečaja eura u odnosu prema kuni, odnosno analizirati sezonske oscilacije tog trenda.

Tablica 2. Univarijantna analiza vremenske serije uz pomoć modela REF II

Mjeseci	Prijestup		Analiza relevantnosti	
	trend rasta (R) (%)	trend pada (F) (%)	WE	IV
siječanj	8,0	4,0	-0,693150	0,027726
veljača	10,0	9,0	-0,105360	0,001054
ožujak	13,0	6,5	-0,690400	0,044752
travanj	6,0	7,0	0,154151	0,001542
svibanj	6,7	9,0	0,295117	0,006788
lipanj	9,0	10,0	0,105361	0,001054
srpanj	9,3	36,0	1,353505	0,361386
kolovoz	5,0	6,5	0,265116	0,004024
rujan	3,0	8,0	0,980829	0,049041
listopad	20,0	2,0	-2,302590	0,414465
studen	5,0	1,0	-1,609440	0,064378
prosinac	5,0	1,0	-1,609440	0,064378
				1,040586

Izvor: izračun autora

Temeljna transformacija u model REF II provedena je na vremenskoj seriji tečaja eura u odnosu prema kuni u razdoblju siječanj 1999 – siječanj 2007. Vremenska serija tečaja na mjesečnim razinama dobivena je kao prosjek dnevnih srednjih vrijednosti tečaja HNB-a po mjesecima za navedeno razdoblje. Tako formirana vremenska serija transformirana je u notaciju REF II prema formulama/postupcima (1), (2), (3), (4), (5).

Za analizu sezonskih oscilacija tečaja potrebna je tablica sljedeće strukture.

Ti se pokazatelji mogu kombinirati s tradicionalnim varijablama koje se upotrebljavaju prilikom razvoja modela bodovanja koji uzimaju u obzir ponašanje dužnika, npr. s brojem stanja prijestupa u nekom razdoblju, ukupnim brojem dana kašnjenja u nekom razdoblju, brojem proteklih dana od izlaska iz posljednjeg stanja prijestupa i sl.

Prilikom mjerenja relevantnosti atributa i univarijantne analize obično se oslanjamo na težinu dokaza i vrijednost informacija (engl. *weight of evidence* – WE i *information value* – IV) (Siddiqi, 2005; Thomas, 2002).

Tablica 3. Analiza sezonskih oscilacija tečaja eura u odnosu na kunu pomoću REF II modela (u %)

Mjesec	Pad tečaja	Rast tečaja
siječanj ^a	11,1	77,8
veljača	37,5	62,5
ožujak	62,5	37,5
travanj	87,5	12,5
svibanj	100	0
lipanj	87,5	12,5
srpanj	87,5	12,5
kolovoz	37,5	62,5
rujan	37,5	62,5
listopad	37,5	62,5
studeni	37,5	62,5
prosinac	37,5	62,5

^a 11,1% u siječnju 1999. neodređeni je trend jer u podacima nisu bile dostupne vrijednosti tečaja za prosinac 1998.

Izvor: izračun autora

Iz priložene je tablice vidljivo kako s vrlo velikom sigurnošću možemo ustvrditi da će tijekom svibnja euro u odnosu prema kuni oslabjeti, te se taj trend s nešto manjom vjerojatnošću proteže i tijekom lipnja i srpnja, a takav je bio i u travnju. Ta analiza pokazuje zone sezonskih oscilacija. Ovisno o stupnju pouzdanosti, možemo prihvatiti hipoteze o sezonskim oscilacijama. Ako je taj stupanj pouzdanosti veći od 85%, prihvaćamo hipo-

tezu o sezonskoj oscilaciji tečaja u razdoblju travanj – srpanj. S obzirom na to da model REF II može mnogo preciznije transformirati vremensku seriju, tu analizu možemo učiniti još preciznijom uzimajući u obzir površinu ispod krivulje i precizne otklone trendova. S obzirom na to da nam je cilj prikazati metodologiju uključivanja vremenskih serija u modele bodovanja, ponajprije ćemo se usmjeriti na tu jednostavnu transformaciju. Prikazana se metodologija, naravno, može iskoristiti i u kompleksnijoj transformaciji. Na temelju dobivenih rezultata možemo postaviti hipotezu o utjecaju oscilacije tečaja na urednost otplate kredita s valutnom klauzulom u eurima.

Osim utjecaja oscilacija tečaja, možemo mjeriti i utjecaj stope nezaposlenosti određenog sektora u kojemu prevladavaju korisnici kredita te na isti način u model možemo uključiti i ostale makroekonomske faktore poput likvidnosti, kretanja trendova cijena na malo i sl.

Iako tradicionalne tehnike razvoja bodovanja podrazumijevaju primjenu tih pokazatelja, važno je napomenuti kako je nakon njihovih izračuna za razumijevanje i profiliranje ponašanja rizičnog klijenta bitno spoznati logičnu povezanost analiziranih faktora. Mogućnost logičnog tumačenja faktora rizika presudna je komponenta prilikom izbora varijabli za kreiranje modela. Uglavnom se teži logičnom tumačenju ovisnosti svake varijable koje ulazi u model.

Ta je metodologija zbog sezonskih obilježja teško primjenjiva za procjenu utjecaja temporalnih varijabli na stanje prijestupa.

Kako bi se premostio taj problem, potrebno je vremensku seriju (stanja prijestupa po mjesecima) transformirati u model REF II. Vremenska serija stanja prijestupa kreirana je kao agregacijska funkcija broja stanja prijestupa i slučajeva u kojima nije zabilježen prijestup na mjesečnoj razini. Kako bismo prikazali metodologiju, mjesečne trendove stanja prijestupa unatrag sedam godina transformiramo u notaciju REF II, što omogućuje provedbu analize relevantnosti atributa na tradicionalan način unutar modela bodovanja pomoću WE i IV mjera. Tablica 1. ilustrira proces univarijantne analize vremenske serije uz pomoć modela REF II, prema formulama/postupcima (1), (2), (3), (4), (5).

4. Predviđanja na temelju sezonskih čimbenika

Metodologiju grupiranja i stvaranja pseudovarijabli (engl. *dummy variables*) pri razvoju modelâ bodovanja koji ne uzimaju u obzir sezonske utjecaje moguće je primijeniti i u modelima s uključenim sezonskim utjecajima.

Na temelju sezonskih varijabli mogu se kreirati pseudovarijable, i to tako da pseudovarijabla koja se, primjerice, odnosi na razdoblje siječanj – ožujak poprima vrijednost 1 kad se odnosi na to razdoblje, a u protivnome poprima vrijednost 0.

Kreirane varijable ulaze kao prediktori u model nakon provedbe korelacijske analize i sudjeluju u kreiranju modela bodovanja.

Iako su kao primjeri uzete situacije koje korespondiraju sa sezonalnošću vremenske serije, za predviđanje prijestupa moguće je na temelju vremenskih atributa ugraditi i svojevrsni mehanizam *što-ako* (engl. *what-if*). Analiziramo li, primjerice, stopu likvidnosti unutar nekog sektora koji kao poslodavac dominira u promatranom portfelju, tada je tu varijablu moguće predočiti u vremenskom obliku.

Univarijantom analizom tako stvorene vremenske serije, sukladno prethodno prikazanom primjeru možemo analizirati koliki je utjecaj pada stope likvidnosti u nekom sektoru na makroekonomskoj razini na urednost otplate kredita unutar određenog portfelja.

Taj nam podatak može pomoći u budućim predviđanjima (što je i zadatak modelâ bodovanja koji uzimaju u obzir ponašanje dužnika) kako bismo preciznije odredili rizičnost našeg portfelja pri nekoj budućoj promjeni stope likvidnosti u određenom sektoru. Tako modeli ponašanja koji uzimaju u obzir ponašanje dužnika osim za svoju temeljnu zadaću mogu poslužiti i za kreiranje scenarija ako na taj način u njih integriramo dovoljan broj makroekonomskih varijabli. Pri tome sve makroekonomske varijable ne moraju nužno poprimiti vremenski karakter. Pristup koji se oslanja na stvaranje scenarija omogućuje prepoznavanje modela ponašanja i osjetljivosti segmenata našeg portfelja, što nam može biti putokaz pri donošenju strateških odluka. Takav je pristup u skladu s metodologijom testiranja portfelja u ekstremnim poslovnim uvjetima (engl. *stress testing*).

5. Ocjena pouzdanosti modelâ bodovanja koji uzimaju u obzir ponašanje dužnika uz uključena sezonska predviđanja

Kako bi se ispitala hipoteza o boljim rezultatima bodovanja kad se uključe sezonska predviđanja, izgrađen je model bodovanja bez predviđanja sezonskih varijabli i model u kojemu su predviđene sezonske varijable. Za to su korišteni nizovi podataka (Pyle, 2001; Thomas, 2002; Modelandmine, 2007; Kdnuggest, 2007).

Svaki uzorak podataka podijeljen je u uzorak za treniranje i uzorak za testiranje u omjeru 80:20. Trend stanja prijestupa tumači se kao vremenska serija i pretvoren je u model REF II.

Univarijantna analiza pokazala je sezonsku prirodu te varijable sa značajnom signifikantnošću na temelju IV vrijednosti. Nakon kreiranja pseudovarijabli na temelju značajnih obilježja provedena je korelacijska analiza.

Na uzorku za treniranje pomoću binominalne logističke regresije napravljena su dva modela bodovanja, s tim da za razvoj prvog modela nije korištena vremenska varijabla, a za razvoj drugoga jest.

Na uzorku za testiranje ispitana je pouzdanost svakog pojedinog modela, a rezultati analize modela iz prvog niza podataka prikazani su u sljedećoj tablici.

Tablica 4. Testiranje rezultata modeliranja s predviđanjem sezonskih varijabli ili bez tog predviđanja (u %)

Vrsta testa	Model bodovanja kredita		Razlika u rezultatu
	bez predviđanja sezonskih varijabli	s predviđanjem sezonskih varijabli	
Kolmogorov-Smirnov test	61,22	62,83	+1,61
Ginijev indeks	71,4	73,51	+2,08

Izvor: izračun autora

Iz priložene je tablice vidljivo kako model u kojemu su predviđene sezonske varijable daje pouzdanije rezultate od modela u koji one nisu uključene.

Testirana su još tri podataka iz spomenutih izvora i dobiveni rezultati potvrđuju postavljenu hipotezu.

Potrebno je uzeti u obzir činjenicu da je učinak sezonalnosti očit u izlaznom razdoblju, koje je kraće ili dulje od godine dana, jer unutar izlaznog razdoblja točno za godinu dana eliminiramo utjecaj sezonalnosti.

6. Zaključak

Iako se prilikom izgradnje modelâ bodovanja koji uzimaju u obzir ponašanje dužnika nastoje minorizirati sezonski utjecaji maksimalnim proširivanjem izlaznih razdoblja (engl. *outcome window*), takav je pristup neprikladan kada želimo na kraći rok (npr. na tri mjeseca) ili na rok koji se ne podudara s kalendarskim završecima godina predvidjeti urednost otplate korisnika rizičnih proizvoda. U tom su nam slučaju za preciznije predviđanje važni sezonski faktori i utjecaji.

Nadalje, sezonski faktori mogu poboljšati rezultate modela bodovanja, što je vidljivo iz prikazane ocjene pouzdanosti modelâ u kojima jesu, odnosno u kojima nisu predviđeni sezonski utjecaji. Modeli u kojima su predviđeni sezonski utjecaji daju bolje rezultate, posebice unutar kraćih promatranih izlaznih razdoblja.

Predloženi model potencira sezonske faktore umjesto da ih neutralizira, što je posebno važno za izlazna razdoblja manja od godine dana, te se na taj način omogućuju preciznija predviđanja.

Predložena metodologija, nadalje, daje rješenje kako u modele bodovanja integrirati vremenske serije kao prediktore, uvodeći u njih i makroekonomske i mikroekonomske faktore. To je posebno važno jer modeli bodovanja pridonose boljem razumijevanju strukture portfelja, obrazaca ponašanja njihovih pripadnika te daju odgovore na to kako određeni faktori (pa i makroekonomski) direktno utječu na konkretan portfelj.

Ta se metodologija može iskoristiti i pri razvoju modela bodovanja za procjenu rizika od potencijalnih prijevara u financijskom poslovanju, te u svim vrstama modela bodovanja koji uzimaju u obzir ponašanje dužnika iz domene financijskog sektora vezanih za standarde Basel II.

Metodologija se ne mora nužno odnositi na procjene rizika u užem smislu, već se može iskoristiti i za donošenje strategije nastupa na tržištu. Tako se prilikom osmišljavanja marketinških kampanja kojima se želi privući nove klijente u financijsku instituciju prilikom određivanja ciljne skupine možemo voditi i kriterijima profila rizičnosti. Na taj način možemo težiti optimizaciji rizičnosti portfelja na dulji rok ako je to naša strateška odluka.

LITERATURA

Basel Comitee on Banking Supervision, 2001. *Overview of the new capital accordp.* Basel, Bank for International Settlements.

Berry, Michael, J. A. and Linoff, G., 1997. *Data mining techniques for marketing sales and customer support.* San Francisco: John Wiley & Sons Inc.

Han, J. and Kamber, M., 2001. *Data mining-concepts and techniques.* London: Morgan Kaufmann publishers.

Klepac, G. i Panian, Ž., 2006. *Poslovna inteligencija.* Zagreb: Masmedia.

Klepac, G., 2005. *Otkrivanje zakonitosti temeljem jedinstvenoga modela transformacije vremenske serije.* Doktorska disertacija. Varaždin: Fakultet organizacije i informatike.

Lee, Y. [et al.], 1997. *Solving Data Mining Problems Through Pattern Recognition.* New York: Prentice Hall

Modelandmine, 2007. <http://www.modelandmine.com/dataprep.htm>.

Naem, S., 2005. *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring.* San Francisco: John Wiley&Sons Inc.

Ozsu, T. and Chen, L., 2003. "Multi-Scale Histograms for Answering Queries over Time Series Data" in: *Proceedings of the 20th International Conference on Data Engineering, March 30 - April 2.* Boston, MA.

Pyle, D., 2001. *Data preparation for data mining.* London: Morgan Kaufmann publishers.

Thomas, L., Edelman, D. and Crook, J., 2002. *Credit Scoring and Its Applications.* Philadelphia: SIAM.

Wang, W. and Yang, J. Yu P., 2001. "Mining long sequential patterns in a noisy environment" [online]. *IBM research report.* Available from: [http://www-sal.cs.uiuc.edu/~hanj/pdf/ww1sigmod02_1.pdf].

Williams, J. Graham, Weiqiang Lin, Mehmet A. Orgun 2002. "An Overview of Temporal Data Mining". *Proceedings of the 1st Australian Data Mining Workshop, Canberra, 2002.* **Kdnugget, 2007.** <http://www.kdnuggets.com/datasets/index.html>.

G o r a n K l e p a c : Integrating Seasonal Oscillations into Basel II Behavioral Scoring Models

Abstract

The article introduces a new methodology of temporal influence measurement (seasonal oscillations, temporal patterns) for behavioral scoring development purposes. The article shows how significant temporal variables can be recognized, and how they can be integrated in the behavioral scoring models with the aim of improving model performance.

Behavioral scoring models are integral parts of Basel II standard on Internal Ratings-Based Approaches (IRB). The IRB approach is much more precise in its reflection of bank individual risk profile (Basel Committee on Banking Supervision, 2001).

A solution of the problem of how to analyze macroeconomic and microeconomic factors represented in time series and integrate them into behavioral scorecard models will be shown in the paper with the use of the REF II model.

Keywords: credit scoring, REF II, time series analyze, data mining, temporal influence, seasonal oscillation, Basel II