

סדרת ניירות עבודה
WORKING PAPER SERIES

מס' 44 No.

**שימוש בשיטות סטוכסטיות
לביון וחיוץ של סדרות עתיות**
**Using stochastic methods for
interpolation and extrapolation
of time series**

יורי גובמן*

Yury Gubman*

אייר תשס"ט, מאי 2009 May

* הלשכה המרכזית לסטטיסטיקה – אגף מדען ראשי

* Central Bureau of Statistics – Chief Scientist Department

הוצאת הלשכה המרכזית לסטטיסטיקה, רח' כנפי נשרים 66, פינת רח' בקי,

ת"ד 34525, ירושלים 91342

טל': 02-6592666; פקס: 02-6521340

אתר הלמ"ס באינטרנט: www.cbs.gov.il

דואר אלקטרוני: info@cbs.gov.il

הלשכה המרכזית לסטטיסטיקה (הלמ"ס) מעודדת מחקר המבוסס על נתוני הלמ"ס. פרסומי תוצאות מחקרים אלו אינם פרסומים רשמיים של הלמ"ס, והם לא עברו את הביקורת שעוברים פרסומים רשמיים של הלמ"ס. הדעות והמסקנות המתבטאות בפרסומים אלו, כולל בפרסום זה, הן של המחברים עצמם ואינן משקפות בהכרח את הדעות והמסקנות של הלמ"ס. פרסום מחדש של העבודה, כולה או מקצתה, טעון אישור מוקדם של המחברים.

רחוב כנפי נשרים 66 פינת רחוב בקי, גבעת שאול, ת"ד 13015, ירושלים 95464 טלפון: 02-6592666, פקס' 02-6521340

דואר אלקטרוני: info@cbs.gov.il כתובת האתר: www.cbs.gov.il

תקציר

במקרים רבים, נתונים אודות תופעות כלכליות-חברתיות זמינים בתדירות נמוכה בלבד, לעיתים גם בפרקי זמן באורך שונה, וקיימת דרישה לנתונים אלה בתדירות גבוהה יותר. בסטטיסטיקה הרשמית, הנתונים אמורים להיות זמינים מדי תקופה קצובה מראש (חודש, רבעון, שנה), כאשר נדרשת ההתאמה של נתונים אלה לנתון הזמין בתדירות נמוכה. בלשכות סטטיסטיות, התאמה זו חייבת להיות מלאה ומדויקת, על מנת למזער סטייה בין סדרות האומדנים השונות. בנוסף, קיים צורך למנוע סתירות בין מקורות מידע שונים המתפרסמים על ידי הלשכה הסטטיסטית.

לצורך בניית הסדרה העונה על הדרישות הנ"ל נהוג להשתמש בסדרה עתית אחרת אשר זמינה בתדירות הרצויה ומתארת תופעה כלכלית-חברתית דומה, הנקראת אינדיקטור. בניית סדרה עתית בתדירות גבוהה על סמך האינדיקטור ועל סמך הנתונים בתדירות נמוכה, בתוך הטווח של הנתונים בתדירות נמוכה, ידועה כבעיית ביון סדרות עתיות. אותה הפעולה הנעשית מחוץ לטווח הנתונים בתדירות נמוכה, ידועה כבעיית חיוץ. במקרה האחרון, אנו משתמשים בידע היסטורי על התופעה הנאמדת ובידע עדכני על התופעה הקרובה אשר זמינה בתדירות גבוהה לבניית התחזית. תחזית זו נתונה לשינוי עם קבלת הנתון העדכני בתדירות נמוכה.

בעבודה הנוכחית, נדונו שיקולים מנחים לבחירת האינדיקטור לסדרה הנאמדת. הוצעה שיטת בדיקה תיאורית פשוטה העונה על השאלה אודות מידת הדמיון בין התופעה המתוארת על ידי האינדיקטור לבין התופעה הנאמדת אשר ידועה בתדירות נמוכה בלבד. נחקר טיב הקירוב של שיטות נומריות של דנטון על ידי מודל הרגרסיה. נבדקו חלופות שונות בבחירת מודל רגרסיה לביון וחיוץ בסדרה. נערך ניתוח נרחב של שיטת הרגרסיה תוך שימוש במודלים ARMA לשאריות: (1) השומרים על התנודות (moving preservation models), (2) מודל ARMA לשאריות הנאמד מהנתונים המקוריים. במקרה (2), מוצע אלגוריתם לאמידה של מודל ARMA ולבחירת מודל לביון וחיוץ. נערכו השוואות של השיטות המוצעות עם השיטה הקיימת בלמ"ס על סמך ניתוח של סדרות נבחרות של חשבונות לאומיים ואומדנים חודשיים של סקר כוח אדם. נבחנו קריטריונים לבחירת השיטה המתאימה לכל סדרה עתית העוברת תהליך של ביון וחיוץ.

מוצע להשתמש בשיטת הרגרסיה אם חותך משתנה לביון וחיוץ בסדרות השנתיות והרבעוניות של חשבונות לאומיים הנאמדות על סמך נתונים הזמינים פעם במספר שנים. המודל לחותך מבוסס על הנחות כלכליות אודות התנהגות ההטיה כפונקציה של המרחק מהנתון הרב-שנתי.

מילות מפתח: ביון, חיוץ, שיטת הרגרסיה, תהליכי ARMA השומרים על תנודות.

תודות

לדרי' דמיטרי רומנוב – מדען ראשי, על ההערות וההארות המועילות שניתנו בשלב כתיבת העבודה.
לגבי' סולי פלג – ראש אגף מקרו כלכלה, על ההסברים המקיפים על הפקת אומדני חשבונות לאומיים.
לגבי' גליה יוחאי – ראש תחום תעשייה ועסקים, על שיתוף הפעולה בהטמעת השיטה ורעיונות לשיפורה.
למר מרק פלדמן – ראש תחום עבודה, אגף מיקרו-כלכלה, על ההסברים המקיפים על הפקת האומדנים
החודשיים של סקר כוח אדם.

תוכן עניינים

7.....	1. מבוא
9.....	2. שיטות ביון וחיוץ.....
9.....	2.1 שיטות של דנטון.....
10.....	2.2 שיטת הרגרסיה.....
12.....	3. סוגיות מתודולוגיות - קביעת מודל רגרסיה ומודל ARMA לשאריות.....
13.....	3.1 קירוב השיטות של דנטון באמצעות קביעת מודל ARMA.....
13.....	3.2 קביעת מודל רגרסיה.....
14.....	3.3 מודלים ARMA השומרים על תנודות.....
15.....	3.4 אמידת מודל ARMA לשאריות מהנתונים.....
17.....	3.5 ביון וחיוץ של סדרות שנתיות.....
19.....	3.6 benchmarking בסדרות סכ"א.....
20.....	4. נתונים מקוריים ודיאגנוסטיקות לטיב האמידה.....
20.....	4.1 תאור נתונים.....
21.....	4.2 שלבי המחקר.....
23.....	4.3 דיאגנוסטיקות לבדיקת טיב ביון, חיוץ וניכוי עונתיות.....
25.....	5. ממצאים עיקריים.....
25.....	5.1 שלב מקדים - בחירת אינדיקטורים.....
25.....	5.2 שלב א' - ניתוח 3 הסדרות הנאמדות על סמך אומדני מע"מ (case study).....
28.....	5.3 שלב ב' - ניתוח 7 הסדרות הנאמדות על סמך מדדי תעשייה.....
29.....	5.4 בחירת מודל לביון וחיוץ בסדרות חודשיות של סכ"א.....
30.....	6. הרחבת המחקר - שימושים נוספים בלמ"ס.....
31.....	7. מסקנות והמלצות.....
34.....	8. רשימה ביבליוגרפית.....
36.....	נספחים.....

1. מבוא

במקרים רבים, נתונים אודות תופעות כלכליות-חברתיות זמינים בתדירות נמוכה בלבד, לעיתים גם בפרקי זמן באורך שונה. עם זאת, קיימת דרישה לנתונים אלה בתדירות גבוהה יותר. בסטטיסטיקה הרשמית, הנתונים אמורים להיות זמינים מדי תקופה קצובה מראש (חודש, רבעון, שנה), כאשר נדרשת ההתאמה של נתונים אלה לנתון הזמין בתדירות נמוכה. כך, למשל, סדרות של אומדנים דמוגרפיים אמורים להיות שווים לתוצאות המפקד עבור השנה בא נערך המפקד. בלשכות סטטיסטיות נדרשת, ברוב המקרים, התאמה מלאה ומדויקת, על מנת למזער סטייה בין סדרת האומדנים בתדירות גבוהה לבין התופעה האמיתית הנצפית בתדירות נמוכה בלבד. בנוסף, קיים צורך למנוע סתירות בין מקורות מידע שונים המתפרסמים על ידי הלשכה הסטטיסטית.

לצורך בניית הסדרה העונה על הדרישות הנ"ל נהוג להשתמש בסדרה עתית אחרת אשר זמינה בתדירות הרצויה ומתארת תופעה כלכלית-חברתית דומה. לסדרה זאת קוראים סדרת האינדיקטור. בניית סדרה עתית בתדירות גבוהה על סמך האינדיקטור ועל סמך הנתונים בתדירות נמוכה, בתוך הטווח של הנתונים בתדירות נמוכה, ידועה כבעיית ביון סדרות עתיות. אותה הפעולה הנעשית מחוץ לטווח הנתונים בתדירות נמוכה, ידועה כבעיית חיוץ. במקרה האחרון, אנו משתמשים בידע היסטורי על התופעה הנאמדת ובידע עדכני על התופעה הקרובה אשר זמינה בתדירות גבוהה לבניית התחזית. תחזית זו נתונה לשינוי עם קבלת הנתון העדכני בתדירות נמוכה. יש לציין, כי בעיות ביון וחיוץ של סדרות עתיות מהוות מקרים פרטיים של בעיית ה-Benchmarking (ראה Gubman and Burck, 2005).

ביון וחיוץ של סדרות עתיות יכולים להתבצע באמצעות שיטות שונות. כך, נהוג להבדיל שיטות נומריות, המבוססות על מזעור פונקציית המטרה המוגדרת מראש, לבין שיטות סטוכסטיות, המבוססות על מודלים לסדרה עתית או/ו לשאריות. בלשכות הסטטיסטיות, השיטות הנומריות הנפוצות ביותר הינן שיטת של דנטון (ראה סעיף 2.1) והשיטה הסטוכסטית הנפוצה ביותר הינה שיטת הרגרסיה (ראה סעיף 2.2).

בלשכה המרכזית לסטטיסטיקה (למ"ס), ביון וחיוץ מתבצע במספר רב של סדרות עתיות. בשנת 2005 שודרגה מערכת של ביון וחיוץ בתחום חשבונות לאומיים, והוטמעה התוכנה BENCH, אשר פותחה בלשכה הסטטיסטית הקנדית ב-1994. תוכנה זו מבוססת על שיטת הרגרסיה שהיא שיטה סטוכסטית לביצוע ביון וחיוץ. לפי מסקנת המחקר שנערך לפני שדרוג תהליך הביון והחיוץ של סדרות עתיות בלמ"ס (ראה גובמן ובורק, 2005), קירוב לשיטה הפרופורציונאלית של דנטון נקבע כמודל הסטנדרטי לכל הסדרות בהן מתבצעים ביון וחיוץ. לפי מסקנת המחקר, מודל זה הוא המודל הטוב ביותר לסדרות הלמ"ס לבין השיטות הנומריות השונות שנבדקו. בפרק המסקנות צוין, כי קיים צורך לעקוב אחרי התפתחות הסדרות ולחקור אפשרויות שימוש במודלים נוספים על מנת להעלות טיב האומדנים המתקבלים.

עם זאת, אפשרויות רבות הזמינות תוך שימוש בשיטת הרגרסיה אינן מנוצלות היום בלמ"ס. שימוש בשיטת זו מאפשר ניצול מידע נוסף אודות ההתנהגות של סדרה עתית ושאריות המודל. כך, ניתן להשתמש בידע על קיום עונתיות בסדרה הנאמדת, מבנה תלות שאריות ומבנה הסטייה בין האינדיקטור לבין הסדרה הנאמדת. בנוסף, התנהגותן של סדרות עתיות המתארות תופעות חברתיות-כלכליות שונות אינה דומה לאורך זמן, ולכן שימוש באותו מודל לכל הסדרות מתחומי מחקר שונים נראה לא מוצדק. שימוש בשיטת הרגרסיה מאפשרת התאמת מודל ייחודי לכל סדרה עתית או קבוצת סדרות עם מאפיינים דומים, בדומה לנעשה בתהליך ניכוי עונתיות. אפשרויות אלו לא נבדקו בעבודות הקודמות, והשיטה הקיימת היום בלמ"ס לא מאפשרת התייחסות אליהן. בנוסף, עד היום כל הסדרות השנתיות וחלק מהסדרות הרבעוניות עוברות תהליך זה לפי השיטה המיושנת, המבוססת על תהליך איטרטיבי אשר מתבצע ידנית בסביבת EXCEL.

בעבודה הנוכחית, נדונו שיקולים מנחים לבחירת האינדיקטור לסדרה הנאמדת. הוצעה שיטת בדיקה תיאורית פשוטה העונה על השאלה אודות מידת הדמיון בין התופעה המתוארת על ידי האינדיקטור לבין התופעה הנאמדת אשר ידועה בתדירות נמוכה בלבד. נציין, כי קיימת חשיבות רבה לבדיקה זו, היות ובחירה לא נכונה של סדרת האינדיקטור עלולה לגרום להורדה משמעותית בטיב האומדנים המתקבלים. במידה ומתברר, כי התנהגות האינדיקטור שונה משמעותית מזו של הסדרה בתדירות נמוכה, מופרת הנחת היסוד של שיטת הביון והחיוץ על סמך האינדיקטור. במקרה זה, יש להחליף אינדיקטור, או, אם הדבר אינו אפשרי מסיבות נושאות, להתחשב בעבודה זו בבניית מודל לשאריות או/ו לטעות המדידה. היום בלמ"ס לא נעשית בחינת האינדיקטור לקיום הנחת דמיון לסדרה הנאמדת, מה שמטיל ספק על טיב האומדנים המתקבלים ועלול לייצור סתירות בין מקורות מידע שונים.

נחקר טיב הקירוב של שיטות נומריות של דנטון על ידי מודל הרגרסיה. נבדקה אפשרות שימוש במודל כפלי וכדאיות אמידת החותך במודל הרגרסיה. נערך ניתוח נרחב של שיטות סטוכסטיות המבוססות על מודלים ARMA השומרים על התנודות (moving preservation models). הוצע אלגוריתם המאפשר שימוש בשיטה סטוכסטית לביון וחיוץ בסדרות עתיות. נערכה השוואה של השיטות המוצעות עם השיטה הקיימת. נדגיש, כי כל השיטות המוצעות אינן דורשות קיום הנחות חזקות, אשר מגבילות את השימוש בשיטה או/ו לא ניתנות לבדיקה.

מודלים ARMA השומרים על תנודות הוצגו ב- (2004) Cholette and Chhab ו- Cholette and Dagum (2006). שימוש במודל אוטורגרסיבי מסדר ראשון השומר על תנודות (AR(1)) במודלים סטוכסטיים נדון ב- (2006) Chen and Wu, תוך השוואה עם האומדנים של השיטות הקלאסיות של דנטון. בכל העבודות הנ"ל הניתוח בוצע באמצעות סימולציות.

יש לציין, כי שימוש מעשי במודלים ARMA השומרים על התנודות מסדר כלשהו לא נידון בספרות די הצורך. כך, לא נדונו עד כה סוגיות של שימוש במודלים ARMA מסדר גבוה, התאמת מודל השומר על תנודות לסדרה ספציפית, וניתוח של תוצאות המתקבלות בסדרות חברתיות-כלכליות תוך שמירה על טיב אומדני ניכוי עונתיות ומזעור עדכונים בסדרה. הנושא האחרון רלוונטי במיוחד לסדרות ישראליות, המושפעות מקיום חריגים כתוצאה ממשברים כלכליים ואירועים ביטחוניים, וחשופות לאי-סדירות גבוהה. בפרט, שימוש במודלים סטוכסטיים בנתוני חשבונות לאומיים לא נדון בספרות, ועד כה בלשכות סטטיסטיות נהוג להשתמש בסדרות אלו בשיטות של דנטון או בקירוב אליהם.

בעבודה הנוכחית, דיון מיוחד הוקדש לבעיית החיוץ, תוך הצעת מדדים לטיב החיוץ. נושא זה הינו בעל חשיבות רבה בסטטיסטיקה הרשמית עקב הדרישה למזעור היקף העדכונים בסדרות עם קבלת הנתון העדכני בתדירות נמוכה. לבסוף, נקבעו קריטריונים לבחירת המודל הטוב ביותר לביצוע ביון וחיוץ בסדרות עתיות.

סעיף 2 מתאר את שיטות נומריות של דנטון ואת שיטת הרגרסיה לפתרון בעיות ביון וחיוץ. סעיף 3 מציג סוגיות מתודולוגיות בנושאי ביון וחיוץ, תוך שימת דגש על הקשר בין תהליך זה לבין תהליכים אחרים אותם עוברים הסדרות הנאמדות, ובפרט תהליך אמידת השפעת החגים העבריים וימי הפעילות הנהוגים במשק הישראלי, ניכוי עונתיות ואמידת המגמה. מוצע אלגוריתם לאמידת מודל ARMA מנתונים מקוריים, ולאמידה של סדרות שנתיות ורבעוניות תוך קביעת מודל לחותך הלוקח בחשבון את ההנחות הכלכליות המקובלות. סעיף 4 מתאר את הנתונים שהתקבלו לבדיקה, שלבי המחקר, והמדדים המוצעים לבדיקת טיב האומדנים המתקבלים. סעיף 5 מציג את הממצאים העיקריים. בסעיף 6 מובאות המסקנות.

2. שיטות לביצוע ביון וחיוץ

קיים מספר רב של שיטות נומריות וסטוכסטיות לביצוע ביון וחיוץ (ראה Dagum and Cholette, 2006). בעבודה זו, נתמקד בשתי השיטות הנפוצות ביותר: שיטות נומריות של דנטון ושיטת רגרסיה.

2.1 שיטות של דנטון

הרעיון המקורי של דנטון (1971) היה לפתח שיטה לביון וחיוץ על סמך סדרת האינדקטור, כך ש:

1. תובטח שמירה מקסימאלית על התנודות בטווח קצר בסדרה המקורית בעלת תדירות גבוהה (סדרת האינדקטור),

2. תובטח התאמה מדויקת בין הסדרה הנאמדת לבין הנתונים בתדירות נמוכה, ולא יישארו סתירות בין מקורות מידע שונים לאחר ביצוע תהליך.

נגדיר: מטריצה L - מטריצה המגדירה את מגבלת ההתאמה (הקשר בין סדרת האינדקטורים $\{Y_t\}$ לבין הנתונים בעלי תדירות נמוכה $\{Z_m\}$), ומטריצה A המגדירה את פונקציית ההפסד. אזי, ניתן לנסח את בעיית ה-benchmarking לפי השיטה של דנטון באופן הבא (בכתיבה מטריציונית):

$$\begin{aligned} \min_{\eta} \{ (\eta - Y)' A (\eta - Y) \\ \text{s.t. } Z = L\eta \end{aligned} \quad (\text{א1})$$

כאשר η מסמן את הסדרה הנאמדת. באופן כללי, המגבלה ניתנת לכתיבה:

$$Z_m = \frac{1}{p_m} \sum_{t \in m} \eta_t \quad (\text{ב1})$$

כאשר m מגדיר תקופת הכיסוי של הנתון בתדירות נמוכה (מספר חודשים או רבעים), ו- p_m - אופן המגבלה (עבור דרישת ההסתכמות, $p_m = 1, \forall m$; עבור סדרות המדד, בהן נדרשת התאמה בין הממוצע של אומדנים בתדירות גבוהה לבין הנתון בתדירות נמוכה, $p_m = \frac{1}{m}$). הורת המגבלה מגדירה את מבנה של מטריצה L . הגדרת המבנה של מטריצה A נדונה בהמשך הסעיף, בהתאם לקביעת פונקציית מטרה של דנטון. הסדרה המותאמת $\{\hat{\eta}_t\}$ מתקבלת על ידי מינימיזציה של פונקציית הלגרנזיאן $L(\eta, \lambda)$:

$$\min_{\eta} L(\eta, \lambda) = \min_{\eta} \{ (\eta - Y)' A (\eta - Y) + \lambda' (Z - L\eta) \} \quad (\text{א2})$$

כאשר λ הנו וקטור של כופלי לגרנז'. הפתרון של (1) ניתן על ידי:

$$\hat{\eta} = Y + A^{-1} L' (L A^{-1} L')^{-1} (Z - LY) \quad (\text{ב2})$$

ניתן לחשוב על צורות שונות של פונקציית הפסד של השיטה הדנטונית, כמו הגישות המוצעות ב-Denton (1971), Fernandez (1981), Helfand, Monsour and Trager (1977), Cholette (1979, 1984). כמו כן ניתן להרחיב את השיטה הדנטונית להתאמת מספר סדרות בבת אחת תוך שמירת הקשרים ביניהם, כמו שהוצע על ידי Di Fonzi and Marini (2003).

במסגרת הבדיקה הנוכחית אנו נסקור את השיטות של דנטון השימושיות בנתוני חשבונות לאומיים (הצורה הכללית של המגבלה הינה כמו בנוסחה 1 לעיל):

א) השיטה הפרופורציונאלית: הרעיון של השיטה זו הוא לשמור על הפרופורציה בין שני נתונים עוקבים בסדרת האינדיקטורים:

$$\frac{Y_t - \eta_t}{\eta_t} - \frac{Y_{t-1} - \eta_{t-1}}{\eta_{t-1}} \equiv \frac{Y_t}{\eta_t} - \frac{Y_{t-1}}{\eta_{t-1}}, \quad t \in m, \quad m = 1, \dots, M, \quad t = 1, \dots, T$$

על כן, פונקצית ההפסד עבור הפרשים הראשוניים במקרה זה תראה:

$$\min_{\eta} \sum_{t=2}^T \left[\left(\frac{Y_t}{\eta_t} \right) - \left(\frac{Y_{t-1}}{\eta_{t-1}} \right) \right]^2 \quad (3)$$

בעבודה הנוכחית, נשתמש בתוצאה של Di Fonzi and Marini (2003) לפתרון בעיה (3) עבור הסדרה הבודדת. בשיטה הפרופורציונאלית של דנטון, מטריצת ה- A מ- (1) ניתנת על ידי $A = YDD^t Y^t$, כאשר:

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}, \quad D, \text{ ממימד מתאים.}$$

ב) שמירה על שיעור גידול: בשיטה זו נשמור על שיעורי הגידול בסדרת האינדיקטור:

$$\frac{Y_t - Y_{t-1}}{Y_{t-1}} - \frac{\eta_t - \eta_{t-1}}{\eta_{t-1}} \equiv \frac{Y_t}{Y_{t-1}} - \frac{\eta_t}{\eta_{t-1}}, \quad t \in m, \quad m = 1, \dots, M, \quad t = 1, \dots, T$$

לכן פונקצית ההפסד עבור הפרשים הראשוניים תראה:

$$\min_{\eta} \sum_{t=2}^T \left[\frac{Y_t}{Y_{t-1}} - \frac{\eta_t}{\eta_{t-1}} \right]^2 \quad (4)$$

ניתן לשנות נוסחאות לעיל כך שפונקציות ההפסד יתחשבו בתלות של התצפית t בתצפיות בזמנים $t-1$ ו- $t-2$:

2. על ידי כך נקבל שיטות דנטון עבור הפרשים שניים. במקרה זה נקבל פונקצית הפסד מורכבת יותר. שיטות אלה נקראות "נומריאליות" היות והן לא מניחות שום מודל סטטיסטי לנתונים או לשגיאות. למעשה, מדובר משפחת השיטות לא פרמטריות. השיטות של דנטון מספקות מסגרת פשוטה ונוחה לביצוע ביון וחיוץ בסדרות עתיות. אומדני דנטון הנם עמידים ולא תלויים בהנחות על מודל כלשהו עבור הסדרה הנאמדת או על התפלגות השאריות. הסיבות המרכזיות לשימוש כמעט בלעדי בשיטות אלו בנתוני חשבונות לאומיים הם: (1) סטיות תקן של סדרות האינדיקטור בדרך כלל אינן ידועות, (2) חשש מצורך בהנחות פרמטריות לגבי הסדרות שקשה לבדוק אותם, (3) חשש מהגדלת סטיית התקן של הסדרה הנאמדת, כאשר בדרך כלל אין משמעות לאמידה ופרסום סטיות תקן בנתוני חשבונות לאומיים. בנתונים אלה, שיטות אחרות לביון וחיוץ כמעט ולא נבדקו, ואפשרות מעבר לשיטות מתחרות לא נדון. עבור רוב החוקרים שימוש בשיטות דנטון לביון וחיוץ של נתוני חשבונות לאומיים מהווה הנחת היסוד המתחילה את המאמר (ראה Di Fonzi and Marini, 2003, Bloem, Dippelsman, and Maehle, 2001, ועוד). המחקר בלמי"ס בנושא זה בשנת 2005 יצא מנקודת הנחה דומה (ראה גובמן, 2005, גובמן ובורק, 2005).

עם זאת, השיטות האלו לא לוקחות בחשבון את התכונות הסטטיסטיות של הסדרה העתית, כמו מבנה הקורלציה וטעויות הדגימה. יתר על כן, שיטות של דנטון מניחות כי גם הסדרה בתדירות גבוהה (האינדיקטור), וגם הנתונים בתדירות נמוכה אינם חשופים לשונות. הנחה זאת בוודאי לא נכונה לגבי סדרות האינדיקטור. בנוסף, שיטות נומריות של דנטון מגבילות את המשתמש בקביעת פונקציות מטרה:

לא ניתן לקבוע פונקציה מטרה הלוקחת בחשבון הפרשים מסדר גבוה מ-2, כמו למשל הפרשים עונתיים (מסדר 4 בסדרות רבעוניות, מסדר 12 בסדרות חודשיות).

2.2 שיטת הרגרסיה

מודל רגרסיה, אשר הוצע על ידי (Cholette and Dagum (1991, 1994), כולל שתי משוואות:

$$\begin{aligned} Y_t &= c + \eta_t + e_t, \quad E(e_t) = 0, \quad E(e_t e_{t-k}) = \sigma_{e_t} \sigma_{e_{t-k}} \rho_k, \quad t = 1, \dots, T \\ Z_m &= \left(\sum_{t \in m} \eta_t \right) / p_m + w_m, \quad E(w_m) = 0, \quad E(w_m^2) = \sigma_{w_m}^2, \quad m = 1, \dots, M \end{aligned} \quad (5)$$

המשוואה הראשונה היא מודל רגרסיה על הנתון המקורי הנצפה Y_t , כאשר c קבוע (חותך), η_t הוא הפרמטר שיש לאמוד, ו- e_t הם טעויות אשר יכולות להתפרש כטעויות דגימה. הטעויות יכולות להיות הטרוסקדאסטיות, דהיינו, השונות $\sigma_{e_t}^2$ יכולות להשתנות עם הזמן. אוטוקורלציות ρ_k מתייחסות לתהליך סטציונרי הפיך ARMA המסופק על ידי המשתמש.

נוכיר, כי מודל ARMA לסדרה e_t ניתן לכתיבה:

$$\phi_r(B)e_t = \theta_l(B)a_t \quad (6)$$

כאשר a_t מסמן "רעש לבן" – סדרה בעלת תוחלת אפס ושונות σ_a^2 , B - אופרטור הזזה (אחורנית): $BY_t = Y_{t-1}$, r - סדר הפולינום האוטורגרסיבי ϕ_r , l - סדר הפולינום של ממוצעים נעים θ_l . הפולינומים האלה ניתנים לכתיבה:

$$\begin{aligned} \phi_r(B) &= 1 - p_1 B - p_2 B^2 - \dots - p_r B^r \\ \theta_l(B) &= 1 - q_1 B - q_2 B^2 - \dots - q_l B^l \end{aligned} \quad (7)$$

כאשר ϕ_i ו- θ_j הינם המקדמים של הפולינומים הנ"ל, $i = 1, \dots, p; j = 1, \dots, q$. מקובל לרשום מודל ARMA (6) באופן הבא: $(P \ Q) (p \ q)$, כאשר p מסמן סדר הפולינום האוטורגרסיבי (AR) לא עונתי, P - סדר הפולינום האוטורגרסיבי (AR) עונתי, q - סדר פולינום ממוצעים נעים (MA) לא עונתי, ו- Q - סדר פולינום ממוצעים נעים (MA) עונתי. s מסמן מספר העונות (בעבודה הנוכחית גודל זה שווה ל-4, היות והסדרות הינן רבעוניות).

המשוואה השנייה במודל של Cholette and Dagum לעיל מתארת את המגבלות המוטלות על η_t על ידי הסדרה בעלת תדירות נמוכה Z_m עם טעויות w_m , עבור נקודת זמן m . כמו במשוואה הראשונה, טעויות אלו יכולות להתפרש כטעויות דגימה במקרה של נתוני סקרים. במקרים בהם נדרשת התאמה מדויקת, נניח $w_m = 0, \forall m$ ונקבל התאמה מלאה של הסדרה הנאמדת לנתונים בתדירות נמוכה. p_m מסמן את מספר התצפיות של סדרה בעלת תדירות גבוהה המכוסות על ידי נתון Z_m , עבור סדרות מדדים; עבור סדרות מסוג מלאי $p_m \equiv 1, \forall m$. לדוגמה, עבור הנתונים הרבעוניים שממוצע שלהם אמורה להסתכם לנתון השנתי, $p_m = 1/4$, לכל נתון רבעוני השייך לשנה m .

אם נחליף מודל רגרסיה חיבורי במודל כפלי (לוג-חיבורי) במשוואה הראשונה, נקבל את המודל הבא (Cholette, 1994):

$$\ln Y_t = \ln(c_t \times \eta_t \times e_t) = \ln c + \ln \eta_t + \ln e_t \quad (8)$$

בהמשך העבודה, בכל מקום שמוזכר המודל הכפלי, הכוונה למודל המתואר בנוסחה האחרונה, ואין הבדל בין מודל כפלי ולוג-חיבורי. הפתרון מתקבל לפי שיטת הריבועים הפחותים המוכללים (Generalized Least Squares, GLS). במקרה החיבורי, האומדנים מתקבלים על ידי:

$$\hat{\eta} = Y + V_e L^T (LV_e L^T)^{-1} (Z - LY) \quad (9א)$$

והמטריצה של שונותיות משותפות $V_{\hat{\eta}}$ של האומדן $\hat{\eta}$ הינה:

$$V_{\hat{\eta}} = V_e - V_e L (LV_e L^T)^{-1} LV_e \quad (9ב)$$

כאשר X הינה מטריצת התכנון המתאימה, ו- V_e מסמן מטריצת השונותיות המשותפות של וקטור טעויות $\{e_t\}$. האומדים (9א) הם (Best Linear Unbiased Estimator) BLUE, ההוכחה ראה ב- Cholette and

Dagum, 1994. באותו המאמר מובאים פתרונות עבור המודל הכפלי (הלוג-חיבורי), ראה נוסחה (8). שיטת הרגרסיה היא השיטה הסטוכסטית הנפוצה ביותר בלשכות הסטטיסטיות, כאשר ניתן לקרב את השיטות של דנטון על ידי קביעה מתאימה של הפרמטרים במודל רגרסיה. נתבונן בנוסחאות (2ב) ו- (9א), ונשים לב כי ההבדל היחיד הוא בכך שבנוסחה (9א) מטריצה A מוחלפת במטריצת השונותיות המשותפות V_e . לכן, קביעה מתאימה של תהליך ARMA לשאריות, ושל סטיות התקן, ייתנו את הקירוב לשיטות של דנטון. יש להדגיש, כי פעולה זו מביאה לקבלת קירוב בלבד, וכן לא ניתן לטעון טענה הפוכה (אומדני דנטון הינם אומדני רגרסיה תחת איזה הנחות, ולכן אף הם BLUE, ראה דיון ב- (Chen and Wu, 2006)). ההנחות, שתחתן מודל הרגרסיה מספקת קירוב לאומדני דנטון לפי שיטות: "פרופורציונאלי" (3) ו"שמירה על קצב גידול" (4) הינן כדלקמן:

1. הנתון בתדירות נמוכה הינו "אמיתי" ($w_m = 0, \forall m$).

2. לטעויות e_t שונות קבועה:

▪ במונחי CV (סטיית התקן חלקי הנתון עצמו, כפול 100): $CV_t = CV, \forall t$, עבור והשיטה הפרופורציונאלית ושיטת השמירה על קצב הגידול;

3. מודל רגרסיה הנו:

▪ חיבורי ללא חותך עבור השיטה הפרופורציונאלית,

▪ כפלי ללא חותך עבור שיטת השמירה על קצב הגידול;

4. טעויות e_t מתנהגות לפי המודל של

▪ "הליכה מקרית" (מודל ARMA (1,0)(0,0) עם מקדם 0.999) עבור המודלים עם הפרשים ראשוניים,

▪ מודל ARMA (2,0)(0,0) עם מקדמים 0.99 (מסדר ראשון) ו- 1.98 (מסדר שני) המודלים עם הפרשים שניים.

הוכחה מפורטת של העובדה שתחת ההנחות (1) – (4) מתקבל קירוב של שיטות של דנטון הרלוונטיות נמצא ב- Cholette and Dagum, 2006. נזכיר כאן, כי נכון להיום משתמשים בלמ"ס בקירוב לשיטה הפרופורציונאלית של דנטון (עם הפרשים ראשוניים בנתוני חשבונות לאומיים). הקירוב נעשה על ידי קביעה המתאימה של הפרמטרים במודל הרגרסיה ובמודל ARMA לשאריות, כפי שמובא לעיל.

3. סוגיות מתודולוגיות - קביעת מודל רגרסיה ומודל ARMA לשאריות

בפרק זה, ננתח את השיטה הקיימת היום בלמ"ס, ונציע שיטות חלופיות לביצוע ביון וחיוץ, אשר עשויות לשפר את האמידה ללא צורך בקיום הנחות מגבילות או כאלו שלא ניתנות לבדוקה. כל השיטות הנדונות בפרק זה לא דורשות שדרוג התוכנה הקיימת, התקנת תוכנות חדשות או הכשרה נוספת לעובדי למ"ס.

3.1 קירוב השיטות של דנטון באמצעות קביעת מודל ARMA

כפי שצוין בסעיף 2.2, על ידי הקביעה המתאימה של הפרמטרים ניתן לקרב את השיטות של דנטון באמצעות מודל רגרסיה. ההוכחה המתמטית של עובדה זו חורגת ממסגרת העבודה הנוכחית. עם זאת, נביא את טיוטת ההוכחה (כאמור, הוכחה מלאה מובאת ב-Dagum and Cholette, 2006).

- ראשית, עבור מודל רגרסיה נתון (חיבורי או כפלי, בהתאם לשיטה של דנטון, ראה לעיל), רושמים ביטוי לשאריות e_t .

- שנית, רושמים את הדרישה הנובעת מפונקציית המטרה של השיטה הספציפית של דנטון (למשל, שמירה על אחוזי שינוי).

- שלישית, מראים, כי בנקודת האופטימום, בהינתן המודל ARMA של "הליכה מקרית", פונקציית המטרה של שיטת דנטון מגיעה, בקירוב, למינימום.

כך, הוכחנו כי תחת התנאים (1) – (4) של סעיף 2.2 הדרישות של אחת השיטות של דנטון מתמלאות, בנקודת האופטימום. לכן, כפי שמציינים המחברים, ניתן לקרב שיטות של דנטון על ידי קביעת פרמטרים מתאימה במודל הרגרסיה.

יש לציין, כי הוכחה הינה חד-כיוונית, כלומר ממנה לא נובעת שקילות בין השיטות. כמו כן ניתן להראות, כי מטריצת התכנון A בנוסחה (2) בשיטות של דנטון ומטריצת התכנון המתקבלת על ידי קביעת תהליך של "הליכה מקרית" לשאריות אינן זהות. בנוסף, קיימים הבדלים בין שיטות האמידה, כפי שהוסבר בפרק 2, ראה נוסחאות (2), (9א), (9ב). כתוצאה, פונקציות מטרה בשני המקרים אינן זהות, ונקודות הפתרון לא יהיו שוות. לכן, קיים צורך לחקור את ההבדל ביניהן. נציין, כי השוואה זו לא מופיעה בספרות.

קיימת משמעות מעשית לבדיקת טיב הקירוב. השיטות הנומטריות של דנטון קלות מאוד לתכנות, אפילו בגיליון EXCEL, והשימוש בהן נפוץ מאוד בלשכות הסטטיסטיות ברחבי העולם. בדיקת טיב הקירוב של שיטות נומריות תוך שימוש בשיטת הרגרסיה נחוצה לשם השוואה בין התוכנה BENCH לבין התוכנות המתחרות, המבצעות מזעור פונקציות מטרה של דנטון (כמו למשל תוכנת ECOTRM המומלצת על ידי EUROSTAT, ראה Barcellan and Buino, 2002). בנוסף, נוכל להסיק על כדאיות שדרוג מערכת ביון וחיוץ בסדרות שנתיות, בהינתן הנתונים הרב-שנתיים. נזכיר, כי נכון להיום ביון וחיוץ בנתונים אלה מתבצע לפי שיטות נומריות המתוכננות ב-EXCEL על ידי עובדי תחום חשבונות לאומיים.

3.2 קביעת מודל רגרסיה

הנחה (3) בסוף סוף סעיף 2.2 קובעת, כי אין לאמוד חותך על מנת לקרב שיטות של דנטון בעזרת שיטת רגרסיה. עם זאת, בנתוני חשבונות לאומיים, קיימים הבדלי רמות בין האינדיקטור לבין הסדרה הנאמדת. במקרים רבים, ניתוח השאריות הנאמדות $\hat{e}_t = Y_t - \eta_t$ מראה כי גודל השארית תלוי ברמת הסדרה הנאמדת, ולכן מודל עם חותך כפלי יכול להתאים. ההנחה (3) לעיל נראית כבלתי מוצדקת וסותרת את

מבנה הנתונים. בעבודה זו, מערכה השוואה בין מודלים חיבוריים לכפליים (לוג-חיבוריים), עם וללא חותך. בנוסף, שימוש במודל כפלי יכול להיות עקבי עם שימוש במודל כפלי לפירוק עונתי בתהליך ניכוי עונתיות, ברוב סדרות ה"מ"ס (ובכל הסדרות של חשבונות לאומיים). נזכיר כאן, כי תהליך ניכוי עונתיות מתבצע בל"מ"ס באמצעות התוכנה X-12-ARIMA (X12).

3.3 מודלים ARMA השומרים על תנודות

מודלים ARMA, בהם המקדמים בפולינום אוטורגרסיבי או/ו בפולינום של ממוצעים נעים בנוסחה (7) קרובים לדטרמיניסטיים, נקראים מודלים השומרים על התנודות (moving preservation models), ראה (Dagum and Cholette, 2006, Cholette and Chhab, 2004). מקדמים דטרמיניסטיים מסדר ראשון הינם המקדמים אשר קרובים ל-1 בערך מוחלט.

כך, לדוגמה, לפי ההוכחה שתיארנו בסעיף 3.1 לעיל, קביעת מודל של "הליכה מקרית" לשאריות - $(0\ 0)$ עם מקדם קרוב ל-1, גורמת לשמירה על אחוזי שינוי מנתון לנתון בסדרה הנאמדת, ביחס לסדרת האינדיקטור. באופן דומה, תהליך ARMA $(0\ 0)$ עם מקדמים: $p_1=1.98$, $p_2=-0.999$ שומר על התנודות מסדר ראשון ושני בסדרה הנאמדת, ביחס לסדרת האינדיקטור. לכן, המודלים הנ"ל (שהם מודלים של "הליכה מקרית" מסדר ראשון ושני, בהתאמה) הינם מקרים פרטיים של מודלים השומרים על תנודות.

על ידי קביעת מודל ARMA לשאריות הרגרסיה לביון וחיוץ נקבעת מטריצת השונויות המשותפות של השאריות הנ"ל. בתהליך האמידה באמצעות שיטת GLS, שימוש במטריצה הנוצרת מתהליך ARMA השומר על תנודות יגרום לשמירת התנודות בטווח המוגדר על ידי הפרמטרים של מודל ARMA בסדרה הנאמדת, ביחס לסדרת האינדיקטור. נשים לב, כי בניגוד לשיטות של דנטון, ניתן לשמור בשיטה זו תנודות בטווח בינוני וארך, כמו למשל תנודות עונתיות. מדובר בהרחבה של שיטות נומריות של דנטון למקרה כללי כאשר לא תמיד ניתן לרשום את פונקציית המטרה בצורה סגורה של תבנית ריבועית. כך, מודל $(1\ 0)$ עם מקדמים 0.99 ו-0.99 מגדיר פונקציית מטרה השומרת על תנודות בטווח קצר ועל תנודות עונתיות בסדרה נאמדת, ביחס לסדרת האינדיקטור. באופן דומה, ניתן להשתמש במודלים עם מקדמי MA (נוסחה 7) הקרובים לדטרמיניסטיים. לכן, שימוש בתוכנת BENCH המבוססת על מודל רגרסיה, מאפשר לקבוע מגוון רחב של פונקציות מטרה הלוקחות בחשבון מבנה תלות מורכב יותר מאשר המודל של "הליכה מקרית".

אחד היתרונות הבולטים של שיטות המבוססות על השימוש במודלים השומרים על התנודות הוא בכך שהם אינם דורשים הנחות פרמטריות כלשהן להתנהגות הסדרה או/ו שאריות. מודל ARMA קובעת, במקרה זה, את פונקציית מטרה לאופטימיזציה, ולא ההתפלגות. במובן זה, מדובר בשיטה פרמטרית למחצה (semi-parametric), כאשר מגדירים מודל רגרסיה (כפלי או חיבורי), אך לא מניחים הנחות מגבילות על התפלגות השאריות. יצוין, כי ישנן דוגמאות למודלים ARMA אחרות בעלות התכונה של שמירת תנודות בטווח קצר ותנודות העונתיות (ראה Dagum and Cholette, 2006, פרק 4.6). עם זאת, תכונות האומדנים המתקבלים תוך שימוש במודלים אלה בתהליך הביון והחיוץ לא נחקרו די הצורך בספרות, ושימוש בהם אינו נפוץ. נראה, כי קיימת חשיבות רבה לבדיקת התכונות של האומדנים המתקבלים, ובפרט טיב הביון, החיוץ, וניכוי העונתיות.

3.4 אמידת מודל ARMA לשאריות מהנתונים

כידוע מתיאוריה סטטיסטית, ידיעת ההתפלגות האמיתית של שאריות e_t במודל (5) אמורה להעלות את טיב האמידה: ידע זה אמור להביא להקטנת שונויות של אומדני המודל, ולהעלאת עוצמה ויעילות סטטיסטית של המבחנים למובהקותם (ראה, למשל, דיון ב- Cholette and Hillmer, 1990, Chhab and 2004). בעבודות אלו, מוצע להשתמש במבנה הקורלציה של סדרת האינדקטור לשם בניית מטריצה של השונויות המשותפות של שאריות e_t . אומנם, נתונים אלה אינם זמינים ברוב לשכות סטטיסטיות בעולם (בעיה זו נדונה בהרחבה ב- Cholette and Dagum, 1997, Chhab and 2004). בעבודה זו, מוצע לאמוד התפלגות השאריות מהנתונים הזמינים, ולבדוק טיב אומדני הרגרסיה לביון וחיוץ תוך שימוש בשיטה סטוכסטית של הרסיה. על ידי השוואת אומדנים אלה עם אומדני השיטה הפרופורציונאלית של דנטון ועם אומדנים המתקבלים תוך שימוש בתהליכים ARMA השומרים על התנודות נרצה לבדוק את הטענה הרווחת בספרות כי אין מקום לשימוש בשיטות סטוכסטיות בתהליך ביון וחיוץ של נתוני חשבונות לאומיים (ראה Barcellan and 2005, Buino 2002).

בלי הגבלת הכלליות נתייחס למודל חיבורי ללא חותך. במקרה הכפלי יש להחליף הפרש ליחס, ומקרה של מודל עם חותך יש להחסיר ממוצע מסדרת השאריות (לחלק בחותך הנאמד במקרה כפלי). נשים לב, כי סדרת השאריות $\{e_t\}$ מתקבלת כהפרש בין סדרת האינדקטור לבין הסדרה הלא ידועה שיש לאמוד אותה $Y_t - \eta_t = e_t$, תחת מגבלה של שורה שנייה בנוסחה (5). היות וסדרת האינדקטור מתארת תופעה כלכלית חברתית דומה לזו המתוארת על ידי הסדרה בתדירות נמוכה, ניתן להתייחס לסדרת השאריות כאל **טעות המדידה**. במילים אחרות, אילו ההגדרות, שיטות איסוף הנתונים ותהליך האמידה של סדרת האינדקטור היה תואם את ההגדרות של הנתונים בתדירות נמוכה המתארים את התופעה הנאמדת, ניתן היה לבצע ביון וחיוץ על ידי העתקת האינדקטור לרמה של הנתונים בתדירות נמוכה (עד כדי טעות מסוג של "רעש לבן" שאינו תלוי בזמן).

אם היינו יודעים את האומדן ל- η , היינו יכולים להתאים מודל ARMA לתהליך של סטיית האינדקטור מהסדרה הנאמדת $\{\hat{\eta}_t\}$. מודל זה יהיה אינפורמטיבי יותר ממודל של "הליכה מקרית", עליו מבוססים הקירובים לשיטות של דנטון. הדבר יהיה אפשרי אם סדרה עתית בתדירות נמוכה לא כוללת ערכים חריגים במיוחד, ואינדקטור שנבחר אכן מודד תופעה דומה לתופעה הנאמדת. השערה זו נבדקת בקלות על ידי השוואת אחוזי שינוי משנה לשנה בסדרה בתדירות נמוכה ובסדרת האינדקטור, במידה ונתונים בתדירות נמוכה הינם שנתיים (לפרטי השיטה ראה סעיף 4 להלן).

יצוין, כי קיימות מגבלות לשימוש בשיטה סטוכסטית זו. לא נוכל להתאים מודל ARMA לשאריות באופן מובהק במקרים הבאים: (1) התנהגות השאריות אכן קרובה ל"רעש לבן" או למודל של "הליכה מקרית"; (2) קיימים חריגים קיצוניים בנתונים בעלי תדירות נמוכה. במקרה (2), חוסר מובהקות במודל הנאמד עלול לגרום לאי-יציבות האומדנים המתקבלים. במקרים אלה, שימוש בשיטה הסטוכסטית המוצעת לא יעלה את טיב האומדנים המתקבלים.

נציע כאן אלגוריתם המשתמש באומדן של השיטה הפרופורציונאלית של דנטון לחישוב האומדן הראשוני לשאריות:

1. קבל אומדן ראשוני $\hat{\eta}_t^{(0)}$ לסדרה $\{\eta_t\}$ תוך שימוש בשיטה הפרופורציונאלית של דנטון (ראה הערות ל-1). נשתמש בנוסחת Di Fonzi and Marini (2003), ראה דיון בסעיף 2.1).
2. חשב שאריות $\hat{e}_t^{(0)} + c = Y_t - \hat{\eta}_t^{(0)}$, עבור המודל החיבורי לטעות, ו- $\hat{e}_t^{(0)} \times c = Y_t / \hat{\eta}_t^{(0)}$, עבור המודל הכפלי. נמרכז את השאריות המתקבלות על ידי החסרת הממוצע/חלוקה בממוצע שלהן. סדרת השאריות שהתקבלה הינה סטציונארית (ראה הוכחה עבור מקרה הכפלי שהוא הנפוץ בלמ"ס בהערה 2 להלן), ולכן ניתן לעבור ל-3.
3. התאם מודל ARMA לסדרת השאריות שהתקבלה בסעיף 2. לשם כך, ייבחר המודל הטוב ביותר, לפי קריטריון AIC (נוסחה (13) להלן), מדד SSE וקריטריון המובהקות של המודל ושל המקדמים הנאמדים (ראה הערה 3 להלן). ייבחנו מודלים בהם סדר פולינום אוטורגרסיבי לא-עונתי p אינו עולה על 2, וסדר פולינום אוטורגרסיבי עונתי P , סדר פולינום של ממוצעים נעים לא-עונתי q וסדר פולינום של ממוצעים נעים עונתי Q – אינם עולים על 1. נציין, כי שימוש במודלים מסדר גבוה יותר מוסיף קושי באמידה עקב הגדלת מספר הפרמטרים ולא משפר את טיב האומדנים המתקבלים (ראה, למשל, Findley et al., 2000).
4. נבצע ביון וחיוץ של סדרה עתית באמצעות שיטת הרגרסיה, תוך קביעת המודל ARMA הנאמד ב- (3) כמודל לשאריות הרגרסיה בנוסחה (5) בסעיף 2.2.

הערות

1. בחירת השיטה הפרופורציונאלית של דנטון נעשתה משיקולים הבאים: (א) השיטה הינה לא פרמטרית ולא דורשת הנחות כלשהן לגבי התנהגות הסדרה הנאמדת או/ו על התנהגות השאריות; (ב) תחת שיטה זו, מובטחת שמירה על התנודות בטווח קצר בסדרה הנאמדת ביחס לאינדיקטור, בהינתן המגבלה המוטלת על ידי הנתונים בתדירות נמוכה. לכן, במובנים הנ"ל, סדרת השאריות $\{\hat{e}_t^{(0)}\}$ תוכל לשמש כקירוב מספיק טוב לסדרת השאריות $\{e_t\}$.
2. נזכיר, כי התנאי להתאמת תהליך ARMA לסדרה עתית הוא שהסדרה הינה סטציונארית. לכן, לפני ביצוע תהליך ההתאמה ובחירת מודל לשאריות, יש לערוך מבחן סטטיסטי לסטציונאריות תוך שימוש בפרוצדורה PROC ARIMA ב-SAS, כפי שמתואר ב-Dickey (2005).
3. בסך הכל נבדקים 23 מודלים ARMA כאשר עם $p \leq 2, P, q, Q \leq 1$, כפי שמבואר לעיל. נסמן ב- L את לוג-הנראות של מודל ARMA, וב- m את מספר הפרמטרים הנאמדים. אזי, Akaike Information Criterion (AIC) ניתן לכתובה על ידי:

$$AIC = 2m - 2L \quad (13)$$
 מדד SSE הינו סכום הסטיות הריבועיות (Sum of Squared Error). מובהקות הפרמטרים של המודל נקבעה על סמך מבחן t . בעבודה זו, נבחרו מודלים בעלי מדדים AIC ו-SSE הנמוכים ביותר, מה שמצביע על התאמה טובה יותר של המודל. העדפנו מודל בו כל הפרמטרים הנאמדים הינם מובהקים.

3.5 ביון וחיוץ של סדרות שנתיות

ניתן להרחיב שימוש בשיטת הרגרסיה ולבצע בעזרתה תהליך ביון וחיוץ של סדרות שנתיות של חשבונות לאומיים. כאמור לעיל (ראה מבוא), תהליך זה טרם שודרג בסדרות השנתיות בלמ"ס. בעבודה זו הנחנו, כי הנתונים השנתיים ידועים והם גם ודאיים (במובן ששונותם שווה לאפס). עם זאת, בפועל סדרות אלו נאמדות מהנתונים. כפי שצינו לעיל, הנתונים השנתיים המגיעים מסקר תעשייה זמינים בפיגור של שנתיים, וקיים צורך בחיוץ הסדרות השנתיות בשנתיים האחרונות (ראה הסברים בתיאור הנתונים המקוריים, סעיף 4.1). במקרים אחרים, הנתונים השנתיים מפקדיים זמינים פעם במספר שנים (נתוני מרשם עסקים, לוח תשומה-תפוקה). אז, קיים צורך להשלים סדרה בין השנים האלו (בעיית ביון), ולאמוד את הסדרה השנתית מעבר לנתון המפקדי האחרון (בעיית חיוץ). לשם כך, נעשה שימוש בשיטה נומרית שפותחה בלמ"ס, אשר מניחה סטייה הולכת וגדלה עם הזמן בקצב גיאומטרי בין סדרת האינדיקטור לבין הנתון האמיתי הידוע. האלגוריתם למציאת הפתרון הינו איטרטיבי:

$$A_t' = A_t \times \left(\frac{A_B'}{A_B}\right) \wedge \left(\frac{t-b}{B-b}\right), b \leq t \leq B \quad (14)$$

כאשר A_t מסמן אומדן ראשוני (לפני התיקון) עבור שנה t , A_t' מסמן אומדן מתוקן עבור אותה השנה, B מסמן שנת ה-benchmark (שנה עבורה הנתון רב-שנתי הינו זמין) הקרוב ביותר לשנה t , ו- b שנת ה-benchmark הקודם. ההנחה על סטייה גיאומטרית נובעת מניתוחים כלכליים שנעשו בתחום חשבונות לאומיים אודות התנהגות הסדרות השנתיות. האמידה מתבצעת באמצעות שיטה נומרית הקרובה לשיטת דנטון, תוך שימוש באלגוריתם איטרטיבי המתוכנן בגיליונות EXCEL. נשים לב, כי אותה סדרת האינדיקטור משמשת לביון וחיוץ של סדרה שנתית, ולביון וחיוץ רבעוני. לאור זאת, ניתן להציע מודל לאמידת הסדרה השנתית והסדרה הרבעונית בו זמנית. נתייחס למקרה הכפלי של מודל (5), עבור הסדרה מסוג "מלאי" (דהיינו, כאשר הנתונים בתדירות גבוהה אמורים להסתכם לנתון המתאים בתדירות נמוכה). נרשום משוואות במודל רגרסיה (5) בצורה מטריציונית:

$$\begin{aligned} Y &= c \times \eta \times e \\ Z &= J \times \eta \end{aligned} \quad (15)$$

כאשר J מסמן מטריצת התכנון. מבנה מטריצת ה- J הנובע מדרישת ההסתכמות המדויקת:

$$J = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 1 & 1 & 1 \end{bmatrix} \quad (16)$$

כאשר מספר השורות שווה למספר שנים זמינות (מימד של וקטור Z), ומספר עמודות שווה למספר רבעים זמינים (מימד של וקטור Y). היות והחותך מתאר סטייה בין הנתון האמיתי לאינדיקטור, הנחת הסטייה הגיאומטרית קובעת צורת החותך c . מהנוסחה האיטרטיבית לחישוב האומדנים השנתיים לעיל נובע, כי מספיק לקבוע את A_t' . לצורך כך, ניתן להשתמש באומדן של דנטון לפי שלב 1 של האלגוריתם 3.3 (בניית מטריצת התכנון המתאימה ראה בהמשך). חסרון של שיטה זו בכך שהיא מצריכה הרצה נוספת של אלגוריתם לביון וחיוץ. לחילופין, ניתן להעביר את סדרת האינדיקטור לסדרה שנתית, תוך הכפלתה בקבוע K המתאים רמת האינדיקטור השנתי לרמת הסדרה הרב-שנתית של benchmarks. מקדם זה יכול להתקבל באופן הבא:

$$K = \frac{1}{N_b} \sum_{l=1}^{N_b} \frac{A_l}{Y_l} \quad (17)$$

כאשר הסכימה נעשית על פני השנים עבורם זמין הנתון הרב-שנתי האמיתי A_l , Y_l מסמן אינדיקטור שנתי ו- N_b מסמן מספר השנים הנ"ל. לכן, הנוסחה לחותך תהיה:

$$c^* = \alpha \times \left[\left(K \times \frac{Y_b}{A_b} \right) \wedge \left(\frac{t-b}{B-b} \right) \right], b \leq t \leq B \quad (18)$$

כאשר α מסמן חותך במודל רגרסיה, ו- Y_b - אינדיקטור שנתי בשנה B . קיבלנו מודל רגרסיה עם חותך כפלי שאינו קבוע. הנוסחה מבטיחה, כי בשנות ה- benchmark ערך משתנה החותך יהיה 1 (היעדר סטייה). ניתן להציע שתי מודלים חלופיים לפתרון הבעיה:

1. אמידה של סדרה **שנתית** על סמך הנתונים הרב שנתיים:

$$\begin{aligned} Y^* &= c^* \times \eta^* \times e^* \\ Z^* &= J^* \times \eta^* \end{aligned} \quad (19)$$

כאשר Y^* , Z^* הינן הסדרות של אינדיקטורים שנתיים ונתונים המתקבלים בתדירות נמוכה, בהתאמה. נזכיר, כי הסדרה Z^* התקבלה על ידי השמטה של השנים עבורם הנתון השנתי מהווה הערכה לפי השיטה הנומריית הקיימת היום בתחום של חשבונות לאומיים. מטריצה לאמידת הסדרה השנתית באופן ישיר J^* תהיה בעלת N_b שורות ומספר העמודות השווה למימד הוקטור של אינדיקטורים שנתיים Y^* . במקומות המתאימים לשנים עבורם הנתון הרב-שנתי זמין, נציב ערך 1, ובשאר המקומות - 0. היות וכל השורות במטריצה J^* הינן בלתי תלויות, מטריצה $(J^*)^T \times J^*$ הינה רגולארית, ממימד $N_b \times N_b$, מה שמבטיח קיום פתרון בשיטת הריבועים הפחותים המוכללים (GLS).

ברור, כי במקרה זה מודל ARMA לשאריות e_t^* עם הפרשים עונתיים אינו שימושי, היות וסדרה שנתיית אינה עונתית לפי ההגדרה. מודל סטוכסטי מסעף 3.3 לא מתאים במקרה זה בגלל מספר קטן של תצפיות שנתיות. המודל של "הליכה מקרית" מסדר שני עלול לעוות נקודת מפנה במחוז-מגמה, במיוחד בתקופת החיוץ (למשל, כניסה למיתון). הדבר נובע מכך שבתקופת החיוץ המודל מניח היצמדות למגמה הקיימת בתקופת הביון, לפחות בשתי התצפיות המחויצות הראשונות. לכן, עדיף במקרה זה לבחור במודל ARMA (0 0) (1 0) עם מקדם $p=0.999$. לאחר מכן, נבחר מודל ARMA השומר על תנודות (ראה מדדים לבחירת מודל בסעיף 4.1), או מודל ARMA הנאמד משאריות (ראה אלגוריתם בסעיף 3.3). בעזרת מודל זה, ובהינתן סדרה Z שנאמדה לעיל:

$$Z = \hat{\eta}^* \quad (5)$$

2. אמידה של סדרה **רבעונית** תוך שימוש באינדיקטור רבעוני Y וסדרה של נתונים רב-שנתיים Z^* . במקרה זה, מטריצה J^{**} הקובעת את המגבלה הרב-שנתית תתקבל על ידי השמטת ממטריצה J (16) כל השורות המתייחסות לשנים שעבורם נתון רב-שנתי אינו זמין. כמו בסעיף 1, אי-תלות של השורות במטריצה זו מבטיח קיום הפתרון בשיטת ה- GLS. המודל לחותך תהיה:

$$c_{t_q}^{**} = \alpha \times \left[\left(K \times \frac{Y_B}{A_B} \right) \wedge \left(\frac{t_q - b_q}{B_q - b_q} \right) \right], b_q \leq t_q \leq B_q \quad (20)$$

במקרה זה, ישנה ערך בכל רבעון המסומן t_q . ערכי B_q ו- b_q יהיה הרבע הראשון של שנת ה-benchmark הנוכחי, והרבע האחרון של שנת ה-benchmark הקודם, בהתאמה. בשנות ה-benchmark ערך משתנה החותך יהיה 1 (היעדר סטייה גיאומטרית ברבעים השייכים לשנות ה-benchmark). המודל ניתן לכתיבה על ידי:

$$\begin{aligned} Y &= c^{**} \times \eta \times e \\ Z^* &= J^{**} \times \eta \end{aligned} \quad (21)$$

מודל ARMA לשאריות e_t תיקבע על פי קריטריונים לבחירת מודל מסעיף 4.3.2 להלן, תוך השוואה בין המודל השומר על תנודות הטוב ביותר, לבין המודל הסטוכסטי (ראה סעיף 4.4 לעיל). הסדרה השנתית תתקבל על ידי הפעלת מטריצת התכנון (16) על הסדרה הרבעונית הנאמדת:

$$\hat{Z} = J \times \hat{\eta}$$

הערה: אמידת K באמצעות נוסחה (17) והכנסתו למודל אינה הכרחית, שכן ניתן לאמוד חותך זה כחלק מאמידת החותך במודלים של גרסיה (19) ו-(21). בשני המקרים לעיל קיבלנו מודל גרסיה עם חותך לא קבוע. בין התוכנות הקיימות היום בלשכות הסטטיסטיות, רק התוכנה (GBS) Generalized Benchmarking System נותנת מענה לבעיה זו ומספקת פתרון ללא צורך בתכנות נוספת. התוכנה פותחה בלשכה הסטטיסטית הקנדית בשנת 2004 וכוללת, בנוסף לכל האפשרויות הכלולות בתוכנה BENCH, מספר רב של אפשרויות נוספות (ראה Cholette and Chhab, 2004, ראה גם דו"ח פנימי – גובמן, 2007). התוכנה מופצת חינם, ומופעלת בסביבת SAS.

3.6 benchmarking בסדרות סכ"א

עבור 4 הסדרות העיקריות של סכ"א – מעוסקים ובלתי-מועסקים, לפי מין, מופקים אומדנים חודשיים מהסקר הרבעוני. קיים צורך להתאים אומדנים אלה לנתונים רבעוניים, אשר מדויקים הרבה יותר, ונאמדים על סמך מדגם רבעוני מייצג. לשם כך, פותחה והוטמעה מערכת ל-benchmarking בסדרות הני"ל, כפי שמפורט בהרחבה ב- גובמן ובורק, 2005. נזכיר, כי במקרה של אומדנים חודשיים של סכ"א מדובר בצורך לשפר אומדנים חודשיים תוך שימוש באומדנים רבעוניים. נשים לב, כי בעיה זו שקולה לבעיית ביון וחיוץ של סדרה עתית. במקרה זה, הסדרה בתדירות נמוכה הינה רבעונית מסוג מדד (כלומר, נדרשת התאמה מדויקת של הממוצע לנתון הרבעוני), התדירות הרצויה הינה חודשית, והאומדנים החודשיים של סכ"א מהווים אינדיקטור לתופעה הנאמדת. לכן, השיטות לפתרון הינן זהות. באמידת סדרות סכ"א חודשיות, היום משתמשים בתוכנת BENCH תוך קביעת מודל חיבורי של גרסיה ותהליך "הליכה מקריית" מסדר שני לשאריות (כלומר, קירוב לשיטה הפרופורציונאלית של דנטון עם הפרשים שניים). היות והתנהגות של הסדרות הני"ל משתנה מסדרה לסדרה, שימוש באותו מודל בכל הסדרות נראה בלתי סביר, בדומה לשימוש באותו המודל של "הליכה מקריית" מסדר ראשון לכל הסדרות של חשבונות לאומיים. כך, הסדרה "נשים בלתי מועסקות" הינה בעלת עונתיות חזקה ויציבה, בעוד שבסדרה "גברים מועסקים" השפעת המרכיב העונתי הינה חלשה, וגורמי העונתיות קרובים ל-100.

שימוש במודלים עונתיים של ARMA (נאמדים משאריות או מודלים השומרים על תנודות) עשוי לשפר את טיב האמידה ולמזער עדכונים עוד יותר. יש לציין, כי הלמ"ס מפרסמת מידי חודש אומדני המגמה לסדרות הנגזרות מ- 4 הסדרות הנ"ל (אחוז בלתי מועסקים, אחוז שייכים לכוח העבודה ועוד), כאשר המגמות מחושבות לאחר ביצוע ביון וחיוך וניכוי עונתיים. לכן, מזעור העדכונים בתהליך החיוך ישפיע ישירות על יציבות הסדרות המתפרסמות, וכתוצאה מכך יעלה את טיבם ומהימנותן בעיני המשתמשים.

4 נתונים מקוריים ודיאגנוסטיקות לטיב האמידה

4.1 תאור נתונים

לשם הבדיקה הנוכחית התקבלו 10 סדרות של חשבונות לאומיים:

מספר	שם סדרה	שם האינדיקטור	תקופת העיבוד	ניכוי עונתיות
1	תוצר בענף התקשורת	נתוני מע"מ בענף התקשורת	החל מ- 1.1998	הרצה X12
2	תוצר בענף המסחר	נתוני מע"מ בענף המסחר	החל מ- 1.1998	הרצה X12
3	תוצר בענף המחשבים	נתוני מע"מ בענף המחשבים	החל מ- 1.1998	הרצה X12
4	תמורה למשרות בתעשייה	מדד עלות עבודה (TAL50)	החל מ- 1.1995	חלוקה בגורמים רבעוניים
5	שעות עבודה למעשה בתעשייה	מדד שעות עבודה למעשה (THP50)	החל מ- 1.1995	חלוקה בגורמים רבעוניים
6	מספר משרות בתעשייה	מדד משרות (TMB50)	החל מ- 1.1995	חלוקה בגורמים רבעוניים
7	מספר משרות לשכיר בתעשייה	מדד משרות שכיר בתעשייה (TMG50)	החל מ- 1.1995	חלוקה בגורמים רבעוניים
8	תוצר תעשייתי	מדד ייצור תעשייתי (TPR50)	החל מ- 1.1995	חלוקה בגורמים רבעוניים
9	פדיון תעשייתי	מדד פדיון במחירים שוטפים (TRV50)	החל מ- 1.2003	חלוקה בגורמים רבעוניים
10	תוצר תעשייתי הענף כרייה וחציבה	מדד ייצור תעשייתי בענף כרייה וחציבה	החל מ- 1.1995	חלוקה בגורמים רבעוניים

בסדרות (5) ו- (6), ממוצע הנתונים הרבעוניים הנאמדים אמור להיות שווה לנתון השנתי, ובשאר הסדרות הנתונים הרבעוניים אמורים להסתכם לנתון השנתי הידוע.

יצוין כי הרצות של ביון וחיזוף של הסדרות (1) - (3) מתבצעות בתחום חשבונות לאומיים, כאשר סדרות מע"מ משמשות כאינדיקטורים. ניכוי עונתיות בסדרות אלו מתבצע על ידי הרצת התוכנה X12. הסדרות (4) - (10) נאמדות על סמך מדדי תעשייה המופקים באגף עסקים-כלכלה, כאשר הנתונים השנתיים הינם אומדני סקר תעשייה. מדדי תעשייה מתפרסמים בלמ"ס בתדירות חודשית, ולכן, על מנת לקבל סדרת אינדיקטורים רבעוניים מחושבים ממוצעים תלת-חודשיים. יש להדגיש, כי אומדני סקר תעשייה זמינים בפיגור של שנתיים, והאומדנים השנתיים לשנים 2005 ו-2006 מהווים בעצמם אומדני חיזוף של הסדרות השנתיות הרלוונטיות, על סמך אותם מדדי תעשייה שמשמשים כאינדיקטורים בביון וחיזוף רבעוני.

סדרות (4) - (10) הנאמדות על סמך מדדי תעשייה אינם עוברים תהליך ניכוי עונתיות באופן ישיר. במקום זאת, גורמים כוללים מחושבים על סמך ההרצה של מדדים חודשיים עד סוף כל רבע, וסדרה מנוכת עונתיות מתקבלת על ידי החלקה של הסדרה המקורית המתקבלת מתהליך ביון וחיזוף בגורמים אלה. בסדרה (9) לא מתבצע תהליך ניכוי עונתיות, בגלל מספר קטן של תצפיות זמינות. אופן חישוב הגורמים הרבעוניים לניכוי עונתיות בסדרות (4) - (10), וניתוח מפורט של יתרונות השיטה הזו מופיע בפרסום טכני (בורק, סלמה, גובמן, 2006). סוף תקופת העיבוד הוא רבע רביעי של שנת 2006, בכל הסדרות.

בנוסף, התקבלו 4 סדרות של סכ"א (אומדנים חודשיים): "גברים מועסקים", "נשים מועסקות", "גברים בלתי מועסקים" ו"נשים בלתי מועסקות". ממוצע הנתונים החודשיים המותאמים בסדרות אלו אמור להיות שווה לנתון הרבעוני הזמין, לכל רבע. ניכוי עונתיות ואמידת המגמות בכל הסדרות החודשיות של סכ"א מתבצע על ידי הרצת X12 באופן ישיר. תקופת העיבוד לצורך העבודה הנוכחית הינו מ-1.1998 ועד 12.2007, לכל ארבעת הסדרות.

4.2 שלבי המחקר

בשלב מקדים, נבדקה ההתאמה בין הנתונים השנתיים אודות התופעה הנאמדת לבין האינדיקטור שנבחר על ידי היחידות הנושאות, עבור כל 10 הסדרות של חשבונות לאומיים ו-4 הסדרות של סכ"א.

בשלב ראשון, נערך case-study על סמך 3 סדרות של חשבונות לאומיים הנאמדות על בסיס אומדני מע"מ (סדרות (1) - (3)). נבדקו טיב הקירוב של השיטה הפרופורציונאלית של דנטון תוך שימוש בשיטת הרגרסיה וכדאיות אמידת החותך. נבדק שימוש במודלים ARMA השומרים על תנודות, תוך שימוש הן במודל חיבורי והן במודל כפלי. נבחרו המודלים הטובים ביותר, במובן מדדי טיב הביון והחיזוף (ראה סעיף 4.3.1 להלן), ועבורם בוצעו הרצות של התוכנה X12. על סמך מדדי טיב אמידת השפעת החגים וימי הפעילות, ניכוי עונתיות ואמידת המגמה (ראה סעיף 4.3.2 להלן), נקבעו מודלים השומרים על מבנה גורמי החג וימי הפעילות ועל מבנה העונתיות, תוך השגת קירבה מירבית בין סדרת האינדיקטור מנוכת העונתיות לבין הסדרה הנאמדת מנוכת העונתיות. בנוסף, חושבו מדדי הקירבה המקבילים עבור המגמות. לבדיקת יציבות האומדנים נעשו הרצות עוקבות של ביון וחיזוף כאשר הנתונים בתדירות נמוכה מסופקים עד השנים: 2000, 2001, 2002, 2003, 2004, 2005. על סמך הרצות אלו נבחן טיב החיזוף, בנפרד מטיב הביון, ונאמדו היקפי העדכונים הנדרשים תחת 3 המודלים הטובים ביותר. לכל סדרה, נבחר המודל המתאים ביותר במובן המדדים המתוארים בסעיף 4.3 להלן. יצוין, כי בעבודה הקודמת (גובמן, 2005), לא נבחנה יציבות הסדרות הנאמדות, ולא נבדקו מדדי טיב החיזוף. לבסוף, לשלושת הסדרות הותאם מודל ARMA לשאריות, והופקו מדדי טיב הביון, החיזוף וניכוי העונתיות בדומה למתואר לעיל.

בשלב השני, ניתוח מקביל נערך ל-7 הסדרות של חשבונות לאומיים הנאמדות על סמך מדדי תעשייה (הסדרות (4) - (10)). בבדיקת הסדרות הנ"ל נעזרנו בתוצאות שהתקבלו בשלב הראשון של המחקר. הרצות של ניכוי העונתיות בסדרות אלו לא בוצעו, היות והרצת X12 לא מתבצעת בסדרות הנ"ל, ומדדי X12 לא

מופקים (ראה סעיף 4.1). כתוצאה מך שניכוי עונתיות מתבצע על ידי חלוקה בגורמים הנאמדים על סמך סדרת האינדיקטור, השיטה שמספקת מדדי טיב בין וחיוץ הטובים ביותר, תספק אוטומטית גם מדדים הטובים ביותר לניכוי עונתיות (במובן המדדים המפורטים בסעיף 4.3.2).

היות וסטיית התקן של סדרות האינדיקטורים אינן זמינות בכל הסדרות המשמשות כאינדיקטורים לביון וחיוץ של נתוני חשבונות לאומיים, נקבע כי מקדם ההשתנות (Coefficient of Variation, CV), השווה לסטיית התקן של הנתון מחולקת לנתון עצמו, כפול 100, יהיה שווה לאחוז אחד, לכל הסדרות ולכל התצפיות. אז, סטיית התקן של הסדרה הנאמדת (לאחר העברתה למונחי CV) תהווה מדד לתוספת בסטיית התקן של הנתון כתוצאה מביצוע תהליך ביון וחיוץ, ונעדיף שיטה הממזערת תוספת זו. בנוסף, קביעת מקדם ההשתנות שווה ל- 1% נראית סבירה בסדרות שנבדקו, וביון וחיוץ של אומדני חשבונות לאומיים נעשה היום תחת ההנחה זו. הסרת הנחה, כאשר סטיות התקן הרלוונטיות יהיו זמינות בלמ"ס, לא תשנה מסקנות המחקר (ראה דיון בסוגיה זו בהקשר ל- Benchmarking באומדני סכ"א, גובמן ובורק, 2005).

נערכה בחירת המודל הטוב ביותר לביון וחיוץ בסדרות החודשיות של סכ"א, על סמך המדדים המסכמים של טיב הביון והחיוץ. במקרה זה, ניתן לבנות קירוב למקדם ההשתנות CV של הסדרה החודשית המשמשת כאינדיקטור. לכל רבע, סטיות התקן של 4 הסדרות שבניתוח הינן ידועות. בנקודת הפתרון, ממוצע תלת-חודשי של אומדנים חודשיים השייכים לרבע מסוים שווים לנתון הרבעוני. לכן, לכל חודש, ניתן לקרב סטיית התקן על ידי: $s_t \approx \sqrt{3} \times s_q$, כאשר s_t מסמן את סטיית התקן של הנתון החודשי, ו- s_q - של הנתון הרבעוני. חלוקת ה- s_t בנתון עצמו עבור חודש t נותנת מקדם ההשתנות CV לאותו החודש. בעבודה הקודמת בנושא (ראה גובמן ובורק, 2005), נחקרו סדרות של מקדמי ההשתנות עבור 4 הסדרות החודשיות של סכ"א. מהניתוח מתברר, כי מדד ה- CV המתקבל אינו משתנה בצורה משמעותית על פני זמן. לכן, קיימת הצדקה להשתמש בממוצע מדדי CV על פני תקופת העיבוד כמקדם ההשתנות לצורך ביון וחיוץ, לכל אחת מהסדרות. נקבעו ערכים הבאים למדד ה- CV:

שם סדרה	ממוצע CV על פני תקופת העיבוד, באחוזים
גברים מועסקים	1.08
נשים מועסקות	2.07
גברים בלתי מועסקים	7.32
נשים בלתי מועסקות	7.43

בעבודה הנוכחית, השתמשנו באומדנים הנ"ל לביצוע הביון והחיוץ, לשם הבטחת יכולת השוואה עם האומדנים המופקים בלמ"ס. למעשה, כדאי לבדוק הנחה על CV קבוע בסדרות החודשיות פעם נוספת. כל ההרצות של ביון וחיוץ תוך שימוש במודל רגרסיה בוצעו באמצעות התוכנה BENCH. הרצות של השיטה הפרופורציונאלית של דנטון בוצעו בתוכנת SAS, תוך שימוש בנוסחת Di Fonzi and Marini (2003). התאמת מודלים ARMA לשאריות בוצעו ב- SAS, תוך שימוש בתוכנה שנכתבה על ידי המחבר; ניתן לקבל תוכנה זו לפי דרישה.

4.3 דיאגנוסטיקות לבדיקת טיב ביון, חיוץ וניכוי עונתיות

4.3.1 דיאגנוסטיקה לבחינת האינדקטור

נציע שיטה גראפית פשוטה לבחינת דמיון של סדרת האינדקטור לתופעה הנאמדת. תחילה, נעביר את האינדקטור לאותה התדירות בה זמינים הנתונים בתדירות נמוכה. בסדרות של חשבונות לאומיים, בהן הנתונים בתדירות נמוכה זמינים מדי שנה, נעביר את האינדקטור לשנתי על ידי סכימה או מיצוע. בסדרות חודשיות של סכ"א, ניקח ממוצע תלת-חודשי של האומדנים החודשיים לפני התאמה. נחשב אחוזי שינוי מנתון לנתון בסדרה בתדירות נמוכה ובאינדקטור באותה התדירות. נבנה תרשים של אחוזי שינוי אלה בשתי הסדרות, ונבדוק את הקרבה ביניהן. ייתכנו מספר מצבים:

1. שתי הסדרות של אחוזי שינוי קרובות זו לזו, עד כדי סטייה לא שיטתית – במקרה זה נסיק כי האינדקטור עונה על דרישות המודל, וניתן להשתמש בו לביצוע ביון וחיוץ.
2. קיים הבדל בין הרמות של שתי הסדרות, אך אחוזי השינוי עצמם נראים דומים מאוד – במקרה זה נבחר במודל כפלי עם חותך, כאשר החותך "יספוג" הבדלי רמות.
3. קיימים הבדלים משמעותיים בין שתי הסדרות, הן ברמה והן במגמה (אין התאמה באחוזי שינוי). ביצוע ביון וחיוץ ללא התחשבות בממצא זה עלול להוריד את טיב האומדנים המתקבלים וליצור הבדלים בין מקורות מידע שונים. קיימות מספר אפשרויות להתמודד עם הבעיה: (א) במידה והדבר אפשרי מבחינה נושאת, יש להחליף אינדקטור; (ב) במידה והחלפת האינדקטור אינה אפשרית, יש למדל את הסטייה באמצעות בחירה נכונה של מודל ARMA לשאריות (בפרט, מודל ARMA סטוכסטי כמתואר בסעיף 3.4 לעיל) או באמצעות קביעת מודל לחותך (למשל, על ידי שימוש במודל החותך הלא קבוע כמתואר בסעיף 3.5 לעיל).

4.3.2 דיאגנוסטיקות לטיב ביון וחיוץ

על מנת להשוות מודלים שונים לביצוע benchmarking נשתמש בדיאגנוסטיקות הבאות:

1. סטיית התקן הממוצעת: ממוצע סטיות התקן של האומדנים $\hat{\eta}_t$ בטווח העיבוד:

$$Astd = \frac{1}{T} \sum_{t=1}^T std(\hat{\eta}_t)$$

2. מדד יחס לשמירת תנודות בטווח קצר Cm:

$$Cm = \frac{1}{T-1} \sum_{t=2}^T \left| \frac{\hat{\eta}_t / \hat{\eta}_{t-1}}{Y_t / Y_{t-1}} - 1 \right| = \frac{1}{T-1} \sum_{t=2}^T \left| \frac{\hat{\eta}_t}{Y_t} \times \frac{Y_{t-1}}{\hat{\eta}_{t-1}} - 1 \right|$$

3. מדד חיבורי לשמירת תנודות בטווח קצר Ca:

$$Ca = \frac{1}{T-1} \sum_{t=2}^T |(\eta_t - \eta_{t-1}) - (Y_t - Y_{t-1})|$$

4. הפרש, בערך מוחלט, בין מדד החלקלקות של הסדרה הנאמדת לבין מדד החלקלקות של האינדקטור.

5. מדד להיקף העדכונים בסדרה הנאמדת כתוצאה מהוספת נתון עדכני בתדירות נמוכה:

$$R = \sum_t |\hat{\eta}_t^{(m)} - \hat{\eta}_t^{(m-1)}|$$

כאשר $\{\hat{\eta}_t^{(m-1)}\}$ מסמן את הסדרה הנאמדת מהרצה בה הנתונים בתדירות נמוכה זמינים עד לתקופה $m-1$, ו- $\{\hat{\eta}_t^{(m)}\}$ מתקבל מהרצה עם הוספת הנתון העדכני בתדירות נמוכה עבור התקופה m .

נעדיף מודל כאשר הסדרה הנאמדת בעזרתו ממזערת את חמשת המדדים לעיל. נשים לב, כי המדדים (1) – (4) הינם נורמות המודדות מרחקים בין אחוזי שינוי של שתי הסדרות המשוות, ומדד (5) מודד מרחק בין שתי הסדרות המתקבלות משתי הרצות עוקבות. נציין כי המדד C_m הוא הרלוונטי בעבודה זו היות ומעוניינים לשמור על אחוזי שינוי מנתון לנתון, ולא על ההפרשים.

4.3.3 דיאגנוסטיקות לטיב ניכוי עונתיות

התוכנה X12 מייצרת מגוון רחב של דיאגנוסטיקות לטיב התאמת מודל רגרסיה לאמידת השפעת חגים וימי פעילות (אמידת גורמי התאמה מראש), התאמת מודל ARIMA להרחבת הסדרה המותאמת מראש על ידי תחזיות, ניכוי עונתיות ואמידת המגמה. אנו נשתמש במספר דיאגנוסטיקות עיקריות, כפי שמפורט להלן:

1. ערך מבחן חי בריבוע (Chi-Square) לבדיקת מובהקות מודל הרגרסיה לאמידת גורמי התאמה מראש (ערך גבוה יותר מצביע על התאמה טובה יותר של מודל הרגרסיה. בעבודה הנוכחית נשתמש במודל של שתי קבוצות רבעים לאמידת השפעת החגים וימי הפעילות (ראה פרסום אינטרנט "גורמים עונתיים וגורמי ההתאמה מראש לשנת 2007 והמגמות לשנים 2006-2001);
 2. ממוצע טעויות בערך מוחלט של תחזיות בתוך המדגם עבור שלוש שנים אחרונות המצביע על טיב התאמת המודל (ערך קטן יותר מצביע על התאמה טובה יותר);
 3. מבחן חי בריבוע (Chi-Square) המצביע על מובהקות מודל ARIMA (הסתברות על פי הערך הסטטיסטי של חי בריבוע גבוהה יותר מצביעה על מובהקות גבוהה יותר);
 4. מבחן F קיום עונתיות;
 5. סטיית תקן של הגורמים העונתיים וסטיית תקן של הגורמים הבלתי סדירים, כאשר אם מתקיים: $\frac{std_D10}{std_D13} > 1$, אזי התנודות העונתיות גוברות על התנודות הבלתי סדירות;
 6. התרומות היחסיות (בממוצע) של מרכיבי הסדרה לאחוז שינוי בין שני נתונים עוקבים בסדרה המקורית;
 7. מדד QCD המצביע על מספר רבעים הדרושים למגמה-מחזור על מנת להתגבר על אי-סדירות;
 8. מדד מסכם לטיב ניכוי עונתיות Q.
- בנוסף מחושבים שני המדדים הבאים:
9. מדד החלקלקות של הסדרה מנוכת העונתיות, כאשר ערך קטן יותר של המדד מצביע על טיב גבוה יותר של ניכוי עונתיות:

$$S = \frac{1}{T-1} \sum_{t=2}^T \left| \left(\frac{Y_t}{Y_{t-1}} - 1 \right) \times 100 \right|$$

כאשר Y_t הוא ערך סדרה בזמן t . מדד זה חושב לסדרות האינדקאטורים, סדרות לאחר ביון וחיוץ, סדרות מנוכות העונתיות והמגמות.

מדדים אלה מחושבים עבור הסדרות (1) – (3) בלבד. בסדרות (4) – (10) מדדים אלה לא מחושבים היות ואין צורך בהרצת X12 לקבלת סדרה מנוכת עונתיות. לכן, טיב מדדי ביון וחיוץ יקבע את טיב ניכוי עונתיות בסדרות אלו. ככל שהתנודות בסדרה הנאמדת קרובות יותר לתנודות בסדרת האינדיקטור, וככל שהיקף העדכונים קטן יותר, כך טיב ניכוי עונתיות יעלה, היות ויש הצדקה חזקה יותר להשתמש בגורמים הנאמדים מסדרת האינדיקטור.

נוכיר, כי אנו מעוניינים לשמור על תכונות הסטטיסטיות של הסדרה העתית המקורית במהלך ביצוע ביון וחיוץ. לכן, נעדיף מודל עבורו הדיאגנוסטיקות המתוארות לעיל הן הקרובות ביותר לאלו המתקבלות מהרצה של הסדרה המקורית.

5. ממצאים עיקריים

5.1 שלב מקדים – בחירת האינדיקטורים

לאור האמור לעיל (ראה מבוא, וכן סעיף 4.3.1 לעיל), יש לוודא כי סדרת האינדיקטור מתארת את תופעה חברתית-כלכלית דומה לזו המתוארת על ידי הסדרה בתדירות נמוכה. תרשימים 1 - 10 מתייחסים לסדרות של חשבונות לאומיים ומציגים אחוזי שינוי משנה לשנה בסדרות השנתיות (סדרות ה- Benchmarks) ובסדרות האינדיקטור המתאימות. ניתן להסיק, כי בסדרות (2), (4), (5), (6), (7), (8) ו- (9) תנאי בחירת האינדיקטור מתקיים, הוא אכן מתאר תופעה כלכלית דומה לזו המתוארת על ידי הסדרה השנתית. עם זאת ניתן לראות, כי בסדרה "תוצר בענף המסחר", קיים הבדל באחוזי שינוי שנתיים בין האינדיקטור לבין הסדרה בתדירות נמוכה בשנת 2001, בה התקבל נתון חריג במיוחד בסדרה השנתית.

בסדרות (1), (3) ו- (10) (ראה תרשימים 1, 3 ו- 10) קיימים הבדלים משמעותיים בין אחוזי שינוי באינדיקטור השנתי לבין הנתונים המתאימים בתדירות נמוכה. מהשיחות עם עובדי היחידות הנושאות הסקנו שלפחות בשלב זה יימשך השימוש באינדיקטורים הקיימים. ברור, כי במקרים אלה מופרת ההנחה הבסיסית של השיטה הקיימת, לפיה האינדיקטור מתאר תופעה חברתית-כלכלית דומה לזו המתוארת על ידי הסדרה בתדירות נמוכה ואנו מעוניינים להעתיק, עד כדי המגבלה של התאמה מלאה, את ההתפלגות של האינדיקטור אל הסדרה הנאמדת. הפרת ההנחה הנ"ל תגרום לכך שההערכת האוטומטית שכזו תגרום להורדת טיב האומדנים ולסתירות בין מקורות מידע שונים המתפרסמים על ידי הלמ"ס. בעבודה הנוכחית, נשתמש בשיטה לאמידת מודל ARMA סטוכסטי למידול ההטיה לפי האלגוריתם המוצע בסעיף 3.4 לעיל, ועל ידי מידול הסטייה נתגבר על בעיית האי-התאמה.

5.2 שלב א' – ניתוח 3 הסדרות הנאמדות על סמך אומדני מע"מ (case study)

נוכיר, כי שלב זה כולל ניתוח פרטני של הסדרות (3) - (1): "תוצר בענף תקשורת", "תוצר בענף מסחר", ו"תוצר בענף מחשבים".

5.2.1 בחירת מודל רגרסיה ומודל ARMA

לוח 1א מראה, כי שימוש בשיטה נומרית של דנטון לפי הנוסחה של Di Fonzi and Marini מביא לתוצאות הרבה פחות טובות במובן השמירה על התנודות בטווח קצר. לכן, הקירוב לשיטה של דנטון בתוכנה BENCH מהווה שיפור של השיטה הנומרית המקורית. מסקנה זו עולה בקנה אחד עם מסקנת המחקר של Chen and Wu (2006).

כמו כן, לוח 1א מראה כי המודל הנמצא כעת בשימוש הלמ"ס - מודל רגרסיה חיבורי ללא חותך, עם מודל ARMA לשאריות (0 0) (1 0), ועם $p1=0.999$ (מודל 1) - מספק מדדי טיב ביון וחיוץ המצביעים על שמירה

על התנודות בטווח קצר בסדרה הנאמדת. ניתן לראות, כי הוספת חותך למודל חיבורי פוגע בטיב האמידה (מדד הכפלי לשמירת אחוזי שינוי עולה), בעוד שמעבר למודל כפלי לא משפיע על המדדים העיקריים - מדד כפלי לשמירת אחוזי שינוי והפרש, בערך מוחלט, בין מדד החלקלקות של סדרת האינדיקטור לבין מדד החלקלקות של הסדרה הנאמדת. במודל הכפלי, לא קיים הבדל במדדי טיב האמידה בין המודלים עם וללא חותך. יש לציין, כי החותך מובהק בכל המודלים (חיבורי וכפלי). היות והחותך מובהק סטטיסטית בכל הסדרות שבניתוח, בהמשך נתייחס למודלים הכוללים אמידת חותך.

בנוסף ניתן לראות, כי מעבר למודל כפלי מוריד סטיית התקן הממוצעת של האומדנים בסדרה "תוצר בענף המסחר", בעוד שבשני הסדרות האחרות הדבר מעלה את סטיית התקן הממוצעת.

לוח 1ב מנתח תוצאות של שימוש במודלים ARMA שונים השומרים על התנודות. ניתן לראות, כי המודל הנמצא בשימוש בלמ"ס נמצא בין המודלים הטובים ביותר במובן מזעור המדד הכפלי לשמירת אחוזי שינוי ומזעור ההפרש בין מדדי החלקלקות של האינדיקטור ושל הסדרה הנאמדת. בכל הסדרות, שימוש במודלים חיבוריים מביא לגידול בערכי המדדים לטיב ביון וחיוץ, כלומר להורדת טיב התהליך. הדבר מצביע על יתרון של המודלים הכפליים בשמירת תנודות בטווח קצר, בהשוואה למודלים החיבוריים המקבילים. לכן, נעדיף שימוש במודלים כפליים עם חותך, ובהשמך הניתוח לא נתייחס למודלים גרסיה אחרים.

בשלב זה, מודלים אשר מספקים את התוצאות הטובות ביותר, במונחי מדדי טיב ביון וחיוץ, סומנו במספרים מ-1 ועד 9, כאשר מודל המהווה קירוב לשיטה הפרופורציונאלית של דנטון, אשר נמצא בשימוש הלמ"ס, סומן כמודל 1. מודלים אשר לא משפרים את טיב האמידה ביחס למודל 1 באף אחד מהמדדים לא מוספרו ולא נלקחו להמשך הניתוח, ולכן לא נתייחס אליהם בהמשך. עם זאת, המדדים העיקריים לטיב הביון והחיוץ מוצגים בלוחות 1א, 1ב ו-1ג עבור כל המודלים שנבדקו במסגרת עבודה זו.

מלוחות 1א ו-1ב ניתן לראות, כי ישנם מודלים המספקים מדדי טיב הביון והחיוץ העולים על אלה המתקבלים מהמודל הפשוט (0 0)(0 1). בכל שלושת הסדרות, במודלים 3, 4, 5, ו-7 לפחות אחד מהמדדים קטן יותר (ולכן טוב יותר) מהמדדים המקבילים של מודלים 1 ו-2. המדדים המתקבלים ממודלים 1, 2, ו-6 הינם דומים. נשים לב, כי לפי המדדים לשמירת התנודות בטווח קצר, המודלים הטובים ביותר לסדרות "תוצר בענף תקשורת" ו"תוצר בענף מסחר" הינם 3, 5 ו-7; לסדרה "תוצר בענף מחשבים" - מודלים 4 ו-5. יש לציין, כי סטיית התקן הממוצעת במודל 7 גבוהה משמעותית מאלו המתקבלות ממודלים האחרים.

בנוסף, נבחן מספר מודלים ARMA נוספים מסדר גבוה יותר (ראה לוח 1ג). מכל המודלים שנבחנו, המודלים 8 ו-9 מספקים המדדים המצביעים על השמירה על התנודות בטווח קצר ביחס לסדרת האינדיקטור. עם זאת, ניתן לראות כי שימוש במודלים 8 ו-9 אינו משפר את טיב הביון והחיוץ, בהשוואה למודלים הטובים ביותר מלוח 1ב. ניתן להסיק, כי שימוש במודלים ARMA מסדר גבוה אינו משפר את תוצאות האמידה, מה שעולה בקנה אחד עם התוצאות של המחקרים הקודמים (ראה דיון בסעיף 3.4 לעיל). בהמשך העבודה נתייחס למודלים 1, 2 (מלוח 1א) ו-3, 4, 5, 6 ו-7 (מלוח 1ב), ומודלים 8 ו-9 מלוח 1ג.

5.2.2 ביון, חיוץ וניכוי עונתיות

לוחות 2א, 2ב ו-2ג מציגים את הדיאגנוסטיקות העיקריות לטיב ניכוי עונתיות ב-3 הסדרות שבניתוח. בנוסף למדדים רגילים, מוצגים גם המדדים הכפליים לשמירת אחוזי שינוי בסדרה בתהליך ביון וחיוץ. בלוחות אלה, מדדי טיב הביון והחיוץ מחושבים גם לסדרות מנוכות עונתיות ולמגמות (נציין, כי לא בוצעו הרצות ביון וחיוץ לא לסדרות מנוכות עונתיות ולא למגמות).

עבור הסדרה "תוצר בענף תקשורת", במודלים 7 ו-9 מדדים C_m לסדרות מנוכות העונתיות הינם הגבוהים ביותר, ובמודלים 3, 7, 8, 9 מדדים אלה הם הגבוהים ביותר עבור המגמות. כמו כן, ניתן לראות כי

במודלים 6, 8 ו-9 חל שיבוש במבנה השפעת החגים וימי העילות ובמבנה עונתיות בתהליך ביון וחיוץ – חל שינוי משמעותי בערך סטטיסטי המבחן חי בריבוע למובהקות של הרגרסיה לאמידת גורמי התאמה מראש. ניתן להסיק, כי בסדרה זו, מודלים 1, 2, ו-5 שומרים בצורה הטובה ביותר על התנודות בטווח קצר, על מבנה השפעת החגים וימי הפעילות ועל מבנה עונתיות. תוצאות אלו מתקבלות מניתוח אחוזי שינוי מנתון לנתון ומדדי החלקלקות בסדרות מקוריות המתקבלות מתוכנה BENCH, סדרות מנוכות העונתיות והמגמות.

ניתוח דומה עבור הסדרה "תוצר בענף מסחר" (ראה לוחות 2ב) מראה כי המודלים הטובים ביותר לסדרה זו הינם 1, 5 ו-7, ועבור הסדרה "תוצר בענף מחשבים" (ראה לוחות 2ג) נבחר מודלים 1, 2 ו-4. ניתן להסיק, כי בסדרות שנבדקו, המודלים מסדר גבוה לא עונות על דרישות שהצבנו לפנינו אודות השמירה על התנודות בטווח קצר, על מבנה החגים וימי הפעילות, מבנה עונתיות ועל מדדי טיב ניכוי העונתיות. כמו כן, המודלים בהם סטיית התקן הממוצעת הינה גבוהה, אינם יציבים ומעוותים את מבנה השפעת החגים וימי הפעילות ומבנה עונתיות. יש לשים לב, כי תופעה זאת מתרחשת למרות שמדד הכפלי לשמירת אחוזי שינוי C_m נמוך מאוד (דהיינו, טוב מאוד במובן השמירה על התנודות), עבור הסדרות המקוריות המותאמות.

5.2.3 ניתוח עדכונים וטיב החיוץ

לוח 3 מציג הרצות ביון וחיוץ באמצעות שיטת הרגרסיה, עם מודלים 1 - 7, כאשר הסדרות הותאמו עם ובלי הנתון של שנת 2006 - הנתון השנתי האחרון אשר היה זמין בתקופת המחקר. חושבו העדכונים בסדרות הנובעים מהוספת הנתון השנתי של 2006: סך הכל עדכונים בסדרה ועדכונים לשלוש השנים האחרונות. נזכיר, כי אנו מעוניינים למזער את היקף העדכונים בסדרה עקב הוספת הנתון השנתי העדכני, על מנת למנוע שינויים דרמטיים בסדרות ועל מנת להבטיח עקביות ויציבות באומדני הלמ"ס, לרבות סדרות מנוכות עונתיות ומגמות. בנוסף, מזעור עדכונים יביא להעלאות טיב ומהימנות התחזיות של הגורמים העונתיים וגורמי החג וימי הפעילות המתפרסמות על ידי הלמ"ס מידי שנה (ראה פרסום אינטרנט הגורמים העונתיים וגורמי ההתאמה מראש לשנת 2007 והמגמות לשנים 2001-2006). כמו כן נשים לב, כי מזעור היקף העדכונים פירושו - העלאת טיב החיוץ.

נראה, כי מודל 5 ממזער את מספר והיקף העדכונים בסדרה "תוצר בענף התקשורת". יצוין, כי המודל הקיים (מודל 1), מספק מדדים המצביעים על מספר קטן יחסית של עדכונים בסדרה. עבור המודלים 1, 5, ו-7 נערך ניתוח יציבות פרטני, כאשר לכל אחד מהמודלים בוצעו הרצות עם הנתונים השנתיים עד שנת 2000, 2001, 2002, 2003, 2004, 2005 ו-2006. ניתן להסיק, כי מודל 5 מספק אומדנים מחויצים הטובים ביותר בין כל המודלים שנחקרו.

באותו אופן, ניתוח סדרה "תוצר בענף מסחר" (לוח 4) מראה כי מודל 5 הוא העדיף בסדרה זו, כאשר היקף העדכונים יורד, בהשוואה למודל 1, ביותר מפי 2. בסדרה "תוצר בענף ממחשבים" (לוח 5) מודל 4 הוא הטוב ביותר, עם היקף העדכונים קטן פי 5 ויותר, בהשוואה למודל 1. יש לציין כי גם השימוש במודל 2 בסדרה זו מביא לשיפור משמעותי בטיב בחיוץ ובהיקף העדכונים.

5.2.4 בחירת מודל ARMA לשאריות

בהתאם לאלגוריתם שהוצע בסעיף 3.4 לעיל, נאמד מודל ARMA לשאריות, לפי הקריטריונים שפורטו בהערה 3 בסעיף זה. תחילה, נבדקה השערת סטציונאריות, עבור כל הסדרות שבניתוח - הן בשלב א', הן בשלב ב' של המחקר האמפירי. לא ניתן לדחות השערת סטציונאריות, עבור כל הסדרות שנבחנו. המודלים שנבחרו הפרמטרים הנאמדים והדיאגנוסטיקות לטיב הביון, החיוץ וניכוי העונתיות תוך שימוש במודלים אלה מובאים בלוח 6. אנו מעוניינים להשוות את מודל זה עם המודלים השומרים על תנודות (ראה דיאגנוסטיקות בסעיף 5.2.3 לעיל). ניתן לראות, כי מדד הכפלי לשמירת התנודות בטווח קצר והפרש בערך מוחלט בין מדדי החלקלקות של האינדיקטור ושל סדרה הנאמדת דומים לאלה המתקבלים מהמודלים הטובים ביותר השומרים על תנודות בסדרות הנחקרות. בסדרות "תוצר בענף תקשורת" ו"תוצר בענף מסחר" סטיית התקן הממוצעת גדולה מזו המתקבלת במודלים המתחרים. עם זאת, מספר והיקף העדכונים בסדרה כתוצאה מהוספת הנתון העדכני של שנת 2006 אינו גדול במיוחד; בשלושת המקרים היקף העדכונים תחת מודל זה ותחת מודל 1 הינם דומים, כאשר קיים יתרון קל למודל הטוב ביותר עם מודל ARMA השומר על תנודות (מודל 5 בסדרות "תוצר בענף תקשורת" ו"תוצר בענף מסחר", ומודל 4 בסדרה "תוצר בענף מחשבים"). מלוח 6 ניתן לראות, כי טיב ניכוי העונתיות נשמר תחת המודלים הסטוכסטיים שנאמדו משאריות, בהשוואה למדדים המתקבלים מניכוי עונתיות בסדרת האינדיקטור. השוואת לוח 7 עם לוחות 3 - 5 ראה יתרון חשוב של שיטה סטוכסטית - יציבות של היקפי העדכונים על פני הרצות שונות, תוך קביעת תקופות עיבוד שונות. מלוחות אלה ניתן לראות, כי תחת השיטה הסטוכסטית לקביעת מודל ARMA, סדרות העדכונים בתקופת החיוץ (דהיינו, בשנה האחרונה) הינן יציבות יותר מאלו המתקבלות ממודל 1, וקרובות לאלו המתקבלות לפי השיטה הטובה ביותר תחת המודל ARMA השומר על התנודות. הדבר נובע מכך שתחת מודל זה, נעשה שימוש באינפורמציה נוספת על ההתנהגות האמיתית של טעות מדידה, בעוד שהשיטות האחרות לא לוקחות מידע זה בחשבון. יצוין עוד, כי יתרון זה יבוא לידי ביטוי רק במקרה בו ניתן לאמוד מודל ARMA לשאריות באופן מובהק, לרבות מובהקות המקדמים.

5.3 שבל ב' - ניתוח 7 הסדרות הנאמדות על סמך מדדי תעשייה

5.3.1 בחירת מודל ARMA השומר על תנודות

בשלב זה, בוצעו הרצות של 7 הסדרות של חשבונות לאומיים המבוססות על מדדי תעשייה (סדרות (4) - (10) בטבלה בסעיף 4.1). נבדקו מודלים ARMA נוספים השומרים על תנודות (שמספרם 10, 11, 12, 13). תחת מודל 2, ישנו שיפור לפחות באחת המדדים לטיב הביון והחיוץ, בכל 7 הסדרות שנבדקו. כמו כן, עבור הסדרות הני"ל, מודלים 7, 11 ו-12 משפרים מדדי טיב הביון והחיוץ ביחס למודל 1, אך סטיית התקן הממוצעת הינה גבוהה ומגיעה ל 1% מרמת הסדרה. נזכיר כאן כי סטיית התקן של הנתונים המקוריים בתדירות גבוהה נקבעה כ 1% מהנתון. לכן, תוצאה זו מצביעה על הגדלת סטיית תקן כתוצאה מביצוע ביון וחיוץ בכ- 100%, ביחס לסטיית התקן המקוריות. כפי שראינו בשלב א', הגדלת סטיית התקן גורמת לחוסר יציבות בסדרה ולפגיעה במדדי טיב ניכוי עונתיות. לכן, לא נתייחס למודלים אלה בהשמך הניתוח. נשים לב, כי עבור כל הסדרות, סטיית התקן הממוצעת הקטנה ביותר מושגת במודל 1. להלן מסקנות הניתוח לבחירת מודל לביון וחיוץ בסדרות (4) - (10):

1. בסדרה "תמורה למשרות" מודל 2 הוא המודל הטוב ביותר מבין אלה שנבדקו.
2. בסדרה "שעות עבודה למעשה בתעשייה" מודלים 5 ו-10 משפרים את טיב הביון והחיוץ בהשוואה למודל 1, תוך שמירה על סטיית התקן הממוצעת קטנה יחסית. בסדרה זו, יש לבחור במודל 10

- בגלל סטיית התקן הממוצעת קטנה יותר, והפרש בין מדדי החלקלקות בין סדרת האינדקטור לבין הסדרה הנאמדת קטן יותר.
3. בסדרה "מספר משרות בתעשייה" שימוש במודל 2 משפר את האומדנים המתקבלים, ביחס לשיטה הקיימת.
4. בסדרה "מספר משרות שכיר בתעשייה" נעדיף מודל 2 המשפר את מדד הכפלי לשמירת תנודות בטווח קצר בסדרה הנאמדת.
5. בסדרה "תוצר תעשייתי" מודלים 2, 3 ו-5 מספקים תוצאות טובות במובן מזעור הפרש בין מדדי החלקלקות בסדרת האינדקטור ובסדרה הנאמדת, כאשר עליה בסטיית התקן הממוצעת של הסדרה הנאמדת אינה משמעותית, ביחס למודל 1. עם זאת, המדדים הכפליים לשמירת אחוזי שינוי בסדרה הנאמדת במודלים 3 ו-5 גבוהים (טובים פחות), בהשוואה לזה המתקבל ממודל 2. במצב זה, נעדיף את מודל 2.
6. בסדרה "פדיון תעשייתי" מודלים 4, 5, ו-13 משפרים במידה משמעותית את מדדי טיב הביון והחיוץ בסדרה זו, ביחס לשיטה הקיימת בלמ"ס. בהתחשב בכל מדדי הטיב שנבדקו, יש לבחור את מודל 13 בסדרה זו.
7. בסדרה "תוצר תעשייתי בענף הכרייה וחציבה" מודלים 1 ו-2 מספקים דיאגנוסטיקות דומות מאוד, עם עדיפות קלה למודל 2 הממזער את המדד הכפלי לשמירת תנודות בטווח הקצר בסדרה הנאמדת.

5.3.2 בחירת מודל ARMA לשאריות

תוצאות האמידה של מודל ARMA לשאריות לפי האלגוריתם המוצע בסעיף 3.3 מובאים בלוח 8. נשים לב, כי בסדרות (4), (5), (8) ו-(9) המקדמים של מודלים ARMA הנאמדים אינם מובהקים. הדבר נובע, בפרט, מקיום תצפיות חריגות בנתונים מקוריים, כפי שצינו בפרק 5.1. בסדרה (9), כפי שראינו בתרשים 9, אחוזי השינוי בין נתון לנתון בסדרת האינדקטורים ובסדרה השנתית המתארת את התופעה הנאמדת קרובים מאוד, ובמצב זה התנהגות השאריות קרובה ל"רעש לבן".

בסדרות (4), (5), (6), (7) ו-(9) שימוש במודל סטוכסטי לשאריות הנאמד מתוך הנתונים ממזער את ההפרש בערך מוחלט בין מדדי החלקלקות בסדרת האינדקטור ובסדרה הנאמדת, ומדד כפלי לשמירת תנודות בטווח קצר. במובן זה, זהו המודל הטוב ביותר מבין כל המודלים שנחקרו במסגרת עבודה זו. בסדרות (8) ו-(10) לא קיימת עדיפות למודל סטוכסטי, כאשר שני המדדים הנ"ל קרובים לאלה המתקבלים תוך שימוש בשיטות השומרות על תנודות. חשוב להדגיש כי בסדרות בהן מקדמי המודל ARMA לא היו מובהקים ((4), (5), (8) ו-(9)), סטיית התקן הממוצעת הינה גבוהה יחסית, ומגיעה ל-1% מרמת הסדרה הנאמדת בסדרות. במצב זה, נעדיף שימוש במודלים ARMA השומרים על תנודות שנבחרו בסעיף 5.3.1 לעיל.

5.4 בחירת מודל לביון וחיוץ בסדרות חודשיות של סכ"א

נזכיר, כי אנו מעוניינים לשפר את האומדנים החודשיים של 4 הסדרות העיקריות של סכ"א תוך התאמתם לנתוני הסקר המתקבלים מדי רבע. במונחים של ביון וחיוץ, סדרות האומדנים החודשיים המנופחים ישמשו באינדקטור, והנתונים המקבילים של סכ"א המתקבלים על סמך המדגם הרבעוני ישמשו כעוגן (נתון בתדירות נמוכה, benchmark). השוואת אחוזי שינוי מנתון לנתון מראה, כי אחוזי השינוי של

האינדיקטור ושל הסדרה בתדירות רבעונית דומים בכל 4 המקרים. נסיק, כי אומדנים חודשיים יכולים לשמש כאינדיקטורים ב- 4 הסדרות שבניתוח.

ניתן להשיג שיפור משמעותי ביחס לשיטה הקיימת (מודל IS) תוך שימוש במודל כפלי עם אמידת החותך ומודלים ARMA השומרים על תנודות. בולט יתרונם של המודלים 3, 4, 8, 9 ו-10, כאשר המדד הכפלי לשמירת התנודות לא משתנה, ועיקר השיפור מושג בהקטנה משמעותית של סטיית התקן הממוצעת. כך, בסדרה "נשים מועסקות" מעבר למודל 8 מקטין את סטיית התקן הממוצעת בסדרה מ- 127.21 ל- 3.93 בלבד, ותוצאות דומות מתקבלות בשאר הסדרות. לוח 9 מראה, כי ניתן להקטין עדכונים ולהעלות משמעותית את טיב החיוץ בסדרות הנבדקות. בסדרה "נשים מועסקות" טעות החיוץ ברבע רביעי של שנת 2007 ירדה מ- 31,326 בשיטה הקיימת ל- 2,364 במודל 4. שיפור דומה ניתן לראות בשאר הסדרות: 11,924 בשיטה הקיימת לעומת 514 במודל 4 ו- 690 במודל 6 בסדרה "גברים מועסקים", 25,260 בשיטה הקיימת לעומת 801 במודל 4 בסדרה "נשים בלתי מועסקות", ו- 1,114 בשיטה הקיימת לעומת 108 במודל 3 ו- 141 במודל 8 בסדרה "גברים בלתי מועסקים". כתוצאה מהורדת סטיית התקן בנתונים הנאמדים על ידי השימוש במודלים ARMA השומרים על תנודות, ניתן לשפר משמעותית את יציבות הסדרות המתקבלות ולמזער סתירות בין שני מקורות מידע (הודעה לעיתונות אודות המגמות בסדרות העיקריות של סכ"א בתדירות חודשית ופרסום אומדני סכ"א מדי רבע). בסדרות החודשיות של סכ"א שימוש במודל ARMA הנאמד מהנתונים מספק תוצאות סבירות, כאשר מדדי טיב האומדנים המתקבלים דומה מאוד לאלה המתקבלים מהשיטה הקיימת. נסיק, כי טענות על נחיתות של שיטות סטוכסטיות ביחס לשיטה של דנטון (או קירובים אליה) נראית לא נכונה גם בסדרות החודשיות של סכ"א. ניתן לראות, שלא בכל המקרים המדדים לביון והמדדים לחיוץ מסכימים, ולכן היחידה הנושאת תבחר את המודל לפי מטרותיה. ניתן לסכם את ההמלצות בטבלה הבאה:

שם סדרה	מודל מומלץ לביון	מודל מומלץ לחיוץ
נשים מועסקות	9	4
גברים מועסקים	9	4
נשים בלתי מועסקות	8	4
גברים בלתי מועסקים	8	3

משמעות הבחירה של מודל הטוב ביותר לביון – הבטחת היציבות ותוך תקופת העיבוד הנתונה. בחירת המודל לפי מדדי טיב החיוץ תקטין היקף העדכונים בסדרה הנאמדת עם קבלת הנתון הרבעוני העדכני.

6 הרחבת המחקר – שימושים נוספים בלמ"ס

בנוסף לאומדני חשבונות לאומיים ונתוני סכ"א, ישנו מספר תחומים נוספים בלמ"ס בהם ניתן להשתמש בשיטת הרגרסיה לביון וחיוץ. להלן מוצעים שימושים נוספים לשיטות המוצעות בעבודה זו:

- תחזיות בתחום מערכת החינוך. מדי שנה, נבנים תחזיות לחמש השנים הבאות של מספר מורים, מספר תלמידים, צרכי הכשרה ועוד. חלק מהסדרות אלו נמדדים (או נאמדים) בהווה בשיטות שונות על סמך נתונים מנהליים ואומדנים דמוגרפיים. התחזיות נאמדות על סמך שיטת ההשלכה של אחוזי שינוי בהווה לתקופות עתידיות, או על סמך תחזיות של מודל רגרסיה (ליניארית או לוגיסטית). במקרים רבים, קל יותר לחזות את המדד הנבדק בתדירות נמוכה יותר מאשר שנתית (למשל, תחזית לחמש שנים סה"כ).

ניתן לאמוד תחזיות באופן הבא. במקרה זה, אינדיקטור יכול להיות סדרת אומדנים דמוגרפיים, סיכום נתונים מנהליים או להיות תוצאה של מודל סטטיסטי או כלכלי ("אינדיקטור סינטטי"). ניתן למדוד את התופעה הנאמדת בתדירות נמוכה בדרכים שונות. כך, ניתן להתייחס לתחזיות רב-שנתיות הנבנות ממודל, ליעדים של המערכת, או לתסריטים שונים של ההתפתחות העתידית. התאמת האינדיקטור לנתונים אלה תוך שימוש במודל רגרסיה (5) לביון וחיוץ מאפשרת ניתוח פשוט של התערבויות שונות במערכת החינוך כתוצאה משינויים מבניים או אחרים המוצעים על ידי קובעי המדיניות ללא צורך להניח הנחות פרמטריות בכל שלבי האמידה.

- יצירת אומדנים בתדירות חודשית או רבעונית על סמך הסקרים השנתיים. כך, ניתן לאמוד סדרות של צריכה של קבוצות מוצרים שונות תוך שימוש במודל הרגרסיה (5) לביון וחיוץ בעזרת האלגוריתם מסעיף 3.4 לקביעת מודל ARMA לשאריות. במקרה זה, ניקח סדרות מדדי מחירים לקבוצות המוצרים הרלוונטיות או/ו נתוני פדיון בענפים הרלוונטיים באינדיקטורים, ונתונים שנתיים מסקר הוצאות משקי בית אודות צריכת המוצרים המתאימים ייקבעו כ- benchmarks בתדירות נמוכה (שנתית במקרה הנדון). אומדנים אלה יכולים להיות שימושיים במחקר הכלכלי.

7. מסקנות והמלצות

המסקנות העיקריות של המחקר הנוכחי:

- יש לשדרג את מערכת הביון והחיוץ של סדרות עתיות בלמ"ס, תוך מעבר לשימוש במודלים ARMA השומרים על תנודות, ובמקרים בהם ניתן לאמוד באופן מובהק את מודל ARMA מאומדן ראשוני לשאריות, יש לעבור למודל המתואר בסעיף 3.3.
- בסדרות שנתיים, מומלץ לשדרוג מערכת ביון וחיוץ לפי השיטה המוצעת בסעיף 3.5, בגלל המגבלות הרבות של השיטה הנומרית המופעלת בתחום חשבונות לאומיים, ואודות היתרונות של שיטת הרגרסיה שתוארו בעבודה זו.

להלן מובא פירוט המסקנות:

1. לפני קביעת המודל לביון וחיוץ, יש לבדוק ההתאמה בין סדרת האינדיקטור לבין הסדרה השנתית המתארת את התופעה הנאמדת. במקרה של אי-התאמה בין שתי הסדרות, יש להפנות תשומת לב של היחידה הנושאית לכך, ולנסות למצוא אינדיקטור מתאים יותר. במידה והדבר בלתי אפשרי מבחינה נושאית, יש לקחת עובדה זו בחשבון בבחירת מודל ARMA לשאריות או/ו בקביעת מודל לחותך. המחקר מראה, כי בתקופות בהן נראית אי-התאמה מסוג זה, טיב האומדנים המתקבלים יורד, במיוחד בשיטה הקיימת היום בלמ"ס.
2. שימוש בשיטת הרגרסיה משפר משמעותית את אומדני הביון והחיוץ בהשוואה לשיטות נומריות, גם כאשר משתמשים במודלים ARMA פשוטים יחסית של "הליכה מקרית" לשאריות, מסדר ראשון ושני. לכן, שימוש במודלים אלה מהווה לא רק קירוב לשיטות נומריות של דנטון, אלא גם שיפור בטיב האומדנים המתקבלים.

3. שימוש במודל הרגרסיה הכפלי עם אמידת החותך משפר את האומדנים המתקבלים, בכל הסדרות שנבדקו. מעבר למודל זה עשוי להעלות את טיב האומדנים ויבטיח עקביות עם שיטת ניכוי העונתיות בלמ"ס המבוססת על השימוש במודל הכפלי לפירוק העונתי, ברוב הסדרות.

4. שימוש במודלים ARMA לשאריות השומרים על תנודות עשוי לשפר משמעותית את טיב האמידה, במובנים של (1) שמירה על התנודות בסדרה הנאמדת ביחס לאינדיקטור, (2) שמירה על מבנה השפעת החגים וימי הפעילות ועל מבנה העונתיות בסדרה הנאמדת ביחס לאינדיקטור, (3) מזעור העדכוניים עם הוספת הנתון העדכני בתדירות נמוכה, ביחס למודל 1 הנמצא היום בשימוש הלמ"ס. מודלים אלה הינם מודלים סטוכסטיים אשר "מחקים" את פונקציות מטרה מורכבות יותר מאשר אלו של דנטון (נוסחה 1). בכל הסדרות שנבדקו, נמצא מודל ARMA השומר על תנודות אשר משפר טיב ביון, חיוץ וניכוי עונתיות, בהשוואה למודל 1 הנמצא היום בשימוש הלמ"ס. השיפור כתוצאה משימוש במודלים סטוכסטיים השומרים על תנודות ניכר במיוחד באומדני החיוץ, והוא מתבטא במזעור של היקף העדכוניים בסדרות הנאמדות. יש לשים לב, כי ישנם מקרים בהם קיימת סתירה בין מדדי טיב הביון למדדי טיב החיוץ, כפי שראינו בחלק מהסדרות הנאמדות על סמך מדדי התעשייה ובסדרות החודשיות של סכ"א. במקרה זה, שיקול כלכלי-נושאי יקבע את המודל המתאים. עבור הסדרות שנבדקו, המודלים הטובים ביותר ששומרים על תנודות מתוארים בסעיפים 5.2, 5.3 ו-5.4.

5. בכל הסדרות שנחקרו, בהם ניתן להתאים באופן מובהק את מודל ARMA לשאריות לפי האלגוריתם בסעיף 3.3, מדדי טיב של השיטה הסטוכסטית המוצעת דומים לאלה המתקבלים מהקירוב לשיטה של דנטון. יתר על כן, המודל הנ"ל מספק הפתרון הטוב ביותר לבעיית הביון בסדרות בהן ניתן לאמוד מודל ARMA לשאריות באופן מובהק. עם זאת, כאשר קיים קושי להתאים מודל ARMA לפי האלגוריתם בסעיף 3.3, סטיית התקן של האומדנים המתקבלים עולה משמעותית, מה שמוריד את יציבות האומדנים ואת טיב ניכוי העונתיות. במקרה זה, כדאי להשתמש במודל השומר על התנודות. בנוסף, אמידת המודל לשאריות לפי האלגוריתם המוצע יכולה להיות מושפעת מהחריגים בסדרת האינדיקטור או/ו מנתונים חריגים בסדרה בתדירות נמוכה. לכן, בכל הסדרות שנבדקו, העדפנו את המודל השומר על תנודות על פני המודל הנאמד. אמידת מודל ARMA לשאריות תוך איתור ואמידת תצפיות חריגת הינו נושא למחקר עתידי. פיתוח שיטה סטוכסטית העמידה לחריגים עשוי לשפר את יציבות הסדרה הנאמדת, להוריד סטיות התקן שלה, וכתוצאה מכך להעלות את טיב האומדנים המחויצים.

6. מודל סופי לביצוע ביון וחיוץ נקבע על סמך מגוון רחב של מדדים. עם זאת, ישנם שלושת המדדים העיקריים: (1) מדד כפלי לשמירת אחוזי שינוי בסדרה הנאמדת (מסופק על ידי תוכנה לביון וחיוץ בפלט ההרצה), (2) הפרש בין מדדי החלקלקות בין האינדיקטור לסדרה הנאמדת, (3) סטיית התקן הממוצעת בתקופת העיבוד. מזעור שלושת המדדים הנ"ל יבטיח יציבות האומדנים, שמירה על מבנה השפעת החגים וימי פעילות, שמירה על מבנה עונתיות וטיב ניכוי עונתיות, ומזעור מספר והיקף העדכוניים בסדרה עם הוספת הנתון העדכני בתדירות נמוכה. כאשר אין הסכמה בין שלושת המדדים, יש לערוך השוואה בין מדדי ניכוי עונתיות וניתוח יציבות.

ניתן להחליף מדד (3) בסטיית התקן של הנתון האחרון בתקופת העיבוד (בדרך כלל, שם מתקבלת סטיית התקן המקסימאלית בתקופת העיבוד הנתונה). מדד זה מסופק אף הוא על ידי התוכנה לביון וחיוץ. העבודה הנוכחית מראה, כי מספיק לבדוק מודלים (2) - (5), (7) - (13), ואין צורך לבדוק מודלים ARMA מסדר גבוה.

7. מומלץ לשדרג מערך הביון והחיוץ של הסדרות השנתיות, תוך שימוש במודל של חותך משתנה המוצע בסעיף 3.5.

8. מומלץ לשדרג מערכת ביון וחיוץ בסדרות הלמ"ס, ולהטמיע תוכנה GBS לביצוע התהליך. הטמעת תוכנה זו תאפשר ניצול היתרונות הרבים הטמונים בה, ובפרט שדרוג מערכת הביון והחיוץ של הסדרות השנתיות בתחום של חשבונות לאומיים. בנוסף, בעבודה Gubman and Burck, 2005, המלצנו להתאים מערכת של סדרות עתיות תוך שמירה על הקשרים החשבוניים ביניהן, באמצעות שימוש בשיטת הרגרסיה. שיטה זו לא הוטמעה בלמ"ס עד כה, אך שימוש בה עשוי לפתור בעיית הסתירה בין מקורות מידע שונים, במיוחד בין מערכות של סדרות מנכות העונתיות והמגמות. הקשרים החשבוניים מופרים כתוצאה מהפעלה ישירה של תהליך של ניכוי עונתיות באמצעות X12 על מערכת של סדרות עתיות.

9. יש לעקוב אחרי התפתחות של סדרות הלמ"ס. מומלץ לערוך בדיקות המודלים בסדרות עיקריות לפחות פעם בשנתיים, על מנת לעדכן, לפי צורך, את הפרמטרים להרצת הביון והחיוץ.

- Barcellan, R., (2005) *The use of benchmarking techniques in the compilation of the European quarterly national accounts: situation and perspectives*, Workshop on Frontier in Benchmarking Techniques and Their Applications to Official Statistics, Luxembourg
- Barcellan, R., and Buino, D., (2002) *ECOTRIM Interface, User Manual*, Eurostat
- Bloem, A.M., Dippelsman, R.J., Maehle, N.O., (2001), *Quarterly National Accounts Manual – Concepts, Data Sources, and Compilation*, International Monetary Fund, Textbook, May 2001
- Chen, Z. G., Cholette, P.A., and Dagum, E.B., (1997) A nonparametric method for benchmarking survey data via signal extraction, *Journal of the American Statistical Association*, Vol. 92, No. 440, 1563-1571
- Chen, Z.G., and Wu, K.H., (2006) Comparison of benchmarking methods with and without a survey error model, *International Statistical Review*, Vol. 74, 283-304
- Cholette, P.A., and Dagum, E.B., (1994), *Benchmarking socio-economic time series data: a unified approach*, Time series research and analysis, Statistics Canada, Working paper 89-006E
- Cholette, P.A., (1994), *Bench program, manual*, Working paper TSRA-90-008, Methodology Branch, Statistics Canada.
- Cholette, P.A., and Chhab, N., (2004) *Users' manual of the Generalized Benchmarking System*, Statistics Canada
- Dagum, E.B., and Cholette, P.A. (2006) *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*, Lecture Notes in Statistics, No186, USA
- Denton, F.T., (1971), Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization, *Journal of the American Statistical Association*, Vol.66, No. 333, 365-377
- Dickey, D.A., (2005) *Stationarity issues in time series models*, 30th SAS Users Group International (SUGI 30) Conference, paper 192-30
- Di Fonzo, T., and Marini, M. (2003), *Benchmarking systems of seasonally adjusted time series according to Denton's movement preservation principle*, Università di Padova, Working paper 2003.09
- Findley, D.F., Monsell, B.C., Bell W.R., Otto, M.C., Chen, B.C., (2000), *New capabilities and methods of the X-12-ARIMA seasonal adjustment program*, U.S. Bureau of the Census, USA.
- Gubman, Y., Burck, L., (2005), *Benchmarking of Israeli economic time series*, Workshop on Frontier in Benchmarking Techniques and Their Applications to Official Statistics, Luxembourg

Hillmer, S.C., and Trabelsi, A., (1990) Benchmarking time series with reliable benchmarks, *Applied Statistics*, Vol. 39, No. 3, 367-379

הגורמים העונתיים וגורמי ההתאמה מראש לשנת 2007 והמגמות לשנים 2001-2006,

פרסום אינטרנט: <http://www.cbs.gov.il/publications/tseries/seasonal07/presentationh07.pdf>

בורק לאויזה, סלמה אסתר, גובמן יורי, (2006) *ביון וחיוץ באמצעות תוכנת BENCH וניכוי עונתיות*

בסדרה "תוצר תעשייתי", דו"ח טכני, הלשכה המרכזית לסטטיסטיקה, ירושלים

גובמן יורי, (2005) *Benchmarking בסדרות של חשבונות לאומיים*, נייר טכני, הלשכה המרכזית

לסטטיסטיקה, ירושלים

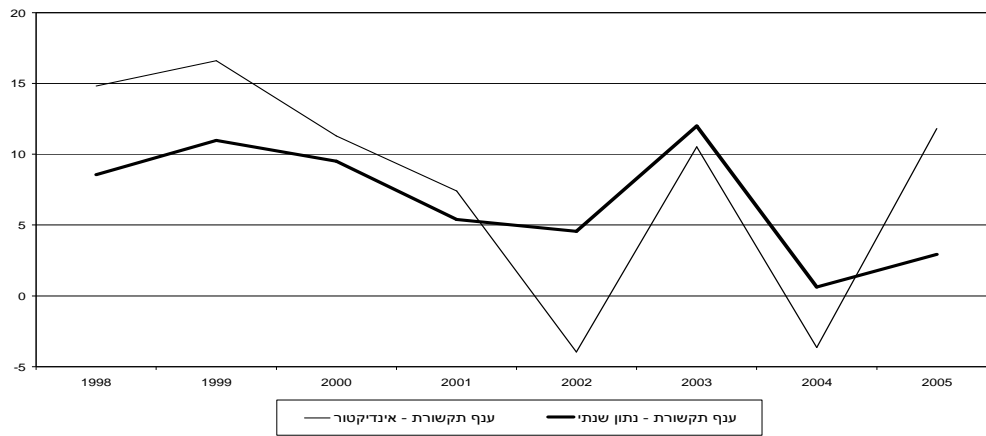
גובמן יורי, בורק לואיזה, (2005) *Benchmarking וניכוי עונתיות של סדרות סקר כוח אדם*, נייר טכני,

הלשכה המרכזית לסטטיסטיקה, ירושלים

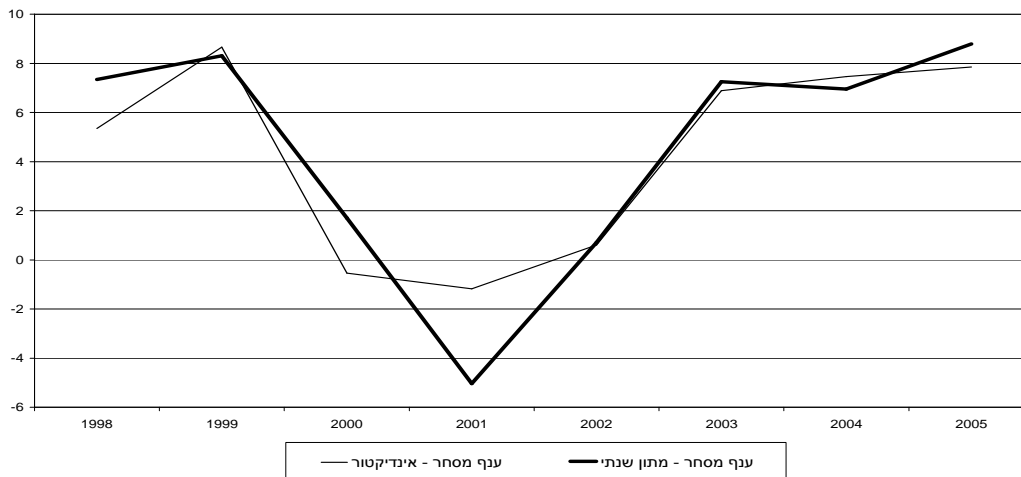
שנתון סטטיסטי לישראל, הלשכה המרכזית לסטטיסטיקה

**נספחים – לוחות ותרשימים
(ניתן לקבל חומר נוסף מהמחבר)**

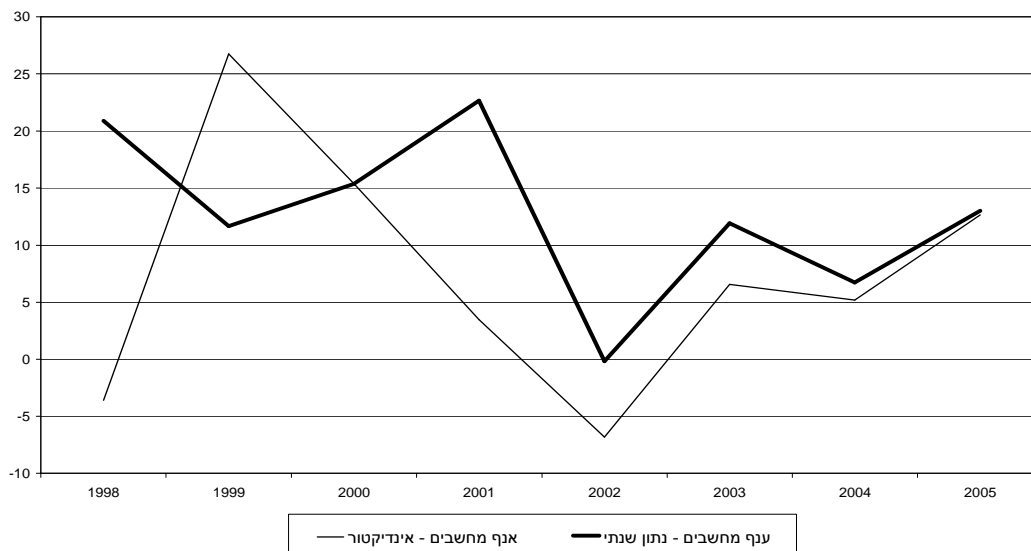
תרשים 1: תוצר בענף התקשורת - אחוזי שינוי בנתונים שנתיים ובאינדקטור



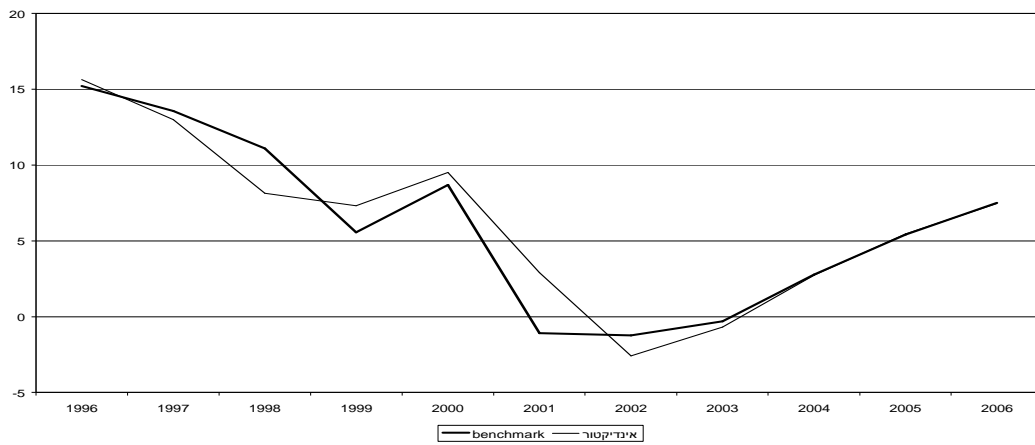
תרשים 2: תוצר בענף המסחר - אחוזי שינוי בנתונים שנתיים ובאינדקטור



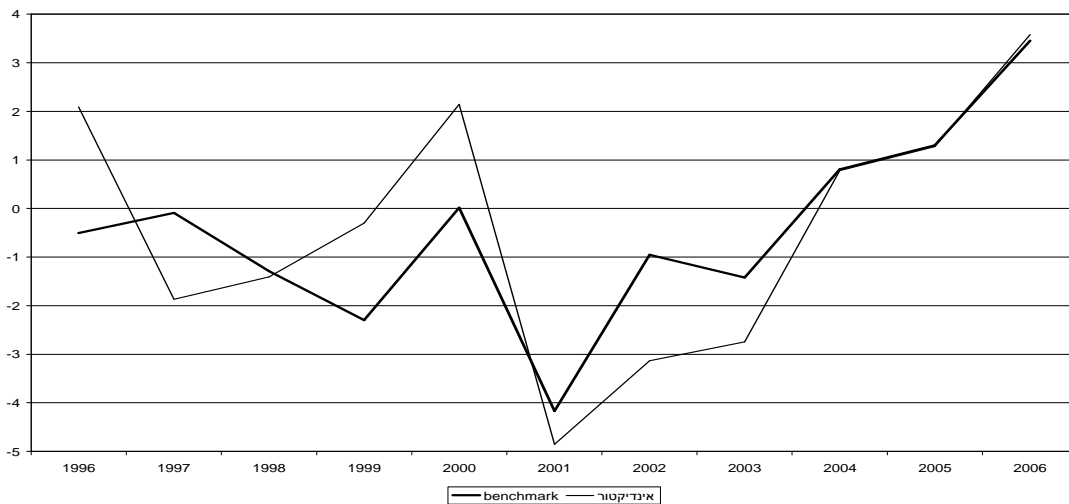
תרשים 3: תוצר בענף המחשבים - אחוזי שינוי בנתונים שנתיים ובאינדקטור



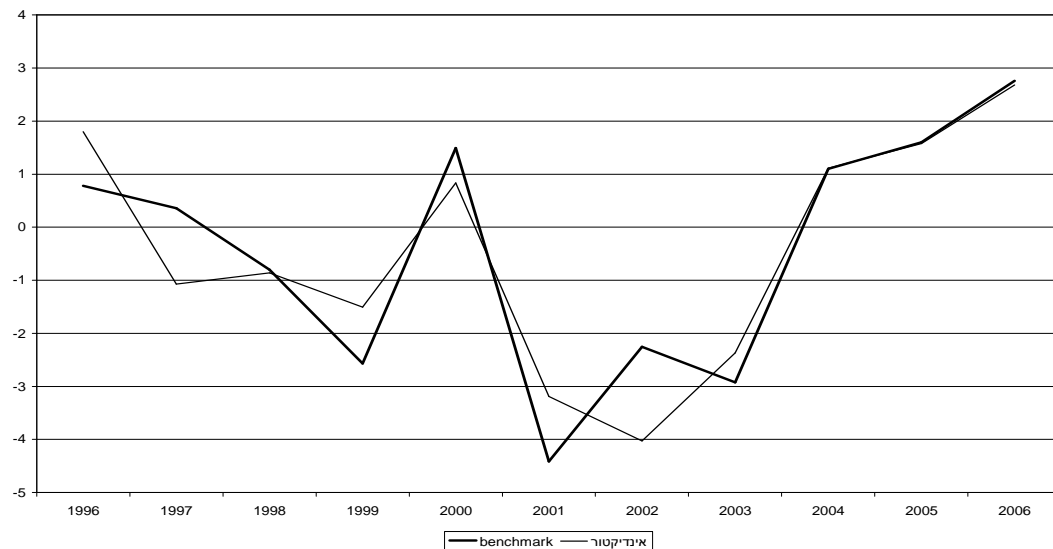
תרשים 4 : אחוזי שינוי משנה לשנה בסדרות : מדד עלות עבודה (האינדיקטור) ותמורה למשרות (benchmark)



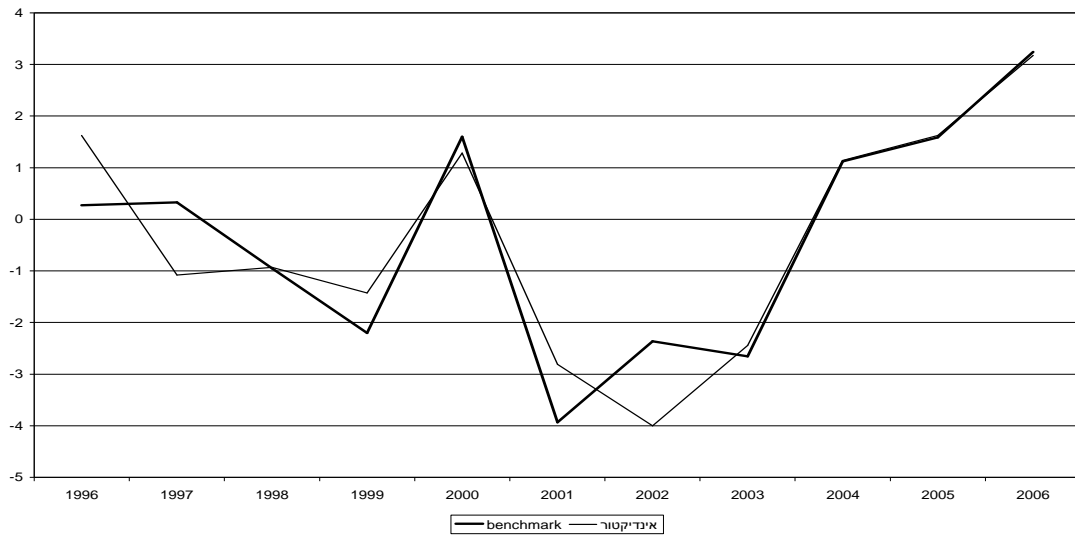
תרשים 5 : אחוזי שינוי משנה לשנה בסדרות : מדד שעות עבודה למעשה (האינדיקטור) ושעות עבודה למעשה בתעשייה (benchmark)



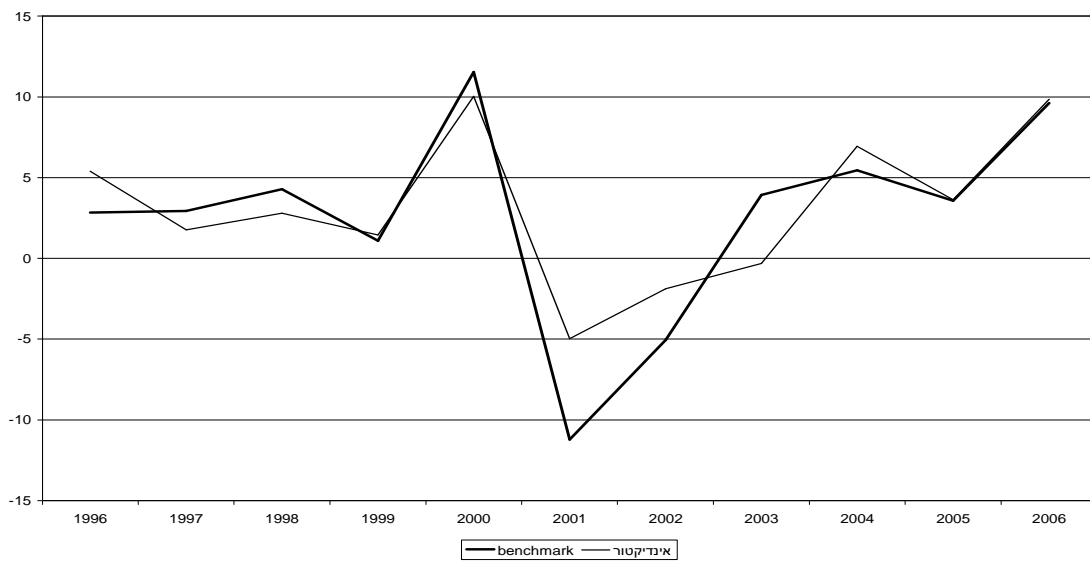
תרשים 6 : אחוזי שינוי משנה לשנה בסדרות : מדד משרות (האינדיקטור) ומספר משרות בתעשייה (benchmark)



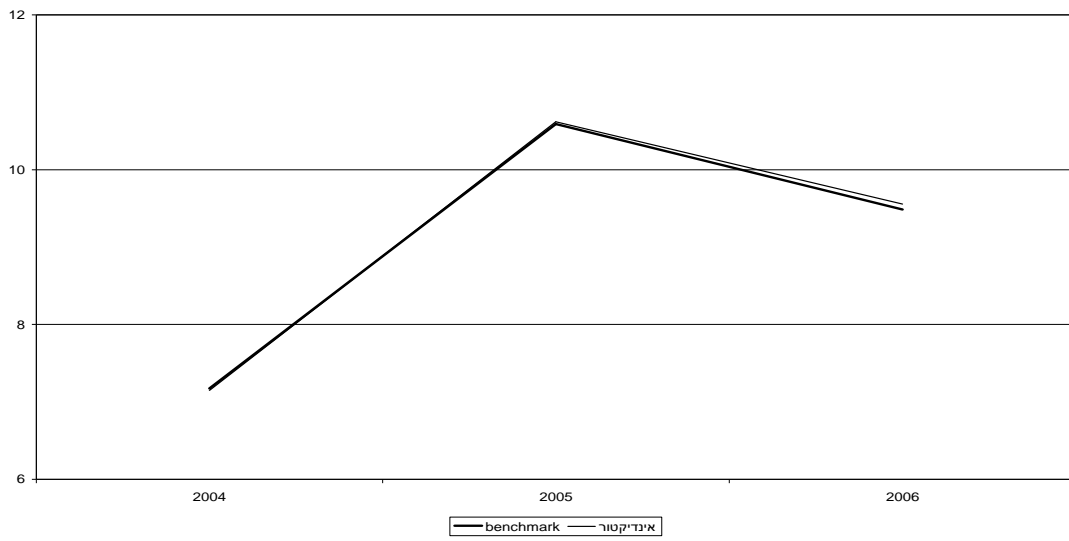
תרשים 7 : אחוזי שינוי משנה לשנה בסדרות : מדד משרות שכיר (האינדקטור)
ומספר משרות שכיר בתעשייה (benchmark)



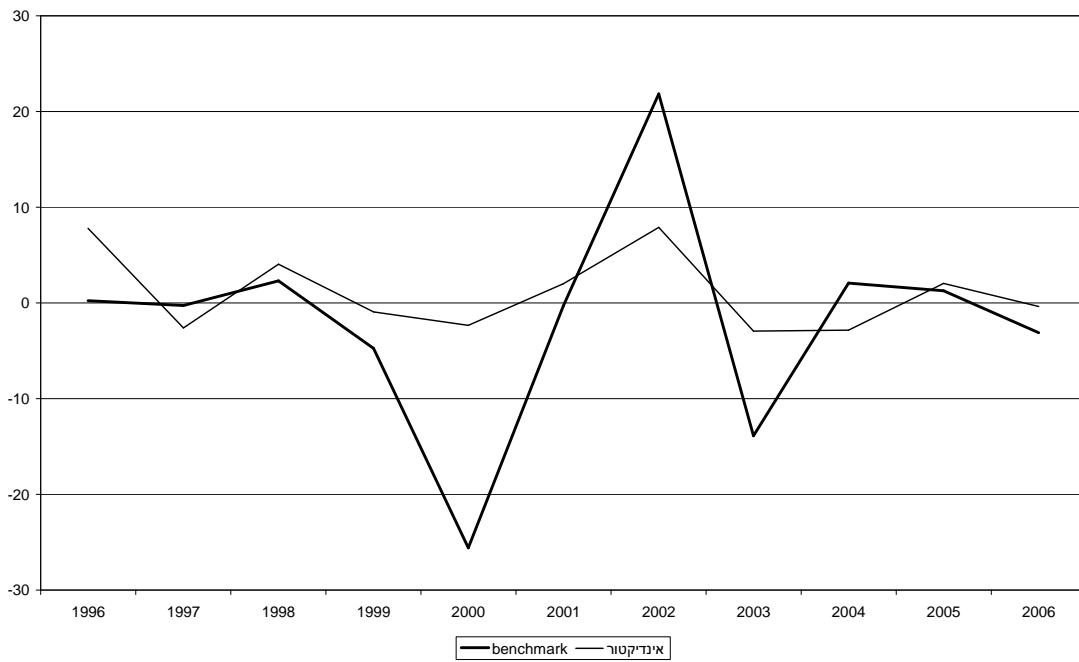
תרשים 8 : אחוזי שינוי משנה לשנה בסדרות : מדד ייצור תעשייתי (האינדקטור)
ותוצר תעשייתי (benchmark)



תרשים 9 : אחוזי שינוי משנה לשנה בסדרות : מדד פדיון במחירים שוטפים (האינדקס) ופדיון תעשייתי (benchmark)



תרשים 10 : אחוזי שינוי משנה לשנה בסדרות : מדד ייצור תעשייתי בענף כרייה וחציבה (האינדקס) ותוצר בענף כרייה וחציבה (benchmark)



לוח 1א: מדדים מסכמים לטיב ביון וחיזוי של 3 סדרות תוצר באמצעות שיטה פרופורציונלית של דנטון ו בשיטת רגרסיה עם מודל "הליכה מקרית" לשאריות: $p=0.999, ARMA(1,0)(0,0)$

מדדים					מודל	חותך	מספר מודל	שם סדרה
הפרש, בערך מוחלט, בין מדדי חלקיקות של סדרות מקוריות - מותאמת מול אינדיקטור	מדד חיבורי לשמירת אחוזי שינוי	מדד כפלי לשמירת אחוזי שינוי	סטיית תקן ממוצעת	מובהקות החותך				
0.49	50.658	0.021	---	---	דנטון פרופורציונלי	ללא חותך	1	תוצר בענף תקשורת
0.26	33.439	0.013	0.603	---	מודל רגרסיה - חיבורי	ללא חותך*		
1.31	29.029	0.018	0.619	כן		עם חותך		
0.26	33.318	0.013	0.811	---	מודל רגרסיה - לוג-חיבורי	ללא חותך		
0.26	33.269	0.013	0.811	כן		עם חותך		
3.43	526.922	0.076	---	---	דנטון פרופורציונלי	ללא חותך	1	תוצר ענף מסחר
0.05	449.307	0.004	6.117	---	חיבורי	ללא חותך*		
1.24	357.361	0.013	6.639	כן		עם חותך		
0.05	449.377	0.004	2.269	---	לוג-חיבורי	ללא חותך		
0.04	449.377	0.004	2.269	כן		עם חותך		
1.17	371.976	0.031	---	---	דנטון פרופורציונלי	ללא חותך	1	תוצר בענף מחשבים
0.43	363.625	0.025	0.472	---	חיבורי	ללא חותך*		
6.32	190.535	0.073	0.583	כן		עם חותך		
0.42	363.762	0.025	1.555	---	לוג-חיבורי	ללא חותך		
0.42	363.762	0.025	1.555	כן		עם חותך		

* נמצא בשימוש הלמ"ס

לוח 1ב: מדדים מסכמים לטיב ביון וחיוץ של 3 סדרות תוצר באמצעות שיטת רגרסיה עם מודלים ARMA שונים לשאריות							
שם סדרה	מספר מודל	מודל רגרסיה	מודל ARMA לשאריות	מדדים			
				מובהקות החותר	סטיית תקן ממוצעת* שניית תקן	מדד כפלי לשמירת אחוזי שינוי	מדד חיבורי לשמירת אחוזי שינוי
תוצר בענף תקשורת	3	חיבורי עם חותר	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	0.217	0.019	31.521	1.39
			(1,0)(1,0) $p_1=0.999, P_1=0.9$	0.108	0.024	42.117	1.81
			(1,1)(1,0) $p_1=0.999, q_1=0.999, P_1=0.999$	2.411	0.021	21.274	0.29
			(1,1)(0,1) $p_1=0.999, q_1=-0.999, Q_1=0.9$	0.272	0.024	42.213	1.83
			(0,1)(0,1) $q_1=-0.999, Q_1=-0.9$	14.817	0.017	29.403	1.04
	4	לוג-חיבורי עם חותר	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	0.256	0.014	39.786	0.18
			(1,0)(1,0) $p_1=0.999, P_1=0.9$	0.169	0.014	37.383	0.23
			(1,1)(0,1) $p_1=0.999, q_1=-0.999, Q_1=0.9$	0.872	0.013	34.873	0.18
			(1,1)(1,0) $p_1=0.999, q_1=0.999, P_1=0.999$	0.306	0.024	67.273	0.15
			(0,1)(0,1) $q_1=-0.999, Q_1=-0.9$	20.019	0.014	39.645	0.10
תוצר בענף מסחר	3	חיבורי עם חותר	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	2.347	0.046	177.641	4.57
			(1,0)(1,0) $p_1=0.999, P_1=0.9$	1.289	0.011	516.638	0.75
			(1,1)(0,1) $p_1=0.999, q_1=-0.999, Q_1=0.9$	3.083	0.017	560.180	1.18
			(1,1)(1,0) $p_1=0.999, q_1=0.999, P_1=0.999$	2.411	0.008	434.146	0.40
			(0,1)(0,1) $q_1=-0.999, Q_1=-0.9$	150.250	0.064	199.298	6.25
	4	לוג-חיבורי עם חותר	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	0.731	0.005	449.746	0.06
			(1,0)(1,0) $p_1=0.999, P_1=0.9$	0.458	0.005	447.289	0.10
			(1,1)(0,1) $p_1=0.999, q_1=-0.999, Q_1=0.9$	2.433	0.004	451.055	0.03
			(1,1)(1,0) $p_1=0.999, q_1=0.999, P_1=0.999$	0.872	0.007	449.547	0.26
			(0,1)(0,1) $q_1=-0.999, Q_1=-0.9$	55.861	0.004	447.829	0.06
תוצר בענף מחשבים	3	חיבורי עם חותר	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	0.278	0.090	171.802	7.58
			(1,0)(1,0) $p_1=0.999, P_1=0.9$	0.103	0.113	160.668	8.66
			(1,1)(0,1) $p_1=0.999, q_1=-0.999, Q_1=0.9$	0.250	0.049	238.745	3.77
			(1,1)(1,0) $p_1=0.999, q_1=0.999, P_1=0.999$	0.167	0.006	480.140	2.55
			(0,1)(0,1) $q_1=-0.999, Q_1=-0.9$	11.483	0.091	178.513	7.17
	4	לוג-חיבורי עם חותר	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	0.369	0.028	362.191	0.58
			(1,0)(1,0) $p_1=0.999, P_1=0.9$	0.236	0.029	353.774	0.12
			(1,1)(0,1) $p_1=0.999, q_1=-0.999, Q_1=0.9$	1.244	0.028	360.447	0.25
			(1,1)(1,0) $p_1=0.999, q_1=0.999, P_1=0.999$	0.403	0.054	411.228	0.47
			(0,1)(0,1) $q_1=-0.999, Q_1=-0.9$	28.414	0.026	374.432	0.63

* סטיית התקן חושבו בהנחה כי בסדרות האינדיקטורים CV=1

לוח אג: מדדים מסכמים לטיב בין וחיוץ של 3 סדרות תוצר באמצעות שיטת רגרסיה מודלים ARMA לשאריות מסדר גבוה							
שם סדרה	מספר מודל	מודל רגרסיה	מודל ARMA לשאריות	מדדים			
				מובהקות החותר	סטיית תקן ממוצעת*	מדד כפלי לשמירת אחוזי שינוי	מדד חיבורי לשמירת אחוזי שינוי
תוצר בענף תקשורת	8	לוג-חיבורי עם חותר	$(2,1)(0,0) p1=1.98, p2=-0.999, q1=-0.999$	0.103	0.014	40.432	0.06
	9		$(2,1)(0,1) p1=1.98, p2=-0.999, q1=-0.999, Q1=-0.9$	0.025	0.017	55.161	0.32
			$(2,1)(1,0) p1=1.98, p2=-0.999, q1=-0.999, P1=0.999$	0.239	0.050	130.214	2.52
			$(2,0)(0,1) p1=1.98, p2=-0.99, Q1=-0.999$	0.078	0.018	55.114	0.32
			$(2,0)(1,0) p1=1.98, p2=-0.99, P1=0.999$	0.350	0.075	217.227	5.53
			$(2,0)(1,1) p1=1.98, p2=-0.99, P1=0.999, Q1=-0.999$	0.292	0.782	2173.865	77.37
			$(2,1)(1,1) p1=1.98, p2=-0.99, q1=-0.999, P1=0.999, Q1=-0.999$	0.214	0.463	1222.421	44.22
תוצר בענף מסחר	8	לוג-חיבורי עם חותר	$(2,1)(0,0) p1=1.98, p2=-0.999, q1=-0.999$	0.277	0.005	449.836	0.16
	9		$(2,1)(0,1) p1=1.98, p2=-0.999, q1=-0.999, Q1=-0.9$	0.119	0.007	454.408	0.27
			$(2,1)(1,0) p1=1.98, p2=-0.999, q1=-0.999, P1=0.999$	0.661	0.020	441.704	1.24
			$(2,0)(0,1) p1=1.98, p2=-0.99, Q1=-0.999$	0.169	0.007	454.363	0.27
			$(2,0)(1,0) p1=1.98, p2=-0.99, P1=0.999$	0.980	0.035	477.938	2.41
			$(2,0)(1,1) p1=1.98, p2=-0.99, P1=0.999, Q1=-0.999$	0.914	0.303	1996.542	29.52
			$(2,1)(1,1) p1=1.98, p2=-0.99, q1=-0.999, P1=0.999, Q1=-0.999$	0.639	0.169	1052.103	15.85
תוצר בענף מחשבים	8	לוג-חיבורי עם חותר	$(2,1)(0,0) p1=1.98, p2=-0.999, q1=-0.999$	0.150	0.029	361.893	1.01
	9		$(2,1)(0,1) p1=1.98, p2=-0.999, q1=-0.999, Q1=-0.9$	0.069	0.030	364.425	1.13
			$(2,1)(1,0) p1=1.98, p2=-0.999, q1=-0.999, P1=0.999$	0.336	0.077	440.685	3.31
			$(2,0)(0,1) p1=1.98, p2=-0.99, Q1=-0.999$	0.094	0.030	364.415	1.13
			$(2,0)(1,0) p1=1.98, p2=-0.99, P1=0.999$	0.492	0.131	520.238	3.93
			$(2,0)(1,1) p1=1.98, p2=-0.99, P1=0.999, Q1=-0.999$	0.461	0.225	1017.493	17.87
			$(2,1)(1,1) p1=1.98, p2=-0.99, q1=-0.999, P1=0.999, Q1=-0.999$	0.319	0.101	405.702	1.97

* סטיות התקן חושבו בהנחה כי בסדרות האינדיקטורים CV=1

לוח 2א: דיאגנוסטיקות של ניכוי עונתיות בסדרה "תוצר בענף תקשורת", לאחר ביון וחיזוי										
מודל 9	מודל 8	מודל 7	מודל 6	מודל 5	מודל 4	מודל 3	מודל 2	מודל 1	אינדיקטור	
0.0171	0.0143	0.0137	0.0195	0.0132	0.0138	0.0142	0.0132	0.0132	---	דיאגנוסטיקה לשמירת תנודות Cm במנוכה
0.0150	0.0128	0.0126	0.0121	0.0121	0.0125	0.0127	0.0124	0.0124	---	דיאגנוסטיקה לשמירת תנודות Cm במגמה
127.31	155.12	154.07	55.60	163.56	162.13	156.23	165.94	165.04	155.00	סטטיסטי Chi-Square למובהקות הרגרסיה
מודל של ARIMA X12										
(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)
6.59	5.69	4.62	4.36	5.10	5.32	5.64	5.02	5.13	6.92	מודל ARIMA ממוצע טעויות בערך מוחלט של תחזויות בתוך המדגם: שלוש שנים אחרונות
66.74	52.90	35.66	49.71	44.53	38.05	50.70	34.95	38.21	27.82	הסתברות Chi בריבוע
טיב ניכוי עונתיות										
17.06	21.73	22.33	8.98	25.43	24.07	22.00	24.79	24.49	25.60	מבחן F לעונתיות (לוח F8B)
2.12	2.08	2.15	1.63	2.10	2.09	2.08	2.05	2.05	2.07	סטית תקן של גורמים עונתיים (לוח D10)
0.98	0.73	0.82	0.96	0.72	0.78	0.82	0.72	0.72	0.77	סטית תקן של גורמים בלתי סדירים (לוח D13)
תרומות יחסיות של:										
2.14	1.89	1.70	13.73	2.22	1.93	1.89	2.25	2.25	1.63	I - אי-סדירות (לוח E3)
37.72	29.65	25.23	34.06	25.80	27.50	29.25	25.98	26.03	37.90	C - מגמה-מחזור (לוח D12)
35.32	39.81	44.90	40.19	42.65	41.50	40.16	43.18	43.16	35.38	S - עונתיות (לוח D10)
24.83	28.65	28.16	12.02	29.33	29.07	28.70	28.59	28.56	25.09	P - השפעת הגים יומי פעילות (לוח D18)
1	1	1	1	1	1	1	1	1	1	QCD
0.43	0.42	0.45	0.83	0.42	0.42	0.42	0.42	0.42	0.41	Q

לוח 2ב: דיאגנוסטיקות של ניכוי עונתיות בסדרה "תוצר בענף מסחר", לאחר ביון וחיזוי										
מודל 9	מודל 8	מודל 7	מודל 6	מודל 5	מודל 4	מודל 3	מודל 2	מודל 1	אינדיקטור	
0.0070	0.0051	0.0046	0.0062	0.0044	0.0045	0.0050	0.0044	0.0044	---	דיאגנוסטיקה לשמירת תנודות Cm במנוכה
0.0067	0.0049	0.0042	0.0040	0.0042	0.0043	0.0048	0.0041	0.0041	---	דיאגנוסטיקה לשמירת תנודות Cm במגמה
212.18	224.09	223.35	183.47	229.32	219.24	231.87	219.00	219.36	229.44	סטטיסטי Chi-Square למובהקות הרגרסיה
מודל של ARIMA X12										
(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)
1.28	1.23	1.39	1.57	1.26	1.26	1.22	1.28	1.29	1.48	מודל ARIMA ממוצע טעויות בערך מוחלט של תחזויות בתוך המדגם: שלוש שנים אחרונות
27.88	10.22	6.74	6.15	7.50	6.61	9.51	7.10	6.98	18.86	הסתברות Chi בריבוע
טיב ניכוי עונתיות										
100.91	108.36	123.77	95.86	113.49	113.05	108.26	108.82	109.27	108.48	מבחן F לעונתיות (לוח F8B)
2.16	2.15	2.21	2.37	2.15	2.19	2.15	2.16	2.16	2.14	סטית תקן של גורמים עונתיים (לוח D10)
0.56	0.54	0.54	0.70	0.53	0.55	0.54	0.54	0.54	0.54	סטית תקן של גורמים בלתי סדירים (לוח D13)
תרומות יחסיות של:										
2.81	2.08	4.38	5.34	4.11	4.34	2.04	4.36	4.36	3.16	I - אי-סדירות (לוח E3)
23.01	22.10	19.00	16.70	20.42	20.18	21.89	20.17	20.26	16.82	C - מגמה-מחזור (לוח D12)
53.90	54.67	56.99	53.44	54.81	55.17	54.84	54.78	54.72	57.62	S - עונתיות (לוח D10)
20.28	21.15	19.62	24.52	20.65	20.30	21.23	20.70	20.66	22.40	P - השפעת הגים יומי פעילות (לוח D18)
1	1	1	1	1	1	1	1	1	1	QCD
0.25	0.24	0.29	0.43	0.27	0.28	0.23	0.25	0.25	0.25	Q

לוח 2ג: דיאגנוסטיקות של ניכוי עונתיות בסדרה "תוצר בענף מחשבים", לאחר ביון וחיזוי										
מודל 9	מודל 8	מודל 7	מודל 6	מודל 5	מודל 4	מודל 3	מודל 2	מודל 1	אינדיקטור	
0.0307	0.0289	0.0257	0.0389	0.0269	0.0276	0.0289	0.0262	0.0267	---	דיאגנוסטיקה לשמירת תנודות Cm במנוכה
0.0294	0.0301	0.0267	0.0264	0.0281	0.0292	0.0300	0.0289	0.0294	---	דיאגנוסטיקה לשמירת תנודות Cm במגמה
29.84	28.39	34.17	41.64	35.93	29.90	28.44	34.39	34.42	30.53	סטטיסטי Chi-Square למובהקות הרגרסיה
מודל של ARIMA X12										
(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)
7.62	8.05	8.61	9.16	8.80	7.90	8.01	7.89	7.88	8.02	מודל ARIMA ממוצע טעויות בערך מוחלט של תחזויות בתוך המדגם: שלוש שנים אחרונות
30.70	27.00	24.51	41.30	44.35	24.17	27.14	29.35	26.47	34.78	הסתברות Chi בריבוע
טיב ניכוי עונתיות										
39.93	38.72	31.92	19.62	36.27	37.96	38.82	41.73	42.01	40.38	מבחן F לעונתיות (לוח F8B)
9.90	9.83	10.10	9.07	10.13	9.72	9.83	10.19	10.19	9.89	סטית תקן של גורמים עונתיים (לוח D10)
4.06	4.05	4.32	4.19	4.12	4.01	4.05	3.87	3.86	3.83	סטית תקן של גורמים בלתי סדירים (לוח D13)
תרומות יחסיות של:										
5.35	3.36	4.40	11.43	4.20	3.78	3.36	4.16	4.17	6.49	I - אי-סדירות (לוח E3)
7.36	7.46	6.56	6.30	6.46	7.64	7.41	6.32	6.47	5.45	C - מגמה-מחזור (לוח D12)
80.28	82.08	81.59	70.09	82.27	80.29	82.11	82.45	82.34	80.63	S - עונתיות (לוח D10)
7.01	7.10	7.45	12.19	7.08	8.29	7.12	7.07	7.02	7.44	P - השפעת הגים יומי פעילות (לוח D18)
2	2	2	2	2	2	2	2	2	2	QCD
0.57	0.52	0.57	0.87	0.52	0.54	0.53	0.52	0.52	0.59	Q

לוח 3: סדרה תוצר בענף תקשורת - חיוץ של שנת 2006

שנה	רבע	מודל 1			מודל 2			מודל 3			מודל 4			מודל 5			מודל 7		
		ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט
1998	1	1930.5	1930.5	0.0	1931.5	1931.4	0.1	1947.4	1940.8	6.6	1946.8	1947.2	0.4	1899.5	1897.4	2.1	1897.4	1888.7	8.7
	2	2020.0	2020.0	0.0	2020.3	2020.2	0.1	2020.4	2020.1	0.3	2022.5	2024.1	1.6	2018.6	2018.5	0.1	2018.0	2019.3	1.3
	3	2044.5	2044.5	0.0	2044.1	2044.1	0.0	2036.1	2039.2	3.1	2040.1	2039.7	0.4	2059.9	2060.9	1.0	2064.0	2067.8	3.8
	4	2082.1	2082.1	0.0	2081.3	2081.3	0.0	2073.2	2076.9	3.7	2067.7	2066.1	1.6	2099.1	2100.4	1.3	2097.7	2101.3	3.6
1999	1	2086.7	2086.8	0.1	2086.3	2086.3	0.0	2085.6	2088.0	2.4	2077.7	2078.4	0.7	2097.4	2098.1	0.7	2070.1	2069.7	0.4
	2	2109.0	2109.0	0.0	2109.0	2109.0	0.0	2111.2	2111.8	0.6	2109.7	2111.5	1.8	2109.9	2110.0	0.1	2090.2	2088.4	1.8
	3	2244.2	2244.2	0.0	2244.5	2244.5	0.0	2245.6	2244.5	1.1	2252.7	2252.2	0.5	2239.3	2239.0	0.3	2256.7	2256.9	0.2
	4	2327.5	2327.5	0.0	2327.8	2327.7	0.1	2325.1	2323.3	1.8	2327.5	2325.4	2.1	2320.8	2320.3	0.5	2350.5	2352.6	2.1
2000	1	2220.3	2220.2	0.1	2220.1	2220.1	0.0	2215.9	2214.5	1.4	2216.8	2217.7	0.9	2216.5	2216.2	0.3	2228.3	2230.5	2.2
	2	2334.6	2334.6	0.0	2334.5	2334.4	0.1	2330.6	2330.0	0.6	2336.0	2338.3	2.3	2333.2	2333.1	0.1	2337.5	2339.9	2.4
	3	2497.1	2497.1	0.0	2497.2	2497.2	0.0	2497.9	2498.4	0.5	2503.9	2503.3	0.6	2498.2	2498.3	0.1	2497.0	2496.3	0.7
	4	2676.8	2676.8	0.0	2677.0	2677.0	0.0	2684.3	2685.8	1.5	2671.9	2669.4	2.5	2680.8	2681.1	0.3	2665.9	2661.9	4.0
2001	1	2620.2	2620.3	0.1	2620.2	2620.3	0.1	2626.1	2627.8	1.7	2613.1	2614.3	1.2	2621.1	2621.4	0.3	2611.1	2612.1	1.0
	2	2695.7	2695.9	0.2	2695.5	2695.6	0.1	2700.1	2701.1	1.0	2696.9	2699.8	2.9	2697.5	2697.6	0.1	2701.7	2708.3	6.6
	3	2757.9	2757.9	0.0	2757.8	2757.8	0.0	2755.7	2755.1	0.6	2767.2	2766.3	0.9	2758.4	2758.3	0.1	2765.5	2766.5	1.0
	4	2580.6	2580.4	0.2	2581.0	2580.8	0.2	2572.7	2570.4	2.3	2577.4	2574.1	3.3	2577.5	2577.2	0.3	2576.2	2567.6	8.6
2002	1	2775.9	2775.3	0.6	2778.0	2777.4	0.6	2775.9	2772.4	3.5	2769.3	2769.8	0.5	2778.8	2778.2	0.6	2780.2	2770.8	9.4
	2	2753.3	2752.8	0.5	2755.2	2754.7	0.5	2752.7	2750.1	2.6	2755.5	2758.1	2.6	2751.7	2751.3	0.4	2768.2	2768.0	0.2
	3	2818.4	2818.4	0.0	2817.9	2818.0	0.1	2818.4	2819.2	0.8	2828.5	2827.7	0.8	2815.7	2815.7	0.0	2815.2	2823.2	8.0
	4	2882.1	2883.2	1.1	2878.6	2879.6	1.0	2882.7	2888.0	5.3	2876.3	2874.1	2.2	2883.6	2884.5	0.9	2866.0	2867.7	1.7
2003	1	2718.0	2720.5	2.5	2714.3	2716.7	2.4	2705.9	2714.0	8.1	2704.1	2708.3	4.2	2713.2	2714.9	1.7	2718.8	2707.0	11.8
	2	2784.0	2786.1	2.1	2782.9	2785.0	2.1	2779.6	2786.0	6.4	2780.5	2786.7	6.2	2787.3	2789.0	1.7	2793.4	2783.7	9.7
	3	3085.1	3085.0	0.1	3086.8	3086.9	0.1	3090.1	3088.9	1.2	3097.9	3096.9	1.0	3089.6	3089.7	0.1	3080.7	3088.1	7.4
	4	3154.8	3150.3	4.5	3157.9	3153.3	4.6	3166.3	3152.9	13.4	3159.3	3149.9	9.4	3151.8	3148.2	3.6	3149.0	3163.0	14.0
2004	1	3124.5	3113.4	11.1	3126.5	3115.2	11.3	3147.4	3125.3	22.1	3121.1	3110.3	10.8	3122.8	3115.7	7.1	3131.7	3137.1	5.4
	2	3180.5	3170.7	9.8	3181.2	3171.3	9.9	3192.0	3174.4	17.6	3182.8	3176.1	6.7	3172.6	3165.9	6.7	3185.2	3184.9	0.3
	3	3407.4	3407.7	0.3	3406.7	3406.8	0.1	3399.7	3403.2	3.5	3417.5	3416.1	1.4	3403.9	3403.8	0.1	3397.2	3392.7	4.5
	4	3438.5	3459.2	20.7	3436.5	3457.6	21.1	3411.8	3448.1	36.3	3429.4	3448.4	19.0	3451.6	3465.5	13.9	3436.8	3436.2	0.6
2005	1	3286.1	3333.6	47.5	3284.1	3334.3	50.2	3244.8	3304.2	59.4	3258.6	3318.0	59.4	3308.4	3334.1	25.7	3310.4	3332.8	22.4
	2	3236.8	3275.3	38.5	3236.0	3277.3	41.3	3209.3	3261.0	51.7	3220.0	3273.4	53.4	3261.6	3284.3	22.7	3252.4	3288.3	35.9
	3	3345.8	3342.3	3.5	3346.5	3342.7	3.8	3353.2	3351.5	1.7	3356.9	3352.2	4.7	3347.6	3346.1	1.5	3325.3	3350.1	24.8
	4	3364.1	3281.6	82.5	3366.2	3278.4	87.8	3425.4	3315.9	109.5	3397.2	3289.1	108.1	3315.2	3268.2	47.0	3344.7	3261.4	83.3
2006	1	3830.9	3606.0	224.9	3834.5	3600.7	233.8	3979.9	3681.7	298.2	3911.4	3632.5	278.9	3705.1	3593.5	111.6	3848.6	3554.9	293.7
	2	3611.7	3307.9	303.8	3616.2	3305.9	310.3	3826.9	3356.5	470.4	3742.7	3338.5	404.2	3432.6	3291.7	140.9	3647.7	3292.6	355.1
	3	3783.1	3400.0	383.1	3789.0	3402.2	386.8	4086.2	3385.7	700.5	3982.0	3406.3	575.7	3554.1	3398.8	155.3	3824.3	3452.8	371.5
	4	3713.2	3305.9	407.3	3720.2	3311.1	409.1	4085.5	3196.0	889.5	3946.9	3242.7	704.2	3469.0	3335.9	133.1	3686.4	3319.5	366.9
סך העדכונים				1545.1			1577.7			2730.9			2277.1		682.3			1675.0	
סך העדכונים - 3 שנים				1319.1			1340.0			2358.6			1963.0		540.9			1387.2	

לוח 4: סדרה תוצר בענף מסחר - חיוץ של שנת 2006

שנה	רבע	מודל 1			מודל 2			מודל 3			מודל 4			מודל 5			מודל 7		
		ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט
1998	1	6469.9	6469.9	0.0	6462.1	6462.1	0.0	6400.0	6402.0	2.0	6438.4	6439.2	0.8	6479.7	6479.9	0.2	6446.3	6444.0	2.3
	2	6412.9	6412.9	0.0	6412.6	6412.6	0.0	6402.7	6402.8	0.1	6400.5	6400.1	0.4	6411.2	6411.2	0.0	6426.5	6426.8	0.3
	3	6709.0	6709.0	0.0	6712.8	6712.8	0.0	6740.6	6739.6	1.0	6720.4	6719.3	1.1	6703.0	6702.9	0.1	6714.6	6715.6	1.0
	4	6791.5	6791.5	0.0	6795.8	6795.8	0.0	6840.0	6838.9	1.1	6824.0	6824.7	0.7	6789.4	6789.3	0.1	6795.9	6796.8	0.9
1999	1	6728.8	6728.8	0.0	6730.4	6730.4	0.0	6759.6	6758.9	0.7	6744.4	6745.3	0.9	6725.6	6725.6	0.0	6722.0	6721.8	0.2
	2	6573.6	6573.6	0.0	6573.7	6573.7	0.0	6584.9	6584.8	0.1	6563.1	6562.6	0.5	6577.4	6577.4	0.0	6552.1	6551.1	1.0
	3	7359.7	7359.7	0.0	7359.0	7359.0	0.0	7347.6	7347.9	0.3	7347.2	7346.0	1.2	7363.4	7363.4	0.0	7352.8	7352.8	0.0
	4	7658.2	7658.2	0.0	7657.2	7657.2	0.0	7628.1	7628.7	0.6	7665.5	7666.4	0.9	7653.9	7653.9	0.0	7693.4	7694.6	1.2
2000	1	7195.9	7195.9	0.0	7195.8	7195.8	0.0	7179.9	7180.3	0.4	7202.7	7203.8	1.1	7197.2	7197.2	0.0	7238.1	7238.7	0.6
	2	7520.5	7520.5	0.0	7520.8	7520.8	0.0	7512.0	7512.2	0.2	7508.9	7508.3	0.6	7515.9	7515.9	0.0	7523.2	7522.8	0.4
	3	8192.0	8192.0	0.0	8192.1	8192.1	0.0	8198.7	8198.5	0.2	8183.7	8182.2	1.5	8189.2	8189.2	0.0	8159.6	8159.2	0.4
	4	7762.6	7762.6	0.0	7762.2	7762.2	0.0	7780.3	7779.9	0.4	7775.6	7776.7	1.1	7768.6	7768.6	0.0	7749.9	7750.2	0.3
2001	1	7774.6	7774.6	0.0	7775.2	7775.1	0.1	7774.2	7773.7	0.5	7784.2	7785.4	1.2	7769.9	7769.9	0.0	7780.0	7780.7	0.7
	2	7679.3	7679.3	0.0	7680.5	7680.4	0.1	7682.3	7682.0	0.3	7669.0	7668.2	0.8	7684.9	7684.9	0.0	7680.2	7680.2	0.0
	3	7941.6	7941.6	0.0	7941.8	7941.8	0.0	7942.6	7942.8	0.2	7933.4	7931.7	1.7	7946.4	7946.4	0.0	7946.1	7945.8	0.3
	4	7796.8	7796.8	0.0	7794.9	7795.0	0.1	7793.1	7793.7	0.6	7805.7	7806.9	1.2	7791.1	7791.1	0.0	7786.0	7785.7	0.3
2002	1	7433.6	7433.6	0.0	7431.2	7431.4	0.2	7449.2	7450.1	0.9	7437.0	7438.5	1.5	7440.2	7440.3	0.1	7384.8	7383.7	1.1
	2	7336.5	7336.6	0.1	7335.9	7336.0	0.1	7341.3	7341.9	0.6	7325.2	7324.6	0.6	7334.4	7334.4	0.0	7304.6	7304.2	0.4
	3	7443.1	7443.1	0.0	7444.3	7444.2	0.1	7435.5	7435.3	0.2	7438.7	7436.9	1.8	7439.1	7439.1	0.0	7476.9	7477.8	0.9
	4	7410.6	7410.5	0.1	7412.4	7412.1	0.3	7397.8	7396.5	1.3	7423.0	7423.9	0.9	7410.1	7410.0	0.1	7457.5	7458.1	0.6
2003	1	7150.1	7149.8	0.3	7151.0	7150.3	0.7	7131.0	7129.0	2.0	7155.3	7156.0	0.7	7143.8	7143.7	0.1	7164.3	7163.6	0.7
	2	7155.6	7155.4	0.2	7155.9	7155.4	0.5	7149.9	7148.5	1.4	7145.7	7144.2	1.5	7155.1	7154.9	0.2	7155.9	7155.5	0.4
	3	7755.2	7755.2	0.0	7754.9	7754.9	0.0	7764.7	7765.0	0.3	7750.1	7748.1	2.0	7758.0	7758.0	0.0	7750.3	7751.1	0.8
	4	7772.3	7772.8	0.5	7771.5	7772.6	1.1	7787.7	7790.7	3.0	7782.1	7784.9	2.8	7776.3	7776.6	0.3	7762.8	7763.0	0.2
2004	1	7725.7	7726.9	1.2	7724.3	7727.1	2.8	7735.7	7740.6	4.9	7727.6	7732.7	5.1	7729.5	7730.1	0.6	7719.5	7718.6	0.9
	2	7604.4	7605.4	1.0	7603.3	7605.7	2.4	7605.3	7609.1	3.8	7592.3	7594.2	1.9	7606.5	7607.0	0.5	7607.0	7607.4	0.4
	3	8186.4	8186.4	0.0	8186.5	8186.5	0.0	8180.7	8179.9	0.8	8181.6	8179.2	2.4	8185.6	8185.6	0.0	8192.9	8193.8	0.9
	4	8481.8	8479.5	2.3	8484.2	8479.0	5.2	8476.6	8468.7	7.9	8496.7	8492.2	4.5	8476.7	8475.6	1.1	8479.0	8478.4	0.6
2005	1	8082.5	8077.1	5.4	8088.0	8075.8	12.2	8087.3	8074.2	13.1	8096.8	8084.1	12.7	8077.7	8075.5	2.2	8073.0	8074.0	1.0
	2	8344.6	8339.7	4.9	8349.6	8338.6	11.0	8352.5	8340.5	12.0	8343.8	8329.9	13.9	8339.5	8337.4	2.1	8338.2	8343.9	5.7
	3	9095.4	9095.5	0.1	9095.3	9095.6	0.3	9097.6	9097.7	0.1	9090.5	9088.3	2.2	9095.2	9095.2	0.0	9087.9	9096.0	8.1
	4	8699.9	8710.1	10.2	8689.5	8712.5	23.0	8685.0	8710.0	25.0	8691.2	8720.2	29.0	8709.9	8714.2	4.3	8723.4	8708.4	15.0
2006	1	9030.9	9056.5	25.6	9004.6	9062.7	58.1	8986.6	9049.2	62.6	8987.6	9058.8	71.2	9051.7	9061.7	10.0	9132.5	9073.9	58.6
	2	9083.0	9120.2	37.2	9041.0	9125.5	84.5	9007.9	9114.4	106.5	8997.6	9106.1	108.5	9112.7	9126.7	14.0	9221.9	9143.0	78.9
	3	9393.9	9440.3	46.4	9334.5	9440.1	105.6	9283.2	9441.5	158.3	9282.7	9433.2	150.5	9425.5	9441.2	15.7	9525.3	9443.7	81.6
	4	9561.6	9612.8	51.2	9485.0	9601.6	116.6	9413.9	9624.6	210.7	9434.1	9631.7	197.6	9586.1	9600.2	14.1	9653.5	9569.3	84.2
סך העדכונים				186.7			425.0			624.1			627.0			65.8			350.9
סך העדכונים - 3 שנים				160.4			364.8			538.1			527.8			53.8			303.3

לוח 5: סדרה תוצר בענף מחשבים - חיוץ של שנת 2006

שנה	רבע	מודל 1			מודל 2			מודל 3			מודל 4			מודל 5			מודל 7		
		ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט	ללא סה"כ 2006	עם סה"כ 2006	הפרש בערך מוחלט
1998	1	1666.0	1666.0	0.0	1685.3	1685.3	0.0	1524.5	1527.1	2.6	1607.7	1607.6	0.1	1800.3	1806.7	6.4	1764.8	1773.3	8.5
	2	2456.7	2456.7	0.0	2472.1	2472.1	0.0	2422.5	2423.1	0.6	2423.8	2423.8	0.0	2486.1	2487.0	0.9	2529.4	2528.0	1.4
	3	2020.0	2020.0	0.0	2013.7	2013.7	0.0	2076.8	2075.7	1.1	2058.2	2058.2	0.0	1958.3	1955.6	2.7	1967.3	1965.0	2.3
	4	2806.7	2806.7	0.0	2778.2	2778.2	0.0	2925.6	2923.5	2.1	2859.7	2859.8	0.1	2704.7	2700.1	4.6	2687.9	2683.2	4.7
1999	1	2374.8	2374.8	0.0	2355.5	2355.5	0.0	2414.8	2413.7	1.1	2427.6	2427.6	0.0	2304.8	2302.6	2.2	2350.5	2352.9	2.4
	2	2614.9	2614.9	0.0	2614.0	2614.0	0.0	2626.8	2626.5	0.3	2588.8	2588.7	0.1	2614.0	2613.7	0.3	2644.6	2648.5	3.9
	3	2755.7	2755.7	0.0	2767.5	2767.5	0.0	2740.5	2741.0	0.5	2750.7	2750.6	0.1	2795.4	2796.4	1.0	2770.6	2770.8	0.2
	4	3073.8	3073.8	0.0	3082.2	3082.2	0.0	3037.0	3038.0	1.0	3052.2	3052.3	0.1	3105.0	3106.5	1.5	3053.5	3047.0	6.5
2000	1	2550.5	2550.5	0.0	2544.6	2544.6	0.0	2557.3	2557.9	0.6	2573.1	2573.1	0.0	2576.7	2577.5	0.8	2493.8	2488.7	5.1
	2	3141.1	3141.1	0.0	3134.5	3134.5	0.0	3151.9	3152.3	0.4	3098.8	3098.7	0.1	3140.2	3140.4	0.2	3076.3	3079.4	3.1
	3	3033.6	3033.6	0.0	3035.4	3035.4	0.0	3029.9	3029.7	0.2	3038.2	3038.2	0.0	3019.4	3019.1	0.3	3044.5	3050.4	5.9
	4	3355.2	3355.2	0.0	3365.8	3365.8	0.0	3341.2	3340.4	0.8	3370.3	3370.4	0.1	3344.0	3343.3	0.7	3465.7	3461.8	3.9
2001	1	3424.2	3424.2	0.0	3437.8	3437.8	0.0	3406.6	3405.6	1.0	3490.7	3490.6	0.1	3419.7	3418.8	0.9	3584.4	3574.9	9.5
	2	3246.0	3246.0	0.0	3252.7	3252.7	0.0	3238.4	3237.8	0.6	3215.8	3215.7	0.1	3247.5	3247.0	0.5	3309.7	3307.8	1.9
	3	3301.8	3301.8	0.0	3297.7	3297.7	0.0	3309.2	3309.4	0.2	3287.5	3287.5	0.0	3302.9	3302.9	0.0	3204.8	3208.6	3.8
	4	3966.4	3966.4	0.0	3950.2	3950.2	0.0	3984.3	3985.6	1.3	3944.5	3944.6	0.1	3968.4	3969.7	1.3	3839.5	3847.1	7.6
2002	1	4276.2	4276.3	0.1	4259.7	4259.8	0.1	4261.8	4264.0	2.2	4315.8	4315.8	0.0	4255.4	4258.1	2.7	4311.0	4325.5	14.5
	2	3951.6	3951.6	0.0	3946.8	3946.9	0.1	3946.2	3947.7	1.5	3904.7	3904.6	0.1	3948.7	3951.2	2.5	4024.2	4029.1	4.9
	3	3854.2	3854.2	0.0	3860.3	3860.3	0.0	3861.0	3860.7	0.3	3848.7	3848.7	0.0	3863.9	3864.2	0.3	3822.2	3814.4	7.8
	4	5014.0	5013.9	0.1	5029.2	5029.1	0.1	5027.1	5023.6	3.5	5026.8	5026.9	0.1	5028.0	5022.5	5.5	4938.6	4926.9	11.7
2003	1	3850.0	3849.8	0.2	3855.7	3855.6	0.1	3870.3	3865.8	4.5	3904.8	3904.8	0.0	3868.5	3859.7	8.8	3837.0	3838.2	1.2
	2	4013.6	4013.5	0.1	4014.7	4014.6	0.1	4022.4	4018.6	3.8	3978.0	3977.9	0.1	4021.6	4012.6	9.0	4011.3	4016.7	5.4
	3	4051.2	4051.2	0.0	4049.0	4049.0	0.0	4043.8	4044.1	0.3	4038.2	4038.2	0.0	4046.5	4045.4	1.1	4051.2	4050.7	0.5
	4	5148.2	5148.5	0.3	5143.7	5143.9	0.2	5126.4	5134.5	8.1	5141.9	5142.0	0.1	5126.3	5145.3	19.0	5163.4	5157.4	6.0
2004	1	3908.9	3909.5	0.6	3906.8	3907.2	0.4	3888.4	3899.4	11.0	3952.7	3952.6	0.1	3880.4	3910.7	30.3	3896.3	3896.3	0.0
	2	4571.9	4572.6	0.7	4571.4	4571.8	0.4	4559.2	4570.3	11.1	4535.9	4535.7	0.2	4545.3	4580.7	35.4	4528.6	4537.9	9.3
	3	4525.0	4525.1	0.1	4525.9	4526.0	0.1	4530.1	4530.4	0.3	4516.7	4516.7	0.0	4524.0	4530.3	6.3	4529.6	4529.8	0.2
	4	6089.5	6088.1	1.4	6091.4	6090.4	1.0	6117.7	6095.3	22.4	6090.0	6090.4	0.4	6145.6	6073.8	71.8	6140.9	6131.4	9.5
2005	1	4189.3	4186.6	2.7	4189.4	4187.5	1.9	4221.0	4191.5	29.5	4231.8	4232.1	0.3	4271.3	4164.1	107.2	4157.8	4163.6	5.8
	2	4686.9	4684.1	2.8	4686.4	4684.4	2.0	4716.1	4683.5	32.6	4651.3	4651.6	0.3	4776.9	4655.4	121.5	4620.3	4629.4	9.1
	3	4785.0	4784.5	0.5	4784.6	4784.2	0.4	4787.5	4780.7	6.8	4774.8	4774.9	0.1	4802.4	4778.3	24.1	4842.4	4793.6	48.8
	4	6715.2	6721.2	6.0	6716.0	6720.3	4.3	6651.7	6720.6	68.9	6718.4	6717.8	0.6	6525.7	6778.6	252.9	6755.9	6789.9	34.0
2006	1	5143.5	5156.8	13.3	5146.2	5155.7	9.5	5023.6	5164.9	141.3	5204.5	5202.4	2.1	4785.1	5248.7	463.6	4834.9	5165.5	330.6
	2	5688.6	5710.8	22.2	5694.0	5709.9	15.9	5452.5	5724.5	272.0	5675.3	5671.5	3.8	5082.2	5807.1	724.9	5146.0	5705.7	559.7
	3	5656.0	5683.1	27.1	5663.7	5683.2	19.5	5296.4	5687.9	391.5	5677.9	5673.0	4.9	4911.0	5687.3	776.3	5190.4	5739.3	548.9
	4	6444.8	6478.8	34.0	6456.3	6480.6	24.3	5871.4	6452.2	580.8	6489.3	6482.5	6.8	5512.3	6286.4	774.1	5840.8	6419.0	578.2
סך העדכונים				112.2			80.4		1606.8		21.0		3461.6			2256.8			
סך העדכונים - 2006				96.6			69.2		1385.6		17.6		2738.9			2017.4			

לוח 6: מדדי ביון, חיזוי וניכוי עונתיות בסדרות (1) - (3) תוך שימוש במודל ARMA הנאמד			
שם סדרה	(1) תוצר בענף תקשורת	(2) תוצר בענף מסחר	(3) תוצר בענף מחשבים
מדדים של ביון וחיזוי			
מודל ARMA לשאריות שנבחר	(2 1)(1 0)	(2 1)(0 1)	(2 0)(0 1)
מקדמים של המודל	$p_1=1.57^*$, $p_2=-0.67^*$, $q_1=-0.67^*$, $P_1=-0.14$	$p_1=1.05^*$, $p_2=-0.78^*$, $q_1=0.11$, $Q_1=-0.49^*$	$p_1=1.77^*$, $p_2=-0.77^*$, $Q_1=-0.055^*$
דיאגנוסטיקה לשמירת תנודות Cm	0.013	0.007	0.027
הפרש בין מדדי חלקיקות של הסדרה הנאמדת ובאינדיקטור, בערך מוחלט	0.062	0.054	0.949
סטיית תקן ממוצעת	5.194	55.775	0.944
סכום העדכונים בהרצה עד 2006, ביחס להרצה ללא הנתון השנתי האחרון	1157.8	213.6	418.1
סכום העדכונים בשלום שנים האחרונות בהרצה עד 2006, ביחס להרצה ללא הנתון השנתי האחרון	1106.6	197.7	409.8
סכום העדכונים בשנה האחרונה (בתקופת החיזוי) בהרצה עד 2006, ביחס להרצה ללא הנתון השנתי האחרון	906.0	146.5	356.5
מודל רגרסיה לאמידת גורמי חג ימי פעילות			
סטטיסט Chi-Square למובהקות מודל רגרסיה	162.46	226.21	29.50
מודל ARIMA של X12			
מודל ARIMA	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)
ממוצע טעויות בערך מוחלט של תחזויות בתוך המדגם: שלוש שנים אחרונות	5.40	1.72	7.89
הסתברות Chi בריבוע	49.01	27.71	27.17
טיב ניכוי עונתיות			
מבחן F לעונתיות (לוח F8B)	23.85	99.40	39.72
סטיית תקן של גורמים עונתיים (לוח D10)	2.06	2.07	9.90
סטיית תקן של גורמים בלתי סדירים (לוח D13)	0.75	0.57	3.99
תרומות יחסיות של:			
I - אי-סדירות (לוח E3)	2.25	2.32	4.31
C - מגמה-מחזור (לוח D12)	27.44	21.86	7.29
S - עונתיות (לוח D10)	40.64	54.20	81.29
P - השפעת חגים ימי פעילות (לוח D18)	29.67	21.62	7.12
QCD	1	1	1
Q	0.45	0.23	0.54

* - מקדם ARMA מובהק ברמת מובהקות של 5%

לוח 7: סך העדכונים בתקופת החיזוי (שנה אחרונה) בסדרות (1) - (3) תוך שימוש בשיטת הרגרסיה עם מודל ARMA הנאמד			
שם סדרה	(1) תוצר בענף תקשורת	(2) תוצר בענף מסחר	(3) תוצר בענף מחשבים
תקופות ההרצות המשוות			
עד IV.2001 - עם וללא נתון שנתי של 2001	430.8	823.1	2119.1
עד IV.2002 - עם וללא נתון שנתי של 2002	587.7	591.4	1718.7
עד IV.2003 - עם וללא נתון שנתי של 2003	606.6	472.5	739.2
עד IV.2004 - עם וללא נתון שנתי של 2004	122.7	357.8	991.5
עד IV.2005 - עם וללא נתון שנתי של 2005	882.2	352.6	294.9
עד IV.2006 - עם וללא נתון שנתי של 2006	906.0	146.5	356.5

לוח 8: מדד טיב הביון והחיזוי בסדרות (4) - (10) תוך שימוש בשיטת הרגרסיה עם מודל ARMA הנאמד						
שם סדרה	מודל ARMA הנבחר	מקדמים של מודל ARMA הנבחר	מדד כפלי לשמירת תנודות	הפרש בין מדדי חלקיקות של הסדרה הנאמדת ובאינדיקטור, בערך מוחלט	סטיית תקן ממוצעת	סך העדכונים בשנה האחרונה - בהרצות עד IV.2006 עם וללא נתון שנתי של 2006
(4) תמורה למשרות בתעשייה	(1 0)(0 0)	$p_1=0.143$	0.0028	0.0472	88.931	560.4
(5) שעות עבודה למעשה בתעשייה	(1 0)(0 0)	$p_1=0.135$	0.0031	0.0887	1462.388	7614.7
(6) מספר משרות בתעשייה	(2 0)(0 0)	$p_1=0.596^*$, $p_2=-0.289^*$	0.0023	0.0127	2.777	3.6
(7) מספר משרות לשכיר בתעשייה	(2 0)(0 0)	$p_1=0.528^*$, $p_2=-0.302^*$	0.0022	0.0077	2.748	4.4
(8) תוצר תעשייתי	(2 0)(0 0)	$p_1=0.218$, $p_2=-0.151$	0.0072	0.1552	148.913	2861.8
(9) פדיון תעשייתי	(0 1)(0 0)	$q_1=-0.187$	0.0001	0.0031	289.325	118.4
(10) תוצר תעשייתי הענף כרייה וחציבה	(2 0)(0 0)	$p_1=1.613^*$, $p_2=-0.637^*$	0.0226	0.4064	0.546	56.8

* - מקדם ARMA מובהק ברמת מובהקות של 5%

לוח 9: היקף העדכונים וטיב החיוץ במודלים נבחרים לסדרות חודשיות של סכ"א

סדרה	מס' מודל	מודל ARMA לשאריות	עדכונים*	טעות חיוץ**
נשים מועסקות	1S	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	36897	31326
	3	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	3952	3310
	4	(1,0)(1,0) $p_1=0.999, P_1=0.9$	2951	2364
	8	(2,1)(0,0) $p_1=1.98, p_2=-0.999, q_1=-0.999$	4148	3512
	9	(2,1)(0,1) $p_1=1.98, p_2=-0.99, q_1=-0.999, Q_1=0.9$	4783	4059
	10	(2,0)(0,1) $p_1=1.98, p_2=-0.99, Q_1=0.999$	5765	3331
גברים מועסקים	1S	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	14048	11924
	3	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	1330	1081
	4	(1,0)(1,0) $p_1=0.999, P_1=0.9$	667	514
	8	(2,1)(0,0) $p_1=1.98, p_2=-0.999, q_1=-0.999$	1344	1136
	9	(2,1)(0,1) $p_1=1.98, p_2=-0.99, q_1=-0.999, Q_1=0.9$	813	690
	10	(2,0)(0,1) $p_1=1.98, p_2=-0.99, Q_1=0.999$	2687	1024
נשים בלתי מועסקות	1S	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	29964	25260
	3	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	2988	2475
	4	(1,0)(1,0) $p_1=0.999, P_1=0.9$	1612	801
	8	(2,1)(0,0) $p_1=1.98, p_2=-0.999, q_1=-0.999$	3342	2789
	9	(2,1)(0,1) $p_1=1.98, p_2=-0.99, q_1=-0.999, Q_1=0.9$	6411	5202
	10	(2,0)(0,1) $p_1=1.98, p_2=-0.99, Q_1=0.999$	9354	4833
גברים בלתי מועסקים	1S	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	1312	1114
	3	(2,0)(0,0) $p_1=1.98, p_2=-0.99$	156	108
	4	(1,0)(1,0) $p_1=0.999, P_1=0.9$	941	423
	8	(2,1)(0,0) $p_1=1.98, p_2=-0.999, q_1=-0.999$	167	141
	9	(2,1)(0,1) $p_1=1.98, p_2=-0.99, q_1=-0.999, Q_1=0.9$	3160	2642
	10	(2,0)(0,1) $p_1=1.98, p_2=-0.99, Q_1=0.999$	5847	2543

* עדכונים חושבו כסכום הפרשים בערך מוחלט בין הסדרות מהרצות עם וללא הנתון של קבע רביעי של שנת 2007
 ** טעות החיוץ חושבה כהיקף העדכונים בתקופת החיוץ בלבד, דהיינו ברבע רביעי של שנת 2007.

Abstract

In recent years, there is an increasing demand for reliable high-frequency socio-economic series. However, in many cases in official statistics, low-frequency data is available. To estimate a series with desired frequency, it is necessary to use other time series as indicators, where those series are of high frequency, and describe socio-economic phenomena which is similar to the original low-frequency series. For this purpose, interpolation and extrapolation procedures should be applied. In most cases, the numerical Denton approach or its approximations are used.

The current paper deals with the considerations underlying the selection of indicators for the estimated series. A simple descriptive method of examination is proposed, which responds to the question regarding the extent of similarity between the indicator and the original low-frequency time series.

Using the Cholette and Dagum (1991) regression framework, a number of stochastic methods for interpolation and extrapolation of time series were studied.

A comprehensive analysis of wide class of ARMA moving preservation models was carried out. It was shown, that for each time series, the appropriate moving preservation model can be fitted. Using one model for all series may increase revisions in the estimated series and decrease the quality of the seasonal adjustment procedure. A stochastic method for interpolation and extrapolation, based on estimation of the residual distribution from the available data was proposed. Goodness of fit statistics, calculated for the proposed stochastic method, for the method based on the ARMA moving preservation model, and for the traditional Denton method were compared. It was shown, that applying the stochastic methods significantly increases the quality of seasonal adjustment procedure, decreases revisions and improves moving preservation in the estimated series.

A regression model with the non-constant intercept (Dagum and Cholette, 2006) was proposed for interpolation and extrapolation of low frequency macro-economic data, where one figure covers a number of years

Key words: Interpolation, extrapolation, regression method, ARMA moving preservation methods.

**Published by the Central Bureau of Statistics, 66 Kanfe Nesharim St.,
Corner Bachi St., P.O.B 34525, Jerusalem 91342, Israel
Tel. 972-2-6592666; Fax: 972-2-6521340
Internet Site: www.cbs.gov.il
E-Mail: info@cbs.gov.il**

The Central Bureau of Statistics (CBS) encourages research based on CBS data. Publications of this research are not official publications of the CBS, and they have not undergone the review accorded official CBS publications. The opinions and conclusions expressed in these publications, including this one, are those of the authors and do not necessarily represent those of the CBS. Permission for republication in whole or part must be obtained from the authors.

WORKING PAPER SERIES

NO.44

Using stochastic methods for
interpolation and extrapolation of time series

Yury Gubman*

May 2009

*Central Bureau of Statistics – Chief Scientist's Department