

סטטיסטיקה של הכנסות ממקורות מנהליים:

תיאור בסיסי הנתונים

פורמן אורלי

תחום מחקר כלכלי, אגף מדען ראשי

יולי 2007

תוכן עניינים

רקע	
מבוא	
1. קובץ מס הכנסה שכירים	
1.1 תיאור הקובץ	
1.1.1 נתוני הכנסה ומס בקובץ מס הכנסה שכירים	
1.1.2 תיאור רשומה	
1.1.3 סוגי המשרות	
1.2 השלמה, טיוב ובדיקות	
1.3 קבצים קיימים	
2. קובץ מס הכנסה עצמאים	
2.1 תיאור הקובץ	
2.2 תיאור רשומה	
2.3 קבצים קיימים	
2.4 השלמה, טיוב ובדיקות	
2.4.1 בדיקת אוכלוסיית העצמאים	
2.4.2 בדיקת גובה והרכב ההכנסות	
2.4.3 בדיקת רכיב השכר	
3. בסיס נתונים מזווג עובד-מעביד	
3.1 העיקרון	
3.2 המבנה	
3.2.1 מבנה טבלאות עובד	
3.2.1.1 מאפיינים קבועים	
3.2.1.2 מאפיינים שוטפים	
3.2.2 מבנה טבלאות מעביד	
3.2.2.1 מאפיינים קבועים	
3.2.2.2 מאפיינים שוטפים	
4. בסיס נתוני הכנסות פרט	
4.1 תהליך הבנייה	
4.2 מבנה בסיס הנתונים	
5. הצעה לפרסום "סטטיסטיקה של הכנסות ממקורות מנהליים"	
5.1 מהקיים בעולם	
5.2 הגדרות והסברים	

תודות

לדרי **דימיטרי רומנוב** על הליווי, ההכוונה וההארה לאורך כל הפרוייקט.
ליורי חומנקו על התמיכה והסיוע הטכני.

רקע

השימוש בקבצים מנהליים כמקור נתונים למטרות סטטיסטיות צובר תאוצה בשנים האחרונות וזוהי מגמה בינלאומית.¹ קבצים מנהליים הם קבצי נתונים הנאספים על ידי גורם שלישי (ספק הנתונים), כאשר הנתונים הנאספים נועדו בראש ובראשונה לשרת את מטרות הגוף המנהל אותם ולא למטרות סטטיסטיות. יתרונות חשובים גלומים בשימוש בקבצים מנהליים למטרות סטטיסטיות:

- חסכון בעלויות הנובעות מתכנון וביצוע סקרים, וכן בעיבוד הנתונים, שכן לטיפול בקובץ מנהלי נדרש צוות מצומצם יותר יחסית לצוות הנדרש להפקת אומדני הסקר.
- הפחתת נטל ההשבה.
- כיסוי גדול יותר של אוכלוסיה – למעשה מדובר ב"מרשם" של כל האוכלוסיה הרלוונטית, שמאפשר חקירה של קבוצות קטנות יותר (שלא תמיד ניתן לבצע מתוך נתוני הסקר המדגמי).
- ניצול מקורות מידע קיימים מאפשר הפקה מהירה של סטטיסטיקה לעומת תהליך עריכת סקר, המתפרש על חודשים ארוכים ואף שנים.
לצד יתרונות אלה עומדים גם חסרונות:
- נתונים מנהליים לא יכולים להוות תחליף מושלם לסקרים.
- ככלל, אין הגדרה סטטיסטית רצויה של האוכלוסיה המכוסה על ידי מקור מינהלי, ואין בקרה על תהליך איסוף הנתונים ואיכותם.
- אין ללמ"ס שליטה בתוכן המאגר ומענה לשאלות מחקריות מתבקשות (שאלות הסקר מנוסחות לפי רצוננו, שדות קובץ מנהלי לא).
- קבצים מנהליים לא יכולים לעמוד ברשות עצמם בלבד, אלא יכולים לסייע בהשלמת התמונה בצורה מקיפה יותר, בין אם במסגרת שילוב של מספר קבצים מנהליים ממקורות שונים, ובין אם כתוספת לנתוני סקרים.
דיון מקיף במעלות ומגרעות של השימוש בנתונים מנהליים נמצא במאמרו של יצחקי (2007).
בסיכום הדיון הפורה בסוגיות אלה בסמינר הסטטיסטיקאים האירופאים נאמר (הדגשה שלנו):

"The Conference discussed efficiency in terms of trade-off between outputs and resources used. The session dealt with some organizational and methodological approaches to generating efficiency, such as centralisation versus decentralisation, and active management of non-response. **The use of administrative data was considered as a way to reduce costs. However, the related problems need to be carefully managed and the quality has to be continuously monitored.** The session also considered the use of new, more efficient methods to increase the cost-efficiency of statistical production like record linkage, standardisation of corporate methods, reorganization of IT, and redesigning of business processes.

Furthermore, the following points were made:

¹ ראייה לכך היא דיון בשימוש במאגרי מידע מינהליים בנושא מרכזי בסמינר הסטטיסטיקאים האירופאים השנה. ישראל תרמה רבות לדיון זה (יצחקי, 2007).

- (a) data collection is the most expensive stage of the statistical cycle, so its optimization is vital;
- (b) analysis of the costs of all phases of statistical production is an important tool for decision-making;
- (c) not only costs of the statistical office should be taken into account but also those of the respondents;
- (d) in countries where high quality administrative registers exist and there is a positive public perception of their use in statistics, they are an efficient method of data collection;**
- (e) for some countries, the use of administrative data can be challenging. It may be difficult to ensure public compliance or overcome negative perceptions surrounding confidentiality of administrative data;**
- (f) it must also be borne in mind that the processing of data from administrative registers, quality control and assurance require specific efforts and constant development of personnel competence. An important factor for success is to maintain a continuous dialogue with agencies responsible for administrative data;**
- (g) linking or matching different administrative data sources can provide efficiency gains, but poses a high level of risk because of the related privacy and confidentiality concerns; one solution can be to mine statistical aggregates directly from registers instead of importing data;**
- (h) it was suggested to create a CES task force to focus on issues surrounding the use of administrative data;**
- (i) reengineering and standardising business processes is a valuable opportunity to achieve productivity gains and make the work processes more transparent;
- (j) the right skills and competencies of staff are a key to increasing efficiency and productivity."

(Source: UN Statistical Commission and Economic Commission for Europe, Conference of European Statisticians, Report of the fifty-fifth plenary session – ECE/CES/72).

בשנים האחרונות התקבלו בלמ"ס קובצי מס הכנסה שכירים ועצמאים, קבצים מנהליים שיעודם המקורי הוא לצורכי מיסוי, אולם נתונים אלה משמשים כמקור מידע ייחודי על הכנסות הפרט. העשרת קבצים אלה על ידי זיווג למקורות מנהליים נוספים הקיימים בלמ"ס, דוגמת מרשם התושבים ומרשם העסקים, הניבה מאגר נתונים האוצר בתוכו מידע נרחב ביותר אודות שוק העבודה מצד העובד ומצד המעביד. לצורך הערכת טיב הנתונים המצויים בקבצים מנהליים אלה ומידת השימושיות שלהם לצרכים הסטטיסטיים, נערכו בדיקות רבות. הבדיקות העלו כי מדדים ברי השוואה, אותם ניתן להפיק מקובצי מס הכנסה, תואמים במידה רבה את המדדים המתקבלים מאומדני הסקרים (פורמן, 2006).

על הפרק כעת תיעוד ממצא של מקורות מידע מנהליים אלה ופיתוח מסגרת להפקת סטטיסטיקה שוטפת על בסיסם.

מבוא

על פי פקודת מס הכנסה, כל אדם שיש לו הכנסה חייבת במס בישראל חייב בהגשת דו"ח שנתי לרשויות המס. הדו"ח אמור לכלול פירוט כל הכנסות הפרט והניכויים במקור שנוכו מההכנסה. מרבית תושבי ישראל, ובתוכם המועסקים כשכירים, פטורים מהגשת הדו"ח למעט שכירים בעלי הכנסה גבוהה במיוחד, בעלי שליטה בחברות בע"מ, עוסקים עצמאים, ופרטים שנדרשו לכך מפורשות על ידי פקיד השומה. הדו"ח השנתי של יחיד מוגש על גבי **טופס 1301** של רשות המיסים. הגשת דו"ח שנתי מתאפשרת גם לשכירים שאינם חייבים על פי חוק בהגשת הדו"ח, למשל לשם תיאום מס ותביעת החזרי מס – הדו"ח במקרה זה מוגש על גבי טופס 135 (דו"ח שנתי מקוצר). מעבידים ו"מנכים במקור" אחרים² מחוייבים בדיווח (ממוכן) על גבי **טופס 126** ובו פרטים על המשכורת השנתית ותשלומים אחרים ששולמו לכל אחד מעובדיהם וניכויי מס שבוצעו מתשלומים אלה.

בלמ"ס התקבלו שני סוגי קבצים, קובץ מס הכנסה שכירים (הנגזר מדיווח לפי טופס 126) וקובץ מס הכנסה עצמאים (מקבץ שדות מתוך טופס 1301). קבצים אלה עברו תהליך קליטה, טיוב, השלמה ובדיקות השוואתיות.

הגדרת האוכלוסיה לכל דו"ח אינה יחידנית בעבור כל פרט, משמע ייתכנו פרטים אשר יופיעו מספר פעמים בקובץ שכירים וכן מצב בו יימצאו בקובץ שכירים וגם בקובץ עצמאים; ההתייחסות לכל כפילות ואופן הטיפול מפורט בפרקים בהמשך.

² מי שחייב לנכות מס במקור אפילו אם אין בינו לבין מקבל התשלום יחסי עובד- מעביד, דוגמת תשלום לקבלנים.

1. קובץ מס הכנסה שכירים

1.1 תיאור הקובץ

קובץ מס הכנסה שכירים הוכן על בסיס דיווח מעבידים לפי טופס 126 עבור הלשכה המרכזית לסטטיסטיקה על ידי שירות עיבודים ממוכנים של משרד האוצר (להלן, שע"מ), המשמש יחידת מחשוב מרכזית של רשות המיסים. טופס 126 הינו דו"ח המוגש על ידי המעביד לרשות המיסים ובו פרטים על המשכורת השנתית ששילם לכל אחד מעובדיו וניכויי מס שבוצעו ממשכורת זו. פרטי העובד ממולאים על פי דיווחי העובד בתחילת שנת מס, כולל הצהרה בדבר משרות עבודה נוספות, פטורים וזיכויים לצורך חישוב חבות המס, הצהרה זו מגדירה למעשה את סוג משרת העבודה.

קובץ מס הכנסה שכירים הינו קובץ ברמת משרה: לכל פרט ייתכנו מספר משרות ואף משרות מסוגים שונים אצל אותו מעביד או לחילופין אצל מספר מעבידים. "סוג משרה" הוא משתנה המבטא את סוג הקשר בין עובד למעביד ומפרט את מהות התשלום (פירוט בסעיף 1.1.3). יודגש כי משתנה זה אינו מבטא את היקף המשורה במובן הרגיל (משרה מלאה או חלקית). כל מעביד ממלא את הטופס אודות כל עובדיו על כל משרותיהם וכל הטפסים יחד מאוגדים לקובץ מס הכנסה שכירים לשנת מס נתונה.

בקובץ שהוכן עבור הלמ"ס בנוסף למידע על כל התשלומים שבוצעו במהלך שנת המס, ישנם תשלומים שבוצעו אחרי תום שנת המס (ועד מועד בניית הקובץ) בגין אותה שנת המס. מבנה זה מבטא תמונה "על בסיס מצטבר" של הכנסות הפרט – אלה שהתקבלו בשנת המס ואלה שצמחו בשנת המס אך התקבלו מאוחר יותר. כפועל יוצא, מצרף ההכנסה של הפרט הנגזר מקובץ מס הכנסה שכירים לא זהה בהכרח להכנסה חייבת במס המוגדרת "על בסיס מזומן".

1.1.1 נתוני הכנסה ומס בקובץ מס הכנסה שכירים

רמת פירוט נתוני ההכנסה והמס בקובץ היא גבוהה מאחר וזו תואמת את צורכי רשויות המס, סוגי הכנסה שונים ממוסים אחרת ולכן יש משמעות לפירוט. לצרכים סטטיסטיים עניין מרכזי הוא בסך הכנסות הפרט (ממשרת עבודה). הנתון הרלוונטי הוא "סה"כ משכורת ותשלומים", זהו משתנה שהוא סכום חמישה רכיבי הכנסה המפורטים בקובץ:

- א. כל רכיבי הכנסה שחייבים במס בשיעור רגיל;
- ב. שכר עבודה במשמרות שנייה ושלישית בתעשייה ממנו נוכה מס בשיעור מופחת;
- ג. חלקיות הפרשות המעביד להבטחת זכויות סוציאליות של העובד (לקרן השתלמות, לקופת גמל לתגמולים, לתגמולים בקופת גמל לקיצבה) העולה על התקרה המאושרת וחייבת במס;
- ד. תשלומים לעובד שנוכה מהם מס בשיעור מיוחד על-פי הוראות פקיד השומה, בדרך כלל כתוצאה מתיאום מס;
- ה. תשלומים הפטורים מניכוי מס.

בקובץ מס הכנסה שכירים קיים פירוט של שלושת המסים הישירים שנוכו משכרו של העובד על ידי המעביד: סך הניכוי למס הכנסה מכל סוגי התשלומים, דמי ביטוח לאומי (חלקו של העובד בלבד) ודמי ביטוח בריאות. נתונים אלה מאפשרים חישוב הכנסה נטו.

1.1.2 תיאור רשומה

לוח 1: פירוט הרשומה בטופס 126

שם שדה	תאור שדה	פירוט והסבר לשדה
1	Koteret_mis_tik	מספר תיק ניכויים
2	Koteret_mis_agid	מספר אגיד
3	Prat_itur_kesher	מספר תעודת זהות
4	Koteret_shana_male	שנת המס המדווחת
5	Prat_shem_mishpacha	שם משפחה
6	Prat_shem_prati	שם פרטי
7	Prar_ktovet	רח' / שכונה ומספר בית
8	Prat_ishuv	שם עיר/ יישוב
9	Prat_min	מין
10	Prat_mazav	מצב משפחתי
11	Prat_shnat_leda	שנת לידה
12	Prat_mis_kid_18	מספר ילדים בחזקת העובד שבשנת המס טרם מלאו להם 18 שנה ומזכים בנקודת זיכוי מלאה
13	Prat_mis_kid_g	מספר ילדים הנמצאים בחזקת העובד ומזכים במחצית הנקודה
14	Prat_sug_misra	סוג משרה
15	Prat_tkufat_avoda	תקופת העבודה : החודשים בהם עבד במהלך השנה המדווחת
16	Prat_chodhe_avoda	סה"כ מספר חודשי עבודה במהלך שנת המס המדווחת
17	Prat_shem_zug	שם פרטי של בן/בת זוג

18	Prat_zehut_zug	מספר תעודת זהות של בן/בת זוג	
20	Prat_avoda_zug	האם לבן/בת זוג יש הכנסה 1 - יש הכנסה 2 - אין הכנסה	
21	Pratz_mis_nikud_m	מספר נקודות זיכוי מספר נקודות הזיכוי שניתנו לעובד כולל בשל תושבות ובשל נסיעה לעבודה (*100).	
22	Pratz_schum_nikud_m	סכום נקודות הזיכוי סכום (ערך) נקודות הזיכוי שניתנו לעובד בפועל (ש"ח).	
23	Pratz_nikud_nosaf_m	קוד הזכאות לנקודות זיכוי נוספות 1 - חייל משוחרר 2 - עולה חדש 3 - בן משפחה נטולת יכולת 4 - מזונות לבן זוג לשעבר 5 - מזונות לילד בחזקת בן הזוג לשעבר 6 - לבן/בת הזוג אין הכנסה 7 - אחר	
24	Pratz_ahuz_anaha_m	אחוז הנחה- תושב איזור מיוחד אחוז ההנחה לתושבי ישובים מיוחדים הזכאים להנחה	
25	Pratz_mis_ehidot_m	מספר יחידות מס מספר יחידות המס שנלקחו בחשבון בחישוב המס (25 כפול חודשי עבודה)	
26	Pratz_shiur_mas_m	שיעור מס מיוחד על ידי תאום או בעבודה חלקית שיעור המס שנקבע במסגרת תאום. למס במשרה חלקית זהו שיעור המס התקף.	
27	Pratt_mas_ragil	תשלומים מהם נוכה מס רגיל ולא פורטו בנפרד כולל את כל הסכומים והתשלומים החייבים בשיעורי מס רגילים כולל: חלק משכר משמרות החייב במס רגיל, הענקות בגין פריון עבודה, שעות נוספות, כונויות, תשלומים לכיסוי הוצאות למיניהם, הבראה, ספרות מקצועית, קצובת נסיעה לעבודה, ביגוד, דיור, החזקת רכב, החזקת טלפון וכד'.	
28	Pratt_mishmarot	שכר עבודה במשמרות שכר עבודה במשמרת (שניה או שלישית) בתעשייה, ממנו נוכה מס מופחת עד לתקרה שאושרה.	
29	Pratt_hisht_gemil	סכומי ההפרשה לקרן השתלמות, לקופת גמל לתגמולים/ לקצבה מהם נוכה מס הפרשות מעל התקרה המאושרת שחויבו במס.	

	תשלומים שנוכה מהם מס בשיעור מיחד על פי הוראת פקיד השומה	Pratt_mas_meuhad	30
תשלומים פטורים ממס לרבות החזר הוצאות מקרן השתלמות, החלק הפטור של קיצבה שהוונה וההוצאות מוכרות אצל עובדים ישראלים בחו"ל בהתאם לסוג המשרה.	תשלומים פטורים מניכוי מס	Pratt_ptor_mmas	31
סה"כ התשלומים לעובד בשנת המס, סכום שדות 25-29	סה"כ המשכורת והתשלומים	Prat_masc_vtash	32
לקופת גמל לתגמולים או לקיצבה	תשלומים המהווים בסיס לתגמולים, לקצבה	Prat_tagmul_kizba	33
	סכומים פטורים מתשלום מס מעסיקים	Pratt_ptor_maasik	34
	תשלומים פטורים מדמי ביטוח לאומי (שאינם נכללים בטופס 102)	Pratt_ptor_bl_lo102	35
	תשלומים החייבים בדמי ביטוח לאומי עד למקסימום שפורסם	Pratt_hiuv_ad_max	36
	תשלומים פטורים מדמי ביטוח לאומי מעל המקסימום שפורסם	Pratt_ptor_mal_max	37
סה"כ השכר ששולם לעובדים שאינם מבוטחים בביטוח זיקנה ושארים	תשלומים שלא נוכו מהם דמי ביטוח לאומי אינם מבוטחים בזיקנה ושארים	Pratt_lonuke_zikna	38
סה"כ המס שנוכה ממשכורת העובד מכל סוגי התשלומים	סה"כ הניכוי למס הכנסה	Pratt_mas_hahnasa	39
	משיכה מקופת גמל וקרן השתלמות לפני המועד החוקי תמורה/ רווח ממכירת ני"ע	Pratt_hiuv_gemel	40
	דמי ביטוח לאומי המנוכים מהעובד	Pratt_dmey_bl	41
	דמי ביטוח בריאות המנוכים מהעובד	Pratt_dmey_briut	42
הסכום שנוכה מהמשכורת לקופת גמל לתגמולים או לקיצבה כאשר גם המעביד מפריש לקופה (זיכוי 25%)	תשלומים לקופת גמל (עמית שכיר)	Pratt_hefresh_geml	43

44	Pratt_nuka_ lekizba	תשלומים לקופת גמל לקיצבה/ תרומות למוסדות מאושרים	הסכום שנוכה מהמשכורת לקופת גמל לקיצבה מקיפה (זיכוי 35%). הסכום שתרם העובד למוסדות מאושרים (זיכוי 35%)
45	Pratt_zikuy_ meuhad	זיכוי מיוחד לפי אישור פקיד השומה או הוראות הנציב	
46	Pratt_lobasis_ ltag	תשלומים שאינם מהווים בסיס להפרשות המעביד לתגמולים או לקיצבה	
47	Pratt_geml_bli_ maavid	תשלומים לקופת גמל כעמית עצמאי	הסכום שנוכה מהמשכורת לחיסכון בקופת גמל לתגמולים כשהמעביד אינו מפריש לקופה
48	Pratt_baada_ hefresh	משכורת שבעדה הפריש המעביד לקרן השתלמות	מלוא המשכורת בפועל שעבורה הפריש המעביד לקרן השתלמות (ולא המשכורת "הקובעת" עד התקרה)
49	Pratt_hisht_ mehkar	הסכום שנוכה מהמשכורת כהשתתפות בהוצאות מחקר ופיתוח	לצורך ניכוי ממס בגין מימון מחקר מדעי
50	Pratt_nikuy_ meuhad	ניכוי מיוחד לפי אישור פקיד השומה או הוראות הנציב	
51	Prat_yom_avoda	יום תחילת עבודה	
52	Prat_hodesh_ avoda	חודש תחילת עבודה	
53	Prat_shnat_ avoda	שנת תחילת עבודה	

1.1.3 סוגי המשרות

לוח 2: פירוט סוגי המשרות, לפי הקודים המופיעים בקובץ מס הכנסה שכירים

קוד סוג משרה	פירוט סוג המשרה	פירוט חודשי העבודה
01	<u>משרה יחידה</u> (מלאה או חלקית) על פי הצהרת העובד, משרה בה מחשב המעביד את חבות המס לפי שיעורים שוליים ומעניק נקודות זיכוי על פי החוק.	קיים
02	<u>משרה נוספת</u> (מלאה, חלקית או יומית) משכורת עובד שלא הצהיר על העסקתו/ אי העסקתו במקום אחר או שהצהיר על הכנסות בגינן מקבל מדרגות מס ונקודות זיכוי. על הכנסות ממשרה זו חל שיעור המס 50% או שיעור אחר לפי אישור פקיד שומה	קיים
03	<u>משרה יומית</u> חישוב המס בה לפי לוח ניכויים יומי	קיים

04	<u>קיצבה (פנסיה) קיצבה מאת מעביד לשעבר או מקופת גמל בגין עבודה בעבר</u>	קיים
05	<u>הפרשי שכר לשנים קודמות יש לדווח על ההפרשים לכל שנת מס ברשומה נפרדת</u>	לא קיים
06	<u>פיצויי פרישה תשלומים בגין פיצויי פרישה כולל פדיון ימי מחלה, תוך אבחנה בין סכום החייב לבין סכום הפטור</u>	לא קיים
07	<u>פדיון חופשה אם שולם במסגרת ניתוק יחסי עובד מעביד (אחרת יירשם במשרה מסוג 01 02 03)</u>	לא קיים
08	<u>קיצבת שארים אם שולמה על פי חוק או הסכם קיבוצי</u>	קיים
09	<u>היוון קיצבה תוך אבחנה בין החלק החייב לחלק הפטור</u>	לא קיים
10	<u>משיכה מקרן השתלמות לשכירים לפני תום שש שנים ממועד ההפקדה הראשונה בחשבון, בין אם נוכח מס במקור ובין אם לאו.</u>	לא קיים
11	<u>משיכה מקופת גמל לפני תום תקופת החסכון</u>	לא קיים
13	<u>תשלומים לבעל שליטה שאינם משכורת או שכר עבודה</u>	קיים
14	<u>תמורה ממכירת ניירות ערך אשר ההקצאה מאושרת על פי סעיף 102 לפקודה, הכנסה זו תדווח כמשיכה מקופת גמל ומקרן השתלמות, אחרת תדווח כהכנסה ממשכורת</u>	לא קיים
15	<u>רווח הון ממכירת ניירות ערך כאשר ההקצאה על פי סעיף 102 לפקודה, הרווח מדווח כמשיכה מקופת גמל ומקרן השתלמות</u>	לא קיים
16	<u>תשלומים לעובדים ישראלים בחו"ל מדווח בשקלים, תשלום המס מפורט בין אם שולם בארץ ובין אם בחו"ל</u>	קיים

כאשר המעביד משלם לעובד בשנת המס גם משכורת וגם קיצבה עליו לציין זאת ברשומות נפרדות, תוך ציון סוג המשרה. אותו הדין לגבי פרישה במהלך השנה, כל סוג משרה ברשומה. לשכיחות רשומות של משרות אלה בקבצי השנים 2002-2003, ראה לוח 4 בהמשך.

1.2 השלמה, טיוב ובדיקות

- קובץ הנתונים עובר תהליך טכני ארוך ומדוקדק על מנת לוודא כי הנתונים איכותיים ותואמים את הצרכים הסטטיסטיים של הלמ"ס. להלן הבדיקות וההשלמות שנעשות לקובץ זה:
- בדיקת כמות רשומות: כדי לוודא כי הקובץ הגיע משע"מ בשלמותו והוא מכיל רק את נתוני השנה המדוברת;
 - המרת השדות לסוג ופורמט המתאים כפי שהוגדר בבסיס הנתונים, נומרי גדול, קטן וכ"ו. בשלב זה נעשית בדיקה לטיב הנתונים, כגון מספרים חריגים וערכים לא לגיטימיים;
 - בכדי להשלים נתונים דמוגרפים לפרטים מתבצע קישור למרשם תושבים לפי מפתח מספר ת.ז., שנת הלידה, מין, שם פרטי או/ו שם משפחה. רשומות של פרטים שלא נמצאו במרשם נמחקות. כ-99% נמצאים במרשם תושבים;
 - ביקורת לוגית לשדות שונים: שנה וחודש תחילת עבודה, סך חודשי העבודה. הצורך בבדיקות אלה נובע מכך שרשויות המס אינן מגדירות את כל השדות בטופס 126 כשדות חובה, וסביר שתוכן שדות שכאלה אינו מפוקח דיו בשלב קליטת הדיווחים בשע"מ. נתונים

נוספים מתוקנים ומחושבים ועליהם ארחיב בסעיף 3, בתיאור בניית בסיס הנתונים עובד-מעביד.

בעבר נערכה השוואה בין מדדים שונים מתוך סקרי הלמ"ס לנתונים מקובץ מס הכנסה שכירים. המדדים שנבחנו הם: משרות שכיר, שכר למשרת שכיר, שכירים ושכר ממוצע לשכיר. הממצאים סוכמו בנייר טכני מס' 14 "קובץ מס הכנסה שכירים – ניתוח השוואתי של מדדי תעסוקה ושכר בין מקורות שונים", למ"ס.

1.3 קבצים קיימים

קובץ שנתי מוזמן על ידי הלמ"ס משע"מ כשנה וחצי לאחר תום שנת המס, בכדי לאפשר שלמות מירבית של הנתונים, שכן דיווח מעבידים בטופס 126 לשלטונות המס וקליטתם בשע"מ מתפרש על פני יותר משנה.

עד למועד כתיבת מסמך זה נקלטו, נבדקו, הושלמו וטוייבו קבצים לשנים 1999-2005. קבצים לשנים 1996-1998 התקבלו אך טרם טופלו.

לוח 3: מספר רשומות, פרטים ושכירים, בקבצי מס הכנסה שכירים, 1999-2005

שנה	מספר רשומות שהתקבלו	מספר רשומות לאחר טיוב	רשומות שהן משרות עבודה	מספר פרטים בעלי משרות עבודה	מספר פרטים
1999	4,609,145	4,433,577	3,623,481	2,497,544	2,705,821
2000	4,948,720	4,674,748	3,700,130	2,567,450	2,815,340
2001	5,017,579	4,781,522	3,767,089	2,614,117	2,882,498
2002	5,030,135	4,799,935	3,804,403	2,666,660	2,926,700
2003	4,795,388	4,606,657	3,661,548	2,598,064	2,892,598
2004	4,747,839	4,624,268	3,687,651	2,609,125	2,900,247
2005	4,731,018	4,597,115	3,477,802	2,625,660	2,911,152

לוח 4: התפלגות משרות בקובץ מס הכנסה שכירים, 2002-2003

סך משכורת ותשלומים				מספר משרות				
אחוזים		מיליוני ש"ח		אחוזים		מספר		
2003	2002	2003	2002	2003	2002	2003	2002	
85.3	86.3	180,717.9	190,965.9	72.4	71.7	3,333,124	3,441,606	01
2.6	2.7	5,450.1	5,883.9	4.3	4.4	197,485	213,357	02
0.7	0.7	1,403.8	1,528.2	2.8	3.1	130,939	149,440	03
7.3	6.9	15,475.0	15,173.1	6.7	6.0	310,652	287,960	04
0.2	0.3	503.0	772.1	5.5	6.9	252,824	329,424	05
2.1	1.5	4,545.6	3,412.1	2.2	2.1	101,729	98,572	06
0.3	0.3	667.2	626.5	3.9	3.8	178,271	183,568	07
1.2	1.0	2,610.6	2,305.8	2.0	1.7	90,305	80,028	08
0.0	0.1	0.3	215.7	0.0	0.0	6	896	09
0.0	0.0	-	-	0.0	0.0	11	12	10
0.0	0.0	0.1	0.1	0.1	0.1	2,434	6,610	11
0.0		-		0.0		99		14
0.0	0.0	11.2	22.0	0.1	0.0	3,791	579	15
0.2	0.1	391.0	270.6	0.0	0.0	1,216	971	16

2. קובץ מס הכנסה עצמאים

2.1 תיאור הקובץ

על פי הגדרת מס הכנסה, עצמאי הוא אדם שאין מי שינכה בעבורו מס במקור, אלא הוא אמור לנכות לעצמו מס במקור. טופס 1301 הוא הדו"ח השנתי על הכנסות בארץ ובחו"ל שמוגש למס הכנסה על ידי הפרט או בא כוחו. קובץ מס הכנסה עצמאים הוא קובץ ברמת פרט המכיל את כל בעלי תיק עצמאים ופירוט הנתונים בו הוא על סמך הדו"ח השנתי המעודכן ביותר שהוגש לרשויות המס. הרשומות בקובץ הנתונים כוללות תיקים פעילים בלבד (תיקים שטרם נסגרו יכילו בשדות המשתנים השונים אפסים, למעט הכנסה ממשכורת). האוכלוסייה בקובץ כוללת מספר קבוצות המוגדרות לפי **סוג תיק**, כאשר ההבדל ביניהן הוא אופי ניהול ספרי העסק:

- עוסקים יחידים, עוסקים זעירים, רוכלים וחקלאים (יכולים להיות גם מעסיקים);
- מנהלי חברה בעלי שליטה- לרוב אין להם הכנסה עסקית;
- בעלי מקצועות חופשיים;
- עצמאים גדולים (בהתאם למחזור עסקים).

רשומה מכילה פרטי זיהוי ומידע ממספר סוגים :

1. ההכנסה של "בן זוג רשום" (הגדרה בהמשך), על מרכיביה השונים : הכנסה עסקית, הכנסה ממשכורת, הכנסה הונית, הכנסה ללא יגיעה אישית ותשלומי מס הכנסה;
2. סוג תיק, נתונים על העסק, מלאי, רווח גולמי, מחזור, ניכויים במקור, פרטי הפסד וענף כלכלי;
3. נתוני הכנסה אודות בן/בת הזוג של "בן הזוג הרשום" : הכנסה עסקית, הכנסה ממשכורת, הכנסה הונית, הכנסה ללא יגיעה אישית ותשלומי מס הכנסה.

"בן זוג רשום" – על פי הוראות פקודת מס הכנסה, "פקיד השומה רשאי לקבוע בהודעה לכל שני בני זוג, כי אחד מהם הוא בן-זוג רשום לענין חוק זה, באשר הכנסתו החייבת בשנת המס שקדמה בשנתיים לשנת המס הנדונה לראשונה לענין זה היתה למעלה מ-50% מסך כל ההכנסה החייבת של שני בני הזוג".

בן/בת הזוג של "בן זוג רשום" – פרט המזוהה כבן הזוג של "בן זוג רשום", כפי שמופיע במרשם התושבים של שנת המס הנידונה. פרטי ההכנסות של בן/בת הזוג קיימים בדו"ח המוגש לרשויות המס, אולם בקובץ שהתקבל בלמ"ס מספר הזיהוי חסר ועל כן היה צורך בהשלמתו. נתון זה הושלם ממרשם התושבים של תחילת השנה העוקבת של שנת המס עליה דווח.

2.2 תאור רשומה

הנתונים המופיעים בקובץ מתבססים על הדיווח בטופס 1301 על נספחיו (הטופס נמצא בנספח 3). הרשומה כוללת פרטים ונתונים על "בן הזוג הרשום" לפי ההגדרה לעיל ובמידה והתיק מנוהל כתיק מאוחד של בני הזוג, מכיל גם נתוני הכנסה אודות בן/בת הזוג. הנתונים הם במונחים שנתיים, מחירים שוטפים.

לוח 5: מבנה רשומה בקובץ מס הכנסה עצמאים.

מספר שדה בטופס	הסבר/ ערכים	שם שדה	שם משתנה	
	מספר התיק הוא למעשה מספר הזהות של בעל התיק	תיק	TIK	1
		שנת המס	SM	2
	20- פטור מנהול ספרים 30- מנהל בחברה בעל שליטה 40- עוסק יחיד ועוסק זעיר (מע"מ) 41 – מערכת חשבונות חד צידית פשוטה 42-מנהל מערכת חשבונות חד צידית 43- חקלאי שלא מנהל חשבון בשטח 52- מנהל מערכת חשבונות חד צידית מרכבת וגם בעלי מקצועות חופשיים (רו"ח, רופאים, עו"ד וכו') 53- מנהל חשבונות בשיטה הכפולה	סוג תיק, לפי אופן ניהול ספרי העסק	STIK	3
	4 ספרות, מסווג על ידי פקידי רשויות המס לפי ספר ענפי הכלכלה של הלמ"ס	ענף כלכלי	ANAF	4
	פקיד שומה המטפל בתיק	מפקח	MEFD	5
	על פי סיווג שלטונות המס	סעיף שומה	SEIFD	6
	אופן חישוב חבות המס: תיק של יחיד, תיק מאוחד של בני הזוג, תיק בנפרד כאשר כל אחד מבני הזוג מנהל תיק בנפרד או תושב חוץ	סוג חישוב	SHISD	7
	מספר זהות, ראה סעיף 2.2	זהות "בן זוג רשום"	ZEHUTD	8
	1-זכר 2-נקבה	מין "בן זוג רשום"	MIND	9
39	תושב ספר/ שטח פיתוח/ יישוב חוץ, מזכה בהנחת מס	אחוז ספר "בן זוג רשום"	SFRD	10
מחושב	חישוב גובה חיוב המס מכל ההכנסות החייבות	מס הכנסה "בן זוג רשום"	MASD	11
	הכנסה כוללת לצורכי מס, כל סוגי ההכנסה למעט הכנסה פטורה	הכנסה כוללת "בן זוג רשום"	HACKD	12
209 (כולל 109)	סה"כ הכנסות ורווחים פטורים ובלתי חייבים במס, כולל הכנסת נכה.	הכנסה פטורה "בן זוג רשום"	HACPD	13
109	הכנסות פטורות ממס על פי חוק- נכה 100%	הכנסה נכה "בן זוג רשום"	HACND	14

150, 250, 149, 151, 162, 120 152	1. הכנסה מיגיעה אישית, מעסק או משלח יד 2. תקבולים והחזרים מהמוסד לביטוח לאומי החייבים במס ואינם נכללים בסעיף אחר 3. הכנסה מהשכרת נכס ששימש בעסק למעלה מ-10 שנים	הכנסה עסקית "בן זוג רשום"	HACED	15
258, 358, 068 158	הכנסות ממשכורת/ שכר עבודה, עבודה במשמרות בתעשייה, הסכום החייב במס של מענקי פרישה וגמלאות	הכנסה ממשכורת "בן זוג רשום"	HACSD	16
	רווח אינפלציוני חייב+רווח ראלי של כל נספחי רווח הון (אחרי קיזוז הפסדים) כולל רווחי הון מניירות ערך	הכנסה הונית "בן זוג רשום"	HACHD	17
212, 211, 210, 215, 123, 213 216, 214, 217 222, 225, 227 208	הכנסות החייבות בשיעורי מס מיוחדים- הכנסה מחו"ל, דמי מפתח, פטנט, דיבידנד, ריבית על אג"ח	הכנסה הונית אחרת "בן זוג רשום"	HACHAD	18
203, 204, 205 201, 202	הכנסות חייבות בשיעורי מס רגילים- מנכס בית, מ"חברת בית", מחברה משפחתית, הכנסות מרכוש אחר	הכנסה ללא יגיעה אישית "בן זוג רשום"	HACLD	19
083	תושב ספר/ שטח פיתוח/ יישוב חוץ	אחוז ספר בן זוג	SFRBZD	20
מחושב	חישוב גובה חיוב המס מכל ההכנסות החייבות	מס הכנסה בן זוג	MASBZD	21
	הכנסה כוללת לצורכי מס, כל סוגי ההכנסה למעט הכנסה פטורה	הכנסה כוללת בן זוג	HBZKD	22
170, 270, 169, 175, 178, 220 171	1. הכנסה מיגיעה אישית, מעסק או משלח יד 2. תקבולים והחזרים מהמוסד לביטוח לאומי החייבים במס ואינם נכללים בסעיף אחר 3. הכנסה מהשכרת נכס ששימש בעסק למעלה מ-10 שנים	הכנסה עסקית בן זוג	HBZED	23
272, 372, 069 172	הכנסה ממשכורת/ שכר עבודה, עבודה במשמרות בתעשייה, הסכום החייב במס של מענקי פרישה וגמלאות	הכנסה ממשכורת בן זוג	HBZSD	24
	רווח אינפלציוני חייב+רווח ראלי של כל נספחי רווח הון (אחרי קיזוז הפסדים) כולל רווחי הון מניירות ערך	הכנסה הונית בן זוג	HBZHD	25
308, 316, 317, 313, 314, 315 310, 311, 312	הכנסות החייבות בשיעורי מס מיוחדים- הכנסה מחו"ל, דמי מפתח, פטנט, דיבידנד, ריבית על אג"ח	הכנסה הונית אחרת בן זוג	HBZHAD	26

27	HBZLD	הכנסה ללא יגיעה אישית בן זוג	הכנסות חייבות בשיעורי מס רגילים- מנכס בית, מ"חברת בית", מחברה משפחתית, הכנסות מרכוש אחר	303, 304, 305, 301, 302
28	HEFMD	הפסד מועבר	יתרה להעברה לשנת המס הבאה (עוקבת)- הפסדים מעסק וממשלח יד	079
29	HEFKD	הפסד שקוזז	הפסד עסקי שקוזז מהכנסות השנה	279
30	HEFED	הפסד מניירות ערך במונחי מס	יתרה המועברת לשנת המס הבאה (עוקבת)	160
31	MLAYD	מלאי	ערך המלאי בעסק העיקרי בסוף שנת המס	097
32	REVGD	רווח גולמי	סך הרווח הגולמי בעסק העיקרי	195
33	NIKSD	ניכוי במקור משכר	מס הכנסה שנוכה במקור ממשכורת, משכר עבודה ומקצבאות, סה"כ "בן זוג רשום" ובן/ בת הזוג	042
34	NIKED	ניכוי במקור מעסק	סה"כ הסכומים שנוכו במקור (ללא מקדמות) מהכנסות אחרות הכלולות בדו"ח זה	040 (כולל ניכוי יהלומן שדה 091)
35	NIKRD	ניכוי במקור מריבית	ניכוי במקור מריבית על הכנסות מריבית על פקדונות ותוכניות חסכון	043
36	MACZD	מחזור מעסק עקרי	המחזור בעסק העיקרי (ללא מע"מ)	293
37	MACMD	מחזור	סך המחזור ללא מע"מ מעסק או משלח יד והכנסות אחרות בשיעורים רגילים, מכל העסקים, בארץ ובחו"ל בשנת המס	294
38	SHVCD	חיוב במס שבח	סה"כ מס שבח שנקבע בשומת מס שבח (קרן בלבד)	041
39	MAAMD	שדה מע"מ		103
40	ZHULD	זיכוי חו"ל	זיכוי על מס ששולם בחו"ל, כדי להגיע למס המגיע של השומה יש צורך לחבר מס מגיע+ מס מגיע ב"ז ולהחסיר את זיכוי חו"ל.	

מתוך קובץ עצמאים במבנהו הני"ל לא ניתן לחשב תשלומי דמי ביטוח לאומי (כעצמאי וכשכיר), מה שעשוי למנוע השוואתיות מול הכנסה נטו (הכנסה ברוטו בניכוי תשלומי חובה) בקובץ שכירים. כדי להשלים את החסר, הוצאה בקשה מתאימה לשע"מ, ונתונים אלה יושלמו בעתיד, תוך זקיפת דמי ביטוח בריאות.

2.3 קבצים קיימים

בלמ"ס נקלטו, נבדקו, הושלמו וטוייבו קבצי מס הכנסה עצמאים לשנים 1999-2004. לוח 6 מציג את אוכלוסיית הקובץ כאשר רשומה מתייחסת לנתוני "בן זוג רשום".

לוח 6: ספירת רשומות בקבצי מס הכנסה עצמאים, 1999-2004

שנה	מספר רשומות שהתקבלו	מספר רשומות שנמצאו במרשם תושבים	מספר רשומות שיש להן הכנסה חייבת במס	מספר רשומות שיש להן הכנסה עסקית
1999	323,988	322,079	306,454	201,590
2000	342,026	340,090	322,708	211,889
2001	357,523	355,581	336,010	221,256
2002	371,494	369,500	346,967	230,536
2003	385,528	383,412	357,706	237,399
2004	389,321	387,422	361,541	245,629

לוח 7א: התפלגות אוכלוסיית העצמאים לפי סוג התיק, 1999-2004.

סוג תיק	1999	2000	2001	2002	2003	2004
סה"כ	322,079	340,090	355,581	369,500	383,412	387,422
20	1,008	1,139	1,227	1,344	1,507	2,068
30	79,991	83,009	85,088	86,653	85,981	79,975
40	799	878	946	951	917	1,079
41	18,366	19,447	20,286	21,087	22,163	22,824
42	143,462	153,717	163,048	171,821	181,870	189,801
43	17,194	17,444	17,623	17,785	18,179	17,332
52	54,412	57,569	60,464	62,984	65,799	67,695
53	6,847	6,887	6,899	6,875	6,996	6,648
אחוזים	100.0	100.0	100.0	100.0	100.0	100.0
20	0.3	0.3	0.3	0.4	0.4	0.5
30	24.8	24.4	23.9	23.5	22.4	20.6
40	0.2	0.3	0.3	0.3	0.2	0.3
41	5.7	5.7	5.7	5.7	5.8	5.9
42	44.5	45.2	45.9	46.5	47.4	49.0
43	5.3	5.1	5.0	4.8	4.7	4.5
52	16.9	16.9	17.0	17.0	17.2	17.5
53	2.1	2.0	1.9	1.9	1.8	1.7

לוח 27: התפלגות לפי סוגי התיק לעצמאים בעלי הכנסה חיובית, 2004-1999

סוג תיק	1999	2000	2001	2002	2003	2004
סה"כ	306,454	322,708	336,010	346,967	357,706	361,541
20	931	1,040	1,136	1,243	1,411	1,926
30	76,693	79,247	80,806	81,787	80,956	75,414
40	765	830	904	898	877	1,028
41	17,477	18,477	19,176	19,950	20,810	21,469
42	135,637	144,973	153,195	159,966	167,736	175,065
43	15,672	15,864	15,996	16,098	16,321	15,503
52	52,998	55,993	58,583	60,847	63,302	35,129
53	6,281	6,284	6,214	6,178	6,293	6,007
אחוזים	100.0	100.0	100.0	100.0	100.0	100.0
20	0.3	0.3	0.3	0.4	0.4	0.5
30	25.0	24.6	24.0	24.3	22.6	20.9
40	0.2	0.3	0.3	0.3	0.2	0.3
41	5.7	5.7	5.7	5.9	5.8	5.9
42	44.3	44.9	45.6	47.6	46.9	48.4
43	5.1	4.9	4.8	4.8	4.6	4.3
52	17.3	17.4	17.4	18.1	17.7	18.0
53	2.0	1.9	1.8	1.8	1.8	1.7

2.4 השלמה, טיוב ובדיקות

השלמת הנתונים הדמוגרפיים בקובץ מס הכנסה עצמאים נעשה לפי מספר זהות של "בן זוג רשום", הוא בעל התיק לפי הגדרת מס הכנסה. כל קובץ שנתי הוקבל למרשם התושבים הנכון לתחילת השנה העוקבת (לשנת 2001 אין נתוני מרשם ולכן השנים 2000 ו-2001 זווו למרשם שנת 2002). הפרטים אשר נקלטו מהמרשם הם: שם משפחה, שם פרטי, מין, תאריך לידה, ארץ לידה, תאריך עליה, לאום, מצב משפחתי, תאריך שינוי מצב משפחתי, סמל יישוב המגורים, תאריך פטירה ומספר תעודת הזהות של בן/ בת הזוג, אם קיים במרשם, ללא תלות במצב המשפחתי של "בן הזוג הרשום". 99.5% מבעלי התיקים נמצאו במרשם התושבים.

על מנת לעמוד על טיב הנתונים בקבצי מס הכנסה עצמאים שהתקבלו לראשונה בלמ"ס, נבחר קובץ הנתונים לשנת 2002 כקובץ מייצג, עליו נערכו שלוש הבדיקות הבאות:

1. בדיקת אוכלוסיית עצמאים פעילים;

2. בדיקת גובה והרכב הכנסות;

3. בדיקת מהימנות דיווח על שכר.

שתי הבדיקות הראשונות נערכו תוך השוואה מצרפית מקובץ עצמאים לסטטיסטיקה של מגזר העצמאים בדוחות מינהל הכנסות המדינה, שם קיים מודל המס שבסיס הנתונים שלו נבנה

עקרונית מאותם מאגרי מידע של שע"מ המשמשים גם את הלמ"ס. אולם הגדרת אוכלוסיית קובץ עצמאים (פעילים) אינה בהכרח זהה לזו הקיימת במודל המס של מינהל הכנסות המדינה. הבדיקה השלישית נערכה באמצעות קישור בין קובץ עצמאים לקובץ מס הכנסה שכירים, במטרה לבחון את מידת החפיפה בין שני מקורות מידע אלה.

2.4.1 בדיקת אוכלוסיית העצמאים

לוח 8: אוכלוסיית העצמאים לפי מין ומצב משפחתי, 2002 (אלפים)

הפרש באחוזים	מודל המס של מינהל הכנסות המדינה	קובץ עצמאים	
4.7	352	369.5	סה"כ
4.6	302	316.6	גברים
5.5	50	52.9	נשים
			מצב משפחתי
13.8	38	44.1	רווק/ה
3.6	278	288.3	נשוי/ה
4.6	25	26.2	גרש/ה
14.3	9	10.5	אלמן/ה
	2	0.4	אחר או לא ידוע

אוכלוסיית העצמאים הפעילים בקובץ שברשות הלמ"ס גדולה ב-17 אלף איש (4.7%) מזו שמשמשת את מינהל הכנסות המדינה למודל המס שלו. התפלגות לפי מצב משפחתי מלמדת שפערים גדולים יחסית קיימים בקרב הרווקים והאלמנים. עם זאת, ההפרשים הקיימים אינם גדולים, בהתחשב בעיתוי שונה של שליפת הנתונים (קבצי הלמ"ס התקבלו מאוחר יותר, לכן כלולים בהם דוחות של עצמאים שהוגשו באיחור).

2.4.2 בדיקת גובה והרכב הכנסות

לוח 9 מגלה מספר הבדלים בין התפלגויות של הכנסה ומס של העצמאים בשני המקורות:

- בממוצע, הכנסה ברוטו (חייבת במס) גבוהה ב-6.2% באוכלוסיית מודל המס של מינהל הכנסות המדינה;
 - בתשעת העשירונים הראשונים ההכנסה נמוכה יותר – כאשר תשלומי מס הכנסה גבוהים יותר – בקובץ עצמאים של הלמ"ס לעומת בסיס נתונים של מינהל הכנסות המדינה;
 - בעשירון העליון ההכנסה הממוצעת דומה, בשעה שתשלומי מס הכנסה בקובץ עצמאים נמוכים יותר לעומת מודל המס של הכנסות המדינה.
- ניתן לשער כי ההבדלים הנ"ל נובעים, לפחות בחלקם, מהעובדה כי קובץ עצמאים של הלמ"ס אינו כולל הפרשי שומה, שאמורים להגדיל הכנסה חייבת במס בעיקר בעשירונים הנמוכים. סיבה אפשרית נוספת היא הרכב ההכנסות: שיעור גבוה יותר של רווחי הון, הממוסים בשיעור מס שולי (גבוה יותר משיעור מס קבוע שחל על הכנסות הוניות אחרות), מעלה את שיעור המס השולי הממוצע. זהו נושא הבדיקה הבאה.

לוח 9: השוואת התפלגות הכנסה ומס של עצמאים בין קובץ מס הכנסה עצמאים¹ ודו"ח מינהל הכנסות המדינה², 2002

התפלגות מס הכנסה, אחוזים		התפלגות הכנסה, אחוזים		הכנסה חייבת במס, ש"ח לחודש ³		עשירון הכנסה ברוטו
דו"ח מינהל	קובץ עצמאים	דו"ח מינהל	קובץ עצמאים	דו"ח מינהל	קובץ עצמאים	
0.0	0.0	1.0	0.7	16,776	10,424	1
0.0	0.0	2.1	1.8	33,372	27,338	2
0.0	0.1	2.9	2.7	46,716	40,129	3
0.2	0.5	3.8	3.5	61,212	52,704	4
0.9	1.4	4.9	4.5	77,988	68,008	5
2.2	2.6	6.2	5.8	99,012	87,806	6
4.4	4.5	7.9	7.6	127,104	114,923	7
8.3	8.2	10.6	10.3	170,136	155,284	8
16.8	17.0	15.9	15.5	256,056	233,779	9
67.1	65.6	44.7	47.5	717,408	715,168	10
				160,572	150,555	ממוצע

מקור: קובץ עצמאים למ"ס ודו"ח מינהל הכנסות המדינה לשנת 2005.
הערות ללוח:

- (1) נתונים עבור בעלי הכנסה חיובית בלבד, מתוך קובץ 2002.
- (2) נתוני דוחות ושומות של עצמאים לשנת מס 2002.
- (3) הכנסה שנתית חלקי 12, דהיינו אינה מתואמת למספר חודשי עבודה בפועל.

לוח 10: הרכב הכנסה החייבת במס של עצמאים, קובץ עצמאים של הלמ"ס לעומת נתוני מינהל הכנסות המדינה, 2002

דו"ח מינהל הכנסות המדינה					קובץ עצמאים של הלמ"ס						
הכנסה הונית	הכנסה הונית אחרת	הכנסה שלא מיגיעה אישית	משכורת	הכנסה מעסק או משלח יד	הכנסה הונית	הכנסה הונית אחרת	הכנסה שלא מיגיעה אישית	משכורת	הכנסה מעסק או משלח יד	הכנסה כוללת שנתית	עשירון
0.1	0.0	4.6	21.8	73.5	0.2	0.4	4.9	20.6	73.8	100	1
0.1	0.0	3.1	26.1	70.8	0.1	0.3	3.6	25.3	70.7	100	2
0.1	0.0	2.4	33.3	64.2	0.1	0.3	2.7	33.3	63.6	100	3
0.1	0.0	2.5	35.6	61.9	0.1	0.3	2.6	36.1	61.0	100	4
0.1	0.0	2.4	38.7	58.7	0.1	0.4	2.6	40.6	56.3	100	5
0.1	0.0	2.5	43.0	54.3	0.1	0.5	2.6	46.1	50.7	100	6
0.1	0.0	2.8	49.7	47.3	0.1	0.5	2.7	52.0	44.7	100	7
0.2	0.1	3.1	52.5	44.1	0.2	0.7	2.8	55.6	40.7	100	8
0.2	0.2	2.9	60.9	35.8	0.6	1.0	2.8	61.7	33.8	100	9
2.0	11.2	3.4	53.1	30.3	10.4	13.0	3.3	48.8	24.2	100	10
1.0	5.0	3.1	50.5	40.5	5.1	6.5	3.0	49.7	35.5		ממוצע

מקור: קובץ מס הכנסה עצמאים לשנת 2002; דו"ח מינהל הכנסות המדינה 2005 (נתוני דוחו"ת ושומות עצמאים לשנת) 2002.

בלוח 10 ניתן לראות כי מרכיב ההכנסה העסקית בעשירונים הגבוהים גדול יותר בדו"ח מינהל הכנסות המדינה, מרכיב ההכנסה ההונית כולל הונית אחרת גבוה יותר בקובץ עצמאים ובעיקר בעשירון העליון.

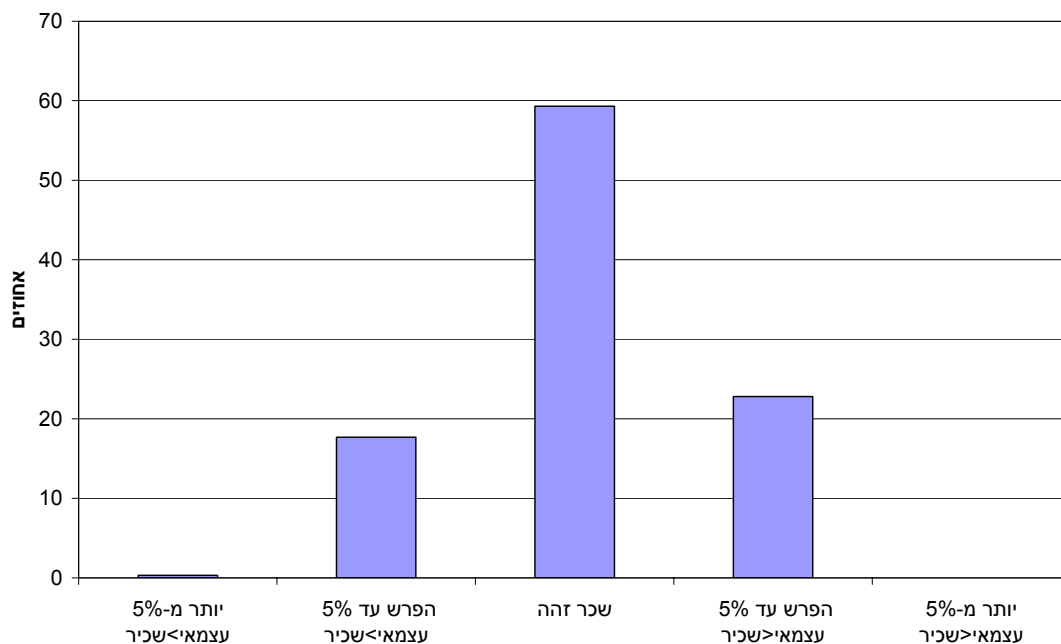
2.4.3 בדיקת רכיב השכר

כאמור, מידע בקובץ מס הכנסה שכירים מקורו בדיווח על ידי המעבידים, ואילו קובץ עצמאים נשען על דיווח עצמי של הפרטים החייבים בזאת (או מייצגים כמו רו"ח, עו"ד, יועץ מס). הרשומות בקובץ מס הכנסה שכירים הן ברמה של משרה, כאשר פרט יכול להחזיק במספר משרות במהלך השנה, בין אם במקביל ובין אם בשל החלפת מקומות עבודה. בקובץ מס הכנסה עצמאים הרשומה היא ברמת תיק. רכיב אחד מנתוני ההכנסה הכוללת המופיע בקובץ עצמאים הוא הכנסה ממשכורת, ונתונים אלה אמורים להיות מדווחים כהכנסה מעבודה שכירה (במשרות 01, 02, 03) בטופס 126 על ידי כל המעבידים של הפרט. בכדי לבחון האם זהו אכן המצב, הושוותה ההכנסה השנתית ממשכורת בין שני הקבצים, עבור כל הפרטים שאותרו בשני המקורות. בקובץ מס הכנסה שכירים סוכמו כל ההכנסות השנתיות מכל משרות העבודה בהן החזיק פרט במהלך שנת 2002, ותוצאה זו הושוותה לנתון הכנסה ממשכורת כפי שדווחה בקובץ עצמאים.

לוח 11: השוואת נתוני שכר בקבצי מס הכנסה שכירים ועצמאים, 2002

מספר רשומות	
369,500	סה"כ בעלי תיק בקובץ עצמאים
176,474	מזה: נמצאים בקובץ מס הכנסה שכירים 2002
164,044	מזה: בעלי משרות עבודה
97,253	שכר זהה בין שני המקורות

תרשים 1: התפלגות הפרשים בנתוני שכר של בעלי תיק בין קובץ עצמאים לקובץ שכירים, 2002



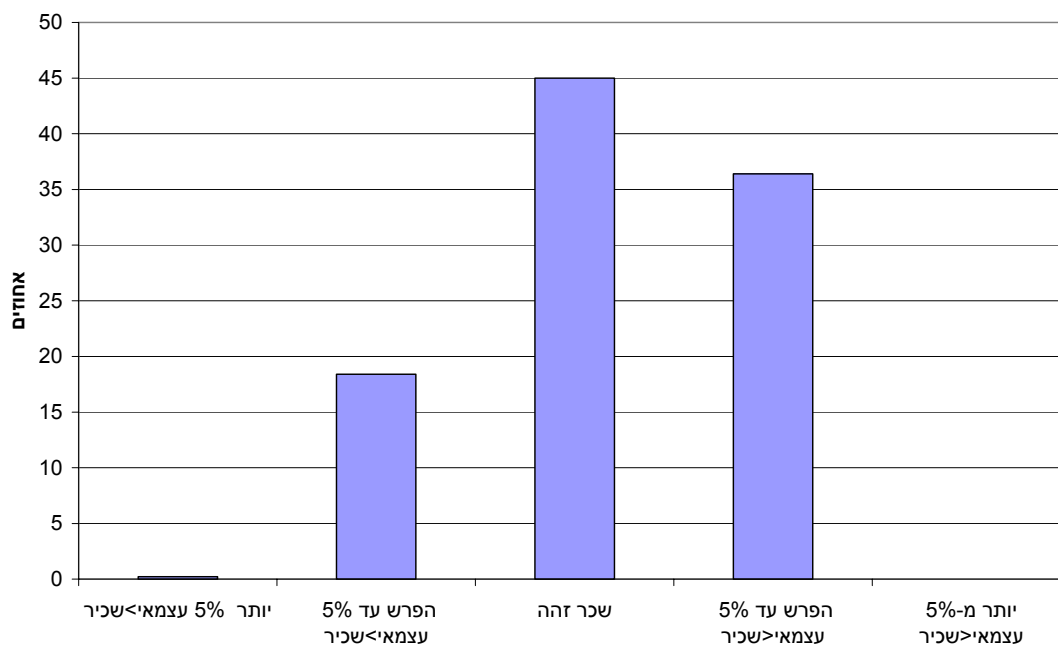
מסקנה: 48% מבעלי תיק עצמאי נמצאו בקובץ מס הכנסה שכירים. למעל 99% מהעצמאים אשר נמצאו בקובץ שכירים, פער השכר מעבודה בין דו"ח עצמאי לטופס 126 אינו עולה על 5%.

בדיקה זהה נעשתה גם לבני הזוג המופיעים בקבצי מס הכנסה של העצמאים.

לוח 12 : השוואת נתוני שכר של בן/בת זוג של "בן זוג רשום" בקבצי מס הכנסה שכירים ועצמאים, 2002

מספר רשומות	
320,729	סה"כ בן/בת זוג של "בן זוג רשום"
188,708	מזה: נמצאים בקובץ מס הכנסה שכירים 2002
179,826	מזה: בעלי משרת עבודה
80,849	שכר זהה

תרשים 2: התפלגות הפרשים בנתוני שכר בן/בת זוג של "בן זוג רשום" בין קובץ עצמאים לקובץ שכירים, 2002

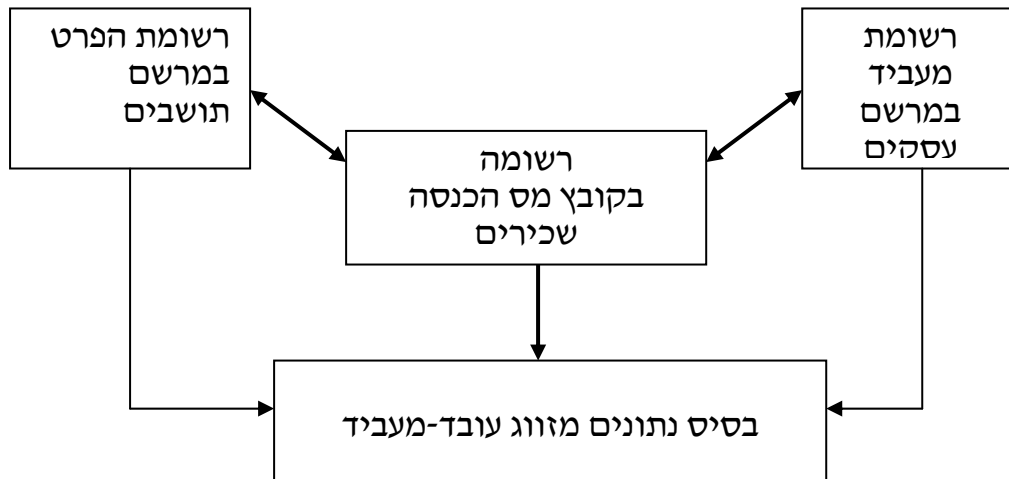


59% מקרב בני הזוג נמצאו בקובץ מס הכנסה שכירים. למעל 99% מביניהם פער השכר בין המדווח עצמאית למדווח בטופס 126 לא עולה על 5%. בבדיקה זו למעשה קיבלנו תימוכין גם לזיווג מספר ת.ז. שנעשה לבני הזוג של "בן-זוג רשום" לפי מרשם התושבים.

3. בסיס נתונים מזווג עובד-מעביד

הרציונאל שבבניית בסיס נתונים מזווג עובד-מעביד המתבסס על קובץ מס הכנסה שכירים, מקורו ברצון לנצל מקורות מידע מנהליים קיימים והשלמת מידע חסר לצורכי מחקר שוק העבודה על שני צדדיו, הן מצד העובד והן מצד המעביד. בסיס הנתונים מתייחס לאוכלוסיית השכירים, למעסיקים ולמשרות השכיר לפי מאפיינים שונים, בנקודת זמן מסויימת ולאורך שנים.

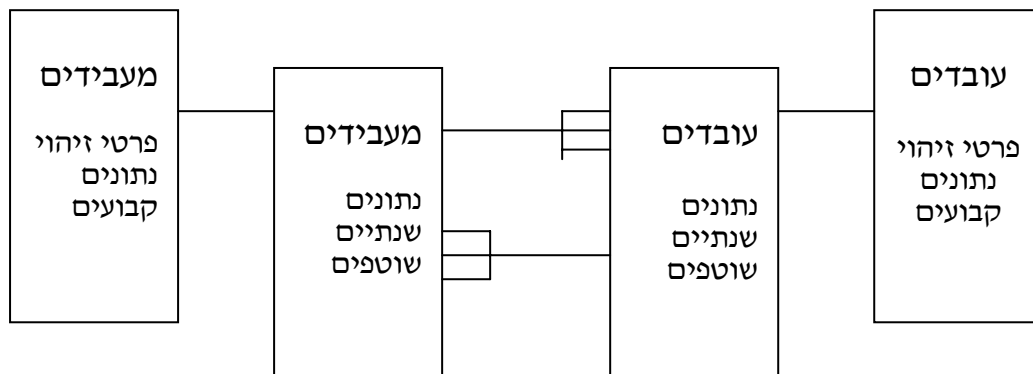
3.1 העיקרון



תהליך הבנייה כלל מספר שלבים: איתור וזיהוי עובדים במרשם התושבים בעזרת מספר הזיהוי; טיוב והשלמת נתוני הפרט; איתור וזיהוי מעבידים מול מרשם העסקים בעזרת מספר תיק ניכויים; השלמת נתוני המעביד. הקובץ הראשון שזווג הוא לשנת 1999 ואילו נקודת הפתיחה של מרשם עסקים 2003 ולכן היה צורך בהשלמת זיהוי מעבידים ידני עבור כל אחד מהקבצים השנתיים לעסקים שנסגרו או שינו מספר תיק ניכויים בין 1999 ל-2003. פירוט התהליך יובהר בהמשך בתיאור טבלאות בסיס הנתונים עובד-מעביד.

3.2 המבנה

נתוני העובדים השכירים מאוגדים בשתי טבלאות, באחת הנתונים הקבועים של העובדים, השנייה מכילה פירוט נתונים שנתיים. נתוני המעבידים מאוגדים אף הם בשתי טבלאות, האחת לנתונים הקבועים, השנייה מכילה את הנתונים השוטפים.



הקשר בין טבלת עובדים-נתונים שנתיים שוטפים ובין טבלת מעבידים-נתונים שנתיים שוטפים אינה חד-חד-ערכית מאחר ולמעביד על-פי רוב ישנם מספר עובדים, כמו כן לעובד יכולים להיות מספר מעבידים באותה שנה, אם בשל ריבוי משרות או בשל החלפת מקום עבודה.

3.2.1 מבנה טבלאות עובד

כאמור, רשומות של שכירים אותרו וזוהו במרשם התושבים באמצעות מספר זהות (כפוף לבדיקת תקינות ספרת הביקורת), כ-99% פרטים זוהו במרשם התושבים לפי מספר ת.ז, שנת לידה, מין ושם פרטי או/ו שם משפחה. פרטים שמספר ת.ז. שלהם לא נמצא במרשם התושבים לא הוכנסו לטבלאות העובד.

לאחר מכן הושלמו נתונים דמוגרפיים בטבלאות מאפיינים קבועים ומשתנים. כל קובץ שכירים שנתי זווג למרשם התושבים העדכני לסוף אותה שנה, קרי תחילת שנה עוקבת, למעט קובץ הנתונים של שנת 2000 שזווג למרשם 2002 בשל חוסר מרשם 2001. הנתונים שנלקחו ממרשם תושבים: שם משפחה, שם פרטי, מין, תאריך לידה, ארץ לידה, תאריך עליה, לאום, מצב משפחתי, תאריך שינוי מצב משפחתי, סמל יישוב המגורים ותאריך פטירה.

3.2.1.1 מאפיינים קבועים

בטבלה זו מופיעים כל הפרטים אשר לפחות באחת מן השנים 1999-2005 הופיע בקובץ מס הכנסה שכירים, כל פרט מופיע פעם אחת בלבד ומכיל את כל הנתונים הקבועים של הפרט:

לוח 13: מבנה רשומה בטבלת מאפיינים קבועים של עובד

שם משתנה	מהות השדה	הסבר/ מקור	ערכים
Prat_itur_kesher	מספר תעודת זהות	מספר בעל 9 ספרות, בדוק מול מרשם תושבים	
Mt_shemmishpacha	שם משפחה	מרשם תושבים	
Mt_shemprati	שם פרטי	מרשם תושבים	
Mt_kodmin	מין	מרשם תושבים	1-גבר 2-אשה
Mt_taarichleda	שנת לידה	מרשם תושבים	YYYYMMDD
Mt_koderetzleda	ארץ לידה	מרשם תושבים	נספח 2
Mt_taarichaliya	שנת עליה	מרשם תושבים	YYYY
Mt_kodleom	לאום	מרשם תושבים	נספח 1
Mt_taarichptira	תאריך פטירה	מרשם תושבים	YYYYMMDD

עם הגעת נתונים לשנת מס נוספת, לאחר תהליך קליטתם, נערכת בדיקה מול טבלת מאפיינים קבועים של העובדים, באם הפרט לא נמצא בה קודם, מתווספת רשומה חדשה לטבלה עם הנתונים הקבועים ממרשם התושבים המתאים.

3.2.2.1 מאפיינים שוטפים

טבלה המכילה את כל המשרות של כל השכירים בעבור כל שנת מס בנפרד, פרטים אישיים הניתנים לשינוי, כמו כתובת מגורים ומצב משפחתי, נלקחים כל שנה מחדש ממרשם התושבים המתאים. נעשית בחינה לסך חודשי העבודה בכל משרה ובודקים כי סה"כ החודשים מתאים לפירוט החודשים (ולא עולה על 12), מטויב משתנה תאריך תחילת העבודה.

לוח 14: מבנה רשומה בטבלת מאפיינים משתנים של עובד

שם משתנה	מהות השדה	הסבר/ מקור	ערכים
Koteret_mis_tik	מספר תיק ניכויים מס הכנסה		9 ספרות
Control_Shana_shotef	שנת המס		1999-2005
Prat_ktovet	כתובת	כתובת, טופס 126	
Prat_ishuv	יישוב	יישוב, טופס 126	לפי פירוט ספר היישובים
Prat_sug_misra	סוג משרה		פירוט הערכים בסעיף 1.1.3
Prat_hodesh_avodastart	חודש תחילת העבודה	ערך חסר -1	
Prat_shnat_avodastart	שנת תחילת העבודה	ערך חסר -1	
Prat_itur_kesher	מספר תעודת זהות		
Prat_shem_mishpacha	שם משפחה		
Prat_shem_prati	שם פרטי		
Hodesh1...hodesh12	חודשי עבודה, ינואר עד דצמבר	שניים עשר משתנים נפרדים	0-לא עבד -1 עבד
Worked_minths	סך חודשי עבודה	סכימה אופקית של חודשי העבודה	0-12
Pratt_dmey_bl	דמי ביטוח לאומי המנוכים מהעובד	סכום שנתי	ש"ח
Pratt_dmey_briutl	דמי ביטוח בריאות המנוכים מהעובד	סכום שנתי	ש"ח
Pratt_hisht_geml	סכומי הפרשה לקרן תשלומים/קופות גמל מעל תקרה מאושרת מהם מנוכה מס	סכום שנתי	ש"ח
Pratt_hiuv_ad_max	תשלומים החייבים בדמי ביטוח לאומי (עד התקרה)	סכום שנתי	ש"ח

ש"ח	סכום שנתי	סה"כ הניכוי למס הכנסה (מכל סוגי התשלומים)	Pratt_mas_hahnasa
ש"ח	סכום שנתי	תשלומים מהם נוכה מס בשיעור מיוחד (אישור פקיד שומה)	Pratt_mas_meuhad
ש"ח	סכום שנתי	תשלומים מהם מנוכה מס רגיל	Pratt_mas_ragil
ש"ח	סכום שנתי	סה"כ הכנסה- משכורת ותשלומים	Pratt_masc_vtash
ש"ח	סכום שנתי	שכר מעבודה במשמרות	Pratt_mishmarot
ש"ח	סכום שנתי	תשלומים פטורים מדמי ביטוח לאומי	Pratt_ptor_bl_lo102
ש"ח	סכום שנתי	תשלומים פטורים מדמי ביטוח לאומי (מעל גובה מסוים)	Pratt_ptor_mal_max
ש"ח	סכום שנתי	תשלומים הפטורים מתשלום מס	Pratt_ptor_mmas
10-רווק 20-נשוי 30-גרש 40- אלמן	מרשם תושבים	מצב משפחתי	Mt_kodmatzavishi
YYYYMMDD	מרשם תושבים	תאריך השינוי במצב המשפחתי	Mt_taarichmatzavishi
לפי פירוט ספר היישובים	מרשם תושבים	יישוב מגורים	Mt_semelyishuv

3.2.2 מבנה טבלאות מעביד

לאחר תהליך הקליטה, הושלמו נתונים למעבידים ממרשם עסקים, הזיווג נעשה לפי מספר תיק ניכויים מס הכנסה והוספו הנתונים הבאים: סוג עסק, תאריך פתיחה וענף כלכלי. אחר עסק שלא נמצא במרשם עסקים בין אם בשל סגירה או לחילופין שינוי מספר תיק, נערך חיפוש ידני ברשומות ביטוח לאומי ע"מ לקשר את הרשומה בין שנים ולהשלים ענף כלכלי. נכון להיום לא קיימת טבלה המקשרת בין מספרי תיקים שהוחלפו, תיק שהוחלף מוגדר כתיק אחר, דהיינו החלפת מספר תיק גורמת לכאורה להיווצרות מעביד חדש. קישור אוטומטי של תיק ישן לתיק חדש תושלם בעתיד.

הנתונים לגבי פדיון ותשומות חודשיים נקלטו מדו"חות העוסקים למע"מ. במרשם עסקים לכל עסק ישנם ארבעה מספרים מזהים: מספר יישות, תיק ביטוח לאומי, תיק ניכויים מס הכנסה, תיק מע"מ ובעבור כל קומבינציה ביניהם ניתן מספר עסק למ"ס ייחודי (ESEK_LMS).

בבניית בסיס הנתונים המפתח היחיד הקיים הוא מס' תיק ניכויים מס הכנסה, וייתכנו מקרים בהם לתיק ניכויים מס הכנסה אחד יהיו כמה מספרי עסק למ"ס, כלומר הקישור אינו חד חד ערכי. בשל כך השלמת נתוני הפידיון בוצעה כדלקמן: עבור תיק ניכויים מס הכנסה אשר קיים קישור חד חד ערכי למספר עסק למ"ס, הנתונים נלקחים מדו"חות העוסקים למע"מ. לתיקים בעלי הקישור המרובה נעשתה בדיקה אם כל נתוני הפדיון והתשומות החודשיים במהלך שנת המס זהים, אם כן – נלקחה רשומה אחת והכפילויות הושמטו (תיק מע"מ הוא זהה ואילו אחד או יותר המספרים המזהים האחרים שונה), אם נתוני הפדיון ו/או התשומות שונים נעשית סכימה של כל חודש בנפרד והסכום הוא הנתון המתאים (תיק מע"מ שונה). לאחר טיפול זה נבנו שתי טבלאות: האחת כוללת את הנתונים הקבועים של מעבידים, והשנייה את הנתונים השוטפים.

3.2.2.1 מאפיינים קבועים

בטבלה זו מופיעים כל המעבידים אשר לפחות באחת מן השנים 1999-2005 הופיעו בקובץ מס הכנסה שכירים, כל מעביד מופיע פעם אחת בלבד, לפי מספר תיק ניכויים מס הכנסה.

לוח 15: מבנה רשומה בטבלת מאפיינים קבועים של מעביד

שם משתנה	מהות השדה	הסבר/ מקור	ערכים
Koteret_mis_tik	מספר תיק ניכויים		9 ספרות
Taarich_pticha	תאריך פתיחה	מקור: מרשם עסקים	YYYYMMDD
Shem_maav	שם מעביד	מקור: טופס 126	
Anaf_kalkali	ענף כלכלי	מקור: מרשם עסקים	
Semel_yishuv	יישוב	סמל יישוב מיקום העסק. מקור: טופס 126	ספר היישובים

3.2.2.1 מאפיינים שוטפים

טבלה המכילה את כל המעבידים בעבור כל שנת מס, וכן פרטים הניתנים לשינוי. מספר העובדים הוא משתנה תחשיבי המשקלל את מספר המשרות ומספר חודשי העבודה אצל כל מעביד, סך הוצאות השכר אף הוא תחשיבי וזוהי למעשה סכימה של כל הוצאות המשכורת והתשלומים של המעביד לכל עובדיו בעבור משרות עבודה לשנת מס נתונה.

לוח 16: מבנה רשומה בטבלת מאפיינים משתנים של מעביד

שם משתנה	מהות השדה	הסבר/ מקור	ערכים
Koteret_mis_tik	מספר תיק ניכויים	9 ספרות	
Anaf_kalkali_maam	ענף כלכלי מע"מ	ענף כלכלי על פי דיווח מע"מ (נועד בעיקר להשלמה לעסקים שלא נמצא ענף כלכלי)	
Mis_ovdim	מספר עובדים	במונחי שנת עבודה למשרה	סכימה של מספר חודשי העבודה של כל עובדי המעביד וחלוקה ב-12
Masc_vtach	סך הוצאות שכר	סה"כ משכורת ותשלומים ששילם המעביד לכל עובדיו	משרות עבודה
N_divuah	מספר דיווחים	מספר הדיווחים למע"מ במהלך השנה	1-12
Pidion1...pidion12	פדיון ינואר עד דצמבר	שניים עשר משתנים נפרדים, על פי דיווח למע"מ	
Tsumot1...tsumot12	תשומות ינואר עד דצמבר	שניים עשר משתנים נפרדים, על פי דיווח למע"מ	

4. בסיס נתוני הכנסות פרט

גולת הכותרת של שילוב קובצי מס הכנסה הוא קובץ הכנסות פרט. קובץ אשר כל רשומה בו היא רשומת פרט, וכולל בתוכו את כל הפרטים במשק (המדווחים/מדווחים למס הכנסה) שנמצאו בקובצי שכירים ועצמאים.

בסיס נתונים מזווג עובד-מעביד, הנבנה ברמת משרה, ובסיס נתוני עצמאים, הנבנה ברמת תיק במס הכנסה, היוו ייסוד להקמת בסיס נתונים מאוחד שנתי לכל העובדים בעלי הכנסה מיגיעה אישית במשק, ברמת הפרט.

בסיס נתונים מכיל את פרטי ההכנסה ותשלומי החובה, תוך אבחנה האם מעבודה, אם לאו. פירוט הנתונים שונה בין שכירים ועצמאים: לשכיר קיים נתון ייחודי אודות חודשי העבודה השכירה ומספר משרות העבודה בהן החזיק במהלך השנה ואילו לעצמאי קיים פירוט רב יותר על מרכיבי ההכנסה החייבת במס.

4.1 תהליך הבנייה

ההנחה הבסיסית שקבעה את סדר איתור הרשומות בבסיס נתונים מזווג עובד-מעביד ובסיס נתוני עצמאים ולפיה נבחר מקור הנתונים הנמצאים בשני בסיסי נתונים אלה היא, שדיווח ישיר של הפרט לשלטונות המס על גבי טופס 1301 (כעצמאי) מהימן יותר מאשר אוסף הדיווחים העקיפים לגבי שכיר באמצעות מעבידיו. מתוקף הנחה זו, אם הנתון זמין הן מדוח עצמאי והן מקובץ מס הכנסה שכירים, הוא יילקח מהמקור הראשון.

שלב ראשון

המרת מידע מקובץ מס הכנסה שכירים ממבנה משרות למבנה של שכירים, תוך אבחנה בין משרות עבודה למשרות אחרות. לכל בעל משרות עבודה מרובות נעשתה סכימה של הכנסות מעבודה ותשלומי החובה, סכימה "אנכית" של חודשי עבודה (שלא יעלה על 12) וכך מספר משרות העבודה. לפרט שהיו לו משרות אחרות (שלא משרות עבודה), נעשתה סכימה של הכנסה ותשלומי מס הכנסה בגינן.

חודשי עבודה-סכימה "אנכית" של חודשי העבודה (עבד/לא עבד בכל אחד מ-12 החודשים) בכל משרות העבודה (מסוג 01,02,03).

הכנסה ממשכורת-סח"כ ההכנסות ממשכורת העבודה 01,02,03.

מס הכנסה מעבודה-סח"כ תשלומי מס הכנסה ממשכורת העבודה 01,02,03.

ענף כלכלי- לפי משרה עיקרית (הכנסה וחודשי עבודה). משרה עיקרית לפי חודשי עבודה היא המשרה בה החזיק הפרט זמן ממושך ביותר מכל משרותיו האחרות; משרה עיקרית לפי הכנסה היא המשרה שהניבה את השכר הגבוה ביותר מכל משרותיו של הפרט.

הכנסות אחרות-סך כל ההכנסות מכל המשרות שאינן מסוג 01,02,03.

מיסים אחרים-סח"כ תשלומי מס הכנסה מכל המשרות שאינן מסוג 01,02,03.

בסוף התהליך התקבלה רשומה ברמת פרט, על כל הכנסותיו ותשלומי מס השנתיים.

שלב שני

איתור עצמאי (לפי ת.ז. "בן זוג רשום") שנמצא כשכיר בבסיס נתונים מזווג עובד-מעביד ; אם נמצא, הרשומה מקבלת סמל איתור $st=1$.

שלב שלישי

עצמאי שלא נמצא בבסיס נתונים מזווג עובד-מעביד – הרשומה מקבלת סמל איתור $st=2$.

שלב רביעי

בני/ בנות זוג של בעלי התיק בקובץ עצמאים, הנמצאים בבסיס נתונים מזווג עובד-מעביד, הרשומה מקבלת סמל איתור $st=3$.

שלב חמישי

בני/ בנות זוג של בעלי התיק שלא נמצאו בבסיס נתונים מזווג עובד-מעביד, הרשומה מקבלת סמל איתור $st=4$.

שלב שישי

שכירים שלא אותרו בקובץ עצמאים – לא כבעלי תיק ולא כבני/ בנות זוגם – הרשומה מקבלת סמל איתור $st=5$.

בסוף התהליך התקבלה טבלה אחת שבה כל רשומה בה היא פרט, על כל הכנסותיו ותשלומי החובה השנתיים. מבנה זה נוח לא רק להפקת סטטיסטיקה של הכנסת מהמשתנים הכלולים בטבלה, אלא גם כאמצעי זיהוי וקישור רשומות לטבלאות שכירים ועצמאים, במידה ויהיה צורך לנתונים נוספים הנמצאים שם. יתרון נוסף של מבנה שכזה הוא נכונות לקישור משפחות מינהליות, הרי שני בני הזוג יכולים להימצא הן בקבצי שכירים והן בקבצי עצמאים ואיתורם במקורות אלה עשוי להיות מסובך ועתיר שגיאות.

4.2. מבנה בסיס הנתונים

נתוני הכנסות פרט נלקחו בחלקם מבסיס נתוני עצמאים ובחלקם מבסיס נתונים עובד-מעביד, להלן פירוט המשתנים הנמצאים בבסיס נתוני הכנסות פרט ומקורם.

לוח 15: מבנה בסיס נתוני הכנסות פרט, לפי סמל איתור בקבצי שכירים ועצמאים

סמל איתור					מהות/ הסבר	שם משתנה
5	4	3	2	1		
U	V	V	V	V	מספר ת.ז (9 ספרות)	Teudatzehut
U	V	V	V	V	הכנסה משכר	Hahnasasahar
U					הכנסה אחרת	Hahnasa_aheret
	V	V	V	V	הכנסה כוללת	Hahnasa_kolelet
	V	V	V	V	הכנסה עיסקית	Hahnas_iskit
	V	V	V	V	הכנסה הונית	Hahnasa_honit
	V	V	V	V	הכנסה הונית אחרת	Hahnasa_honit_aheret
	V	V	V	V	הכנסה ללא יגיעה אישית	Hahnasa_lelo_yishit
	V	V	V	V	הכנסה פטורה	Hahnasa_ptura
	V	V	V	V	הכנסת נכה	Hahnasat_nehe
U	V	V	V	V	מס הכנסה	Masmisahar
U					מס הכנסה מהכנסה אחרת	Mas_aher
U		U		U	ביטוח לאומי	Dmey_bl
U		U		U	ביטוח בריאות	Dmey_briut
	V		V	V	ענף כלכלי	Anaf_kalkali
U		U		U		Anaf_hodesh
U		U		U		Anaf_sahar
U		U		U	חודשי עבודה	Work_month
U		U		U	מספר משרות	Mispar_misrot

הערה: כאשר המקור הוא קובץ עצמאים-V, כאשר המקור הוא קובץ שכירים-U

לוח 16: אוכלוסיית בסיס נתוני הכנסות פרט, 1999-2004

סה"כ	סמל איתור					שנה
	5	4	3	2	1	
2,903,238	2,412,650	29,229	137,373	168,188	155,798	1999
3,023,772	2,505,758	30,873	145,116	177,559	164,466	2000
3,101,300	2,559,405	32,184	152,191	186,618	170,902	2001
3,154,340	2,592,317	32,622	157,909	195,018	176,474	2002
3,138,941	2,560,852	35,698	156,870	210,645	174,876	2003
3,154,579	2,572,142	39,294	153,824	215,038	174,281	2004

5. הצעה לפרסום "סטטיסטיקה של הכנסות ממקורות מנהליים"

הלוחות המוצגים בהמשך הינם הדגמה קטנה של אפשרויות העיבוד הגלום בבסיסי נתונים עליהם נכתב עד כה. הגיוון חובק תחומים ונושאים רבים, וכבר ניתן לראות את השימושים הרבים הנעשים בבסיסים אלה על ידי יחידות שונות בלמ"ס, בין אם לצורכי השלמת מידע (במפקד 2008, בבנייה ועדכון של המדד החברתי-כלכלי) ובין אם למחקר ופיתוח נושאים. מתוך שלל הנתונים הזמינים להצגה, נבחרו אלה המאפשרים השוואה לאומדנים המתפרסמים כיום בלמ"ס המופקים מנתוני הסקרים ונתוני המוסד לביטוח לאומי, אך בעיקר נתונים חדשים ייחודיים, אלה שלא יכולנו עד כה לפרסם בהיעדר גישה למקורות מנהליים של מס הכנסה.

5.1 מהקיים בעולם

בארצות רבות בעולם נהוג לפרסם נתונים סטטיסטיים המתבססים על מקורות מנהליים, בין אם על ידי הלשכות לסטטיסטיקה ובין אם על ידי גופים רשמיים אחרים של המדינה. להלן פירוט חלק מהמדינות והקישור לאתרי האינטרנט:

1. GBA (Gemeentelijke Basisadministratie): <http://www.gba.nl/>

הולנד: קובץ מנהלי הכולל אינפורמציה ביחס למאפיינים הדמוגרפיים והרכב משקי הבית של כל הפרטים בהולנד. הקובץ נאגר מהקבצים המינהליים הנמצאים בידי הרשויות המוניציפליות.

2. "SSB Banenbestand" (Social Statistical Database of Jobs):

<http://www.cbs.nl/en-GB/default.htm>

הולנד: קובץ מנהלי המשלב מידע ביחס לכלל קשרי התעסוקה ממגוון מקורות. שני המקורות העיקריים הינם נתוני המוסד לביטוח לאומי (Verzekeringsadministratie werknemers, VZA) וקבצים פיסקליים (Fibase) המציגים נתונים אודות מיסי הכנסה ששולמו בעבור תעסוקת הפרטים.

3. <http://www.ssb.no/english/>

נורבגיה: הלשכה המרכזית לסטטיסטיקה של נורבגיה. נתוני עובד-מעביד (מקושרים) – המבוססים על נתונים מס מנהליים ומסדי נתונים מנהליים שנתיים אודות ענפי התעשייה.

4. http://www.insee.fr/en/home/home_page.asp

צרפת: המכון הלאומי למחקרים סטטיסטיים וכלכליים. נתונים מנהליים של מרשם האוכלוסין ותעסוקה בצרפת.

5. <http://www.dst.dk/HomeUK.aspx>

דנמרק: הלשכה המרכזית לסטטיסטיקה של דנמרק. IDA: קובץ לונגיטודינלי הכולל מידע ביחס לכל הפרטים בני 15-74 (מאפיינים דמוגרפיים, השכלה, ניסיון בשוק העבודה, קביעות והכנסה) וביחס לכל המועסקים בכל המגזרים בדנמרק בין השנים 1980-2001.

6. http://www.scb.se/default_2154.asp

שוודיה: הלשכה המרכזית לסטטיסטיקה של שוודיה. נתונים מנהליים מקושרים של דמוגרפיה ותעסוקה בשוודיה.

7. <http://www.statistica.md/pubs.php?lang=en>

מולדביה: הלשכה המרכזית לסטטיסטיקה של מולדביה.
דוגמאות: נתונים מינהליים אודות דמוגרפיה (מרשם התושבים), מרשם פיסקלי (פנסיות, מיסוי עובדים).

קבצים מקובצים: REN IM-REN US (מיזוג קבצי מרשם עם נתונים פיסקליים).

8. <http://www.statistics.ge/>

גרוזיה (גיאורגיה): הלשכה המרכזית לסטטיסטיקה של גרוזיה (גיאורגיה).
דוגמאות: נתונים מינהליים אודות שוק העבודה, הוצאה פרטית לחינוך, הוצאה פרטית לבריאות, מרשם האוכלוסין – וקיום קבצים מקושרים.

9. <http://www.stat.kz/>

קזחסטן: הלשכה המרכזית לסטטיסטיקה של קזחסטן.
דוגמאות: נתונים מינהליים אודות הפרטים המקבלים סיוע לצורך יזמות, נתונים מינהליים (נתוני פרט) משוק העבודה ומערך התמיכות.

5.2 הגדרות והסברים ללוחות

משרות עבודה – רשומות בעלות סוג משרה 01,02,03, מתוך טבלאות עובד של בסיס נתונים מזווג עובד-מעביד.

משרה יחידה/עיקרית – (סוג משרה 01) משרה שבה מחשב המעביד את חבות המס לפי שיעורים שוליים ומעניק נקודות זיכוי לפי החוק. מותנה בהצהרת העובד שזאת משרה עיקרית או יחידה שלו.

משרה נוספת – (סוג משרה 02) משרה נוספת (על-פי הצהרת העובד), בה ינוכה מס בשיעור מרבי, או בשיעור אחר, לפי אישור פקיד השומה, ולא מובאות בחשבון הטבות מס.

משרה יומית – (סוג משרה 03) משרה בה המס מחושב על פי טבלת מיסוי יומי.

משרה אחרת – רשומות מסוג משרה אחר מתייחסות לתשלומים בגין יחסי עובד-מעביד בעבר (קצבה, פיצויי פיטורין, רווח ממכירת מניית עובד), או לתשלומים מיוחדים;

שכיר – פרט שהייתה לו במהלך שנת המס לפחות משרת עבודה אחת.

שכר למשרת עבודה – השכר הוא על בסיס שנתי. לצורך חישוב השכר החודשי למשרת עבודה נעשה שימוש בשכר השנתי אשר חולק במספר חודשי העבודה בכל משרה מתאימה.

חודשי עבודה לשכיר – סכימה "אנכית" של חודשי עבודה בכל משרות העבודה של הפרט: במידה ועבד בחודש מסויים במשרה אחת לפחות, אזי עבד ובאם לא עבד באף אחת מהמשרות אזי לא עבד; סכימה זו מוגבלת ל-12 חודשי עבודה לשנת מס.

שכר חודשי לשכיר¹ – החישוב מתבצע על ידי חלוקת השכר השנתי בכל משרת עבודה במספר חודשי העבודה בשנה בהם הועסק במשרה זו. במידה ועבד בחודש נתון – תזקף המנה שהתקבלה כשכרו החודשי לאותה משרה, אחרת – שכרו בחודש זה במשרה זו הוא אפס. על מנת לקבל את הכנסתו החודשית הכוללת מכל מעסיקיו נעשתה סכימת השכר

¹ שכר ממוצע לשכיר נמוך מהשכר המתקבל מסקר הכנסות מאחר ותקופת ההתייחסות כאן היא שנה לעומת סקר הכנסות בו התקופה היא שלושה חודשים (פורמן, 2005).

החודשי מכל משרותיו של הפרט. למעשה נעשתה החלוקה שכר למרות הידיעה כי קיימת שונות בשכר לאורך השנה.

ענף כלכלי – הענף הכלכלי העיקרי לכל שכיר נבחן בשני אופנים: האחד לפי המשרה בה החזיק זמן ממושך ביותר מכל משרותיו; בשני נבחנה המשרה אשר הניבה את השכר הגבוה ביותר. בלמעלה מ-95% הענף הכלכלי לפי שני האופנים זהה. הניתוח נעשה לפי ענף כלכלי שנקבע לפי מבחן השכר.

מספר עובדים (לקביעת קבוצות גודל המעביד) – סה"כ המשרות בעסק בשקלול מספר חודשי עבודה. סכום חודשי העבודה במשרות העבודה של כל עובדי המעביד מחולק ב-12. **סך תשלומי שכר** – סכימה של כל הוצאות המשכורת והתשלומים של המעביד לכל עובדיו (במשרות עבודה).

עצמאים – במסמך זה השימוש במונח עצמאים הוא לפי הגדרת מס הכנסה המובאת בסעיף 2.1 עצמאי על פי הגדרת מס הכנסה הוא אדם שאין מי שינכה בעבורו מס במקור, אלא הוא אמור לנכות לעצמו מס במקור. טופס 1301 הוא הדו"ח השנתי על הכנסות בארץ ובחו"ל שמוגש למס הכנסה על ידי הפרט או בא כוחו. סוגיות הקשורות להגדרת העצמאים נדונות בעבודתה של שלומית דרור-כהן.

כלל העובדים – כלל הפרטים אשר מופיעים בבסיס נתונים מס הכנסה שכירים או/ו בבסיס נתונים מס הכנסה עצמאים, אלה אוחדו לבסיס נתונים אחד הוא בסיס נתוני הכנסות הפרט.

הכנסה משכר ברוטו – לשכירים, הכנסות הנצמחות ממשרות העבודה.

הכנסה אחרת ברוטו – הכנסות ממשרות שאינן משרות עבודה, נובעות מיחסי עובד-מעביד לשעבר כמו קצבה, פיצויי פרישה וכו'.

הכנסה כוללת ברוטו – לעצמאים, סכום כל ההכנסות החייבות במס על שיעוריו השונים. ישנה הפרדה בין הכנסה מעבודה הכוללת הכנסה משכר והכנסה עסקית, לבין הכנסות אחרות.

הכנסה נטו – הכנסה כוללת ברוטו לעצמאי או הכנסה משכר ברוטו לשכיר לעיל בניכוי תשלומי החובה שהם מס הכנסה, ביטוח לאומי וביטוח בריאות.

ביבליוגרפיה

- דו"ח מנהל הכנסות המדינה 2005, מיסוי ישיר על יחידים- ממצאים עיקריים ממודל המס. שלומית דרור-כהן (2007), חקירת מגזר העצמאים והכנסתם בסקרים, למ"ס.
- אורלי פורמן (2005), שכירים-קובץ מס הכנסה חקירה ויישום. נייר טכני מס' 14, למ"ס.
- Yitzhaki.S. (2007), Pros and Cons for using Administrative records in Statistical Bureaus. United Nations Economic Commission for Europe, Conference of European Statisticians, Geneva.