

**פרויקט מסכם לבוגר במתמטיקה שימושית (B.Sc)
יבגני קרבצ'נקו**

The effect of telecommunication development on economic growth

**השפעה של התפתחות בטלקומוניקציות על צמיחה
כלכלית**

The curator: Mr. Nir Levi

המנחה: מר ניר לוי

The Presenter: Evgeniy Kravchenko

המגיש: קרבצ'נקו יבגני

תוכן עניינים

1	תקציר.....	3
2	הקדמה.....	4
3	סקירת ספרות.....	5
4	סקירת נתונים.....	7
5	מתודולוגיה.....	11
6	הצגת המודלים.....	12
7	תוצאות ומסקנות.....	20
8	סימולציה.....	21
9	השוואה בין התוצאות של שני המודלים: 2SLS ו-GMM.....	23
10	סיכום ומסקנות נוספות.....	23
11	נספח.....	24
12	ביבליאוגרפיה.....	27

1. תקציר

ענף הטלקומוניקציה הנו ענף חשוב מאוד בחיי היומ-יום של כל אחד ואחד מאיתנו. קשה לתאר את חיינו ללא טלפונים ניידים, אינטרנט, טלוויזיה וכו'. ענף הטלקומוניקציה התפתח והתרחב בצורה משמעותית החל מהמאה ה-19 והוא ממשיך להתפתח בקצב מהיר מאוד עד עצם היום הזה.

שיפור בענף הטלקומוניקציה הביא לתקשורת טובה יותר בין אנשים, בין חברות ובין מדינות. בעזרת הפיתוחים בטלקומוניקציה ישנו קיצור תהליכים וקיצור מרחקים. בעידן שלנו המידע נגיש מאוד לכל אחד מאיתנו. וזה כמובן מעלה את רמת הידע, כאשר עלייה ברמת הידע מביאה לעלייה ברמת החיים.

כמובן נשאלה השאלה על ידי רבים האם קיים קשר בין שיפור באמצעי תקשורת לבין רמת החיים. למראית העין, ישנו קשר ישיר בין שני הדברים. אם, למשל, ניקח מדינה לא מפותחת וענייה באפריקה, אז ניתן לראות כי גם רמת התקשורת נמוכה מאוד וגם רמת החיים מהנמוכות בעולם. ולעומת זאת, אם ניקח מדינה מפותחת, כמו ארצות הברית, אז ניתן לראות כי גם התקשורת מפותחת מאוד במדינה זו וגם רמת החיים מהגבוהות בעולם.

עבודה זו מתמקדת באמידת הקשר בין מידת הפיתוח של תשתיות התקשורת במדינות ה-OECD לבין שיעור הצמיחה בתוצר לנפש באותן המדינות, בשנים 2001-2009. כאשר שיעור הצמיחה בתוצר לנפש מהווה מדד לרמת החיים ומאפשר לנו לבצע השוואה בין המדינות השונות.

בעבודה זו השתמשנו בין היתר בשיטות אמידה מתקדמות המיועדות לפאנל, שמהווה מדגם של מדינות רבות על פני תקופת זמן לא ארוכה. התמקדנו במודלים שיודעים להתמודד עם פאנל דינאמי. דינאמיות היא עובדה כי המשתנה התלוי מופיע כאחד המשתנים המסבירים בפיגור של תקופה אחת.

2. הקדמה

בעבודה זו אנו נבחן כיצד תשתית התקשורת משפיעה על הצמיחה הכלכלית, כאשר המדד לצמיחה כלכלית הנו הצמיחה בתמ"ג.

ההשקעה בתשתיות התקשורת יכולה להוביל לצמיחה כלכלית במספר דרכים:

ראשית, ההשקעה בתשתית התקשורת עצמה מובילה לצמיחה, כי המוצרים כגון: כבלים, מתגים וכו' מביאים לעלויות בביקוש למוצרים ולשירותים המשמשים בייצור שלהם. למשל, ברגע שיש ליותר אנשים מכשיר סלולרי או אינטרנט אז כתוצאה מכך ליותר אנשים משתלם להשתמש בטכנולוגיה הזו (יותר אפליקציות ויותר תכנים אינטרנטיים), יותר מכך, ברגע שיש שיפור בתשתיות יש אפשרות לחברות התקשורת לספק שירות לקהל רחב יותר. בנוסף, ברגע שיש תשתית קיימת, העלות השולית להוספת משתמש חדש יורדת. לכן, קיימת השפעה הן על צד הביקוש והן על צד ההיצע.

בנוסף, היכן שמערכת הטלפון לא מפותחת מספיק, תקשורת בין חברות מוגבלת. עלויות העסקה, עלויות איסוף המידע ועלויות חיפוש אחר שירותים הן גבוהות. מסיבה זו, ככל שמערכת הטלפון משתפרת, עלויות של ביצוע העסקה יורדות בצורה משמעותית, ותפוקתן של פירמות מסויימות בסקטורים מסויימים של הכלכלה תצמחנה. לכן, ההשקעה בתשתית התקשורת והשירותים הנגזרים מספקים יתרונות משמעותיים: הנוכחות שלהם מאפשרת ייצור יעיל יותר. יתר על כן, החשיבות של אפקט זה עולה, ככל שאינטנסיביות המידע של תהליך הייצור עולה. כך, השקעה בתקשורת עלולה להוביל לשיפור בתחומים אחרים.

אנו נתרכז במדינות חברות בארגון הכלכלי הבינלאומי OECD. אנו נבדוק את ההשערה האם שיפור והשקעה בתשתיות התקשורת מביאים לצמיחה כלכלית שמתבטאת בצמיחה בתמ"ג. אנו נבדוק את ההשערות הללו לפי נתונים היסטוריים וננסה להסיק מסקנות לגבי העתיד.

ניתן לראות כאן בעיה של סיבתיות הפוכה, כלומר, לא בטוח שדווקא שיפור בטכנולוגיה משפיע על שיפור ברמת החיים ולא ההיפך. אנו לא נטפל בבעיה זו בפרויקט שלנו, אלא נתייחס לאי תלות בין מרכיבים שונים במשוואות שנטפל בהן. אנו נתייחס להשקעה בתשתיות התקשורת כאל משתנה אקסוגני במשוואת התמ"ג.

בפרויקט זה נעבוד עם שני מודלים אקונומטריים: המודל הראשון הינו מודל 2SLS עם אפקטים קבועים בשיטת First Difference, ואילו המודל השני שבו אנו משתמשים נקרא GMM עבור פאנל דינאמי שגם הוא מתמודד עם אפקטים קבועים. את המודלים הללו אנו נסקור בהמשך.

3. סקירת ספרות

מחקרים רבים נעשו בניסיון למצוא את הגורמים העיקריים לצמיחה ברמת החיים, שמתבטאת בעליית התמ"ג. מחקרים רבים בדקו, למשל, האם יש קשר בין צמיחה כלכלית בארה"ב לצמיחה כלכלית בעולם כולו. אבל מחקרים מעטים עסקו דווקא בקשר בין טלקומוניקציה לבין צמיחה ברמת חיים. ברצוני לסקור את המחקרים שנעשו עד כה בנושא זה ולהציג את המסקנות העיקריות אליהן הגיעו החוקרים.

המחקר הראשון שעסק ספציפית בנושא של פרויקט שלנו נעשה בשנת 1980 ע"י Hardy. הוא מראה כי ישנה תרומה עצומה מהשקעה בתשתיות הטלקומוניקציה. אבל המחקר הזה לא התייחס לבעיית הסיביות ההפוכה ולמתאם סדרתי בהפרעות האקראיות שקיים בין משתנה מסביר למשתנה מוסבר. מחקרים הבאים הראו כי כאשר שתי הבעיות הללו נשלטות בצורה אקונומטרית התרומה מההשקעות יורדת בצורה משמעותית.

ניתוח שלם יותר של הקשר בין תקשורת לבין צמיחה כלכלית מסופק ע"י Norton(1992). באמצעות שימוש בנתונים מ-47 מדינות בתקופה של 1957-1977, הוא בודק את ההשפעה של המספר הממוצע של טלפונים בין השנים 1957 ל-1977 על ממוצע שיעור הצמיחה השנתי, כאשר הוא מבצע פיקוח על המספר של טלפונים ב-1957 ובנסוף מבצע פיקוח על מספר משתני מקרו. נרטון מגלה כי התקשורת הוא משתנה חיובי ומובהק ומסכם כי קיומה של תשתית התקשורת מפחית את עלויות העסקאות.

מחקר אחר, שבוצע ב-1996 ע"י Pablo T. Spiller ו-Shane Greenstein בדק את ההשפעה של תשתית הטלקומוניקציה (מספר כבלי טלפון ומספר כבלי ISDN) על ביצועים כלכליים בארה"ב. המסקנה שאלה הגיעו שני החוקרים היא כי השקעה בתשתית התקשורת אחראית על חלק ניכר מהצמיחה בהכנסות העסקים בתחום שירותי התקשורת.

מחקר נוסף שנעשה בנושא הנו:

Telecommunications and economic development בשנת 2001 שבוצע ע"י : Roller and Waverman. החוקרים הגיעו למספר מסקנות: ראשית, ישנו קשר חיובי בין התפתחות בטלקומוניקציה לצמיחה כלכלית. שנית, מדינה מסויימת צריכה לצבור "מסה קריטית" בתשתית התקשורת על מנת ששיפור בטלקומוניקציות יתחיל להשפיע על צמיחתה הכלכלית. כלומר, אם נסתכל על מדינות מסויימות באפריקה, אשר בהן תשתית התקשורת כמעט ולא מפותחת, הן ימשיכו להיות עם רמת חיים נמוכה עד שהן לא יצליחו לצבור "מסה קריטית" בתשתית התקשורת. כלומר, על המדינות הללו להשקיע הון התחלתי גבוה בתשתיות התקשורת בהתחלה, על מנת שיהנו מתרומה מהשקעה זו בצמיחה כלכלית. שלישית, למדינות היותר מפותחות תהיה תרומה נמוכה יותר מהשקעה בתשתית הטלקומוניקציה מאשר למדינות המתפתחות.

כמו כן, מסקנה נוספת שנתקלתי בה במהלך סקירת המחקרים הקודמים היא כי השפעה של הטלקומוניקציה על הצמיחה הכלכלית תלויה לא רק במצב ההתפתחות בה נמצאת מדינה כלשהי, אלא גם באיזור בו המדינה נמצאת (ישנו הבדל משמעותי בין מדינות באפריקה למדינות באירופה).

בפרויקט שלנו, אנו נבצע מחקר משמעותי כאשר אנו נטפל בבעית האינדוגניות ובבעית המתאם הסדרתי. השיטות שנשנתמש בהן הן שיטות אקונומטריות מתקדמות, שאותן נסקור בהמשך. אנו ננסה לחזק את הדעה הרווחת עד כה כי ישנו קשר חיובי בין טלקומוניקציה לצמיחה כלכלית.

4. סקירת נתונים

מאגר הנתונים איתו עבדנו הוא מאגר שניתן להשיג באתר הרשמי של סטטיסטיקה של מדינות האיחוד האירופי. המאגר הזה כולל את הנתונים משנת 2001 עד 2009 במדינות האיחוד האירופי ומדינות חברות בארגון OECD.

עבדנו עם רשימה של 30 מדינות חברות בארגון הכלכלי הבין לאומי OECD:

אוסטרליה, אוסטריה, בלגיה, קנדה, צ'כיה, דנמרק, אסטוניה, פינלנד, צרפת, גרמניה, יוון, הונגריה, איסלנד, אירלנד, איטליה, יפן, לוקסמבורג, מקסיקו, הולנד, ניו זילנד, נורבגיה, פולין, פורטוגל, סלובקיה, ספרד, שבדיה, שווייץ, טורקיה, בריטניה וארצות הברית.

ישראל לא נכללה ברשימה זו כיוון שהיא חברה בארגון OECD רק החל מ-2010 ועדיין לא נצברו מספיק נתונים.

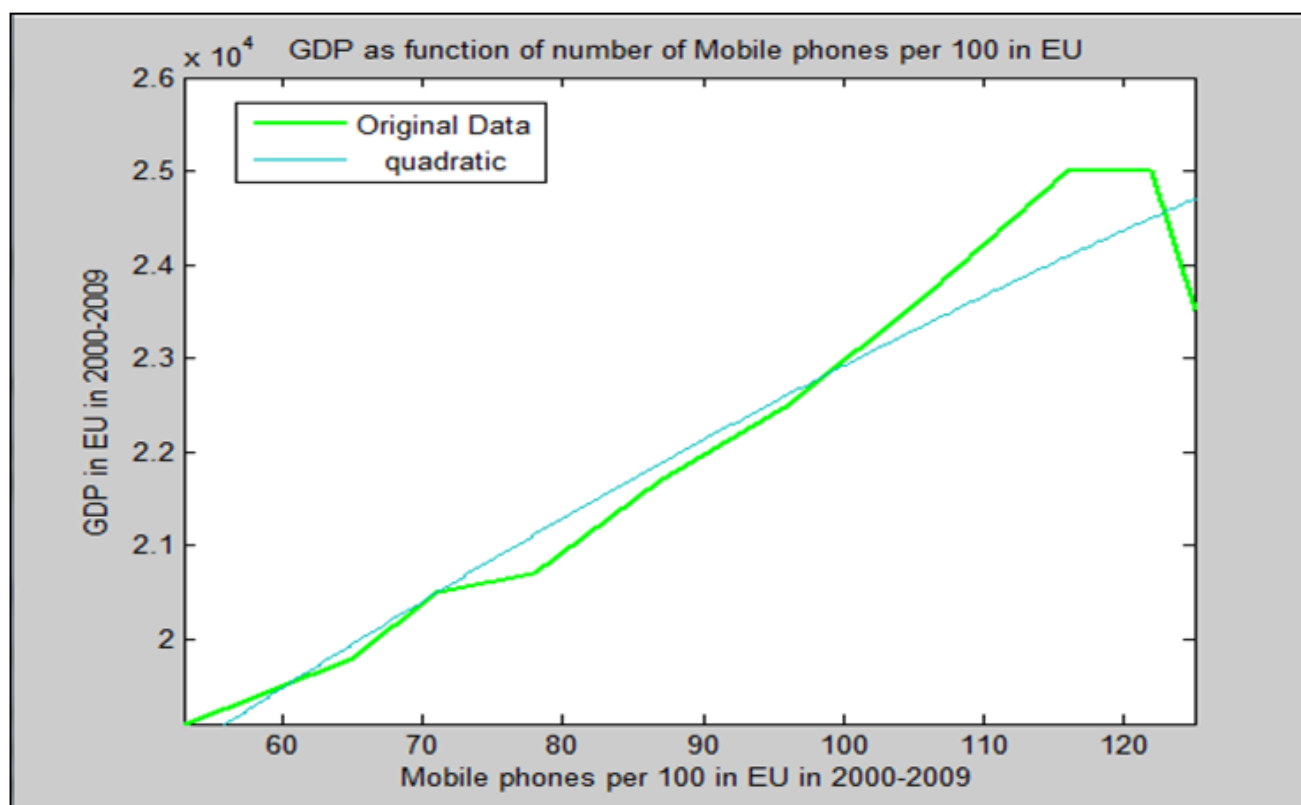
בחרנו לעבוד עם הנתונים מתוך המאגר לשנות 2001-2009 בקטגוריות הבאות:

תמ"ג כולל, מספר פלאפונים פר 100 תושבים, השקעה של הממשלה בתחום הטלקומוניקציה פר תושב, מספר כולל של דרכי גישה לטלקומוניקציה פר 100 תושבים, הכנסה מטלקומוניקציה במליוני דולרים, מספר מחשבים מחוברים לאינטרנט פר domain (הכוונה לסיומת של אתר במדינה מסוימת: למשל בישראל: .il), מספר משקי בית עם טלוויזיה באלפים, ומספר כולל של מנויי פס רחב פר 100 תושבים.

עבור הנתונים הללו אנו מניחים אקסוגניות (אי תלות עם הפרעה אקראית) על מנת שנוכל לעבוד עם המודלים שלנו.

ברצוננו לאמוד את תרומת המשתנים האלה לצמיחה בתמ"ג. בנוסף, בחרנו לעבוד עם עוד שני משתנים מסבירים, שהם צמיחה בתמ"ג בפיגור של תקופה אחת ותמ"ג בפיגור של תקופה אחת.

(1) נדגים את המגמה על פי הגרף הבא:

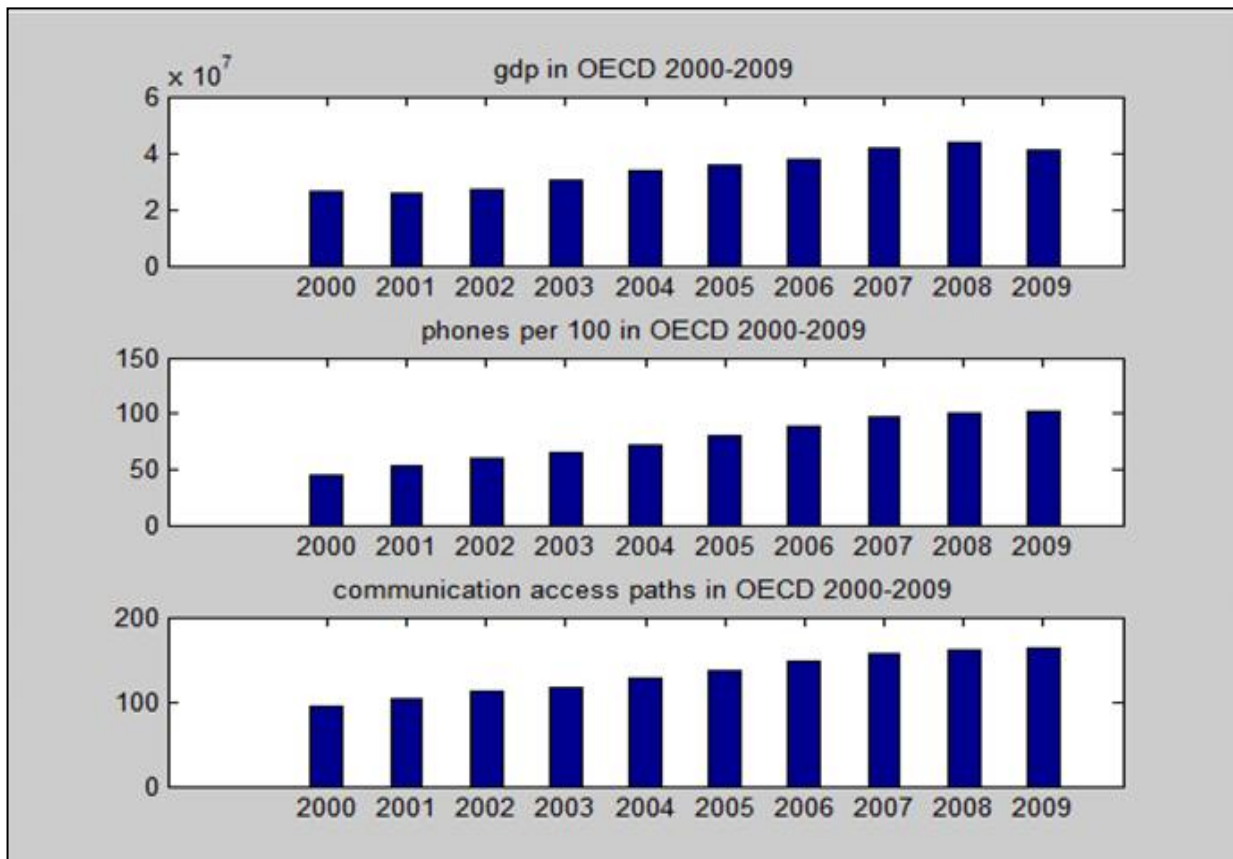


זהו גרף שמתאר את הקשר בין מספר פלאפונים פר 100 איש באיחוד האירופי בשנים 2000 עד 2009 לבין תמ"ג לנפש באיחוד האירופי באותן השנים.

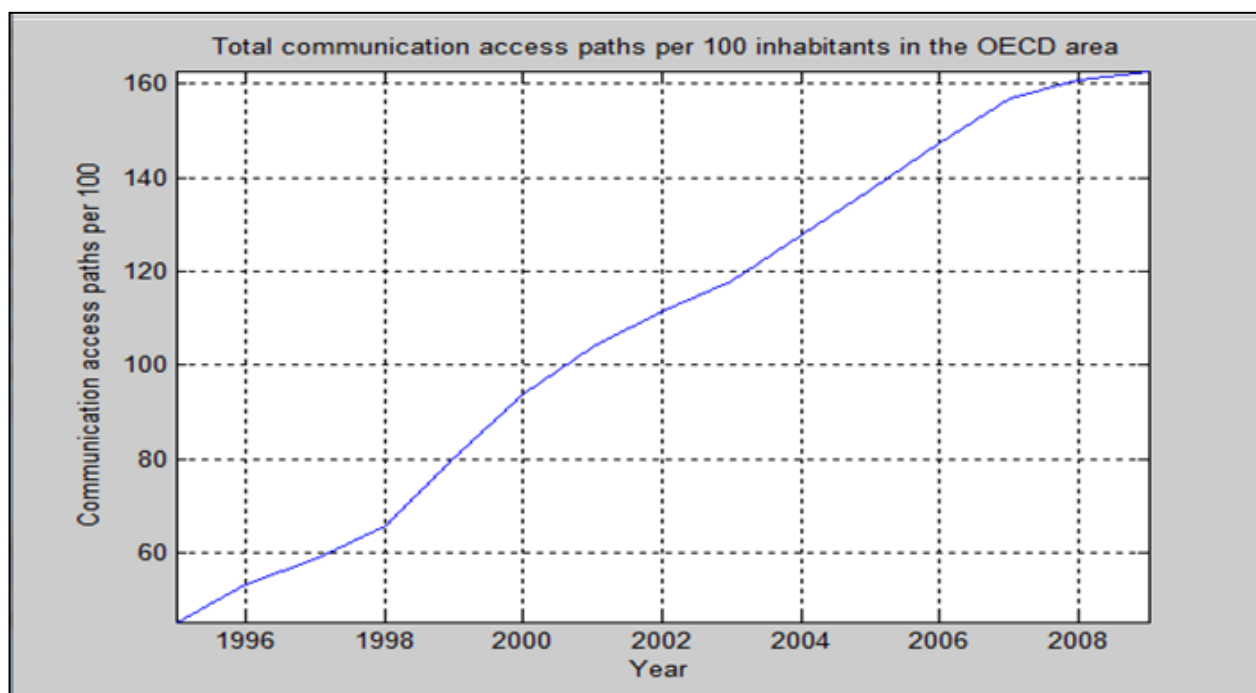
הקו הירוק מתאר את הנתונים המקוריים והקו הכחול מייצג את הקירוב הריבועי לנתונים הללו.

ניתן לראות כאן קשר כמעט לינארי בין מספר הפלאפונים פר 100 תושבים לתמ"ג לנפש. הירידה בשנת 2008-2009 קרתה, לפי עניות דעתי, עקב המשבר הכלכלי העולמי הידוע, אז התמ"ג לנפש ירד בצורה ניכרת ולמרות שמספר הפלאפונים לנפש המשיך לעלות, התמ"ג הושפע חזק יותר מגורמים כלכליים אחרים שגרמו להאטה ואף לירידה בתמ"ג לנפש.

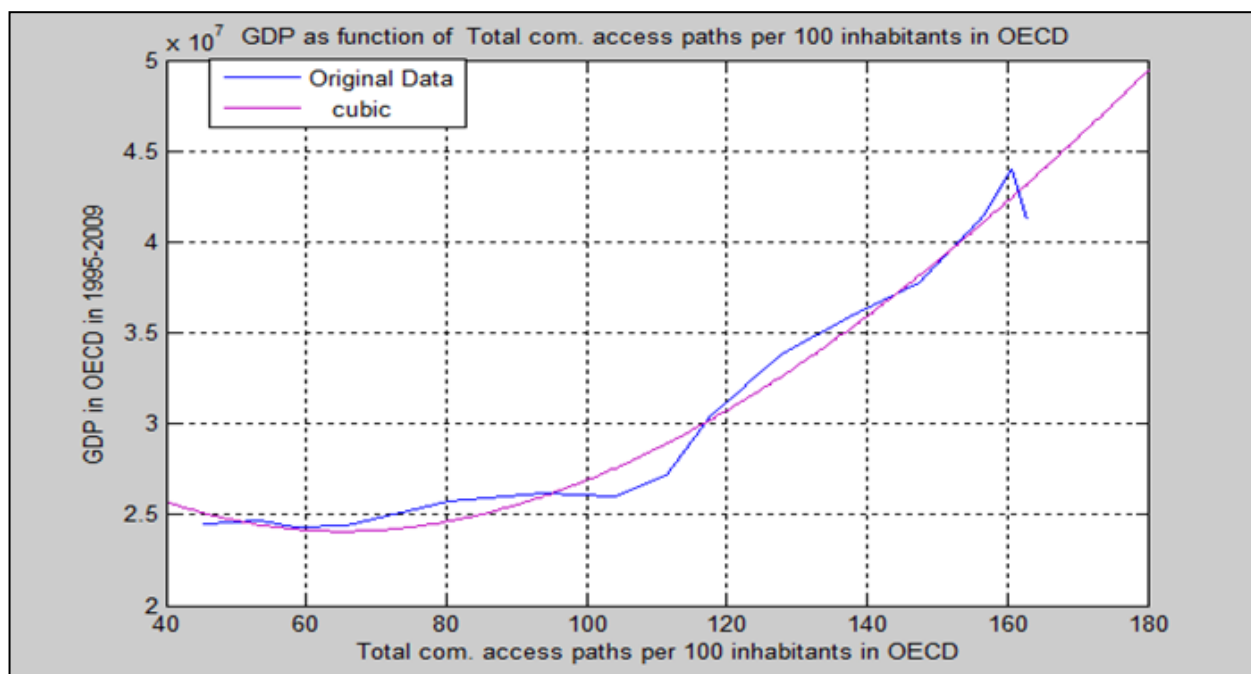
(2) גרף המקלות הבא מראה כי ישנה מגמה דומה לשלושת הרכיבים המופיעים בו. ניתן לראות כי במדינות החברות בארגון OECD בשנים 2000-2009 גם ה-GDP הכולל, גם מספר הפלאפונים פר 100 תושבים וגם מספר נקודות גישה פר 100 תושבים נמצאים במגמת עלייה.



(3) כעת נציג מגמה כללית יותר באמצעות נתונים משנת 1995 עד 2009: גם כאן ניתן לראות עלייה במספר נקודות גישה משנת 1995, כלומר רואים התרחבות הולכת וגוברת במספר נקודות גישה טלקומוניקציוניות.



(4) וכעת נראה את הקשר בין מספר נקודות הגישה פר 100 תושבים ל-GDP הכולל:



הגרף הכחול מתאר את הנתונים המקוריים והגרף הסגול מתאר את הקירוב הקובי. ניתן לראות כי, ככל שעולה מספר נקודות הגישה כך עולה גם רמת החיים שמתבטאת ב-GDP.

5. מתודולוגיה

הגדרה - משתנה אקסוגני (exogenous variable): משתנה מסביר המצוי מחוץ למערכת ואשר יש לו השפעה על המערכת. בניתוח כלכלי מתייחסים למשתנים האקסוגניים- לצורך הדין באותה מערכת כלכלית - כנתון. אלה הם המשתנים שאינם מוסברים על ידי המודל הכלכלי עצמו, אלא נכפים עליו על ידי כוחות חיצוניים ומשתנים אלה אינם נעים, בעיקרם, בשל התנועה הפנימית של המערכת להבדיל ממשתנים אינדוגניים. ההגדרה של אקסוגניות שקולה להגדרה כי: $cov(x_i, \varepsilon_i) = 0$, כלומר שאין מתאם בין המשתנה המסביר להפרעה האקראית. (השונות המשותפת בין משתנה אקסוגני להפרעה האקראית שווה ל-0).

הגדרה - משתנה אינדוגני (endogenous variable): משתנה מסביר הנקבע והמוסבר בתוך המסגרת של המודל הכלכלי הנתון. כלומר, זהו גורם שביסוד שלו נמצא התהליך הכלכלי הפנימי. משתנה זה, להבדיל ממשתנה אקסוגני, נע ביחס לאלמנטים אחרים במודל. הוא מתואם עם ההפרעה האקראית במודל.

הגדרה - fixed effect: היחידות הנמדדות (מדינות למשל) הינן הטרוגניות, לכל יחידה נמדדת ישנו חותך אישי (אפקט קבוע) שמיוחד דווקא ליחידה זו. למשל, אם אנחנו אומדים את המשוואה הבאה:

$$Y_i = \alpha + \alpha_i + \beta X_i + \varepsilon_i$$
 אז ההטרוגניות מתבטאת במשתנה α_i , שהוא מרכיב ייחודי לכל יחידה נאמדת. על מנת להיפטר מהטרוגניות עלינו לבצע פעולה שנקראת First Difference. בפעולה זו אנו מחסרים בין כל שתי משוואות עוקבות וכך נפטרים מההטרוגניות במשוואות הנאמדות.

הגדרה - שיטת OLS:

$$Y_i = \alpha + \alpha_i + \beta X_i + \varepsilon_i \quad \text{אנו מניחים מודל הבא:}$$

שיטת OLS - זוהי שיטה אקונומטרית בסיסית לאמידת מודלים כלכליים. בשיטת OLS מתקבלים אומדים בלתי מוטים ויעילים אם ורק אם מתקיימות 5 ההנחות הקלאסיות.

5 ההנחות הקלאסיות הן:

- נכונות המודל התאורטי. ההנחה הזו דורשת שברשותינו המודל הנכון.
- $E(\varepsilon_i) = 0$. תוחלת ההפרעה האקראית הנה 0. הנחה זו היא לצורך נוחות. מספיק לדרוש כי התוחלת של הפרעה אקראית תהיה קבועה.
- $V(\varepsilon_i) = \sigma^2, \sigma^2 < \infty$. הנחה זו דורשת כי השונות של הפרעה אקראית תהיה קבועה וסופית.
- $cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$. חוסר מתאם סדרתי – ההפרעות האקראיות אינן מתואמות.
- $cov(x_i, \varepsilon_i) = 0$. זוהי הנחת האקסוגניות של המשתנים המסבירים.

$$\beta_{OLS} = (X'X)^{-1} X'Y \quad \text{את אומדי OLS מוצאים לפי הנוסחא:}$$

נוסחא זו מתקבלת לאחר פתרון בעיית המזעור של סכום ריבועי השגיאות, שמחושבות כהפרשים בין הערכים החזויים של המשתנה התלוי לערכים האמיתיים שלו (הידועים מתוך המדגם הנתון).

שיטת OLS היא שיטה פשוטה ליישום אך היא דורשת כי כל 5 ההנחות הקלאסיות יתקיימו. בדרך כלל לא כל ההנחות הקלאסיות מתקיימות ולכן ישנן שיטות יותר מתקדמות שבהן אנו נשתמש בפרויקט שלנו. במודל שאנו אומדים ישנה אינדוגניות בגלל המרכיב הקבוע על פני זמן (החותך האישי) שיש לכל מדינה, כלומר מתקיים: $\text{COV}(x_i, \alpha_i) \neq 0$: ההפרעה האקראית ε_i כוללת בתוכה את α_i - החותך האישי של כל מדינה, כי כל מה שלא ניתן לאמידה נכנס בתוך ההפרעה האקראית. לכן כעת ישנה בעיית האינדוגניות.

כעת נעבור ליישום הפרויקט שלנו. המודל הראשי שעליו מתבססות המסקנות העיקריות הנו מודל GMM. כהקדמה למודל GMM נתבונן במודל 2SLS.

6. הצגת המודלים

מודל 2SLS: (2-Stage Least Square)

הצגה כללית של מודל 2SLS לאמידת מודל כלכלי:

מודל 2SLS יותר כללי ממודל OLS כיוון שהוא לא דורש את קיום כל ההנחות הקלאסיות. מודל זה אינו דורש את הנחת האקסוגניות של משתנים מסבירים. מודל OLS הבסיסי לא יכול להתגבר על בעיית ההטרוגניות בחותך, שכן לכל מדינה ישנו מרכיב ייחודי α_i . אך ניתן להשתמש בשיטת האפקטים הקבועים עבור OLS המאפשרת להפטר מהחלק ההטרוגני על ידי הפעלת שיטת ההפרשים. אולם מרגע שעברנו לאמידת משוואת ההפרשים יצרנו מתוך בנייה אינדוגניות במודל. ומסיבה זו גם שיטת OLS עם אפקטים קבועים אינה מתאימה לאמידת המודל. שכן, הצמיחה בשנה שעברה, שהיא משתנה מסביר במודל וגם המשתנה התלוי שלנו בפיגור של תקופה אחת, מתואמת עם ההפרעה האקראית בהפרשים. כעת נציג כיצד שיטת 2SLS מצליחה לפתור את הבעיות המתוארות כאן.

אמידה במודל 2SLS מתבצעת בשני שלבים:

- אומדים את שני המשתנים האינדוגניים: הצמיחה בפיגור והתוצר בפיגור. (פיגור של תקופה אחת)
- מבצעים אמידה עם שימוש במשתני עזר, שהם משתנים אקסוגניים ביחס ל-Y, כלומר שאינם מתואמים עם ההפרעה האקראית. בנוסף, משתנה עזר חייב להיות בעל מתאם גבוה עם המשתנה המסביר.

על מנת לחשב את אומדי 2SLS יש להשתמש בנוסחא המתאימה:

$$\hat{\beta}^{2SLS} = \left(X'Z (Z'Z)^{-1} Z'X \right)^{-1} X'Z (Z'Z)^{-1} Z'Y$$

ולאחר צימצומים מתקבלת נוסחא קצרה יותר:

$$\hat{\beta}^{2SLS} = \left(X'Z (Z'Z)^{-1} Z'X \right)^{-1} X'Z (Z'Z)^{-1} Z'Y.$$

כאשר בנוסחא זו: X זוהי מטריצות של משתנים מסבירים שבחרנו, במקרה שלנו בחרנו את המשתנים הבאים: המשתנים האינדוגניים - לוגריתם הצמיחה בפיגור ולוגריתם התוצר בפיגור, (פיגור בתקופה אחת) ו- 7 משתנים אקסוגניים, ראה נספח: "רשימת משתנים במודל 2SLS".

ו- Z זוהי מטריצת משתני עזר שבה המשתנים האנדוגניים, לוגריתם הצמיחה בפיגור ולוגריתם התוצר בפיגור, מוחלפים במשתני עזר אקסוגניים: לוגריתם הצמיחה ולוגריתם התוצר כאשר שניהם בפיגור של שתי תקופות, בהתאמה. (נתונים בפיגור של שתי תקופות – מתקיימת הנחת המתאם הגבוה עם המשתנים המסבירים האינדוגניים), ראה נספח: "רשימת משתנים במודל 2SLS".

אנו משתמשים בשיטת Fixed Effect שמשתמשת בהפרשים בין תקופות עוקבות (First Differences). באמצעות שיטה זו אנו מתגברים על המרכיבים הייחודיים לכל מדינה (α_i ו- α_0) על ידי חיסור בין כל שתי משוואות עוקבות.

First Difference result :

$$\begin{aligned} & \begin{cases} \ln gr_{it} = \alpha_0 + \alpha_i + \alpha_1 \ln gr_{it-1} + \alpha_2 \ln gdp_{it-1} + \alpha_3 \ln phones_{it} + \dots + \alpha_9 \ln broad_subs_{it} + \varepsilon_{it} \\ \ln gr_{it-1} = \alpha_0 + \alpha_i + \alpha_1 \ln gr_{it-2} + \alpha_2 \ln gdp_{it-2} + \alpha_3 \ln phones_{it-1} + \dots + \alpha_9 \ln broad_subs_{it-1} + \varepsilon_{it-1} \end{cases} \\ \Rightarrow \Delta \ln gr_{it} &= \alpha_1 \Delta \ln gr_{it-1} + \alpha_2 \Delta \ln gdp_{it-1} + \alpha_3 \Delta \ln phones_{it} + \dots + \alpha_9 \Delta \ln broad_subs_{it} + \Delta \varepsilon_{it} \end{aligned}$$

לאחר שביצענו הפרש בין כל 2 תקופות עוקבות לכל אחת מ-30 מדינות, קיבלנו מערכת משוואות בהפרשים. מצד אחד, ע"י שימוש בשיטת האפקטים הקבועים (Fixed Effects) נפטרו מהחותר האישי (ההטרוניות), ומצד שני, לאחר הפעלת שיטת ההפרשים קיבלנו מתוך בנייה מתאם סדרתי בהפרעות האקראיות בהפרשים, ובנוסף, יצרנו בעיית אנדוגניות, כיוון שכעת ישנו מתאם בין המשתנה המסביר, לוגריתם הצמיחה של שנה שעברה (שהוא משתנה תלוי בפיגור של תקופה אחת) לבין ההפרעה האקראית בהפרשים. ניתן להתגבר על הבעיה בשיטת 2SLS, אולם היא פחות יעילה משיטות אמידה אחרות שנציג בהמשך, שכן בשיטת 2SLS נגיע לאומדים עם שונות גבוהה.

כך נוצרה בעיה של אנדוגניות במשוואה הנאמדת שעברה למונחים של הפרשים (First Difference):

$$\text{cov}(\Delta \ln gr_{it-1}, \Delta \varepsilon_{it}) = \text{cov}(\Delta \varepsilon_{it-1}, \Delta \varepsilon_{it}) = \text{cov}(\varepsilon_{it-1} - \varepsilon_{it-2}, \varepsilon_{it} - \varepsilon_{it-1}) \neq 0$$

בעיית המתאם הסדרתי בהפרעות בהפרשים, נובעת מ- $\text{cov}(\Delta\epsilon_{it-1}, \Delta\epsilon_{it}) \neq 0$

קיים איבר משותף ולכן השונות המשותפת אינה שווה לאפס ומתקיים: $\text{cov}(\epsilon_{it-1}, -\epsilon_{it-1}) \neq 0$

בעיית מתאם סדרתי זו לא הייתה קיימת לפני כן, שכן הנחנו שמתקיים: $\text{cov}(\epsilon_{it}, \epsilon_{it-1}) = 0$

לסיכום, מודל 2SLS - מייצר אומדים עקיבים - מתכנסים לגודלם האמיתי כאשר המדגם שואף לאינסוף. אך, האומדים שמתקבלים אינם יעילים, כלומר השונות שלהם אינה מינימאלית. והוספה של משתנים מסבירים ומשתני עזר נוספים לא תעזור להתגבר על בעיית המתאם הסדרתי שנוצר. הוספת משתני עזר נוספים, תביא לניצול פחות יעיל של הנתונים כיוון שאנו כופים את אותם משתני עזר לכל תקופה. אנו לא יוצרים קבוצות משתני עזר נפרדות לכל תקופה כפי שנעשה את זה בשיטת GMM.

הנתונים שעבדנו איתם מיוצגים בקובץ Excel: ראה נספח - טבלה מס 1.

רשימה זו כוללת נתונים ל-30 מדינות על פני 7 שנים בהן נאספו הנתונים. נאלצנו לצמצם את מספר התקופות כיוון שהשתמשנו בשיטת ההפרשים. לכן הנתונים מתחילים משנת 2003 ולא משנת 2001. משתנה העזר שלנו הוא בפיגור של שתי תקופות לכן מתחילים לאמוד משנת ה-2003. אלו הם אותם משתנים מסבירים שבחרנו לעבוד איתם במודל השני – מודל GMM.

על מנת לראות את הקוד של התכנית בעזרתה חישבנו את האומדים עבור מודל 2SLS ראה נספח, תכנית מס' 1.

הגדרה - מודל GMM (General Method of Moments)

מודל GMM בשילוב עם אפקטים קבועים יוצר אומדים עקיבים ויעילים תוך ניצול אופטימאלי של הנתונים שבידינו, ראה מאמר של Arrelano and Bond (1991). הוכח בעבר כי האומדים שמתקבלים על ידי שיטת GMM יותר יעילים ביחס לאומדים המתקבלים לפי שיטת 2SLS, כלומר שונותם קטנה יותר. הייחודיות של שיטת GMM על פני שיטת 2SLS היא התאמת משתני עזר ייחודיים לכל תקופה בנפרד. כעת נציג את השיטה הבסיסית של GMM ללא האפקטים הקבועים.

פונקצית המטרה שלנו היא: $Q_T(\beta) = g_T(\beta)' W_T g_T(\beta)$

פונקצית מטרה זו היא סקלר ואנו מחפשים את וקטור האומדים $\hat{\beta}$ שימזער את פונקצית המטרה, כאשר W_T זוהי מטריצת משקולות. במודל GMM יש להגדיר את המומנטים, שאלה המשוואות שאנו רוצים להביא לאפס, ע"י מציאת וקטור $\hat{\beta}$ אופטימאלי. במדגם המומנטים (המשוואות) מסומנים כך:

$$g_T(\hat{\beta}) = \frac{1}{T} \begin{bmatrix} Z'_1(Y - X\hat{\beta}) \\ Z'_2(Y - X\hat{\beta}) \\ \vdots \\ Z'_P(Y - X\hat{\beta}) \end{bmatrix} = \frac{1}{T} Z'(Y - X\hat{\beta})$$

g – תנאי אורתוגונליות: אוסף של משוואות שרוצים להביא ל-0.

X – מטריצת המשתנים המסבירים המקוריים שבחרנו אותם כדי להסביר את המשתנה התלוי Y.

Z – מטריצת משתני העזר שהינם משתנים אקסוגניים, שאינם מתואמים עם ההפרעה האקראית. (בשיטת אמידה זו ישנם שני סוגים של משתני עזר: אותם משתני עזר שאנו מניחים כי הם באמת אקסוגניים והם משותפים לכל התקופות ומשתני עזר נוספים שאנו מתייחסים אליהם כאל אקסוגניים כי הם נקבעו בעבר הרחוק, predetermined, והם מותאמים לכל תקופה בנפרד).

P – מספר מומנטים (משוואות) במודל.

K – מספר משתנים מסבירים במודל, 9 במקרה שלנו.

N – מספר מדינות (פאנלים), 30 במקרה שלנו.

T_i – מספר תקופות למדינה ה- i.

$\hat{\varepsilon} = Y - X\hat{\beta}$ – זוהי ההפרעה האקראית בעת האמידה.

פירוט השיטה: ברצוננו לאמוד מודל עם מערכת של P משוואות (מומנטים). על מנת להבטיח זיהוי של האומדים, אנו דורשים כי התנאי הבא יתקיים: $K \leq P$. אם מתקיים כי $K < P$ יש לנו עודף של מומנטים, כלומר מתקיים זיהוי ביתר. ישנה בעיה להביא כל כך הרבה משוואות בו זמנית לאפס, כי עבור ערכי β מסוימים קבוצת מומנטים מסוימת תתקרב לאפס ואילו אחרות לא יהיו קרובות מספיק לאפס. לשם התגברות על בעיית הזיהוי הזו, אנו צריכים מטריצת משקולות שמעניקה לכל משוואה משקל ביחס הפוך לשונותה הנאמדת של המשוואה. ולבסוף מתקבל סקלר שהוא פונקציה המטרה של ה-GMM שאת הסקלר הזה אנו מעוניינים להביא לאפס. אלמלא מטריצת המשקולות, במקום להביא סקלר לאפס היה עלינו להביא וקטור של מומנטים (כל מומנט מופיע כסקלר) לאפס ואז צריך קריטריון שיקבע אילו רכיבים בוקטור חשובים יותר ואילו פחות.

אנו דורשים כי יתקיימו תנאי האורתוגונליות בין Z להפרעה האקראית:

$$\text{cov}(Z', Y - X\beta) = 0$$

$$\text{cov}(Z', Y - X\beta) = \text{cov}(Z', \varepsilon) = E(Z' * \varepsilon) - E(Z') * E(\varepsilon) = 0$$

$$E(\varepsilon) = 0: \text{ידוע כי}$$

לכן נשאר לדרוש כי יתקיים

$$E(Z' * \varepsilon) = 0$$

$$\begin{cases} E[Z_1'(Y - X\beta)] = 0 \\ \vdots \\ E[Z_p'(Y - X\beta)] = 0 \end{cases} \quad \text{בידינו } P \text{ משוואות ואנו דורשים כי לכל אחת יתקיים:}$$

ברצוננו למזער את פונקציית המטרה שלנו לכן אנו נגזור את פונקציית המטרה שלנו לפי וקטור β ונשווא את תוצאת הגזירה לאפס:

$$Q_T(\beta) = g_T(\beta)' W_T g_T(\beta) = (T^{-1} Z'(Y - X\beta))' W_T (T^{-1} Z'(Y - X\beta))$$

$$= T^{-2} (Y'Z W_T Z'Y - 2\beta' X'Z W_T Z'Y + \beta' X'Z W_T Z'X \beta)$$

$$\frac{\partial Q_T(\beta)}{\partial \beta} = -2T^{-2} X'Z W_T Z'Y + 2T^{-2} X'Z W_T Z'X \beta = 0$$

$$\hat{\beta}_{GMM}(W_T) = (X'Z W_T Z'X)^{-1} X'Z W_T Z'Y$$

קיבלנו צורה כללית עבור האומדים במודל הליניארי של שיטת GMM, שעליה מתבססות הוריאציות השונות של שיטת GMM, אותן יישמנו. בכל אחת משיטות GMM מתבצעים שני שלבים של אמידה. כאשר בשלב הראשון מחשבים את האומדים באמצעות מטריצת משקולות שמתקבלת לפי ניחוש מושכל ובשלב השני משתמשים במטריצת משקולות שמחושבת על פי הנתונים שבידינו ועל פי התוצאות של השלב הראשון. הכלל לחישוב מטריצת משקולות של השלב השני במודל הליניארי של GMM הנו: משוואה עם שונות קטנה יותר מקבלת משקל גבוה יותר במטריצת המשקולות החדשה, כיוון ששונות קטנה יותר מעידה על דיוק רב יותר.

אם ניזכר בשיטת 2SLS, אז ניתן לראות כי בשיטה זו האומדים מחושבים לפי הנוסחא:

$$\hat{\beta}^{2SLS} = (X'Z (Z'Z)^{-1} Z'X)^{-1} X'Z (Z'Z)^{-1} Z'Y.$$

ניתן לראות כי מטריצת המשקולות במקרה זה שווה ל: $W_T = (Z'Z)^{-1}$

במקרה של מודל 2SLS מטריצת משקולות מעניקה משקל גבוה יותר לערכים גבוהים יותר כיוון שאנו נקבל טעות קטנה יותר במושגים של אחוזים עבור ערך גבוה יותר של התצפיות. לדוגמא: בן אדם מקבל משכורת גבוהה והטעות באמידה הנה 5 יחידות כסף, בן אדם נוסף מקבל משכורת נמוכה והטעות באמידה הנה 5 יחידות כסף. עבור בן אדם עם משכורת גבוהה יותר אותה הטעות פחות משמעותית לעומת הבן אדם עם משכורת נמוכה יותר.

מודל First Difference GMM

החוקרים (Arrelano and Bond (1991) פיתחו שיטת אמידה המיועדת למדגמים מסוג פאנל כאשר המודל הינו מודל דינאמי: משתנה תלוי בפיגור מופיע כאחד המשתנים המסבירים. במודל הזה מטריצת X מכילה את המשתנים המסבירים בהפרשים. כאן אנחנו אומדים מערכת של P משוואות (מומנטים), ומספרן גבוה ממספר המשתנים המסבירים במודל (זיהוי ביתר). יש לנו 7 תקופות בהן אנחנו מבצעים אמידה ולכל תקופה אנו מייצרים שני משתני עזר ייחודיים ובנסוף לכל התקופות ישנם גם משתני עזר אקסוגניים משותפים. כלומר, יש בידינו $P=21=7+7*2$ משוואות ו $K=9$ משתנים מסבירים, ומתקיים כי $9 < 21$ וזהו זיהוי ביתר. היה באפשרותינו להגדיל את כמות משתני עזר ייחודיים לכל תקופה ככל שאנחנו מתקדמים עם השנים ולצרף לתקופה הבאה גם את משתני עזר של תקופה קודמת, אך בחרנו לא לנצל שיטה זאת כיוון שהיא יוצרת סירבול רב.

$$\hat{\beta}_{FD-GMM} = ((\Delta X'Z)V^{-1}(Z'\Delta X))^{-1}(\Delta X'Z)V^{-1}(Z'\Delta Y)$$

כאשר:

ΔX - מטריצת המשתנים המסבירים בהפרשים.

Z - מטריצת משתני עזר שבה המשתנים האינדוגניים מוחלפים במשתני עזר אקסוגניים:

משתנים בפיגור של שתי תקופות משמשים כאקסוגניים. משתנים שהם אקסוגניים מופיעים בהפרשים במונחי התקופה הנוכחית, החל מ- $\Delta \ln Phones_{i3}$.

$$\Delta X_i = \begin{bmatrix} \Delta \ln GR_{i2} & \Delta \ln GDP_{i2} & \Delta \ln Phones_{i3} & \cdots \\ \Delta \ln GR_{i3} & \Delta \ln GDP_{i3} & \Delta \ln Phones_{i4} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Delta \ln GR_{iT-1} & \Delta \ln GDP_{iT-1} & \Delta \ln Phones_{iT} & \cdots \end{bmatrix}$$

*נקודות מצינויות 6 משתנים אקסוגניים נוספים בהפרשים שנכנסו למטריצת משתני עזר.

$$Z_i = \begin{bmatrix} \ln GR_{i1} & \ln GDP_{i1} & 0 & 0 & \cdots & 0 & 0 & \Delta \ln Phones_{i3} & \cdots \\ 0 & 0 & \ln GR_{i2} & \ln GDP_{i2} & \cdots & 0 & 0 & \Delta \ln Phones_{i4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \ln GR_{iT-2} & \ln GDP_{iT-2} & \Delta \ln Phones_{iT} & \cdots \end{bmatrix}$$

- $W_T = V^{-1} = (Z'KZ)^{-1}$ זוהי מטריצת משקולות לפני ייעול שמשתמשים בה בשלב ראשון,

$$K = \begin{pmatrix} K_1 & \dots & 0 \\ : & \ddots & : \\ 0 & \dots & K_p \end{pmatrix}, K_i = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

K - זוהי מטריצה שמאפשרת לנו לשקלל את משתני העזר, תוך התחשבות בשונות של משוואות המומנטים, בשל קיומו של מתאם סדרתי מסדר ראשון בהפרעות כשהן במונחים של הפרשים. זוהי מטריצת הניחוש המושכל שהוצגה מקודם. מטריצה זו מגדירה קשר בין ההפרעות האקראיות בתוך משוואת הפרשים. הוכח כי מטריצה זו נותנת את התוצאות הטובות ביותר בשלב הראשון.

בשלב השני של מודל זה מטריצת המשקולות החדשה מעניקה משקל נמוך יותר למשוואה שיש לה שונות גבוהה יותר, זאת בהתבסס על השגיאות הנאמדות.

$e_i = Y - X\hat{\beta}_{FD-GMM}$ - אלה הן השגיאות הנאמדות ובאמצעותן נבנית מטריצת המשקולות של השלב השני.

$$\hat{\beta}_{Efficient-FD-GMM} = \left((\Delta X'Z) \left(\sum_{i=1}^N Z_i' e_i e_i' Z_i \right)^{-1} (Z' \Delta X) \right)^{-1} (\Delta X'Z) \left(\sum_{i=1}^N Z_i' e_i e_i' Z_i \right)^{-1} (Z' \Delta Y)$$

מודל System GMM

מודל זה מבוסס על המודל של Arellano and Bover (1995). במודל זה אנו אומדים מערכת משוואות System GMM, כאשר אחת המשוואות הנה המשוואה המקורית והמשוואה השנייה הנה משוואת הפרשים. שתי המשוואות ניתנות לאמידה באותה קלות בה נאמדת המשוואה הבודדת במודל הקודם. אולם הפעם, משתני עזר נבחרים בצורה שונה מזו של המודל GMM הקודם: מסתבר שמשתני עזר הטובים ביותר למשוואה המקורית הנם המשתנים בהפרשים, ואילו משתני העזר הטובים ביותר עבור משוואת הפרשים הם דווקא המשתנים המקוריים. בדומה למודל הקודם, כאן ישנם 9 משתנים מסבירים וסה"כ $2*(2*7+7) = 42$ מומנטים וזהו זיהוי ביתר (הכפלה נוספת ב-2 מתבצעת כיוון שכעת יש לנו מערכת של שתי משוואות: System GMM). האומדים נכפים להיות זהים לשתי המשוואות ומחושבים בצורה הבאה:

$$\hat{\beta}_{SYS-GMM} = \left((\Delta X_S' Z_S) W_{SYS}^{-1} (Z_S' \Delta X_S) \right)^{-1} (\Delta X_S' Z_S) W_{SYS}^{-1} (Z_S' \Delta Y_S)$$

אנו אומדים את מערכת המשוואות הבאה:

$$\begin{cases} \Delta \ln_GR_{it} = \beta_1 \Delta \ln_GR_{it-1} + \beta_2 \Delta \ln_gdp_{it-1} + \beta_3 \Delta \ln_phones_{it} + \dots + \beta_9 \Delta \ln_broad_subs_{it} + \Delta \varepsilon_{it} \\ \ln_GR_{it} = \alpha + \beta_1 \ln_GR_{it-1} + \beta_2 \ln_gdp_{it-1} + \beta_3 \ln_phones_{it} + \dots + \beta_9 \ln_broad_subs_{it} + \varepsilon_{it} \end{cases}$$

כפי שנאמר, החוקרים כופים את אותם האומדים לשתי המשוואות, למעט החותך, שבחרנו שלא לאומדו.

כעת נפרט כיצד נבנות מטריצות משתני העזר במודל זה.
מטריצת משתני עזר למשוואת הפרשים (Difference):

$$Z_{Di} = \begin{bmatrix} \ln GR_{i1} & \ln GDP_{i1} & 0 & 0 & \cdots & 0 & 0 & \Delta \ln Phones_{i3} & \cdots \\ 0 & 0 & \ln GR_{i2} & \ln GDP_{i2} & \cdots & 0 & 0 & \Delta \ln Phones_{i4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \ln GR_{iT-2} & \ln GDP_{iT-2} & \Delta \ln Phones_{iT} & \cdots \end{bmatrix}$$

*נקודות מציינות 6 משתנים אקסוגניים בהפרשים נוספים שנכנסו למטריצת משתני עזר.

מטריצת משתני עזר למשוואה המקורית (Levels):

$$Z_{Li} = \begin{bmatrix} \Delta \ln GR_{i1} & \Delta \ln GDP_{i1} & 0 & 0 & \cdots & 0 & 0 & \ln Phones_{i1} & \cdots \\ 0 & 0 & \Delta \ln GR_{i2} & \Delta \ln GDP_{i2} & \cdots & 0 & 0 & \ln Phones_{i2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \Delta \ln GR_{iT-2} & \Delta \ln GDP_{iT-2} & \ln Phones_{iT} & \cdots \end{bmatrix}$$

במודל FD-GMM ישנו $Z = Z_D$, ואילו במודל System-GMM המטריצה הבאה משמשת כמטריצת

$$Z_{SYS} = \begin{bmatrix} Z_D & 0 \\ 0 & Z_L \end{bmatrix} \text{ משתני עזר למערכת של שתי משוואות:}$$

כעת מתייחסים למדגם כאילו ונוספו לו תצפיות חדשות. התצפיות החדשות הן הנתונים המקוריים (Levels) שמתווספים לנתונים שבהפרשים. $X_{SYS} = [X; X_Level]$. בעצם קיבלנו עוד אינפורמציה מהנתונים שהיו בידינו ולכן זהו ניצול יעיל יותר של המידע הקיים.

$$W_{SYS} = \begin{bmatrix} K & H_1 \\ H_2 & I \end{bmatrix} \text{ מטריצת המשקולות של השלב הראשון הנה:}$$

W_{SYS} - זוהי מטריצת ניחוש מושכל שהוצעה על ידי שני החוקרים. כאשר K זוהי מטריצה שמגדירה קשר בין ההפרעות האקראיות בתוך משוואת הפרשים, H_1 and H_2 אלה הן מטריצות שמגדירות את הקשר בין ההפרעות של שתי המשוואות, כלומר בין ההפרעה בהפרשים להפרעה במשוואת המשתנים המקוריים. ואילו I זוהי מטריצה שמגדירה את הקשר בין ההפרעות האקראיות בתוך משוואת המשתנים המקוריים ומכיוון שלא הנחנו מתאם סדרתי היא מנרמלת ל- I .

2 Step System GMM: זהו שלב שבו מטריצת משקולות נבנית על פי הנתונים שבידינו. מטריצת משקולות מחושבת הפעם כפונקציה של השאריות שהתקבלו בשלב הראשון. ומכיוון שאנו משתמשים בנתונים אמיתיים ולא ניחוש מושכל אנו מגיעים לתוצאות עוד יותר מדויקות.

7. תוצאות ומסקנות

במודלים שאמדנו ניסינו להבין מהי ההשפעה של משתנה מסביר כלשהו על המשתנה המוסבר – צמיחה בתמ"ג. ברגע שאנו משתמשים במשתני עזר, אנו מנסים להסביר כיצד שינוי אקסוגני במשתנים המסבירים האינדוגניים שלנו משפיע על המשתנה התלוי. כיוון שבחרנו במשתנים שלנו במונחים של לוגריתמים, משמעות של כל אומד הנה: בכמה אחוזים יגדל בממוצע המשתנה המוסבר אם היינו מגדילים את המשתנה המסביר באופן אקסוגני ב- 1%.

הצגת תוצאות ומסקנות עיקריות של מודל 2SLS:

לאחר הרצת התכנית קיבלנו את התוצאות הבאות (ראה נספח – טבלה 2, עמודה 1):

קיבלנו כי הקשר בין לוגריתם הצמיחה בשנה שעברה (משתנה מסביר) ללוגריתם הצמיחה השנה (משתנה תלוי) הנו שלילי. פירושו כי כל עלייה אקסוגנית של 1% בצמיחה של שנה שעברה מקטינה בממוצע ב- 0.8579% את הצמיחה השנה. למרות זאת סטיית התקן הנה גבוהה מאוד: 1.7598, מה שמטיל ספק גדול באמינות האומד הזה.

אותה מסקנה בדיוק לגבי לוגריתם התמ"ג של שנה שעברה.

ההשפעה החיובית הגדולה ביותר הנה ממספר הפלאפונים שישנם ל-100 תושבים. עלייה אקסוגנית ב-1% במספר הפלאפונים השנה תגרום לעלייה ב-0.3598% בצמיחה. בנוסף סטיית התקן שהתקבלה נמוכה הרבה יותר, לעומת שני המשתנים המסבירים הראשונים, אך היא עדיין לא קטנה מספיק ושווה ל-0.1128. תיאורטית, ניתן להסיק מכך שאם היינו מכפילים השנה את כמות הפלאפונים בבעלותנו פי 2, ואת שאר המשתנים המסבירים היינו מחזיקים קבועים, אז היינו מקבלים עלייה בצמיחה ב-35.98%!

ההשפעה החיובית השנייה בגודלה הנה הרווח מענף הטלקומוניקציה. כלומר אם רווח מענף הטלקומוניקציה יגדל בצורה אקסוגנית ב-1% הוא יגרום לצמיחה של 0.1347%, שגם זאת עלייה נאה מאוד בצמיחה. סטיית התקן של האומד הנה 0.1249.

תוצאות ומסקנות של מודל First Difference GMM:

ראה נספח, טבלה 2, עמודה 2.

המסקנה העיקרית שברצוני להתייחס אליה היא לגבי סטיות התקן של האומדים. כיוון שזה הוא שלב ביניים לפני המודל העיקרי בעל האומדים המדויקים ביותר, ניתן לדלג על משמעות האומדים. סטיות התקן של רוב האומדים יותר קטנות לעומת המודל 2SLS, מה שמלמד אותנו על דיוק טוב יותר.

תוצאות ומסקנות של מודל System GMM:

ראה נספח, טבלה 2, עמודה 3.

שלב זה נותן תוצאות לא דומות לשני המודלים הקודמים. אנו מנצלים את המידע שבידינו בצורה יותר רחבה ולכן מקבלים תוצאות מדויקות יותר. סטיות תקן אשר מתקבלות במודל זה קטנות הרבה יותר בהשוואה למודלים הקודמים.

הפעם, כפי שציפינו, הקשר בין לוגריתם הצמיחה בשנה שעברה (משתנה מסביר אינדוגני) ללוגריתם הצמיחה השנה (משתנה מוסבר) הנו חיובי. פירושו כי כל עלייה אקסוגנית של 1% בצמיחה של שנה שעברה מגדילה בממוצע ב- 0.7221% את הצמיחה השנה. וסטיית התקן הקטנה, 0.1309, מעידה על הדיוק הרב של אומד זה, לעומת המודלים הקודמים, בהם קיבלנו סטיית תקן גדולה יחסית עבור אומד זה.

הקשר בין לוגריתם התמ"ג בשנה שעברה לבין לוגריתם הצמיחה השנה שלילי, אך זניח. כל עלייה אקסוגנית של 1% בתמ"ג תביא לירידה של 0.0298% בצמיחה השנה. וזה הוא שיעור זניח. כלומר, על פי מודל זה, לתמ"ג בשנה שעברה כמעט ואין קשר לצמיחה השנה. האומד שקיבלנו מדויק ובעל סטיית תקן קטנה של 0.0364.

תוצאות ומסקנות של מודל 2 Step System GMM:

בשלב זה מתקבלים האומדים המדויקים ביותר. ניתן לשפוט על כך לפי השונויות הקטנות ביותר. תוצאות אלה לא שונות בהרבה מהשלב הראשון של System GMM, כיוון שהשינוי היחיד כאן היא מטריצת המשקולות שנבנתה על פי התוצאות של השלב הראשון.

8. סימולציה

סימולציה עבור מודל 2SLS:

כדי להבין טוב יותר את האומדים שקיבלנו אנו נריץ את המודל עם כמות קטנה יותר של אותם משתנים אקסוגניים:

B2s1s =	SD =
-0.4065	1.4304
-0.7571	0.4673
0.0146	0.0101
0.1381	0.1224
0.1132	0.1027
0.0555	0.0610
-0.0020	0.0218
0.0063	0.0107

זוהי תוצאה שהתקבלה לאחר הרצת המודל ללא המשתנה האקסוגני בעל ההשפעה הגדולה ביותר: מספר הפלאפונים פר 100 תושבים. ניתן לראות תוצאה מעניינת כי ההשפעה השלילית של משתנה המסביר המסביר הראשון ירדה בצורה ניכרת. ולא ניכר שינוי גדול במשתנים אקסוגניים אחרים. בנוסף לא ניכר שינוי משמעותי בסטיות התקן שהתקבלו, שעדיין נשארו גדולות.

סימולציה עבור מודל 2 Step System GMM:

ניתן להריץ מודל זה מספר פעמים לפי בקשתו של משתמש. כאשר בכל הרצה נבחרות 30 מדינות בצורה אקראית עם החזרה ובסופו של דבר אנו מחשבים את הממוצע של כל האומדים שהתקבלו בהרצות השונות ואת סטיות התקן של אומדים אלה. וכך אנו מקבלים תמונה כללית מאוד על האומדים שבידינו. בנוסף ברצוננו לבדוק האם לאחר הרבה הרצות ישנה התייצבות של האומדים סביב ערך מסוים.

	2 runs average	10 runs average	100 runs average	1000 runs average	5000 runs average	10000 runs average
ln_gr_t	0.6581	0.685	0.6851	0.6847	0.6844	0.6936
	[0.1386]	[0.1601]	[0.189]	[0.2018]	[0.2005]	[0.1964]
ln_gdp_t	-0.0003	-0.0532	-0.0462	-0.0415	-0.0445	-0.0438
	[0.0181]	[0.0534]	[0.0704]	[0.0792]	[0.0792]	[0.0791]
ln(Cellular mobile subscriptions per 100 capita)	0.3838	0.3127	0.1777	0.1555	0.1567	0.1564
	[0.0770]	[0.2572]	[0.2650]	[0.2698]	[0.2645]	[0.2651]
ln(Public telecommunication investment per capita)	-0.0102	0.0173	0.005	0.0077	0.0075	0.008
	[0.0456]	[0.0250]	[0.0241]	[0.0258]	[0.0264]	[0.0264]
ln(Total communication access paths per 100 capita)	-0.3254	-0.2252	-0.0851	-0.0686	-0.0676	-0.0686
	[0.0870]	[0.2622]	[0.2949]	[0.3089]	[0.2991]	[0.3009]
ln(Telecommunication revenue in USD Millions)	0.0373	0.0188	0.0447	0.0416	0.0435	0.0429
	[0.0679]	[0.0381]	[0.0561]	[0.0624]	[0.0621]	[0.0619]
ln(Internet host by domain)	-0.0468	-0.0481	-0.0562	-0.0571	-0.0568	-0.0571
	[0.0084]	[0.0310]	[0.0258]	[0.0267]	[0.0254]	[0.0252]
ln(Television households in Thousands)	0.0072	0.0755	0.0551	0.0547	0.0554	0.056
	[0.0931]	[0.0701]	[0.0595]	[0.0677]	[0.0652]	[0.0658]
ln(Total broadband subscriptions per 100 capita)	0.0281	0.029	0.0215	0.0212	0.0207	0.021
	[0.0138]	[0.0106]	[0.0195]	[0.0223]	[0.0222]	[0.0222]

האומדים שקיבלנו אינם נבדלים בהרבה מהתוצאות של המודל הנתון לאחר הרצה אחת בלבד. ראה נספח, טבלה 2 עמודה 4. כלומר ניתן להתייחס לתוצאות אלה כאל תוצאות מדויקות. בשורה הראשונה מופיעים האומדים עצמם ובשורה מתחת בסוגריים סטיות תקן מדגמיות של האומד. ניתן לראות כאן התייצבות של האומדים לאחר הרבה הרצות סביב ערך מסוים שכמעט ומפסיק להשתנות לאחר 100 הרצות.

סימולציה נוספת שהחלטנו לבצע הנה: להריץ את המודל עם צירופים מסויימים של משתנים מסבירים אקסוגניים ולראות איזה צירוף מביא לשונות מינימאלית של האומדים המסבירים האינדוגניים.

Combination of Exogenic Variables ->	phones & invest	phones & access	invest & access
ln_gr_t	0.4959	0.5375	0.5912
	[0.0090]	[0.0068]	[0.0119]
ln_gdp_t	0.0043	0.001	0.0098
	[0.0005]	[0.0005]	[0.0003]

השונות הקטנה ביותר התקבלה כאשר הרצנו את המודל עם הצירוף של לוגריתם של מספר הפלאפונים פר 100 תושבים ולוגריתם של מספר נקודות גישה פר 100 תושבים. מסקנה מעניינת היא כי כמות גדולה של משתנים מסבירים מגדילה את השונות של כל האומדים.

9. השוואה בין התוצאות של שני המודלים: 2SLS ו-GMM.

ישנם פערים בין מודל 2SLS לבין GMM והתוצאות אינן דומות, ההבדל העיקרי נובע ממבנה מטריצת משתני העזר. כזכור, שיטת GMM מעניקה משתני עזר נפרדים לכל תקופה. על מנת שמודל ה-GMM יתכנס בחזרה ל-2SLS, צריך לדרוש שמטריצת משתני העזר תהיה בפורמט הרגיל של מטריצת IV (Instrument Variable) ללא הענקת משתנה עזר ייחודי לכל תקופה ואז התוצאות הינן דומות, במקרה של GMM קוראים לפעולה collapse. אולם, אז מפסידים את היעילות – כי משתמשים באותה האינפורמציה, אך בצורה שלא מאפשרת להעניק לכל תקופה משתנה עזר אישי.

10. סיכום ומסקנות נוספות

צמיחה בתמ"ג – זהו המדד העיקרי המאפשר לאמוד את מצב הכלכלה במדינה מסויימת. ניסינו לבדוק האם וכיצד התפתחות בתקשורת משפיעה על הצמיחה הכלכלית. הגענו לכמה מסקנות מעניינות לגבי שאלה זו. ברצוננו להציג את המסקנות על פי המודל 2 Step System GMM כי זהו המודל הכי מדוייק ומנצל את האינפורמציה בצורה מיטבית ויעילה ביותר. המסקנות מתייחסות בעיקר למדינות המפותחות, עקב המחסור בנתונים מהמדינות המתפתחות.

כפי שציפינו, הצמיחה בפיגור היא בעלת השפעה מובהקת והכי ניכרת. אולם, התמ"ג של שנה שעברה בעל השפעה זניחה, אך שלילית על הצמיחה השנה. המשתנה האקסוגני בעל ההשפעה הכי גדולה על הצמיחה הנו מספר הטלפונים הניידים פר 100 תושבים, שכן אם היינו מגדילים את מספר הפלאפונים ברשותנו אפילו ב-10% בצורה אקסוגנית, היינו מצפים להגדלה בצמיחה של 1.6%.

בנוסף, קיבלנו הוכחה לכך שבמדינות המפותחות התשתית של טלקומוניקציה כבר מאוד מבוססת ולכן הגדלת ההשקעה בתשתיות הטלקומוניקציה לא תגרום לעלייה משמעותית בצמיחה. אפילו אם ממשלה במדינה מפותחת תחליט להכפיל את ההשקעה בתשתיות התקשורת – הדבר צפוי להביא לעלייה של 0.74% בלבד בצמיחה השנה. אם היו בידינו הנתונים מהמדינות המתפתחות – היינו מצפים כי ההשקעה בתשתיות התקשורת תהיה בעלת השפעה מאוד ניכרת וחיובית על הצמיחה באותה מדינה.

ברצוני להציג כיצד שינוי אקסוגני במשתנים מסבירים היה משפיע על הצמיחה בבריטניה, שכלכלתה מזכירה את כלכלת ישראל. על פי הנתונים שבידינו הצמיחה בבריטניה הנה בערך 2.8% בשנה בממוצע. אם באחת השנים צמיחתה תגדל בצורה אקסוגנית ב-10% אז בשנה הבאה זה צפוי לגרום להגדלה בצמיחתה ב-7.35% וצמיחתה תהיה: $1.0735 \times 2.8\% = 3.0058\%$, שזה כמובן מאוד משמעותי. ואם לדוגמא, באחת השנים מספר הפלאפונים פר 100 יגדל בצורה אקסוגנית ב-10% הדבר צפוי לגרום לצמיחה של: $1.01602 \times 2.8\% = 2.845\%$, שכמובן זוהי השפעה קטנה יותר אך עדיין ישנו שיפור בכלכלת המדינה.

הגענו למסקנה כי רוב המשתנים שאמדנו הם בעלי השפעה חיובית על הצמיחה הכלכלית. אולם, חלקם בעלי השפעה זניחה. הסיבה לכך טמונה בעובדה כי המדינות שאיתם עבדנו הן מדינות מפותחות ושינויים בתשתית התקשורת, שהיא כבר מאוד מבוססת, בעלי השפעה קטנה יחסית על הצמיחה הכלכלית. בתור המלצה לעתיד, הייתי מעוניין לבדוק את הנושא הזה עבור מדינות מתפתחות ומדינות מ"העולם השלישי", ולבדוק מהי הדרך היעילה ביותר שתביא לשיפור בכלכלתם ובחיים של האנשים במדינות אלה.

11. נספח

טבלה 1: קובץ Excel המייצג נתונים עבור מודל 2SLS.

dif_In(Total broadband subscriptions per 100 capita)	dif_In(Televisi	dif_In(Interne	dif_In(Teleco	dif_In(Total c	dif_In(Public	dif_In(Cellula	IV.ln_gdp_t2	IV.gr_t2	ln_gdp_t_f	gr_t_f	gr_f	year
0.641324465	0.018440435	0.11376746	0.178569356	0.070773205	0.440745367	0.060585	13.2551	1.056393	0.05253	-0.002457	0.013081	2003
0.783718747	0.017930449	0.104833028	0.204726743	0.097158268	-0.013956922	0.065922	13.30763	1.053936	0.06487	0.013081	-0.010321	2004
0.572773541	0.017467323	0.52637417	0.111896546	0.083911303	0.051240673	0.046309	13.3725	1.067017	0.05515	-0.010321	0.006397	2005
0.299110888	-0.017916303	0.225410965	0.01502168	0.051722503	0.04563375	0.023318	13.42765	1.056696	0.06118	0.006397	0.010557	2006
0.21830283	0.018974197	0.345155557	0.132147035	0.049983858	0.230612635	0.025445	13.48883	1.063093	0.07106	0.010557	-0.001132	2007
0.080650418	0.019041539	0.227436249	0.036824371	0.007645371	0.026575594	-0.00341	13.55989	1.07365	0.07001	-0.001132	-0.050765	2008
-0.039905446	0.018975436	0.057234912	-0.045238046	0.019756779	0.169492977	-0.026515	13.6299	1.072518	0.02152	-0.050765	0.022104	2009
0.310265104	0.003921887	0.243101824	0.227437341	0.040970047	-0.792918019	0.042547	12.36061	1.007841	0.05332	0.046927	-0.022101	2003
0.331821488	0.040455851	0.15879266	0.11966524	0.087733883	0.051266547	0.051468	12.41393	1.054768	0.03215	-0.022101	0.023036	2004
0.302477594	0.008378266	0.484197086	0.029229665	0.044532033	0.771663742	0.043125	12.44608	1.032667	0.05421	0.023036	-0.024937	2005
0.152718806	0.013963068	0.205207784	-0.022826945	0.068433656	-0.017586893	0.033094	12.50029	1.055703	0.0303	-0.024937	0.062555	2006
0.13991988	0.008112783	0.174516366	0.035522664	0.045143686	0.245637683	0.02867	12.53059	1.030766	0.08922	0.062555	-0.048526	2007
0.097366373	-0.009638617	0.105386005	0.007670669	0.054477998	-0.478093309	0.053214	12.61981	1.093321	0.04382	-0.048526	0.004654	2008
0.056554054	0.019814168	0.11460478	-0.111433157	0.054066802	-0.069234229	0.037389	12.66363	1.044795	0.04826	0.004654	-0.0702486	2009
0.299375184	0.022177508	0.454070907	0.237363275	0.052111457	-0.022589306	0.029078	12.58881	1.034446	0.05701	0.024224	-0.046569	2003
0.284165095	0.005887062	0.323195074	0.143901039	0.054784164	0.042555886	0.027735	12.64582	1.05867	0.01203	-0.046569	0.02192	2004
0.156063381	0.006952519	0.324710833	0.047672539	0.04382693	0.064956104	0.027421	12.65785	1.012101	0.03345	0.02192	0.003846	2005
0.206936119	0.007592352	0.23531173	0.042382625	0.032800137	-0.017853273	0.006273	12.6913	1.034021	0.03717	0.003846	0.032639	2006
0.134840815	0.011621414	0.213092533	0.101492144	0.063041653	0.189474711	0.06092	12.72847	1.037867	0.06813	0.032639	-0.021647	2007
0.078887285	0.020628534	0.138384032	0.071420745	0.054797156	0.155461243	0.025108	12.7966	1.070506	0.04771	-0.021647	-0.002301	2008
0.048580195	0.484486089	0.188178361	-0.056689625	0.034588928	0.132815981	0.015054	12.84431	1.048859	0.0455	-0.002301	-0.0472749	2009

טבלה 2: תוצאות עבור 3 המודלים.

	2SLS	FD_GMM	SYS_GMM	2 step SYS_GMM
ln_gr_t	-0.8579 [1.7598]	-0.0712 [0.2268]	0.7221 [0.1309]	0.7355 [0.0494]
ln_gdp_t	-0.9059 [0.5752]	-1.3029 [0.266]	-0.0298 [0.0364]	-0.0342 [0.0158]
ln(Cellular mobile subscriptions per 100 capita)	0.3598 [0.1128]	0.7447 [0.3589]	0.2196 [0.1609]	0.1602 [0.1153]
ln(Public telecommunication investment per capita)	0.0156 [0.0121]	0.0261 [0.0298]	0.0046 [0.0163]	0.0074 [0.0092]
ln(Total communication access paths per 100 capita)	0.0023 [0.1393]	-0.8305 [0.6951]	-0.135 [0.18]	-0.0698 [0.1351]
ln(Telecommunication revenue in USD Millions)	0.1347 [0.1249]	0.2779 [0.0965]	0.038 [0.034]	0.0401 [0.0153]
ln(Internet host by domain)	0.0726 [0.0748]	0.3206 [0.1176]	-0.076 [0.0169]	-0.0803 [0.0122]
ln(Television households in Thousands)	0.0005 [0.0263]	-0.1882 [0.3533]	0.0714 [0.0432]	0.0779 [0.0274]
ln(Total broadband subscriptions per 100 capita)	-0.0017 [0.0136]	-0.0231 [0.0458]	0.036 [0.0144]	0.0341 [0.006]

* עבור כל משתנה ניתן לראות את האומד למעלה ואת סטיית התקן שלו למטה בסוגריים.


```

1 function TSLS_model
2 % the data has to be full without missings
3 trim = xlsread('DataSet_to_use.xls','final_to_use');
4 % The information is built of differences between
5 % 2 consecutive periods of time
6 Y = trim(:,3); % gr_f=gdp / gdp_t: explained variable
7 Exog = trim(:,8:14); % (Exogenous Variables)
8 % differences [ ln(number of phones per 100) ]
9 Endog = trim(:,4,5); % ln_gr_t_f -> ln(growth of gdp lag of 1 year)
10 % ln_gdp_t_f -> ln of GDP lagged for 1 year
11 Inst = trim(:,6,7); % Instrument Variables (Lag of 2 years)
12 %X -> matrix of explanatory variables we have chosen
13 %X = [ones(size(Y,1),1) Exog Endog];
14 %Z -> endogenous variables were changed by exogenic
15 %Z = [ones(size(Y,1),1) Exog Inst];
16 X = [Endog Exog];
17 Z = [Inst Exog];
18 % 2SLS Estimator
19 B2sls = inv( (X'*Z)*inv(Z'*Z)*(Z'*X) )*(X'*Z)*inv(Z'*Z)*(Z'*Y)
20 SSE = (Y-X*B2sls)*(Y-X*B2sls); % goodness of fit
21 sigma = (SSE / (size(Y,1)))^0.5; % sqrt(variance)
22 %standard deviation of 2SLS estimators
23 SD = sigma * diag(inv( (X'*Z)*inv(Z'*Z)*(Z'*X) ))^0.5
24 %standard deviation of the estimators
25 % as the SD grows higher- the 2sls estimator is getting worse
26

```

רשימת משתנים "מודל 2SLS".

$$1. \text{ הגדרה לצמיחה בתמ"ג: } gr_{it} = \frac{gdp_{it}}{gdp_{it-1}}$$

2. משתנה מוסבר – הפרש בין לוגריתמים של צמיחה בתמ"ג בין כל שתי תקופות עוקבות:

$$\Delta \ln gr_{it} = \ln gr_{it} - \ln gr_{it-1} = \ln \frac{gr_{it}}{gr_{it-1}} = gr_t_f$$

3. משתנה מסביר אינדוגני – הפרש בין לוגריתמים של צמיחה בתמ"ג בין תקופות עוקבות שנה

$$\Delta \ln gr_{it-1} = \ln gr_{it-1} - \ln gr_{it-2} = \ln \frac{gr_{it-1}}{gr_{it-2}} = gr_t_f$$

4. משתנה מסביר אינדוגני – הפרש בין לוגריתמים של תמ"ג בין תקופות עוקבות שנה לפני:

$$\Delta \ln gdp_{it-1} = \ln gdp_{it-1} - \ln gdp_{it-2} = \ln \frac{gdp_{it-1}}{gdp_{it-2}} = gdp_t_f$$

5. משתנה מסביר אקסוגני ראשון – הפרש בין לוגריתמים של מס' פלאפונים פר 100 תושבים:

$$\Delta \ln phones_{it} = \ln phones_{it} - \ln phones_{it-1} = \ln \frac{phones_{it}}{phones_{it-1}}$$

ועוד 6 משתנים מסבירים אקסוגניים – כאשר כל אחד מהמשתנים בצורה של משתנה אקסוגני ראשון. משתנים אלו הוצגו בפרק "נתונים".

6. משתנה עזר ראשון – לוגריתם הצמיחה בפיגור של שתי תקופות: $\ln gr_{it-2} = gr_t2$

7. משתנה עזר שני – לוגריתם התמ"ג בפיגור של שתי תקופות: $\ln gdp_{it-2} = gdp_t2$

12. ביבליאוגרפיה

- 1) Telecommunications and Economic Development: Empirical Evidence from Southern Africa. By James Alleman, Carl Hunt, Donald Michaels, Milton Mueller, Paul Rappoport and Lester Taylor.
- 2) Difference GMM Model. By Arellano and Bond, 1991.
- 3) Another look at the instrumental variable estimation of error-components models. By Arellano and Bover, 1995.
- 4) The Role of Telecommunication Infrastructure in the Regional Economic Growth of Africa. By Enowbi Batuo Michael (November 2008).
- 5) Telecommunications Infrastructure and Economic Development: A Simultaneous Approach. By Lars-Hendrik Roller and Leonard Waverman.
- 6) Econometric Analysis of Cross Section and Panel Data. By Jeffrey M. Wooldridge.
- 7) OECD Communications Outlook 2011.