

ДЕКОМПОЗИЦИЯ НА РАЗЛИКАТА В СРЕДНАТА ПРОДЪЛЖИТЕЛНОСТ НА ПРЕДСТОЯЩИЯ ЖИВОТ ЧРЕЗ АНАЛИЗ НА ОБГРЪЩАНЕТО И ВРЪЗКА С МЕТОДА НА DAS GUPTA (1978)

*Петя Брайнова**

Въведение и цел

Декомпозиционните техники имат изключително важно място в областта на демографската статистика, имайки предвид богатия теоретичен инструментариум, който предлагат, както и практическата им полезност в чисто приложен аспект. Методите за декомпозиране се развиват от втората половина на миналия век до днес и за този период множество изследователи от целия свят са дали своя принос. Без да претендираме за изчерпателност, можем да изброим имената на някои от тези изследователи, общото между които е, че са работили и постигнали значими резултати при декомпозирането на разликата между две средни величини и по специално, на разликата в СППЖ между две съвкупности: Kitagawa, Retherford, Cho, Das Gupta, Arriaga, Pollard, Pressat, Keyfitz, Vaupel, Romo, Русев, Andreev, Shkolnikov, Begun, Preston, Beltran-Sanchez и още много други.

Целта на настоящото изследване е да представи анализа на обгръщането - DEA - като декомпозиционна техника и същевременно с това да покаже неговата връзка с метода на Das Gupta (1978) за декомпозиция на разликата в СППЖ. Класическото предназначение на DEA е да сравнява производствената ефективност на хомогенни единици (напр. предприятия, банки, университети, болници и др.). Пример за такова сравнение е икономическият анализ на ефективността на болничните услуги в България за периода 2010 - 2012 г. на авторския колектив Салчев, Атанасов, Димитрова-Савова и Грива.

До настоящия момент DEA е иновативно приложен като декомпозиционна техника от Брайнова (2012) при декомпозирането на разликата в СППЖ между две съвкупности по няколко фактора. В статията си от 2012 г. Брайнова сравнява съвкупностите на мъжете и жените в България през 2008 г., като единиците, подлежащи на анализ, са 12

* Докторант в катедра „Статистика и иконометрия”, УНСС, София;
e-mail: petia_brainova@abv.bg.

възрастово-полови групи - 6 възрастови групи за мъжете и 6 - за жените. Разгледаните декомпозиционни фактори за всяка възрастово-полова група поотделно са делът на хората с индекс на телесната маса (ИТМ) в норма и делът на нетютюнопушещите в съответната група, като източник на данните е официалният източник на статистическа информация за Европа - Евростат.

В настоящото изследване за сравнителни съвкупности са взети населението на Дания и населението на Швеция за 2008 г., като са разгледани четири възрастови групи - 27 - 34, 35 - 44, 45 - 54 и 55 - 64. Обектът на декомпозиране (Брайнова, 2012) е съответната разлика в СППЖ между страните за всяка възрастова група, разгледаните декомпозиционни фактори са класическите демографски характеристики - пол и образование¹, а информационният източник отново е Евростат. В първата част на изложението са представени последователно методът на Das Gupta (1978), анализът на обгръщането DEA, използваният декомпозиционен модел заедно с неговите ограничения и получените ефекти вследствие на DEA декомпозицията. Във втората част на изложението са описани четирите основни случая на DEA декомпозицията² и техните интерпретации, връзката на DEA декомпозицията с метода на Das Gupta, предимствата на анализа на обгръщането и накрая получените изводи и заключения.

1. Метод на Das Gupta (1978)

Методът на Das Gupta (1978) е класически декомпозиционен метод в демографската статистика. Той декомпозира разликата между средни величини на две съвкупности (P_1 и P_2). В настоящото изследване методът е приложен върху разликата в СППЖ между Дания и Швеция по възрастови групи. Формулата на Das Gupta е в сила за краен брой

¹ Причината първата възрастова група да бъде 27 - 34 (вместо 25 - 34 по аналогия на следващите групи) е, че броят на висшистите се стабилизира след 26-годишна възраст. Страните са избрани по признак в Евростат, за тях да има данни по избраните фактори и същевременно наблюдаваните разлики в СППЖ за дела на жените и дела на висшистите да бъдат възможно най-големи за постигане на по-добро диференциране след анализа на обгръщането.

² Четирите основни случая са конструирани с помощта на хипотетични съвкупности, аналогични на използваните хипотетични съвкупности от Das Gupta (1978) при извеждане на неговата декомпозиционна формула.



фактори и в случая ще разгледаме два декомпозиционни фактора - пол и образование, отбелязани съответно с I и J . Формулата на Das Gupta е подобрение на формулите на Kitagawa (1955) и Retherford и Cho (1973), като нейните основни предимства са, че в нея не присъства ефектът на взаимодействието между факторите, нито съществува зависимост от реда на пренебрегването на тези фактори при изчисление на ефектите. За да избегне ефекта на взаимодействието, Das Gupta използва хипотетични съвкупности R_{ij} и r_{ij} , които удовлетворяват следната зависимост:

$$R_{ij} = r_{ij} = \frac{R_{ij} + r_{ij}}{2} = \frac{T_{ij} + t_{ij}}{2} \quad (1)$$

където T_{ij} и t_{ij} са съответно СППЖ, които съответстват на единиците от съвкупности P_1 и P_2 (в случая едноименни възрастови групи в Дания и Швеция³), принадлежащи на i -тата категория на фактор I (пол) и същевременно на j -тата категория на фактор J (образование).

Das Gupta дефинира четири различни ефекта - $I(I, J)$, $J(I, J)$, $I(J, I)$, $J(J, I)$, като означението (I, J) след името на съответния ефект означава, че на първа стъпка се пренебрегва факторът J (образование) и съответно се отделя I ефектът от пола, а на втора стъпка се пренебрегва I (полът) и по този начин се пресмята J ефектът от образованието. За да избегне зависимостта на декомпозиционната формула от реда на пренебрегване на факторите, Das Gupta прави още едно полагане:

$$\begin{aligned} I \text{ effect} &\doteq \frac{1}{2}(I(I, J) + I(J, I)) \\ J \text{ effect} &\doteq \frac{1}{2}(J(I, J) + J(J, I)) \end{aligned} \quad (2)$$

откъдето следва, че

$$r_{..} - R_{..} = I \text{ effect} + J \text{ effect} = \sum_i \sum_j \left(\frac{r_{ij} + R_{ij}}{2} \right) \left(\frac{n_{ij}}{n_{..}} - \frac{N_{ij}}{N_{..}} \right), \quad (3)$$

където N_{ij} и n_{ij} са съответно броят на единиците от съвкупности P_1 (съответната възрастова група на Дания) и P_2 (съответната възрастова група на Швеция), принадлежащи на i -тата категория на фактор I (пол) и

³ За по-голяма прегледност на формулите не е въведен допълнителен индекс, който да се мени по възрастовите групи.

същевременно на j -тата категория на фактор J (образование), а $N_{..}$ и $n_{..}$ са общите обеми съответно на съвкупностите P_1 и P_2 . По този начин Das Gupta елиминира ефекта на взаимодействието, като го разпределя поравно между ефектите от факторите пол и образование и **формулата на Das Gupta придобива следния вид:**

$$t_{..} - T_{..} = I \text{ effect} + J \text{ effect} + \sum_i \sum_j \left(\frac{\frac{n_{ij}}{n_{..}} + \frac{N_{ij}}{N_{..}}}{2} \right) (t_{ij} - T_{ij}). \quad (4)$$

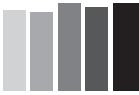
Формула (4) не съдържа ефект на взаимодействието между факторите I и J и същевременно не зависи от реда на тяхното изчисление, където последното събираемо се нарича **остатъчен ефект**⁴. Тя е приложима и при повече от два фактора, като трябва да се има предвид, че при n на брой фактори, всеки от ефектите представлява средна аритметична на $n!$ събираеми, точно колкото е броят на възможните подредби на n елемента.

За практическото приложение на метода на Das Gupta (1978) са изготвени два вида IJ матрици с размерност 3×2 (3 образователни степени (в т.ч. висше образование) и 2 пола):

- Първият вид матрици съдържат абсолютния брой на хората от съответния пол и същевременно притежаващи съответната степен на образование, като те са изготвени за всяка възрастова група във всяка страна поотделно. Това прави четири възрастови групи по две страни, което е равно на осем IJ матрици от първия вид. На базата на тези матрици се изчисляват съответните IJ разпределения, участващи в декомпозиционната формула.

- Вторият вид матрици съдържат средната продължителност на живота на хората на възраст, равна на долната граница на всяка изследвана възрастова група, които са от съответния пол и същевременно притежават съответната степен на образование. Това прави четири възрастови групи по две страни, което е равно на осем IJ матрици от втория вид. На базата на тези матрици се изчисляват съответните IJ специфични средни величини, участващи в декомпозираната разлика.

⁴ Същият остатъчен ефект присъства и във формулата на Kitagawa (1955).



След приложението на метода на **Das Gupta (1978)** върху данните от съответните 16 *IJ* матрици за всяка от четирите възрастови групи поотделно, чрез поетапно пресмятане на всички събираеми, съставлящи декомпозиционни ефекти, се получават резултатите, посочени в табл. 1.

1. Декомпозиционни резултати от метода на **Das Gupta (1978)** по възрастови групи

Възрастови групи	Ефект от пола в години	Ефект от образованието в години	Остатъчен ефект в години	Общ ефект в години (разлика в СППЖ в полза на Швеция)
27 - 34	0.04	-0.29	-2.05	-2.30
35 - 44	0.02	-0.14	-2.05	-2.17
45 - 54	0.01	-0.09	-1.93	-2.02
55 - 64	0.01	-0.03	-1.60	-1.62

За всяка възрастова група стойностите на ефектите от фактора пол са положителни и показват каква би била⁵ разликата в СППЖ между Дания и Швеция, ако дяловете на жените са равни на реалните си стойности, а дяловете на висшистите за двете страни са равни помежду си и също така са равни на средноаритметичната на техните реални дялове. С други думи, ако пренебрегнем ефекта на образованието и вземем предвид ефекта на фактора пол, разликата ще се обърне в полза на Дания и ще приеме стойности близки до 0 (вж. табл. 1, колона 2). По аналогичен начин се интерпретират и ефектите от фактора образование, като при него разликата в СППЖ остава в полза на Швеция, но абсолютната ѝ стойност намалява за всяка от възрастовите групи, като при най-възрастните клони към 0.

Анализ на обгръщането (Data Envelopment Analysis - DEA)

След представянето на класическия декомпозиционен метод на **Das Gupta (1978)** следва описание на анализа на обгръщането (DEA), който също ще бъде приложен като декомпозиционна техника на разликата в СППЖ между две съвкупности.

Анализът на обгръщането (DEA) е непараметричен, базиран на техниката на линейното оптимизиране и измерва ефективността на мно-

⁵ За коректната интерпретация на резултатите е важно да се отбележи, първо, че ефектите от метода на **Das Gupta** не доказват наличието на каквито и да било причинно-следствени връзки и второ, че не е задължително при увеличаване на броя на факторите остатъчният ефект да намалява.

жество еднородни единици⁶. Ефективността е изразена като съотношение на продукцията (outputs) и използваните за нейното производство ресурси (inputs). Най-ефективните единици получават стойност на ефективността равна на 1 (100% ефективност) и стават еталон за останалите единици. Най-общо, функцията на ефективност за всяка единица има вида:

$$Efficiency_{DMU} = \frac{Outputs_{DMU}}{Inputs_{DMU}} = \frac{\sum w_i * output_i}{\sum w_j * input_j}, \quad (5)$$

където i се мени от 1 до броя на продуктите, а j - от 1 до броя на ресурсите. По-точно, **ефективността** е частно между претеглената сума от всички произведени продукти и претеглената сума от всички използвани за тяхното производство ресурси. Разгледаните в анализа ресурси и продукти ще бъдат наричани с общото име „фактори на DEA”. Теглата им (w_i, w_j) се определят от самия анализ при оптимизиране на ефективността на всяка единица, като стойностите на теглата са възможно най-благоприятни за нея. Това означава, че всяка от разглежданите единици получава най-големи тегла на продуктите, които произвежда най-ефективно, или на ресурсите, които използва най-ефективно, в зависимост от ориентацията на модела.

Например в статията на Брайнова (2012) продукцията е сумата от вече изживения среден брой години на всяка възрастова група и средния брой години предстоящ живот, което за краткост ще наричаме „общ брой години живот”. Ресурсите характеризират здравословния начин на живот на всяка група, а именно процентът на непущачите и процентът на хората с индекс на телесна маса (ИТМ) в норма в групата. Хипотезата е, че при увеличаване на дяловете на хора с ИТМ в норма и увеличаване на дяловете на непущачи средната продължителност на предстоящия живот също нараства, т.е. здравословният начин на живот произвежда дълголетие. По този начин DEA влиза в ролята на декомпозиционна техника на разликата в СППЖ между мъжете и жените за всяка от разглежданите възрастови групи.

Имайки предвид, че анализът на обгръщането функционира според алгоритъма на известния „симплекс метод” за решаване на оптимиза-

⁶ На езика на DEA тези единици се наричат единици за вземане на решения, т.нар. Decision making units (DMUs).



ционни задачи, целта на всяка единица в анализа е да произведе възможно най-много продукция с използване на определено количество ресурси (**модел, ориентиран към изхода**) или да използва възможно най-малко ресурси за производството на определено количество продукция (**модел, ориентиран към входа**). Ако една единица има ефективност, равна на 1 (100%), тя се намира на **границата на производствените възможности** (*DEA-frontier*). Тази граница се определя от всички единици, които имат оптимална ефективност. Важно е да се отбележи, че оптималната ефективност се изчислява в контекста на всички разглеждани единици в анализа, без наличие на предварително дефиниран еталон.

Анализът на обгръщането предлага множество модели в зависимост от естеството на задачата, която искаме да решим. Те се определят по своята ориентация (към входа/изхода) и по броя на оптималните нива на ефективност. Авторите Cooper, Seiford и Tone (2007) описват подробно различните DEA модели в едноименната си книга, както и множество примери за приложението и спецификата на всеки от тях. Cook и Zhu (2008) публикуват практическо ръководство с геометричните интерпретации на основните DEA модели.

2. DEA модел за декомпозиция на разликата в СППЖ

За декомпозиране на разликата в СППЖ между възрастовите групи на Дания и Швеция по факторите пол и образование е използван DEA модел, **ориентиран към входа с ограничени множители, който допуска едно ниво на ефективност за оптимално**⁷. Този модел минимизира претеглената сума на ресурсите (пол и образование) при фиксирана претеглена сума на продукта (общ брой години живот), като неизвестните величини са оптималните тегла на ресурсите и продукта. Моделът решава толкова на брой линейни оптимизационни задачи, колкото са на брой сравнителните единици в анализа (в случая 8 задачи по 4 за всяка страна), т.е. решава се отделна задача за всяка възрастова група. В ориентирания към входа модел, ефективността на всяка изследвана единица може да приема стойности от 0 до 1 (0 до 100%), което показва какъв

⁷ Използваният модел (**Input-oriented constant return to scale (CRS) model with restricted multipliers**) определя само едно ниво като 100% ефективно и съпоставя всички единици с него. Границата на производствените възможности е правата линия, определена от всички единици, имащи оптимално съотношение между продукти и ресурси.

процент от използваните ресурси биха оправдали получената продукция, така че производственият процес да бъде ефективен. **В конкретния случай** ефективността на всяка възрастова група показва, какъв трябва да бъде делът на жените и делът на висшистите в нея, така че „произведените“ години живот да бъдат оптимални, като изчислението е направено в контекста на останалите седем възрастови групи.

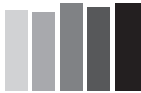
Входните данни за последващата DEA декомпозиция по метода на Брайнова (2012) се съдържат в една матрица с размерност 8×3 (табл. 2), като по редове са разположени възрастовите групи за всяка от страните (4 възрастови групи за 2 страни). В първите две колони са разположени декомпозиционните фактори дял на жените и дял на завършилите висше образование в съответната възрастова група за съответната страна. В последната колона се намира т.нар. общ брой години живот, дефиниран като сума на средната продължителност на предстоящия живот за хората на възраст, равна на долната граница на всяка изследвана възрастова група и стойността на тази долна граница (изживения брой години до момента).

2. Входни данни за DEA декомпозицията

Номер на групата - i	Име на групата	Дял на жените - I	Дял на висшистите - J	Общ брой години живот - t
1	27 - 34 Дания	49.9	33.1	79.5
2	35 - 44 Дания	49.3	31.8	79.7
3	45 - 54 Дания	49.5	29.3	80.2
4	55 - 64 Дания	50.1	25.1	81.3
5	27 - 34 Швеция	49.0	38.5	81.9
6	35 - 44 Швеция	49.0	30.2	82.1
7	45 - 54 Швеция	49.3	27.5	82.4
8	55 - 64 Швеция	49.9	25.3	83.2

3. Метод на замяната - ограничение върху теглата

В случая анализът на обгръщането се използва като декомпозиционна техника и тогава ресурсите (полът и образованието) влизат в ролята на декомпозиционни фактори. В резултат на DEA декомпозицията



за всяка единица (възрастова група в съответната страна) се изчисляват както нейната ефективност, така и оптималните тегла на факторите и продукта. Представената DEA декомпозиция е осъществена по **ориентиран към входа модел с ограничени множители**. Този метод изисква предварително фиксиране на съотношението между някои от теглата. Това ограничение е необходимо, тъй като в противен случай е възможно декомпозиционните ефекти да варират в много широки граници. Последното би довело до проблем с интерпретирането на резултатите, тъй като получените стойности биха противоречали на действителността.

Един възможен подход в този случай е фиксирането на параметър, равен на съотношението между съответните тегла, да бъде извършено по **метода на замяната** (trade-off approach), разработен от руския учен Podinovski. При този подход стойностите на ресурсите и продуктите се менят едновременно, така че промените им са технологично съвместими със съответния производителен процес. В резултат на тези технологично съвместими замени (trade-offs) диференциацията между отделните единици по отношение на производствената им ефективност става по-голяма и същевременно се избягват евентуални нулеви тегла. В своя статия Podinovski (2002) посочва основното предимство на метода на замяната пред други подобни методи - **той запазва радиалната природа на ефективността**. Под радиална природа на ефективността (при моделите, ориентирани към входа) се разбира свойството стойностите на всички ресурси да се намаляват в една и съща пропорция, а именно до достигане на оптимална ефективност на производствения процес, докато стойностите на продуктите остават постоянни. Podinovski допълва, че ако вместо ограничения, базирани на допустими замени, се използват такива от съображения за важността на различните фактори или съображения за тяхната парична равностойност, радиалната природа на ефективността не би се запазила. Повече информация по тази тема може да се намери в статията на Allen et al (1997) и в книгата на Thanassoulis (2001).

На практика методът на замяната работи чрез дефиниране на неравенства между теглата на факторите и продукта, като всяко неравен-

ство съответства на допустима замяна в производителния процес. Ако разгледаме входните данни от табл. 2 и тяхната динамика по възрастовите групи за двете страни, можем да обобщим:

- Когато дялът на жените нарасне средно с 0.2%, общият брой години нараства с 0.5 години.

- Когато дялът на висшистите намалее средно с 3.5%, общият брой години нараства с 0.5 години.

- Когато дялът на жените нарасне средно с 0.2%, дялът на висшистите намалява с 3.5%.

На езика на метода на Podinovski (2002) това може да се представи по следния начин:

$$0.5 * \text{брой_години} - 0.2 * \text{дял_жени} \leq 0$$

$$0.5 * \text{брой_години} + 3.5 * \text{дял_висшисти} \leq 0$$

$$-0.2 * \text{дял_жени} + 3.5 * \text{дял_висшисти} \leq 0$$

Неизвестните променливи в последните три неравенства са всъщност неизвестните тегла на съответните фактори, които се получават в резултат на анализа на обгръщането. Оттук от първото неравенство

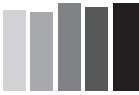
получаваме $\frac{\text{брой_години}}{\text{дял_жени}} \leq 0.4$. От второто получаваме противоречие

с положителните по дефиниция стойности. От третото неравенство

получаваме $\kappa = \frac{\text{дял_висшисти}}{\text{дял_жени}} \leq 0.055$. За ограничително условие е достатъчно да фиксираме съотношение на теглото на дела на висшистите и теглото на дела на жените $\kappa = 0.055$. Нека означим последното ограничително равенство с (6). Така си осигуряваме ненулеви тегла на факторите пол и образование за всяка възрастова група.

Следващата стъпка в анализа на обгръщането е зареждането на входните данни от табл. 2 заедно с ограничение (6) за теглата, на базата на които се пресмятат резултатите (4 параметъра) по избрания DEA модел за всяка от осемте възрастови групи, а именно:

- Стойност на функцията на ефективност
- Оптимално тегло на ресурса „пол“
- Оптимално тегло на ресурса „образование“
- Оптимално тегло на продукта „общ брой години живот“.



3. Резултати след приложение на ДЕА модела

Номер на групата - i	Име на групата	Ефективност	Тегло на пола - w^J	Тегло на обработването - w^J	Тегло на общия брой години
1	27 - 34 Дания	0.9482	0.0193	0.0011	0.0119
2	35 - 44 Дания	0.9630	0.0196	0.0011	0.0121
3	45 - 54 Дания	0.9666	0.0196	0.0011	0.0121
4	55 - 64 Дания	0.9732	0.0194	0.0011	0.0120
5	27 - 34 Швеция	0.9878	0.0196	0.0011	0.0121
6	35 - 44 Швеция	0.9986	0.0197	0.0011	0.0122
7	45 - 54 Швеция	0.9989	0.0197	0.0011	0.0121
8	55 - 64 Швеция	1.0000	0.0195	0.0011	0.0120

Резултатите от приложението на анализа на обгръщането показват, че изследваните осем възрастови групи имат близки стойности на ефективност, като се наблюдава плавно покачване на ефективността с нарастване на възрастта и при двете страни. Като цяло ефективността на групите от Швеция е по-висока в сравнение с Дания, макар че разликите са пренебрежими. Нека разгледаме двете „гранични групи в анализа” - населението на Дания на възраст между 27 и 34 години има ефективност 95% (0.9482) и живее средно до 79.5 години, а населението на Швеция между 55 и 64 години има оптимална ефективност 100% (1) и живее до 83.2 години (табл. 2). При интерпретацията на по-ниската ефективност при „производството“ на години живот на групата „Дания 27 - 34” в сравнение с групата „Швеция 55 - 64” трябва да се има предвид, че дялът на жените и на висшистите са пропорционални на общия брой години живот. Данните от табл. 2 показват, че групата „Дания 27 - 34” живее средно с 3.7 години по-малко от групата „Швеция 55 - 64”, въпреки че висшистите в групата „Дания 27 - 34” са със 7.8% повече от тези в групата „Швеция 55 - 64”⁸. За да достигнат оптимална ефективност в общия си брой години живот (при модел, ориентиран към входа⁹), дяловете на

⁸ Дяловете на жените в двете гранични групи съвпадат (49.9%).

⁹ При модел, ориентиран към изхода, достигането на оптимална ефективност би изисквало повишаване на общия брой години живот при наличните дялове жени и висшисти.

жените и на висшистите от групата „Дания 27 - 34” би трябвало да са равни на 95% от настоящите им стойности. С други думи, в контекста на останалите възрастови групи и техните преживяемости, наличните дялове на жените и висшистите от групата „Дания 27-34” предполагат по-голям брой години общ живот от 79.5, което обяснява неоптималната ефективност на тази група.

4. DEA декомпозиция и получените от нея ефекти

Терминът „DEA декомпозиция” е въведен за първи път от Брайнова (2012). При DEA декомпозицията всяка единица (в случая възрастова подгрупа на съответната страна) е отделен обект на декомпозиционния процес. На следваща стъпка се преминава от ниво „възрастова група в съответната страна” на ниво „възрастова група” и на практика се декомпозира всяка от n -те разлики $t_{n+i} - t_i$ в СППЖ между всеки две едноименни възрастови групи на съответните страни, където i се мени от 1 до n (броя на групите, в случая е $n = 4$). Според Брайнова DEA декомпозицията на разликата в СППЖ между две съвкупности има следния аналитичен вид:

$$t_{n+i} - t_i = A_i + B_i + C_i + D_i, \quad (7)$$

където:

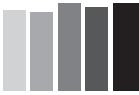
$$A_i \doteq \left(\frac{w_{n+i}^I * I_{n+i} + w_i^I * I_i}{2} \right) (t_{n+i} - t_i),$$

$$B_i \doteq \left(\frac{t_{n+i} + t_i}{2} \right) (w_{n+i}^I * I_{n+i} - w_i^I * I_i),$$

$$C_i \doteq \left(\frac{w_{n+i}^J * J_{n+i} + w_i^J * J_i}{2} \right) (t_{n+i} - t_i),$$

$$D_i \doteq \left(\frac{t_{n+i} + t_i}{2} \right) (w_{n+i}^J * J_{n+i} - w_i^J * J_i),$$

където I_i и J_i са стойностите съответно на факторите I и J (пол и образование) за i -тата група, w_i^I и w_i^J са съответните им оптимални тегла, получени в резултат на анализа, t_i е общият брой години живот за i -тата



група, като i се мени от 1 до n ($n = 4$). Общият брой на изследваните единици е $2n$ - по n на брой единици (възрастови групи) за всяка от двете съвкупности (страни), между които се пресмята разликата в СППЖ.

Можем да направим следната **интерпретация на получените ефекти чрез анализа на обгръщането:**

- **Ефектът A_j** , измерва влиянието на фактора I (пол) при условие, че НЕ се отчита различната му степен на влияние върху СППЖ на двете съвкупности (страни) и същевременно факторът J (образование) е изолиран.

- **Ефектът B_j** , отчита различната степен на влияние на фактора I (пол) върху СППЖ на двете страни. Ако искаме да измерим влиянието на пола и същевременно да отчетем различната му степен на влияние върху СППЖ (при условие, че факторът образование е изолиран), трябва да пресметнем сумата на първите два ефекта ($A_i + B_i$).

- **Ефектът C_j** , измерва влиянието на фактора J (образование) при условие, че НЕ се отчита различната му степен на влияние върху СППЖ на двете страни и същевременно факторът I (пол) е изолиран.

- **Ефектът D_j** , отчита различната степен на влияние на фактора J (образование) върху СППЖ на двете страни. Ако искаме да измерим влиянието на образованието и същевременно да отчетем различната му степен на влияние върху СППЖ (при условие, че факторът пол е изолиран), трябва да пресметнем сумата на последните два ефекта ($C_i + D_i$).

Резултатите от ДЕА декомпозицията на разликата в СППЖ между Дания и Швеция по пол и образование по възрастови групи са посочени в табл. 4.

4. Резултати от ДЕА декомпозицията по възрастови групи

Възрастови групи	А ефект пол	В ефект пол	С ефект образование	Д ефект образование	Общ ефект (разлика в СППЖ в полза на Швеция)
27 - 34	-2.26	0.51	-0.09	-0.51	-2.35
35 - 44	-2.26	-0.12	-0.08	0.12	-2.34
45 - 54	-2.14	-0.14	-0.07	0.14	-2.21
55 - 64	-1.82	0.03	-0.05	-0.03	-1.87

Получените декомпозиционни резултати не са пряко сравними с резултатите от метода на Das Gupta (табл. 1) поради нееднородното естество на получените ефекти и поради зависимостта на резултатите от фиксирания параметър $\kappa = 0.055$. **Следователно възниква необходимостта да се потърси друга зависимост**, която:

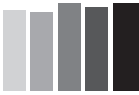
- Да не зависи от параметъра κ .
- Да характеризира DEA декомпозицията.
- Да осигурява сравнимост между резултатите на метода на Das Gupta и DEA декомпозицията.

5. Хипотетични съвкупности

Дотук имаме DEA декомпозиция за всяка възрастова група, но тя не дава като резултат чистите ефекти на факторите (пол и образование) за разлика от метода на Das Gupta. Затова по **аналогия с Das Gupta разглеждаме хипотетични съвкупности**, чиито стойности са средноаритметични от реалните стойности на факторите пол и образование на двете страни за i -тата възрастова група, а именно:

$$Q_i = Q_{n+i} = \frac{I_{n+i} + I_i}{2} \text{ и } R_i = R_{n+i} = \frac{J_{n+i} + J_i}{2}. \quad (8)$$

Вследствие на полагането (8) оптималното тегло w_i^I (w_i^J) на фактора I - пол (J - образование), което се отнася за i -тата възрастова група съвпада при двете страни (Дания и Швеция). С други думи, от полагането (8) за всяко i от 1 до 4 следват равенствата $w_{n+i}^I = w_i^I$ и $w_{n+i}^J = w_i^J$ и след заместването им в ефектите B_i и D_i от формула (7) се получава, че $B_i = D_i = 0$. По този начин получаваме друг декомпозиционен случай на разликата в СППЖ за i -тата възрастовата група между двете страни. Полагането (8) означава, че използваната хипотетична съвкупност се характеризира с едни и същи (средноаритметични) разпределения при двете страни както по фактора I (пол), така и по фактора J (образование) и с реалните стойности на съответния общ брой години живот за всяка възрастова група за съответната страна, а именно t_{n+i} за i -тата възрастова група за Швеция и t_i за i -тата възрастова група за Дания. Получените декомпозиционни резултати в този частен случай са представени в табл 5.



5. Резултати от DEA декомпозицията при използване на хипотетични съвкупности за пола и образованието

Възрастови групи	А ефект пол	В ефект пол	С ефект образование	Д ефект образование	Общ ефект (разлика в СППЖ в полза на Швеция)
27 - 34	-2.26	0.00	-0.09	0.00	-2.35
35 - 44	-2.26	0.00	-0.08	0.00	-2.34
45 - 54	-2.14	0.00	-0.07	0.00	-2.21
55 - 64	-1.82	0.00	-0.05	0.00	-1.87

От сравнението между табл. 4 и 5 се вижда, че ефектите А и С (и общият ефект) се запазват след използването на съответните хипотетични съвкупности, а ефектите В и D се нулират. Дотук беше описан само един от възможните случаи на DEA декомпозиция при използване на хипотетични съвкупности. Следва описание и на останалите три случая.

6. Четири случая на DEA декомпозиция при използване на хипотетични съвкупности. Взаимно изключване и адитивност на случаите.

С помощта на използваните различни хипотетични съвкупности по отношение на независимите декомпозиционни фактори (пол и образование) както поотделно, така и заедно можем да конструираме общо четири взаимноизключващи се случая:

1. **Не ползваме хипотетични съвкупности**, т.е. всички величини в анализа приемат реалните си стойности $I_{n+i}, I_i, J_{n+i}, J_i, t_{n+i}$ и t_i .
2. Използваме хипотетична съвкупност **само по отношение на фактора J (образование)**, т.е. $R_i = R_{n+i} = \frac{J_{n+i} + J_i}{2}$, а всички останали величини (пол и общ брой години живот) приемат реалните си стойности I_{n+i}, I_i, t_{n+i} и t_i .
3. Използваме хипотетична съвкупност **само по отношение на фактора I (пол)**, т.е. $Q_i = Q_{n+i} = \frac{I_{n+i} + I_i}{2}$, а всички останали величини (образование и общ брой години живот) приемат реалните си стойности J_{n+i}, J_i, t_{n+i} и t_i .
4. Използваме хипотетични съвкупности **по отношение и на двата фактора I (пол) и J (образование)**, т.е. $Q_i = Q_{n+i} = \frac{I_{n+i} + I_i}{2}$ и

$R_i = R_{n+i} = \frac{J_{n+i} + J_i}{2}$, а зависимите средни величини (общ брой години живот) приемат реалните си стойности t_{n+i} и t_i ¹⁰.

На практика всяка от тези хипотетични съвкупности уеднаквава стойностите на едноименните възрастови групи на Дания и Швеция по съответния фактор и по този начин изолира неговото влияние в декомпозиционния процес. За всеки един от първите три случая (и за всяка от четирите възрастови групи) се пресмята отделна DEA декомпозиция, като получените декомпозиционни резултати са представени в табл. 6. Непосредствено след числените резултати, поместени в табл. 6, е описана адитивността на отделните случаи, както и интерпретацията на всеки от тях.

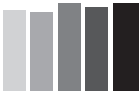
6. Декомпозиционни резултати за случаи № 1, № 2 и № 3 по възрастови групи

Случай № 1 - отчитат се и полът, и образованието	А ефект	В ефект	С ефект	Д ефект	Общ ефект (разлика в СППЖ в полза на Швеция)
27 - 34	-2.26	0.51	-0.09	-0.51	-2.35
35 - 44	-2.26	-0.12	-0.08	0.12	-2.34
45 - 54	-2.14	-0.14	-0.07	0.14	-2.21
55 - 64	-1.82	0.03	-0.05	-0.03	-1.87

Случай № 2 - отчита се само полът	А ефект	В ефект	С ефект	Д ефект	Общ ефект (разлика в СППЖ в полза на Швеция)
27 - 34	-2.26	0.05	-0.09	-0.05	-2.35
35 - 44	-2.26	0.02	-0.08	-0.02	-2.34
45 - 54	-2.14	0.01	-0.07	-0.01	-2.21
55 - 64	-1.82	0.01	-0.05	-0.01	-1.87

Случай № 3 - отчита се само образованието	А ефект	В ефект	С ефект	Д ефект	Общ ефект (разлика в СППЖ в полза на Швеция)
27 - 34	-2.26	0.45	-0.09	-0.45	-2.35
35 - 44	-2.26	-0.13	-0.08	0.13	-2.34
45 - 54	-2.14	-0.15	-0.07	0.15	-2.21
55 - 64	-1.82	0.02	-0.05	-0.02	-1.87

¹⁰ Декомпозиционният случай № 4 е описан в точка 4 и резултатите за него се съдържат в табл. 5.



Ефектите A_i и C_i (които НЕ отчитат различната степен на влияние съответно на факторите пол и образование върху двете съвкупности) приемат едни и същи стойности при всеки от случаите - означени с a_i и c_i в табл. 7. Според табл. 6 между декомпозиционните резултати от различните случаи съществува адитивност по отношение на ефектите B_i и D_i . По-точно, сумата от ефекта B_i от случай № 2 и ефекта B_i от случай № 3 за i -тата възрастова група е равна на ефекта B_i от случай № 1¹¹.

Таблица 7 представлява обобщение на декомпозиционните резултати от четирите разгледани случаи. Освен адитивност по отношение на ефекта В, се наблюдава и аналогична адитивност при ефекта D, като индексът i се мени от 1 до 4 по възрастовите групи при всеки от случаите.

7. Обобщени декомпозиционни резултати за случаи № 1, № 2, № 3 и № 4 по възрастови групи

DEA	Декомпозиционни ефекти				Декомпозирана разлика
Случай №	A_i	B_i	C_i	D_i	
1	a_i	b_i	c_i	d_i	$t_{n+i} - t_i$
2	a_i	$b_i * \pi_i$	c_i	$d_i * \pi_i$	$t_{n+i} - t_i$
3	a_i	$b_i * (1 - \pi_i)$	c_i	$d_i * (1 - \pi_i)$	$t_{n+i} - t_i$
4	a_i	0	c_i	0	$t_{n+i} - t_i$

Като се има предвид как са дефинирани хипотетичните съвкупности, характеризиращи случаите № 1, № 2 и № 3, е възможна следната **интерпретация** на всеки от тях:

Случай № 1 показва декомпозиционните ефекти при условие, че се отчита влиянието **и на двата фактора** пол и образование (I и J) върху общия брой години живот.

Случай № 2 показва декомпозиционните ефекти при условие, че се отчита влиянието **само на фактора пол** (I) върху общия брой години живот, факторът образование (J) е пренебрегнат.

¹¹ Теоретично в сумата от чистите ефекти може да се добави и случай № 4, но това не ни дава нищо ново, тъй като той се характеризира с нулеви ефекти B_i и D_i .

Случай № 3 показва декомпозиционните ефекти при условие, че се отчита влиянието **само на фактора образование** (J) върху общия брой години живот, факторът пол (I) е пренебрегнат.

В резултат на намерената адитивност между случаите и имайки предвид интерпретацията на случаите № 2 и № 3, може да се заключи, че числото $\Pi_i \doteq \frac{\pi_i}{1 - \pi_i}$, получено от табл. 7, оценява съотношението между чистите ефекти от факторите пол и образование за i -тата възрастова група. Намерената числова характеристика Π_i притежава желаните свойства:

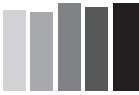
- Не зависи от предварително фиксирания параметър κ^{12} , който задава връзка между теглата на факторите пол и образование.
- Получава се в резултат на DEA декомпозицията.
- Осигурява сравнимост на резултатите между DEA декомпозицията и метода на Das Gupta.

7. Връзка между DEA декомпозицията и метода на Das Gupta

Последното твърдение се нуждае от допълнително разяснение, което да покаже защо намерената числова характеристика Π_i осигурява връзка между двата метода. **Характеристиката Π_i е изведена в резултат на DEA декомпозицията** и приема стойностите $\{0.12, -0.11, -0.6, 0.59\}$ съответно за възрастовите групи $\{27 - 34, 35 - 44, 45 - 54, 55 - 64\}$ (вж. табл. 6). Това потвърждава второто от желаните три свойства. Данните на Евростат¹³, използвани за DEA декомпозицията, са използвани и за да се декомпозира разликата в СППЖ по метода на Das Gupta по същите фактори (пол и образование) и в рамките на същите възрастови групи. Резултатите по метода на Das Gupta са показани в табл. 1. **При използването на тези входни данни стойностите на Π_i по възрастови групи ще се запазят, каквато и стойност да изберем за параметъра κ , което верифицира първото свойство.**

¹² Каквото и да е фиксираното от изследователя ограничение, съотношението Π_i не променя стойността си за всяка от възрастовите групи.

¹³ Методът на Das Gupta изисква данните на входа да бъдат агрегирани по начин, различен от този при DEA декомпозицията, но на практика началните данни, използвани при двата метода, са едни и същи. Малките разлики при разликите в СППЖ по възрастови групи между двата метода се дължат именно на различното агрегиране на входните данни.



Нека вземем сумата на чистите ефекти от факторите пол и образование, пресметнати по метода на Das Gupta (I effect + J effect), и я разпределим в намереното съотношение Π_i . Получените стойности са много близки до оригиналните стойности на чистия ефект от пола I и чистия ефект от образованието J , пресметнати по метода на Das Gupta.

8. Сравнение между чистите ефекти от факторите пол и образование по метода на Das Gupta и разпределението на сумата им в съотношение Π_i

Възрастови групи	Das Gupta (1978)		DEA декомпозиция	
	чист ефект от разпределението по пол	чист ефект от разпределението по степен на образование	сумата от ефектите на пола и образованието - π_i	сумата от ефектите на пола и образованието $(1 - \pi_i)$
27 - 34	0.0	-0.3	0.0	-0.2
35 - 44	0.0	-0.1	0.0	-0.1
45 - 54	0.0	-0.1	0.0	-0.1
55 - 64	0.0	0.0	0.0	0.0

Всичко това означава, че намерената характеристика Π_i може да се използва като оценка на съотношението между чистите ефекти, получени по метода на Das Gupta. С други думи, Π_i задава нетривиална и независима от ограниченията на анализа на обгръщането връзка между DEA декомпозицията и метода на Das Gupta (1978), което е основна цел на настоящото изследване и потвърждава последното трето свойство.

8. Предимства на анализа на обгръщането DEA

Тук са представени накратко основните предимства на анализа на обгръщането, изхождайки от неговите свойства.

Първо предимство на използвания анализ е, че може да работи с **всякакви хомогенни единици** (не само възрастово-полови групи). Това го прави мощен инструмент, адаптивен към много и разнообразни задачи. Освен това, DEA може да работи и с **повече от два ресурса**¹⁴, които в настоящия случай се явяват декомпозиционни фактори. Това разширява изследователските хоризонти и при наличие на необходимите данни допринася за вдълбочеността на проведения анализ.

¹⁴ За да се осигури коректност на анализа, някои автори препоръчват съотношението между сумарния брой на ресурсите и продуктите и броя на единиците в анализа да се мени в определени граници, които варират при различните автори.

Второ предимство е фактът, че DEA позволява използването на ресурси и продукти в **различни мерни единици**. Следователно DEA декомпозицията е възможна, без ограничение от вида на мерните единици на факторите, което улеснява изследвателя и същевременно позволява да включим в анализа повече фактори. В резултат на това полето на приложение на метода се разширява и не на последно място се постига по-голямо разнообразие от получени резултати.

Третото предимство на предложения анализ има чисто технически характер и се отнася за начина на изчисление на решението на оптимизационната задача, което може да се осъществи с помощта на стандартно приложение към Excel, нар. Solver parameters, създадено за решение на линейни оптимизационни задачи, или чрез специална програма, нар. DEAFrontier, създадена от Joe Zhu¹⁵. Увеличаването на броя на изследваните фактори не оказва съществено влияние върху сложността и **скоростта на изчисление на търсените параметри** и при двете приложения.

9. Резултати и заключения

В заключение можем да обобщим получените резултати:

Първо, настоящото изследване верифицира използването на анализа на обгръщането DEA в ролята на декомпозиционна техника. Този резултат е принос сам по себе си, тъй като дава нестандартно приложение на DEA в областта на демографската статистика. Подробно описание на DEA декомпозицията може да се намери в статията на Брайнова (2012), където е изложена декомпозиция на СППЖ между мъжете и жените в България според факторите индекс на телесна маса и тютюнопушене по данни на Евростат за 2008 година.

Второ, конструирани са четири взаимноизключващи се случая на DEA декомпозиция, чиито резултати са адитивни помежду си. Два от тези случая се характеризират с чистото влияние само на един от факторите (при случай № 2 е изолиран факторът образование, а при случай № 3 - факторът пол). Изследването показва как съотношението на техните резултати P_i може да се използва като оценка на съотношението на чистите ефекти от метода на Das Gupta (1978). Важно е да се отбележи, че съотношението P_i е независимо от ограниченията на използвания

¹⁵ Програмата DEAFrontier може да се инсталира върху платформата на Solver parameters и е безплатна при условие, че се използва за научни цели.



DEA модел и при разпределяне на сумата от чистите ефекти на Das Gupta в това съотношение, получените резултати са много близки до самите чисти ефекти на Das Gupta. По този начин намерената характеристика P_i задава търсената връзка между двата използвани подхода - анализа на обгръщането DEA и метода на Das Gupta (1978).

Трето, анализът на обгръщането допринася към декомпозицията и други нови числови характеристики в лицето на функцията на ефективност и оптималните тегла на ресурсите и продуктите за всяка единица (вж. табл. 3). Това обогатява възможностите за интерпретация на декомпозиционните резултати и не на последно място, разкрива нова методологическа гледна точка в областта на декомпозиционните техники.

В заключение анализът на обгръщането е гъвкав и мощен инструмент за анализ на разнообразна статистическа информация с множество различни приложения предимно от практически характер. Основният принос на настоящата статия е, че тя предлага ново приложение на този анализ, което в по-голямата си степен е методологическо. Така DEA е поставен в контекста на декомпозиционните техники, които само по себе си поставя началото на редица изследователски предизвикателства.

ЦИТИРАНА ЛИТЕРАТУРА:

Брайнова, П. (2012). Влиянието на индекса на телесната маса и тютюнопушенето върху разликата в средната продължителност на предстоящия живот, *Статистика*, кн. 1-2/2012, с. 67 - 89.

Калянов, Т. (2011). Декомпозиция на промените в средната продължителност на предстоящия живот, *Население* 1-2/2011, с. 3 - 18.

Русев, Б. (1969). Смъртност и средна продължителност на предстоящия живот, *Статистика*, кн. 2, с. 18 - 35.

Русев, Б. (1977). Измерване влиянието на повъзрастовата смъртност върху средната продължителност на живота, *Икономическа мисъл*, кн. 1, с. 70 - 79.

Салчев П., П. Атанасов, Т. Димитрова-Савова, Х. Грива (2012). Икономически анализ и оценка на ефективността на болничните услуги в България чрез DEA метода в периода 2010 - 2012 г., *Национален център по общественото здраве и анализи* (под печат в *Българско списание за общественото здраве*).

Allen R., A. Athanassopoulos, R. G. Dyson, E. Thanassoulis (1997). Weights Restrictions and Value Judgments in DEA: Evolution, Development and Future Directions. *Annals of Operations Research*; 73, pp. 13 - 34.

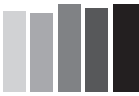
Cook W. D., J. Zhu (2008). *Data Envelopment Analysis, Modeling Operational Processes and Measuring Productivity*.

Cooper W. W., L. M. Seiford, K. Tone (2007). *Data Envelopment Analysis, A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Springer.

Das Gupta, P. (1978). A General Method of Decomposing a Difference between Two Rates into Several Components, *Demography*, Vol. 15, No.1, pp. 99 - 112.

Das Gupta, P. (1993). *Standardization and Decomposition of Rates: A User's Manual*. Current Population Reports Series P-23, No. 186. U.S. Bureau of the Census, Washington, DC.

Eurostat, Database Population on 1 January by age, sex and educational attainment (ISCED 1997).



Kitagawa, E. M. (1955). Components of a Difference between Two Rates, *Journal of American Statistical Association*, Vol. 50, No. 272, pp. 1168 - 1194.

Podinovski, V. V. (2002). Weight restrictions and radical measures of efficiency, *Warwick Business School Research paper № 352*, University of Warwick, UK.

Retherford R. D., L. J. Cho (1973). Comparative Analysis of Recent Fertility Rates in East Asia, *International Union for Scientific Study of Population, Proceedings of 17-th General Conference of the IUSSP*, Vol. 2, pp. 163 - 181.

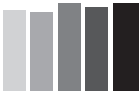
Thanassoulis, E. (2001). *Introduction to the Theory and Application of Data Envelopment Analysis*. Kluwer Academic Publishers, Dordrecht.

ДЕКОМПОЗИЦИЯ РАЗНИЦЫ В СРЕДНЕЙ ПРОДОЛЖИТЕЛЬНОСТИ ПРЕДСТОЯЩЕЙ ЖИЗНИ С ПОМОЩЬЮ ОБОЛОЧЕЧНОГО АНАЛИЗА ДАННЫХ И СВЯЗИ С МЕТОДОМ DAS GUPTA (1978)

*Петя Брайнова**

РЕЗЮМЕ Представлен так называемый „оболочечный анализ данных” - Data Envelopment Analysis (DEA), как прием декомпозиции и указана его связь с методом Das Gupta (1978). Метод Das Gupta декомпозирует разницу в средней продолжительности предстоящей жизни (СППЖ) между двумя совокупностями по нескольким факторам, при этом кроме чистых эффектов факторов, в формуле присутствует и остаточный эффект. Метод DEA декомпозирует разницу в СППЖ (общий эффект) по факторам с учетом или без учета их различной степени воздействия на СППЖ исследуемых совокупностей. В итоге этого, полученная декомпозиция не содержит чистых эффектов и зависит от предварительно фиксированного параметра, равного соотношению между весами факторов, что порождает проблему сопоставимости между DEA декомпозицией и методом Das Gupta. В настоящем исследовании предлагается методологическое решение этой проблемы посредством числовой характеристики, независимой от ограничений анализа.

* Докторант на кафедре статистики и эконометрии, Университет национального и мирового хозяйства (УНМХ); e-mail: petia_brainova@abv.bg.



A DECOMPOSITION OF THE DIFFERENCE IN LIFE EXPECTANCY THROUGH DATA ENVELOPMENT ANALYSIS AND ITS RELATIONSHIP WITH THE METHOD OF DAS GUPTA (1978)

*Petya Braynova**

SUMMARY The current article presents a nonparametric method called Data Envelopment Analysis (DEA) in the role of a decomposition technique and shows its relationship with the classic method of Das Gupta (1978) for decomposition of the difference in life expectancy (LE) between two populations. According to the data, classified by two factors (I and J) Das Gupta's method decomposes the difference in LE as the sum of the net I-effect, the net J-effect and the residual effect. DEA decomposes the difference in LE (total effect) by both factors - I and J - with or without taking into account the different extent of influence they have over the LE of the explored populations. As a result, a decomposition formula is produced, none net effects are derived and also the results are dependent of a fixed parameter, which is a ratio of the factor weights. Hence, a problem concerning the comparability between DEA-decomposition technique and Das Gupta's method arises. The current article provides a methodological solution to this problem of the form of numerical characteristics independent of the limitations of the analysis.

* A doctoral candidate at the "Statistics and Econometrics" Academic Department to the University of National and World Economy, Sofia; e-mail: petya_brainova@abv.bg.