**Open access is not a panacea, even if it’s radical – an empirical study on the role of shadow libraries in closing the inequality of knowledge access.**

1. A háttérirodalomhoz a párhuzamok miatt ajánlom: ***The death of distance revisited: cyber-place, physical and relational proximities***

**EMMANOUIL TRANOS, PETER NIJKAMP tanulmányát. (ennek magyar fordítását mellékelem, de meg tudjátok keresni az angolt is). Az internet (IP) aktivitás erős térbeli meghatározottságát és agglomerálódását. Ez egybeesik azzal a megállapításotokkal, hogy a fejlettebb térségek esetében az internetpenetráció is nagyobb, ami több letöltéshez vezethet.**

<https://ideas.repec.org/p/tin/wpaper/20120066.html>

related articles:

Tranos, E. (2011): The topology and the emerging urban geographies of the Internet backbone and

aviation networks in Europe: a comparative study. *Environment and Planning A,* 2., 378–392.

Tranos, E. (2013): *The geography of the internet: cities, regions and the internet infrastructure in Europe.*

Edward Elgar

Tranos, E., Gillespie, A. (2009): The spatial distribution of Internet backbone networks in Europe: a

metropolitan knowledge economy perspective. *European Urban and Regional Studies,* 4.,

423–437.

2. Emerging country effect: „In low-income countries, increasing GDP causes much larger shadow library use than in high-income countries.” Mindez a lagging behind hatással is magyarázható. A gyorsabb felzárkózáshoz növelni kell a tudástermelést és az ahhoz szükséges a tudásbázist szélesíteni. A legális tudástartalom elérhetősége nem bővül olyan gyorsan,mint azt a gyorsabb gazdasági növekedés igényei megkívánják, ezért ezt az illegális shadow források bevonásával kompenzálják. Illetve releváns az a megállapításotok, hogy tudásgazdaságba történő beruházásoknál először a hard tudásinfrastruktúra kiépítése élvez prioritást az adatbázis és publikációs adatbázisok megvásárlásával szemben. (illetve a ezek az adatbázisok relatíve drágábbak a publikációs platformok (cégek) nagyon magas profitrátája és monopóliuma miatt.

# Érdems lenne összevetni az Európán kívüli emerging gazdaságokat a CEE-vel. Anno hallgattam Andreas Nölke előadását (cikke a témában: „Dependent versus state-permeated capitalism: two basic options for emerging markets” in [International Journal of Management and Economics](https://www.researchgate.net/journal/2299-9701_International_Journal_of_Management_and_Economics) 54(4):269-282 ·: megállapította, hogy kutatásra és oktatásra CEE kevesebbet költ , mint a BRIC országok. Kíváncsi lennék hogy más aspektusokkal is magyarázhatók e a két országcsoport különbségei pl. a tudástermelésben és abszorpcióban.

1. ’ access to knowledge is only one element in the complex infrastructural mix which then produces demand for the knowledge shadow libraries may offer.’

„. In low-income countries, higher investment may lead to fast-growing knowledge absorption capacity, which may not be met with appropriate infrastructural support. This means that low-income countries generate less shadow library usage in general, but within that group, larger investment into knowledge-intensive activities has greater positive impact on usage.”

Egy kicsit relevánsnak érzem itt a 2007-es dél-dunántúli R& D vizsgálatunkat. A 20 régiós EUs vizsgálatban a fő kérdés az volt, hogy a K+F beruházások növelése vezet e nagyobb innovációhoz és gazdasági növekedéshez. A fő megállapítás az volt, hogy a perifériákon hiába nő a K+F ráfordítás, ha a befektetés a perifériák „ritka aktor sűrűségű” közegébe érkezik, ahol nincsennek nagyvállalatok (demand oldalról) és így a tudástermelés bővülése gazdasági oldalról fenntarthatatlan (és kormányzati oldalról sincsennek hatalmas források).

A szegényebb országok esetében a nemcsak a hiányzó infrastrukturális támogatás lehet a magyarázó (az internetellátottság sem), hanem tudástermelői külső kapcsolatok hiánya (nemzetközi tudáshálózatokba való integráltság) és a gyenge gazdasági háttér is kontraproduktív lehet, ami nem serkenti első körben a shadow library aktivitást sem. Az alacsonyabb GDP/per capita, income kisebb tudasabszorpciós kapacitást generál.

Az erősebb tudástermeléshez és hatékonyabb trudásabszorpcióhoz kell egy kiforrott intézményrendszer, ahol a tudás akkumulacionak minimum pár évtizedes hagyománya, stabil intezmenyrendszere és hosszú távon kiszámítható finanszírozási háttere és kiforrott protokollja van. Tehát nagyon igaz, hogy nemcsak az open/shadow access és az internethozzáférés, és nem is csak pénz kérdés a tudomány, de ez utóbbi az egyik esszenciális előfeltétel (olyan mint az FDI befektetéseknél a közlekedési infrastruktúra: ha nincs akkor biztosan nem jön a befektető, de ha van attól még nem biztos hogy tényleg lesz is befektetés). Tehát csak a ráfordítások növelése és az open access (shadow access) még önmagában nem vezet tudásakkumulacióhoz. De az eseti többletráfordításokkal finanszírozott eseti projektek megnövelhetik a Shadow letöltési intenzitást. De mivel nincs kiszámítható intézményi környezet és ezek a régiók érzékenyebben reagálnak a külső gazdasági sokkokra is, ezért a tudásakkumuláció folyamata is gyakran megszakad, tehát nincs egy fenntartható növekedési pálya.

De a szétaprózott (több állás) és fenntartható/kiszámítható kutatói életpályamodell hiánya több CEE/SEE országban korlátozza – ahogy írjátok is- a tudásakkumulaciót.

1. Az európai modellnél a Tudástermelés és abszorpciós kapacitásokhoz a R& D personnel adatok mellett lehetne nézni a ***PHD val rendelkezők számát***, illetve a ***patent*** adatokat lehetne használni.

**Az internet proficiency**: alternatív változó lehet a ***mobil banki applikációt*** használók aránya/száma (a Fintech terndek miatt a hagyományos webes internetbanki használattól erre mozdult el a bankpiac).

**5 A módszertannal kapcsolatosan kétfajta kommentet fűznék a cikkhez.**

Elsőként a fő problémám írnám le. A cikk több modellt épít a vizsgált kérdés megválaszolására, melyből a Model 1 egyszerű lineáris regresszió, míg a többi modell Poisson-regresszió. A cikk azt írja, mivel az eredményváltozó a nemnegatív egész számok halmazán alapuló (count data) megfigyeléshalmaz, ezért a Poisson-regresszió alkalmazhatósága alátámasztott. Az valóban igaz, hogy Poisson-regresszió akkor alkalmazható csak, ha a megfigyelések ebből a számhalmazból származnak, ugyanakkor emellett az is szükséges, hogy az eredményváltozó megfigyelései Poisson-eloszlást kövessenek, amely jelen adatállomány esetén nem áll fenn (vagy erre a cikk nem tér ki). Habár a harmadik ábrán egy olyan oszlopdiagram rajzolódik ki, amelynek alakja hasonlít a Poisson-eloszlás megfelelő paraméterű súlyfüggvényéhez, ez mégsem az. A harmadik ábra első részén ugyanis nem sűrűségfüggvény látható, hiszen nem relatív gyakoriságokat láthatunk rajta, csupán egymás mellé rendezett megfigyeléseket csökkenő sorrendben. Így a Poisson-regresszió alkalmazását nem látom alátámasztottnak. További problémám ezzel, hogy a Poisson-regresszió alkalmazása feltételezi az eredményváltozó megfigyeléseinek egymástól való függetlenségét. Mivel a cikk elemez térbeli kapcsolatokat is az eredményváltozó esetén, így önmagában megsérti a Poisson-regresszió alkalmazhatóságának feltételeit.

Amennyiben mégis ez a regressziós módszer maradna, amit én kifejezetten ellenjavallok, akkor további problémák merülnek fel a modellépítéssel. A Poisson-regresszió az eloszlás tulajdonságaiból kiindulva tipikusan log-lin regresszió, tehát a modell a következőképp futtatható:

mert ekkor kapjuk vissza a Poisson-regresszió exponenciális alakját:

A lineáris formulából jól látszik, hogy a cikkben futtatott modellek egyike sem felel a fenti felírásnak, hiszen a jobb oldalon logaritmált és nem logaritmált változók egyaránt szerepelnek. Így ennek átgondolása szintén javallott lenne.   
A random-effect model esetén semmilyen arra utaló jel nincsen, hogy miért ezt választották a szerzők szemben a fixhatásos vagy első differencia modellekkel. Ennek indoklása is szükséges lenne.  
Alapvetően én ebben az esetben lineáris modelleket alkalmaznék, az adatok tulajdonságaihoz azt gondolom, az a megfelelő.

A második típusú megjegyzések a kisebb, vagy gyorsabban korrigálható csoportba tartoznak.  
Elsőként, a szignifikancia szinteketek szokásosan 1, 5, 10%-on választjuk meg. A cikk is így elemez, viszont a regressziós eredmények táblái nem ezeket a számokat tartalmazzák, érdemes lenne ezt egységesíteni.  
Jóllehet, panel eszközöket is alkalmaznak a szerzők a random hatásos modellek esetén, ugyanakkor az adatállomány szerkezetéről semmit nem tudunk (mekkora keresztmetszet, hány időszak, stb.) Ez annak eldöntéséül is szolgálna, hogy a kiválasztott panel módszer megállja-e a helyét. Továbbá mindenképpen szükséges egy leírás az adatokról. Leíró statisztika is csak az eredményváltozóról készül, a többiről nem, holott ez is igen fontos információkat tartalmaz.

Az előző bekezdés már csak azért is lenne igen hasznos, mert az adatok típusától függhet az alkalmazott módszer. Több helyen a cikkben említésre kerül a megfigyelések száma, amely összevetve az országok számával azt sejteti, hogy az időtáv nem feltétlenül elég hosszú egy random hatásos modell „értelmes” becsléséhez. Az információk hiányában azonban ez csak sejtés.  
A szöveget át kell önteni „olvasható formába”. Azt gondolom, nem szerepelhetnek olyan mondatok egy cikkben, amelyek olyan elemeket tartalmaznak, minthogy a „VIF minden esetben <2”. Ez helyesen leírva a „VIF minden esetben kettőnél kisebb értéket vett fel” formát ölti.  
A regressziós táblákból hiányoznak az illeszkedési (R-négyzet, stb.) mutatók. Szokás szerint a megfigyelések számát és valamilyen modellilleszkedés mutatót tüntetünk fel a táblákban, a devianciák nem hordoznak annyi információt ezekben az esetekben.