

Experimentálna ekonómia

Prednáška 8: Experiment zo štatistickej perspektívy

Matej Lorko

matej.lorko@euba.sk

Materiály: www.lorko.sk/lectures

Referencie:

- Weimann, J., & Brosig-Koch, J. (2019). *Methods in experimental economics*. Springer International Publishing. Chicago
- Jacquemet, N., & l'Haridon, O. (2018). *Experimental economics*. Cambridge University Press.

Experiment zo štatistickej perspektívy

- Ak má byť výskumná otázka zodpovedaná experimentom a s využitím štatistických metód, potom musí byť experiment navrhnutý tak, aby na túto otázku odpovedal čo najlepšie. „Čo najlepšie“ znamená, že formálna metóda analýzy musí byť kompatibilná so štatistickou povahou vygenerovaných dát.
- Inými slovami, výskumná otázka, návrh experimentu, výsledné vygenerované dáta a štatistická analýza musia byť navzájom zladené.
- Zle navrhnutý experiment vedie k slabému vedeckému výsledku – ani tá najsofistikovanejšia metóda analýzy to nemôže zmeniť. Na druhej strane, dobre vybraná analytická metóda môže z dobre navrhnutého experimentu vyvodiť ešte zaujímavejšie zistenia.

Experiment zo štatistickej perspektívy

- Zo štatistického pohľadu je vhodné priebeh experimentálnej štúdie rozdeliť na fázu návrhu a fázu realizácie. Fáza návrhu, ktorá sa realizuje ako prvá, pozostáva z nasledujúcich úloh a typických problémov:
- Operacionalizácia výskumnej otázky: Aké dáta potrebujeme v experimente zbierať, aby bolo možné odpovedať na výskumnú otázku? Dajú sa tieto dáta jasne definovať ako premenné? Ako by sa mali tieto premenné merať? Ktorá z nich je závislá premenná? Ktoré z nich sú nezávislé premenné?
- Štruktúrovanie štatistického dizajnu: Ktoré premenné má experimentátor manipulovať a akým spôsobom (výber treatmentu)? Ktoré premenné môžeme ovládať a ako možno minimalizovať nežiaduce variácie závislej premennej? Ako by sa mala vybrať vzorka subjektov? Koľko subjektov potrebujeme, aby sme dokázali určiť, či hľadaný efekt „existuje“ s dostatočnou pravdepodobnosťou? Na akých úrovniach sa budú sa premenné merať (within vs. between-subject)? Ako často a kedy by sme mali tieto premenné merať? Ktoré premenné sú kvalitatívne a ktoré kvantitatívne?
- Prevod výskumnej otázky do štatistickej hypotézy alebo štatistického modelu: Aký formálny vzťah by mohol existovať medzi pozorovanou variáciou závislej premennej a variáciou nezávislých premenných? Ktoré premenné majú fixné efekty a ktoré náhodné efekty?

Výber vhodnej štatistickej metódy analýzy

- Aký je účel štatistickej analýzy: poskytnúť popisnú reprezentáciu dát a efektov treatmentu? Urobiť štatistický záver týkajúci sa populácie, z ktorej sa vzorka čerpá (inferencia)? Robiť predpoveď na základe odhadovaného modelu? Aké sú hlavné štatistické charakteristiky experimentálneho dizajnu alebo výsledných dát (odpovede z predchádzajúcich otázok)? Aké analytické metódy možno použiť vzhľadom na hlavné štatistické charakteristiky?
- Počítačové spracovanie dát: Chýbajú v dátach niektoré hodnoty? Existujú odľahlé pozorovania (outliers)? Existujú subjekty, ktoré zjavne urobili náhodné rozhodnutia (nepochopili experiment)?
- Vytváranie nových premenných z (kombinácie) už zhromaždených premenných (napr. skupinové priemery); Vytvorenie zoznamu premenných s popisom.
- Analýza dát: Popis dát pomocou kľúčových ukazovateľov; Grafické znázornenie dát; prispôsobenie štatistického modelu dátam a odhadom parametrov modelu; Diagnostika modelu; Vyvodzovanie záverov; Predpovede pomocou odhadovaného modelu.
- Závery: Dajú sa účinky treatmentu štatisticky overiť? Dokáže model dobre vysvetliť pozorované dáta? Sú potrebné ďalšie experimentálne treatmenty?

Premenné

- Aby bolo možné experimentálne testovať výskumné otázky, je potrebné vygenerovať rôzne typy premenných. Predpokladajme napríklad, že výskumná hypotéza znie, že „sumy ponúkané v ultimátnej hre sú nižšie, ak prvý hráč hrá proti počítaču namiesto človeka (a vie o tom)“. V tomto prípade je závislou premennou čiastka, ktorú ponúka prvý hráč.
- Experimentátor očakáva, že nezávislá premenná bude mať vplyv na závislú premennú, ale nie naopak. V súlade s našou výskumnou hypotézou očakávame, že binárna premenná „počítačový protivník“ (áno/nie) bude mať vplyv na ponúkané sumy. V kontrolovanom experimente sú hodnoty týchto nezávislých premenných nastavené experimentátorom. Vo vyššie uvedenom príklade experimentátor meria závislú premennú „ponúknuté sumy“ raz pod hodnotou „áno“ a raz pod hodnotou „nie“, aby bolo možné porovnanie oboch podmienok a testovanie výskumnej hypotézy. V tomto prípade sa nezávislá premenná nazýva aj intervenčná premenná, pretože jej hodnoty predstavujú „intervencie“ (treatmenty), za ktorých sa pozoruje závislá premenná.
- Ak má štúdia vyvodiť kauzálny záver o závislých a nezávislých premenných (a to je hlavným účelom kontrolovaných experimentov), je potrebné zvážiť aj iné premenné, ktoré môžu mať na výsledok vplyv. Musíme byť schopní vylúčiť, že rozdiel medzi treatmentami, ktorý nameráme, bol spôsobený inými vplyvmi, ako iba výlučne nezávislou (treatment) premennou, ktorú v experimente manipulujeme. Rôzne rušivé premenné môžu rozmazať kauzalitu medzi závislými a nezávislými premennými, ak majú „skrytý“ vplyv na závislú premennú, ktorá nie je explicitne súčasťou experimentu.
- Bohužiaľ, existujú aj rušivé premenné, ktoré nemožno kontrolovať. Sú to predovšetkým individuálne charakteristiky subjektov, napr. IQ. Samozrejme, nie všetky možné nekontrolovateľné premenné sú pre experiment relevantné, pretože mnohé nemajú žiadnu súvislosť so závislou premennou. Napriek tomu by sme mali dôkladne zvážiť, čo s vysokou pravdepodobnosťou ovplyvní našu závislú premennú a zároveň sa môže líšiť od subjektu k subjektu, pričom zostáva mimo našej kontroly.

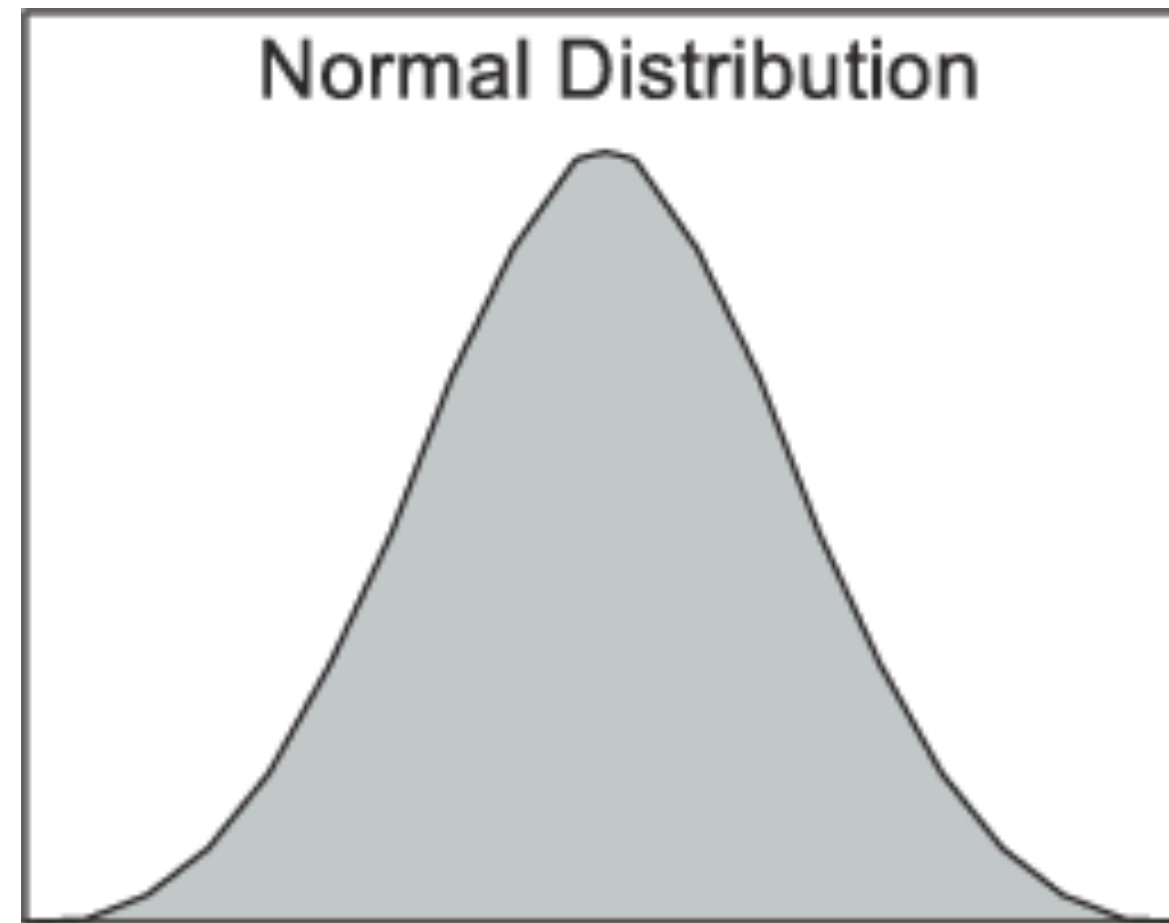
Kontrola, randomizácia a veľkosť vzorky

- Bez ohľadu na to, či je nekontrolovaná rušivá premenná merateľná alebo nie, jej vplyv na našu závislú premennú by sa mal z experimentu čo najviac odstrániť; inak už nie je možné urobiť jednoznačný kauzálny záver o vplyve intervenčnej premennej. 100% kontrola takýchto premenných je sotva možná, keďže mnohé z nich sú nielen nemerateľné, ale dokonca neznáme a ich vplyv je teda „skrytý“.
- Napriek tomu existuje jednoduchý štatistický trik, ktorý dokáže zmierniť ich dopad. Základnou myšlienkou je vytvoriť skupiny subjektov, v ktorých sú možné rušivé premenné rozdelené čo najrovnomernejšie. Robí sa to náhodným zaradením každého subjektu do jednej zo skupín (randomizácia). V tomto procese by sa malo zabezpečiť, aby skupiny pozostávali z dostatočne veľkého počtu nezávislých subjektov.
- Celkovo vzaté, v laboratórnom experimente je ústrednou premennou závislá premenná. Zmeny v tejto premennej sú spôsobené vplyvom vysvetľujúcich premenných a rôznych rušivých premenných. Ak sa má pozorovaná zmena v závislej premennej pripísať zmene vo vysvetľujúcej premennej vyvolanej experimentátorom, treba zvážiť tri najdôležitejšie pojmy:
 - Kontrola (všetky nežiaduce vplyvy, ktoré možno udržiavať konštantné, by sa mali udržiavať konštantné);
 - Randomizácia (vytváranie skupín, ktoré sú v priemere homogénne - ponechaním na náhodu, ktorý subjekt je zaradený do ktorej skupiny);
 - Veľkosť vzorky (alebo replikácia) - zabezpečiť dostatočný počet nezávislých pozorovaní pre každý treatment, t.j. dostatočne veľké skupiny subjektov, ktoré systematicky nevykazujú rovnaké správanie.

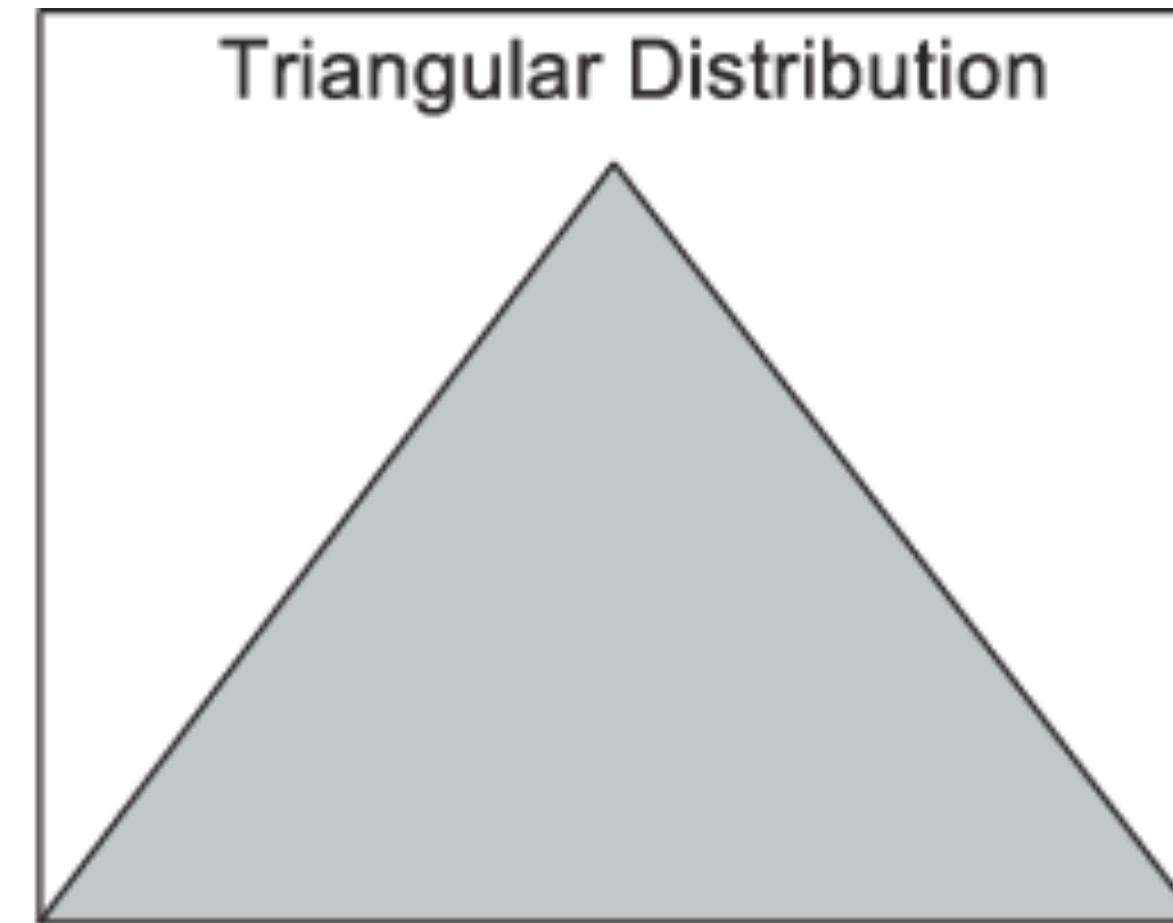
Náhodné premenné a ich distribúcia

- Pri štatistickom modelovaní vzťahu medzi premennými sa závislá premenná interpretuje ako náhodná premenná. Hodnoty náhodnej premennej ktoré sú najpravdepodobnejšie a tie, ktoré sú menej pravdepodobné, určuje jej rozdelenie. Funkcia hustoty diskkrétnej náhodnej premennej udáva pravdepodobnosť, s akou nastane určitá hodnota. Napríklad výsledky hodu kockou sú rovnomerne rozdelené, každý s príslušnou pravdepodobnosťou $1/6$.
- V prípade spojitej náhodnej premennej, ako je dĺžka času rozhodnutia, nie je možné určiť pravdepodobnosť individuálnej hodnoty. Ak existuje nekonečný počet hodnôt, pravdepodobnosť jednej hodnoty musí byť nekonečne blízka nule. Z tohto dôvodu je pri spojitých premenných možné indikovať špecifické pravdepodobnosti len pre rozsahy hodnôt, pričom celková plocha pod funkciou hustoty je vždy 1. Kumulatívna (spojitá) distribučná funkcia je, matematicky povedané, integrálom spojitej funkcie hustoty. Hodnota funkcie v bode x teda udáva pravdepodobnosť, s ktorou náhodná premenná nadobudne hodnotu menšiu alebo rovnú x .
- Väčšina štatistických rozdelení má určité parametre, ktoré určujú tvar funkcie hustoty. Tri najdôležitejšie parametre sú očakávaná hodnota, rozptyl a stupne voľnosti. Očakávaná hodnota je priemerom všetkých hodnôt, ak náhodnú vzorku (teoreticky) vytiahneme nekonečne veľa krát. Napríklad, keďže existuje rovnaká pravdepodobnosť hodu každého čísla na (normálnej) kocke, očakávaná hodnota je $1/6 \cdot (1 + 2 + 3 + 4 + 5 + 6) = 3,5$. Rozptyl je stredná kvadratická odchýlka všetkých realizácií očakávanej hodnoty a teda predstavuje informáciu o “rozhodení” náhodnej premennej. Čím je rozptyl väčší, tým je funkcia hustoty širšia a viac plochá.
- Matkou všetkých distribúcií je normálne rozdelenie. Jeho parametrami sú očakávaná hodnota μ a rozptyl σ^2 . Hustota pravdepodobnosti je zvonovitá a symetrická okolo μ , kde má najvyššiu funkčnú hodnotu hustoty. Iné dôležité rozdelenia nie sú parametrizované priamo pomocou očakávanej hodnoty a rozptylu, ale aj nepriamo pomocou takzvaných stupňov voľnosti, ktoré ovplyvňujú očakávanú hodnotu a/alebo rozptyl. Napríklad (Studentovo) t-rozdelenie má takéto stupne voľnosti, pričom tvar funkcie hustoty sa stále viac a viac približuje funkcii hustoty štandardného normálneho rozdelenia so zvyšujúcimi sa stupňami voľnosti.

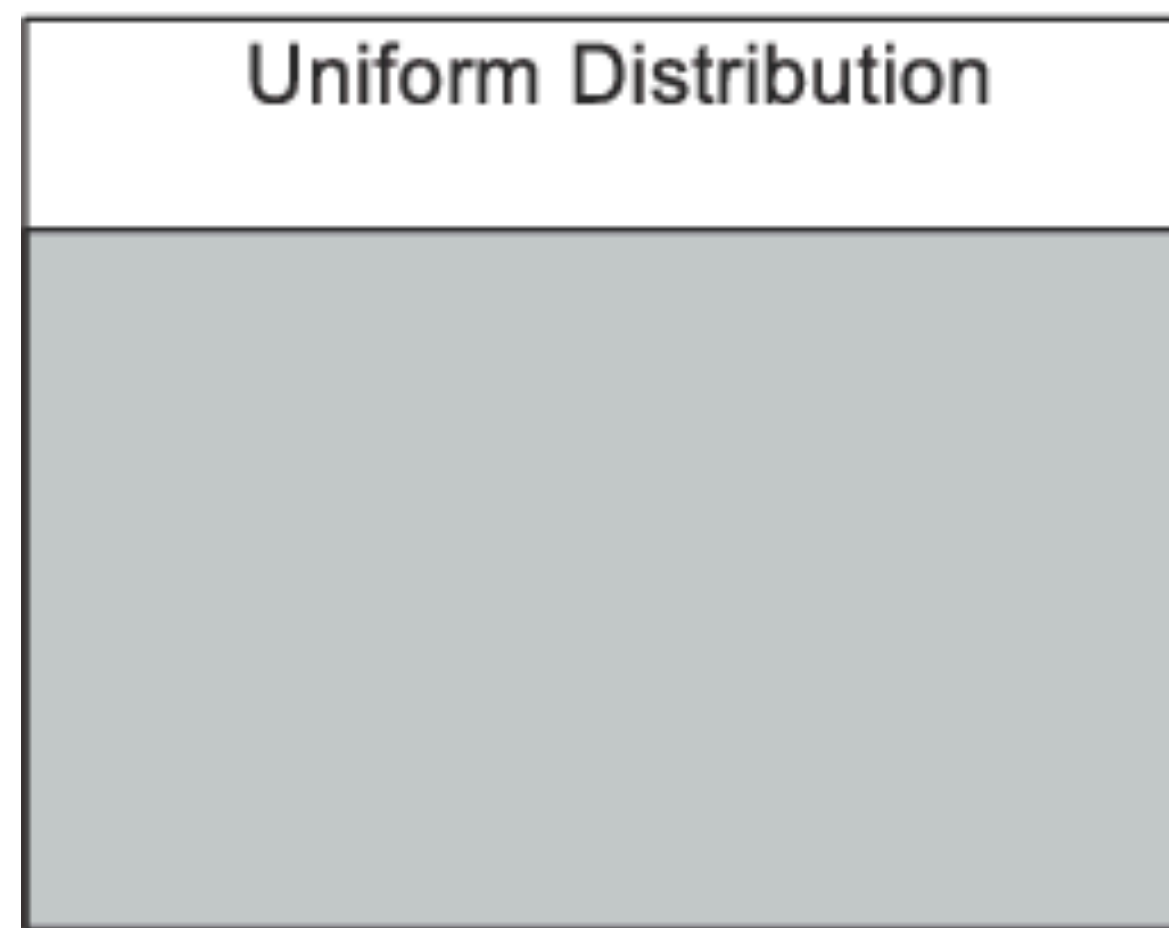
Náhodné premenné a ich distribúcia



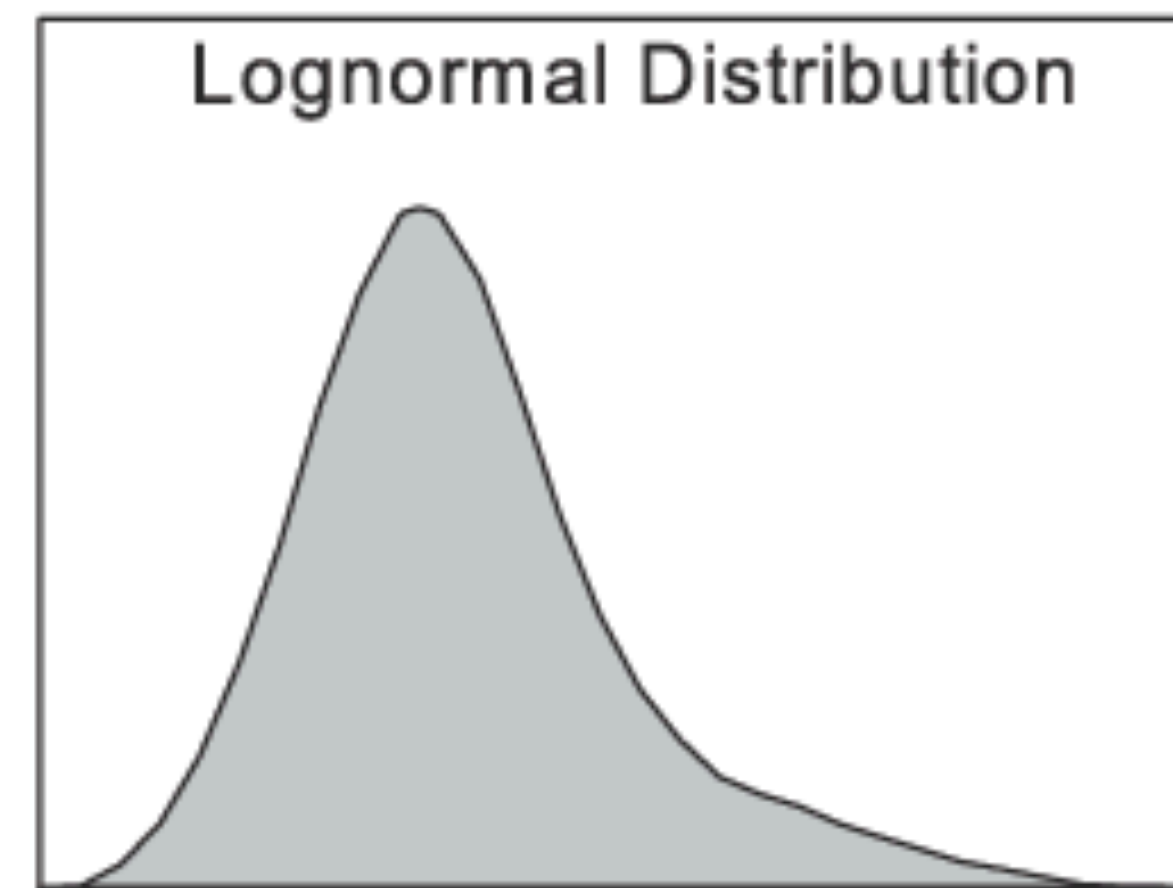
Normal distribution



Triangular distribution



Uniform distribution



Log-normal distribution

Tvorba štatistického designu

- Výber určitého počtu subjektov z celkovej populácie sa v štatistike označuje ako vzorkovanie. Je potrebné sa zamyslieť nad veľkosťou vzorky, t. j. otázkou „Koľko subjektov vyberiem zo špecifikovanej populácie?“ Žiaľ, v experimentálnej praxi sa na túto otázku často odpovedá len na základe rozpočtu, podľa hesla: „Jednoducho berieme toľko subjektov, koľko dokážeme zaplatiť, bez ohľadu na to, či je toto číslo dosť veľké alebo malé“. Napríklad v neurovedách sú laboratórne časy extrémne drahé, takže veľkosti vzoriek sú (často musia byť) v jednocifernom rozsahu.
- Takéto malé vzorky sú však problematické najmä z pohľadu inferenčnej štatistiky. Pravdepodobnosť, že test štatistickej hypotézy správne identifikuje skutočný efekt ako prítomný (toto sa nazýva sila testu), drasticky klesá s menšími vzorkami. Inými slovami, aj keď v skutočnosti existuje v populácii relatívne silný a vedecky relevantný efekt, bude prinajlepšom rozpoznatelný ako „náhodný artefakt“ a nie ako štatisticky významný efekt. Na druhej strane existuje aj „príliš veľká“ vzorka, pretože veľké vzorky môžu spôsobiť, že testy štatistických hypotéz budú príliš citlivé. To znamená, že aj tie najmenšie, možno vedecky nevýznamné účinky sa stávajú štatisticky významnými.
- Štatistická významnosť by sa preto nemala zamieňať s vedeckou významnosťou. V závislosti od veľkosti vzorky môžu byť obe úplne odlišné. Je to preto, že štatistická významnosť je silne ovplyvnená veľkosťou vzorky, zatiaľ čo skutočný účinok, ktorý sa má zistiť v populácii, nie je.
- Ak je jasné, že (dostatočne veľká) náhodná vzorka nie je dostupná ale stále je potrebná reprezentatívna vzorka, potom je vhodnou možnosťou stratifikovaný výber. Populácia sa najskôr rozdelí na subpopulácie (vrstvy), pričom subjekty v rámci každej subpopulácie majú aspoň jednu spoločnú vlastnosť, ktorá ich odlišuje od subjektov ostatných subpopulácií. Z každej vrstvy sa potom vyberie náhodná vzorka. Každá z týchto vzoriek musí tvoriť rovnaký podiel z celkového počtu všetkých vzoriek ako každá vrstva v celkovej populácii.

Tvorba štatistického designu

- Ako sa líšia experimentálne treatmenty?
- Experimentálne treatmenty je možné klasifikovať podľa počtu faktorových premenných a ich typu, ako aj podľa počtu možných hodnôt. V najjednoduchšom faktoriálnom dizajne sa mení iba jedna premenná. Ak ide o binárnu premennú len s dvoma hodnotami alebo úrovňami, hovoríme o 1×2 faktoriálnom dizajne. 1×2 faktoriálny design sa dá vyhodnotiť obzvlášť jednoducho, pretože zvyčajne stačí porovnať iba stredné hodnoty závislých premenných. V ideálnom prípade je tento rozdiel spôsobený samotným treatmentom, a preto sa nazýva (jednoduchý) efekt treatmentu. Kvantitatívny rozdiel medzi týmito dvoma hodnotami sa nazýva veľkosť efektu treatmentu. Na druhej strane, ak má faktorová premenná viac ako dve úrovne, stredné hodnoty závislej premennej môžu byť porovnávané párovo pre každé dve úrovne alebo súčasne pre všetky úrovne.
- Návrh s dvoma faktormi je podstatne zložitejší ako návrh s jedným faktorom. Napríklad, ak chceme experimentálne preskúmať, ako ovplyvňujú správanie faktory „hra proti počítaču“ (nie/áno alebo 0/1) a „experimentátor vie, kto som“ (nie/áno alebo 0/1), potom ide o 2×2 faktoriálny dizajn.
- V dizajne opakovaných meraní každý subjekt absolvuje niekoľko meraní, buď v jednom a tom istom treatmente v rôznych časoch (pozdĺžny dizajn) alebo v rôznych treatmentoch, samozrejme aj v rôznych časoch (cross-over design). Sekvencia treatmentov, ktorými subjekt prechádza, je opäť náhodná.
- Hlavným štatistickým problémom pri viacnásobných meraniach je vzájomná závislosť pozorovaní. Vo faktoriálnom dizajne 1×2 s viacerými meraniami dostaneme kontrolnú skupinu (meraní na úrovni 1) a intervenčnú skupinu (meraní na úrovni 2), ktoré spolu súvisia. Účinok meraný pomocou závislej premennej teda už nemožno jednoznačne pripísať treatmentu, pretože by to mohol byť aj časový alebo sekvenčný účinok (napr. učenie, oboznámenie sa, únava). Často nám v tomto prípade pomáha vyvažovanie (balancovanie), t.j. vytvorí sa dve homogénne skupiny a jedna skupina sa meria v poradí 1, potom 2 a druhá v poradí 2, potom 1.
- Výhodou opakovaných meraní sú nižšie náklady v dôsledku menšieho počtu subjektov, nižšie rozpätie chýb, čo vedie k vyššej štatistickej sile ako pri porovnateľných between-subject dizajnoch, a možnosť merania efektov v čase (dynamika). Nevýhody takéhoto návrhu spočívajú v tom, že zahŕňa podstatne zložitejšie metódy analýzy v dôsledku závislosti pozorovaní a slabších príčinných súvislostí v dôsledku sekvencie, času a efektov prenosu.

Popis dát

- Najlepším spôsobom, ako sa naučiť písať o dátach, je prečítať si dátové sekcie v iných článkoch a venovať sa druhu informácií, ktoré obsahujú. Vaša sekcia s dátami by mala obsahovať minimálne toto:
- Identifikujte zdroj dát. Zahrňte vetu, ktorá explicitne hovorí, odkiaľ vaše dáta pochádzajú.
- Popíšte dáta. Mali by ste napríklad uviesť počet pozorovaní, počet vzoriek, časové obdobie, počas ktorého sa dáta zhromažďovali, spôsob zberu dát atď.
- Uvedte silné a slabé stránky dát. Aké sú vaše dáta v porovnaní s inými dátami použitými v literatúre? Poskytuje váš výskum viac pozorovaní alebo novších pozorovaní než iné výskumy? Boli údaje zhromaždené spoľahlivejším spôsobom? Prečo sú tieto dáta pre vašu štúdiu obzvlášť vhodné (alebo nie)? Všimnite si všetky vlastnosti dát, ktoré môžu ovplyvniť vaše výsledky. Boli určité skupiny populácie nadmerne alebo nedostatočne zastúpené? Existuje skreslenie spôsobené prirodzeným úbytkom (attrition) alebo samo-výberom (selection bias)? Zmenil sa počas zhromažďovania dát spôsob merania?
- Vysvetlite všetky vykonané výpočty alebo úpravy. Dáta vám niekedy niečo neposkytú priamo; možno ste museli pridať / odčítať / vynásobiť / rozdeliť dva zadané údaje, aby ste získali tretí. Popíšte, ako ste vybrali svoju vzorku. Museli ste napríklad vylúčiť určité druhy pozorovaní?

Deskriptívne štatistiky

- Dátové sekcie v článkoch často obsahujú tabuľku deskriptívnych štatistík - teda štatistík ktoré určujú relevantnosť vzorky. Zvyčajne zahŕňajú priemer (napr. priemerný príjem, priemerný vek, priemerné roky školskej dochádzky atď.) a štandardnú odchýlku. V prípade kategorických údajov (napríklad rasy) však neuvádzate priemer; namiesto toho použijete percento pozorovaní v každej skupine.
- Očakávaná hodnota - Priemerná hodnota atribútu vzorky, založená na opakovanom výbere vzoriek z populácie.
- Štandardné chyby - smerodajná odchýlka alebo miera variability / rozptylu vo vzorke. Čím väčšia je vzorka, tým menšia je štandardná chyba.
- Distribúcie vo vzorke - Teoretické (nepozorované) rozdelenie atribútu, ktoré umožňuje výpočet intervalov spoľahlivosti a testy hypotéz.
- POZNÁMKA: Priemer a štandardná odchýlka fungujú dobre pre normálne rozdelenie (v tvare zvonovej krivky). Ak natrafíme na inú distribúciu, môže byť užitočnejšie použiť na opis centrálnej tendencie (očakávanej hodnoty) medián alebo modus.

Grafické znázornenie dát

- Dobre zostavený graf dokáže odpovedať na niekoľko otázok naraz:
- Centrálna tendencia: Kde leží stred distribúcie?
- Rozptyl alebo variácia: Ako veľmi sú pozorovania rozťahnuté alebo koncentrované?
- Tvar distribúcie: Má distribúcia iba jediný vrchol (jedna koncentrácia pozorovaní v relatívne úzkom rozmedzí hodnôt), alebo má vrcholov viac?
- Chvosty: Približne aké percento pozorovaní leží na koncoch (chvostoch) distribúcie?
- Symetria alebo asymetria (nazývaná tiež nesúmernosť): Majú pozorovania tendenciu hromadiť sa na jednej strane distribúcie, zatiaľ čo na druhej ich je relatívne málo? Alebo má každá strana distribúcie zhruba rovnaký počet pozorovaní?
- Odľahlé hodnoty (outliers): Existujú hodnoty, ktoré sa v porovnaní s väčšinou javia ako veľmi veľké alebo veľmi malé?
- Porovnanie: Ako sa dve distribúcie líšia z hľadiska tvaru, natiahnutia a centrálnej tendencie?
- Vzťahy: Je možné, že hodnoty jednej premennej súvisia s hodnotami inej?

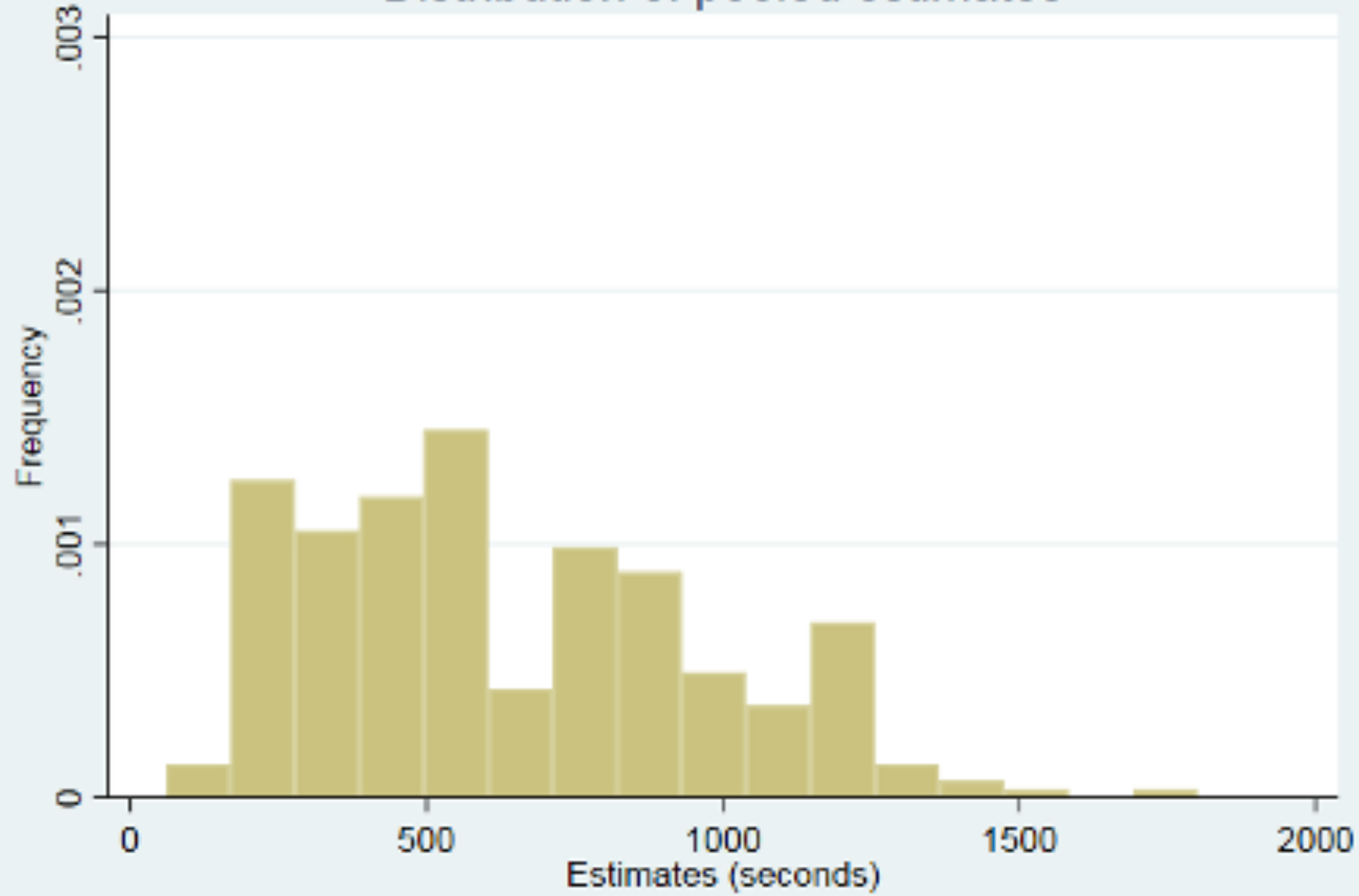
Výber správneho grafu

TABLE 11-10 Typical Presentation and Exploratory Graphs

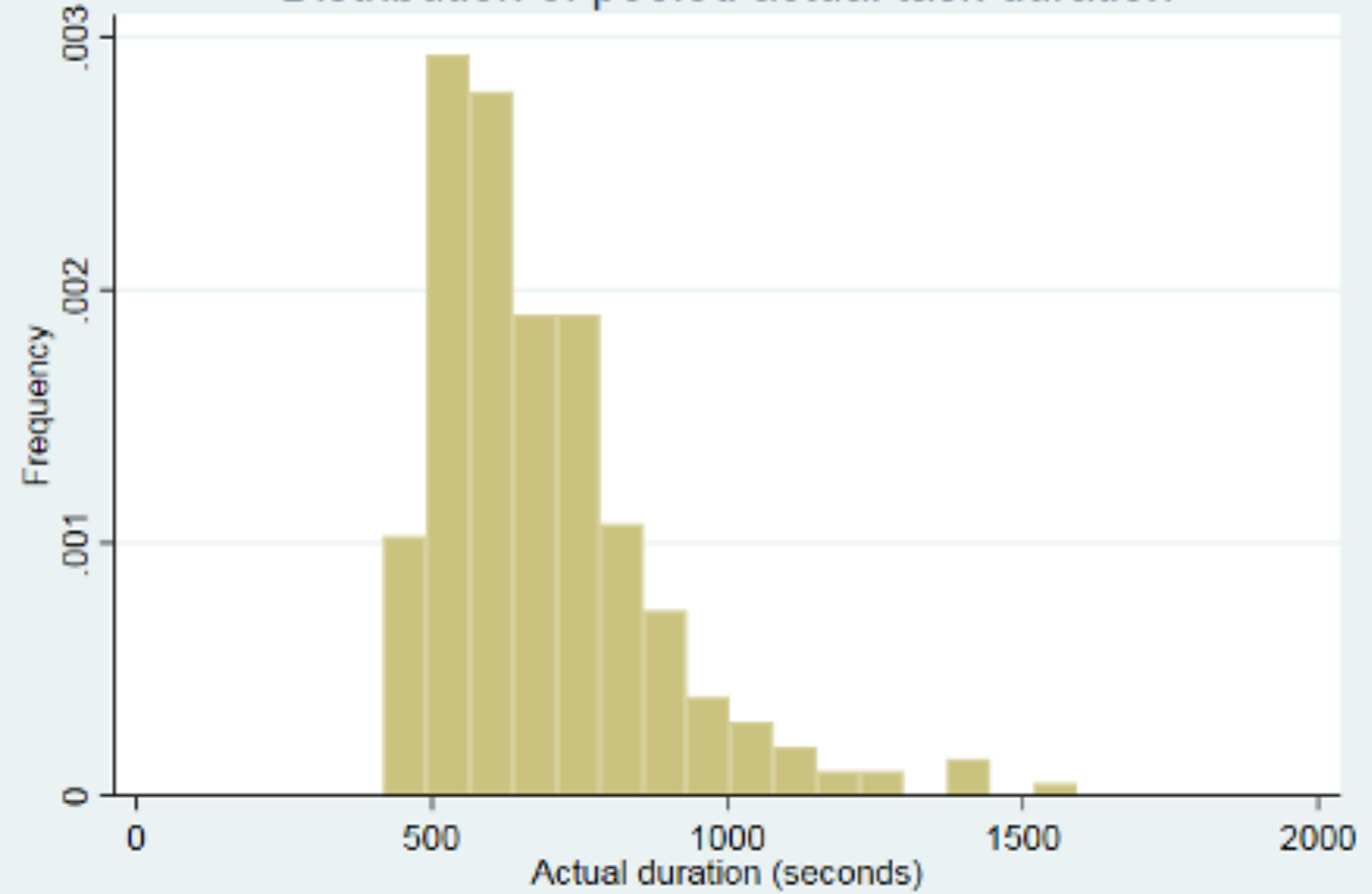
Type of Graph	What Is Displayed	Most Appropriate Level of Measurement	Number of Cases	Comments
Bar chart	Relative frequencies (percentages, proportions)	Categorical (nominal, ordinal)	3-10 categories	Common presentation graphic
Dot chart	Frequencies, distribution shape, outliers	Quantitative (interval, ratio)	<i>Less than 50 cases</i>	Displays actual data values
Histogram	Distribution shape	Quantitative	$N > 50$ cases	Essential exploratory graph for interval or ratio variables with a large number of cases
Boxplot	Distribution shape, summary statistics, outliers	Quantitative	$N > 50$ cases	Can display several distributions; actual data points, an essential exploratory tool
Time series plot	Trends	Quantitative (percentages, rates)	$10 < N < 100$	Common in presentation and exploratory graphics

Histogram

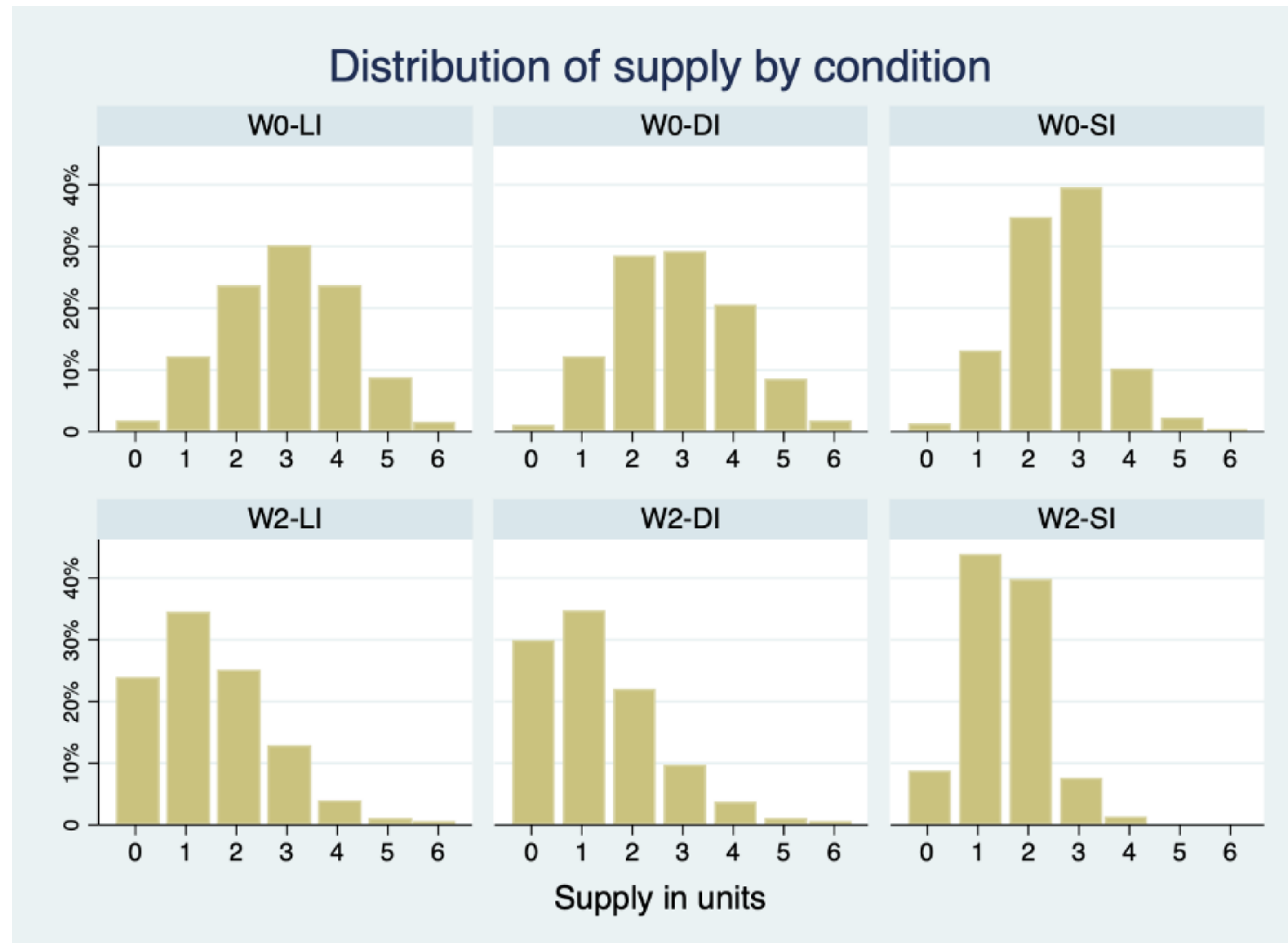
Distribution of pooled estimates



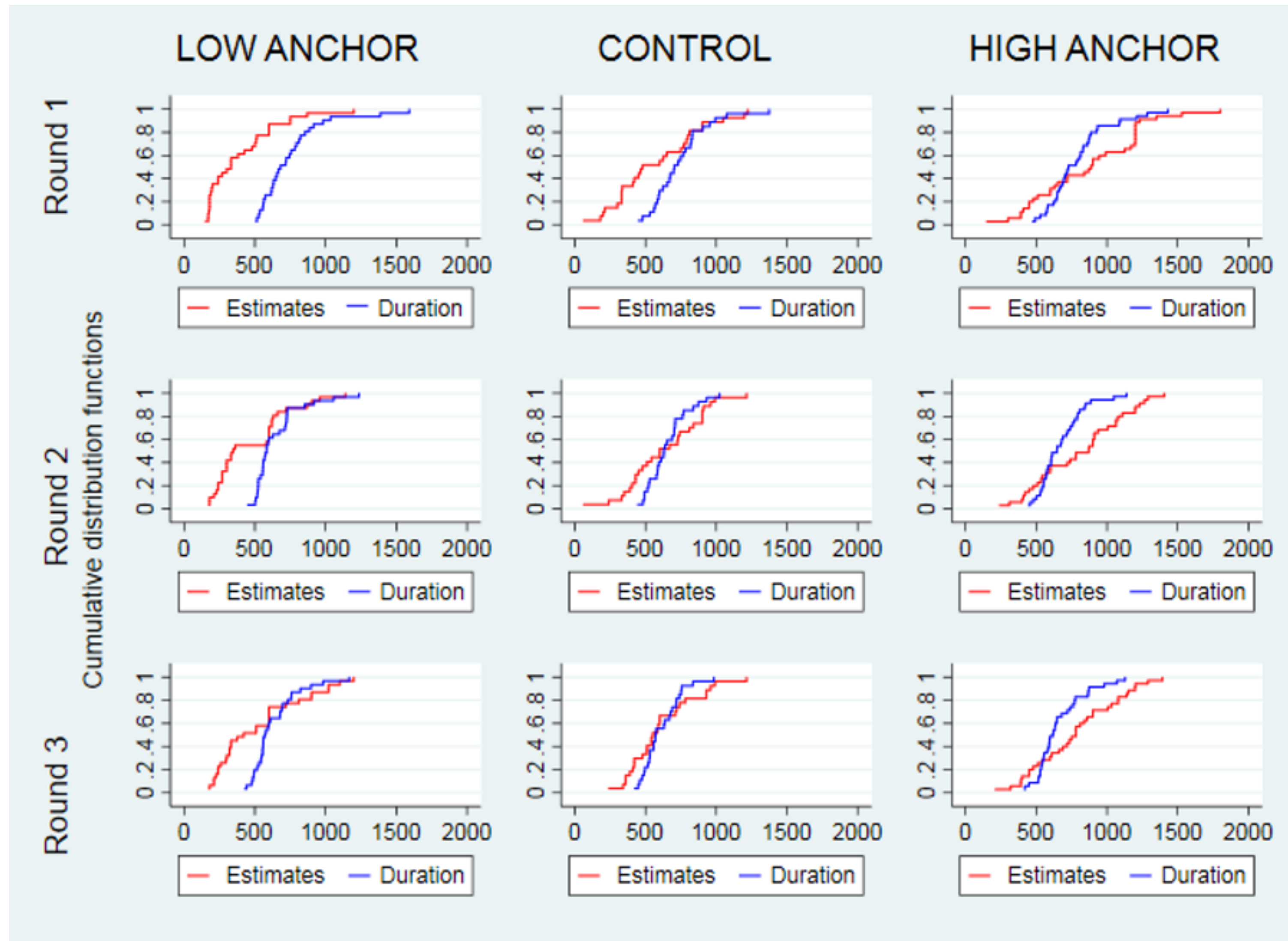
Distribution of pooled actual task duration



Histogram

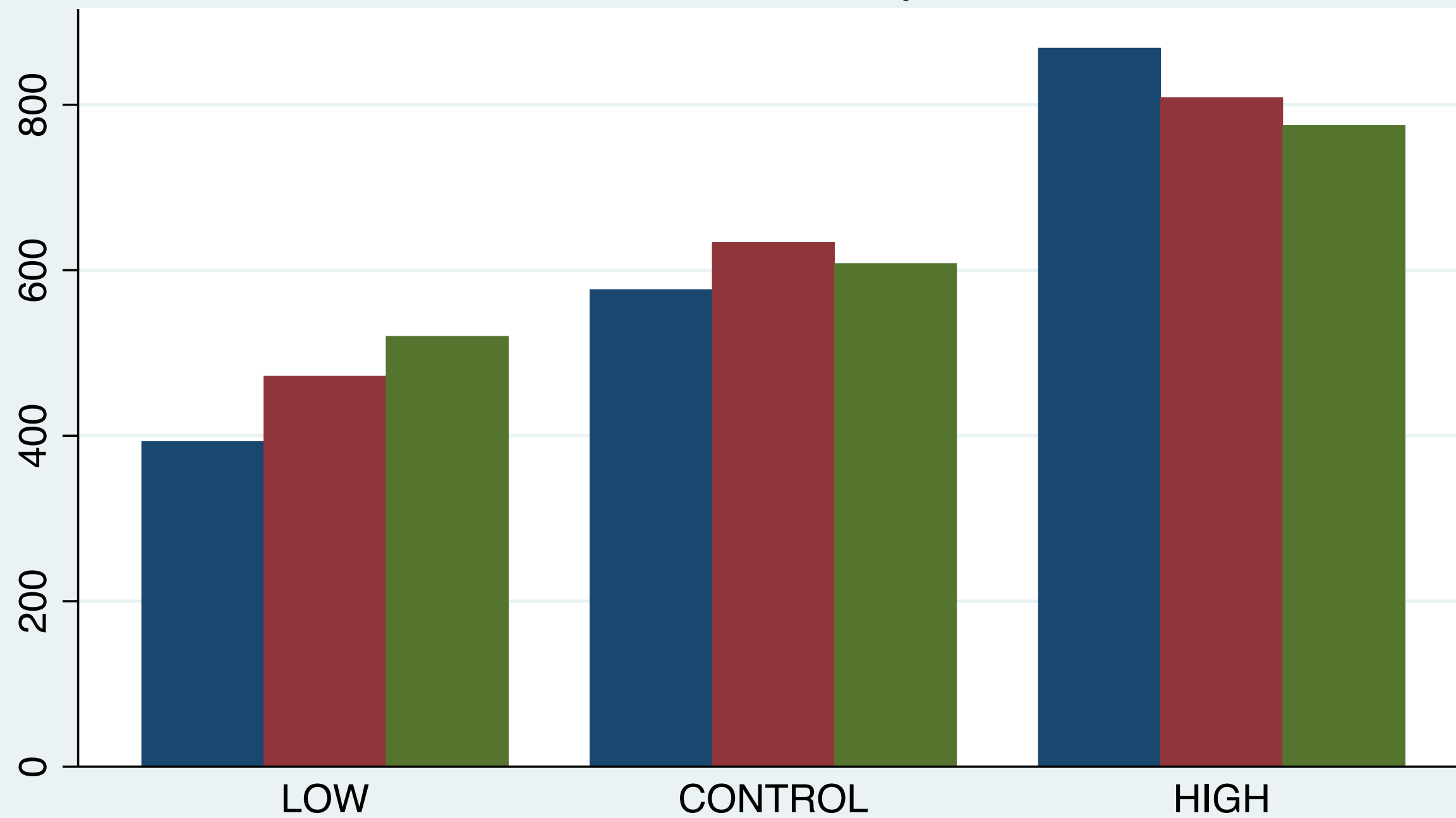


Kumulatívna distribúcia



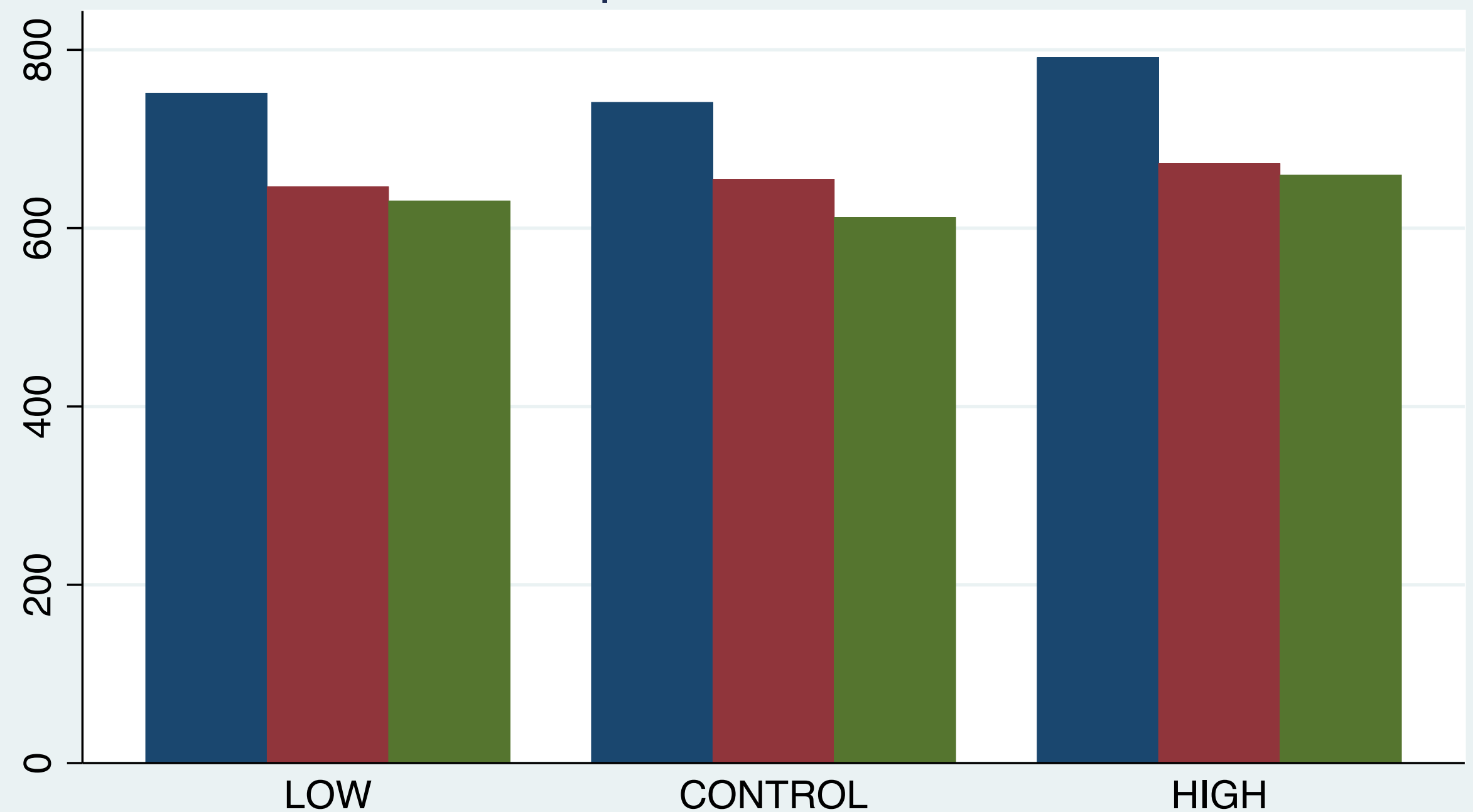
Stípcový graf

Estimates of task completion time



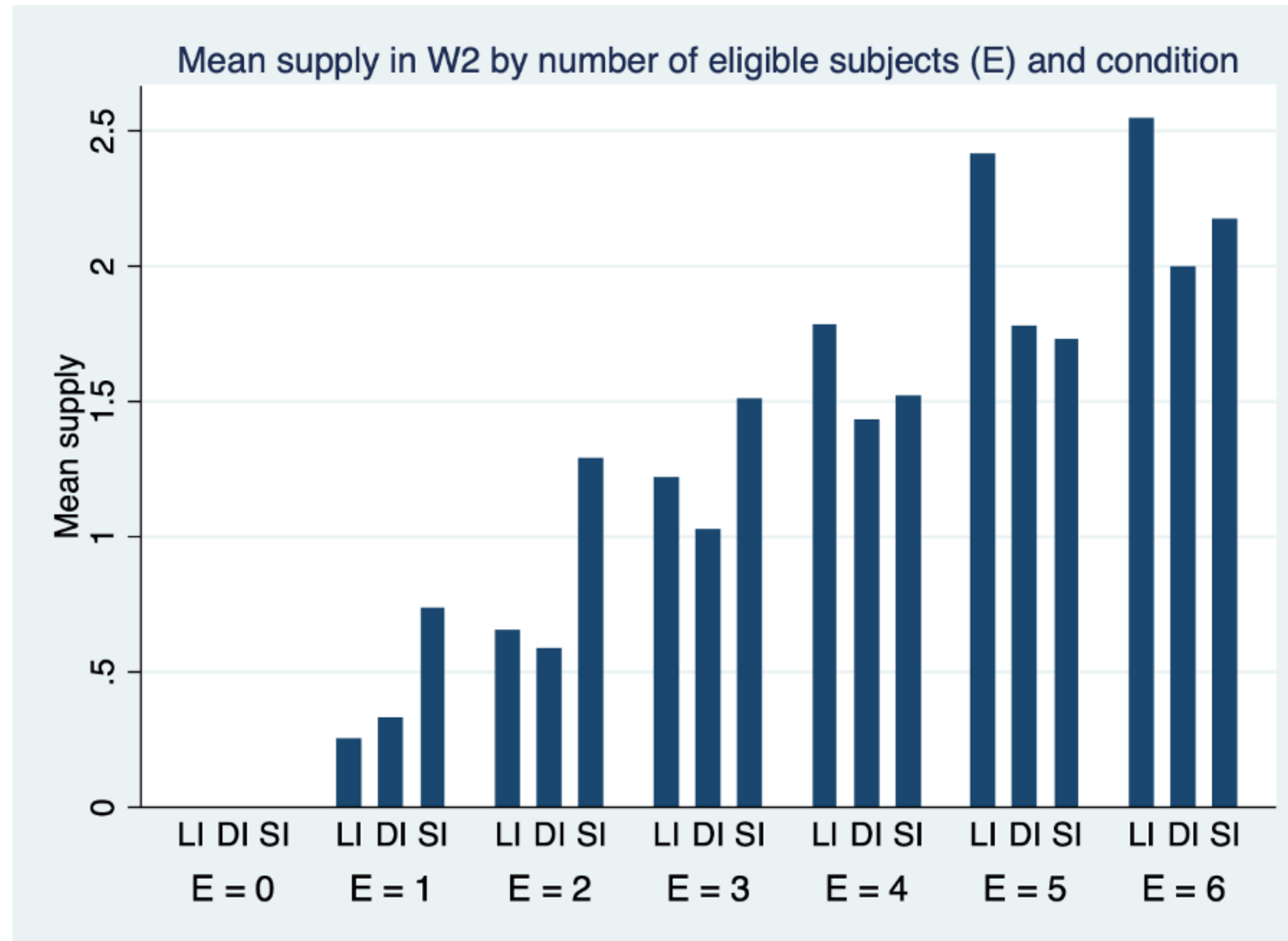
Mean of estimates Round 1 Mean of estimates Round 2
Mean of estimates Round 3

Real completion times in seconds

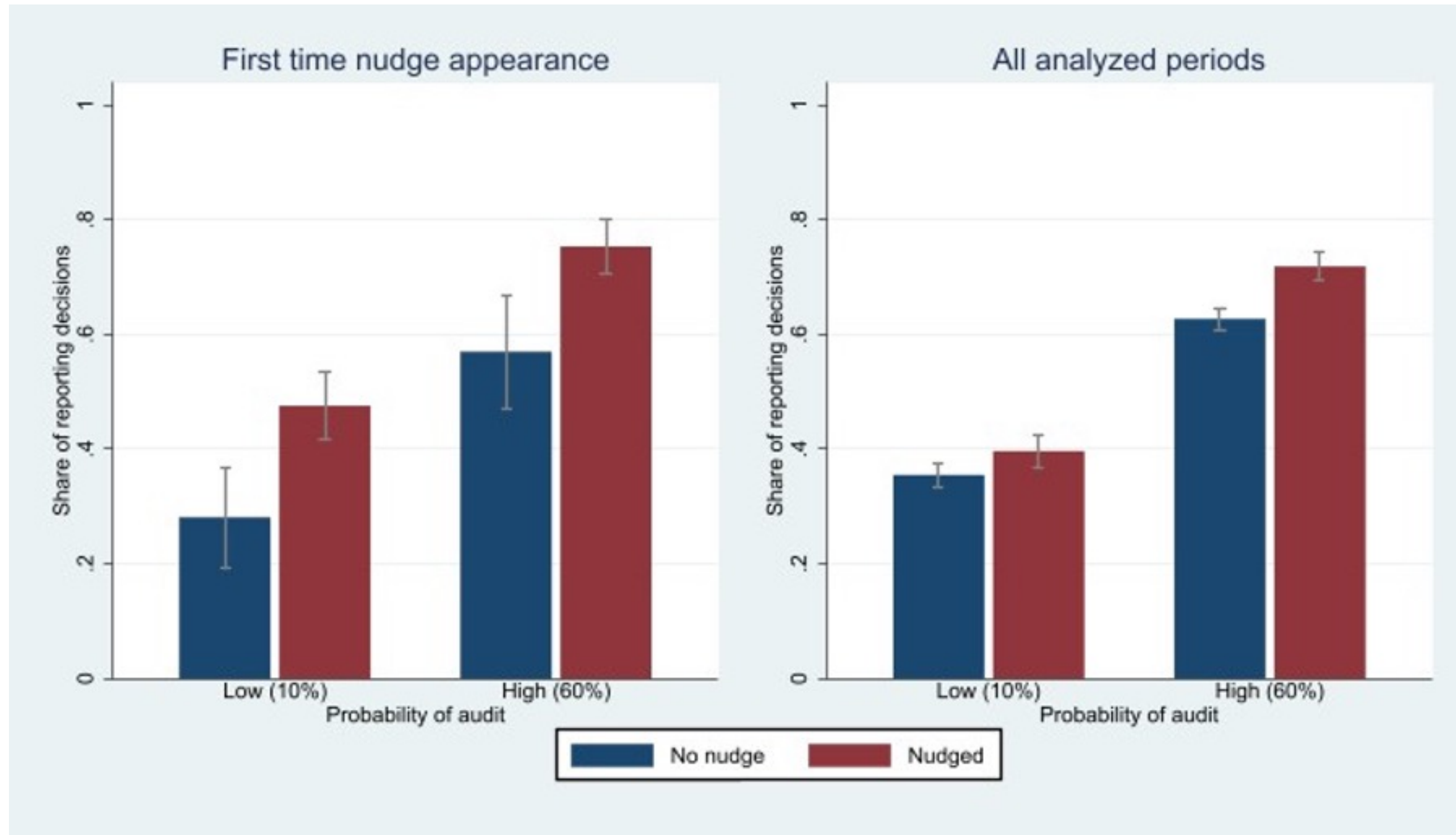


Mean of completion time R1 Mean of completion time R2
Mean of completion time R3

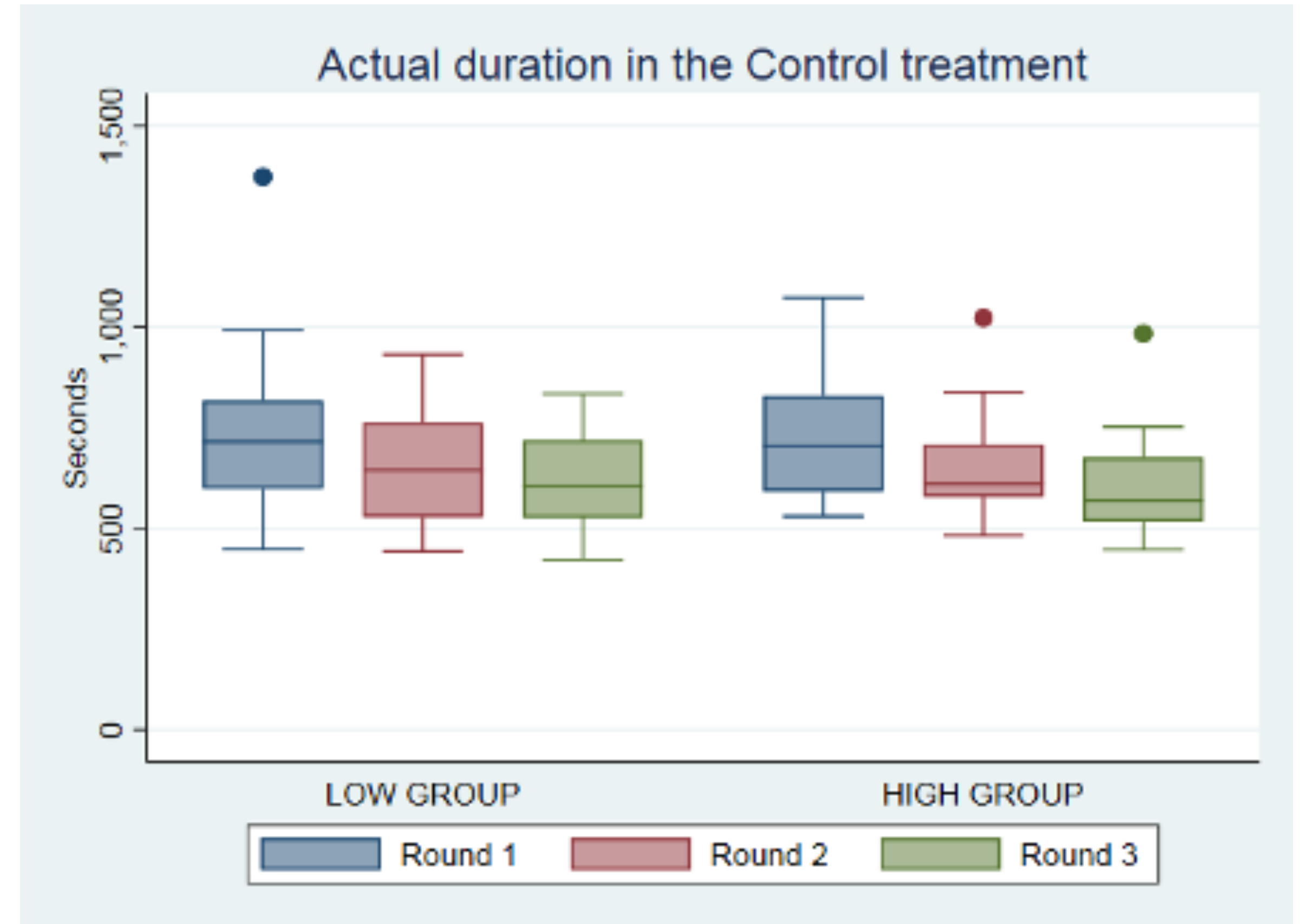
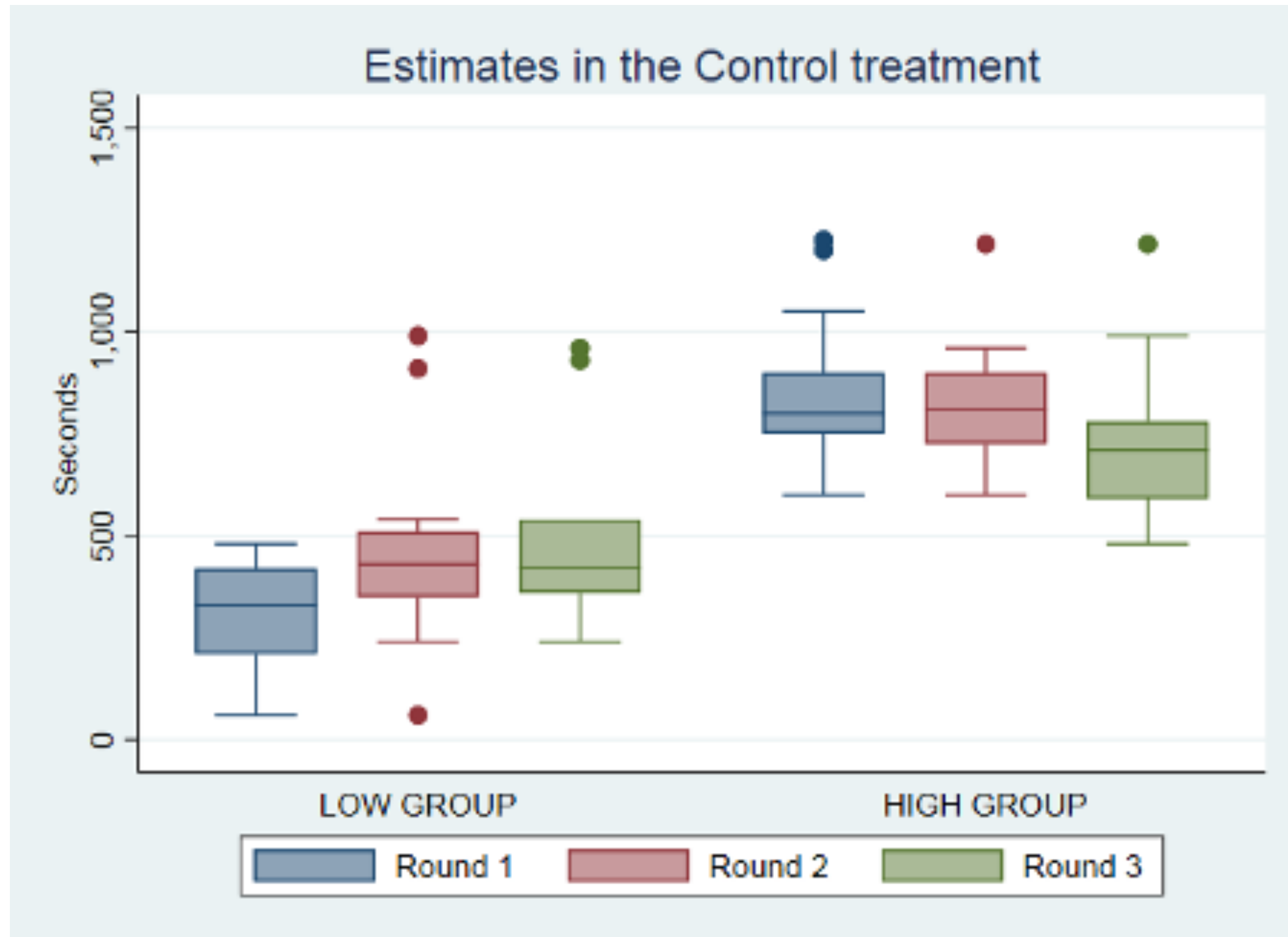
Stípcový graf



Stípcový graf

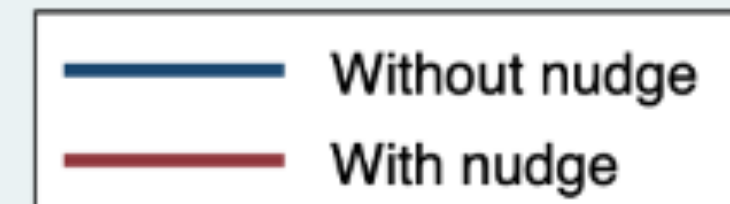
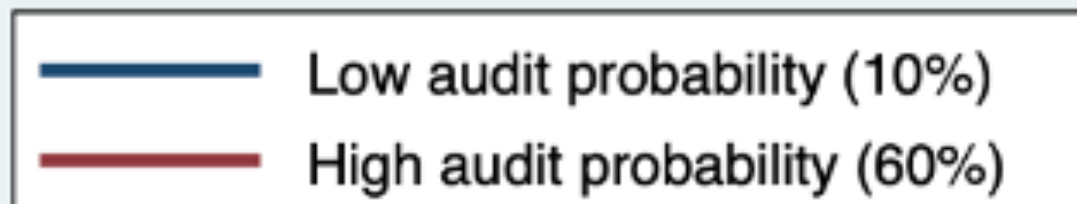
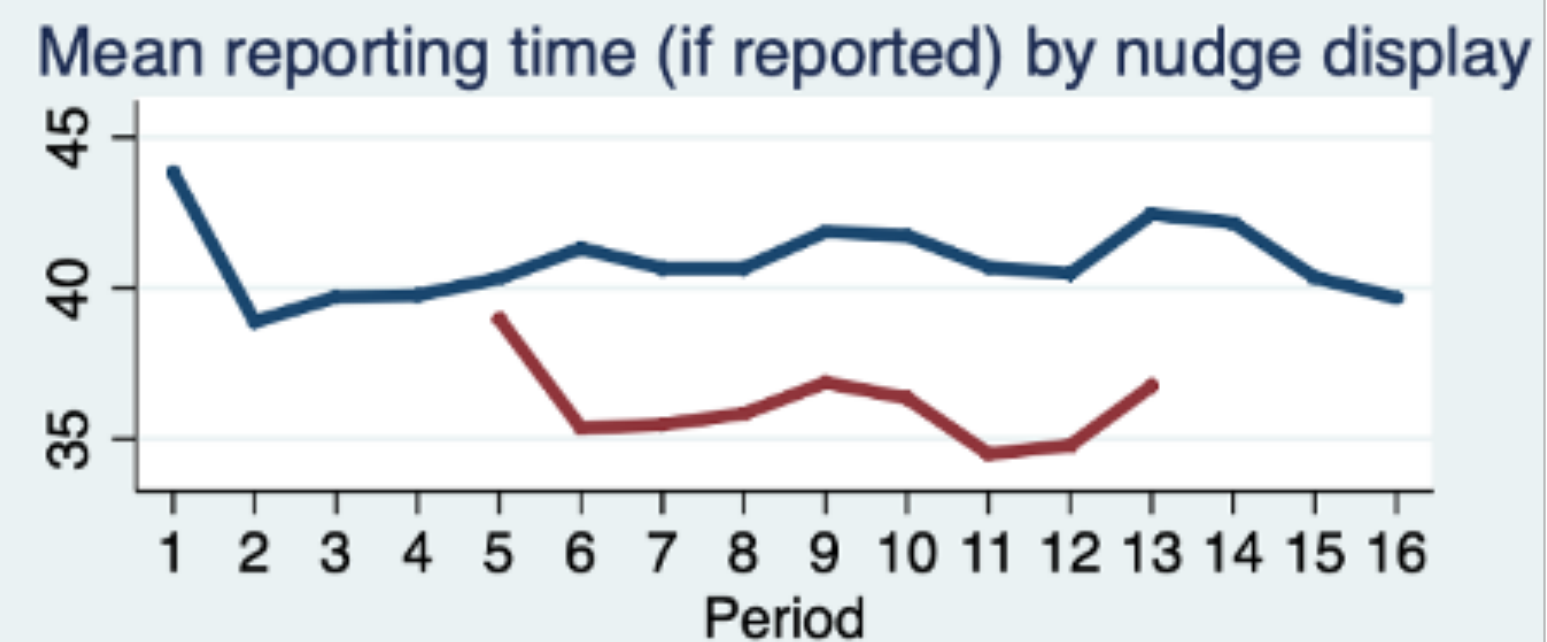
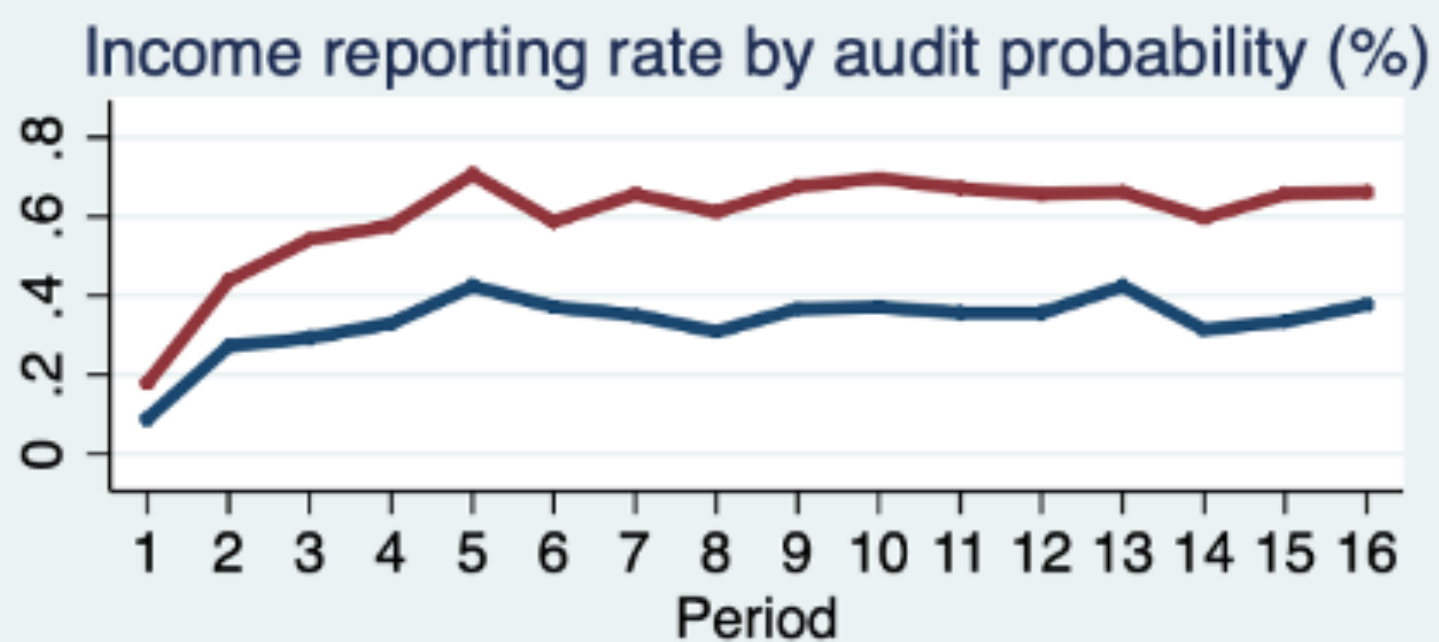
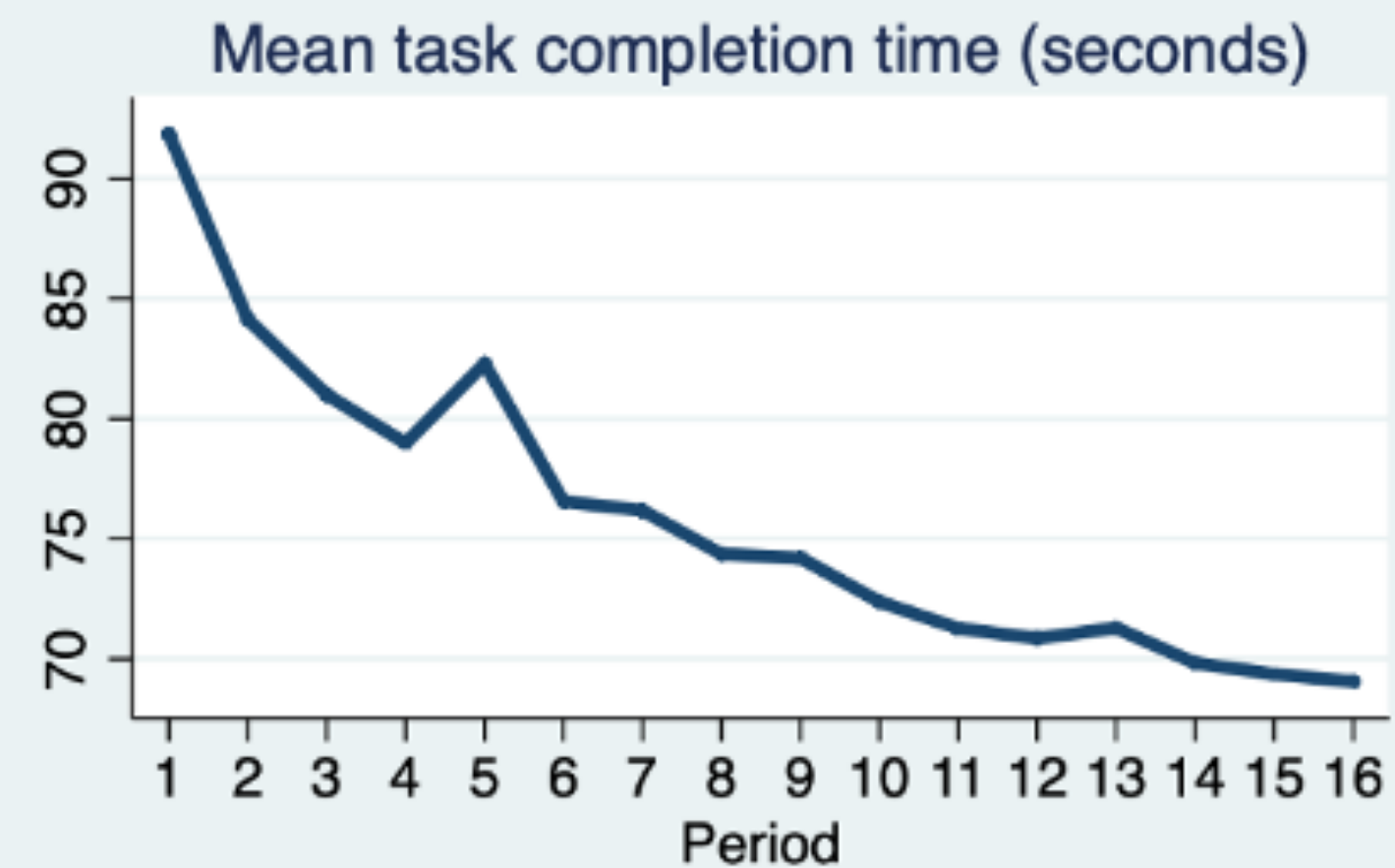


Krabicový graf

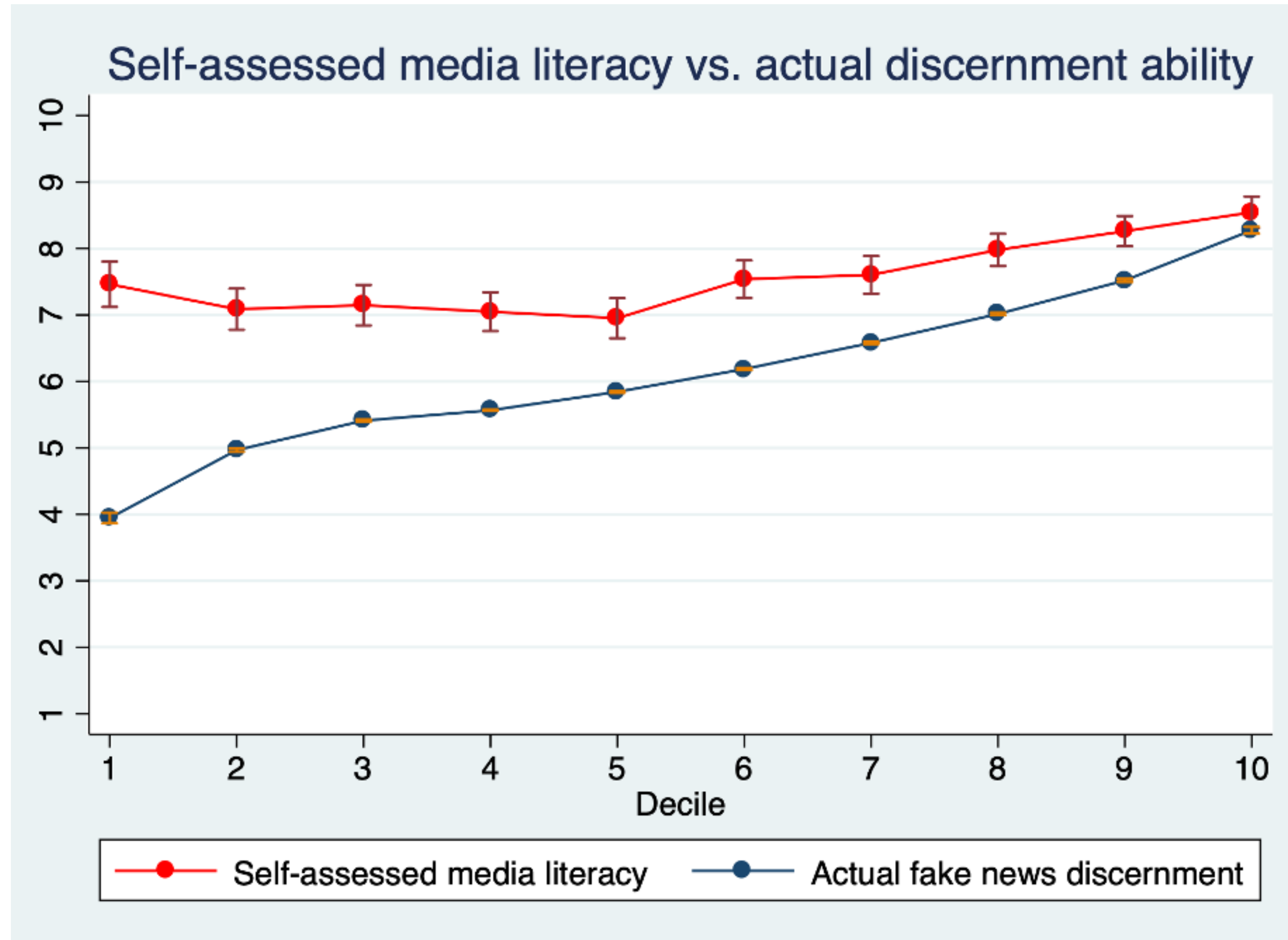


Spojnicový graf

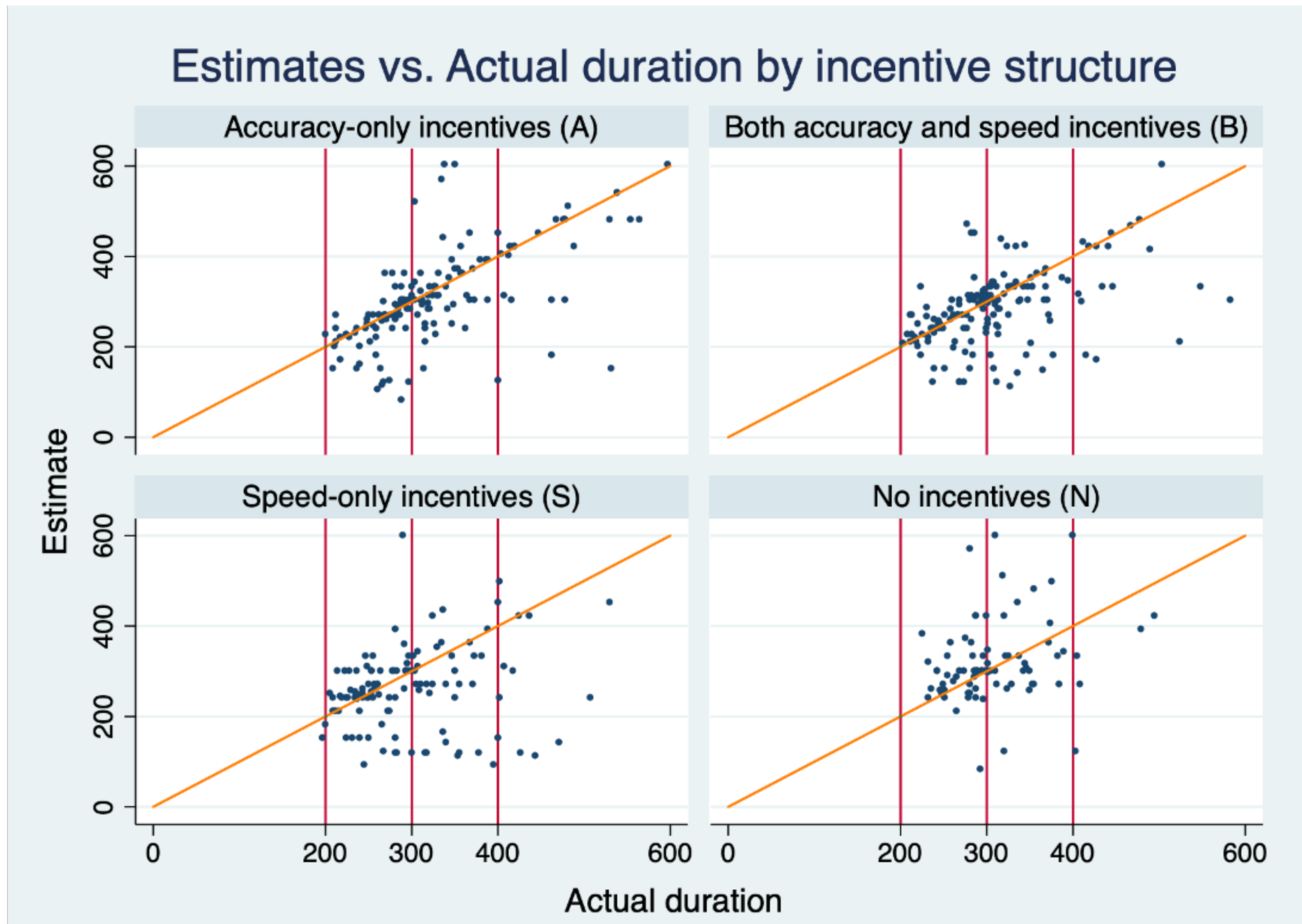
Task performance and income reporting by period



Spojnicový graf

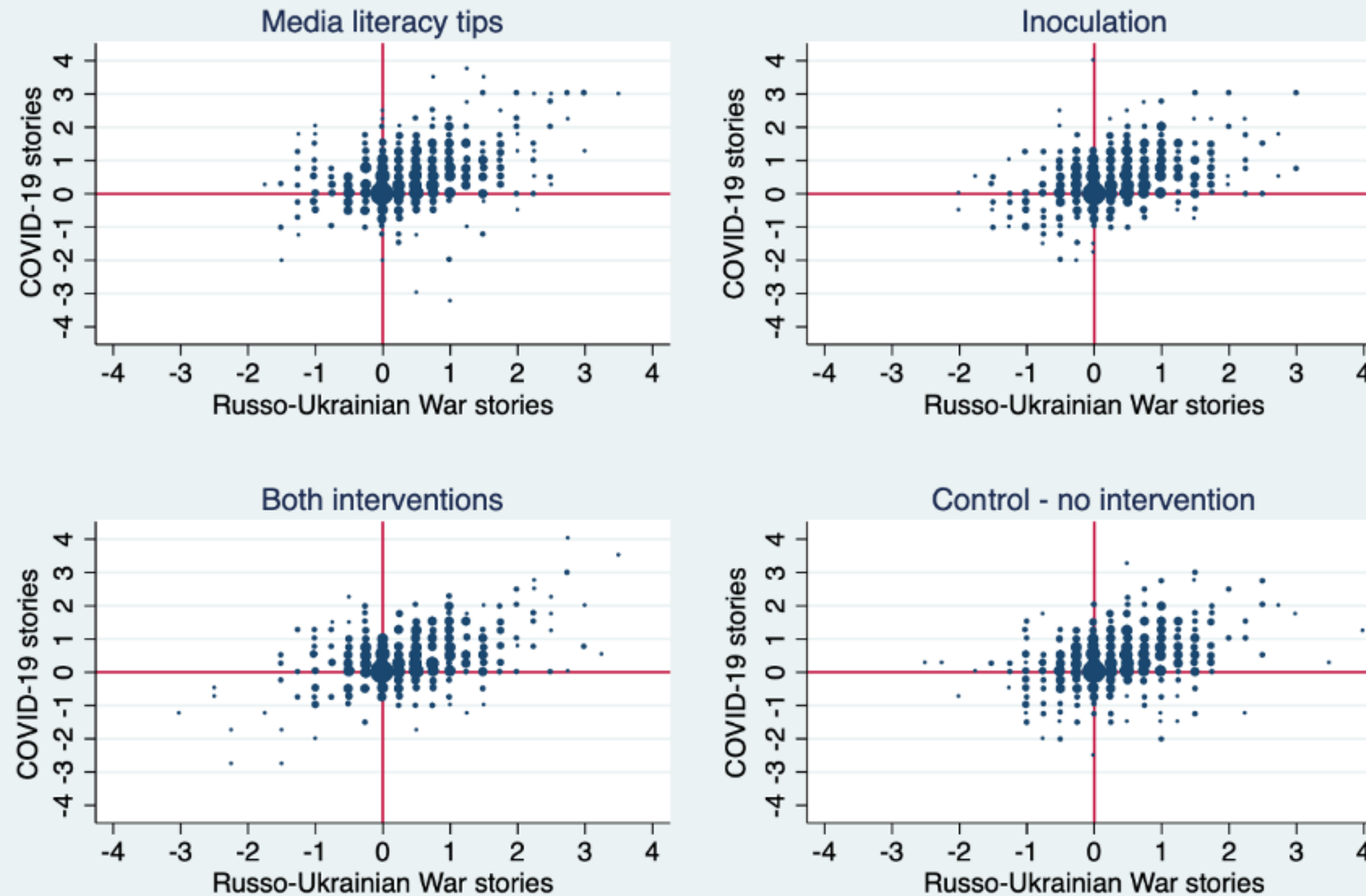


Bodový diagram



Bodový diagram

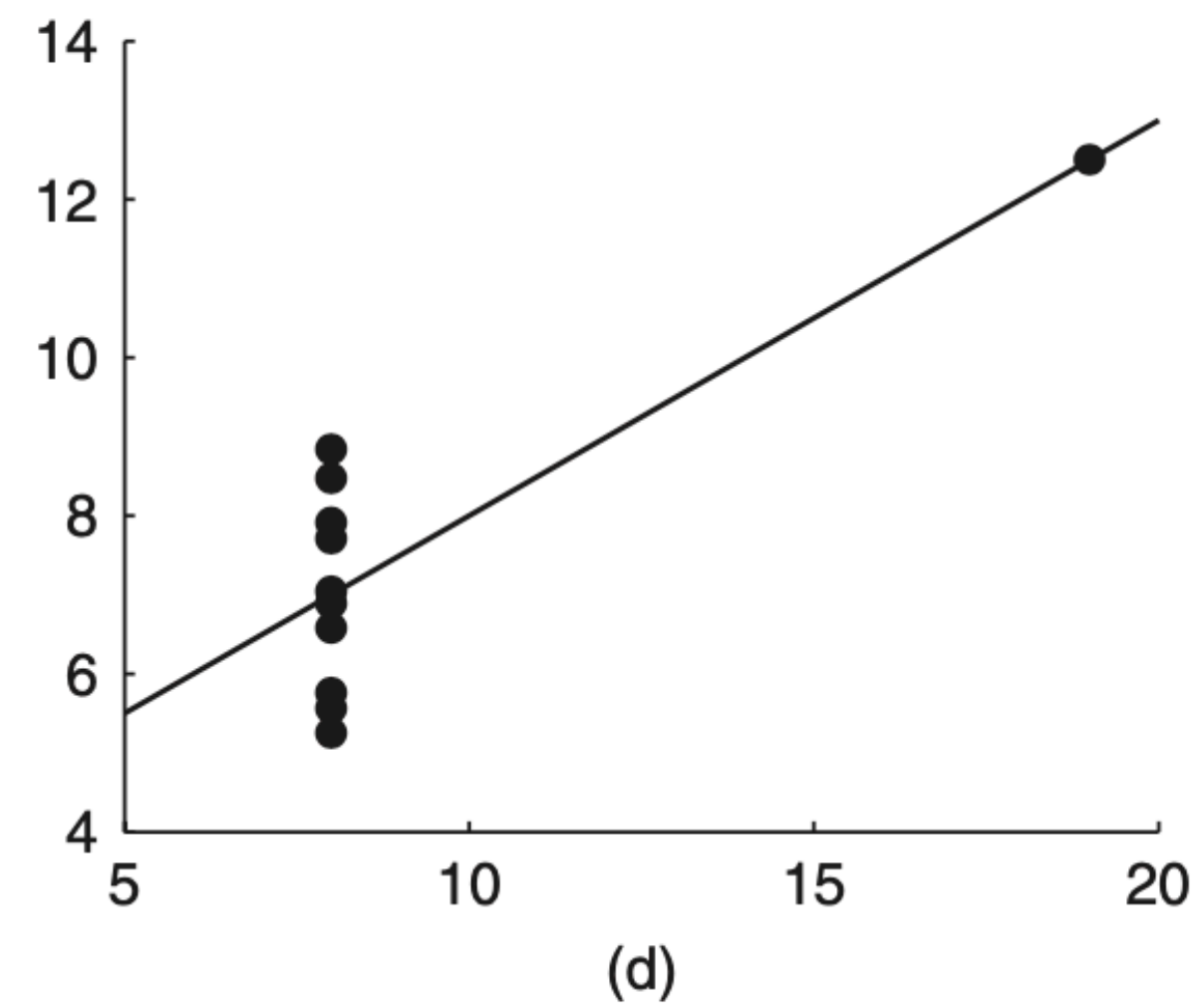
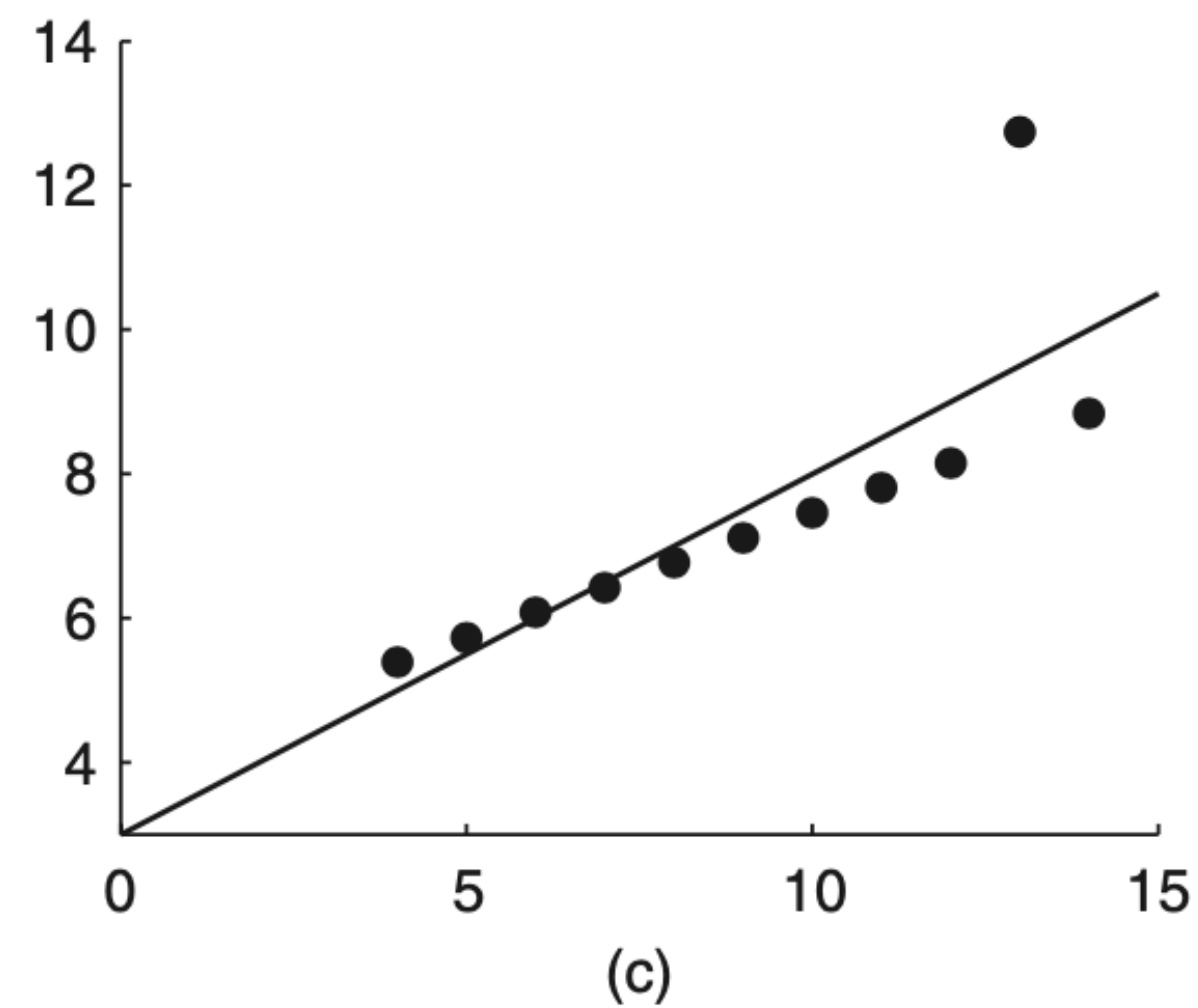
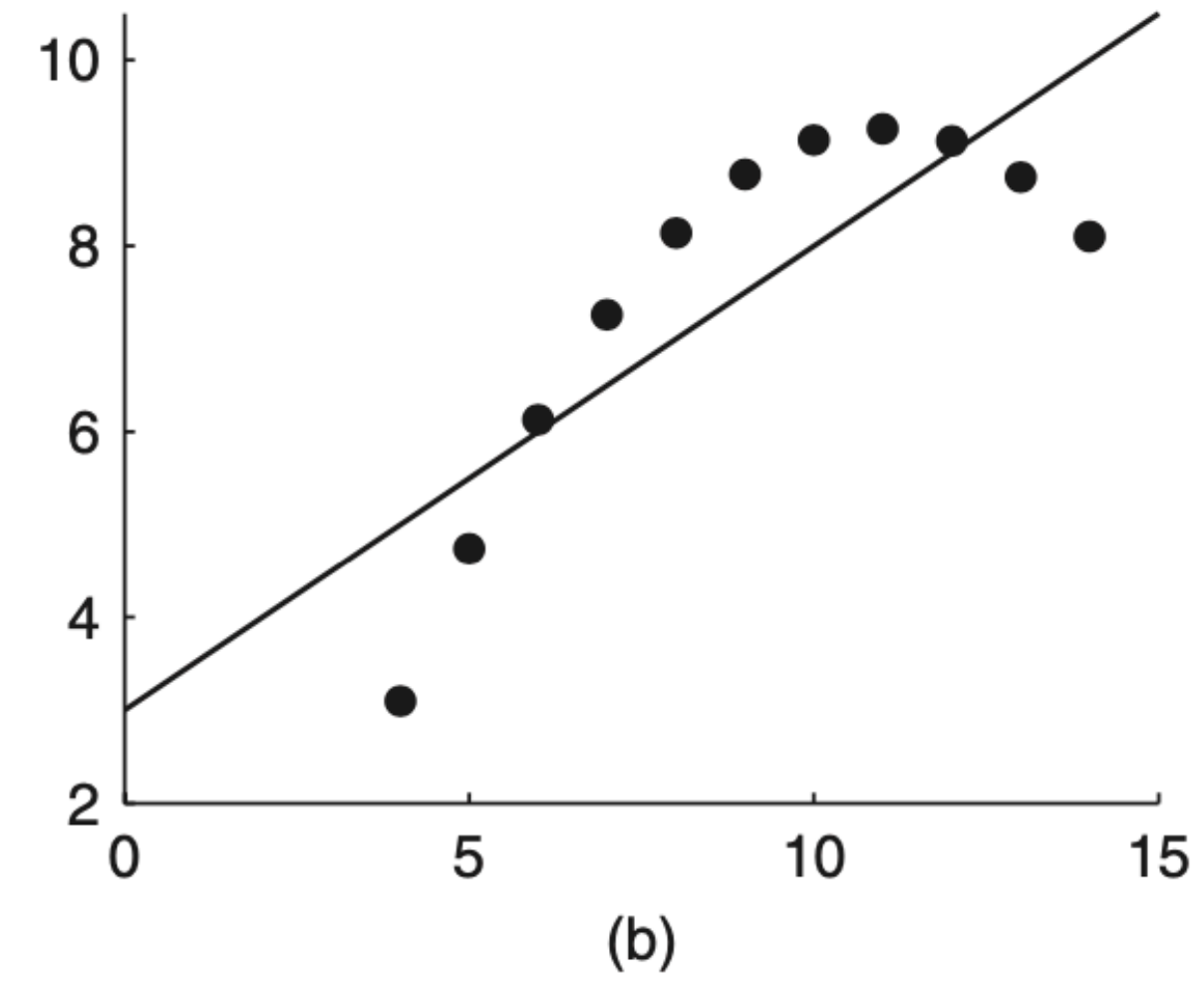
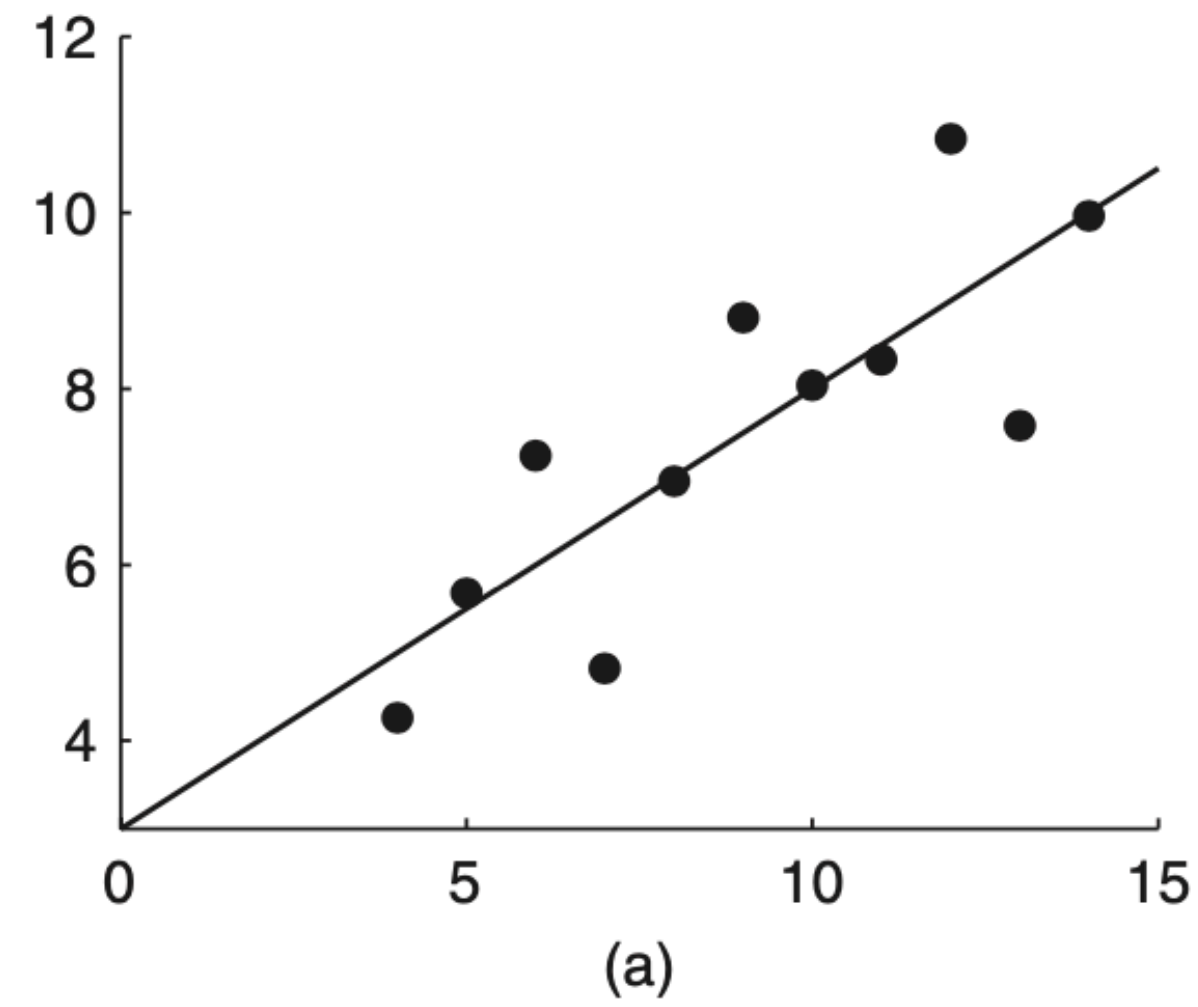
Fake news discernment by intervention



Pearsonova korelácia

- Korelácia meria stupeň asociácie medzi dvoma premennými. Základným nástrojom je tzv. Pearsonova korelácia, ktorá meria stupeň linearity dvojrozmerného vzťahu medzi dvoma premennými.
- Pearsonova korelácia je štandardizovaná kovariancia (t. j. pomer kovariancie medzi dvoma premennými k súčinu ich rozptylov). Pearsonova korelácia je ohraničená medzi -1 (dokonalý negatívny lineárny vzťah) a 1 (dokonalý pozitívny lineárny vzťah).
- Pearsonovu koreláciu neovplyvňujú zmeny v rozsahu (napr. vynásobenie premenných 2) ani v umiestnení (napr. pridanie konštanty) premenných. Dve úplne nezávislé premenné majú korelačný koeficient rovný 0. Zistenie nulovej korelácie medzi dvoma premennými však možno interpretovať ako indikáciu nezávislosti len v špeciálnych prípadoch, napr. pri dvojrozmernom normálnom rozdelení. Použitie Pearsonovej korelácie by tu bolo zavádzajúce vzhľadom na jej základný predpoklad linearity.
- Obrázok na ďalšom slajde predstavuje Anscombovo kvarteto, sériu štyroch súborov údajov, ktoré zostavil Anscombe (1973), aby zdôraznil význam vizualizácie údajov pred vykonaním štatistickej analýzy. Každý súbor údajov pozostáva z 11 dvojíc bodov s Pearsonovou koreláciou 0,816. Priemer prvej premennej (na osi x) je vždy 9, pričom výberový rozptyl je 11. Priemer druhej premennej je približne 7,50, pričom výberový rozptyl je od 4,122 do 4,127. Pearsonova korelácia má zmysel len v paneli a), pre normálne rozdelené údaje s lineárnym vzťahom. V paneli b) je vzťah jednoznačne nelineárny. Panely c) a d) ukazujú, ako je Pearsonova korelácia citlivá na odľahlé hodnoty. V paneli c) je bez odľahlých hodnôt korelácia 1. Stačí jedna jediná odľahlá hodnota, aby sa znížila na 0,816. V paneli d) je korelácia bez odľahlej hodnoty 0, ale stačí jedna odľahlá hodnota, aby sa zvýšila na 0,816.

Pearsonova korelácia



Korelácia vs. Kauzálnosť

- Silu asociácie medzi dvoma premennými možno posudzovať z kvantitatívneho a kvalitatívneho hľadiska. Z kvalitatívneho hľadiska je najsilnejší vzťah kauzálny. To znamená, že hodnota jednej premennej spôsobuje zmenu hodnoty druhej premennej. Napríklad sila prenášaná z nohy na futbalovú loptu je jednou z príčin, prečo lopta letí tak ďaleko.
- Korelácia medzi premennými je kvalitatívne slabšia forma vzťahu a existuje vtedy, keď možno len pozorovať, že zvýšenie jednej premennej je sprevádzané zvýšením alebo znížením druhej premennej. Ani dokonalá korelácia nemusí nevyhnutne znamenať, že premenné sú aj kauzálne prepojené.
- Napríklad možno pozorovať, že množstvo vlasov, ktoré majú muži na hlave, a ich príslušný príjem sú navzájom inverzné, t. j. čím menej vlasov majú muži, tým vyšší je ich príjem. Ak by tu existoval kauzálny vzťah, všetci muži by si pravdepodobne oholili vlasy v nádeji, že sa stanú bohatšími. Skutočný kauzálny vzťah sa dá ľahko zistiť, ak sa zahrnie tretia premenná, vek.
- Čím je muž starší, tým má viac pracovných skúseností, a teda aj vyšší priemerný príjem. Zároveň je v povahe veci, že vypadávanie vlasov u mužov súvisí aj s vekom. Vek má teda príčinný vplyv na množstvo vlasov aj na priemerný príjem pracujúcich mužov. Príčinná súvislosť vždy znamená koreláciu, ale nie každá korelácia znamená príčinnú súvislosť. Inak povedané, ak dve premenné nie sú korelované, nemôže existovať ani kauzálny vzťah. Avšak aj keď neexistuje kauzálny vzťah, môže existovať korelácia.
- Jednoduché štatistické modely len merali silu vzťahu, a preto poskytujú čisto kvantitatívne informácie o vzťahu medzi premennými. Vzťah kvantifikovaný štatistickým modelom je vždy kauzálny len do tej miery, do akej to umožnil vopred vykonaný experimentálny design. Pre kauzalitu v experimente sú rozhodujúce tri faktory: kontrola, opakovanie a randomizácia. Experimenty, v ktorých nie je možná randomizácia, sa nazývajú kvázi-experimenty. V takýchto experimentoch je oveľa ťažšie odvodiť kauzálne vzťahy, ale existujú špeciálne štatistické modely a metódy odhadu, ktoré uľahčujú určenie kauzality (regresné diskontinuitné dizajny). Patrí k nim najmä odhad inštrumentálnych premenných a metóda rozdielov v rozdieloch (DiD).