

Metody měření vlivu nepodmíněného trestu odnětí svobody na recidivu (aneb přiřazujeme pachatelům tresty náhodně)

Mgr. Jakub Drápal, MPhil.
doktorand, Právnická fakulta, Univerzita Karlova¹

Abstrakt: Mnohé účely trestání počítají s efektivitou vězení (převýchova, inkapacitace či odrazení). Existuje paleta metod, jak tuto efektivitu zjišťovat. Tento článek popisuje experimentální i kvaziexperimentální metody, které lze použít k analýze efektivity uvěznění, jejich výhody i nedostatky a rozebírá, které metody lze použít v České republice. Významná část je věnována popisu eticky přijatelného náhodného přiřazování trestů vybraným pachatelům, jako jediné plně spolehlivé metodě, kterou lze zjistit efektivitu nepodmíněného trestu odnětí svobody. Dále je v článku ukázáno, že trestní systém v České republice naplňuje podmínky pro realizaci přirozeného experimentu, který vzniká, když je větší množství případů náhodně přiřazeno různě přísným soudcům. Článek je zakončen popisem kvaziexperimentálních metod, metod použitelných v důsledku amnestie z roku 2013 a zdůrazněním obecných limitací empirických právních výzkumů.

Summary: Imprisonment is justified by various aims of punishment that try to influence the punished person (e.g. re-education, deterrence or incapacitation). Several approaches are available to study the effectiveness of these aims. This article describes not only different experimental and quasi-experimental methods, their advantages and disadvantages, but also to what extent they can be used in the Czech Republic. The author first suggests an ethical and practical way how to randomly assign sanctions to offenders in order to discover the effectiveness of imprisonment. Further, a natural experiment that originates when cases are randomly allocated to differently punitive judges is described and the author presents evidence that such natural experiment emerges in the Czech Republic. Lastly, the article discusses quasi-experimental methods and natural experiment and regression discontinuity design resulting from the amnesty of 2013 and emphasizes the limits of empirical legal research.

Úvod

Na výzkumnou otázku, jaký je vliv nepodmíněného trestu odnětí svobody na recidivu, se pokoušely odpovědět tisíce výzkumů, a přesto nemáme plně spolehlivou odpověď (Villettaz et al. 2015). Čím to je a jak bychom mohli ideálně změřit vliv nepodmíněného trestu odnětí na svobodu na recidivu? Obecným pravidlem ve společenských vědách je, že pokud se snažíme zjistit efektivitu určité intervence, tak musíme porovnat minimálně dvě skupiny, které jsou stejné ve všech aspektech vyjma zkoumané intervence. Pouze pak můžeme mluvit o kauzálním vlivu uloženého trestu na recidivu. V tomto článku rozeberu různé přístupy k zjišťování efektivity určitého trestu, přičemž se zaměřím na ty metody, které jsou obecně považovány za nejspolehlivější. Cílem tohoto článku je kriticky přistoupit k používaným metodám a blíže rozebrat, jak by se daly nejlépe použít v České republice s přihlédnutím k dostupným datům a praktikám.

¹ Děkuji Zuzaně Podané, Marii Vanduchové a Liboru Duškovi za nápady a připomínky k tomuto tématu a rukopisu. Rukopis vznikl za podpory Grantové agentury UK, v rámci projektu 'Analýza efektivity nepodmíněného trestu odnětí svobody z hlediska recidivy', č. 34218. Možnost kontaktu: jakub.drapal@gmail.com.

Nejdříve se věnuji úvahám nad významem experimentálních a kontrolních skupin při realizovaném výzkumu, a to konkrétně tomu, jak různé způsoby jejich vytváření ovlivňují spolehlivost jejich výsledků. Následně popisuji jednotlivé metody a to od těch nejspolehlivějších po ty méně spolehlivé: V úvodu navrhuji, jak by bylo možné a eticky přijatelné uskutečnit experiment, v rámci kterého by soudci náhodně přiřazovali jednotlivé pachatele k odlišným trestům. Dále rozebírám přirozený experiment, který vzniká v situaci, kdy jsou na soudech náhodně přiřazované případy v rámci jednotlivých specializací rozhodovány odlišně přísnými soudci. Poté se věnuji kvaziexperimentálním metodám, které simulují experiment; konkrétně rozebírám dva typy vyvažování a regresi. Na závěr se věnuji specifické situaci, kterou vytvořila amnestie z roku 2013 - v jejím důsledku vznikl nejen přirozený experiment, ale vznikla i situace, ve které je možné využít tzv. regresní diskontinuitu (termín je vysvětlen níže).

Před tím, než se ponořím do výkladu, je na místě poznámka k stylu tohoto článku a k zvolené terminologii. Pokusil jsem se napsat tento text tak, aby mu rozuměl právník nedisponující rozsáhlými znalostmi společenskovední metodologie. Nešermuji proto i takovými základními pojmy jako jsou různé typy validit: interní, externí, deskriptivní, konstrukční či validita statistických závěrů, které se všechny musí brát v potaz při posuzování kvality výzkumu (Farrington 2003; Shadish et al. 2002; Farrington et al. 2002). Pro osoby se znalostí společenskovední metodologie mohou být následující stránky do určité míry trýznivé, ať již z hlediska absence 'řádné' terminologie nebo některých zjednodušujících výroků. Nelze než takové čtenáře poprosit o strpění. Čtenáře bez detailních znalostí společenskovední metodologie pak prosím o strpění s použitím některých termínů, bez kterých by bylo výrazně náročnější tuto kapitolu napsat; když je používám, tak se je vždy snažím řádně vysvětlit.

Experimentální a kontrolní skupina

Představte si, že výsledkem určitého výzkumu bude, že v období dvou let po propuštění z vězení další trestný čin spáchá 50 % odsouzených, kterým byl uložen nepodmíněný trest odnětí svobody, a 35 % odsouzených, kteří dostali jiný trest. Co nám toto tvrzení sděluje? V zásadě velmi málo. Je totiž vysoce pravděpodobné, že skupina odsouzených, kterým byl uložen nepodmíněný trest odnětí svobody, se odlišuje od skupiny pachatelů, kteří dostali podmíněný trest. Roli mohla hrát třeba lítost; v nedávném případě právnička, která tři roky lila svojí nadřazené projímadlo do pití, obdržela nepodmíněný trest odnětí svobody, protože svého činu nelitovala (Pokorný 2017). Výzkum přitom ukazuje, že osoby s odlišným náhledem na svoji trestnou činnost se liší v míře recidivy (Cohen et al. 2012; Tangney et al. 2014). Významně se budou lišit odsouzení jistě i z hlediska předchozí trestní minulosti, protože ta velmi silně ovlivňuje, zda bude obviněnému uložen nepodmíněný či jiný trest; předchozí trestná činnost je přitom považována za jeden z nejlepších prediktorů následné trestné činnosti.

Při posuzování efektivity jakékoli intervence je tak třeba porovnávat skupiny, které jsou shodné ve všech známých i neznámých vlastnostech – či alespoň skupiny tak podobné, jak nám to data dovolí. Skupinu, která podstoupila určitou intervenci, nazýváme

experimentální, a skupinu, která se této intervenci nepodrobila, nazýváme kontrolní.² Pokud se jedná o stejné skupiny, tak mluvíme o experimentu, pokud se jedná o podobné skupiny, tak primárně mluvíme o kvaziexperimentu.³

Cílem jakéhokoli evaluačního výzkumu tak nutně musí být snaha o získání co nejpodobnějších skupin, se kterými bylo zacházeno odlišně z relevantního zkoumaného hlediska. Spolehlivost výsledků vyhodnocení vlivu určité intervence klesá společně s klesající kvalitou vytváření experimentální a kontrolní skupiny; čím méně jsou tyto dvě skupiny podobné, tím méně spolehlivě můžeme říct, že za rozdíl ve výsledcích je zodpovědná právě vyhodnocovaná intervence. Pokud se nám totiž nepodaří získat stejné či srovnatelné skupiny, tedy pokud existují rozdíly mezi experimentální a kontrolní skupinou, právě tyto odlišné vlastnosti mohou mít výrazný vliv na výsledek výzkumu. Tento vliv přitom může být dokonce výraznější než vliv vyhodnocované intervence a v takovém případě argumentační síla jakéhokoli vyhodnocení prudce klesá. Je tak potřeba věnovat výraznou pozornost právě utváření experimentálních a kontrolních skupin.

V některých případech můžeme formulovat vlastnosti účastníků výzkumu, které by mohly (či měly) ovlivnit výsledek intervence. Teoreticky lze zdůvodnit například proč by opakovaní recidivisté měli mít jinou míru recidivy než prvopachatelé: Jejich sociální síť je poničená, byli již označeni jako zločinci či si mohli zvyknout na vězení (byli tzv. prisonizováni). U jiných charakteristik ale nevíme, zda a případně jaký vliv na recidivu mohou mít: je možné, že recidiva Romů je jiná než recidiva ostatních pachatelů; údaje o etnické příslušnosti ale v kontinentální Evropě obvykle nefigurují v databázích na rozdíl od anglo-saského světa, a to zejména kvůli jejich zneužití za druhé světové války (Tonry 1997). Podobných neznámých se v kvaziexperimentálním výzkumu (tedy takovém výzkumu, který nepracuje se stejnými skupinami, ale pouze s podobnými skupinami) může vyskytnout vícero a možnosti spolehlivé interpretace výsledků jsou tak výrazně limitované. Často jsou však z etických a praktických důvodů (jsou jednodušší a levnější) preferovány právě kvaziexperimentální metody.

Lze tak uzavřít, že jediným plně spolehlivým způsobem, jak se vyhnout nekontrolovaným vlivům neznámých charakteristik, je použít experimentální metody. Jediným plně spolehlivým způsobem jak dokázat kauzalitu je tak **náhodné přiřazování případů do kontrolní a experimentální skupiny** výzkumníkem a vyhodnocení vlivu intervence (Bachman & Schutt 2011, p.155; Disman 2011, p.38; Farrington & Welsh 2005; Farrington & Welsh 2006; Shadish et al. 2002, p.13; Weisburd 2003; Sherman 2007; Weisburd et al. 2001; Smith 2013). Této metodě se v angličtině říká „Randomized Controlled Trial“, tedy **randomizovaný kontrolovaný test**. Pokud se dostatečný počet případů (zhruba 200-300 případů pro dostatečně spolehlivé výsledky, byť konkrétní počty záleží na konkrétních podmínkách výzkumu) náhodně přiřadí do experimentální a kontrolní skupiny, tak se považují tyto skupiny za shodné v průměru ve všech aspektech vyjma absolvování intervence, kterou podstoupila pouze experimentální skupina. Pokud tedy náhodně přiřadím polovinu z 600 pachatelů buďto ke krátkému nepodmíněnému nebo delšímu podmíněnému trestu odnětí svobody, tak mohu říct, že obě skupiny jsou ve

² Pro velmi přístupně zpracovaný úvod do společenskovední metodologie v češtině ohledně této otázky doporučuji druhou kapitolu z knihy Miroslava Dismana (2011), popř. první kapitolu anglické knihy od Angrista a Pischkeho (2015). Pro velmi detailní diskuzi problematiky zjišťování kauzálního vztahu za pomoci experimentální či kvaziexperimentálních metod je nejvhodnější odkázat na nejvýznamnější publikaci v této oblasti: knihu Shadisha, Cooka a Campbella (2002).

³ Být hlavní rozdíl mezi těmito je v něčem jiném, jak je naznačeno níže.

všech ohledech stejné vyjma trestu, který jim byl uložen. Jedná se o tzv. zákon velkých čísel (více viz např. Angrist & Pischke 2015, p.18).

Nadřazenost experimentů oproti ostatním metodám, která vedla některé výzkumníky k označení randomizovaných kontrolovaných testů jako zlatého standardu (Farrington 2003; Weisburd et al. 2001; Shadish et al. 2002; Braga 2013), vyvolala kritiku jiných akademiků, kteří poukazovali na některé jejich nedostatky (Tilley 2009; Sampson 2010; Hough 2010; Pawson & Tilley 1997). Kritici zmiňovali, že určité výzkumné otázky není vhodné řešit experimentálně, že kvazi-experimentální metody mohou dosáhnout obdobně kvalitních výsledků jako ty experimentální nebo že experimenty nám nevysvětlují, proč určitá intervence má vliv na některé lidi v specifických kontextech, ale ne na jiné v odlišných podmínkách. Mají pravdu v tom, že experimenty nám neodpovídají na otázku "co funguje pro koho, kdy a kde"; nakonec pak není často možné výsledky z randomizovaných kontrolovaných testů zobecnit na celou populaci (mají nízkou externí validitu). Odpověď na tuto kritiku byla taková, že si zastánci randomizovaných kontrolovaných testů uvědomují, že experimenty nevysvětlují mechanismy kauzálního vlivu, ale že jiné metody mají tak vysoké nedostatky při vysvětlování kauzality, že pouze randomizované kontrolované testy mohou poskytnout skutečnou odpověď na efektivitu určité intervence (Weisburd 2010; Farrington 2003; Weisburd 2003; Sherman & Strang 2004; Shadish et al. 2002). Jiní kritici jsou vůči experimentům vstřícnější a dodávají, že je primárně třeba zlepšit styl, jakým jsou popsány výsledky experimentů (Gill 2011; Perry et al. 2010), popř. že podpora experimentů by neměla vést k nezohlednění výsledků jiných metod (McGloin & Thomas 2013), přičemž někteří volají po upuštění od užití terminologie typu "zlatý standard" (Braga et al. 2013).

Skutečnost, že víme, že určitá intervence má vliv na výsledek, ale neznamena, že lépe poznáme, jak funguje mechanismus, pomocí kterého dochází k působení na účastníky; zjišťování takového mechanismu je přitom potřebné (Sherman & Strang 2004; Pawson & Tilley 1997; Farrington 2003). Často uváděný případ této neznalosti, která ale nemusí bránit používání určité intervence, je léčba kurdějí pomocí citrusů v 18. století britským námořnictvem (Baron 2009). Byť vitamín C byl objeven až v roce 1912, již dříve se vědělo, že citrusy zabírají vzniku kurdějí. Toto poznání může být pro praxi i důležitější než poznání samotného mechanismu účinku. Pro vyléčení kurdějí totiž stačí vědět, že při léčbě kurdějí pomáhá konzumace citrusů; konkrétnímu mechanismu rozumět nemusíme, byť takové porozumění vždy prospěje. Ne vždy ale stačí samotné poznání efektivita bez objevení mechanismu. Naopak až objevení mikroorganismů zodpovědných za přenášení infekcí vedlo k obecnému přijetí požadavků na důkladnější desinfekci, byť efektivita desinfekce byla prokázána již dříve (Tilley 2009). Pro přesvědčení obecné i odborné veřejnosti tak může být potřebné nejen prokázat, že určitá intervence je efektivní, ale i proč. To nicméně neznamena, že by nemělo probíhat vyhodnocování intervencí i bez znalostí mechanismu, proč je intervence efektivní (Farrington 2003).

Čtenář si proto může říct – proč se tedy častěji nezjišťuje efektivita intervencí skrze náhodné přiřazování účastníků do experimentální a kontrolní skupiny? V některých oborech, jako třeba v medicíně, je toto běžným postupem. Ve společenských vědách a zejména v kriminologii se tento přístup začíná prosazovat (tzv. „evidence-based programs“, viz např. Sherman et al. 1998; Farrington 2003; Farrington et al. 2002; Welsh, Braga, et al. 2013, p.2; Welsh, van der Laan, et al. 2013, p.271); často ale nejde účastníky intervence přiřazovat do experimentální a kontrolní skupiny. Jako problematické se může jevit např. náhodné přiřazování trestů jednotlivým pachatelům, byť i v tomto případě existují eticky přijatelné způsoby, jak toto provést; ty jsou rozebrány níže.

Pokud není vhodné či možné použít randomizovaný kontrolovaný test, tak je někdy možné využít situace, kdy jsou vybraní účastníci náhodně přiřazeni určitým intervencím díky zvláštnosti určité situace. Tento případ se nazývá **přirozeným experimentem** (k diskuzi o spolehlivosti této metody viz zejména Fagan, 1990). V takovém případě se zpětně vyhodnocují jednotlivé skupiny, přičemž je ale vždy nutné pečlivě zkontrolovat, zda skutečně došlo k náhodnému přiřazení případů. To je potřebné zejména z důvodu, že výzkumník nemá kontrolu nad přiřazováním osob do jednotlivých skupin.

Níže, po rozebrání pravého experimentu, se budu věnovat tomu, jak lze využít právě vzniku přirozeného experimentu při zjišťování efektivity nepodmíněného trestu odnětí svobody. Jedná se o situaci, kdy jsou trestní případy náhodně přiřazované soudcům (v důsledku čehož vznikají skupiny, které jsou shodné), kteří jsou odlišně přísní, v důsledku čehož posílají odlišné procento pachatelů nepodmíněně do vězení. Jedná se tak o přirozený experiment, kdy jediná odlišnost mezi případy, které řešili dva různé přísní soudci, bude v tom, že jedna skupina zažije v širším měřítku určitou intervenci (vězení), zatímco druhá ji zažije v menším rozsahu. Tento přístup byl již několikrát použitý ve Spojených státech (Berube & Green 2007; Nagin & Snodgrass 2013; Loeffler 2013; Green & Winik 2010; Martin et al. 1993).

Z praktických i etických důvodů jsou bohužel časté případy, kdy není možné náhodně přiřadit účastníky a kdy nenastal přirozený experiment. V takovém případě se používají různé **kvaziexperimentální metody**, které mají za cíl se co nejvíce přiblížit podmínkám experimentu; snaží se tedy identifikovat co nejpodobnější skupiny, kdy jedna podstoupila intervenci a druhá nikoli. Způsobů, které se používají, je několik a jejich konkrétní aplikace v otázce měření účinku nepodmíněného trestu odnětí svobody na pachatele z hlediska recidivy je podrobněji rozebrána níže.

Cílem společenskovědního výzkumu vyhodnocujících efektivitu intervencí je zjistit co nejspolehlivěji, zda určitá intervence „funguje“ či nikoli. Z tohoto důvodu ve společenských vědách existuje výrazný tlak na používání co nejspolehlivějších metod, a to právě randomizovaných náhodných testů. Pro zjednodušení sociologové, psychologové a kriminologové vytvořily seznam metod od těch nejspolehlivějších po ty nejméně spolehlivé, které jsou shrnuty v tzv. Marylandské škále vědeckých metod (Farrington et al. 2002). V této kapitole se pak zabývám jen metodami, které je možné podřadit pod nejspolehlivější úroveň v rámci této školy. Důvod pro takový přístup je jednoduchý: Studií analyzujících efektivitu trestu odnětí svobody bylo nepřeberně mnoho (Villetta et al. (2015) analyzovali přes 3000 studií), ale těch kvalitních je stále málo. Je tak na místě primárně realizovat a věnovat se těm nejspolehlivějším způsobům zjišťování efektivity trestu odnětí svobody; studie používající méně spolehlivé metody takřka nic nepřidají k tomu, co dosud víme.

Proč je potřeba používat co nejkvalitnější metodologie ukazuje working paper z Norska (Bhuller et al. 2016). Jako jedni z velmi mála výzkumníků docházejí ve své práci k závěru, že uvěznění má pozitivní vliv na recidivu, tedy, že ji snižuje. K tomuto závěru docházejí právě na základě pečlivé analýzy přirozeného experimentu. Své výsledky porovnávají s výsledky vícenásobné regrese, která dochází k opačnému závěru (tedy, že vězení vede k vyšší recidivě). Ukazují tak, že méně spolehlivé metody mohou dojít k jinému závěru než k tomu, který se spíše blíží realitě. I vzhledem k tomu, že tento článek dosud nebyl publikován v recenzovaném časopise, je na místě být lehce obezřetný. Při jeho interpretaci je pak třeba zmínit, že jako alternativu k přirozenému experimentu používají jen vícenásobnou regresi, ale nikoli vyvažování, které je považováno za spolehlivější kvaziexperimentální metodu (viz níže). Pokud by toto provedli, tak by mohli naznačit, zda se i jiné, kvalitnější kvaziexperimentální metody, odlišují od výsledků

poskytovaných spolehlivějšími metodami. Nakonec Norové měří na rozdíl od většiny ostatních studií (Drápal 2017) i inkapacitační efekt uvěznění, což mohlo "vylepšit" efekt uvěznění. Další studie v evropských zemích, a mezi nimi i v České republice, by tak měly ověřit, zda k tomuto závěru studie o norském vězeňství dochází z důvodu, že používá spolehlivější metody, že započítává inkapacitaci nebo protože je norské vězeňství výrazně více efektivní než vězeňské systémy v jiných zemích (či vše dohromady).

Před vylíčením jednotlivých metod a možností jejich použití se vyjádřím k možné námitce ze strany orgánů činných v trestním řízení: Ta by mohla zaznít ze strany soudců, popř. policistů, státních zástupců nebo příslušníků vězeňské služby. Oni přece ví, jak jednotlivé tresty působí na pachatele; oni ty pachatele přece vidí a jsou s nimi v kontaktu. Soudce (a podobně i ostatní) nicméně po uložení trestu systematicky nesleduje a neanalyzuje osud pachatelů a neví tedy, zda uložený trest byl účinný či neúčinný z hlediska recidivy či dalšího osudu pachatelů. I pokud by se soudce náhodně s některými pachateli setkával, např. s těmi, kteří v obvodu stejného soudu spáchají další trestné činy a jsou mu přiřazeni k projednání, takový vzorek není v žádném případě reprezentativní, a nemůže proto být zobecněn. Navíc soudci vidí opakovaně pouze ty, kteří neuspěli; nesetkávají se však s těmi, kteří kriminální dráhu opustili. Není možné z něj proto vyvodit závěry o celkové efektivitě nepodmíněného trestu odnětí svobody, natož jeho různých účelů. Je tedy nutné, aby toto dělal někdo za soudce či státní zástupce; s tímto cílem také existuje kriminologie, jako pomocná věda trestního práva.⁴

Metody utváření experimentálních a kontrolních skupin při zkoumání vlivu nepodmíněného trestu odnětí svobody na recidivu

Randomizovaný kontrolovaný test: Náhodné přiřazování trestů pachatelům nezávisle na soudci

Nejspolehlivější metodou, pomocí které lze zjistit, zda je určitá intervence efektivní, je náhodné přiřazování účastníků do experimentální a kontrolní skupiny a měření účinku intervence na tyto skupiny. V případě efektivity trestů by toto znamenalo náhodně vybírat pachatele, kterým budou uloženy přísnější a mírnější tresty. Tento výběr by náhodně činil speciálně pro tento účel navržený počítačový program, nikoli výzkumník či soudce; ani jeden z nich by nemohl ovlivnit výběr trestu v konkrétním případě.

Většina právníků se nejspíše osype, když toto čte: Tresty by přece měly být stanovované na míru pachateli v souladu s principem individualizace; tak jak by bylo možné odůvodnit náhodné určení trestu? Odůvodnění tohoto přístupu však existuje; níže ho detailněji rozebírám. Hlavním důvodem, proč je třeba experimentálně ověřovat efekt trestu odnětí svobody, je, že uložení nepodmíněného trestu odnětí svobody velmi silně omezujeme pachatele. Činíme tak s různými cíli, které nejsou jasně vyjádřené, protože nemáme jasně definovaný účel trestu.⁵ Pokud je naším cílem vězně napravit či jim alespoň zabránit v páchání trestné činnosti po určitou dobu (tedy inkapacitovat), pak musíme zjistit, nakolik tyto teoreticky obhajované účely trestu v praxi reálně fungují. "Pokud nejsme ochotni zjistit, zda určité programy či intervence neškodí, riskujeme, že účastníci těchto programů mohou silně trpět" (Weisburd 2003, p.339).

⁴ Respektive takto o kriminologii mluví právníci; výzkumníci z jiných oborů na to mají jiný názor.

⁵ Srov. nález sp. zn. II. ÚS 2027/17 ze dne 7. 8. 2017, bod 17 a nález sp. zn. I. ÚS 4503/12 ze dne 11. 6. 2014 (N 119/73 SbNU 827), bod 26; podle těchto rozhodnutí by do doby, než zákonodárce blíže stanoví účel trestu, měl trest primárně reprezentovat proporcionalitu k jednání.

Právě proto, že velmi výrazně omezujeme svobodu jednotlivce s určitými cíli, musíme vyvinout významné úsilí k ověření toho, zda tresty dosahují cílených účinků či nikoli. Čím více určitá intervence zasahuje do sféry svobody jednotlivce, tím více energie by měl stát vynaložit na zjištění, zda je nutné takovým způsobem do svobody jednotlivce zasahovat. Toto můžeme vidět na příkladu zajištění osob a věcí a rozhodování Ústavního soudu o nich. Zatímco v případě vazby, kdy je osoba velmi výrazně ovlivněna na své osobní svobodě a soudy proto musí postupovat restriktivně,⁶ při zajištění majetku, což je výrazně méně závažný zásad do práv jednotlivce, Ústavní soud stanovuje orgánům činným v trestním řízení značně širší mantinely, když zdůrazňuje blízkost orgánů rozhodujícího o zajištění k projednávané věci.⁷ Ústavní soud tak v podstatě tvrdí, že pokud jsou narušena ta nejdůležitější práva osoby (zejména právo na svobodu pohybu), tak je třeba postupovat jinak, než pokud jsou omezena méně důležitá práva.

Podobné úvahy o vyvažování různých práv můžeme nalézt zejména v testu proporcionality, který používá Ústavní soud ve chvíli, kdy proti sobě stojí dvě základní lidská práva.⁸ Aplikujme jej na tento experiment; proti sobě de facto stojí právo na osobní svobodu (čl. 8 Listiny základních práv a svobod) a právo, respektive povinnost, státu trestat trestné činy. V rámci testu proporcionality Ústavní soud nejdříve poměřuje vhodnost daného řešení, v tomto případě nepodmíněného trestu odnětí svobody. Ptá se, zda daný prostředek umožňuje dosáhnout sledovaný cíl. Pro zodpovězení této otázky je ale třeba definovat účely tohoto trestu (společenskovední terminologií hypotézy či výzkumné otázky) a ověřit, zda nepodmíněný trest odnětí svobody těchto účelů dosahuje či ne. Některých jistě dosahuje – jeho efektivita se ale dá zjistit pouze z porovnání s efektivitou s dalšími tresty. Proto přichází druhý krok testu proporcionality: kritérium potřebnosti – tedy zda stejného cíle (zde různých účelů trestů) nelze dosáhnout jiným prostředkem. Zde je na místě porovnat právě efektivitu (nepodmíněného) trestu odnětí svobody s ostatními tresty vzhledem k účelům trestů. Pokud by se nepodmíněný trest odnětí svobody ukázal jako stejně či méně efektivní, tak by bylo na místě omezit jeho užívání.⁹

Efektivitu trestů potřebujeme vyhodnotit nejen proto, že cílem práva je být racionálním a efektivním systémem, ale i proto, že to může mít výrazný dopad na rozhodování vrcholných soudních institucí o ukládání trestů. Dovozuje-li Ústavní soud¹⁰ společně s českými akademiky (Novotný 1969, p.8; Solnař & Vanduchová 2009, p.29; Lata 2007, p.67), že účelem trestu v nejobecnějším smyslu má být ochrana společnosti před kriminalitou, pak je nutné ptát se, které tresty skutečně před kriminalitou ochraňují (a zdá vůbec nějaké). Společnost by tak měla podporovat použití co nejspolehlivější metodologie, protože sama společnost sází hodně – zvýšení či snížení míry kriminality. Pokud je trest odnětí svobody neefektivní, či pokud dokonce zvyšuje recidivu, jak naznačují zejména kvaziexperimentální studie (Villettaz et al. 2015; Nagin et al. 2009), pak společnost

⁶ Nález Ústavního soudu sp. zn. I. ÚS 432/01 ze dne 17. 1. 2002 (N 8/25 SbNU 55), nález sp. zn. III. ÚS 612/06 ze dne 30. 11. 2006 (N 215/43 SbNU 393) či nález sp. zn. III. ÚS 1301/13 ze dne 15. 12. 2015.

⁷ Usnesení Ústavního soudu sp. zn. III. ÚS 3930/16 ze dne 11. 4. 2017, usnesení sp. zn. IV. ÚS 1416/16 ze dne 19. 7. 2016, usnesení sp. zn. I. ÚS 2832/15 ze dne 7. 3. 2016 či nález sp. zn. I. ÚS 3502/13 ze dne 17. 4. 2014 (N 63/73 SbNU 209), bod 17.

⁸ Jednotlivé body testu vymezil Ústavní soud již v nálezu sp. zn. Pl. ÚS 4/94 ze dne 12. 10. 1994 (N 46/2 SbNU 57; 214/1994 Sb.).

⁹ V posledním kroku testu proporcionality se pak porovnává závažnost obou v kolizi stojících základních práv, kterou by bylo obtížné aplikovat v případě výběru druhu trestu; to však pro tento účel již není zajímavé, efektivita trestu je důležitá primárně pro zodpovězení na druhý krok testu proporcionality.

¹⁰ Nález Ústavního soudu sp. zn. IV. ÚS 463/97 ze dne 23. 4. 1998 (N 47/10 SbNU 313), nález sp. zn. IV. ÚS 144/05 ze dne 29. 8. 2005 (N 166/38 SbNU 327) či nález sp. zn. III. ÚS 68/04 ze dne 9. 6. 2005 (N 121/37 SbNU 535).

nejedná racionálně a to jak z hlediska humanistické zásady ukládání trestů v minimální možné míře (§ 38 odst. 2 trestního zákoníku) tak z hlediska vysokých nákladů na výkon nepodmíněného trestu odnětí svobody.

Z výše uvedených důvodů je opodstatněné zvolit metody, které nám umožní spolehlivě zjistit, zda je efektivní ukládat nepodmíněný trest odnětí svobody a to i za situace, kdy by byly pro omezený vzorek pachatelů ne zcela naplněny některé zásady ukládání trestů jako přísná individualizace trestu.

Podpurným argumentem, proč připustit omezení (nikoli však opuštění) zásady individualizace trestu ve prospěch náhodného přiřazení trestu u úzkého vzorku, je, že mezi českými soudy existují výrazné rozdíly při ukládání trestů (Drápal 2018). Pokud tedy výsledný trest dnes v mnohém záleží na soudu či soudci (viz dále), který jej ukládá, představa o individualizovaném a naprosto spravedlivém trestu je nekonzistentní. Pokud by se tedy náhodné přiřazování případů ve svém výsledku neodlišovalo od výsledků, které dosahují nejpřísnější a nejmírnější soudci v České republice, nejednalo by se o exces z hlediska praxe ukládání trestů.

Základní podmínkou, kterou by bylo nutné při náhodném přiřazování trestů dodržet, je, že není možné, aby v jeho důsledku bylo pachateli přiřazeno (Farrington 1983, p.290); k účastníkům výzkumu pak musí být přistupováno s respektem a spravedlivě (Shadish et al. 2002, p.281). Pachatel nemůže doplácet na vyhodnocování efektivy jednotlivých trestů. Nelze tak uložit nepodmíněný trest odnětí svobody pachateli, který by mimo experiment dostal jiný, méně závažný druh trestu, a nelze uložit delší trest pachateli, kterému by byl mimo experiment uložen kratší trest. Takovýto experiment je tak eticky přijatelný pouze pokud je částí pachatelů, kteří by dostali přísnější trest, uložen trest mírnější. Malé riziko v takovém případě na sebe bere společnost, která ale zároveň z výsledku experimentu bude nejvíce profitovat. Toto riziko je však minimální, právě proto, že současné výzkumy ukazují na skutečnost, že pachatelé odsouzení k alternativním trestům nepáchají více trestných činů než podobní pachatelé odsouzení k nepodmíněnému trestu odnětí svobody (Villettaz et al. 2015; Nagin et al. 2009).

Nelze však ponechat bez řádného zvážení riziko, které mohou pachatelé trestných činů znamenat, popřípadě postoj společnosti k takovému výzkumu. Náhodně je tak možné rozhodovat jen u trestů, které jsou si blízké z hlediska jejich přísnosti; nebylo by tak například možné, aby vrah byl náhodně přiřazen k alternativnímu trestu. Podobně nelze ignorovat náladu soudců a jejich reakci na podmínky případu: umožní-li to metodologie výzkumu, soudci by měli mít možnost do určité míry ovlivnit výsledek, například z hlediska délky trestů, pokud by dobrovolně přijali závazek neovlivnit druh trestu.

Lze si představit i variantu, kdy by obvinění k takovému experimentování vyjádřili souhlas. Je přitom důvodné domnívat se, že by s tím vyjádřili souhlas - který obviněný by nechtěl, aby mu v případě uložení trestu mohl být snížen trest? Toto řešení by však zároveň mohlo ovlivnit vliv trestu na odsouzené; ti, kterým by byl trest experimentálně snížen, by toto snížení mohli vnímat jako určitou "druhou šanci", určité beneficium. Na rozdíl od lékařských experimentů nelze uložit "placebo" trest (říct některým pachatelům, že jim byl snížen trest, byť jim snížen nebyl, vy v některých případech bylo nemožné a jindy eticky sporné). Ti, kterým by trest nebyl experimentálně snížen, by naopak mohli propadnout určité depresi, že nedostali další šanci, popřípadě by se mohli cítit ukřivdění. Pokud by tedy bylo možné, aby tento experiment byl uskutečněn bez souhlasu obviněných, tak by to bylo vhodnější.

Pro zkoumání efektivity trestů si lze představit několik různých experimentů, které by používaly náhodné přiřazování pachatelů k trestům:

1) **Efektivitu kratšího nepodmíněného trestu odnětí svobody** by bylo možné zkoumat pomocí náhodného přiřazení pachatelů, kterým by soudci plánovali uložit nepodmíněné tresty odnětí svobody v délce do jednoho roku včetně. Ti by byli náhodně přiřazeni buďto k nepodmíněnému trestu odnětí svobody v původní délce či k alternativnímu trestu, popř. k podmíněnému trestu. O délce jiného než nepodmíněného trestu by opět rozhodoval soudce, který by tak mohl vyvážit určitou "mírnost", kterou by jinak viděl v neuložení nepodmíněného trestu odnětí svobody. Výsledek tohoto výzkumu by ukázal, zda krátké tresty odnětí svobody mají vliv na recidivu či nikoli. Zjištění této skutečnosti je důležité, protože významná část nepodmíněných trestů odnětí svobody je kratší či rovná jednomu roku.

2) Odpovědět na otázku, **zda přísnější tresty mají odlišný vliv na recidivu než kratší tresty**, by mohla metoda, podle které by náhodně vybraným pachatelům, které by soudce nepodmíněně poslal na vězení na 1-4 roky, byl snížen trest o pětinu¹¹ délky. Tato metoda by soudcům, zákonodárcům i veřejnosti pomohla zjistit, zda je možné bojovat s kriminalitou či s recidivou jednotlivců pomocí přísnějších trestů; tedy zda je delší tresty odradí od další trestné činnosti, zda jim zabrání páchat určitý počet trestné činnosti či zda platí opak a v důsledku dlouhého nepodmíněného trestu odnětí svobody budou páchat více trestné činnosti. Pokud by delší trest měl negativní či nulový dopad na recidivu, byl by to výrazný argument pro zkrácení ukládaných nepodmíněných trestů odnětí svobody.

3) Rozhodování o trestech je proces, který trvá od jeho uložení do konce jeho vykonání (Hodgson & Soubis 2016, p.248; Padfield 2016, p.250). Termín „rozhodování o trestech“ je výrazně širší než „ukládání trestů“, který odkazuje na prvotní uložení trestu.¹² Při zkoumání efektivity trestů je třeba zkoumat nejen jeho primární uložení, ale i rozhodování o něm, zejména pak institut podmíněného propuštění z výkonu trestu odnětí svobody. V závislosti na výsledku rozhodnutí o podmíněném propuštění totiž trest odnětí svobody může mít odlišný vliv na recidivu, zejména skrze délku pobytu ve vězení.

Měli bychom proto také zkoumat, jaký je **vliv podmíněného propuštění na recidivu**. Konkrétní metodologie by vypadala tak, že odsouzení, které by soudce nechtěl podmíněně propustit, by byli náhodně přiřazeni buďto k podmíněnému propuštění nebo k jeho zamítnutí; do tohoto výzkumu by bylo možné opět zařadit pouze vybrané vězně. Výsledek této studie by mohl pomoci soudcům a zákonodárcům rozhodnout o tom, zda je na místě podmíněně propouštět více či méně vězňů, jelikož bychom byli schopni lépe vyčíslit jejich vliv na kriminalitu i státní rozpočet.

4) **Náhodné přiřazování alternativních trestů**¹³ těm pachatelům, kterým by soudce neplánoval uložit nepodmíněný trest odnětí svobody. Soudce by tak zadal např. do pro

¹¹ Hodnota snížení o jednu pětinu trestu je vybrána lehce arbitrárně. Podobně by mohla být zvolena jedna čtvrtina či jedna desetina. Mělo by se jednat o takové snížení, které sice výrazně nepozmění závažnost trestu, ale zároveň bude simulovat možné snížení trestních sazeb.

¹² Čeština se v tomto liší od angličtiny, ve které je psaná většina současných akademických článků. Anglický výraz „sentencing“ je z tohoto hlediska širší.

¹³ Zde definovaného velmi široce jako peněžitého trestu, obecně prospěšných prací, domácího vězení (ať již s elektronickým či bez elektronického monitoringu), ale i podmíněného odsouzení (s dohledem či bez).

tento účel navrženého počítačového programu, že plánuje uložit jiný než nepodmíněný trest odnětí svobody; tento program by mu náhodně určil druh trestu¹⁴, přičemž o délkách těchto trestů by rozhodoval soudce. Jedním z důvodů studia účinků trestů odlišných od nepodmíněného trestu odnětí svobody je to, že byť je dnes podmíněný trest odnětí svobody nejčastěji ukládaným trestem, tak o jeho účincích víme velmi málo. Porovnání jeho efektivity s efektivitou ostatních alternativních trestů by tak mohlo soudcům nabídnout vodítko, zda více neukládat i ostatní tresty, popřípadě by při přiřazení většího množství případů (a při odpovídajícímu designu studie) bylo možné dojít k závěrům o efektivitě určitých trestů pro detailněji specifikované skupiny pachatelů.

Tyto čtyři způsoby náhodného přiřazování případů mohou odpovědět na jednu ze zásadních otázek, kterou se kriminologové, právníci i politici dlouhodobě zabývají: Jaký vliv mají tresty na pachatele trestných činů z hlediska recidivy? Výsledky těchto výzkumů nám naznačí, zda se z hlediska recidivy vyplatí ukládat určitý alternativní trest, zda ukládat více či méně nepodmíněných trestů odnětí svobody, ukládat delší nebo kratší tresty odnětí svobody nebo vězně ve větší či menší míře z vězení podmínečně propouštět.

Je třeba zmínit, že realita bude o něco problematičtější. Tresty totiž mohou mít odlišný účinek na různé druhy pachatelů (tzv. problém heterogenity, viz např. Sampson, 2010; Smith, 2013); náhodně přiřazovat různé tresty různým skupinám pachatelů by vyžadovalo velké nasazení ze strany soudců, což nelze realisticky očekávat. Tato metoda by tak poskytla primárně základní údaje o těchto otázkách, které by ale byly velmi spolehlivé.

Podstatné by byla zejména skutečnost, aby vybraní (resp. ochotní) soudci náhodně přiřadili **všechny případy**, které by řešili v daném časovém období. Takový přístup by zajistil výrazně vyšší možnost zobecnit výsledky na celou populaci (studie by proto měla vysokou externí validitu), než pokud by si soudce předem vybral jen některé případy, u kterých by umožnil náhodné přiřazení výsledku. V takovém případě by totiž závěry bylo možné vztáhnout jen na případy podobné těm, které byly náhodně přiřazeny.

Dosažení takovýchto výsledků by ale bylo podstatné z jiného důvodu, než jen pro zodpovězení hlavní výzkumné otázky o efektivitě trestů. Výsledky z takového experimentu by bylo možné porovnat s výsledky kvaziexperimentálních metod, které by vycházely ze stejných dat. Pokud by výsledky experimentálních a kvaziexperimentálních metod byly podobné, bylo by možné učinit si lepší úsudek o spolehlivosti kvaziexperimentálních metod, které výrazně převažují v kriminologii. Pokud by byly výsledky podobné, tak by bylo možné používat kvaziexperimentální metody při zodpovídání dalších otázek a předpokládat, že výsledky by byly relativně spolehlivé.

Studií, které využily náhodného přiřazení trestů jednotlivým pachatelům, je bohužel minimálně, byť existují. Z kritického přehledu výzkumů (Villettaz et al. 2015) vyplývá, že do současné doby byly realizovány pouze dva randomizované kontrolované testy: jeden v Idaho (USA, Schneider, 1986) a druhý ve Švýcarsku (Killias et al. 2010; Killias et al. 2000). U obou se však jednalo o malé vzorky účastníků (181 v Idaho a 123 ve Švýcarsku) a o nereprezentativní nepodmíněné tresty odnětí svobody: V Idaho byli mladiství pachatelé odsouzeni k víkendovým trestům odnětí svobody pro mladistvé a ve Švýcarsku k nejvýše 14 dnům odnětí svobody. Tyto experimenty nám proto nabízejí spíše vhled do efektivity velmi krátkých trestů, které nejsou obvyklé v České republice.

¹⁴ Druh trestu by zde musel být omezen na tresty, které lze ukládat většímu množství pachatelů, popřípadě by musel být omezen na pachatele, kterým by bylo možné uložit např. podmíněný trest odnětí svobody, obecně prospěšné práce i peněžitý trest.

Jsem přesvědčený, že mnozí právníci (a zejména soudci) si i po pročtení této části článku říkají, že jim nepřijde správné náhodně přiřazovat tresty pachatelům. Důvodem může být, že jim to nepřijde etické. Pokud se ale experimenty v trestním právu porovnají s experimenty v medicíně, kdy je nový lék záměrně kontrolní skupině nepodáván (v důsledku čehož mohou pacienti i zemřít), uložení lehce odlišného trestu se jeví jako méně morálně problematické. Nad rámec toho se domnívám společně s Davidem Weisburdem (2003), že z morálního hlediska neleží důkazní břemeno na těch, kteří chtějí realizovat randomizovaný náhodný test, ale na těch, kteří jej uskutečnit nechtějí. Pokud není jasné, jaký je efekt určité intervence (jako v případě nepodmíněného trestu odnětí svobody), která je pro dané lidi velmi škodlivá (jako nepodmíněný trest odnětí svobody) a pokud je morálním a zákonným požadavkem ukládat co nejméně citelné sankce, které ještě postačují (viz § 38 odst. 2 trestního zákoníku), je nutné přistoupit k co nejpečlivějšímu vyhodnocení efektivity dané intervence (Weisburd 2003; Shadish et al. 2002, p.272). Tedy k uskutečnění randomizovaného kontrolovaného testu.

Zakončím výklad o experimentování při zjišťování efektivity trestů třemi citáty obhájců této metody. Jako první použiji citát švýcarského kriminologa Martina Killiasa (Killias et al. 2000, p.54), který jako jediný Evropan realizoval podobný experiment: „Odpor k náhodnému přiřazování případů může být krátkozraký, protože nepřipouští zastáncům depenalizace argumentovat naprosto spolehlivými výzkumy. Důkladná vyhodnocení, a závěry z nich vyplývající, mohou chybět v průběhu příštích debat, kdy – i v Evropě – „liberální“ intervence s pachateli trestných činů již možná nebudou podporovány zákonodárci či vládami.“ Již v roce 1983 pak David Farrington (Farrington 1983, p.292), jeden z nejvýznamnějších kriminologů současnosti, uvedl, že podle něj bychom „se měli pokoušet testovat hypotézy a vyhodnocovat intervence za použití experimentálních metod kdykoli jen to je možné.“ Shadish, Campbell a Cook, kteří výrazně ovlivňují směřování uvažování o experimentech svými publikacemi již od šedesátých let, zakončují aktualizovanou verzi své "učebnice": "Pokud nemáme k dispozici žádné spolehlivé experimentální studie hodnotící vliv intervence, riskujeme, že činíme závěry o kauzálních vlivech, které nejsou spolehlivé" (Shadish et al. 2002, p.492). Doufejme proto, že se dočkáme experimentálního vyhodnocení efektivity trestu odnětí svobody i v České republice – čím dříve, tím lépe.

Pokračování tohoto článku bude zveřejněno v čísle 4/2018 časopisu Trestní právo. Popisuje kvaziexperimentální metody (přirozený experiment, nedokonalý přirozený experiment, vyvažování a regrese) a možné využití amnestie z roku 2013 pro analýzu efektivity nepodmíněného trestu odnětí svobody. Závěr druhé části se věnuje tomu, co reálně můžeme měřit pomocí těchto metod a možných limitací všech těchto přístupů. Seznam použité literatury bude uveden na konci druhé části tohoto článku.

V první části tohoto článku, která vyšla v Trestním právu 3/2018, jsem se věnoval popisu důležitosti řádného utváření experimentální a kontrolní skupiny. Aby bylo možné izolovat vliv nepodmíněného trestu odnětí svobody na recidivu od ostatních možných vlivů, nejspolehlivějším způsobem je náhodně přiřadit pachatele k jednotlivým trestům. Lze to udělat i eticky přijatelným způsobem za použití následujících principů: Neuložení přísnějšího trestu než by byl jinak přiřazen, zbytečnému neohrožení společnosti a pouze omezenému snížení trestu pro pachatele v kontrolní skupině.

V druhé části tohoto článku se zabývám kvaziexperimentálními metodami, které lze použít pro měření vlivu nepodmíněného trestu odnětí svobody na recidivu, a jakým způsobem by bylo možné využít amnestii z roku 2013. Závěrem článku pak diskutuji, na co reálně můžeme odpovědět za pomoci těchto metod, jaké jsou jejich limitace a shrnuji obě části článku.

Přirozený experiment: Odlišně přísní soudci rozhodující náhodně přiřazené případy

V některých situacích dochází k náhodnému přiřazování účastníků programu do experimentální a kontrolní skupiny, aniž by to bylo cílem; taková situace se nazývá přirozený experiment. "Přirozený experiment je z definice ta druhá nejlepší metoda po randomizovaném kontrolovaném testu" (Smith 2013, p.53); výsledkem jsou také experimentální a kontrolní skupiny, které se v průměru neliší v ničem jiném než v tom, zda byly podrobeny intervenci či ne. V trestní justici často vzniká přirozený experiment při udílení amnestií, jejichž důsledek na kriminalitu byl opakovaně zkoumán (van der Werff 1981; Drago et al. 2009).

Nejzajímavější metodologie byla v této oblasti použita ve Spojených státech a v Norsku. Výzkumníci použili skutečnosti, že trestní případy jsou náhodně přiřazovány soudcům, kteří jsou odlišně přísní (Berube & Green 2007; Green & Winik 2010; Nagin & Snodgrass 2013; Loeffler 2013; Martin et al. 1993, Bhuller et al. 2016). Toto náhodné přiřazení účastníků je kontrolováno pomocí porovnávání ostatních charakteristik účastníků, které by se neměly odlišovat. Skupina pachatelů, která byla přiřazena mírnějšímu soudci, je tak srovnatelná se skupinou, která byla přiřazena přísnějšímu soudci. Pokud se tak ve výsledku liší pouze tím, že jedna skupina byla více posílána do vězení, protože daný soudce byl přísnější, je možné analyzovat vliv nepodmíněného trestu odnětí svobody na recidivu za pomoci tzv. LATE efektu (Imbens & Angrist 1994; Angrist et al. 1996). Jedná se o tzv. instrumentální proměnné, které využívají situace, kdy je proces rozdělený do dvou po sobě jdoucích procesů: Zatímco v rámci prvního jsou případy náhodně přiřazeny soudcům, v druhém odlišně přísní soudci určují, jaký trest bude uložen. Spojením těchto dvou procesů lze spočítat efektivitu nepodmíněného trestu odnětí svobody.¹⁵

Studie, které použily tento princip, lze spočítat na prstech dvou rukou. Důvodem je zejména to, že tento přirozený experiment nevzniká příliš často, a tam, kde vzniká, nejsou často dostupná data potřebná pro jeho vyhodnocení. Mnoho států totiž neposkytuje (a to ani pro vědecké účely) data umožňující identifikovat jednotlivé soudce, kteří případ rozhodli. V dalších státech neprobíhá náhodné přiřazení případů soudcům, což neumožňuje, aby přirozený experiment vůbec vznikl. Popřípadě ve státech, ve kterých probíhá náhodné (či kvazináhodné) přiřazování pouze části případů soudcům, nejsou dostupné informace o tom, které případy byly přiřazené náhodně a které nikoli.

¹⁵ Pro relativně jednoduché vysvětlení této metody a podmínek, za kterých je možné tuto metodu použít, viz Angrist & Pischke (2015, chap.3).

V České republice je běžné, že na každém soudu existují určité specializace, v rámci kterých jsou případy přiřazovány různým soudcům, kteří řeší tyto specializace. Je přitom běžné, že pro určitou specializaci jsou určeni dva či více soudců. V českém kontextu je tak nutné získat pro využití této metody nejen údaj o soudci, kterému byl případ přiřazen, ale i informaci o tom, v rámci které specializace byl případ přiřazen jednotlivému soudci, protože k náhodnému přiřazování dojde pouze v rámci této specializace. Určitým problémem pro systémy, ve kterých existuje několik specializací, v rámci kterých jsou případy náhodně přiřazované, pak může být, že neexistuje dostatečné množství případů, které lze vyhodnotit. Výhodou, na druhé straně je, že odlišné specializace umožňují sledovat, zda se vlivy intervencí liší podle druhu trestné činnosti (např. obecná specializace a specializace v dopravě).

Jednou z dalších překážek může být i skutečnost, že v některých státech mohou být rozdíly mezi soudci nízké, má-li daný stát vhodně definovanou tzv. 'custody threshold', tedy hranici, kdy je na místě určitému pachateli uložit nepodmíněný trest odnětí svobody. V takovém případě je ale možné použít metodu regresní diskontinuity (regression discontinuity, viz např. Owens & Ludwig 2013; Berk 2013; Shadish et al. 2002, pp.207–245), která využívá skutečnosti, kdy některé případy podrobené intervenci jsou velmi podobné jiným, které jí podrobené nejsou; tuto metodu je možné použít, pokud je jasně stanovena hranice, kdy je (či není) určitý případ podroben intervenci.

Malé množství případů, kdy byla použita tato metoda náhodného přiřazení případů různě přísným soudcům má pak za důsledek, že tato metoda ještě není detailně probádaná. Hlavní nezodpovězenou otázkou je zejména to, zda si soudci vnitřně vytváří odlišné hierarchie trestných činů podle jejich závažnosti a zda v důsledku nich neposílají odlišné typy pachatelů do vězení; v takovém případě by totiž nebyly plně naplněny podmínky přirozeného experimentu (Green & Winik 2010, p.371). Ideální přirozený experiment by vyžadoval, aby přísný soudce Fridrich poslal do vězení všechny pachatele, které do vězení pošle i mírná soudkyně Marie. Bohužel to ale tak nemusí být. Je možné, že Fridrich více posílá do vězení pachatele majetkových trestných činů, kterých je většina, ale je mírnější na pachatele násilných trestných činů, zatímco Marie to dělá naopak. V důsledku tedy Fridrich sice může ukládat více nepodmíněných trestů odnětí svobody než Marie, ale nepošle do vězení všechny ty případy, které do vězení pošle Marie. Stejně tak je možné i to, že určitý soudce je výrazně přísnější na drogové delikty, zatímco je mírnější na násilné; druhý soudce to může mít opačně.

V takových případech tedy nastává otázka, zda jsou porovnávané skupiny skutečně porovnatelné (Shadish et al. 2002, p.13). Na tento problém lze přitom zčásti odpovědět: Můžeme totiž analyzovat míru recidivy u náhodně přiřazených případů soudcům, kteří nejsou odlišně přísní. Pokud se výsledky u skupin, které by měly být shodné, budou lišit, znamená to, že tato metoda je problematická. Pokud zjistíme, že skupiny pachatelé odsouzení stejně přísnými soudci k různým druhům trestů recidivují v rozdílné míře, nemůžeme spolehlivě interpretovat situaci, kdy odlišné skupiny vedou k určitým výsledkům.

Zde se jasně projevují nedostatky a nevýhody přirozeného experimentu. Těmi je zejména skutečnost, že výzkumník nemůže ovlivnit podmínky jeho vzniku, a může se tedy stát, že někdy tento přirozený experiment nenastane. Skutečnost, že výzkumník neovlivňuje vznik experimentu má za neblahý důsledek i to, že nemůže zkoumat specifické výzkumné otázky; naopak vědec musí přizpůsobit výzkumnou otázku možnostem přirozeného experimentu. V případě zkoumání efektivity nepodmíněného trestu odnětí svobody tak přirozený experiment nemůže odpovědět na některé otázky,

třeba zda je z hlediska recidivy efektivnější delší trest než kratší či které alternativní tresty jsou efektivnější.

Přirozeně vznikající experiment tak lze použít pouze na dvě situace ve vztahu k měření efektivity nepodmíněného trestu odnětí svobody. Prvním je porovnání jeho vlivu na recidivu s vlivem alternativních trestů: V rámci specializace jsou případy náhodně přiřazeny různě přísným soudcům. Lze tak rozhodnout, zda je lepší (z hlediska recidivy) uložit alternativní či nepodmíněný trest odnětí svobody. Řešení této situace tak dá odpověď soudcům na to, zda míru kriminality snižují spíše ti soudci, kteří jsou přísní, nebo ti, kteří ukládají méně nepodmíněných trestů odnětí svobody (nebo žádný z nich). Druhá situace, kdy se vytváří přirozený experiment, nastává při rozhodování o podmíněném propuštění. Případy jsou totiž opět náhodně přiřazeny soudcům, kteří rozhodují odlišně o podmíněném propuštění: Někteří propustí více vězňů a někteří méně. Tento druhý případ by nám byl schopný naznačit, zda se z hlediska recidivy vyplatí propouštět více vězňů či nikoli.

V České republice je pravděpodobné, že tyto přirozené experimenty vznikají, a to alespoň na některých soudech. Vzhledem k velkým rozdílům mezi českými soudy při ukládání trestů (Drápal 2018) a vzhledem k zahraničním výzkumům ukazujícím na významné rozdíly mezi soudci při ukládání trestů (např. Anderson et al. 1999; Scott 2010), by nemělo být problematické najít rozdíly při rozhodování mezi českými soudci. Tomu naznačuje i předběžná analýza dostupných dat (viz níže). Metoda přirozeného experimentu tak nabízí cestu k relativně spolehlivým výsledkům, kterou je možné zvolit i v České republice, protože nastavení procesů v České republice odpovídá nárokům tohoto přístupu a stejně tak jsou dostupná i potřebná data. Podmínkou je však náhodné přiřazení případů jednotlivým soudcům v rámci specializace - bez ověření této podmínky jak statistickými, tak kvalitativními výzkumnými metodami by metoda přirozeného experimentu neměla být považována za vysoce spolehlivou.

Za účelem posouzení, zda existují podmínky potřebné pro vznik přirozeného experimentu, jsem analyzoval data poskytnutá Ministerstvem spravedlnosti, které byly vyexportovány ze soudních systémů jednotlivých okresních soudů a propojeny se statistickými listy trestními. Tato data obsahovala údaje o soudci, který případ rozhodl, a o specializaci, v rámci které byl případ přiřazen určitému soudci.

Tato data vykazovala určité nedostatky: Některé soudy nezažádaly specializaci do soudního systému - skutečnost, že u významné menšiny soudů neexistovaly žádné specializace, není možná. Oproti tomu bylo třeba odlišit situace, kdy soudy zadávaly určité specializace do systému (např. týkající se cizinců, dopravních přestupků či s obviněnými ve vazbě), přičemž většině případů nebyla přiřazena žádná specializace. V takových případech jsem předpokládal, že případy bez specializace jsou považované za obecnou specializaci, což v mnohém potvrdila i analýza níže. Dalším nedostatkem bylo, že soudy v rámci specializací často nerozlišovaly mezi mladistvými a dospělými pachateli, byť o mladistvých rozhodují specializované senáty. Z tohoto důvodu byli mladiství pachatelé kompletně vyřazeni z této analýzy.

Skutečnost, zda přirozený experiment v kontinentálních podmínkách skutečně vzniká, nebyla jistá. Soudci totiž o trestu rozhodují až v návaznosti na vinu; mohou existovat, a jistě do určité míry existují, rozdíly z hlediska rozhodování o vině. Pokud by jeden soudce shledával více nevinnými určité typy obviněných než druhý soudce, neumožnilo by to vznik přirozeného experimentu. Obdobně by problémem mohla být skutečnost, že soudci mohou v odlišné míře aplikovat různé typy odklonů, což by opět ovlivnilo otázku ukládání trestů. Skutečnost, že daný přirozený experiment vzniká, tak

musí být potvrzena z hlediska možných statistických rozdílů mezi osobami, odsouzenými jednotlivými soudci; nelze se spolehnout na prostý odkaz na specializace.

Za tímto účelem byly vybrány pro tuto analýzu případy, ve kterých byl pachatel pravomocně odsouzen v letech 2008 a 2009. Jelikož soudce neměl plnou možnost volby při ukládání souhrnného a společného trestu, případy, ve kterých byl tento trest uložen, nebyly analyzovány. Rok 2008 a 2009 byl vybrán, jelikož data o specializacích jsou dostupná od roku 2007 a případný výzkum bude mít zájem o co nejdelší dobu, po kterou bude moci měřit recidivu. Bylo však předpokládáno a z předběžné analýzy vyplynulo, že data za rok 2007 jsou méně spolehlivá, nejspíše protože informační systém byl na některých soudech teprve zaváděn. Celkem tak bylo použito 140.348 trestních odsouzení.

Specializací u jednotlivých soudů bylo celkem 1.211; pro další analýzu byly vybrány pouze ty, které obnášely nad 300 případů, kterých bylo 93. Z těch bylo dále vybráno 74 specializací u jednotlivých soudů, ve kterých se alespoň dva soudci zabývali každý více než 100 případů. Z nich bylo dále vybráno 41 soudů, u kterých bylo potvrzeno (na hladině $p < 0.05$), že charakteristiky pachatelů a případů se nelišily mezi soudci téhož soudu. Konkrétně se jednalo o následující charakteristiky případů: Věk, pohlaví a státní příslušnost odsouzeného, skutečnost, zda se jedná o recidivistu z hlediska zákona, zda se jedná o recidivistu z hlediska kriminální recidivy, jaká byla průměrná závažnost trestného činu, za který byl odsouzený, která byla počítána jako 1/3 mezi dolní a horní hranicí sazby, a zda byl odsouzený shledán vinným z jednoho či více trestných činů. Pro provedení této analýzy byl použit balíček *npmv* v *R*, který umožňuje provedení neparametrické MANOVA.¹⁶

Za použití stejné metody a stejného balíčku bylo zjištěno, že ve 13 z 41 specializací existují statisticky významné rozdíly v tom, jak jednotliví soudci ukládají nepodmíněné tresty odnětí svobody. Jejich kompletní výčet, včetně uvedení soudu a specializace jsou uvedeny v Tabulce 1. Údaje z této tabulky jsou pak samozřejmě zajímavé i pro analýzu ukládání trestů - ukazují, že v některých případech velmi silně záleží na soudci, který stanovuje trest, a tedy nikoli na okolnostech řešených případů.

Tabulka 1: Procento uvězněných u soudců, kterým byly náhodně přiřazeny případy, 2008-2009

Soudce	Procento NEPO trestů	N
<i>Okresní soud v Českých Budějovicích (dopravní spec.)</i>		
1	0,9	111
2	3,4	324
3	7,7	169
4	11	146
<i>Okresní soud v Olomouci (dopravní spec.)</i>		
1	1,4	291
2	4,7	255
<i>Okresní soud v Trutnově (dopravní spec.)</i>		
1	3,1	254
2	7,9	165

¹⁶ Multivariate analysis of variance. Statistická metoda, která zjišťuje, zda existují vztahy mezi skupinami případů s několika různými proměnnými; pro podrobnější vysvětlení viz např. Warne (2014). Neparametrická verze byla zvolena, protože použitá data nesplňovala podmínky pro použití parametrické MANOVA. Pro detaily ohledně tohoto balíčku a metod, které používá, je vhodné odkázat na dokumentaci balíčku (Burchett et al. 2017; Burchett & Ellis 2017).

<i>Okresní soud v Hodoníně (spec. NA¹⁷)</i>		
1	5,8	344
2	6,1	509
3	6,3	415
4	12,7	110
<i>Okresní soud v Prostějově (spec. NA)</i>		
1	10,2	197
2	11,1	180
3	22,4	255
<i>Obvodní soud pro Prahu 1 (spec. NA)</i>		
1	17,4	241
2	18,1	155
3	19,8	121
4	20	265
5	28,2	245
6	29,1	247
7	33,6	125
<i>Obvodní soud pro Prahu 3 (spec. NA)</i>		
1	9,2	142
2	15,2	112
3	25,2	135
<i>Okresní soud v Liberci (spec. NA)</i>		
1	3,7	162
2	3,8	341
3	4,1	341
4	5,3	431
5	10	350
6	10	341
7	13,5	459
<i>Okresní soud v Olomouci (spec. NA)</i>		
1	10,2	205
2	11,2	304
3	11,5	374
4	17,4	121
5	18	345
<i>Okresní soud v Semilech (spec. NA)</i>		
1	9,9	494
2	16,1	483
<i>Okresní soud v Ústí nad Orlicí (spec. NA)</i>		
1	9	267
2	10	269
3	19,7	183
<i>Okresní soud Plzeň – jih (spec. NA)</i>		
1	6,3	142
2	12,1	282
<i>Okresní soud v Rokycanech (spec. NA)</i>		
1	9,5	179
2	17	229 ¹⁸

¹⁷ V těchto případech se jedná o tzv. obecnou specializaci, tedy o trestné činy, které nebyly přiřazeny do žádných ze zbývajících specializací. Obecně se v této kategorii nacházelo největší množství případů.

¹⁸ Zdrojem dat jsou statistické listy trestní soudní poskytnuté Ministerstvem spravedlnosti k výzkumným účelům; vybrány byly případy, ve kterých byl pachatel pravomocně odsouzen v letech 2008 a 2009.

Pro ověření, zda jsou na těchto soudech případy rozhodovány skutečně náhodně, by bylo vhodné dále analyzovat rozvrhy práce u daných soudů a případně provést rozhovory s pracovníky těchto soudů.

Nedokonalý přirozený experiment: Odlišně přísné soudy

Druhou variantou přirozeného experimentu je analyzovat rozdíly v uvěznění mezi jednotlivými soudy, které se od sebe neodlišují v relevantních charakteristikách; z hlediska použitých statistických metod by byl přístup stejný jako v předchozím případě různě přísných soudů.¹⁹ Tato metoda je o něco méně spolehlivá, jelikož případy nejsou jednotlivým soudům přiřazovány náhodně; byť je možné ověřit, zda jsou vybrané charakteristiky okresů a případů podobné, vždy může vzniknout podezření, zda jsou případy řešené na jednotlivých okresech skutečně podobné. V tomto ohledu je tento přirozený experiment spíše kvaziexperimentem. Při výběru případu bude tak záležet nejen na empirickém prokázání, že mezi okresy nejsou významné rozdíly, tak v teoretickém ukotvení tvrzení, že mezi soudy by neměly být výrazné rozdíly, např. z důvodu podobných charakteristik či sousedění.

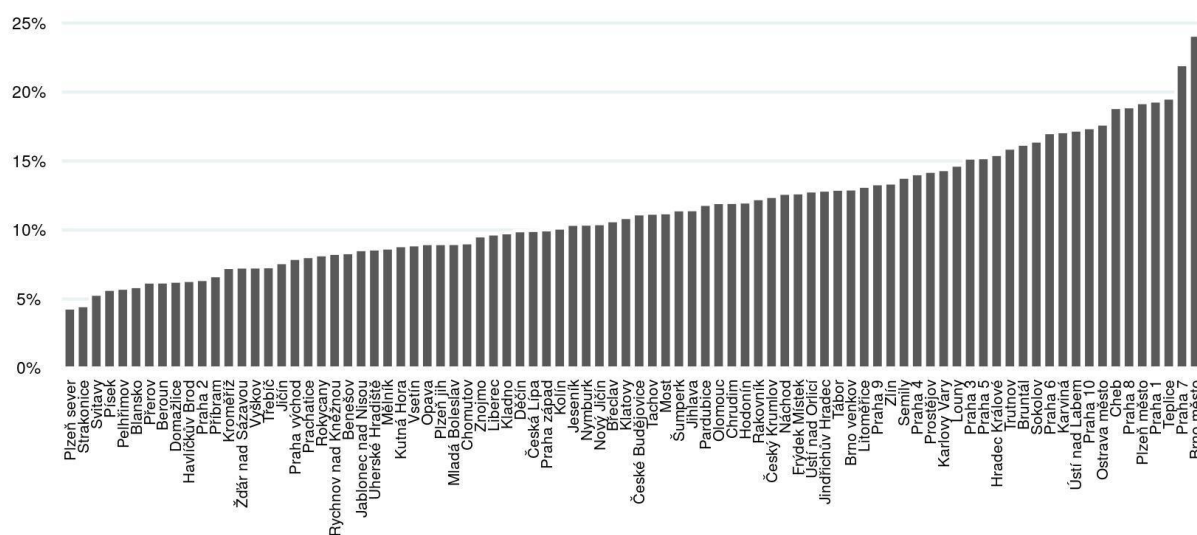
Byť je tato metoda o něco méně spolehlivá než metoda náhodného přiřazení případů různě přísným soudům, skýtá širší možnosti jejího uplatnění. Okresní soudy se totiž výrazně neliší jen v celkovém procentu pachatelů, které uvězní, ale i v procentech pachatelů jednotlivých trestných činů, které uvězní. Grafy 1-4 tak ukazují, jak okresní soudy ukládaly nepodmíněné tresty odnětí svobody obecně a u třech vybraných trestných činů: Opakované krádeže (§ 205 odst. 2 TZ), zanedbání povinné výživy (§ 196 TZ) a Maření výkonu úředního rozhodnutí a vykázání v případě zákazu řízení (§ 337 odst. 1 písm. a) TZ) v letech 2012-2014.²⁰

Graf 1: Procenta uložených nepodmíněných trestů podle okresních soudů, 2012-2014²¹

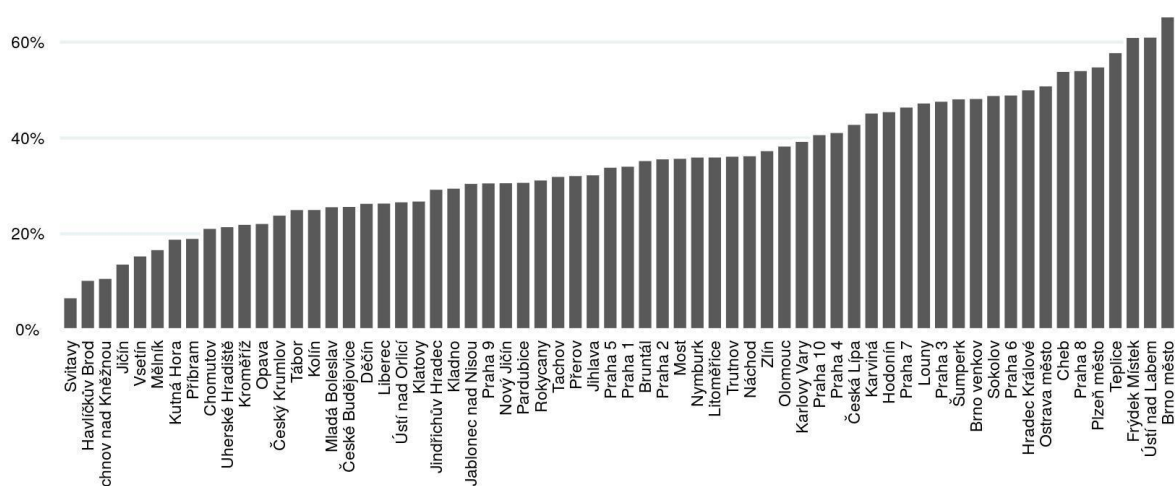
¹⁹ Za tuto myšlenku vděčím Liboru Duškovi.

²⁰ Jedná se o trestní případy zaznamenané ve statistických listech trestních soudních, ve kterých soudy vydaly rozhodnutí v letech 2012-2014; pro větší srovnatelnost byly vybrány pouze ty případy, ve kterých byli souzeni dospělí (věk v době spáchání roven či vyšší 18 roků), kdy případy skončily vyslovením viny, pachatel nebyl odsouzený za nestejnorodý souběh a uložený trest nebyl společný či souhrnný. Zohledněni byli pouze pachatelé odsouzení podle nového trestního zákoníku. Pro přehlednost a spolehlivost byly zahrnuty pouze soudy, které rozhodly určitý počet případů. Pro bližší diskuzi o rozdílech mezi soudy při ukládání trestů a odůvodnění tohoto výběru viz Drápal (2018).

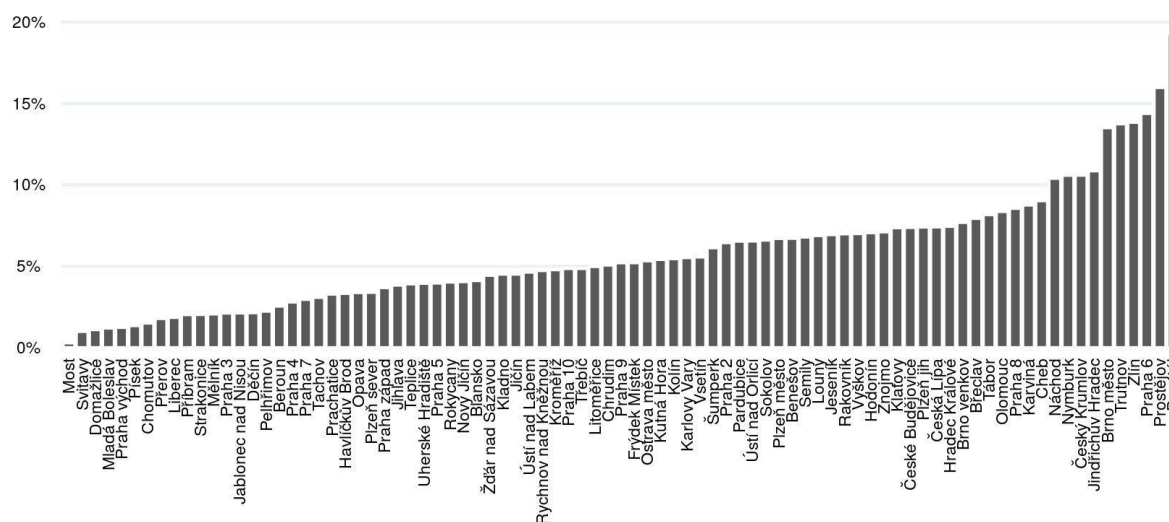
²¹ Zdrojem dat jsou statistické listy trestní soudní poskytnuté Ministerstvem spravedlnosti k výzkumným účelům; bližší specifikace dat je uvedena v článku.



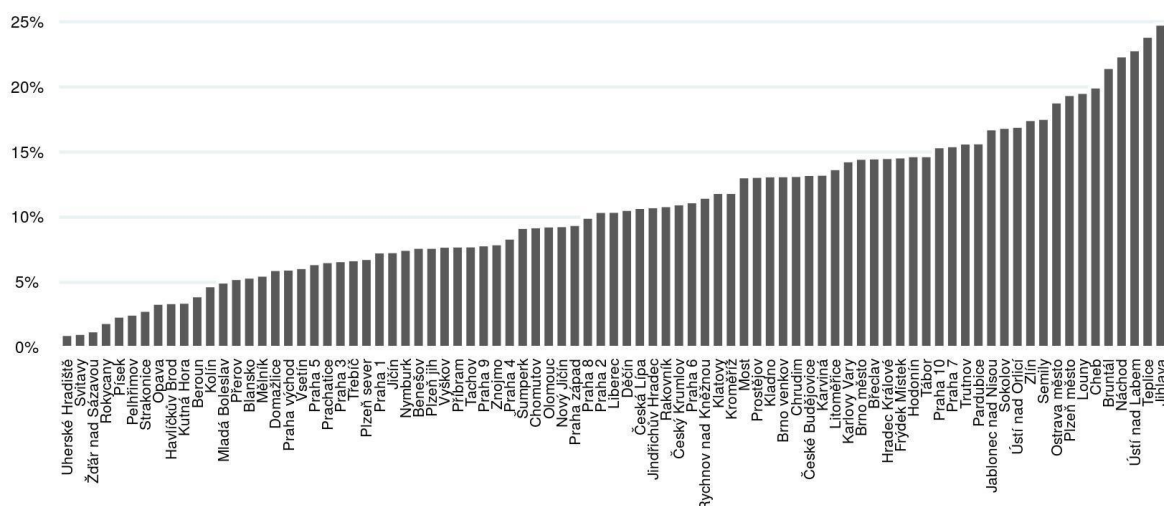
Graf 2: % NEPO trestů za opakovanou krádež (§ 205 odst. 2 TZ), 2012-2014



Graf 3: % NEPO trestů za zanedbání povinné výživy (§ 196 TZ), 2012-2014



Graf 4: % NEPO trestů za maření výkonu zákazu řízení (§ 337 odst. 1 písm. a) TZ), 2012-2014



Z grafů 1-4 je vidět, že bez zohlednění charakteristik případů či okresů rozdíly mezi jednotlivými soudy jsou velmi vysoké - sahají od 1 do 25 % v případě řízení po zákazu či od 5 do 65 % u opakované krádeže. Rozdíly mezi soudy přitom setrvávají významné i ve chvíli, kdy zohledníme vybrané charakteristiky případů, jako např. trestní minulost pachatele (Drápal 2018). Některé rozdíly naznačují, že výše zmíněná metoda by mohla být použita. Minimálně pravděpodobné je, že dvojnásobný rozdíl v uvěznění opakovaných zlodějů mezi Prahou 8 a 9 nebo dvacetinásobný rozdíl v uvěznění osob, které řídí po zákazu v obvodu okresního soudu Žďáru nad Sázavou a Jihlavy, není způsoben pouze odlišnými charakteristikami případů.

Tato o něco méně spolehlivá varianta přirozeného experimentu tak výzkumníkům umožňuje zohlednit nejen celkový vliv nepodmíněného trestu odnětí svobody na následnou recidivu, ale i analyzovat tyto vlivy u různých skupin pachatelů: Zajímavé by bylo například zkoumat pachatele násilných, majetkových, drogových či vybraných trestných činů proti pořádku ve věcech veřejných. Takový přístup by byl v souladu s

principem "realistic evaluation", popř. "theories of change" (srov. Pawson & Tilley (1997; Blamey & Mackenzie (2007)).

Kvaziexperimentální metody: Vyvažování a regrese

Poslední ze zmíněných metod, která je z nich zároveň nejméně spolehlivou cestou, kterou se výzkumník může vydat, je zároveň ta nejvyužívanější. Nedochází k reálnému náhodnému přiřazování případů do experimentální a kontrolní skupiny, ale pouze k fiktivnímu vytváření těchto skupin. Důvodem je zejména to, že realizace experimentu je v některých případech nemožná či velmi obtížná: Experiment se proto simuluje. Výzkumník tak vytváří (kvazi)experimentální a (kvazi)kontrolní skupinu na základě určitých charakteristik. Jeho cílem je přitom vytvořit takové skupiny, které jsou co nejporovnatelnější; tedy takové, které se co nejvíce blíží situaci v rámci experimentu, kdy by byly případy náhodně přiřazované do jednotlivých skupin. Jejich hlavní nevýhodou je nicméně skutečnost, že nikdy nelze říct, zda byly vytvořeny skutečně srovnatelné experimentální a kontrolní skupiny: Nevíme totiž, zda neexistuje určitá proměnná, která sice výrazně ovlivní výsledek (tedy recidivu), ale která byla výzkumníkovi neznámá a kterou tak nemohl použít pro vytvoření podobných skupin (Nagin & Snodgrass 2013, p.602).

Příkladem takové proměnné může být vnitřní postoj pachatele ke spáchanému trestnému činu či k uloženému trestu, které mohou ovlivnit míru recidivy (Tangney et al. 2014; Cohen et al. 2012). Velmi pěkně se k této problematice vyjadřuje Nieuwbeerta, Nagin a Blokland (2009, p.229): „pokud posíláme do vězení pachatele více náchylné k páčání trestné činnosti, míra jejich recidivy bude vyšší než mezi těmi, kterým nebyl uložen nepodmíněný trest odnětí svobody, a to bez ohledu na jakýkoli vliv vězení na vývoj jejich kriminální dráhy.“ Jeden z nejvýznamnějších kvantitativních kriminologů současnosti a výkonný redaktor *Journal of Quantitative Criminology* David Weisburd se v jednom ze svých článků do hloubky zabývá odůvodněními, proč se používají kvaziexperimentální metody a proč naopak mnozí autoři nevnímají jako vhodnou možnost použití experimentu a náhodného přiřazování případů. Dochází ale stejně k závěru, že kvaziexperimentální metody by se měly používat jen tehdy, pokud není možné uspořádat experiment – a primárně by se kriminologové měli snažit o organizaci experimentů (Weisburd 2010).

Mezi nejpoužívanější kvaziexperimentální metody v této oblasti se řadí tzv. vyvažování („matching“), které se snaží vyvažováním charakteristik jednotlivých pachatelů (např. věku, spáchaného trestného činu či předchozí trestní minulosti), vytvořit porovnatelné případy, které zažily odlišné intervence (byly jim uloženy odlišné tresty), a proto umožňují vytvoření kontrolní a experimentální skupiny. Používají se dvě základní varianty vyvažování: V rámci první, která se nazývá „propensity score matching“, se pro jednotlivé pachatele nejdříve vypočítá skóre na základě charakteristik, které byly shledány, že mají vliv na výslednou proměnnou (Shadish et al. 2002, p.162), v tomto případě tedy recidivy. Následně se proti sobě postaví pachatelé se stejným skórem, kdy jeden obdržel nepodmíněný trest odnětí svobody a druhý nikoli, čímž se vytvoří experimentální a kontrolní skupina. Druhým základním způsobem použití vyvažování je tzv. „exact matching“, v rámci kterého jsou jednotlivci, kterým byl uložen jiný trest, propojováni na základě shody ve všech (či většině) měřených charakteristik. Druhá metoda má přitom jednoduchou nevýhodu: Často ve výzkumu potřebujeme použít tak velké množství charakteristik, na základě kterých chceme vytvořit porovnatelné skupiny, že nenajdeme

pachatele, kteří by zároveň měli shodné vlastnosti a zároveň by jeden obdržel nepodmíněný trest odnětí svobody a druhý jiný trest (Nagin et al. 2009, p.138).

Nagin, Cullen a Jonson (2009, p.148) předkládají minimální seznam charakteristik, na základě kterých by mělo docházet k vyvažování. Jedná se o věk, etnicitu, pohlaví, předchozí odsouzení a spáchaný trestný čin, přičemž významná část ze studií nenaplní ani tyto minimální požadavky; při analýze efektivity trestání mladistvých pouze jedna studie vycházelo z tohoto minimálního seznamu charakteristik (McGrath & Weatherburn 2012, p.29). Z tohoto minimálního výčtu lze vyvodit dvě lehce protichůdné myšlenky týkající se etnicity. Citovaní autoři primárně hodnotí situaci ve Spojených státech, kde je etnicita velmi podstatná pro jakýkoli výzkum. Je přitom otázkou, zda tento požadavek nereflakuje spíše americký kontext než požadavek vztahující se na studie realizované po celém světě. Tomuto však lze oponovat tím, že významná část české vězeňské populace je romského etnika²², a pokud naše statistiky běžně nesledují tento ukazatel, tak to může narušit spolehlivost takových výsledků, pokud je kriminální kariéra Romů jiná než ta ne-Romů. Při použití vyvažování (ale i jiných kvazi-experimentálních metod) je ale vždy možným problémem, že byly pominuté některé podstatné vlastnosti účastníků (Weisburd et al. 2001; Weisburd 2010; Shadish et al. 2002).

Druhou nepoužívanější kvaziexperimentální metodou je vícenásobná regrese („multivariate regression analysis“), která se snaží popsat, jaký je vliv několika proměnných (např. věku, předchozí trestní minulosti či vybraného trestu) na recidivu. Myšlenka je taková, že pokud kontroluji charakteristiky pachatele a případu, které mohou ovlivnit recidivu, tak mohu odlišit vliv druhu trestu a vliv těchto charakteristik na recidivu. Tato metoda je nicméně v mnohém problematická: Podle Nagina totiž nedokáže řádně kontrolovat věk pachatele, přičemž ale s ubývajícím věkem výrazně klesá míra páchané trestné činnosti (Nagin et al. 2009, p.155). Dále se v případě regrese počítá s lineárním vztahem mezi recidivou a charakteristiky pachatele (byť je možné analyzovat i nelineární vztahy), zatímco v případě vyvažování tomu tak není, což ve výsledku lépe reflektuje odlišný vliv různých hodnot na recidivu pachatele (McGrath & Weatherburn 2012; Apel & Sweeten 2010). Někteří autoři jdou až tak daleko, že u těchto důvodů regresi samotnou nepovažují za spolehlivou metodu při měření efektivity trestu odnětí svobody (Nieuwbeerta et al. 2009, p.228).

Kvaziexperimentální metody jsou dnes často používány z nouze, protože není možné použít spolehlivější metody. To ale neznamená, že nejsou zlepšovány – naopak jejich výsledky lze považovat za čím dál tím spolehlivější. Při jejich interpretaci je ale třeba brát v potaz jejich nedostatky a to zejména v okamžicích, kdy dosahují jiných výsledků než experimentální studie (k trestu odnětí svobody viz Villettaz et al. 2015; Nagin et al. 2009; obecně k tématu viz Weisburd et al. 2001). Hlavní výhodou kvaziexperimentálních metod je, že lze zpětně analyzovat sebraná data a odpovídat na nejrůznější otázky: od efektivity nepodmíněného trestu odnětí svobody oproti alternativním trestům, přes analýzu efektivity délky nepodmíněného trestu po analýzu efektivity podmíněného propuštění. Na tyto otázky pak lze jednoduše odpovídat pro jakékoli skupiny pachatelů: Prvotrestané, opakované recidivisty, či třeba pro pachatele jednotlivých trestných činů. Nicméně právě proto, že tyto metody jsou jednoduše dostupné v dnešní době velkých dat, je třeba zdůrazňovat potřebu experimentů, které jediné mohou plně a spolehlivě vypovědět o

²² Nezaznamenal jsem výzkum, který by se tomuto tématu věnoval; neexistují tedy jasné čísla. Odhady pracovníků z vězeňství, se kterými jsem mluvil, se pohybují okolo 50 % z vězeňské populace. Zodpovědět tuto otázku jasněji je problematické i z důvodu neexistence jasné definice toho, kteří vězni by měli být považováni za vězně romského etnika.

kauzálním vlivu určitého trestu na pachatele. Kvazi-experimentální metody bohužel svádějí k jednoduchému využití bez snahy použít kvalitnější, ale náročnější metody - experimentální (Shadish et al. 2002, p.486).

To neznamená, že nemají být použité i jiné metody, a to ať k zpřesňování či doplnění těchto poznatků; "randomizované náhodné testy, kvazi-experimenty a pozorování by měly být výslovně považované za doplňky, nikoli za soupeře" (Braga et al. 2013, p.281). Při hodnocení efektivity intervence by však měly být upřednostňovány randomizované kontrolované testy, jelikož ty nejspolehlivěji dokazují kauzální vztah mezi intervencí a výsledkem; a spolehlivé výsledky jsou potřeba, pokud mají být nabídnuty zákonodárcům či praktikům pro změnu přístupu (Weisburd 2010). Byť tedy nejsou jediné, jsou tou preferovanou metodou a mají mít speciální místo při výběru nejspolehlivější metody (Shadish et al. 2002, pp.30, 98). Kvaziexperimentální metody je až následují.

Využití amnestie pro analýzu specifického vlivu uvěznění

V důsledku amnestie z roku 2013 (rozhodnutí prezidenta republiky č. 1/2013 Sb., o amnestii ze dne 1. 1. 2013) byly bezpodmínečně prominuty nevykonané nepodmíněné tresty odnětí svobody, pokud byly pravomocně uloženy před 1. lednem 2013 ve výměře nepřevyšující jeden rok (čl. I. odst. 1). Odsouzení k těmto trestům tak vykonali různou dobu nepodmíněného trestu odnětí svobody - někdo jen jeden den, druhý třeba jedenáct měsíců a dvacet devět dní. Toto prominutí různých délek (oproti původní plánované délce) nám umožní studovat vliv uvěznění na recidivu dvěma různými způsoby: přirozeným experimentem, který v důsledku takového rozhodnutí vzniká, a metodou regresní diskontinuity (regression discontinuity, dále i jen "RD").

Přirozený experiment

Přirozený experiment vzniká v okamžiku, kdy jsou různé intervence náhodně přiřazeny jednotlivým osobám bez zásahu výzkumníka, ale na základě nějaké vnější události - v tomto případě amnestie. Pokud by tak amnestie byla vydána 1. 5. 2013 a nikoli 1. 1. 2013, tak by byli propuštěni odlišní odsouzení, kteří by vykonali jinak dlouhé části ze svých nepodmíněných trestů odnětí svobody. Náhodnost přiřazení různé délky nepodmíněného trestu odnětí svobody (z kriminologického hlediska) tak spočívá právě ve stanovení data účinnosti amnestie.

Není přitom na místě očekávat, že se budou odlišovat pachatelé, kteří vykonali různé délky trestů odnětí svobody, byť pro zvýšení spolehlivosti výsledků bude nutné toto statisticky potvrdit. Nad rámec toho by bylo možné zúžit skupinu odsouzených, kteří by byli analyzováni: Pokud by se třeba vybrali jen ti odsouzení, kterým byl uložen jeden roku odnětí svobody a kteří nevykonali více než 25-40 % z jim uloženého trestu, tak je takřka jisté, že takoví pachatelé nebudou mít odlišné charakteristiky. Je totiž nepravděpodobné, že by byla byť část z těchto odsouzených podmíněně propuštěná z výkonu trestu odnětí svobody (srov. § 88 trestního zákoníku).

Právě posouzením vlivu různé dlouhých trestů na odsouzené, kteří jsou v průměru podobní a liší se jen v tom, jakou část trestu odnětí svobody vykonali, bude možné posoudit vliv různé délky nepodmíněného trestu odnětí svobody na recidivu. Pokud se totiž odsouzení neliší v ničem vyjma délky vykonaného trestu (což vychází z teorie a bylo by ověřeno i statisticky), tak lze sledovat vliv dané intervence na různé charakteristiky - včetně recidivy.

Regresní diskontinuita

Druhou možností, jak využít amnestii pro zjištění vlivu uvěznění na recidivu je analyzovat mezní případy pomocí metody regresní diskontinuity. Ta je založena na porovnávání případů, které jsou takřka totožné, ale je s nimi zacházeno jinak. Lze si představit třeba řidiče, který řídí 59,9 a 60,1 km/h, přičemž pouze ten jedoucí přes 60 km/h bude potrestán; nebo studenty, kteří získají 168,7 a 168,8 bodů v přijímacím řízení, přičemž hranice pro přijetí je právě 168,8 bodů. Takto blízké skupiny jsou de facto totožné; jediný důvod je v intervenci.

Jak by se metoda regresní diskontinuity mohla použít u amnestie? Představme si Viktora a Petra, kteří jsou dobří kamarádi. Viktor byl pravomocně odsouzen koncem roku 2012 k nepodmíněnému trestu odnětí svobody; ani nestačil nastoupit a trest mu byl již prominutý. Petr byl odsouzen k nepodmíněnému trestu odnětí svobody v délce jednoho roku až těsně po Silvestru; když nastupoval do výkonu, tak cestou spatřil šťastného Viktora, jak do sebe leje jednoho ležáka za druhým. V průměru (tedy pokud zohledníme vícero případů), tak by neměl být žádný rozdíl mezi Viktorem a Petrem - rozdíl několika dnů je tak nepatrný, že není důvodné očekávat jakékoli rozdíly mezi nimi (byť i to je vhodné ověřit); to platí zejména u těch osob, u kterých byla rozhodnutí sice vydána v roce 2012, ale právní moc nabyla až v roce 2013; v takovém případě se na ně totiž amnestie nevztahovala (Grivna a Říha 2013).

Mělo by tak být možné porovnávat odsouzené, kteří vůbec nevykonali trest odnětí svobody (jelikož byli akorát pravomocně odsouzeni na úplném konci roku 2012 a do výkonu nenastoupili, popř. do něj nastoupili na velmi krátkou dobu), s odsouzenými, kteří vykonali trest odnětí svobody po svém pravomocném odsouzení z úplného počátku roku 2013 a to v plném rozsahu. Právě protože by mezi těmito odsouzenými neměly existovat žádné rozdíly vyjma trestu, tak bychom měli být schopni zkoumat vliv uvěznění na recidivu. Jsme totiž schopni porovnávat odsouzené, kteří měli vykonávat trest odnětí svobody, ale díky amnestii jej nevykonávali, s takřka totožnou skupinou odsouzených, kteří jej ale vykonat museli.

Tohoto přístupu by navíc mělo být možné použít i pro analýzu vlivu různě dlouhých nepodmíněných trestů odnětí svobody na recidivu. Mělo by totiž být možné zjistit vliv uvěznění na recidivu (pokud by byl k dispozici pro analýzu dostatečný počet amnestovaných a odsouzených) u skupin osob, kterým byl uložen nepodmíněný trest v různé délce (např. 3, 6 a 12 měsíců), popř. u skupin odsouzených s různými charakteristikami (starší a mladší, ženy a muži či odsouzení s různou kriminální minulostí apod.), právě za použití metody regresní diskontinuity. Výsledky těchto přístupů by nám mohly poskytnout lepší informace o tom, zda má kratší nepodmíněný trest (do 1 roku) různý vliv na odlišné skupiny odsouzených.

Co můžeme reálně měřit z dat, které máme?

Na závěr této kapitoly se tak můžeme zeptat: Co reálně můžeme měřit pomocí různých metod a jak to ovlivní naše výroky o efektivitě nepodmíněného trestu odnětí svobody? Z výkladu výše by mělo být zřejmé, že pomocí různých metod můžeme měřit jen omezený okruh výzkumných otázek.

Nejdůležitějším poznatkem je, že nemůžeme změřit efektivitu delších trestů odnětí svobody s alternativními tresty (Villetaz et al. 2015, p.54). Experiment, který by přiřazoval

vrahům náhodně tresty, by byl eticky nepřijatelný z morálního i společenského hlediska. Soudci nesoudí natolik odlišně, aby pachatelé neobdrželi za nejzávažnější kriminalitu odlišné typy trestů, a charakteristiky pachatelů takovýchto trestných činů jsou tak odlišné, že nelze řádně kvaziexperimentálně „sestavit“ kontrolní skupinu pro pachatele vraždy.

Cílem podobných studií tak je primárně odpovědět na otázku, zda je na místě využívat institutu nepodmíněného trestu odnětí svobody více či méně, ale nikoli na otázku, zda je na místě jej využívat či ne. Možnosti aplikace výsledků tohoto výzkumu nejsou malé: Výzkum může naznačit, zda je z hlediska recidivy efektivnější používat nepodmíněné či alternativní tresty (popř. které alternativní tresty) pro velkou část pachatelů, zda je efektivnější ukládat kratší či delší tresty (pro ty, kterým je uložen nepodmíněný trest) či jaký vliv na recidivu má podmíněné propuštění. Z hlediska soudce, státního zástupce i zákonodárce jsou přesně tyto otázky, které si klade: Měli bychom posílat více či méně pachatelů do vězení? Pokud ano, tak na jak dlouho; pokud ne, tak jaký trest bychom jim měli uložit? A je na místě podmíněčně propustit více či méně vězňů, popřípadě za jakých podmínek?²³

Odpovědět na tyto otázky je potřebné zejména v situaci, kdy v České republice mají soudci silnou diskreci a existují výrazné rozdíly při ukládání nepodmíněných trestů odnětí svobody (Drápal 2018); soudci tak mohou výrazněji upravit svoje rozhodování v souladu s výsledky empirických studií. Výsledky tohoto výzkumu tak mohou reálně změnit, nakolik soudci ukládají nepodmíněné tresty odnětí svobody, a to alespoň u té části soudců, kteří rozhodují způsobem, který takový výzkum označí jako neefektivní.

Rozdíly při ukládání trestů přitom paradoxně umožňují analýzu efektivity trestu odnětí svobody. Výše zmíněný přirozený experiment by bez něj nebyl možný, protože jeho princip spočívá v odlišné přísnosti soudců. Vyvažování a případně i regrese je také do velké míry založená na předpokladu, že podobní pachatelé jsou trestáni různě (Mears et al. 2012). Nieuwbeerta, Nagin a Blokland (2009, p.246) to popisují tak, že „‘propensity score matching’ je uskutečnitelné pouze v situacích, ve kterých soudci mají diskreční pravomoc při rozhodování o tom, s kým bude jak zacházeno. Pokud je výběr do určité intervence extrémní, např. pokud všechny osoby s určitou charakteristikou podstoupí danou intervenci, pak není možné použít metodu vyvažování.“ Právě nedokonalý přístup soudů (který je ale přirozený a pokud nepřekračuje určitou míru, tak není problematický) nám tak umožňuje zkoumat efektivitu nepodmíněného trestu odnětí svobody. Institucionalizace této nedokonalosti (nárazové náhodné přiřazení trestů v pečlivě vybraných případech) by nám pak umožnila studovat výrazně spolehlivěji efektivitu trestů

Měření efektivity trestu odnětí svobody bude vždy do určité míry nedokonalé. Prvním nedostatkem je skutečnost, že jednodušeji lze měřit pouze dopad na recidivu u nepodmíněného trestu odnětí svobody, který byl uložen soudem při rozhodnutí o vině a trestu. Nepočítá se tak s pachateli, kterým bylo nařízené vykonat původně uložené podmíněné tresty, nebo s těmi, kterým soudy přeměnily alternativní tresty v nepodmíněný trest odnětí svobody. Bohužel v České republice nemáme jednoduše dostupná data o počtu nařízených či přeměněných trestech na individuální úrovni, které by bylo možné využít k potřebným analýzám. Všechny výsledky (vyjma analýzy efektivity podmíněného propuštění) tak nelze vztáhnout na podmíněné tresty odnětí svobody či přeměněné alternativní tresty. Stejně tak se jako protiklad nepodmíněného trestu odnětí svobody musí

²³ Toto se vztahuje (jako tento celý článek) na efektivitu vztahující se k specifickému odrazení (deterenci), tedy k budoucímu vlivu na jednotlivce, který podstupuje určitou intervenci (trest). Nevztahuje se na generální odrazení skrze vyslovené tresty; toto je výrazně náročnější na studování, viz např. Paternoster (2010).

brát podmíněný trest odnětí svobody. Tento přístup není však v Evropě vzácný: Kvůli nedostatku (jiných) alternativních trestů byla stejná metodologie použita např. ve Španělsku (Cid 2009). Určitým omezením může být i skutečnost, že běžně se zkoumají skupiny rozdělené podle "intent to treat" (rozhodnutí o intervenci, např. uložení trestu), nikoli podle toho, zda je intervence reálně provedená. Pokud existují významné rozdíly mezi uloženými tresty a reálně vykonanými tresty, tak to může omezit interpretovatelnost výsledků.

Na závěr je pak třeba zmínit několik omezení, které se obecně vztahují i ke všemu společenskovědnímu výzkumu. I pokud nám experimenty (popř. jiné metody) nabídnou určitý náhled ohledně reality, je to skrze do určité míry zamlžené sklo (Shadish et al. 2002, p.29); důvodem může být, že jsme schopni měřit pouze minulost a to do výrazné míry konkretizovanou podle výběru trestů a pachatelů (Shadish et al. 2002, p.20), nebo že se vyjadřujeme pouze k vlivu intervence na určitou skupinu, nikoli na jednotlivce (Smith 2013). Výzkumníci doufají, že budoucnost bude obdobná, pokud se výrazně nezmění podmínky – to je ale primárně jejich předpoklad, přání a v některých případech zkušenost s předchozími výzkumy. Znamená to ale, že podrobná měření by se měla opakovat minimálně co deset let (např. jako sčítání lidu). V takovém případě by se empirické závěry opíraly o výzkum provedený v blízké minulosti a případný dopad změn podmínek, které by mohly znehodnotit výsledky, by se zmenšil. Na závěr je pak třeba zmínit, že nelze činit definitivní závěry na základě jedné studie, ať je jakkoli excelentní (Shadish et al. 2002, p.486,496).

Závěr

Úvahy o metodách, které je vhodné v daném kontextu použít, jsou ovlivněné otázkou spolehlivosti dat a mírou spolehlivosti zjištění kauzality (interní validity). V tomto článku jsem popisoval různé přístupy k měření efektivity trestu odnětí svobody, přičemž jsem zdůrazňoval nutnost správného vytváření experimentálních a kontrolních skupin, která jediná zajistí možnost skutečně analyzovat vliv pouze zkoumané intervence. Nejspolehlivější metodologií je experiment, tedy výzkumníkem řízené náhodné přiřazování pachatelů do experimentální či kontrolní skupiny. Výsledkem takového přiřazování je, že při použití dostatečného počtu pachatelů můžeme tyto skupiny považovat za totožné. Výše jsem naznačil podmínky, za kterých by bylo etické a pro společnost i pachatele prospěšné náhodně přiřadit různé tresty skupinám pachatelů. Pouze tento přístup zajistí, že skutečně poznáme, jaký vliv má nepodmíněný trest odnětí svobody na recidivu; žádný jiný přístup nám toto tak spolehlivě nezajistí.

Není-li možné provést náhodné přiřazování trestů pachatelům, je možné použít náhodné přiřazování pachatelů soudcům, které neuskuteční výzkumník, ale který provedl soud. Přirozený experiment vzniká za situace, kdy jsou případy náhodně přiřazeny jednotlivým soudcům (vytvoří se tedy stejné skupiny), kteří jsou odlišně přísní a posílají tak jiný podíl pachatelů nepodmíněně do vězení. Stejný experiment vzniká za obdobných podmínek i při rozhodování o podmíněném propuštění. Byť se jedná o jednu z nejspolehlivějších metod, tato byla použita pouze vzácně z důvodu nedostatku dat a z důvodu odlišně nastavených právních rámců, kdy mnohé státy neoperují s (kvazi)náhodným přiřazováním případů soudcům. Byť tyto metody nabízejí vysoce spolehlivé výsledky za relativně málo snahy, je třeba ji blíže prozkoumávat vzhledem k předpokladům jejího vzniku a nepřiliš častému užití ve světových výzkumech. Přirozený experiment pak vznikl i v důsledku amnestie z roku 2013; ten nám umožní studovat vliv různých délek kratších

nepodmíněných trestů odnětí svobody na recidivu. Spolehlivé výsledky nabízí i metoda regresní diskontinuita, která využívá skutečnosti, že pachatelé odsouzeni těsně před a po vydání amnestie jsou takřka totožní.

Pokud výzkumník nemůže provést experiment a nemůže plně použít přirozený experiment, může experiment zpětně simulovat: Použije kvaziexperimentální metody. Ty se snaží na základě charakteristik pachatelů vytvořit podobné skupiny, které následně porovnají. Jejich nevýhodou je ale skutečnost, že nikdy nemohou zjistit všechny relevantní údaje o pachatelích a nikdy tedy nemůžou dojít k plně spolehlivému závěru o tom, že to byla právě daná intervence (a nikoli nezaznamenaná charakteristika pachatele), která měla vliv na výsledek.

Tento článek tak nelze zakončit jinak než společně s již generacemi kriminologů zvolat, že potřebujeme experimenty i v této oblasti: Potřebujeme v souladu s etickými zásadami náhodně přiřazovat tresty pachatelům trestných činů, abychom zjistili, co můžeme reálně očekávat od různých trestů a jak tomuto přizpůsobit náš trestní systém. V opačném případě si trouláme rozhodovat o životech desítek až stovek tisíc lidí v České republice a miliónů lidí po celém světě, aniž bychom měli pořádné argumenty podkládající naše závěry o (ne)efektivitě nepodmíněného trestu odnětí svobody.

Použitá literatura:

- Anderson, J.M., Kling, J.R. & Stith, K., 1999. Measuring Inter-Judge Sentencing Disparity Before and After the Federal Sentencing Guidelines. *Journal of Service Research*, 42(1), pp.271–307.
- Angrist, J.D., Imbens, G.W. & Rubin, D.B., 1996. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434), pp.444–455.
- Angrist, J.D. & Pischke, J.-S., 2015. *Mastering “Metrics. The Path from Cause to Effect,”* Princeton and Oxford: Princeton University Press.
- Apel, R. & Sweeten, G., 2010. Propensity Score Matching in Criminology and Criminal Justice. In A. R. Piquero & D. Weisburd, eds. *Handbook of Quantitative Criminology*. New York: Springer, pp. 543–562.
- Bachman, R. & Schutt, R.K., 2011. *The Practice of Research in Criminology and Criminal Justice* 4th ed., Thousand Oaks, CA: Sage.
- Baron, J.H., 2009. Sailors’ scurvy before and after James Lind - a reassessment. *Nutrition Reviews*, 67(6), pp.315–332.
- Berk, R.A., 2013. Recent Perspectives on the Regression Discontinuity Design. In A. R. Piquero & D. Weisburd, eds. *Handbook of Quantitative Criminology*. Springer, pp. 563–580.
- Berube, D. & Green, D.P., 2007. The Effects of Sentencing on Recidivism: Results from a Natural Experiment. *Second Annual Conference on Empirical Legal Studies*.
- Bhuller, M. et al., 2016. *Incarceration, Recidivism and Employment*, Cambridge, MA.
- Blamey, A. & Mackenzie, M., 2007. Theories of Change and Realistic Evaluation. *Evaluation*, 13(4), pp.439–455.
- Braga, A.A., 2013. Quasi-Experimentation When Random Assignment Is Not Possible. In B. C. Welsh, A. A. Braga, & G. J. N. Bruinsma, eds. *Experimental Criminology: Prospects for Advancing Science and Public Policy*. New York: Cambridge University Press, pp. 223–250.
- Braga, A.A., Welsh, B.C. & Bruinsma, G.J.N., 2013. Integrating Experimental and Observational Methods to Improve Criminology and Criminal Justice Policy. In B. C. Welsh, A. A. Braga, & G. J. N. Bruinsma, eds. *Experimental Criminology: Prospects for Advancing Science and Public Policy*. New York: Cambridge University Press, pp. 277–297.
- Burchett, W. & Ellis, A., 2017. Package “npmv.” Available at: <https://cran.r-project.org/web/packages/npmv/npmv.pdf>.

- Burchett, W.W. et al., 2017. Nonparametric Inference for Multivariate Data: The R Package nrmv. *Journal of Statistical Software*, 76(4).
- Cid, J., 2009. Is Imprisonment Criminogenic?: A Comparative Study of Recidivism Rates between Prison and Suspended Prison Sanctions. *European Journal of Criminology*, 6(6), pp.459–480.
- Cohen, T.R., Panter, A.T. & Turan, N., 2012. Guilt Proneness and Moral Character. *Current Directions in Psychological Science*, 21(5), pp.355–359.
- Disman, M., 2011. *Jak se vyrábí sociologická znalost*, Praha: Karolinum.
- Drago, F., Galbiati, R. & Vertova, P., 2009. The Deterrent Effects of Prison: Evidence from a Natural Experiment. *Journal of Political Economy*, 117(2), pp.257–280.
- Drápal, J., 2017. *Recidiva jako nebinární ukazatel: zahrnutí závažnosti a frekvence trestné činnosti*, Available at: <http://svoc.prf.cuni.cz>.
- Drápal, J., 2018. Sentencing Disparities in the Czech Republic: Empirical Evidence from Post-Communist Europe. *European Journal of Criminology*, forthcoming.
- Fagan, J., 1990. Natural Experiments in Criminal Justice. In K. L. Kempf, ed. *Measurement Issues in Criminology*. New York: Springer-Verlag, pp. 108–137.
- Farrington, D.P. et al., 2002. Maryland Scientific Methods Scale. In L. Sherman & D. P. Farrington, eds. *Evidence-Based Crime Prevention*. New York: Routledge, pp. 13–21.
- Farrington, D.P., 2003. Methodological Quality Standards for Evaluation Research. *The ANNALS of the American Academy of Political and Social Science*, 587(1), pp.49–68.
- Farrington, D.P., 1983. Randomized Experiments on Crime and Justice. *Crime and Justice*, 4, pp.257–308.
- Farrington, D.P. & Welsh, B.C., 2006. A Half Century of Randomized Experiments on Crime and Justice. *Crime and Justice*, 34(1), pp.55–132.
- Farrington, D.P. & Welsh, B.C., 2005. Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1(1), pp.9–38.
- Gill, C.E., 2011. Missing links: how descriptive validity impacts the policy relevance of randomized controlled trials in criminology. *Journal of Experimental Criminology*, 7(3), pp.201–224.
- Green, D.P. & Winik, D., 2010. Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. *Criminology*, 48(2), pp.357–387.

- Gřivna, T. & Říha, J., 2013. Amnestie prezidenta republiky ze dne 1. 1. 2013 vyhlášená pod č. 1/2013 Sb. *Trestněprávní revue*, (2), pp.29–42.
- Hodgson, J. & Soubis, L., 2016. Understanding the Sentencing Process in France. *Crime and Justice*, 45(1), pp.221–265.
- Hough, M., 2010. Gold standard or fool's gold? The pursuit of certainty in experimental criminology. *Criminology & Criminal Justice*, 10(1), pp.11–22. Available at: <http://journals.sagepub.com/doi/10.1177/1748895809352597>.
- Imbens, G.W. & Angrist, J.D., 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), pp.467–475.
- Killias, M. et al., 2010. How damaging is imprisonment in the long-term? A controlled experiment comparing long-term effects of community service and short custodial sentences on re-offending and social integration. *Journal of Experimental Criminology*, 6(2), pp.115–130.
- Killias, M., Aebi, M. & Ribeaud, D., 2000. Does Community Service Rehabilitate better than Short-term Imprisonment?: Results of a Controlled Experiment. *The Howard Journal of Criminal Justice*, 39(1), pp.40–57.
- Lata, J., 2007. *Účel a smysl trestu*, Praha: LexisNexis CZ.
- Loeffler, C.E., 2013. Does imprisonment alter the life course? Evidence from a natural experiment. *Criminology*, 51(1), pp.137–166. Available at: <http://doi.wiley.com/10.1111/1745-9125.12000>.
- Martin, S.E., Annan, S. & Forst, B., 1993. The special deterrent effects of a jail sanction on first-time drunk drivers: A quasi-experimental study. *Accident Analysis & Prevention*, 25(5), pp.561–568.
- McGloin, J.M. & Thomas, K.J., 2013. Experimental Tests of Criminological Theories. In B. C. Welsh, A. A. Braga, & J. N. Gerben, eds. *Experimental Criminology: Prospects for Advancing Science and Public Policy*. New York: Cambridge University Press, pp. 15–42.
- McGrath, A. & Weatherburn, D., 2012. The effect of custodial penalties on juvenile reoffending. *Australian & New Zealand Journal of Criminology*, 45(1), pp.26–44.
- Mears, D.P., Cochran, J.C. & Bales, W.D., 2012. Gender differences in the effects of prison on recidivism. *Journal of Criminal Justice*, 40(5), pp.370–378.
- Nagin, D.S., Cullen, F.T. & Jonson, C.L., 2009. Imprisonment and Reoffending. *Crime and Justice*, 38(1), pp.115–200.

- Nagin, D.S. & Snodgrass, G.M., 2013. The Effect of Incarceration on Re-Offending: Evidence from a Natural Experiment in Pennsylvania. *Journal of Quantitative Criminology*, 29(4), pp.601–642.
- Nieuwebeerta, P., Nagin, D.S. & Blokland, A.A.J., 2009. Assessing the Impact of First-Time Imprisonment on Offenders' Subsequent Criminal Career Development: A Matched Samples Comparison. *Journal of Quantitative Criminology*, 25(3), pp.227–257.
- Novotný, O., 1969. *O trestu a vězeňství* 2. vydání., Praha: Academia.
- Owens, E.G. & Ludwig, J., 2013. Using Regression Discontinuity Design in Crime Research. In B. C. Welsh, A. A. Braga, & G. J. N. Bruinsma, eds. *Experimental Criminology: Prospects for Advancing Science and Public Policy*. New York: Cambridge University Press, pp. 194–222.
- Padfield, N., 2016. Reflections on sentencing in England and Wales. In A. Kapardis & D. P. Farrington, eds. *The Psychology of Crime, Policing and Courts*. London and New York: Routledge, pp. 231–258.
- Paternoster, R., 2010. How Much Do We Really Know about Criminal Deterrence. *Journal of Criminal Law & Criminology*, 100(3), pp.765–823.
- Pawson, R. & Tilley, N., 1997. *Realistic Evaluation*, London: SAGE.
- Perry, A.E., Weisburd, D. & Hewitt, C., 2010. Are criminologists describing randomized controlled trials in ways that allow us to assess them? Findings from a sample of crime and justice trials. *Journal of Experimental Criminology*, 6(3), pp.245–262.
- Pokorný, M., 2017. Právnicka trávil nadřizenou projímadlem. Litovala jen sebe, tři roky vězení platí, potvrdil soud. *Aktualne.cz*. Available at: <https://zpravy.aktualne.cz/domaci/pravnicka-travila-svoji-nadrizenou-projmadlem-tri-roky-veze/r~7c4277ea445211e7bdee002590604f2e/>.
- Sampson, R.J., 2010. Gold Standard Myths: Observations on the Experimental Turn in Quantitative Criminology. *Journal of Quantitative Criminology*, 26(4), pp.489–500.
- Schneider, A.L., 1986. Restitution and recidivism rates of juvenile offenders: Results from four experimental studies. *Criminology*, 24(3), pp.533–552.
- Scott, R.W., 2010. Inter-Judge Sentencing Disparity After “Booker”: A First Look. *Stanford Law Review*, 63(1), pp.1–66.
- Shadish, W.R., Cook, T.D. & Campbell, D.T., 2002. *Experimental and Quasi-Experimental Design for Generalized Causal Inference*, Boston, MA: Houghton Mifflin Company.
- Sherman, L.W. et al., 1998. *Preventing Crime: What Works, What Doesn't, What's Promising*, Available at: <https://www.ncjrs.gov/works/>.

- Sherman, L.W., 2007. The power few: experimental criminology and the reduction of harm. *Journal of Experimental Criminology*, 3(4), pp.299–321.
- Sherman, L.W. & Strang, H., 2004. Verdicts or Inventions? *American Behavioral Scientist*, 47(5), pp.575–607.
- Smith, H.L., 2013. Research Design: Toward a Realistic Role for Causal Analysis. In S. L. Morgan, ed. *Handbook of Causal Analysis for Social Research*. Springer, pp. 45–73.
- Solnař, V. & Vanduchová, M., 2009. *Systém českého trestního práva: Tresty a ochranná opatření*, Praha: Novatrix.
- Tangney, J.P., Stuewig, J. & Martinez, A.G., 2014. Two Faces of Shame: The Roles of Shame and Guilt in Predicting Recidivism. *Psychological Science*, 25(3), pp.799–805.
- Tilley, N., 2009. Sherman vs Sherman: Realism vs rhetoric. *Criminology and Criminal Justice*, 9(2), pp.135–144.
- Tonry, M., 1997. Ethnicity, Crime, and Immigration. *Crime and Justice*, 21, pp.1–29.
- Villettaz, P., Gillieron, G. & Killias, M., 2015. *The Effects on Re-offending of Custodial vs. Non- custodial Sanctions: An Updated Systematic Review of the State of Knowledge*, Available at:
<http://www.campbellcollaboration.org/library/the-effects-on-re-offending-of-custodial-vs-non-custodial-sanctions-an-updated-systematic-review-of-the-state-of-knowledge.html>.
- Warne, R.T., 2014. A Primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists. *Practical Assessment, Research and Evaluation*, 19(17), pp.1–10.
- Weisburd, D., 2003. Ethical Practice and Evaluation of Interventions in Crime and Justice. *Evaluation Review*, 27(3), pp.336–354.
- Weisburd, D., 2010. Justifying the use of non-experimental methods and disqualifying the use of randomized controlled trials: challenging folklore in evaluation research in crime and justice. *Journal of Experimental Criminology*, 6(2), pp.209–227.
- Weisburd, D., Lum, C.M. & Petrosino, A., 2001. Does Research Design Affect Study Outcomes in Criminal Justice? *The ANNALS of the American Academy of Political and Social Science*, 578(1), pp.50–70.
- Welsh, B.C., Braga, A.A. & Bruinsma, G.J.N., 2013. Experimenting with Crime and Criminal Justice. In B. C. Welsh, A. A. Braga, & G. N. J. Bruinsma, eds. *Experimental Criminology: Prospects for Advancing Science and Public Policy*. New York: Cambridge University Press, pp. 1–11.

Welsh, B.C., van der Laan, P.H. & Hollis, M.E., 2013. Systematic Reviews and Cost-Benefit Analyses. In B. C. Welsh, A. A. Braga, & G. J. N. Bruinsma, eds. *Experimental Criminology: Prospects for Advancing Science and Public Policy*. New York: Cambridge University Press, pp. 253–276.

van der Werff, C., 1981. Recidivism and Special Deterrence. *British Journal of Criminology*, 21(2), pp.136–147.