**Příloha č. 5: Podrobné výsledky kontrafaktuální analýzy**

## Data pro kontrafaktuální analýzu

Data pro kontrafaktuální analýzu byla získána z následujících zdrojů:

* Účetní uzávěrky analyzovaných podniků, agregováno společností BisNode (nyní Dun & Bradstreet Czech Republic); data stahována v srpnu, září a říjnu 2021 (databáze pro finální analýzu k říjnu 2021)
* Registr ekonomických subjektů – využíván průběžně pro základní informace o ekonomických subjektech v analýze (NACE, kategorie velikosti, kraj, subjekt v likvidaci, atd.). Do kontrafaktuální analýzy vstoupila data k říjnu 2021
* Insolvenční rejstřík – data stahována v srpnu, září a říjnu 2021; do finální analýzy vstoupila data k říjnu 2021
* Rejstřík nespolehlivých plátců DPH – data stahována v srpnu, září a říjnu 2021; do finální analýzy vstoupila data k říjnu 2021

## Konstrukce kontrolního vzorku

Kontrolní vzorek byl konstruován zvlášť pro subjekty, které získaly podporu před rokem 2020 a pro subjekty podpořené v roce 2020 a později. Jelikož se charakter problému párování může zásadně lišit pro tyto dvě skupiny, byly uvažovány dva oddělené modely na základě nichž pak bylo provedeno finální spárování. Pro první zmíněnou skupiny byly subjekty párovány na základě dat z roku 2016, pro druhou byly využity informace z roku 2019.

## Popis PSM modelu

Kontrolní vzorek byl vytvořen jako podmnožina z 10525 náhodně vybraných subjektů pomocí metody **Propensity score matching (PSM).**

**Proměnné uvažované pro modelování PSM:**

* NACE kód převedený na písmena, viz tabulka <http://www.nace.cz/>,
* kraj: 14 kategorii,
* typ fyzické osoby, 3 kategorie (obchodní, služby, výrobní),
* počet zaměstnanců – modelování pomocí kategoriální proměnné s 10 kategoriemi:
  + do 5, 6-9, 10-19, 20-24,25-49,50-99,100- 199,200-249,250-499, 500 a vice,
* rok vzniku,
* celková aktiva (AKTIVACELK),
* výkony (VYKONY),
* míza zadlužení (MIRAZADLU) převedená na kategoriální proměnnou s 6 kategoriemi
  + [-Inf,25) [25,50) [50,75) [75,100) [100, Inf),
* roční obrat (OBRAT) – modelováno pomocí logaritmu OBRAT,

Při párování subjektů podpořených v roce 2020 byl každý subjekt zapojený do projektu ANTIVIRUS spárován s nepodpořeným subjektem, který byl taktéž zapojen do programu ANTIVIRUS. Podobně, subjekty nezapojené do ANTIVIRU byly spárovány se subjekty, které také ANTIVIRUS nečerpaly. Párovací model byl pro obě skupiny společný.

Informace o výše uvedených proměnných však nebyly dostupné pro všechny subjekty, jelikož data obsahují řadu chybějících hodnot, a to jak mezi podpořenými tak i nepodpořenými. Např. v letech 2017-2019 bylo podpořeno 505 podniků, z nichž však nemáme dostupnou informaci o jejich výkonu u 133 podniků. Takto došlo k redukci vzorků uvažovaných pro párování, a to v počtech uvedených níže:

Popis vzorků

* za roky 2017-2019
  + 505 podpořených, z nich 369 s dostupnými informacemi o relevantních proměnných potřebných v PSM modelu (uvedených v seznamu výše),
  + 7134 nepodpořených s dostupnou informací,
* od roku 2020
  + 1694 podpořených,
    - z nich 144 již dříve v OP PIK, 858 zapojeno do programu ANTIVIRUS, 836 nezapojeno
    - pro 1349 máme dostatečné dostupné informace o proměnných potřebných v PSM modelu,
    - z nich 696 zapojeno do programu ANTIVIRUS, 653 nezapojeno
  + 7601 nepodpořených s dostupnou informací
    - z nich 2252 zapojeno do programu ANTIVIRUS, 5349 nezapojeno

## Technické detaily

Párování bylo provedeno pro následující volby:

* Glm logit přístup.

Byly zkoušeny i jiné linkové funkce (např. probit), nicméně ty vedly k horším výsledkům, a to zejména s ohledem na diagnostické testy (Kolmogorovův Smirnovův test, t test; podrobnosti viz níže) porovnávající shodu rozdělení klíčových proměnných v podpořeném a kontrolním vzorku.

* Greedy nearest neighbor matching.

Zkoušeny byly i jiné přístupy, ovšem s horším výsledkem, jelikož pro tyto postupy vyšel v následné verifikaci signifikantní rozdíl mezi rozdělení klíčových proměnných v podpořené a kontrolní skupině.

* Každý podpořený spárován s jedním nepodpořeným.

Kategorizace pro některé proměnné lépe umožňuje nelinearitu závislosti PS na dané proměnné. Tento přístup přinesl mnohem lepší výsledky než nekategoriální přístup. V modelu byly uvažovány (zkoušeny) i různé interakce, nicméně nevedlo to k významnému zlepšení modelu (z hlediska statistické významnosti těchto efektů i z hlediska následné diagnostiky). Proto se ve výsledném modelu vyskytují výše uvedené proměnné jen jako hlavní efekty.

V rámci konstrukce PSM byla uvažována i celá řada jiných modelů (podmodelů, nadmodelů, s transformovanými proměnnými, interakcemi apod.), které však vedly k neuspokojivému porovnání hlavních charakteristik vzorků (především ekonomické ukazatele) z hlediska diagnostických testů uvedených v následujícím odstavci.

## Výsledek párování

Na základě výsledného PSM modelu se podařilo spárovat každý podpořený subjekt s jedním nepodpořeným. Následně byly porovnány charakteristiky obou vzorků. Spojité veličiny byly porovnány pomocí Kolmogorovova Smirnovova testu a pomocí dvouvýběrového t-testu. Kategoriální veličiny byly porovnány pomocí chí kvadrát testu. Dále byly při diagnostice využity grafické nástroje, a to graf standardizovaných rozdílů v průměrech před a po spárování, grafy rozdělení klíčových charakteristik a PMS skóre před a po spárování (empirické distribuční funkce a jádrové odhady hustot pro spojité proměnné, sloupcové grafy pro kategoriální proměnné).

Jelikož je samotná evaluace párování spíše technického charakteru, jsou příslušné výsledky uvedeny v samostatné Příloze A, aby neodváděly pozornost od výsledků hlavní statistické analýzy. Zde pouze konstatujeme, že na základě všech těchto diagnostik bylo párování vyhodnoceno jako **úspěšné** a následná statistická analýza by neměla být zatížena vychýlením.

Výsledné párování:

* vzorek 369 podpořených subjektů zapojených do podpory před rokem 2020 a k němu kontrolní vzorek 369 nepodpořených subjektů,
* vzorek 1349 podpořených subjektů zapojených do OPPIK v roce 2020 a k němu stejně velký vzorek kontrolních subjektů.

# Míra přežití

Pro míru přežití sledujeme následující ukazatele

* subjekt v likvidaci v roce 2021,
* subjekt byl zveřejněn jako nespolehlivý do 12/19,
* subjekt byl/je zveřejněn jako nespolehlivý do 10/21,
* subjekt má nějaký záznam v rejstříku ISIR.

Většina sledovaných ukazatelů má velmi malé zastoupení (velmi často nulové), aby bylo možné rozdíl statisticky testovat, viz Tabulky 1 a 2. Ve spárovaných vzorcích navíc nemáme žádné subjekty, které by zaznamenaly první tři události.

Porovnali jsme počet záznamů v ISIR u podpořených a nepodpořených, a to pomocí chí kvadrát testu nezávislosti. Test je nevýznamný jak pro projekty podpořené do roku 2020 tak i pro ty podpořené v roce 2020. Pro podpořené v roce 2020 byl proveden i test podmíněné nezávislosti pomocí Cohranova-Mantelova testu, který testuje podmíněnou nezávislost při dané úrovni proměnné ANTIVIRUS. Také byly provedeny chí kvadrát testy nezávislosti pro obě skupiny Antiviru zvlášť. **Všechny zmíněné testy vychází statisticky nevýznamné. Nelze tedy prokázat vliv podpory na výskyt negativních událostí souvisejících s přežitím firmy.**

V následující Tabulce 1 jsou uvedeny odhady rozdílů v pravděpodobnostech událostí (kontrolní mínus podpoření). Kladné hodnoty odpovídají situaci, kdy je pravděpodobnost negativní události vyšší pro nepodpořené subjekty (tj. lze pozorovat kladný vliv podpory). Pro úplnost byla uvažována porovnání i pro všechny dostupné subjekty (nejen ty spárované). **U většiny případů nelze z důvodu malého množství události vyhodnotit statistickou významnost.**

Pro úplnost bylo porovnání provedeno nejen pro spárované vzorky, ale i pro všechna dostupná data. Zde vychází **pro podpořené v roce 2020 statisticky významný rozdíl u subjektů nezapojených do programu ANTIVIRUS**, a to tak, že pravděpodobnost záznamu je vyšší o 0,34 % pro podpořené**. Nicméně vzhledem k tomu, že tyto dva nespárované vzorky se poměrně dost liší v klíčových ukazatelích** (viz Evaluace párování v příloze)**,** není vhodné z tohoto porovnání nic zásadního usuzovat.

**Obecně tedy spíše nelze tedy prokázat, že by podpora programu OPPIK statisticky významně ovlivnila přežití firem.** **Důvodem je především nedostatečný počet zaznamenaných případů.** Je však možné, že vliv se projeví až v delším časovém horizontu.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Bez rozlišení ANTIVIRUS | | ANTIVIRUS=0 | | ANTIVIRUS=1 | |
|  | Rozdíl C-T | p-hodnota | Rozdíl C-T | p-hodnota | Rozdíl C-T | p-hodnota |
| **Pravděpodobnost likvidity 2021** |  |  |  |  |  |  |
| všichni | 0,09% | NA | 0,10% | NA | 0,07% | NA |
| spárované vzorky pro podporu před 2020 | 0,00% | NA | 0,00% | NA | 0,00% | NA |
| spárované vzorky pro podporu v 2020 | 0,00% | NA | 0,00% | NA | 0,00% | NA |
| **Pravděpodobnost nespolehlivosti do 2019** |  |  |  |  |  |  |
| všichni | -0,08% | NA | -0,04% | NA | -0,12% | NA |
| spárované vzorky pro podporu před 2020 | 0,00% | NA | 0,00% | NA | 0,00% | NA |
| spárované vzorky pro podporu v 2020 | 0,00% | NA | 0,00% | NA | 0,00% | NA |
| **Pravděpodobnost nespolehlivosti do 2021** |  |  |  |  |  |  |
| všichni | -0,13% | NA | -0,13% | NA | -0,12% | NA |
| spárované vzorky pro podporu před 2020 | 0,00% | NA | 0,00% | NA | 0,00% | NA |
| spárované vzorky pro podporu v 2020 | 0,00% | NA | 0,00% | NA | 0,00% | NA |
| **Pravděpodobnost záznamu v ISIR** |  |  |  |  |  |  |
| všichni | -0,12% | 0,2763 | -0,34% | 0,02704 | 0,10% | 0,6151 |
| spárované vzorky pro podporu před 2020 | 0,27% | NA | -0,13% | NA | 0,00% | NA |
| spárované vzorky pro podporu v 2020 | 0,00% | 1 | 0,00% | NA | 0,00% | NA |

Tabulka : Ukazatele míry přežití - rozdíl mezi pravděpodobností v kontrolní skupině a v podpořené skupině. Kladné hodnoty odpovídají situaci, kde je pravděpodobnost výskytu indentifikátoru vyšší u kontrolních.

Přesné počty případů jednotlivých událostí uvádí tabulka na následující straně.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | V RES 2019 | V likvidaci 2019 |  | V RES2021 | V likvidaci 2021 | | Nespolehlivost do konce 2019 | | Nespolehlivost do 2021 | | | Záznam v ISIR | |
|  |  | Počet | Počet | Počet | % | Počet | Počet | % | Počet | % | | Počet | % | Počet | % |
| **Všichni podpoření za celou dobu programu** |  | 5153 | 5092 | 0 | 0,00% | 5153 | 1 | 0,02% | 4 | 0,08% | | 7 | 0,14% | 22 | 0,43% |
|  | Z toho ANTIVIRUS=1 | 2421 | 2416 | 0 | 0,00% | 2421 | 0 | 0,00% | 3 | 0,12% | | 3 | 0,12% | 4 | 0,17% |
|  | Z toho ANTIVIRUS=0 | 2732 | 2676 | 0 | 0,00% | 2732 | 1 | 0,04% | 1 | 0,04% | | 4 | 0,15% | 18 | 0,66% |
| Všichni nepodpoření |  | 10525 | 10524 | 3 | 0,03% | 10525 | 12 | 0,11% | 0 | 0,00% | | 1 | 0,01% | 32 | 0,30% |
|  | Z toho ANTIVIRUS=1 | 2994 | 2994 | 0 | 0,00% | 2994 | 2 | 0,07% | 0 | 0,00% | | 0 | 0,00% | 8 | 0,27% |
|  | Z toho ANTIVIRUS=0 | 7531 | 7530 | 3 | 0,04% | 7531 | 10 | 0,13% | 0 | 0,00% | | 1 | 0,01% | 24 | 0,32% |
| **Podpoření před rokem 2020 (spárovaný vzorek)** |  | 369 | 369 | 0 | 0,00% | 369 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 1 | 0,27% |
|  | Z toho ANTIVIRUS=1 | 166 | 166 | 0 | 0,00% | 166 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 0 | 0,00% |
|  | Z toho ANTIVIRUS=0 | 203 | 203 | 0 | 0,00% | 203 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 2 | 0,99% |
| Kontrolní spárovaný vzorek |  | 369 | 369 | 0 | 0,00% | 369 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 2 | 0,54% |
|  | Z toho ANTIVIRUS=1 | 136 | 136 | 0 | 0,00% | 136 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 0 | 0,00% |
|  | Z toho ANTIVIRUS=0 | 233 | 233 | 0 | 0,00% | 233 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 2 | 0,86% |
| **Podpoření 2020 (spárovaný vzorek)** |  | 1349 | 1349 | 0 | 0,00% | 1349 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 7 | 0,52% |
|  | Z toho ANTIVIRUS=1 | 696 | 696 | 0 | 0,00% | 696 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 3 | 0,43% |
|  | Z toho ANTIVIRUS=0 | 653 | 653 | 0 | 0,00% | 653 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 4 | 0,61% |
| Kontrolní spárovaný vzorek |  | 1349 | 1349 | 0 | 0,00% | 1349 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 7 | 0,52% |
|  | Z toho ANTIVIRUS=1 | 696 | 696 | 0 | 0,00% | 696 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 3 | 0,43% |
|  | Z toho ANTIVIRUS=0 | 653 | 653 | 0 | 0,00% | 653 | 0 | 0,00% | 0 | 0,00% | | 0 | 0,00% | 4 | 0,61% |

Tabulka : Podrobné četnosti ukazatelů míry přežití.

# Zaměstnanost

Pro programy podpořené před rokem 2020 byly analyzovány údaje o zaměstnanosti od roku 2016 do roku 2019. Data z roku 2020 byla chybějící pro příliš mnoho subjektů a navíc v tomto období subjekty vykazovaly jiné trendy než před rokem 2020, což by mohlo způsobit zavádějící a špatné závěry.

Pro podpořené v roce 2020 bylo provedeno porovnání počtu zaměstnanců z roku 2019 a 2020. Bohužel i v tomto vzorku je informace o zaměstnanosti v roce 2020 velmi často chybějící, což bylo vyřešeno analýzou kategoriální proměnné o zaměstnanosti, která byla známa pro výrazně vyšší podíl podniků.

Dále se zdají být problematické i některé nuly v datech. Obrázky Obrázek 1 a Obrázek 2 ukazují náhodně vybrané subjekty a jejich zaměstnanosti. Tyto grafy naznačují, že některé nuly mohou být chybné hodnoty (namísto nich by měl být chybějící údaj). V další analýze však tento aspekt není vzat v úvahu.

Chart, line chart

Description automatically generated

Obrázek 1: Náhodně vybrané podpořené subjekty (data z let, kdy jsou již podpořeny, jsou znázorněny plným kolečkem a nepřerušovanou čarou).

Chart, line chart

Description automatically generated

Obrázek 2: Náhodně vybrané nepodpořené subjekty.

## DiD pro podpořené před rokem 2020

Rokem podpory subjektu se dále rozumí rok, kdy byla podepsána smlouva o podpoře. Pro subjekty podpořené v jednom roce se uvažuje stejný čas podpory (rozdíl v měsíci v roce se neuvažuje, např. všechny subjekty se smlouvou s datem z roku 2017 jsou v jedné skupině).

Vzhledem k odlišným časům zahrnutí do podpory byla provedena jednak společná DiD analýza, tak i dílčí DiD analýzy pro jednotlivé skupiny dle roky udělení podpory zvlášť. Zde lze pak uvažovat porovnání s celou spárovanou kontrolní skupinou nebo s příslušnou podskupinou stejné velikosti (vycházející přímo z PSM párování).

V rámci společného modelu byl uvažován

1. model předpokládající **homogenní změnu**, který předpokládá, že po udělení podpory se zaměstnanost změní stejným způsobem (ve stejném trendu) bez ohledu na to, v jakém roce byla podpora udělena, grafické znázornění viz Obrázek 3,
2. model uvažující **heterogenní efekt** pro různé roky podpory, kdy trend změny v zaměstnanosti může být rozdílný pro jednotlivé roky udělení podpory, viz Obrázek 4.

V případě obou modelů byla uvažována všechna dostupná data o zaměstnanosti od roku 2016 do roku 2019 (tj. ideálně pro každý subjekt 4 údaje o počtu zaměstnanců v jednotlivých letech) a na základě těchto dat byly modely odhadovány.

## Metodologie

V dalším označíme jako Ysti počet zaměstnanců subjektu i v čase t (což je rok údaje mínus 2016) ze skupiny s (pro s nabývající 4 možných hodnot: C nepodpoření, T2017 podpoření v roce 2017, T2018 pro podpořené v roce 2018 a analogicky pro T2019; vše dle podpisu smlouvy). Tuto zaměstnanost budeme modelovat pomocí DID modelu, který připouští lineární efekt času na zaměstnanost a efekt podpory na zaměstnanost.

Model a) předpokládá, jak již bylo uvedeno výše, že je efekt podpory stejný pro skupiny T2017, T2018, T2019. Model pro zaměstnanost Ysti je proto tvaru

kde t je čas (tj. rok mínus 2016) a Dst modeluje čas od přidělené podpory (např. pro skupinu T2018 je tvaru I[t>2]\*(t-2) apod.). Dst je tedy nenulové jen pro podpořené a výše této veličiny se odvíjí od rozdílu času t, ze kterého se bere datový údaj, a roku podpory. Parametry modelu jsou α, β, γs, δ, přičemž **δ odpovídá efektu podpory příslušející jednomu roku**. Poslední člen ε je náhodná chyba.

Model b) připouští různé efekty podpory pro skupiny T2017, T2018 a T2019. Nicméně trend v zaměstnanosti před podporou je uvažován stejný pro všechny (a stejný jako v kontrolní skupině). Model je tedy tvaru

,

kde namísto jednoho společného efektu δ odhadujeme tři různé možné.

## Výsledky

## Model b) je obecnější a zahrnuje model a) jako svůj speciální případ. Je tedy možné testovat, zde je nutné uvažovat různé trendy ve změně zaměstnanosti, nebo se stačí omezit na model se společným trendem pro všechny roky udělené podpory. F-test porovnávající modely a) a b) dává hraniční p-hodnotu 0,05754. Na hladině významnosti 5% je tedy možné uvažovat homogenní model a), nicméně zmíněná p-hodnota je opravdu hraniční. Proto zde uvedeme výsledky pro obě situace.

Odhad modelu a) je znázorněn na obrázku Obrázek 3. **Efekt podpory vychází kladný, téměř 10 zaměstnanců na rok, statisticky významný** (viz Tabulka 3).

Odhadnutý model b) je graficky znázorněn na obrázku Obrázek 4 a jednotlivé hodnoty

DID pro skupiny dle roku podpisu smlouvy jsou uvedeny v tabulce Tabulka 3. Vždy se jedná o efekt podpory odpovídající jednomu roku. **Statisticky významný vychází tento koeficient pro podpořené v roce 2019**, pro podpořené v roce 2018 pak máme hraniční p-hodnotu. Subjektů podpořených v roce 2017 je ve vzorku poměrně málo a tedy možná i z tohoto důvodu **vyšel efekt podpory v této skupině nevýznamný**.

Dále byly provedeny dílčí DiD analýzy, kde byla porovnána vždy jen zaměstnanost rok před podpisem smlouvy a zaměstnanost na konci roku 2019. Jejich výsledky jsou taktéž uvedeny v tabulce Tabulka 3. Vidíme, že **při dílčích analýzách vychází statisticky významný jen přírůstek pro podpořené v roce 2019**. To však může být způsobeno tím, že v této skupině máme nejvíce subjektů a síla příslušných testů je tedy potenciálně vyšší.

**Všechny analýzy vedou ke společnému závěru, že podpora programu OPPIK vede k nárůstu počtu zaměstnanců u podpořených subjektů.**

Tabulka 3: Odhad DiD pro podpořené před rokem 2020

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model |  |  | Odhad | Směr. Chyba | P hodnota |
| homogenní model |  |  | 9,8416 | 2,7083 | 0,0003 |
| heterogenní model | 2017 |  | 3,35881 | 6,2974 | 0,5938 |
|  | 2018 |  | 6,79808 | 3,55595 | 0,0560 |
|  | 2019 |  | 17,18508 | 4,13233 | 0,0000 |
| Dílčí DID | 2017 | spárovaná podskupina | 5,167 | 11,068 | 0,64305 |
|  |  | celá C | 4,7590 | 6,6680 | 0,4756 |
|  | 2018 | spárovaná podskupina | 4,650 | 5,790 | 0,4225 |
|  |  | celá C | 5,267 | 4,263 | 0,21694 |
|  | 2019 | spárovaná podskupina | 12,0274 | 5,7395 | 0,0364 \* |
|  |  | celá C | 11,300 | 5,279 | 0,0325 \* |

Chart, line chart

Description automatically generated

Obrázek 3: Odhadnutý model a) předpokládající stejný trend ve změně zaměstnanosti po udělení podpory. Odhadnuto pro podpořené před rokem 2020.

Chart, line chart

Description automatically generated

Obrázek 4: Odhadnutý model b) umožňující různé trendy v růstu počtu zaměstnanců po udělení podpory pro roky 2017, 2018, 2019. Odhadnuto pro podpořené před rokem 2020.Tento model je v porovnání s modelem a) obecnější a více flexibilní

## DiD pro podpořené v roce 2020

Vzhledem k velkému nedostatku dat číselné (přesné) zaměstnanosti z roku 2020 (viz tabulka níže) nebylo možné (resp. spíše vhodné) porovnávat přímo číselné údaje. Tím bychom totiž mohli dospět ke značně zavádějícím výsledkům, jelikož skupina podniků, pro které data z roku 2020 máme k dispozici, není reprezentativní pro všechny subjekty.

Z tohoto důvodu byla provedena analýza zaměstnanosti kategoriální, pro kterou již máme dostatečné množství údajů. Pro tuto veličinu máme měření dostupné pro většinu subjektů, konkrétně pro 1343 kontrolních a 1323 podpořených.

Tabulka 4: Podpoření v roce 2020: Dostupnost dat o zaměstnanosti na konci roku 2020.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | K dispozici počet zaměstnanců 2020 | | |
|  |  | Počet | Ano | Ne | % k dispozici |
| Podpoření spárovaný vzorek |  | 1349 | 289 | 1060 | 21,42% |
|  | Z toho ANTIVIRUS=1 | 696 | 146 | 550 | 20,98% |
|  | Z toho ANTIVIRUS=0 | 653 | 143 | 510 | 21,90% |
| Kontrolní spárovaný vzorek |  | 1349 | 520 | 829 | 38,55% |
|  | Z toho ANTIVIRUS=1 | 696 | 267 | 429 | 38,36% |
|  | Z toho ANTIVIRUS=0 | 653 | 253 | 400 | 38,74% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | K dispozici počet zaměstnanců 2020 | | |
|  |  | Počet | Ano | Ne | % k dispozici |
| Podpoření spárovaný vzorek | | 1333 | 281 | 1052 | 21,08% |
|  | Z toho ANTIVIRUS=1 | 693 | 145 | 548 | 20,92% |
|  | Z toho ANTIVIRUS=0 | 640 | 136 | 504 | 21,25% |
| Kontrolní spárovaný vzorek | | 1333 | 510 | 823 | 38,26% |
|  | Z toho ANTIVIRUS=1 | 703 | 245 | 458 | 34,85% |
|  | Z toho ANTIVIRUS=0 | 630 | 265 | 365 | 42,06% |
|  |  |  | K dispozici počet zaměstnanců 2020 | | |
|  |  | Počet | Ano | Ne | % k dispozici |
| Podpoření spárovaný vzorek | | 1333 | 281 | 1052 | 21,08% |
|  | Z toho ANTIVIRUS=1 | 693 | 145 | 548 | 20,92% |
|  | Z toho ANTIVIRUS=0 | 640 | 136 | 504 | 21,25% |
| Kontrolní spárovaný vzorek | | 1333 | 510 | 823 | 38,26% |
|  | Z toho ANTIVIRUS=1 | 703 | 245 | 458 | 34,85% |
|  | Z toho ANTIVIRUS=0 | 630 | 265 | 365 | 42,06% |
|  |  |  | K dispozici počet zaměstnanců 2020 | | |
|  |  | Počet | Ano | Ne | % k dispozici |
| Podpoření spárovaný vzorek | | 1333 | 281 | 1052 | 21,08% |
|  | Z toho ANTIVIRUS=1 | 693 | 145 | 548 | 20,92% |
|  | Z toho ANTIVIRUS=0 | 640 | 136 | 504 | 21,25% |
| Kontrolní spárovaný vzorek | | 1333 | 510 | 823 | 38,26% |
|  | Z toho ANTIVIRUS=1 | 703 | 245 | 458 | 34,85% |
|  | Z toho ANTIVIRUS=0 | 630 | 265 | 365 | 42,06% |

Schémata na obrázku Obrázek 5: Počty podniků v jednotlivých kategoriích zaměstnanosti. T pro podpořené a C pro nepodpořené. Řádky odpovídají roku 2021, sloupce roku 2019. znázorňují počty subjektů v jednotlivých kategoriích v roce 2019 a v roce 2021, přičemž se díváme zvlášť na podpořené zapojené do Antiviru (T + Antivirus), kontrolní nepodpořené zapojené do Antiviru (C+Antivirus), podpořené nezapojené do Antiviru   
(T bez Antivirus) a kontrolní nezapojené do Antiviru. Řádky odpovídají kategorii zaměstnanosti v roce 2021, sloupce kategorii v roce 2019. Pro lepší přehlednost jsou navíc počty i barevně zvýrazněny - čím vyšší zastoupení, tím tmavší barva.

Např. tedy první graf ukazuje, že v roce 2019 bylo 46 firem v kategorii do 5 zaměstnanců a tito zde setrvali i v roce 2021. Podobně 11 firem z kategorie 0-5 se přesunulo o kategorii výše a 1 podnik dokonce do kategorie 10-19. Podobně lze číst zbylé údaje. Diagonála odpovídá situaci, kdy nedošlo ke změně v kategorii. Takových firem je nejvíce. Subjekty nad diagonálou si pak v zaměstnanosti polepšily a ty pod diagonálou si naopak pohoršily (počet zaměstnanců klesl do nižší kategorie).

Grafy ukazují, že většina subjektů zůstala ve stejné kategorii zaměstnanosti, v jaké byla v roce 2019, nebo přešla do kategorii o jedno výše nebo níže. Na první pohled je patrné, že subjekty zapojené do programu ANTIVIRUS jsou z hlediska zaměstnanosti jiné než nezapojené subjekty. Rozdíly v přesunech mezi kategoriemi (tj. mimodiagonální prvky) pro podpořené a nepodpořené nejsou z grafů na první pohled výrazné.

Chart

Description automatically generatedChart

Description automatically generatedChart

Description automatically generatedChart, waterfall chart

Description automatically generated

Obrázek : Počty podniků v jednotlivých kategoriích zaměstnanosti. T pro podpořené a C pro nepodpořené. Řádky odpovídají roku 2021, sloupce roku 2019.

## Metodologie

Uvažovali jsme model pro latentní (nepozorovanou) zaměstnanost ve tvaru

*,*

kde γ je hledaný DiD efekt podpory a e je náhodná chyba, o které budeme uvažovat, že má normální rozdělení s nulovou střední hodnotou a neznámým rozptylem. K dispozici máme pouze data Y\*, která odpovídají kategorizované zaměstnanosti (viz tabulky výše). Parametry modelu odhadneme metodou maximální věrohodnosti a významnost parametru γ otestujeme poměrem věrohodností.

## Výsledky

Byly uvažovány dva modely zvlášť pro subjekty zapojené do programu ANTIVIRUS a nezapojené. Výsledný odhad podpory programu spolu s p-hodnotou je uveden v tabulce níže. **V obou případech vychází odhad efektu podporu kladný, nicméně statisticky nevýznamný.** U subjektů zapojených do programu ANTIVIRUS je odhadnutý vyšší efekt než u subjektů nezapojených.

|  |  |  |
| --- | --- | --- |
|  | DiD | p-hodnota |
| ANTIVIRUS=1 | 4,0503 | 0,3343 |
| ANTIVIRUS=0 | 0,3088 | 0,9207 |

Tabulka 5: Výsledek DiD analýzy pro kategoriální zaměstnanost pro podpořené v roce 2020.

**V obou případech je odhadnutý efekt statisticky nevýznamný.** Pro subjekty zapojené do ANTIVIRu vychází kladný, pro nezapojené je odhad blízký nule.

**Možné zdůvodnění výsledků a statistické významnosti:**

* krátké pozorované období, efekt podpory se nestihl ještě projevit,
* pozorujeme pouze kategorizovanou proměnnou, tj. došlo k značné redukci informace,
* efekt podpory je malý natolik, aby se ukázal jako statisticky významný pro daný počet dat a danou zredukovanou dostupnou informaci.

# Ekonomické ukazatele

Porovnáváme DiD pro následující ekonomické ukazatele: VS, PRIDHODN, ON, EBIT, PRODPRACE, TZPVVAS, RENTAKTC.

## Podpoření před rokem 2020

Úprava dat: Z analýzy byly vyloučeny subjekty s hodnotou RENTAKTC rovnou -999. U těchto subjektů byly typicky zbylé ukazatelé zadány jako 0 (přičemž mají být chybějící). Přesné nuly byly z analýzy proto taktéž vyloučeny, a to u všech ukazatelů.

Z podobných důvodů jako u zaměstnanosti, i zde budeme porovnávat údaje z let 2016-2019, jelikož v roce 2020 mohlo dojít (a velice pravděpodobně i došlo) k jinému vývoji ukazatelů než jaké vykazovaly v předchozích letech.

## Metodologie

Pro analýzu DiD byl vyžit stejný přístup jako pro analýzu číselné zaměstnanosti. Konkrétně pro každý subjekt byla využita veškerá dostupná data z let 2016-2019 a pro každý ukazatel byly uvažovány následující modely:

* homogenní model, kde po přidělení podpory dochází ke stejné změně bez ohledu na to, v jakém roce byla podpora přidělena,
* heterogenní model, který připouští různé efekty podpory v závislosti na tom, v kterém roce byla přidělena,
* dílčí DiD analýzy, které berou v úvahu pouze údaj rok před přidělením podpory a údaj z roku 2019.

## Výsledky

U všech ekonomických ukazatelů vychází nevýznamný rozdíl mezi homogenním a heterogenním modelem (3. sloupec v tabulce ). Efekt podpory lze tedy odhadovat společný. **Pro všechny ukazatele ale vychází statisticky vysoce nevýznamný** (2. sloupec tabulky).

|  |  |  |  |
| --- | --- | --- | --- |
|  | DiD | p-hodnota | Stačí homogenní model? |
| VS | 3999,3426 | 0,4043 | 0,7463 |
| PRIDHODN | 3567,4749 | 0,1414 | 0,8741 |
| ON | 1704,6030 | 0,1912 | 0,9445 |
| EBIT | 1114,2231 | 0,3541 | 0,7752 |
| PRODPRACE | 15,8753 | 0,7276 | 0,3889 |
| TZPVVAS | 6161,4535 | 0,3196 | 0,8199 |
| RENTAKTC | -0,4770 | 0,7046 | 0,9900 |

Tabulka 6: Výsledky DiD analýzy pro ekonomické ukazatele pro podpořené před rokem 2020. Sloupec p-hodnota odpovídá p-hodnotě daného DiD efektu, poslední sloupec příslušné p-hodnotě testu, zda stačí uvažovat homogenní model.

**Všechny odhady jsou statisticky nevýznamné.**

## Doplňující výsledky

Zbylé dva přístupy popsané v metodologické části (heterogenní model a dílčí DiD analýzy) lze tedy chápat jako doplňkové a jejich výsledky jsou uvedeny v následující tabulce a mohou sloužit např. k lepšímu náhledu do problému. Ve sloupcích jsou zvlášť DiD odhady pro roku podpory 2017, 2018 a 2019, a to jak bodový odhad efektu, tak jeho p-hodnota významnosti (p vyšší než 0,05 značí statisticky nevýznamný efekt). Pro každý ukazatel máme odhad pro heterogenní model a pro dílčí DiD analýzy provedené v porovnání s kontrolním spárovaným vzorkem a se všemi podpořenými (toto porovnání je však potřeba vnímat jen jako ilustrativní, protože může být zatížené chybou).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 2017 | | 2018 | | 2019 | |
|  |  | Odhad | p | Odhad | p | Odhad | p |
| VS | heterogenní model | -2993,1628 | 0,7941 | 6525,8955 | 0,2945 | 3192,3478 | 0,6594 |
|  | kontrolní spárovaný | -8486,6000 | 0,7089 | 7881,6119 | 0,4508 | 4478,5124 | 0,6913 |
|  | celý kontrolní | -638,5516 | 0,9593 | 5587,0143 | 0,4971 | 3101,6926 | 0,7688 |
| PRIDHODN | heterogenní model | 6180,9058 | 0,2873 | 3379,8075 | 0,2822 | 2755,4902 | 0,4526 |
|  | kontrolní spárovaný | 4063,5722 | 0,6998 | 3496,5294 | 0,3585 | 1988,5275 | 0,7321 |
|  | celý kontrolní | 6506,0607 | 0,3586 | 3371,4234 | 0,4337 | 1931,8131 | 0,7116 |
| ON | heterogenní model | 1821,8484 | 0,5581 | 2002,5841 | 0,2364 | 1210,9521 | 0,5402 |
|  | kontrolní spárovaný | 768,3667 | 0,9089 | 2128,9965 | 0,3500 | 753,5275 | 0,8138 |
|  | celý kontrolní | 2171,9664 | 0,5347 | 1971,3024 | 0,3689 | 659,3486 | 0,8221 |
| EBIT | heterogenní model | 2790,4112 | 0,3322 | 1093,0488 | 0,4832 | 446,0270 | 0,8061 |
|  | kontrolní spárovaný | 2209,3500 | 0,4811 | 1127,1362 | 0,4283 | 615,0108 | 0,8277 |
|  | celý kontrolní | 2902,6441 | 0,4487 | 1390,5217 | 0,5327 | 478,0439 | 0,8482 |
| PRODPRACE | heterogenní model | -80,0986 | 0,4177 | 9,7991 | 0,8717 | 80,5035 | 0,2682 |
|  | kontrolní spárovaný | -132,3333 | 0,7410 | 28,4701 | 0,4353 | -15,2683 | 0,8684 |
|  | celý kontrolní | -110,0935 | 0,3432 | 35,0404 | 0,5503 | -8,7416 | 0,9143 |
| TZPVVAS | heterogenní model | 4911,5624 | 0,7395 | 9212,6215 | 0,2519 | 2190,6147 | 0,8148 |
|  | kontrolní spárovaný | 2395,5611 | 0,9409 | 9087,4753 | 0,4848 | 2343,9100 | 0,8712 |
|  | celý kontrolní | 6947,6187 | 0,6688 | 7608,4126 | 0,4668 | 686,7083 | 0,9592 |
| RENTAKTC | heterogenní model | -0,1642 | 0,9553 | -0,4624 | 0,7776 | -0,6331 | 0,7368 |
|  | kontrolní spárovaný | -2,4008 | 0,0960 | 0,2178 | 0,8957 | 0,4488 | 0,7530 |
|  | celý kontrolní | 0,1238 | 0,9547 | 1,1016 | 0,7402 | 0,4606 | 0,7432 |

Tabulka 7: Doplňující výsledky DiD analýzy pro ekonomické ukazatele pro podpořené před rokem 2020.

Nakonec ještě uvádíme grafy odhadnutých výsledných homogenních modelů pro jednotlivé ekonomické ukazatele, a to na obrázcích 6,7 a 8. Černá křivka odpovídá vývoji ukazatele v kontrolní skupině, červeně jsou podpoření v roce 2017, zeleně podpoření v roce 2018 a modře podpoření v roce 2019. Z grafu je též možné vyčíst průměrné hodnoty v jednotlivých skupinách v roce 2016 a v roce 2019. Dále vidíme, že VS, PRIDHODN, ON, EBIT a TZPVVAS vykazují rostoucí trend ve všech skupinách, zatímco PRODPRACE je u kontrolní skupiny velmi mírně klesající a u podpořených má stagnující až mírně rostoucí trend. RENTAKTC má klesající trend.

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated

Obrázek : Odhadnutý DiD pro VS a PRIDHODN pro podpořené před rokem 2020.

Obrázek : Odhadnutý DiD pro EBIT, ON PPRACE, TZPVVAS pro podpořené před rokem 2020.A picture containing chart

Description automatically generatedChart, line chart

Description automatically generatedChart, line chart

Description automatically generatedChart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

Obrázek : Odhadnutý DiD pro RENTAKTC pro podpořené před rokem 2020.

## Podpoření v roce 2020

Pro velké množství subjektů, podpořených téměř 80 % a kontrolních 60 %, chybí údaj o výši ekonomických ukazatelů v roce 2020. K dispozici tedy máme pouze informace **pro 289 podpořených a 520 kontrolních** pro prvních 6 ukazatelů. Pro RENTAKTC se v datech vyskytovala hodnota -999, kterou jsme nahradili také za chybějící hodnotu. Zde tedy máme data pro 282 podpořených a 507 nepodpořených. V DiD analýze budeme pracovat pouze s těmito subjekty, pro které máme k dispozici hodnoty z obou let 2019 i 2020.

Subjekty zapojené do programu ANTIVIRUS a subjekty nezapojené byly analyzovány zvlášť (pro obě skupiny byl odhadnut separátní model). Výsledky analýzy jsou shrnuty v následující tabulce. **Efekt podpory je na hladině 5 % nevýznamný u všech ukazatelů.** Kladné hodnoty odpovídají tomu, že podpora vedla k navýšení ukazatele oproti hypotetické situaci bez podpory.

**Možné důvody negativních a nevýznamných efektů:**

* odlehlé hodnoty, které mohou zkreslit odhady,
* nedostatečná délka sledovaného období (podpora se nestačila projevit),
* velké množství chybějících hodnot a možný příslušný bias.

Proměnná RENTAKTC obsahuje také jedno značně odlehlé pozorování (-784), které může potenciálně ovlivnit výsledky. Analýza byla proto provedena jak s tímto pozorováním, tak i bez něj, s tím, že větší výpovědní hodnotu má odhad pro data bez tohoto extrému.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | bez ANTIVIRUS | | ANTIVIRUS | |
|  | Odhad | P-hodnota | Odhad | P-hodnota |
| VS | 354,7163 | 0,9772 | 5419,6404 | 0,7177 |
| PRIDHODN | -1132,3089 | 0,7594 | 286,7188 | 0,9657 |
| ON | 77,8927 | 0,9747 | -1,4774 | 0,9997 |
| EBIT | -635,4625 | 0,7655 | 161,8866 | 0,9516 |
| PRODPRACE | 472,5415 | 0,3430 | 109,4479 | 0,1140 |
| TZPVVAS | -2169,4147 | 0,8911 | 6326,7565 | 0,7438 |
| RENTAKTC | -5,9976 | 0,2108 | -0,7303 | 0,7182 |
| RENTAKTC bez odlehlého | -0,3659 | 0,8727 | -0,7303 | 0,7182 |

Tabulka 8: Výsledky DiD analýzy ekonomických ukazatelů pro subjekty podpořené v roce 2020.

Odhadnuté modely jsou znázorněny též na grafech na obrázcích 9,10 a 11. Z nich lze též vyčíst průměrné hodnoty ukazatelů v jednotlivých skupinách v roce 2019 a 2020.

A picture containing chart

Description automatically generatedDiagram

Description automatically generated with medium confidence

Obrázek : Odhadnutý DiD pro VS a PRIDHODN pro podpořené v roce 2020.

A picture containing diagram

Description automatically generatedChart

Description automatically generatedChart, line chart

Description automatically generatedA picture containing diagram

Description automatically generated

Obrázek : Odhadnutý DiD pro EBIT, ON PPRACE, TZPVVAS pro podpořené v roce 2020.

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generatedObrázek : Odhadnutý DiD pro RENTAKTC pro podpořené v roce 2020. Vlevo s jedním odlehlým pozorováním, vpravo bez tohoto extrému.

# Příloha A: Evaluace PSM

Na základě výsledného modelu se podařilo spárovat každý podpořený subjekt s jedním nepodpořeným. Následně byly porovnány charakteristiky obou vzorků. Spojité veličiny byly porovnány pomocí Kolmogorova Smirnovova testu a pomocí dvouvýběrového t-testu. Kategoriální veličiny byly porovnány pomocí chí kvadrát testu. Výsledy jsou uvedeny v tabulkách 9 a 10.

Rozdělení skóre v obou vzorcích pro subjekty podpořené do roku 2020 je ilustrováno graficky na obrázcích 12 a 13. Stejné grafy pro subjekty podpořené v roce 2020 jsou uvedeny na obrázcích 19 a 20. Rozdělení v podpořeném vzorku je znázorněno červenou barvou, ve spárovaném kontrolním vzorku modrou barvou a v souboru všech potenciálních kandidátů na kontrolní vzorek zelenou barvou. Je patrné, že červeně a modře znázorněné rozdělení jsou si velmi blízká na rozdíl od zeleného, které je na první pohled odlišné. Podobný závěr dávají obrázky 13 a 20, kde je patrné, že spárováním došlo ke značnému přiblížení.

Po spárování došlo k značné redukci absolutní standardizované střední vzdálenosti u většiny proměnných (obrázky 14 a 21). To též značí, že párování byl úspěšný krok.

Grafy 15-18 pro podpořené před rokem 2020 a grafy 22-25 pro podpořené v roce 2020 pak ilustrují rozdělení klíčových proměnných v podpořeném vzorku (červená barva), spárovaném kontrolním vzorku (modrá barva) a v souboru všech potenciálních kandidátů na kontrolní vzorek (zelená barva). Pro spojité proměnné je zobrazeno jako porovnání empirických distribučních funkcí (Fn), tak i jádrové odhady hustot (density). Pro kategoriální proměnné je zobrazen graf relativních četností. Z grafů lze pozorovat, že pro některé proměnné došlo k zásadnímu přiblížení rozdělení ve skupině podpořených a kontrolních subjektů.

Tabulky 9 a 10 ukazují, že pro všechny charakteristiky nelze na hladině 0,05 zamítnout hypotézu, že je jejich rozdělení v obou vzorcích stejné, ani že se jejich střední hodnoty liší. Všechny uvedené p-hodnoty jsou totiž vyšší než 0,05 a tedy není detekován žádný statisticky významný rozdíl pro žádnou z uvažovaných klíčových proměnných.

Na základě všech uvedených diagnostik lze tedy označit výsledek PSM za uspokojivý.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | před 2020 | | | | 2020 | | | |
|  | KS test | KS p-hodnota | t-test | t p-hodnota | KS test | KS p-hodnota | t-test | t p-hodnota |
| ROKVZNIKU | 0,049 | 0,772 | -0,397 | 0,691 | 0,036 | 0,360 | -1,058 | 0,290 |
| POCETZAM | 0,043 | 0,879 | 0,115 | 0,908 | 0,044 | 0,151 | 1,295 | 0,195 |
| AKTIVACELK | 0,095 | 0,072 | -0,096 | 0,924 | 0,050 | 0,065 | 0,304 | 0,761 |
| VYKONY | 0,095 | 0,072 | 0,297 | 0,766 | 0,044 | 0,151 | 0,830 | 0,406 |
| MIRAZADLU | 0,057 | 0,589 | -0,017 | 0,987 | 0,070 | 0,003 | 0,612 | 0,541 |
| OBRAT | 0,076 | 0,239 | 0,202 | 0,840 | 0,041 | 0,212 | 2,854 | 0,004 |
| LOGOBRAT | 0,076 | 0,239 | -0,416 | 0,678 | 0,041 | 0,212 | 1,611 | 0,107 |
| PMS skóre | 0,014 | 1,000 | -0,012 | 0,991 | 0,016 | 0,994 | -0,443 | 0,658 |

Tabulka 9: Diagnostické testy pro evaluaci párování: Kolmogorovův Smirnovův test (KS) porovnávající shodu celých rozdělení a t-test, který se zaměřuje na shodu středních hodnot. Všechny p-hodnoty jsou vyšší než 0,05.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | před 2020 | | rok 2020 | |
|  | Chí kvadrát test | p | Chí kvadrát test | p |
| NACE2 | 9,559 | 0,794 | 11,020 | 0,856 |
| KRAJ | 2,118 | 1,000 | 3,871 | 0,992 |
| TYP\_FY | 0,604 | 0,739 | 2,571 | 0,276 |
| POCZAMKAT | 1,665 | 0,976 | 6,136 | 0,632 |
| MZ | 0,442 | 0,979 | 2,529 | 0,640 |

Tabulka 10: Diagnostické testy pro evaluaci párování z hlediska kategoriálních proměnných. Chí kvadrát test a příslušné p-hodnoty, všechny statisticky nevýznamné.

Chart, histogram

Description automatically generated

Obrázek 12: Podpora do roku 2020: Rozdělení PMS skóre ve vzorcích. Znázornění histogramů v jednotlivých vzorcích (červeně podpoření, modře spárování nepodpoření a zeleně všichni nepodpoření) a porovnání empirických distribučních funkcí. Je patrné, že modře a červeně znázorněná rozdělení jsou si blízká, a na rozdíl od zeleného.

A picture containing chart

Description automatically generated

Obrázek 13: Podpora do roku 2020: Znázornění PSM skóre u podpořených, spárovaných nepodpořených a všech potenciálních nepodpořených.

Chart, scatter chart

Description automatically generated

Obrázek 14: Podpora do roku 2020: Graf standardizované střední vzdálenosti před spárováním (bílá kolečka) a po spárování (černá kolečka). Je patrné, že spárované vzorky jsou si bližší z hlediska většiny proměnných.

Graphical user interface, chart

Description automatically generated

Obrázek 15: Podpora do roku 2020: Empirická distribuční funkce spojitých proměnných (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

Chart

Description automatically generated

Obrázek 16: Podpora do roku 2020: Jádrový odhad hustoty spojitých proměnných (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

Chart, histogram

Description automatically generated

Obrázek 17: Podpora do roku 2020: Rozdělení kategoriálních proměnných NACE a Kraj (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

Chart, bar chart

Description automatically generated

Obrázek 18: Podpora do roku 2020: Rozdělení kategoriálních proměnných typ fyzické osoby, počet zaměstnanců kategoriální, míra zadlužení (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

.

Chart, histogram

Description automatically generated

Obrázek 19: Podpora v roce 2020: Rozdělení skóre ve vzorcích. Znázornění histogramů v jednotlivých vzorcích (červeně podpoření, modře spárování nepodpoření a zeleně všichni nepodpoření) a porovnání empirických distribučních funkcí. Je patrné, že modře a červeně znázorněná rozdělení jsou si blízká, a na rozdíl od zeleného.

A picture containing graphical user interface

Description automatically generated

Obrázek 20: Podpora v roce 2020: PSM skóre u podpořených, spárovaných nepodpořených a všech nepodpořených..

Chart, scatter chart

Description automatically generated

Obrázek 21: Podpora v roce 2020: Graf standardizované střední vzdálenosti před spárováním (bílá kolečka) a po spárování (černá kolečka). Je patrné, že spárované vzorky jsou si bližší z hlediska většiny proměnných.

Graphical user interface, chart

Description automatically generated

Obrázek 22: Podpora v roce 2020: Empirická distribuční funkce spojitých proměnných (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

Chart

Description automatically generated

Obrázek 23: Podpora v roce 2020: Jádrový odhad hustoty spojitých proměnných (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

Chart, bar chart, histogram

Description automatically generated

Obrázek 24: Podpora v roce 2020: Rozdělení kategoriálních proměnných NACE a Kraj (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).

Chart, bar chart, waterfall chart

Description automatically generated

Obrázek 25: Podpora v roce 2020: Rozdělení kategoriálních proměnných typ fyzické osoby, počet zaměstnanců kategoriální, míra zadlužení (červeně podpoření, modře spárovaný kontrolní vzorek a zeleně všichni nepodpoření).