

Modelování dopadů reformy financování obcí s využitím syntetických dat

Žijeme v simulaci

Františka Tranová

18.6.2026

Jiří Münich

paq
research



O čem to dnes bude?

- ❶ Pracujeme na návrhu reformy a chceme odhadovat její dopady
 - ❷ Nemáme ale přístup k některým datům
 - ❸ Simulovaná data umožňují provizorně modelovat dopady reforem
- Náš konkrétní příklad:
Rozpočtové určení daní a syntetické mzdy zaměstnanců v obcích

O využití dat syntetických dat

- Syntetická data:

Simulace na základě podložených parametrů

- 1 Modelování dopadů navrhovaných reforem
- 2 Jak se dají využít agregovaná data?
- 3 Potenciál syntetických dat pro veřejnou politiku

Zkoumaný problém:

Rozpočtové určení daní

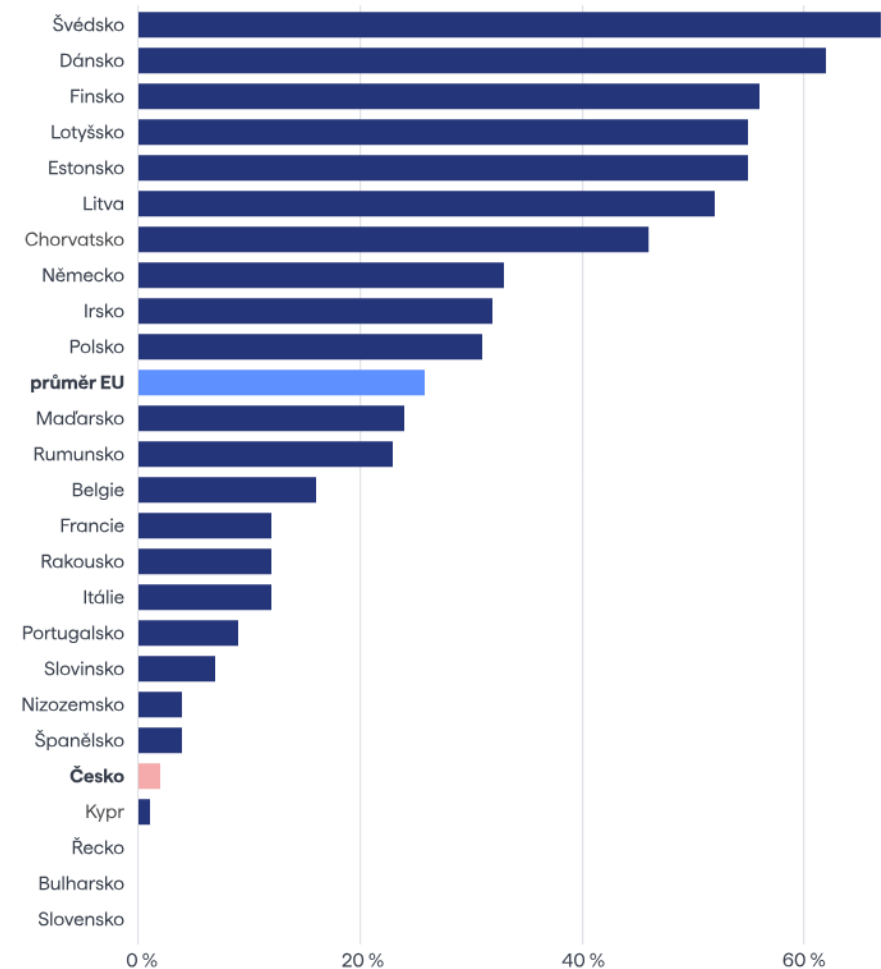
Co je to RUD?

- Formulka, podle které se část vybraných daní rozděluje mezi obce
- Přes 300 miliard Kč
- Současné RUD má mnoho nedostatků
- Mimo jiné:

Nezohledňuje ekonomickou aktivitu v obci

Graf 5 **Příjmy obcí v České republice prakticky neodráží místní ekonomickou aktivitu**

Příjmy z daní z hospodářské činnosti jako procento daňových příjmů obecních rozpočtů v zemích EU



Poznámka Stav k roku 2021
Zdroj Herrmann 2022

Jak se počítá RUD?

- Navrhujeme reformu RUD a to je hodně počítání
- Máme celou knihovnu v R pro počítání různých variant
- Dokonce jsme vytvořili aplikaci

Ale chybí vstupní data

The screenshot displays the 'RUD Manager' application interface. The top navigation bar includes 'RUD Manager', 'Úvod', 'Parametry' (selected), 'Výsledky', 'Scénáře', and 'Data'. The left sidebar, titled 'Nastavení výpočtu', contains a 'Rok výpočtu' dropdown set to '2025', a 'Přesnost (bity)' slider set to '120' (with markers at 64, 120, and 256), and a 'Spustit výpočet' button. The main area, titled 'Model 1 — BFI daně', shows configuration for 'Proměnná v1'. Under 'Proměnná', 'Počet obyvatel' is selected. The 'Transformace (volitelné)' section includes 'Přechodové hodnoty (cuts)' with an empty input field, 'Transformační proměnná (cut_variable)' set to 'Stejná jako proměnná vy', and a note about piecewise transformation. The 'Výjimky (volitelné)' section is also visible.

Co potřebujeme pro reformu RUD?

- Chceme znát dopady navrhovaných reforem kvůli

- Parametrizaci variant

10% RUD dělit podle mezd v obci – moc nebo málo?

- Korelovaným proměnným

Do větších měst dojíždí více lidí za prací, co s tím?

- Dobrému vládnutí

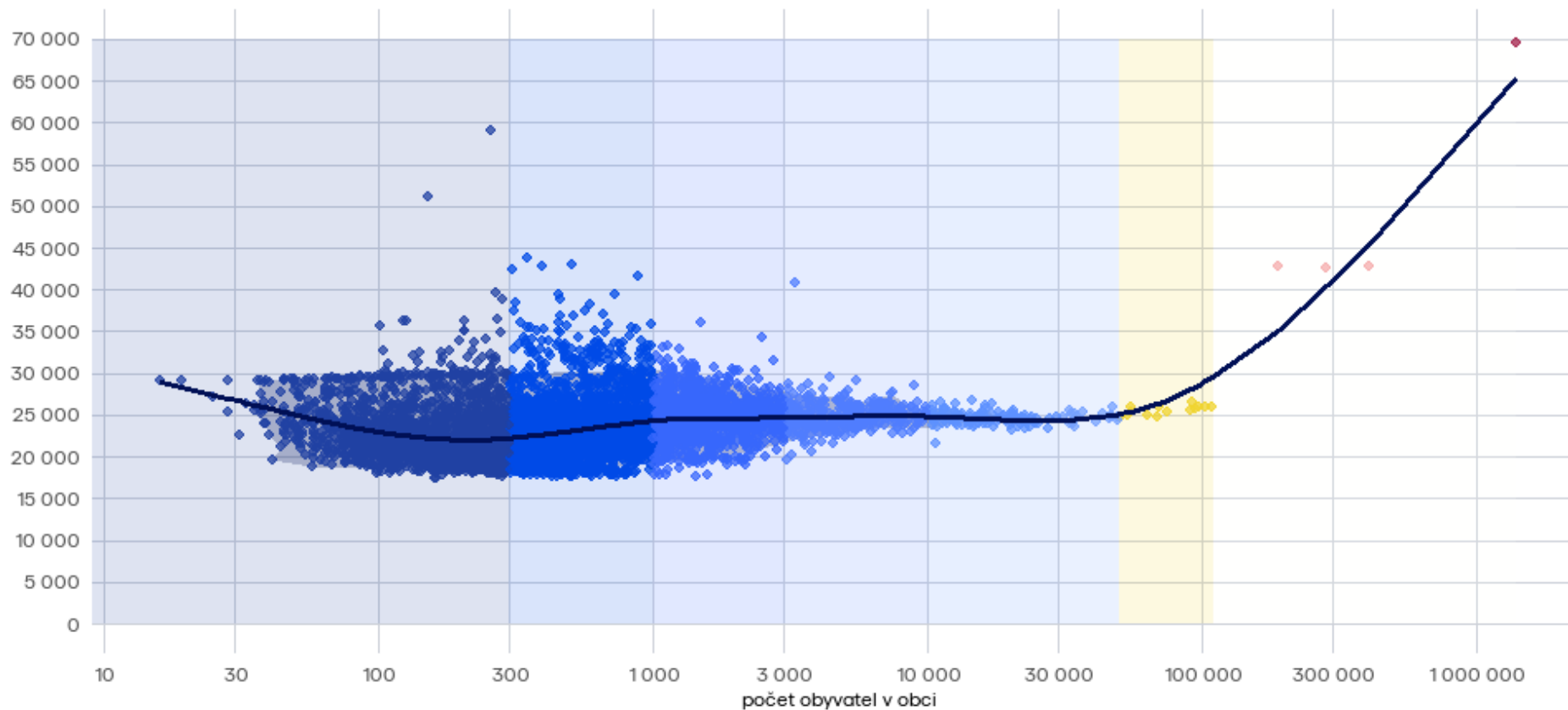
Nemůžete navrhovat reformu bez napočítání dopadů.

- Musíme počítat na nějakých datech

Současné RUD

příjmy z RUD na obyvatele [Kč] (původní varianta)

◆ do 300 ◆ 300 - 1 tisíc ◆ 1 tisíc - 3 tisíce ◆ 3 tisíce - 10 tisíc ◆ 10 tisíc - 30 tisíc ◆ 30 tisíc - 50 tisíc ◆ nad 50 tisíc



Praktický problém:

Nejsou data o mzdách zaměstnanců v obcích

Jaká data máme?

- ❶ **Vzdělání obyvatel** podle bydliště a meziobecní **dojížďka za prací** (SLBD 2021)
- ❷ **Rozdělení platů a mezd** podle vzdělání a kraje (VPSV)
- ❸ **Počty zaměstnanců v obci** z vyhlášky o RUD (ČSSZ, přihlášky k sociálnímu pojištění)

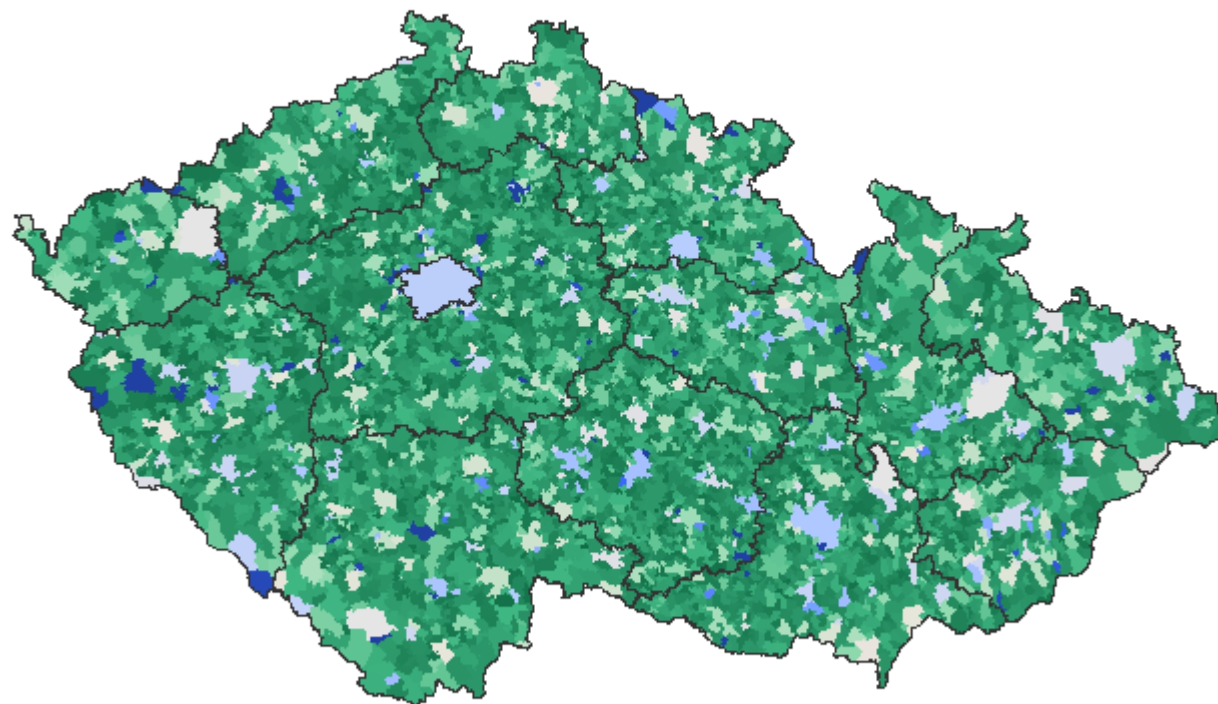
**Jak rozdělit krajské údaje o mzdách
na úroveň obcí?**

Počty zaměstnanců

Stát ví, kolik ve které obci pracuje
zaměstnanců
(zdroj: vyhláška o RUD)

Nevíme ale nic dalšího – skladnice
jako inženýrka

Zaměstnanci na obyvatele



Zdroj: data RUD (decree) + ČSÚ SLDB 2021

Vzdělání podle bydliště

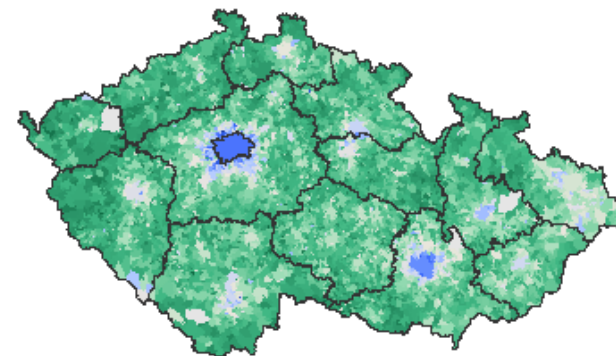
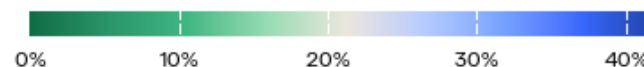
Data mají správnou granularitu...

...ale lidé často pracují jinde, než kde bydlí

A jak se liší plat vysokoškoláka od platu maturanta?

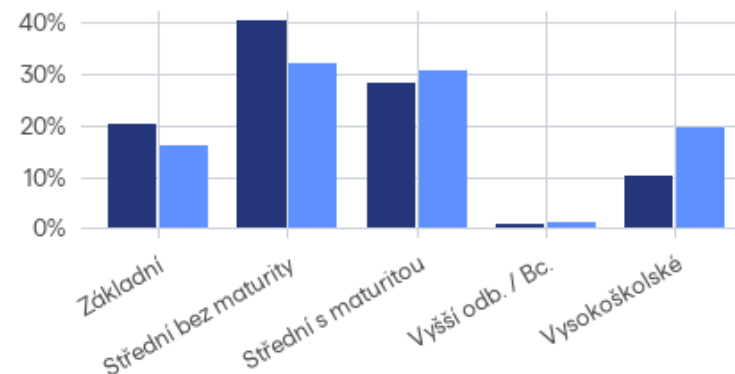
Vzdělání obyvatel v obcích

Podíl vysokoškoláků



Podíl obyvatel

Karviná Ostrava



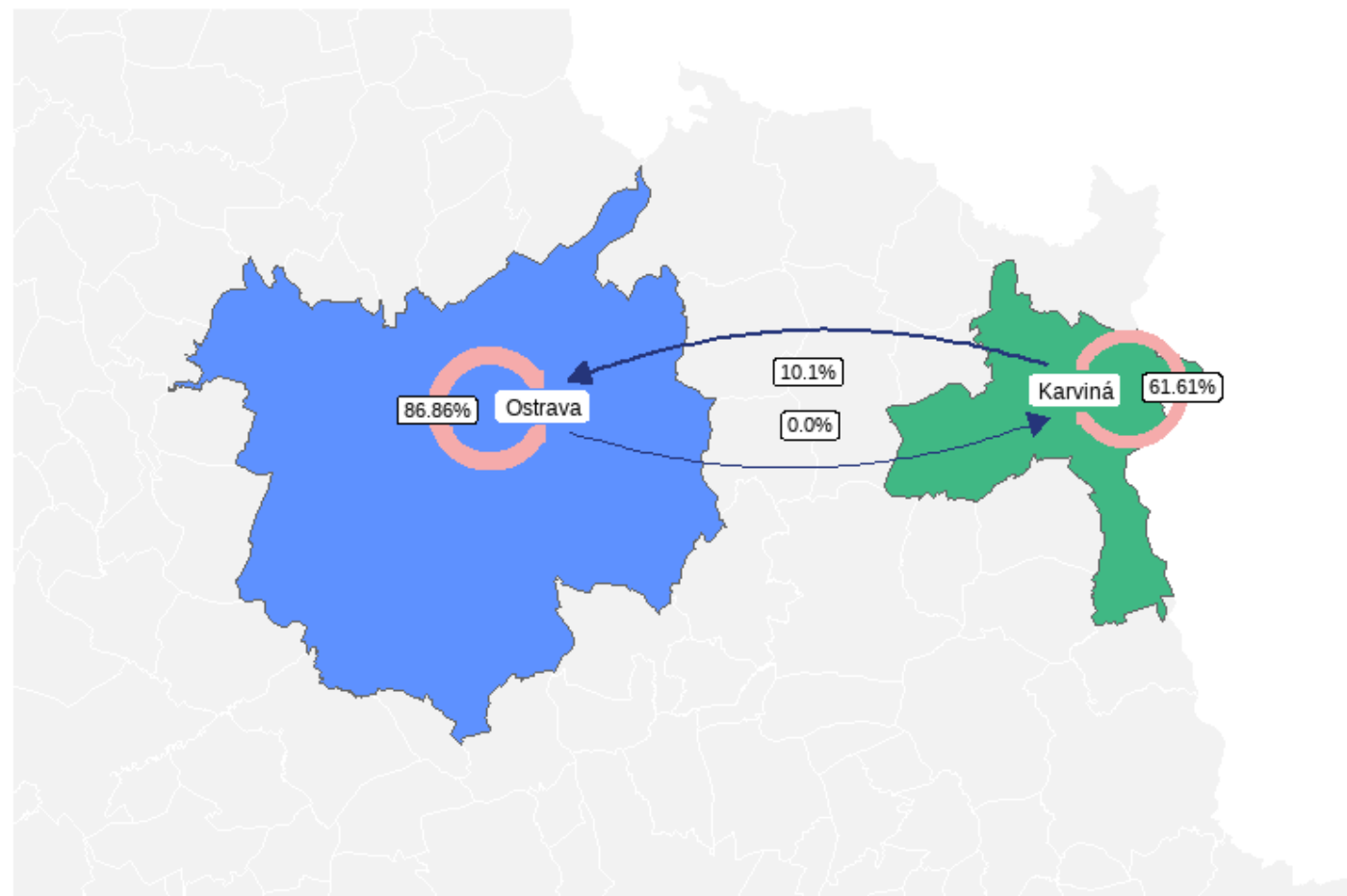
Vzdělání podle místa práce

Podle dat o dojíždce můžeme informace o obyvatelstvu promítnout do složení pracovní síly v obci

Na **100 vysokoškoláků žijících** v Karvině můžeme **počítat 10 (10 %)** do zaměstnanců v Ostravě.

Dojíždka za prací mezi obcemi

Šipka = mezi obcemi · smyčka = uvnitř obce



Ale dojíždí vysokoškoláci za prací stejně jako ostatní?

Ze **souhrnných statistik a akademických** publikací můžeme převážít dojížděku různých skupin

Lidé s vyšším vzděláním například mohou mít tendenci

- Dojíždět na větší vzdálenost
- Do velkých obcí

Diskrece analytictva

Přenesení syntézy a literatury zpět mezi jednotlivá pozorování

Kam dojíždějí vzdělanější lidé

Realokace dojíždějících podle velikosti cílové obce

Velikost obce (počet obyvatel)	Základní	Střední bez maturity	Střední s maturitou	Vyšší odb. / Bc.	Vysokoškolské
> 50 tis.	0.70	0.65	1.00	1.30	1.60
10–50 tis.	0.85	0.85	1.00	1.10	1.20
2–10 tis.	0.95	1.10	1.00	0.98	0.90
500–2 tis.	1.10	1.25	1.00	0.90	0.70
< 500	1.20	1.30	1.00	0.85	0.60

Zdroj: vlastní kalibrace (PAQ) na základě SLDB 2021 a dojížděky ČSÚ

Dojíždka adaptovaná podle výše vzdělání

Můžeme nuancovat na datech založenou dojíždku

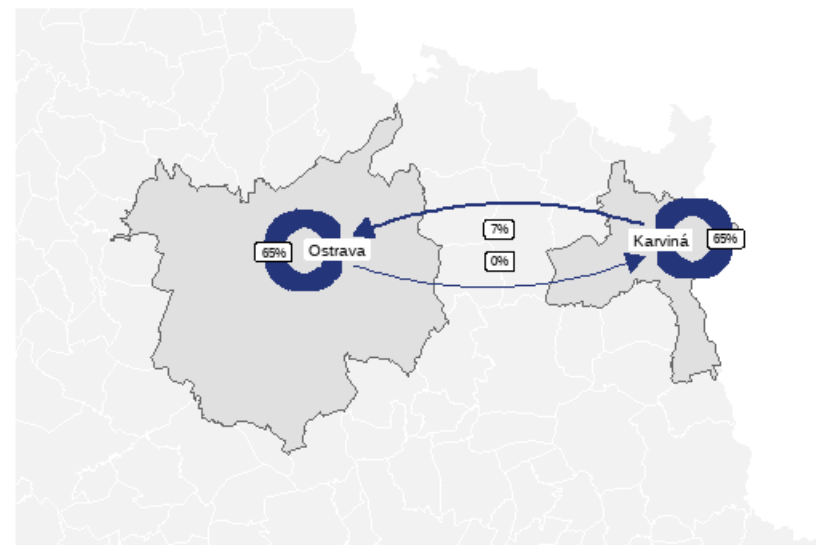
Z Karviné se do ostravských zaměstnanců počítá:

- Jen 7 % lidí se **základní školou**
- Ale 16 % lidí s **vysokou školou**

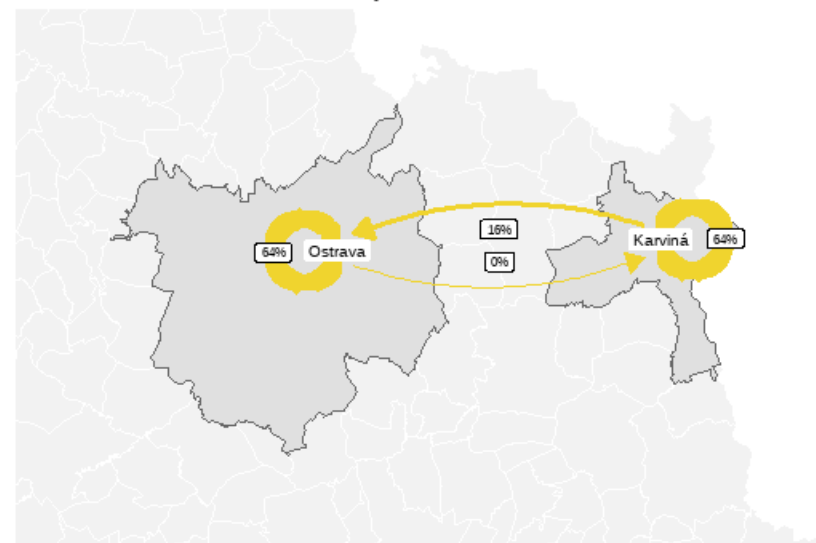
Dojíždka rozložená podle vzdělání

Podíl dojíždějících mezi Karviná a Ostrava podle vzdělání

Základní



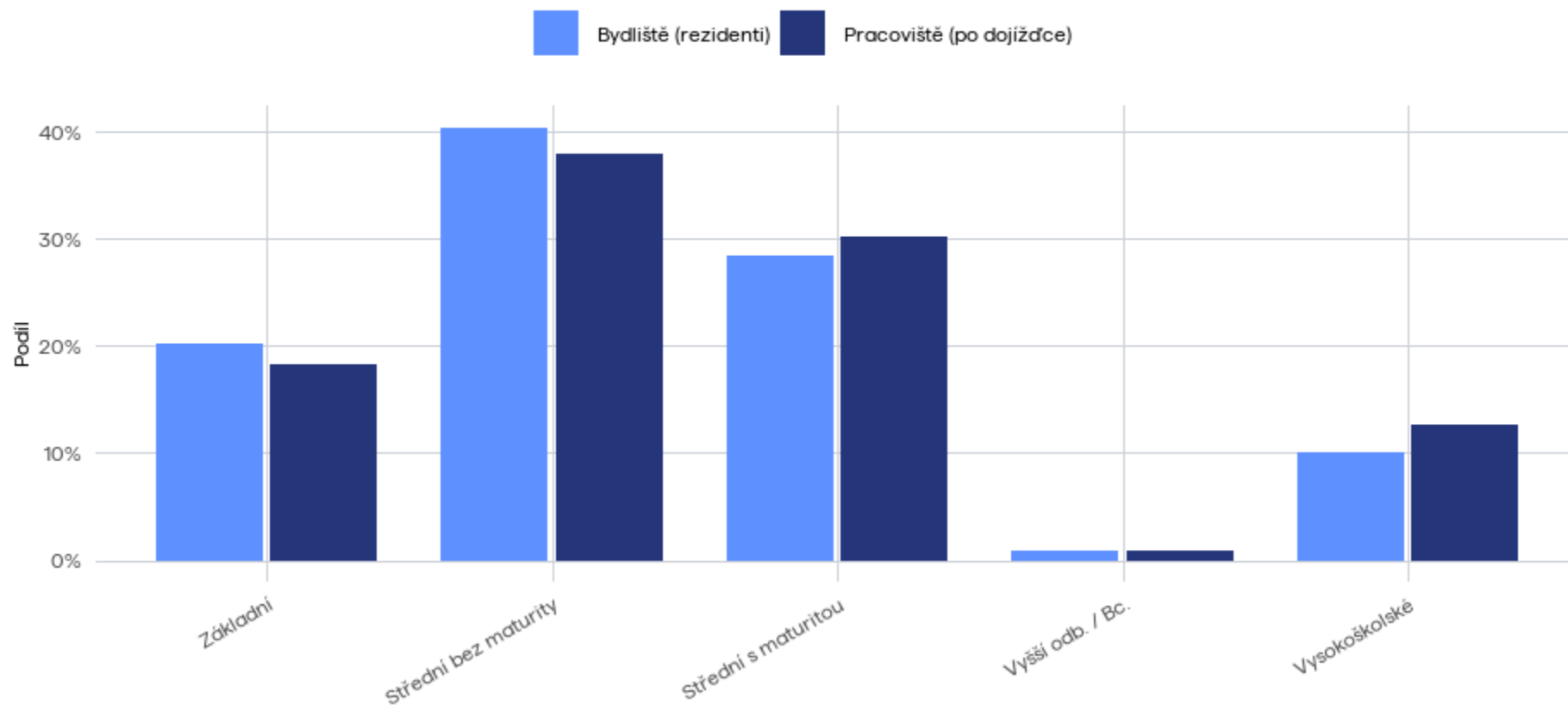
Vysokoškolské



Výsledné rozdělení

Skladba vzdělání: Karviná

Bydliště vs. pracoviště po přerozdělení dojíždkou



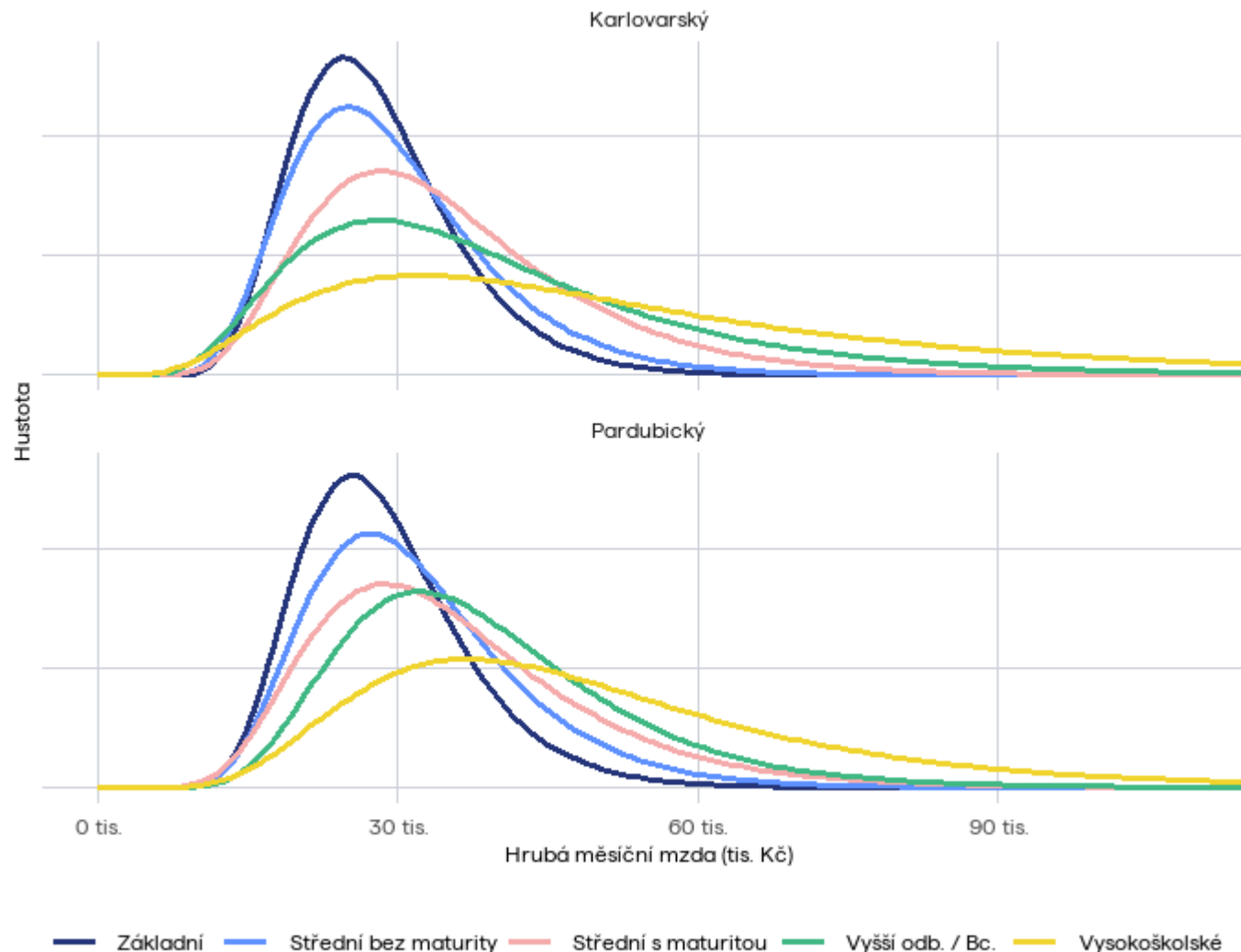
Co víme o mzdách?

Ze souhrnných statistiky (1. a 9. decil, kvartily, průměr) můžeme odhadnout kontinuální log-normální rozdělení mezd

Známe je pro každý kraj a pro úroveň vzdělání

Mzdy podle vzdělání – Karlovarský a Pardubický kraj

Lognormální křivky odhadnuté z mediánu a kvartilů z ISPV



Zdroj: ISPV / MPSV, Regionální statistika ceny práce, 2021

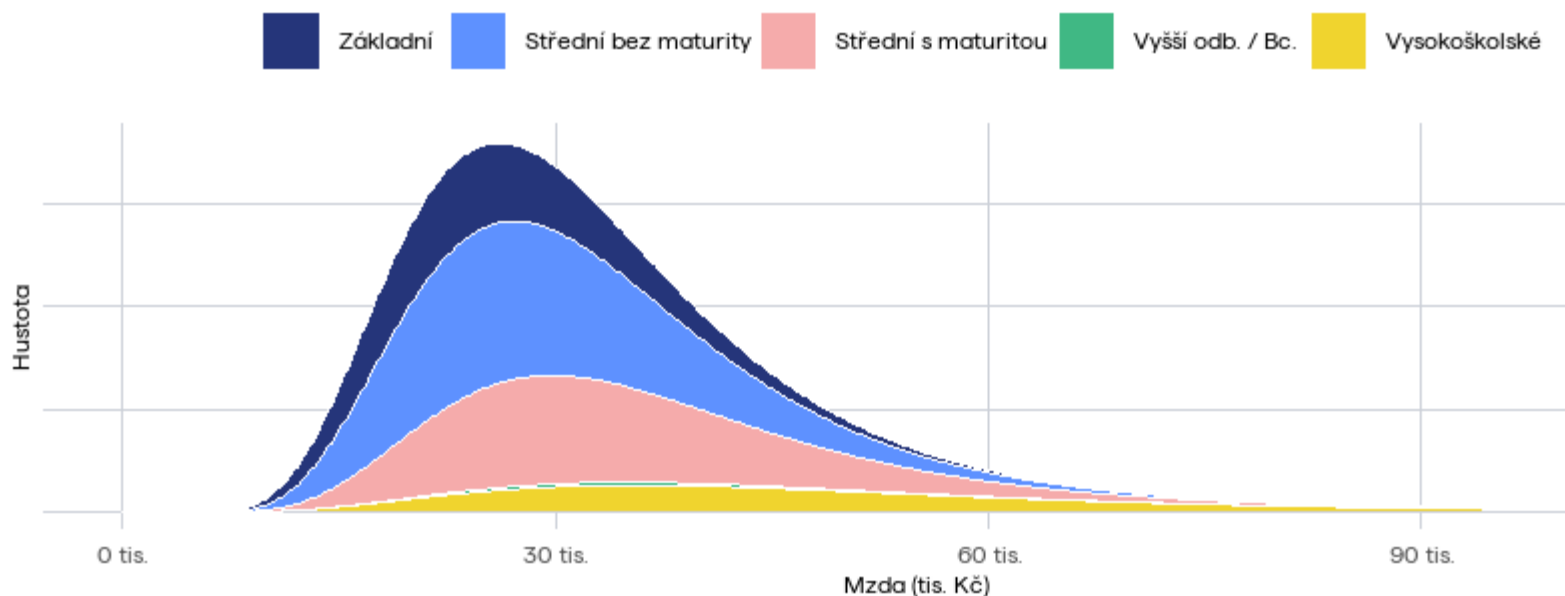
Od kraje k obci

Obecní rozdělení mezd je jen vážené poskládání krajských rozdělení

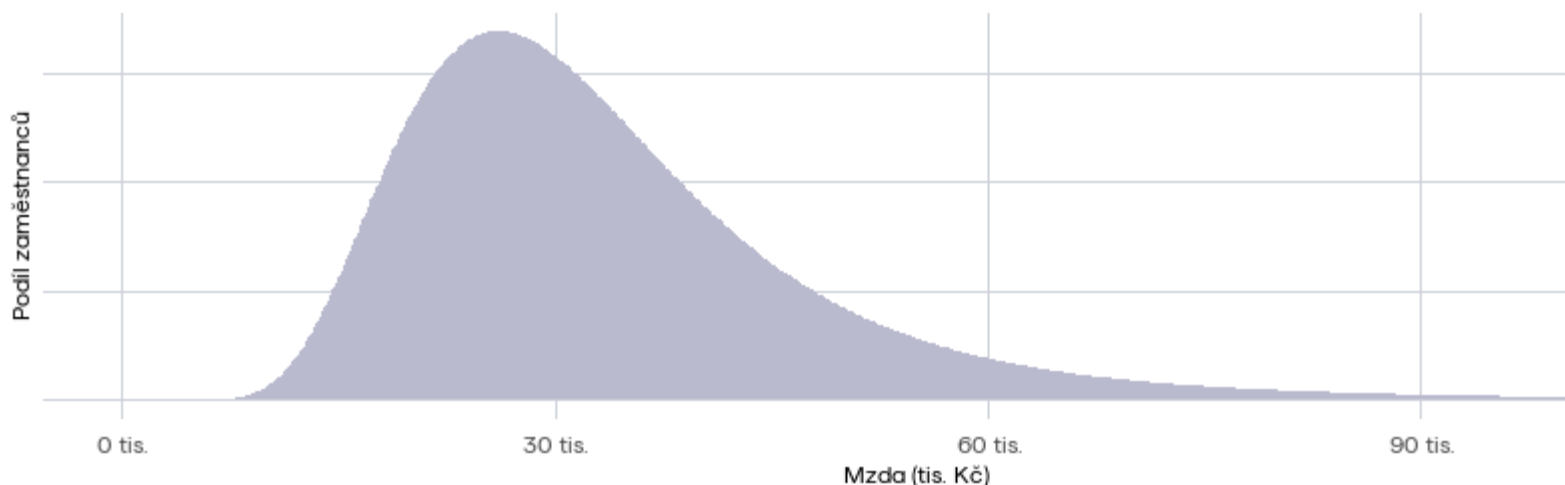
Váhy jsou odhadované zastoupení vzdělanostních skupin zaměstnanců v obci

Ze skladby vzniká mzdová distribuce: Karviná

Rozklad podle vzdělání



Výsledná distribuce



Od rozdělení k počtům

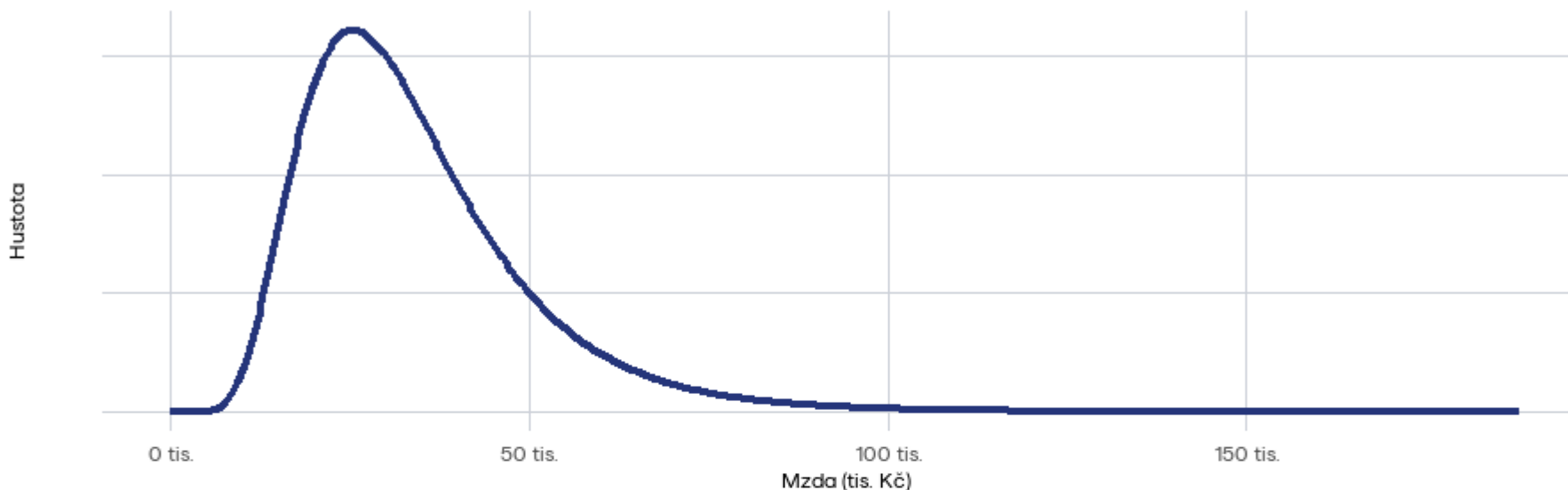
Z pravděpodobnostního rozdělení je pak snadné simulovat data

Stačí znát skutečné počty zaměstnanců:

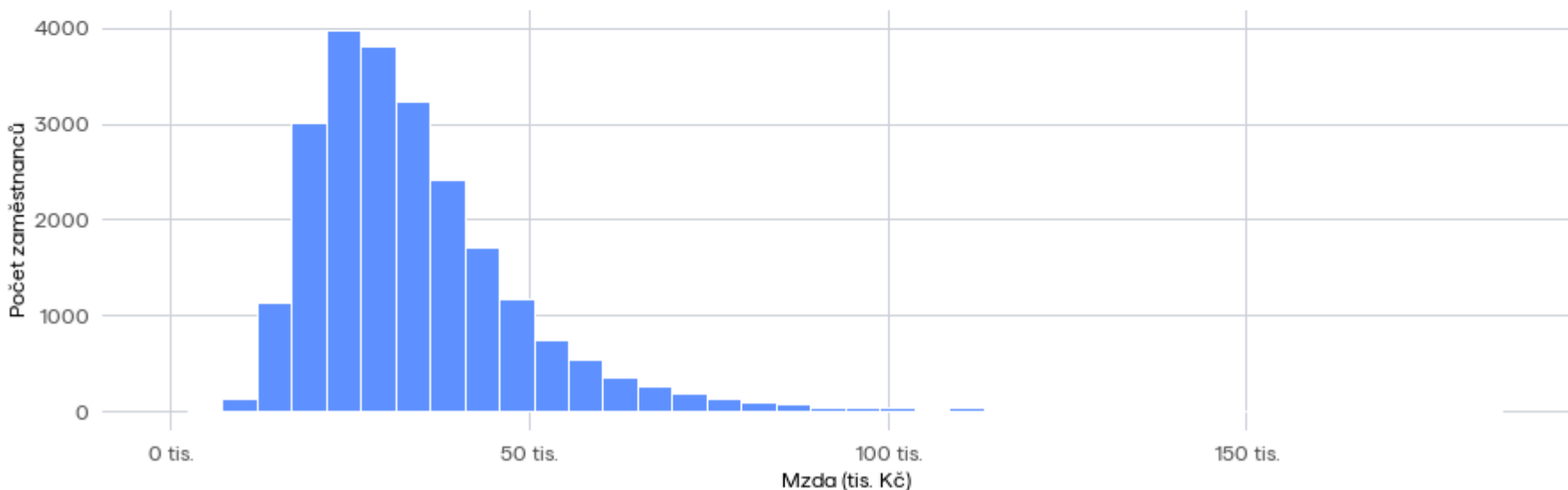
- Z vyhlášky o RUD

Z křivky generujeme zaměstnance: Karviná

Lognormální křivka

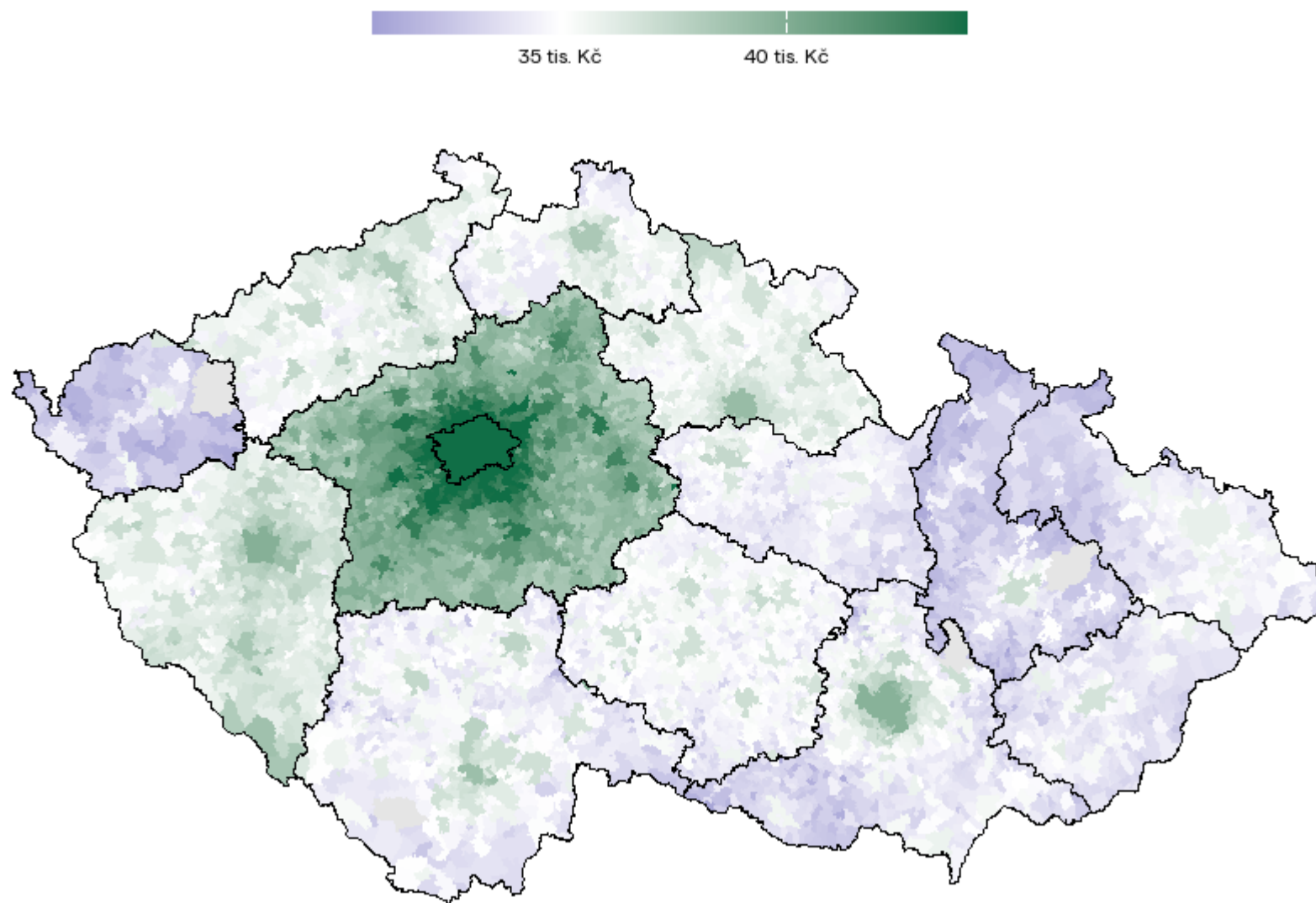


Simulovaní zaměstnanci



Hle, data o mzdách zaměstnanců v obcích

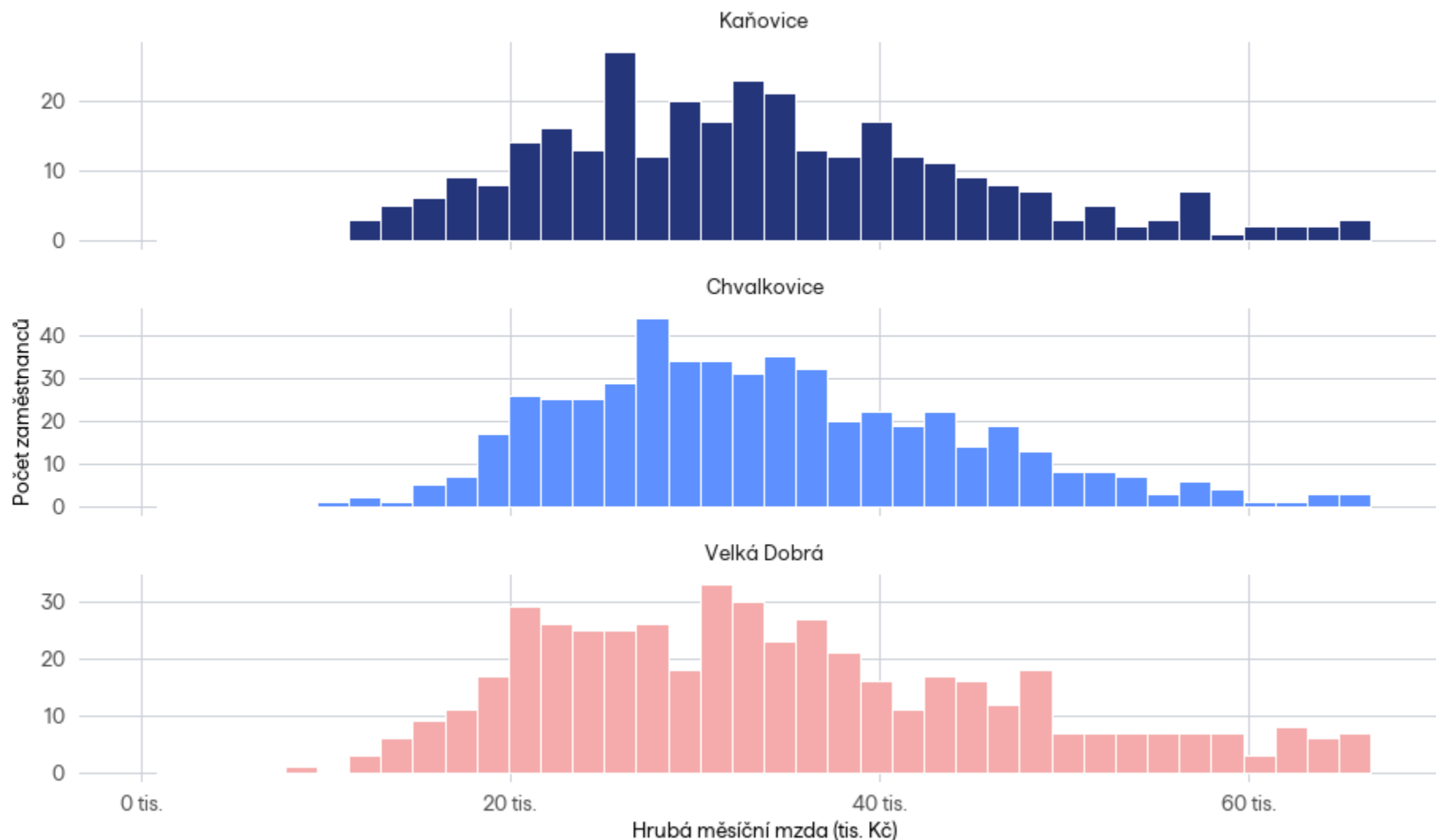
Průměrné mzdy po zahlazení (Fay-Herriot)



Příklady tří menších obcí – rozdíly jsou

Výsledek: mzdy ve třech obcích

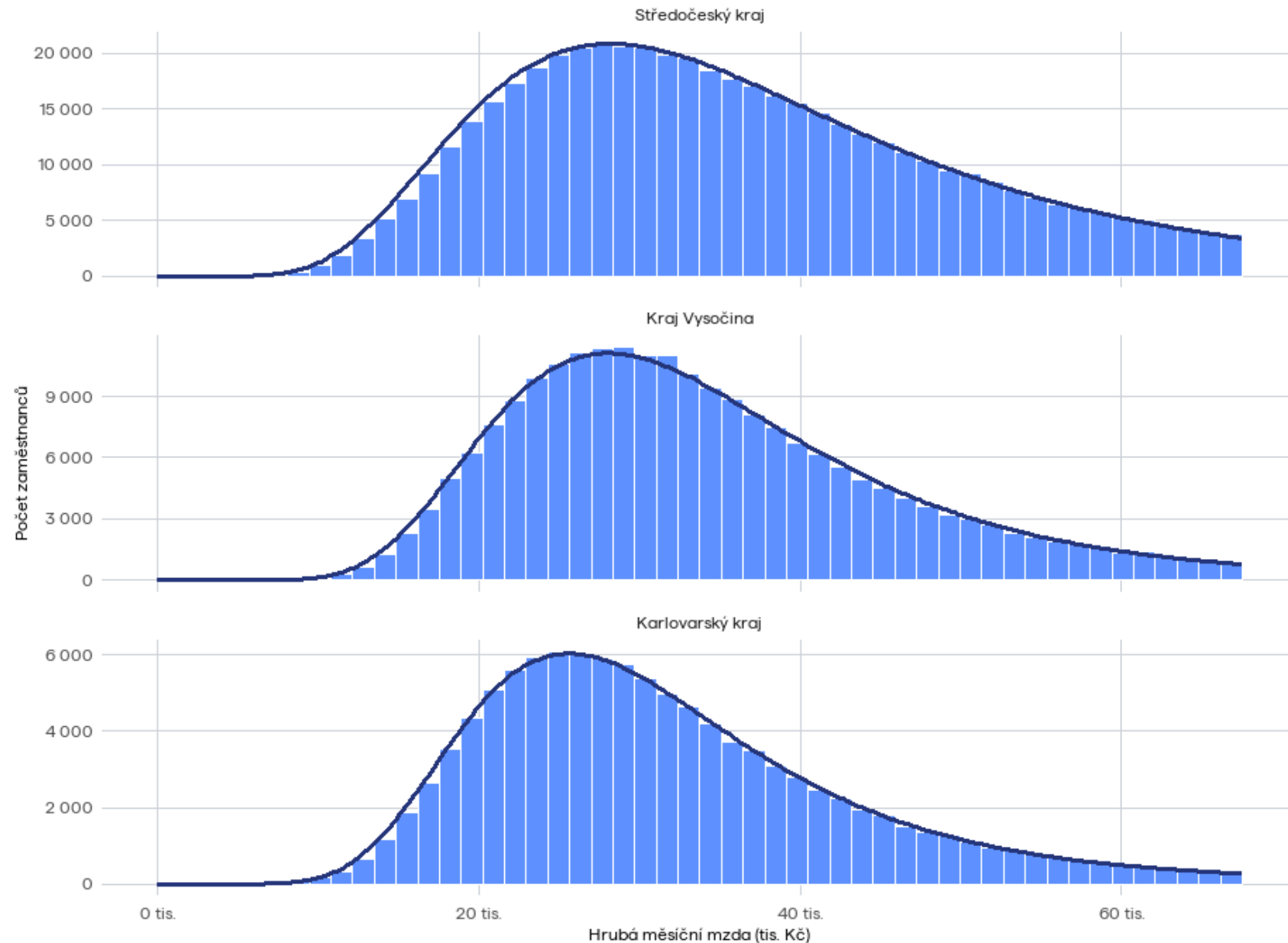
Simulované hrubé měsíční mzdy zaměstnanců



Ale krajské struktury sedí

Kontrola: simulace vs. ISPV

Histogram = simulované mzdy · křivka = ISPV (směs podle vzdělání)



Význam pro RUD:

Co nám taková data umožňují

Můžeme zkoušet sofistikovanější výpočty

Diferenciované křivky podle kraje

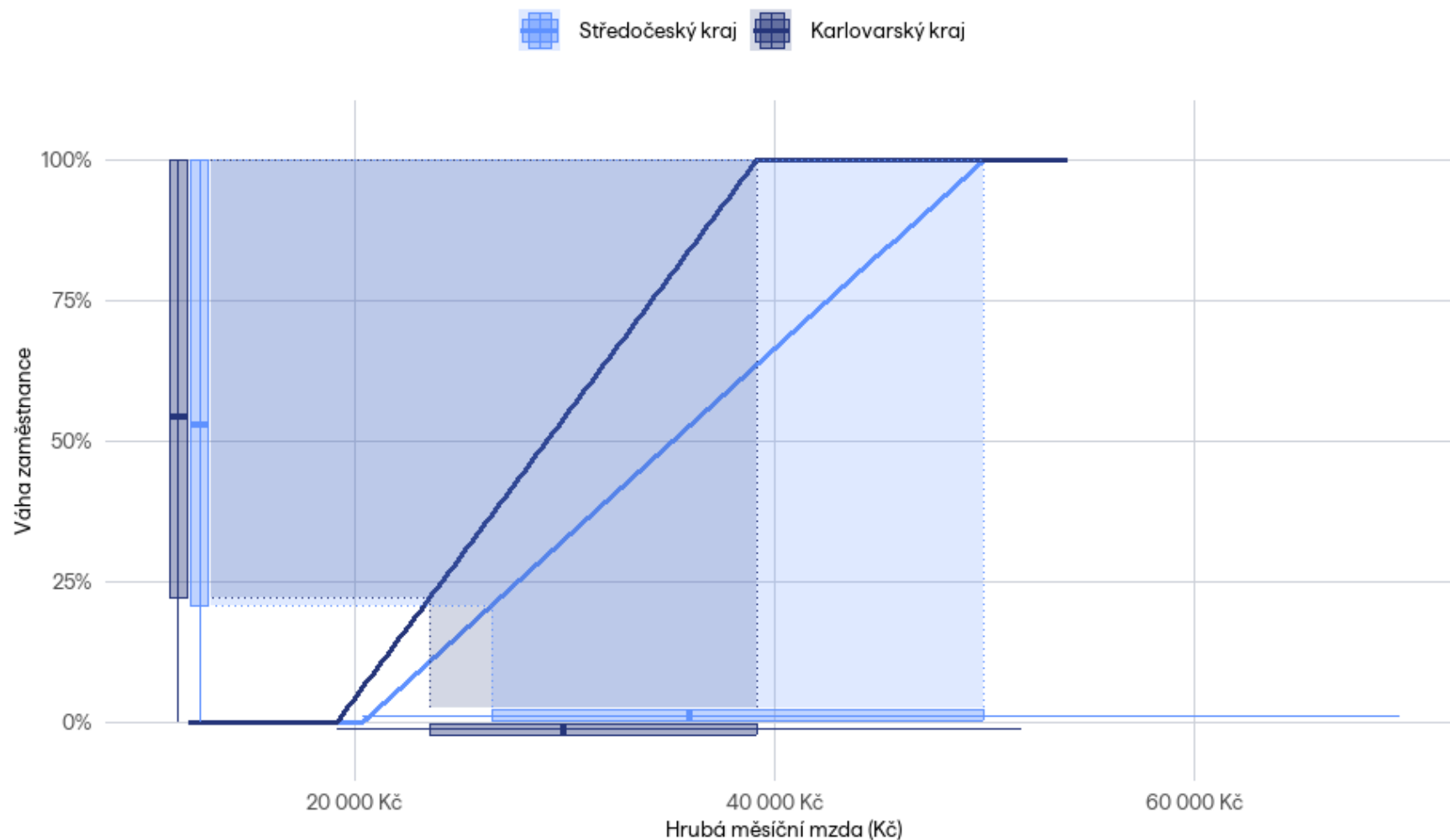
Diferenciované křivky · projekce podle kraje

Jak by vypadalo
komplikované vážení
zaměstnanců v praxi

Ověříme si, že to jde

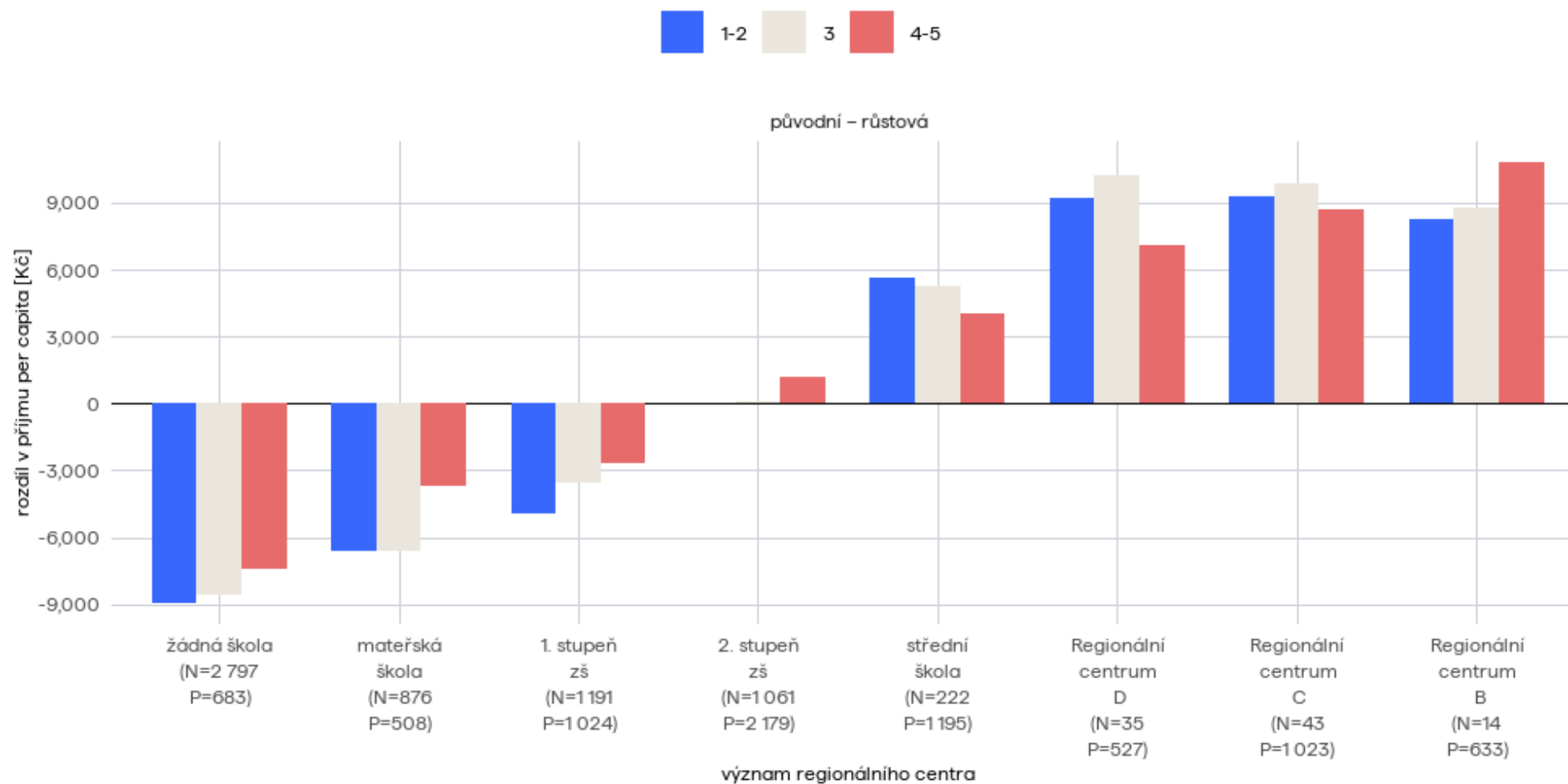
Vychytáme mouchy a
rozpory

Můžeme komunikovat
výsledky a myšlenku
ostatním



Můžeme zkoumat agregované dopady

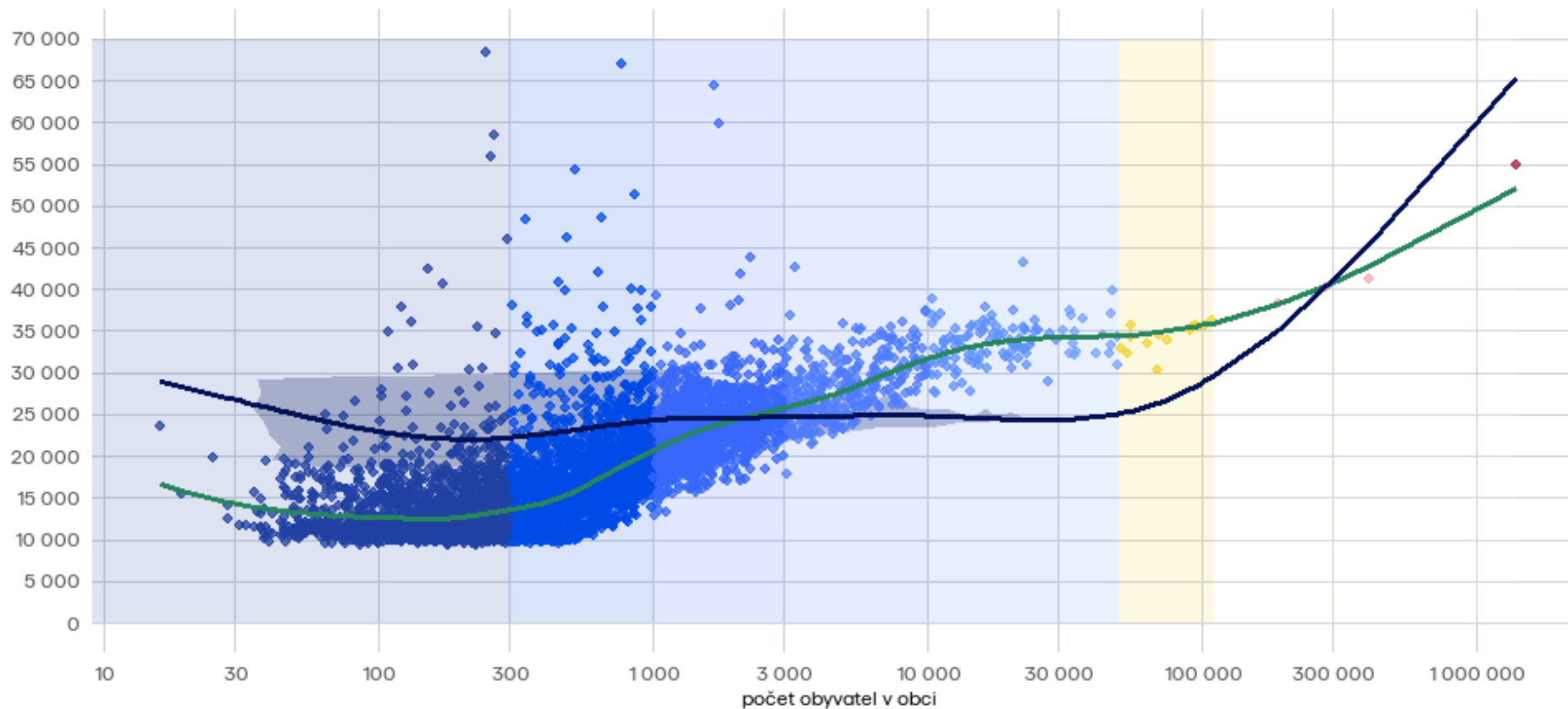
Index socioekonomické rozvinutosti



A posuzovat reformní návrhy všech podob

příjmy z RUD na obyvatele [Kč] (růstová varianta)

◆ do 300 ◆ 300 - 1 tisíc ◆ 1 tisíc - 3 tisíce ◆ 3 tisíce - 10 tisíc ◆ 10 tisíc - 30 tisíc ◆ 30 tisíc - 50 tisíc ◆ nad 50 tisíc



Význam syntetických dat:

Závěr

K čemu to bylo analytikům?

- 1 Ukázali jsme, jak **obcházíme** momentální **absenci** dat
- 2 Na **proof-of-concept** a **ladění modelu** je to dobré
- 3 Simulace může překlenout **prostor** mezi **neformálním návrhem** a **formální spoluprací** nad skutečnými daty

K čemu by to bylo poskytovatelům?

- ❶ Existují **citlivá data**, která se těžko zveřejňují
- ❷ Na **granularitě** ale někdy záleží víc než na přesnosti
- ❸ Hodně **dobré práce** se dá udělat i s daty, která **nejsou pravdivá**, jen **uvěřitelná**

**Když nesmíte dávat pravá data, hodte nám
syntetická**

Kontakt

Jiří Münich

jiri.munich@paqresearch.cz

Františka Tranová

frantiska.tranova@paqresearch.cz