

v databázi za školní rok 2011/2012 přehled o počtu chybějících dat u osmi proměnných použitých v modelu socioekonomického indexu.

**Tabulka č. 8 Rozložení chybějících dat v databázi**

Počet chybějících dat	Počet žáků
0	708 071
1	28 011
2	79 439
3	19 321
4	6 554
5	10
8	6 493

Samozřejmě je nemožné získat odhad socioekonomického indexu pro žáky, u nichž nemáme vůbec žádné informace (6 493 žáků). IRT modely jsou ovšem schopné poskytnout odhad tam, kde je dostupná alespoň jedna informace (841 406 žáků). Při použití IRT modelu tedy ztrácíme pouze necelé 1 % žáků, v porovnání s 16 %, která bychom ztratili, kdybychom nepoužili žádný imputační model pro dosazení chybějících dat.

Na tato kompletní či částečně nekompletní data byl aplikován dvouparametrický IRT model. Vzhledem k velmi proměnlivým korelacím mezi položkou a celkem zvolili výzkumníci dvouparametrický model, který váží proměnnou pro výpočet socioekonomického indexu podle korelace mezi položkou a celkovým souborem všech položek. Proměnné „příjem“, „dosazené vzdělání“ a „ekonomická aktivita“ tak mají v modelu vyšší váhu než ostatní proměnné, jako jsou třeba „rodina“ a „CPAS“.

Tabulka č. 9 ukazuje na úrovni škol a žáků korelace mezi socioekonomickým indexem spočteným pomocí této metody a indikátorem vzdělávacích obtíží, kterým je zde míra odkladů školní docházky a opakování ročníku. Čím vyšší je socioekonomický index, tím nižší je míra odkladů školní docházky a opakování ročníku.

**Tabulka č. 9 Korelace mezi individuálním socioekonomickým indexem (model 3) a vzdělávacími obtížemi (odklady školní docházky a opakování ročníku)**

Úroveň	Primární vzdělávání	Sekundární vzdělávání
Žák	-0,28	-0,30
Škola	-0,65	-0,74

Aby otestovali, zda je tento individuální socioekonomický index vhodnější než územně založený socioekonomický index, agregovali ho výzkumníci přes distrikty a pak přiřadili každému žáku průměr jeho distriktu. Potom porovnali vlastnosti tohoto uměle vypočteného územně založeného indexu a individuálního socioekonomického indexu vypočteného výše popsáním způsobem.

Tabulka č. 10 ukazuje rozklad rozptylu obou dvou indexů. Je vidět, že výpočet socioekonomického indexu přes distrikt ztlačuje variabilitu uvnitř škol. Agregováním individuálního socioekonomického indexu přes distrikty se totiž smazává veškerá socioekonomická variabilita uvnitř distriktů. Například na úrovni primárního vzdělávání se rozptyl individuálního socioekonomického indexu snížil z hodnoty 1,18 (0,28 + 0,90) na 0,28 (0,15 + 0,13). To znamená, že se ztratí 75 % socioekonomické variability, když se každému žáku přiřadí průměrný socioekonomický index jeho distriktu, a to je opravdu citelná ztráta. Dochází k ní převážně na vnitřní úrovni (z hodnoty 0,90 na 0,13), ale objevuje se také na meziškolní úrovni (z hodnoty 0,28 na 0,13).