APPENDIX

Model description

The two-level linear growth curve model with a cross-level interaction effect with cluster-mean education is represented by the following equation:

$$
\begin{aligned}
L1: \sqrt{Y_{tj}} &= \beta_{0j} + \beta_{1j}T_{tj} + \beta_2 EDU_{tj}^{CWC} + \beta_3 lnPOP_{tj}^{CWC} + \beta_4 POV_{tj}^{CWC} + \epsilon_{tj} \\
L2: \beta_{0j} &= \gamma_{00} + \gamma_{01} EDU_j^{CM} + \gamma_{02} lnPOP_j^{CM} + \gamma_{03} POV_j^{CM} + \gamma_{04} TR_j + \mu_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11} EDU_j^{CM} + \mu_{1j}
\end{aligned}
$$

Error terms are all assumed normally distributed:

$$
\begin{aligned}
\epsilon_{tj} &\sim \mathrm{N}(0, \sigma_\epsilon^2) \\
\mu_{0j} &\sim \mathrm{N}(0, \sigma_{\mu_0}^2) \\
\mu_{1j} &\sim \mathrm{N}(0, \sigma_{\mu_1}^2)
\end{aligned}
$$

Consulting the $L1$ part of the equation: β_{0j} are random intercepts, $\beta_k X_{tj}^{CWC}$ are the fixed time-variant coefficients where variables are centered-within-cluster, $\beta_{1j}T_{tj}$ is a time-variant trend variable where the first year is set to 0, and ϵ_{tj} is the level-1 error term. In the $L2$ part of the equation, γ_{00} is the mean municipal level intercept, $\gamma_{0k} X_j^{CM}$ are coefficients for level 1 covariate cluster-means (CM), $\gamma_{04} TR_j$ is a coefficient for median travel time to nearest pharmacy, while μ_{0j} is the intercept variance component. The linear trend variable is modeled as a random effect with μ_{1j} variance component $\gamma_{11} EDU_j^{CM}$. $\beta_2 EDU_{tj}^{CWC}$ is a cross-level interaction between the cluster-mean education level across the time-period and the random linear trend. The term $\beta_2 EDU_{tj}^{CWC}$ was removed in the final model to address the issue of simultaneous growth.

Table A1: Model 1 includes the time-variant education predictor, model 2 is the same as the in-text model. This table aims to show the consequences of simultaneous growth on the estimated trend coefficient and confidence intervals.

	$\sqrt{Dispensed\ prescriptions\ per\ 100\ children}$	
	Model 1	Model 2
Level 1		
Trend	−0.015 (−0.050, 0.019) [.385]	−0.053 (−0.066, −0.039) [<.001]
Poverty	−0.098 (−0.125, −0.071) [<.001]	−0.098 (−0.125, −0.070) [<.001]
Population (ln)	1.562 (0.210, 2.914) [.024]	1.265 (−0.061, 2.592) [.062]
Education	−0.069 (−0.127, −0.010) [.021]	
Level 2		
Education	-0.004 (−0.029, 0.021) [.751]	−0.002 (−0.027, 0.023) [.892]
Population (ln)	0.409 (0.292, 0.527) [<.001]	0.408 (0.290, 0.525) [<.001]
Poverty	−0.085 (−0.130, −0.040) [<.001]	−0.085 (−0.130, −0.041) [<.001]
Travel	−0.0003 (−0.0004, −0.0003) [<.001]	−0.0003 (−0.0004, −0.0003) [<.001]
Trend×Education (L2)	−0.003 (−0.005, −0.0005) [.019]	−0.0034 (−0.006, −0.001) [.005]
Intercept	5.271 (5.072, 5.471) [<.001]	5.459 (5.340, 5.578) [<.001]
Var. Comp.		
Std. Dev. μ_1	.0929	.0927
Std. Dev. μ_0	1.0912	.8647
Misc.		
ρ Comp. Symm.	.000	.000
Groups	426	426
Observations	4,499	4,503
Log Likelihood	−6,431.018	−6,442.764
Akaike Inf. Crit.	12,892.04	12,913.53
Bayesian Inf. Crit.	12,988.21	13,003.3
Note:	95% CI in parentheses. P-values in square brackets.	

Simultaneous growth and MLM interpretation under centering scheme

Model 1 includes all level 1 covariates. Model 2 excludes the group-mean centered education (L1) covariate due to simultaneous growth issues resulting in collinearity between L1 education and trend.

This contrast table shows the effect of simultaneous growth on estimated parameters. The only difference between the models is the removal of the L1 group-mean centered education indicator. Confidence intervals are shown in parentheses.

Group-mean centering level 1 covariates leads to orthogonal relationships between levels; the correlations between level 1 and level 2 covariates are equal to 0. In a model without the uncentered trend variable, excluding level 1 coefficients would not affect level 2 estimates under group-mean centering. In fact, the estimates would be the same regardless of whether level 1 covariates were even in the model [30]. However, since the trend variable is *not* centered, some correlation will exist between levels through correlation with the trend variable, explaining the minor changes in level 2 coefficients. These changes are unsubstantial and only result in minor changes in L2 estimates.

Simultaneous growth leads to a very simple issue of near perfect collinearity between L1 education and the trend variable. This is the reason for the dramatic change in the trend coefficient size and confidence interval. Simply put, the trend effect in model 1 is biased due to collinearity with the L1 education covariate. While there are ways to deal with this problem through *multivariate* growth curve modeling [32], we are primarily interested in the cross-level interaction effect between education traits and the random trend. As such, we prefer the more parsimonious modeling option removing the cluster-mean centered education variable from the level 1 part of the equation.

Interpreting coefficients under centering scheme

Centering and cross-level interactions changes the interpretation of certain coefficients. We base the interpretation on model 2 and focus on three main coefficient interpretations a) the main trend effect and its variance, b) the main trait education effect and c) the cross level interaction term.

Due to grand-mean centering L2 covariates and the inclusion of an interaction term, the main trend effect ($-.015$) is interpreted as the expected square root dispense rate trend for municipalities with a mean level of trait education (21.15%), *ceteris paribus*. This is a random coefficient, and its random parameter μ_1 suggests that the standard deviation from the fixed term is equal to $.919$. The main education effect ($-.002$) is the expected effect of education at $T = 0$ (2006, trend is not centered). This is clearly shown by the very similar intercepts in figure 2 and 3. Lastly, the interaction term (-$0034.$) is the expected decrease in trend for every pp increase in education traits. This model is the basis for figures 2 and 3.

For other L1 coefficients (sans the trend coefficient), a one-unit increase entails a one unit change from a covariates given group mean. The coefficient is thus the average effect of a one unit increase from a given group mean, *ceteris paribus*.

Centering and growth

Notably, we choose not to center the level 1 trend variable for two reasons; firstly, the panels are only slightly imbalanced. Centering the trend variable on the group means practically results in a grand mean centered trend variable (correlation with uncentered trend indicator: $r = .97$), with

Svalestuen S, *et al. BMJ Open* 2022; 12:e058491. doi: 10.1136/bmjopen-2021-058491

no real consequences to the coefficient estimates. The only consequence is on the intercepts and the intercept variance due to the zero point being established in 2011 for all but a few groups. Secondly, the model is a linear random growth curve model. Centering the trend covariate is more of an issue in situations where a polynomial growth curve might be fitted.

Intercept and slope correlation

Intercepts and slopes are negatively correlated at $r = -.597$. This is a natural consequence of bounded data; dispensing rate cannot be less than 0. Municipalities with low starting dispensing rates will naturally not be able to reduce dispensing rates as much as those with higher starting dispensing rates. This is of no particular concern for estimating the interaction term; indeed, the non-significant main education coefficient implies that the intercept variance is not explained by mean population education levels. This is also clear when investigating figure 2 in the main text.

SUPPLEMENTARY FIGURES AND TABLES

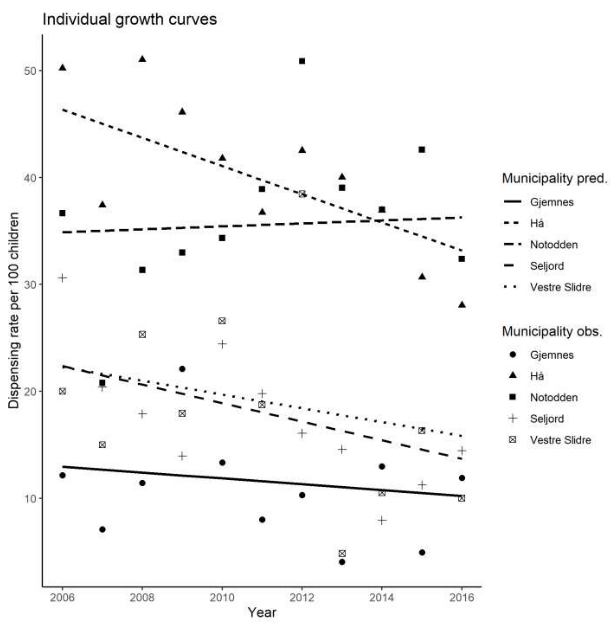

Figure A1: *Linear growth curve predictions and observations from a simple random trend null-model for five random municipalities. Municipalities were randomly sampled from a strata of slope quantiles to ensure that slope variance was represented in the figure. Note that the Y-axis is scaled by min-max observations in the subsample, not the entire distribution.*

Table with transformed and untransformed dispense rates

Table A2: Multilevel growth curve models. Both models include all covariates. Model 1 uses the square-root transformed dispense rates as outcomes. This model is used for prediction (figures 2 and 3) and evaluation of statistical significance. Model 2 uses the dispense rate as the outcome.

	$\sqrt{\text{Dispensed Rx per 100 children}}$	Dispensed Rx per 100 children
	(1)	(2)
Level 1		
Trend	−0.053 (−0.066, −0.039) [<. 001]	−0.608 (-.750, -.466) [< . 001]
Poverty	−0.098 (−0.125, −0.070) [<. 001]	−1.061 (-1.352, -.769) [< . 001]
Population (ln)	1.265 (−0.061, 2.592) [.062]	13.980 (.278, 27.683) [.046]
Level 2		
Education	−0.002 (−0.027, 0.023) [.892]	0.026 (-.239, .291) [. 848]
Population (ln)	0.408 (0.290, 0.525) [< .001]	3.983 (2.767, 5.199) [< .001]
Poverty	−0.085 (−0.130, −0.041) [< .001]	−0.845 (-1.311, -.379) [. 001]
Travel	−0.0003 (−0.0004, −0.0003) [< .001]	−0.003 (-.003, -.002) [< . 001]
Trend × Education (L2)	−0.0034 (−0.006, −0.001) [.005]	−0.041 (-.066, -.017) [. 001]
Intercept	5.459 (5.340, 5.578) [< .001]	32.689 (31.425, 33.952) [< .001]
Var. Comp.		
Std. Dev. μ_1	.0927	.918
Std. Dev. μ_0	.8647	11.54
Misc.		
ρ Comp. Symm.	.000	.000
Groups	426	426
Observations	4,503	4,503
Log Likelihood	−6,442.764	−17,097.230
Akaike Inf. Crit.	12,913.53	34,222.460
Bayesian Inf. Crit.	13,003.3	34,312.240
Note:	*p<0.05; **p<0.01; ***p<0.001	
	95% CI in parentheses. P-values in square brackets.	

Dependent variable distribution before and after square root transformation

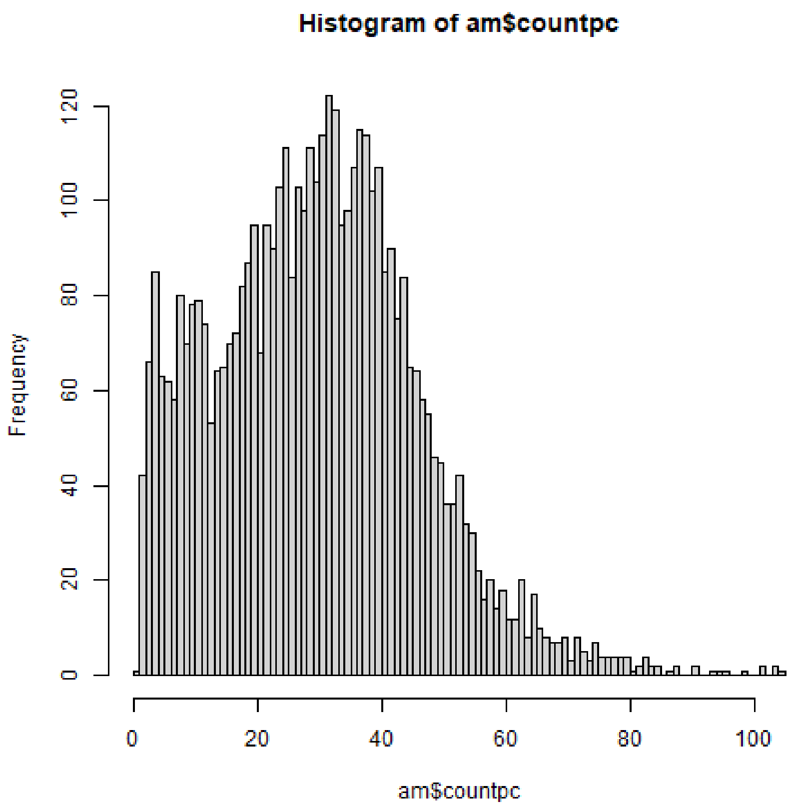

Figure A2: Dispense rate distribution before square root transformation. The distribution is closer to a Poisson distribution, due to the natural bounds of the data.

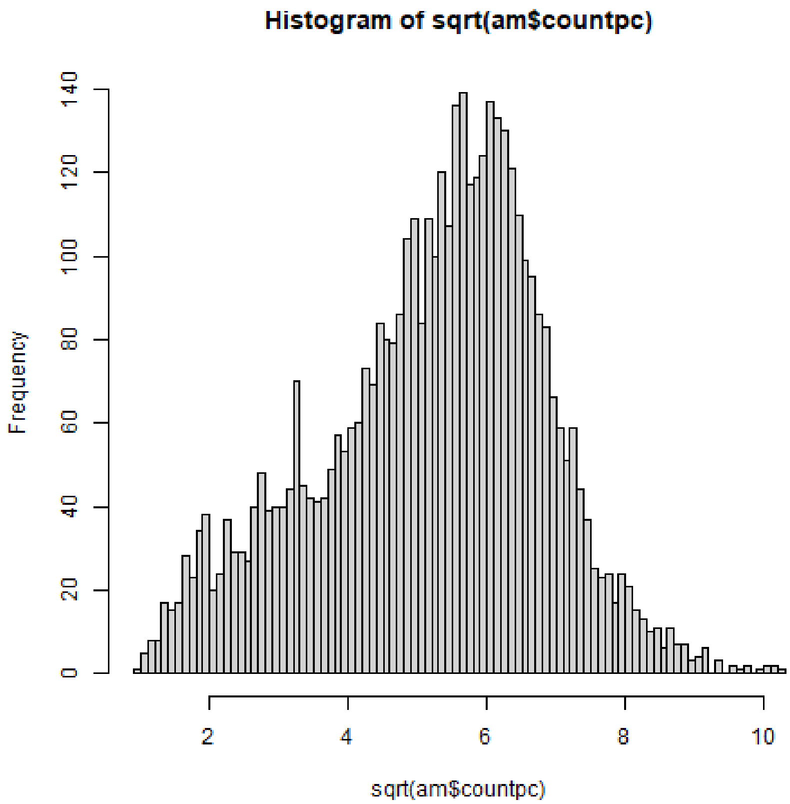

Figure A3: Dispense rate after square root transformation. Where the log-transformation (not shown) aggressively overcorrects the issue, leading to a worse fit than the untransformed version of the model, the square root transformation only moderately corrects the distribution, making residuals more well-behaved than the untransformed model. We emphasize that we performed this transformation to solve a statistical issue particularly present when investigating the residuals vs. the fitted values, and as such were guided by the data rather than theory. However, as the prediction plots, significance tests, and coefficients show, these modeling changes do not affect results in a significant way.

Residual plots main model

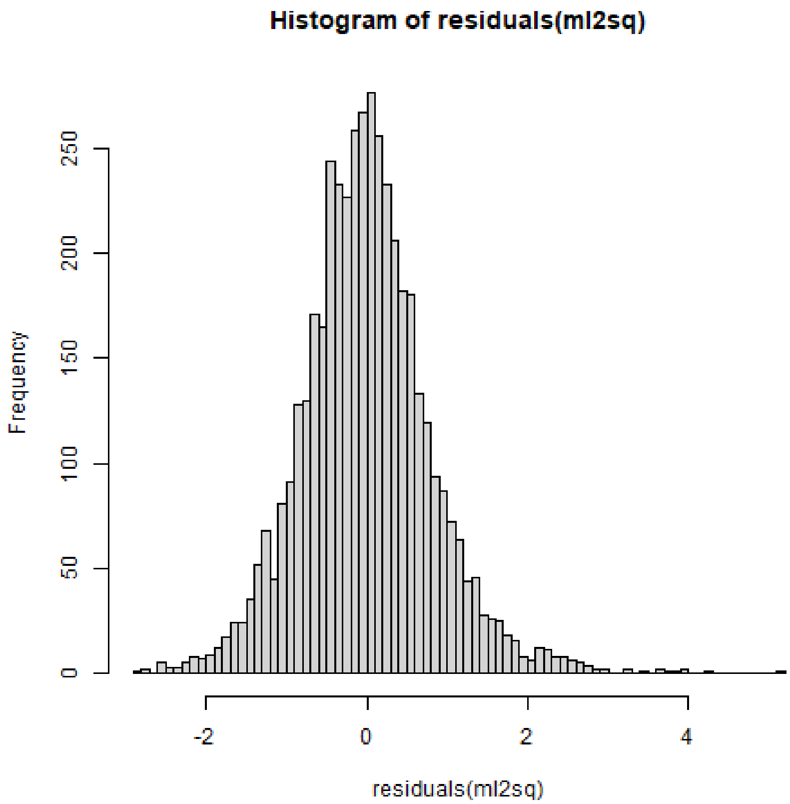

Figure A4: Level 1 Residual distribution after square root transformation of the dependent variable. While a marginally longer tail on positive residuals, we find no particular issues with this distribution.

Svalestuen S, *et al. BMJ Open* 2022; 12:e058491. doi: 10.1136/bmjopen-2021-058491

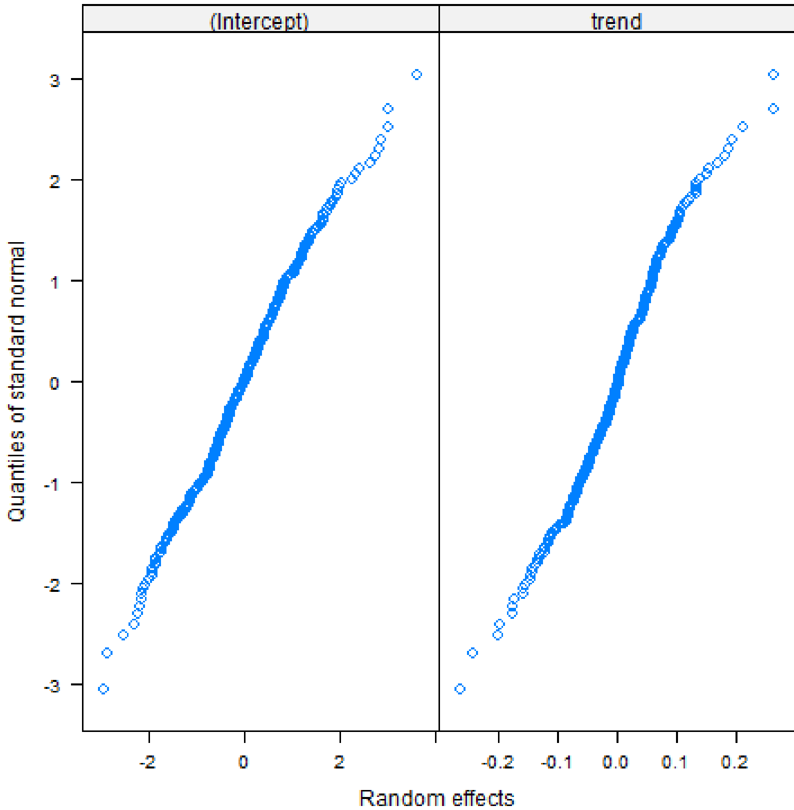

Figure A5: QQ-plot of the random terms in the model. We find that these are approximately normally distributed.

Svalestuen S, *et al. BMJ Open* 2022; 12:e058491. doi: 10.1136/bmjopen-2021-058491

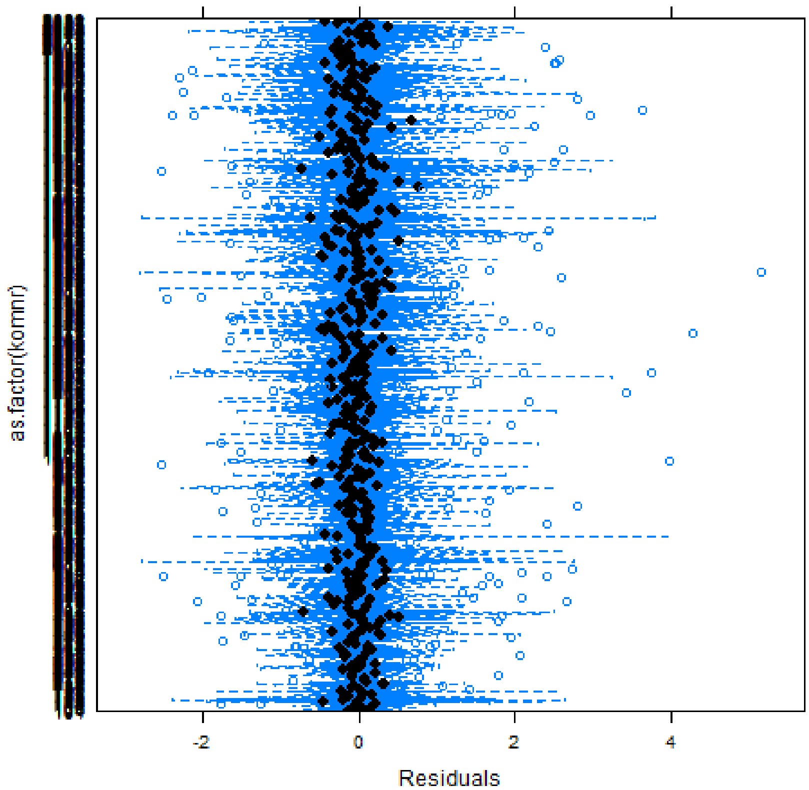

Figure A6: Level-1 residuals by municipality. Residuals seem overall to be centered at 0 with random deviation from this mean. Some differences in variance between municipalities is expected, as the number of repeat observations is relatively small (11).

Svalestuen S, *et al. BMJ Open* 2022; 12:e058491. doi: 10.1136/bmjopen-2021-058491

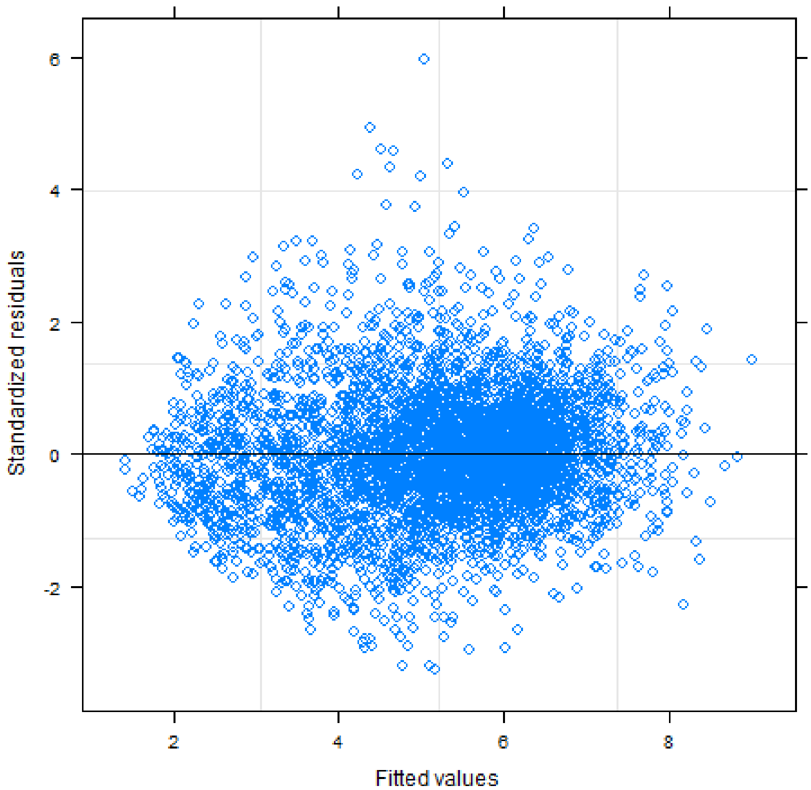

Figure A7: Standardized residuals vs. fitted values plot. We saw some problems with heteroskedasticity in the unadjusted model. While logarithmic transformation aggressively overcorrected the issue, the square root transformation adjusts for the moderate skewness and provides confidence to estimated standard errors.

Svalestuen S, *et al. BMJ Open* 2022; 12:e058491. doi: 10.1136/bmjopen-2021-058491

Full version of summary statistics table

Statistics	N	Mean	St. Dev.	Min	Max
Pooled					
Dispensed Rx/100 chld.	4,519	29.7	16.3	0.9	104.9
Education	4,515	21.2	5.9	9.1	51.9
Population	4,519	11,885	35,479	200	658,390
Poverty	4,518	10.0	2.4	3.7	21.8
Within					
Dispensed Rx/100 child	4,519	0.00	9.58	−40.38	74.42
Education	4,515	0.00	1.87	−5.25	5.97
Population	4,519	0.00	2,180	−60,394	59,5842
Poverty	4,518	0.00	1.07	−3.46	5.76
Between					
Dispensed Rx/100 chld.	428	29.0	13.5	2.8	70.3
Education	428	21.0	5.6	11.2	48.2
Population	428	11,505	34,795	212	598,805
Poverty	428	10.0	2.2	5.1	18.6
Travel (sec.)	426	1,674	1,882	182.0	13,129

Table A3: Summary statistics grouped by levels. Pooled statistics include summary statistics for yearly observations for all municipalities before centering. The dependent variable. The within section shows descriptive statistics for all cluster-mean centered covariates, that is the level 1 parameters in the model. Note the mean 0 ensuring no correlation between level 1 and level 2 covariates. The between section represents the level 2 variables used in the model. These are 428 cluster-means for all covariates excluding travel times, due to municipality mergers before data collection.

Svalestuen S, *et al. BMJ Open* 2022; 12:e058491. doi: 10.1136/bmjopen-2021-058491