Supplementary Material

Out-of-pocket payments and income-related inequality in unmet medical needs: within-country evidence from Central and Eastern Europe

This document collects the appendix tables (A1–A12) and figures (A1–A3) referenced in the main text. Unless stated otherwise, all estimates are two-way (country and year) fixed-effects models with Driscoll–Kraay standard errors (HC0, maxlag = 2), estimated on the balanced CEE-11 panel over 2009–2023 (N = 164). Significance: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Table A1–A3. Annex robustness: alternative specifications of the baseline model.

	(A1) Gap ratio Q1/Q5	(A2) Country-specific trends	(A3) First differences
OOP share (t−1)	−1.077*** (0.208)	−0.083 (0.097)	0.112 (0.143)
N	122	164	153
R^2	0.160	0.780	0.189

Notes: *Columns correspond to Table A1 (gap as the ratio Q1/Q5), Table A2 (baseline plus country-specific linear time trends), and Table A3 (model in first differences). For the gap ratio, 42 observations are undefined (Q5 = 0) and 72 of the 122 remaining have Q5 < 0.5 p.p., inducing denominator instability (see main text).*

Table A4. Cause decomposition of the quintile gap, by reported reason.

	Combined	Too expensive	Waiting list	Too far
OOP share (t−1)	0.481*** (0.161)	0.470*** (0.168)	−0.004 (0.026)	0.015 (0.016)
N	164	164	164	164
R^2	0.387	0.337	0.203	0.133

Notes: *Outcome is the Q1–Q5 gap (percentage points) by reason. Reasons are mutually exclusive main-reason categories; the waiting-list series is reconstructed residually (combined − too expensive − too far).*

Table A5. Gini control: alternative sources and specifications.

	(1) SWIID (baseline)	(2) No Gini	(3) Spliced	(4) Eurostat
OOP share (t−1)	0.470*** (0.168)	0.514*** (0.161)	0.486*** (0.169)	0.241** (0.110)
N	164	164	164	110
R^2	0.337	0.295	0.314	0.170

Notes: *The point estimate on the OOP share is stable across Gini treatments. The Eurostat variant (column 4) is available only from 2014 (N = 110).*

Table A6. Mundlak (correlated random effects) test.

Test	Statistic	df	p-value
Joint significance of country means (Wald, vcovSCC HC0 maxlag = 2)	$\chi^2 = 73.85$	6	6.6×10^{-14}

Notes: *The Mundlak test augments the model with country means of all time-varying regressors and tests their joint significance; rejection favours fixed over random effects. Under serial correlation the classical Hausman statistic is unreliable, so this cluster-robust Wald test is the preferred discriminant. The country means are jointly highly significant, so random effects are rejected and the two-way fixed-effects specification is retained.*

Table A7. Small-cluster inference and population weighting.

Inference / specification	Result
Cluster-robust CR2 test (Satterthwaite)	df ≈ 1 — uninformative at G = 11; inference rests on the bootstrap
Restricted wild cluster bootstrap, Webb weights, 9,999 replications (H_0: $\beta = 0$)	t = 4.56; p = 0.0022

Inference / specification	Result
Population-weighted two-way FE (Driscoll–Kraay, lag 2): β on OOP share (t−1)	0.348*** (0.114); t = 3.07; p = 0.008; N = 164

Notes: *With only 11 clusters the CR2 (Satterthwaite) test has approximately one degree of freedom and is effectively powerless for a persistent regressor, so small-sample inference rests on the restricted wild cluster bootstrap with Webb weights (null imposed; HC1-clustered reference statistic t = 4.56), which confirms significance. The population-weighted specification checks that the result is not an artifact of equal weighting across countries of very different size.*

Table A8. Leave-two-out jackknife (summary).

Sample	Pairs	b range	Sig. 5%	Sig. 10%
Full, 2009–2023	55	[0.222, 0.567]	48/55	54/55
Pre-COVID, ≤2019	55	[−0.120, 0.412]	10/55	10/55
— pre-2019, Latvia dropped by the pair	10	[0.163, 0.412]	10/10	10/10
— pre-2019, Latvia retained in sample	45	[−0.120, 0.257]	0/45	0/45

Notes: *Every pair excludes two countries at once. In the full sample the estimate is robust to any two-country exclusion. In the pre-2019 subsample, significance is entirely conditional on Latvia: all 10 pairs that drop Latvia are significant, whereas none of the 45 pairs that retain Latvia are. Full pair-by-pair estimates are in the replication file etap6_leave_two_out.csv.*

Table A9. Alternative OOP scaling — out-of-pocket per capita.

Specification (no GDP control)	b	SE	p	N
A: OOP per capita, PPS (derived)	−1.016	1.157	0.381	164
B: OOP per capita, US$ (nominal, GHED)	−1.386	1.245	0.267	164

Notes: *Both variants are negative and non-significant. OOP measured as a per-capita level does not reproduce the result obtained for the OOP financing share, consistent with the share (financing intensity) being the theoretically relevant construct.*

Table A10. Alternative OOP scaling — out-of-pocket as % of GDP.

Parametrization	Term	b	SE	p	N
(1) OOP/GDP as main regressor	OOP/GDP (t−1)	0.717	1.621	0.659	164
(2) Baseline + OOP/GDP control	OOP share (t−1)	0.721***	0.136	4.4×10^{-7}	164
	OOP/GDP (t−1), added	−5.891***	1.382	3.8×10^{-5}	164
(3) Decomposition: OOP/CHE + CHE/GDP	OOP share (t−1)	0.417**	0.168	0.014	164
	CHE/GDP (t−1), intensity	−1.017**	0.450	0.026	164

Notes: *The OOP share (OOP/CHE) retains its sign and significance once the expenditure-intensity dimension is partialled out; the raw OOP/GDP ratio is not itself informative for the quintile gap.*

Table A11. Variance inflation factors, within-transformed regressors.

Regressor	VIF (within)
OOP share (t−1)	1.11
ln GDP per capita (PPS)	1.60
At-risk-of-poverty rate	1.35
Unemployment rate	1.15
Population share 65+	1.41
Gini (SWIID)	1.42

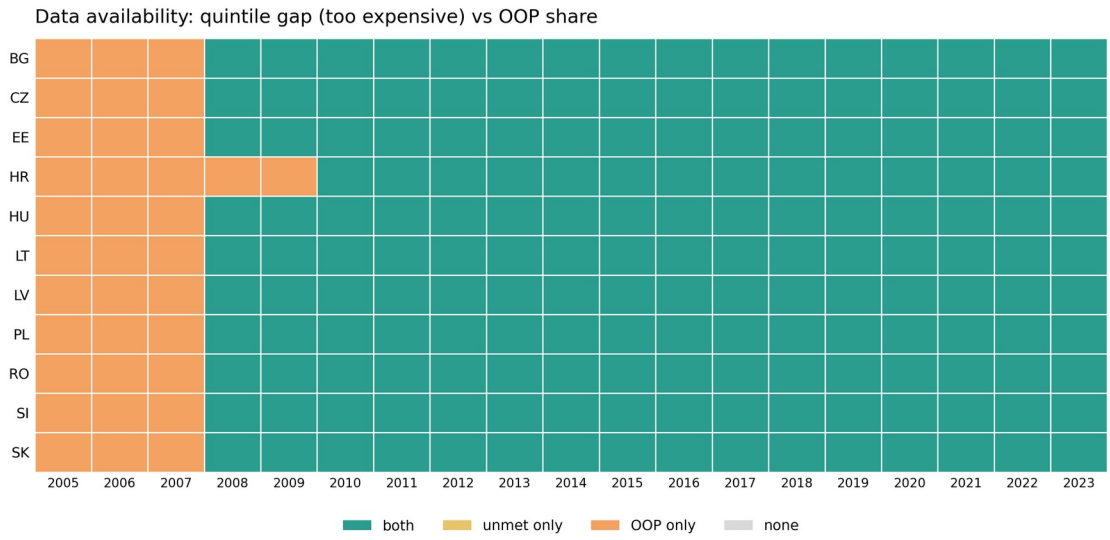
Notes: *VIF computed on regressors after two-way (country and year) within transformation via Frisch–Waugh–Lovell partialling-out — i.e., on the variation that identifies the fixed-effects estimator, not on raw series. All values are far below conventional thresholds (5/10).*

Table A12. Observation-level influence diagnostics (top 10 by Cook's distance).

Country	Year	Leverage	Cook's D	dfbeta (OOP share, t−1)
BG	2009	0.138	0.2363	−0.0018
LV	2010	0.096	0.1773	+0.0313
LV	2011	0.076	0.1411	+0.0197
LV	2017	0.061	0.0849	−0.0574
LV	2013	0.038	0.0635	+0.0003
BG	2011	0.067	0.0635	+0.0033
EE	2010	0.136	0.0627	+0.0033
LV	2021	0.059	0.0501	+0.0327
LV	2023	0.088	0.0468	+0.0344
LV	2014	0.023	0.0378	+0.0123

Notes: *Diagnostics computed on the FWL-residualized regression reproducing the within estimator (coefficient identity with plm asserted to 1×10^{-6}). 15 of 164 observations exceed the conventional 4/n Cook's-distance threshold (0.0244), concentrated in the Latvian episode (2010–2023) and Bulgaria (2009–2011). The largest |dfbeta| for the coefficient of interest is 0.057 (Latvia, 2017).*

Figure A1. Data availability: quintile gap (too expensive) versus OOP share, by country and year.

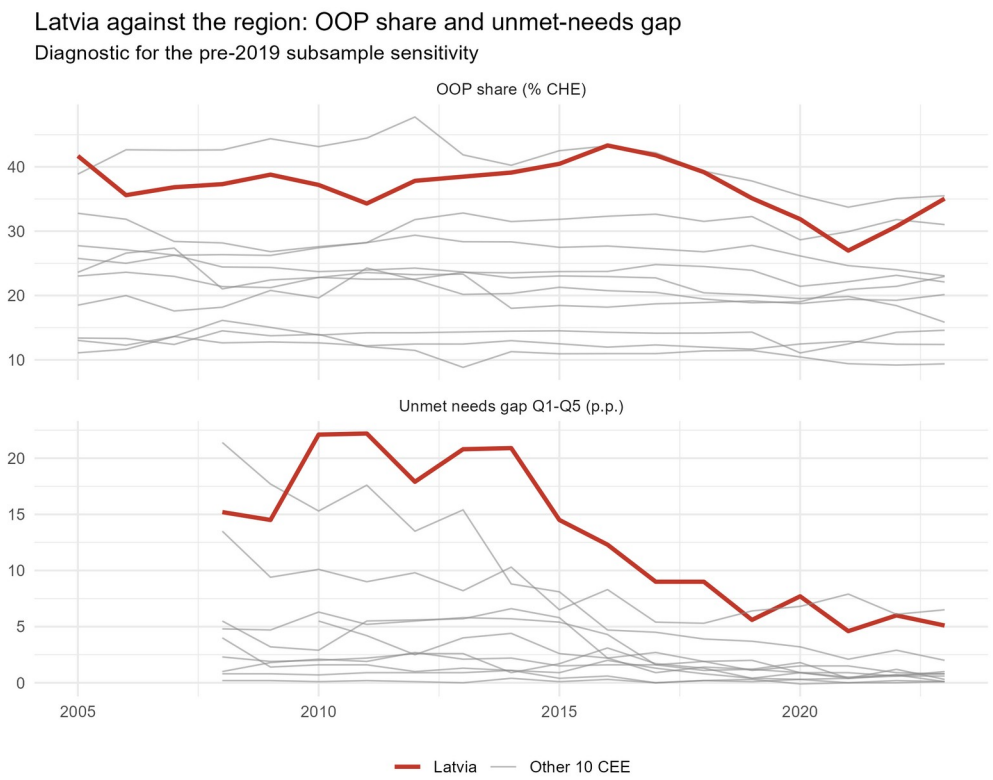


Notes: *"Both" indicates that the outcome (quintile gap) and the OOP share are jointly observed; the estimation window begins in 2009 (Croatia from 2010).*

Figure A2. Quintile gap (Q1 – Q5) in unmet needs, reason "too expensive", by country.

Notes: *Percentage points; shaded band marks the EU-SILC methodological break / pandemic period (2020–2021).*

Figure A3. Latvia against the region: OOP share and the unmet-needs gap.



Notes: *Diagnostic for the pre-2019 subsample sensitivity. Latvia (bold) combines a high OOP share with an unusually rapidly collapsing quintile gap, the pattern that drives the pre-2019 attenuation documented in Table A8.*