Utility performance in the Danube Region: a review of trends and drivers

Danube Water Program

February 2015

Contents

1 Summary and key findings

1. **Measuring utility performance** - WUPI is a best practice indicator that focuses on outputs/outcomes. As a result, it is highly correlated with the utility health indicator APGAR, but only weakly with production cost per m3.

2. **Utility performance and the relationship with tariffs** - Achieving higher WUPI scores is not only possible through higher utility efficiency but in many instances requires higher tariffs. Splitting WUPI into three parts - coverage, quality, management efficiency - we find that particularly coverage performance is related to higher tariffs. Quality is also positively related to tariffs, but to a lesser degree. In contrast, for a given level of coverage and quality, higher management efficiency turns out to be negatively related to tariffs. This suggests that performance-tariff trade-offs are limited to the areas coverage and quality. Higher management efficiency may eventually be even cheaper in terms of tariffs for consumers. From this point of view, increasing management efficiency should be a part of any strategy to improve performance, there seems to be little downside to it.

   While all three components are positively correlated and we would therefore hope that improved performance, e.g. through training, in one area might spill over to other areas, the obtained results suggest that in many instances it will still require higher tariffs.

3. **Performance over time and performance convergence** - Overall utility performance clearly improves over time and there is evidence of performance convergence in the Danube Region. Average WUPI scores improve over time and utilities with lower WUPI scores grow fast than those with initially higher scores. However, the improvements are on average not very large because of very large heterogeneity, between countries but also within countries.

4. **Drivers of utility performance: country level** - On the aggregate institutional level we find that EU membership or the presence of a regulatory agency is no guarantee for improved performance. Although countries that join the EU or put a regulator in place have a slightly higher WUPI after these events, utilities in other countries improved at the same pace or even faster. While this is not to say that EU membership or a regulator generally have adverse effects on utility performance, it seems to suggest that specific programs and adopted policies matter more than just formal status.

5. **Drivers of utility performance: utility level** - On the utility level, the initial result that private participation and larger scope are helpful for performance is not confirmed when looking at over time changes. Switching from public to private or back does not seem to affect performance.

6. **Drivers of utility performance: size** - Zooming in on the question of aggregation, we find that the average effect is very small, ranging from zero to 2 points of WUPI increase. In most cases the effect is not statistically different from zero.

Looking at the sub-components of WUPI, we find that typically aggregations have a positive effect on management efficiency, which improves rather quickly after the aggregation. In contrast, coverage indicators seem to worsen a few years after the aggregations.

However, while all measures of utility size (volume, connection density, number of managed systems) seem to be positively related to WUPI, the analysis shows that the effect of aggregations depends on the type of aggregation. By splitting aggregations into a volume effect, consumer density effect and number of systems effect, we find that most benefits accrue from an increased number of customers whereas the number of served towns has a negative effect. This suggests that the overall effect depends on the type of aggregation. The smaller the number of additional systems and the larger the number of additional customers the better.

In addition, we find that performance gains through aggregation are less likely if the utility is already large before the aggregation. Particularly, the benefits from additional customers through aggregations seem to disappear with increasing size.

7. **Aggregations and tariffs and cost -** Similarly to the results for WUPI, the effect of aggregations on tariffs and costs remains ambiguous. While the average aggregation has a positive effect on the former and a negative effect on the latter, the results are not statistically significant. The impact of aggregations seems to be very heterogeneous, sometimes positive, sometimes negative.

   However, as for WUPI we find that the type of aggregation appears to matter gravely: We find that customer density has a negative effect on tariffs and costs. Also ,increasing customer density has a more cost and tariff dampening effect if the utility is small. The number of systems is positively related to both variables, but is not statistically significant.

2 IBNET utility data

The main data for our analysis are from the International Benchmarking Network (IBNET) database. IBNET is a data repository initiated and maintained by the World Bank with the objective to improve the service delivery of water supply and sewerage utilities through the provision of international comparative benchmark performance information. Given the focus of the Danube Water Program, a World Bank led initiative for water and wastewater services in the Danube region, we use an IBNET sub-sample of 14 Central and Eastern European countries.

The utility coverage by IBNET varies strongly between countries, both in terms of the number of utilities as well as the population living in the service area of the utilities. The number of covered utilities by country and year is exhibited in Table 1. After removing observations with missing or inconsistent information, we have a unbalanced panel of 3506 utilities from 14 countries over 19 years.

Table 1: Utility information by country year

Year	Albania	Bosnia and Herzegovina	Bulgaria	Croatia	Czech Republic	Hungary	Kosovo	FYR Macedonia	Moldova	Montenegro	Romania	Serbia	Slovakia	Ukraine	Total
1995	0	0	0	0	0	22	0	0	0	0	0	0	0	0	22
1996	0	0	0	0	0	22	0	0	41	0	0	0	0	0	63
1997	0	0	0	0	0	22	0	0	41	0	0	0	0	61	124
1998	0	0	0	0	0	22	0	0	41	0	0	0	0	61	124
1999	0	0	0	0	0	22	0	0	41	0	0	0	0	61	124
2000	0	21	0	21	20	23	0	0	41	0	25	0	0	82	233
2001	0	21	0	21	20	23	0	0	41	0	25	0	0	83	234
2002	0	21	0	21	20	23	0	2	41	0	25	0	0	24	177
2003	0	21	0	21	20	23	0	2	41	0	28	0	2	24	182
2004	0	21	19	21	20	23	0	12	41	0	28	0	4	24	213
2005	30	20	19	0	20	24	0	15	41	0	24	0	5	16	214
2006	64	21	19	0	23	20	7	16	41	0	24	0	6	16	257
2007	55	21	19	0	23	20	7	16	41	0	24	26	7	16	275
2008	55	0	19	0	23	0	7	9	39	0	19	27	0	0	198
2009	58	0	30	0	23	0	7	14	39	0	20	27	0	0	218
2010	58	0	30	0	21	0	7	25	39	0	20	28	10	0	238
2011	58	0	30	0	21	0	7	27	39	0	0	28	11	0	221
2012	54	0	30	0	21	0	7	25	39	2	0	30	11	0	219
2013	57	0	0	0	21	0	7	33	39	2	0	1	10	0	170
Total	489	167	215	105	296	289	56	196	726	4	262	167	66	468	3,506

3 WUPI and other performance indicators

WUPI is an aggregate utility performance index based on 10 indicators. The indicators are scaled between 0 and 100 with cut-offs that are at 0 and 100 for actual ratios and at the 10th and 90th percentile for indicators that are not naturally limited on an interval. WUPI is calculated as the average of the 10 scaled indicators, i.e. indicators enter WUPI equally weighted. For utilities providing only water or sewerage, WUPI is based only on the relevant indicators. The indicators can be found in Table 2, distribution plots in Figure 1, and descriptive statistics by country in Table 3.

Table 2: WUPI Indicators

Indicator	**WUPI**	**WUPI water**	**WUPI sewerage**
Water coverage	X	X	
Sewerage coverage	X		X
Treatment coverage	X		X
Hours of service	X	X	
Blockages	X		X
Metering	X	X	
Non-revenue water	X	X	
Staffing	X	X	X
Collection ratio	X	X	X
Cost recovery	X	X	X

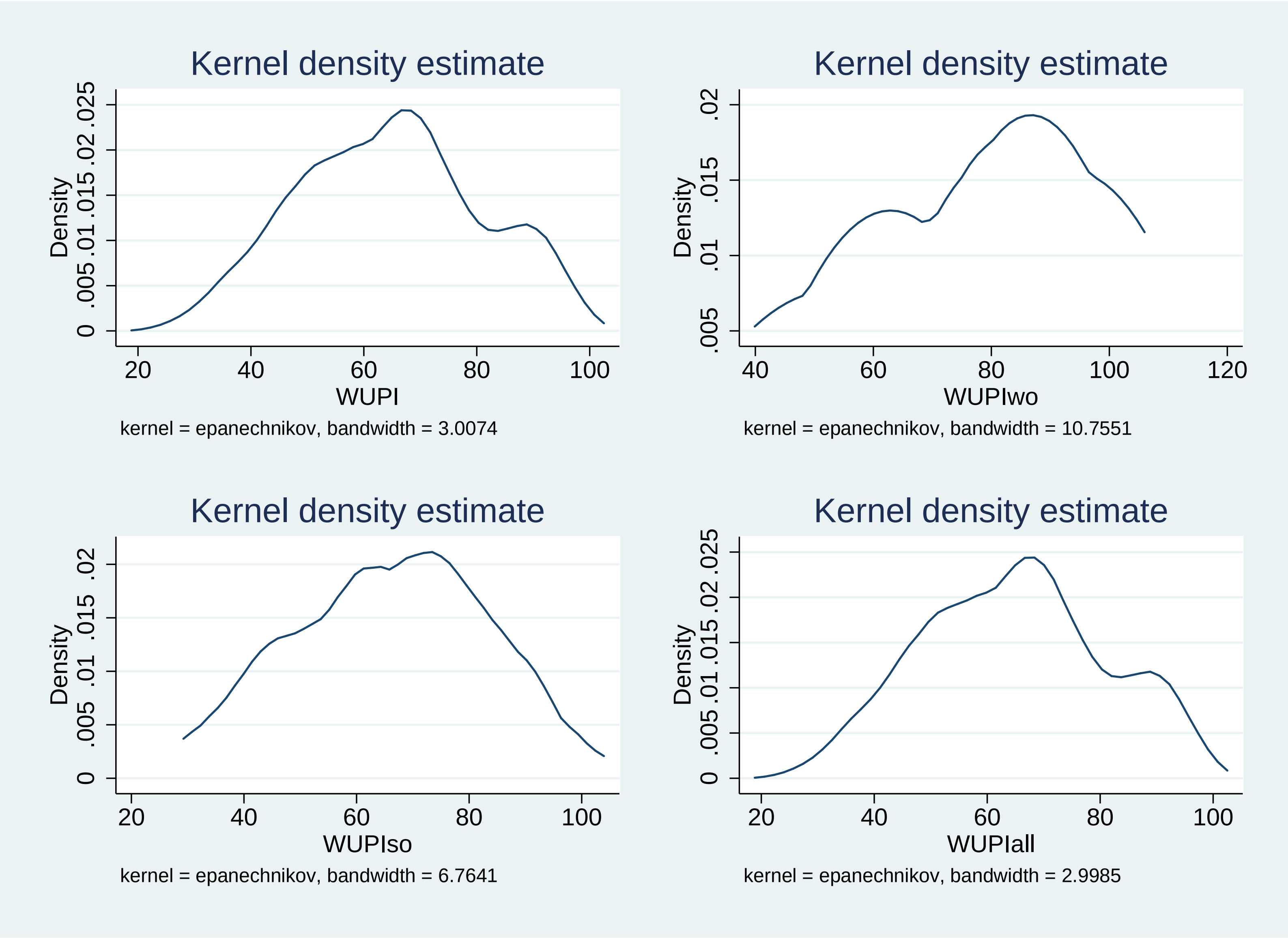

Figure 1: WUPI

Table 3: Descriptive statistics of WUPIall by country

country	**mean**	**p10**	**p25**	**p50**	**p75**	**p90**
Albania	48.3	33.2	39.5	47.6	56.1	62.4
Bosnia and Herzegovina	56.1	41.5	47.2	54.3	65.8	73.7
Bulgaria	66.8	56.7	62.1	68.2	72.2	76.3
Croatia	67.0	55.5	60.7	64.6	75.7	80.2
Czech Republic	90.7	84.3	88.1	90.6	94.3	97.4
Hungary	84.7	75.1	81.4	85.9	88.6	91.9
Kosovo	62.7	55.4	59.1	62.4	66.7	70.0
Macedonia, FYR	67.7	59.3	64.6	68.0	71.1	76.6
Moldova	52.5	37.6	44.2	52.3	62.1	67.8
Montenegro	58.9	57.4	57.7	58.0	60.1	62.3
Romania	73.4	65.0	68.0	72.8	78.5	82.6
Serbia	70.5	54.6	66.2	71.7	76.7	80.0
Slovak Republic	75.7	66.3	70.0	74.0	80.2	92.6
Ukraine	55.4	44.3	48.9	55.4	61.9	67.6
Total	64.6	43.4	52.7	64.9	75.6	88.2

Due to its construction, WUPI is a best practice indicator. For given cost/expenditures, higher values represent better performance. The indicator is therefore similar to the APGAR indicator by IBNET (see van den Berg and Danilenko (2011)). Both these indicators ignore the inputs necessary to achieve this performance. These relationships are highlighted in Table 4: High correlations between WUPI and APGAR, as well as very negative and low correlation with Cost per m3.

Table 4: Cross-correlation table

Variables	WUPIall	apgarscore	Cost_m3_prod
WUPIall	1.000		
apgarscore	0.769	1.000	
Cost_m3_prod	-0.14	-0.056	1.000

3.1 WUPI and missing data

One issue when constructing WUPI is how to deal with missing data. Missing data would typically make calculation of a WUPI score impossible if one or several sub-indicators are missing. The approach we take to calculate a WUPI indicator in the presence of missing data is the following. Firstly, we consider the coverage indicators as essential indicators. If one of these indicators is missing, we abstain from calculating a WUPI indicator. Secondly, we assume that for the best guess for a missing indicator is the average of all other indicators. Since all our indicators enter WUPI with the same weight, the WUPI score for each utility is calculated by:

$$\frac{\sum_{i=1}^{x} I_i}{x} \tag{1}$$

where I_i represents the value of the 1 to x sub-indicators i. In the case of no missing data, the sum of all 10 Indicators is simply divided by 10. Similarly, in the case of 9 available indicators, provided that none of the three coverage indicators missing, the sum of the indicators is divided by 9. We allow for up to three missing indicators.

Table 5: Correlation tables: WUPI and WUPI with 1 missing indicator

	WUPIall
score4	0.9902
score5	0.9771
score6	0.9834
score7	0.9678
score8	0.9769
score9	0.9952
score10	0.9896

The number indicates which sub-indicator is missing.

Table 6: Correlation tables: WUPI and WUPI with 2 missing indicator

	WUPIall
score45	0.9614
score46	0.9675
score47	0.9553
score48	0.9639
score49	0.9856
score410	0.9726
score56	0.9577
score57	0.9242
score58	0.9591
score59	0.9723
score510	0.9600
score67	0.9374
score68	0.9492
score69	0.9743
score610	0.9665
score78	0.9667
score79	0.9576
score710	0.9589
score89	0.9770
score810	0.9589
score910	0.9891

The number indicates which sub-indicators are missing.

Although we have already shown above that WUPI is highly correlated with APGAR, another established performance indicator for water utilities, the question remains how much bias this imputation of missing data generates. Under the assumption that missing data is not the result of strategically non-reporting by utilities, we measure how strongly WUPI scores based on the full set of 10 indicators are correlated with WUPI scores based on a limited subset.

As shown in Tables 5,6, and 7, the correlation between WUPI based on the full set and WUPI where one, two or three indicators are dropped is very high. In the case where we drop one or two indicators, all correlations are above 0.90. Even in the case where we drop 3 WUPI sub-indicators, only 1 out of 35 correlations with 0.88 is below the 0.90 threshold. These findings make us confident that calculating WUPI based on only a subset of the indicators does not introduce significant bias.

Table 7: Correlation tables: WUPI and WUPI with 3 missing indicator

	WUPIall
score654	0.9332
score754	0.9010
score764	0.9155
score765	0.8825
score854	0.9410
score864	0.9267
score865	0.9284
score874	0.9567
score875	0.9310
score876	0.9288
score954	0.9568
score964	0.9576
score965	0.9474
score974	0.9442
score975	0.9126
score976	0.9212
score984	0.9654
score985	0.9610
score986	0.9455
score987	0.9615
score1054	0.9334
score1064	0.9393
score1065	0.9309
score1074	0.9393
score1075	0.9049
score1076	0.9200
score1084	0.9360
score1085	0.9334
score1086	0.9201
score1087	0.9555
score1094	0.9733
score1095	0.9606
score1096	0.9620
score1097	0.9523
score1098	0.9649

The number indicates which sub-indicators are missing.

4 Utility performance and tariffs

As WUPI is a best practice indicator that only considers outputs and outcomes, we turn to the input side by correlating WUPI with water tariffs. While cost may also be interesting,[1] here we focus on tariffs because they are the most direct link to the consumers and their perspective on value for money.

We analyze not only whether tariffs (measured as revenues per capita and converted to international dollars using Purchasing Power Parity (PPP) conversion factors from the World Development Indicators) are related to WUPI, but also its most important subcomponents. For this reason, we split the performance indicator into 3 parts:

- Coverage: Indicator from 0 to 100 containing the average score[2]
    - Water coverage (0-100)
    - Sewerage coverage (0-100)
    - Treatment coverage (0-100)
- Quality: Indicator from 0 to 100 containing the average score[3]
    - Hours of service (0-100)
    - Blockages (0-100)
- Management-Efficiency: Indicator from 0 to 100 containing the average score[4]
    - Metering level (0-100)
    - Non-revenue water (0-100)
    - Staffing level (0-100)
    - Collection ratio (0-100)
    - Cost recovery (0-100)

Table 8: Cross-correlation table

Variables	coverageall	qualityall	mgmtall	WUPIall	tariff3
coverageall	1.000				
qualityall	0.282	1.000			
mgmtall	0.506	0.376	1.000		
WUPIall	0.822	0.574	0.857	1.000	
tariff3	0.305	0.317	0.213	0.328	1.000

---

[1] Using cost per capita instead of tariffs leads to very similar results and the same qualitative conclusions.

[2] Missing values for water coverage and sewerage coverage lead to utilities being excluded from the sample. Missing values for treatment coverage are assumed to be 0 if missing. 2 m3 of primary treatment are assumed to be comparable to 1 m3 secondary and higher treatment.

[3] Missing values for one indicator lead to the quality indicator being calculated based on the remaining score.

[4] Missing values for one indicator lead to the quality indicator being calculated based on the remaining scores.

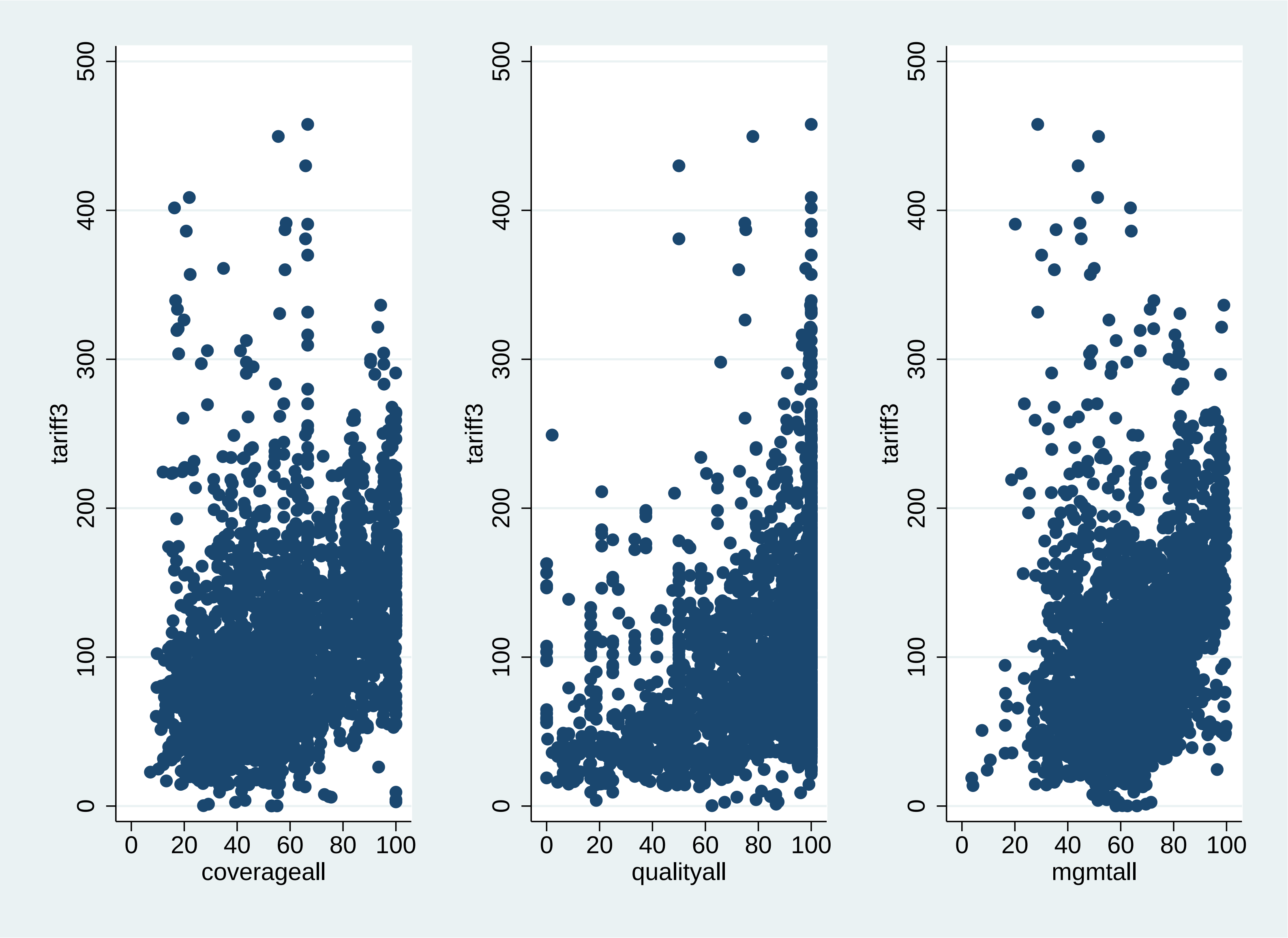

Figure 2: Scatter plot of tariffs and WUPI subcomponents

Table 8 shows the correlations between these 3 components, but also with respect to WUPI and tariffs themselves. Overall the 3 components are positively associated with each other, meaning that utilities with higher coverage are also more likely to have higher management efficiency.

With respect to tariffs, we find that utilities with higher tariffs typically have higher levels of WUPI, but also higher levels of each sub-item. The scatter plots in Figure 2 tells a similar story and suggests a positive relationship between tariffs and either component.[5]

In the next step, we try to measure the relationship between WUPI components and tariffs by using regression methods. We start with bivariate regressions of tariffs on WUPI and its subcomponents separately before regressing tariffs on all three subcomponents simultaneously. Most importantly, this allows us to test possible associations while keeping other utility performance components constant and controlling for country and year effects. The equation to estimate is given by:

$$ln(tariff_{ict}) = \beta_0 + \sum_{k=1}^{3} \beta_k * perf_{k,ict} + \nu_c + \eta_t + u_{ict} \tag{2}$$

where $ln(tariff_{ict})$ is the natural log of tariff in utility i in country c in year t. This quantity is regressed on the three WUPI subcomponents and country (ν_c) and year fixed effects (η_t). u_{ict} is the residual term, which is assumed to be i.i.d. unless mentioned otherwise. Given the wide dispersion and the presence of some obvious outliers - e.g. tariffs over 1000 PPP-dollars - a so-called robust regression method is used (see Li (1985)).[6]

Table 9 shows the results of running a regression of tariffs on the individual components. The coefficient on WUPIall in first column suggests that a 10 point increase in the WUPI score is associated with 6.7% higher tariffs. Repeating the same regressions with other components instead of overall WUPI, we find that the correlation is markedly different between the components. The effect is highest for coverage, with 5.6% for a 10 point increase, 2.2% for quality and -2.2% for the management component.

The result that particularly the management component is not necessarily related to higher tariffs - theoretically we would hope to see that higher productive efficiency should lead to lower costs and therefore lower tariffs - remains if we control for a full set of country-year effects, i.e. we only compare utilities in the same year and the same country with each other. The associated results are shown in Table 10.

This results are also confirmed if we add all three components simultaneously to the regression. The interpretation of these regression is now for a given level of two performance components, what is the correlation of tariffs with the remaining component. For instance, for a given level of coverage and quality performance, what is the correlation between tariffs and management performance. The corresponding results have basically the same interpretation as before: Coverage is most strongly and positively related to tariffs followed by quality. Management efficiency, in contrast, is negatively related to it.

To summarize, we find that having a higher performance in one area is also related to higher performance in other areas. While the obtained results here are only based on correlations, this suggests that increasing one performance area, e.g. through training of staff, may have positive spillovers on other areas.

---

[5] Outliers with tariffs above 500 dollars have been removed for presentation purposes.

[6] The results using ordinary least squares (OLS) are, however, very similar.

Table 9: Regression of tariffs on performance components, country and year fixed effects

	(1) tariff3	(2) tariff3	(3) tariff3	(4) tariff3	(5) tariff3
WUPIall	0.00665***				
	(0.000919)				
coverageall		0.00559***			0.00579***
		(0.000500)			(0.000509)
qualityall			0.00219***		0.00238***
			(0.000435)		(0.000421)
mgmtall				-0.00221***	-0.00337***
				(0.000742)	(0.000747)
_cons	2.763***	2.833***	2.944***	3.123***	2.890***
	(0.147)	(0.141)	(0.146)	(0.147)	(0.145)
N	2754	2754	2640	2754	2640

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

At the same time, the obtained results suggest important trade-offs of performance and tariffs as the costs for consumers. For a fixed degree of utility efficiency, particularly higher coverage is typically associated with higher tariffs. Hence higher coverage, but to a lesser degree also higher quality comes at the cost of higher tariffs. The results are somewhat different for management efficiency: here the relationship to tariffs tends to be negative. One could interpret this result as efficiency potentials within a utility that can be reaped by increasing management efficiency.

5 Utility performance over time and performance convergence

Overall, there is an improvement in WUPI over time. Figure 3 compares the first and last observation of each utility. On average WUPI scores improve by 3.2 points. The low average value is driven by the fact that not all utilities improve their scores over time. On the country level, most countries have a small growth between 0 and 10 points, and only Bulgaria, Macedonia, Slovak Republic, and Ukraine experience a decrease in WUPI scores. Moldova and Bosnia Herzegovina improve significantly by more 10 points (Figure 3).

Table 10: Regression of tariffs on performance components, country-year fixed effects

	(1) tariff3	(2) tariff3	(3) tariff3	(4) tariff3	(5) tariff3
WUPIall	0.0101***				
	(0.000958)				
coverageall		0.00720***			0.00737***
		(0.000502)			(0.000509)
qualityall			0.00213***		0.00234***
			(0.000431)		(0.000408)
mgmtall				-0.00114	-0.00250***
				(0.000781)	(0.000759)
_cons	3.052***	3.416***	4.116***	3.852***	4.030***
	(0.398)	(0.382)	(0.405)	(0.404)	(0.385)
N	2754	2754	2640	2754	2640

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

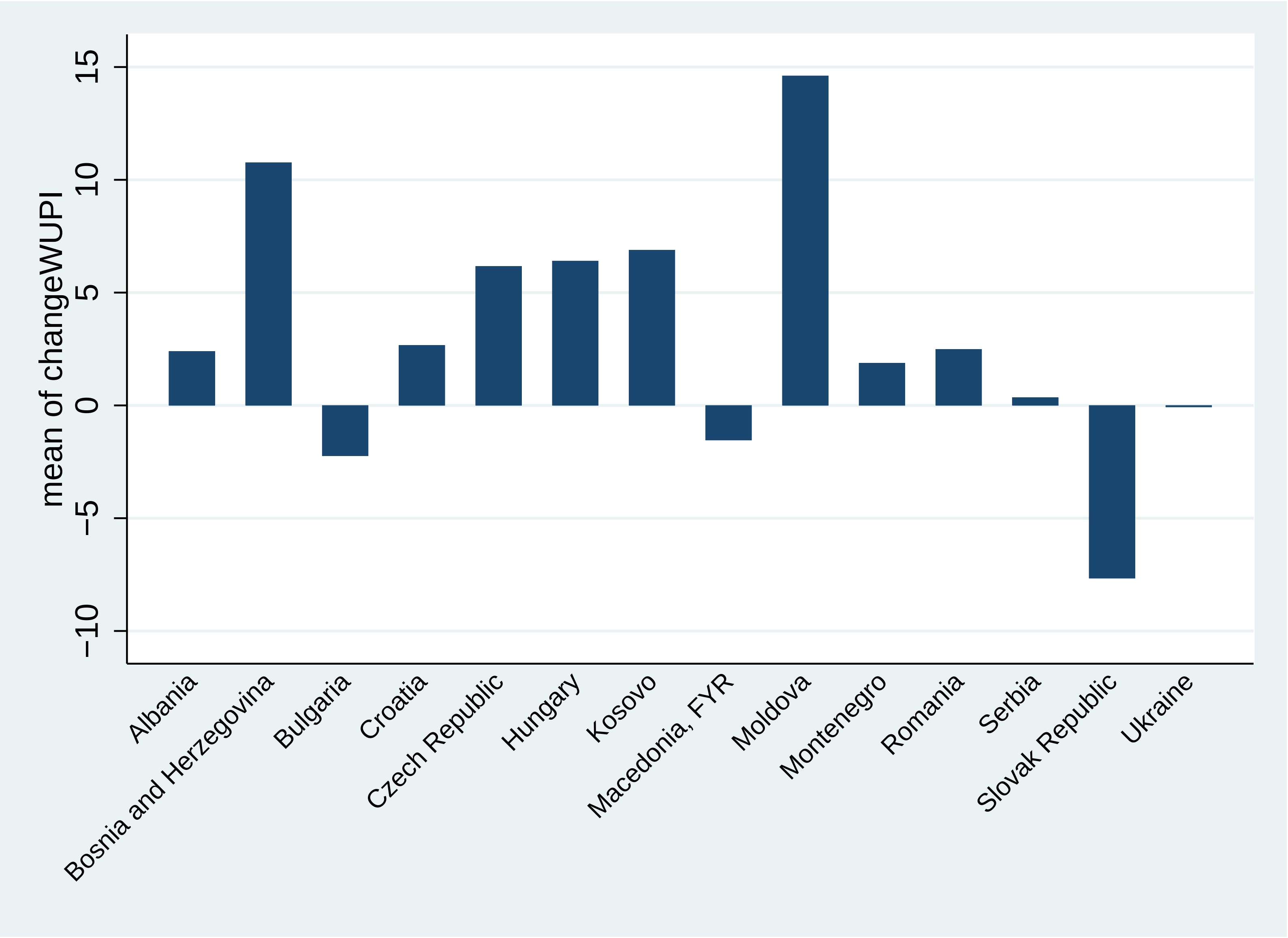

Figure 3: WUPI change, averaged over utilities

A regression with year dummies, while controlling for utility fixed effects, also confirms this positive trend. Purging unsystematic performance variation, Table 11 that there is a very strong and steady upward trend in WUPI scores. Particularly since around year 2000 (yearx6), WUPI scores have increase substantially.

Apart from the overall performance change, we also address the question whether utilities have converged in performance over time. For this reason, we follow the associated empirical literature on convergence (see Sala-i Martin (1996)) and regress the change in WUPI scores on the WUPI score in the first period. This amounts to:

$$ln(WUPI_{ic,t=T}) = \beta_0 + \beta_1 * ln(WUPI_{ic,t=1}) + \eta_t + u_{ict} \tag{3}$$

where $ln(WUPI_{ic,t=T})$ is natural log of the WUPI score of utility i in country c in the last observed period ($=T$) and $ln(WUPI_{ic,t=1})$ is the natural log of the score of the same utility in the first observed period ($t = 1$). Moreover, we add year fixed effects η_t to capture year specific shocks affecting all utilities We prefer collapsing the dataset into two cells per utility to avoid the results being driven by mean-reversion patterns of the data over time. Nevertheless, the results using the full panel are very similar in terms of the convergence pattern. As the errors are potentially correlated within utilities, we use cluster robust standard errors.

A negative relationship, β_1 negative, would indicate that utilities with higher initial performance grow slower, i.e. convergence in performance. The results in Table 12 confirm this type of convergence with a relatively strong catch-up of weaker utilities over time. Figures 4 and 5 present the evidence on convergence graphically.

Table 11: WUPI changes for each year

	(1) WUPIall
yearx2	4.335**
	(1.838)
yearx3	3.868**
	(1.755)
yearx4	1.709
	(1.748)
yearx5	0.582
	(1.749)
yearx6	2.750
	(1.717)
yearx7	2.918*
	(1.716)
yearx8	3.503**
	(1.728)
yearx9	6.042***
	(1.728)
yearx10	7.569***
	(1.724)
yearx11	10.08***
	(1.725)
yearx12	11.37***
	(1.723)
yearx13	12.50***
	(1.721)
yearx14	11.40***
	(1.752)
yearx15	12.16***
	(1.741)
yearx16	12.39***
	(1.737)
yearx17	12.35***
	(1.741)
yearx18	13.52***
	(1.741)
yearx19	15.14***
	(1.760)
_cons	56.10***
	(1.679)
N	2784

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

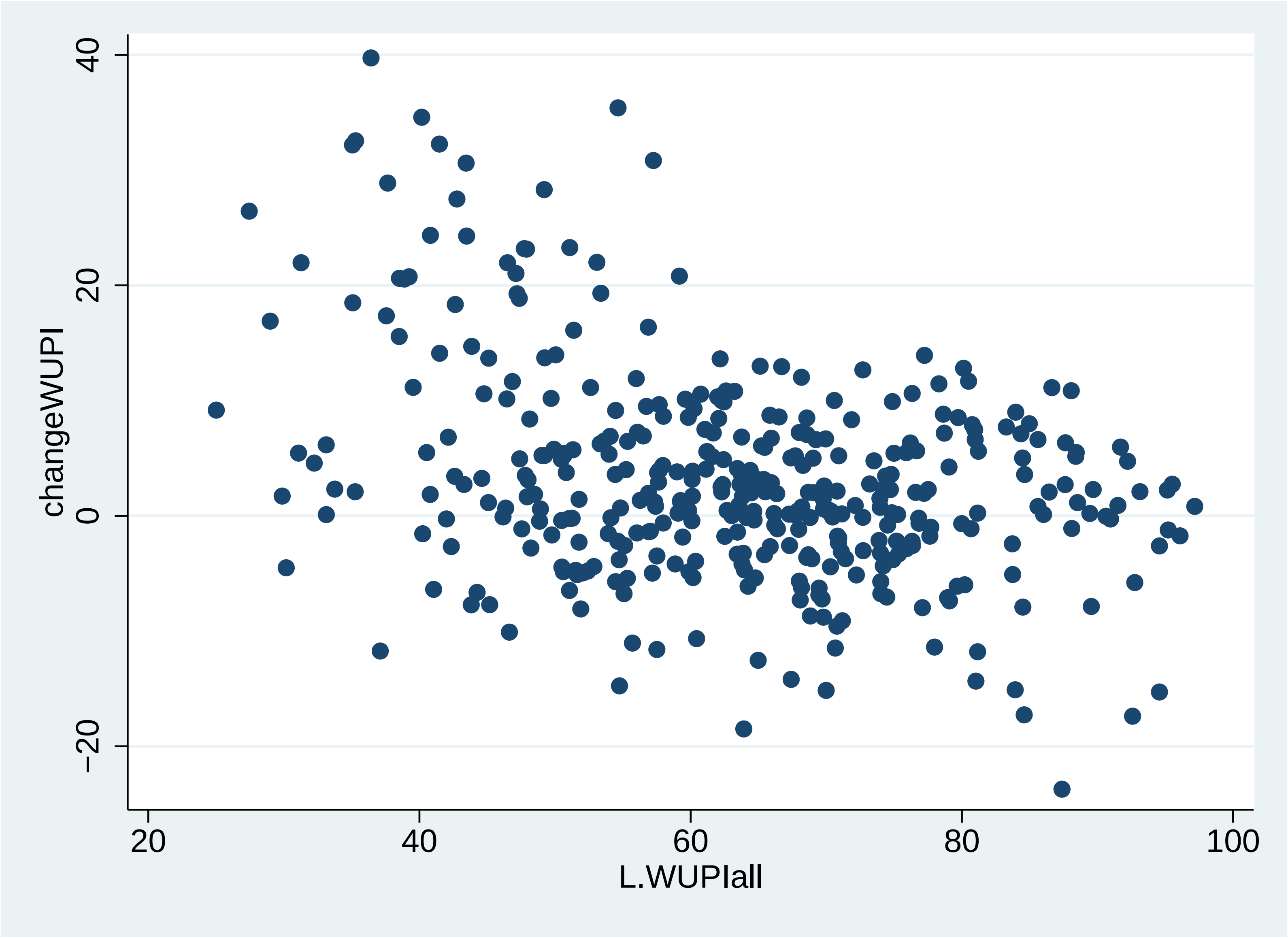

Figure 4: Scatter plot of WUPI change and initial WUPI

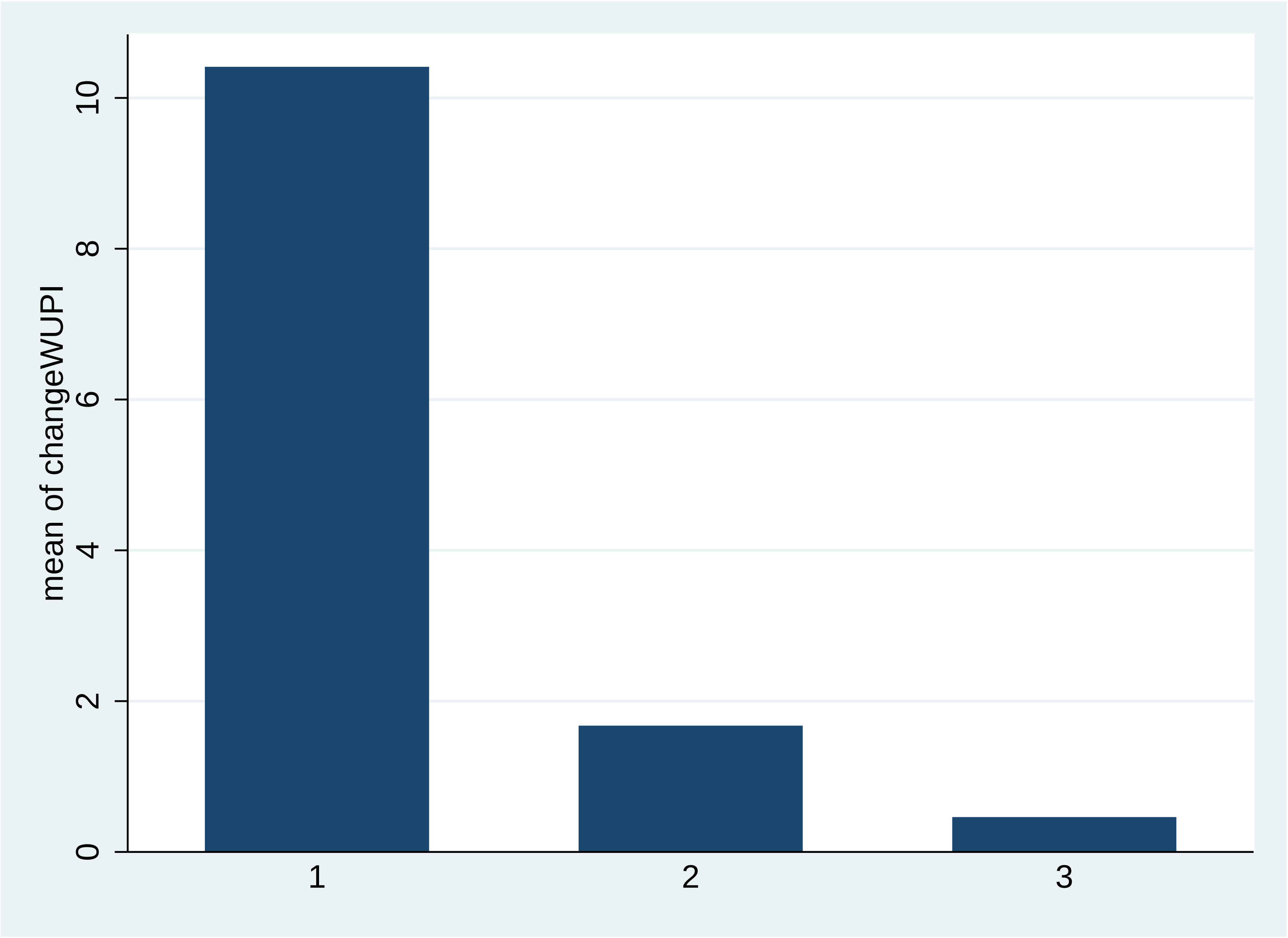

Figure 5: WUPI change for different initial levels of WUPI. Group 1 below 50; Group 2 above 50 below 75; Group 3 above 75

Table 12: Beta-Convergence

	(1) lnd_WUPI	(2) lnd_APGAR
lln_WUPIall	-0.353***	
	(0.0400)	
lln_APGAR		-0.448***
		(0.0458)
N	326	324

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Determinants of utility performance

The approach we take here is to correlate potentially relevant structural and governance variables with WUPI in order to address the question which factors drive utility performance. We first have to distinguish two sets of factors: Firstly, aggregate country level or institutional characteristics like the presence of a regulator or EU membership. Secondly, utility level factors such as the size of a utility or private participation. While the level of observation will remain the utility, the empirical specifications differ with respect to the variation that we use to identify the effects. Most importantly, the effect of utility level factors can be identified while controlling for more rich country and time effects, something which is not easily possible for country level variables.

The first step will therefore focus on the effect of a regulatory agency and EU membership. In the second part, we will look into utility characteristics.

6.1 Country level and institutional factors

First descriptive evidence on the relationship between EU membership and the presence of a regulatory agency can be found in Figures 7 6. Five countries had a regulatory agency in place at some point in time, two added a regulatory agency during the sample period. For EU membership, six countries joined the EU during the sample period, eight have never been members of the EU.

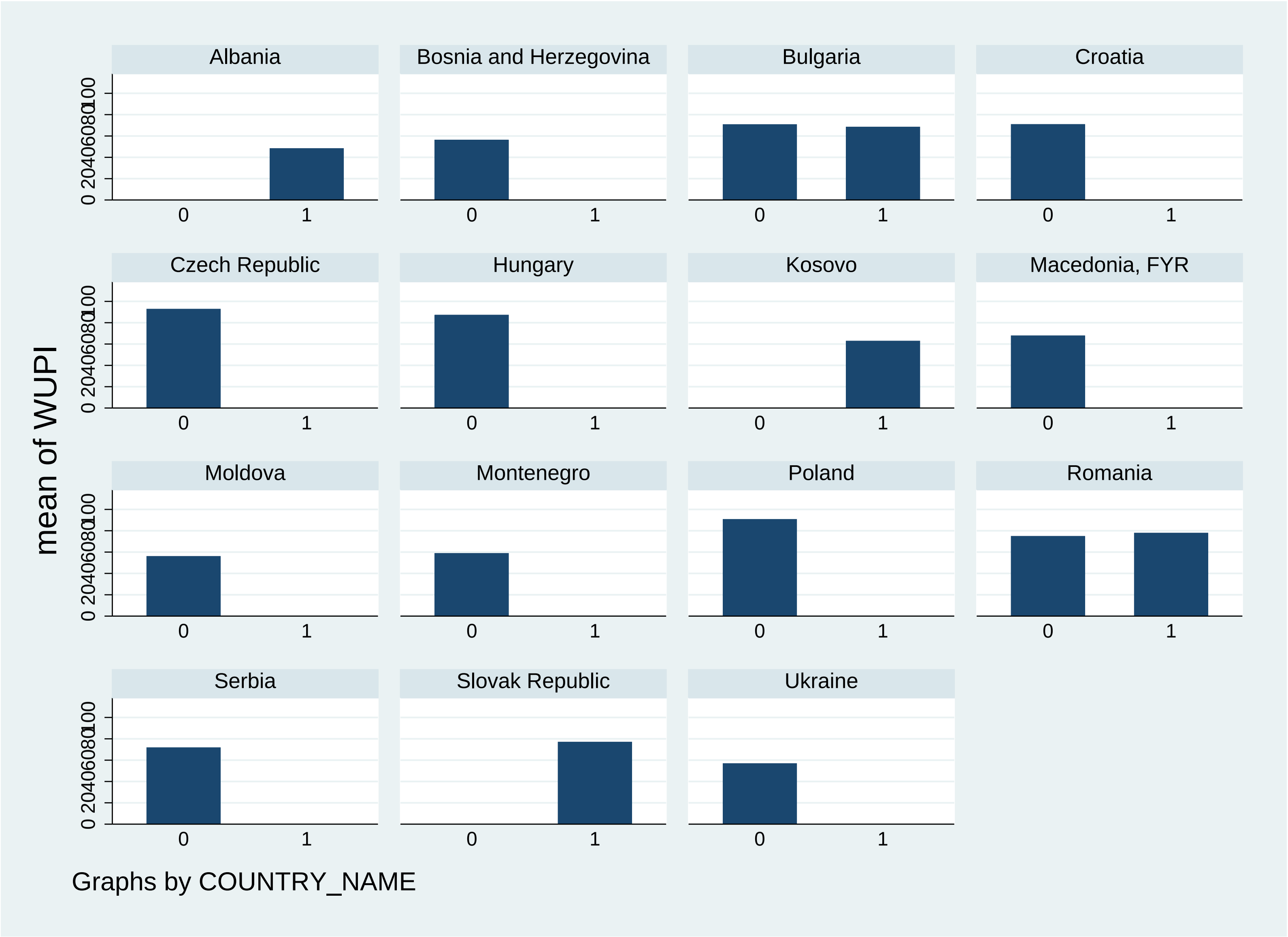

Figure 6: Average WUPI per country, for periods with and without regulator (1=regulator)

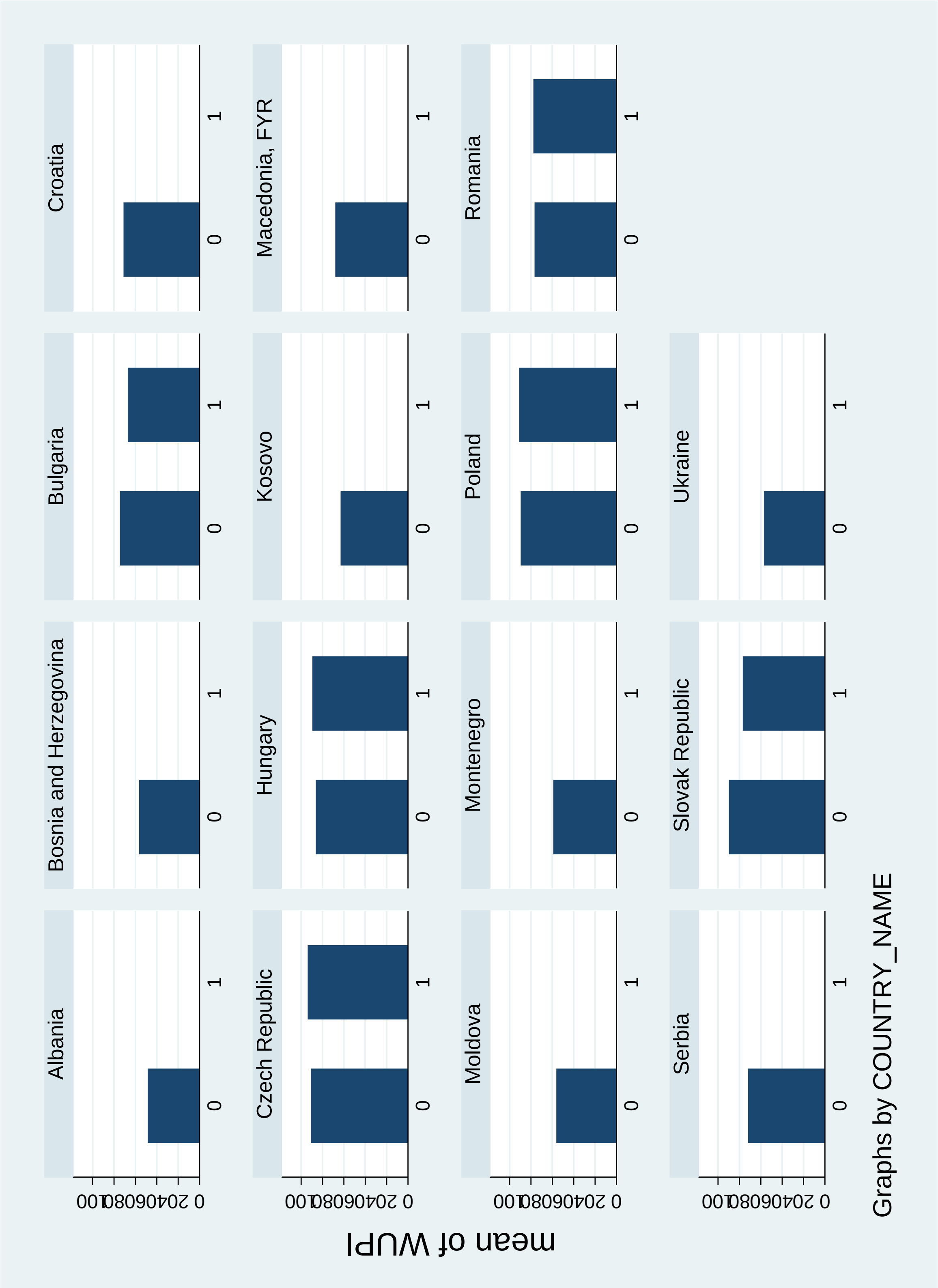

Figure 7: Average WUPI per country, for periods with and without EU membership (1=EU membership)

Table 13: Regression of WUPI on EU membership and regulatory agency

	(1) WUPIall	(2) WUPIall	(3) WUPIall	(4) WUPIall
reg	-9.147***	1.055	-5.114***	-1.878
	(1.405)	(1.348)	(1.413)	(1.562)
EU	22.11***	1.657***	-5.891***	-1.515
	(1.149)	(0.607)	(0.766)	(0.995)
_cons	62.53***	64.08***	55.88***	61.25***
	(0.881)	(0.330)	(1.666)	(1.754)
N	2784	2784	2784	2175
Country FE	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes
Moldova included	Yes	Yes	Yes	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To measure the association between these country level characteristics and utility performance, we correlated WUPI with these two variables and add utility and year fixed effects stepwise. The final specification being:

$$WUPI_{ict} = \beta_0 + \beta_1 * Reg_{ct} + \beta_2 * EU_{ct} + \gamma_i + \eta_t + u_{ict} \tag{4}$$

with $WUPI_{ict}$ representing the WUPI score of utility i in country c in year t. Reg_{ct} and EU_{ct} are dummy variables indicating the presence of a regulatory agency or EU membership. These variables are the same for all utilities in a given country. In all regressions, we cluster errors at the utility level and make them heteroscedasticity robust (Huber-White).

At face value, column 1 in Table 13 shows that utilities in regulated countries have a WUPI that is 12.6 points lower than those without and utilities in EU countries have 19.2 points higher values. To see if the introduction of a regulator or joining the EU had an effect of performance, column 2 compares WUPI scores before and after these events. The results show that the average WUPI before joining the EU was 1.7 points lower than afterwards. Also positive but without statistical significance is the effect of instating a regulatory agency, which increases performance by 1.1 points.

Instead of a simple before-after comparison, we would typically be interested in how utility performance would have changed in the absence of such reforms. This requires a control-group of utilities that are not affected by such reforms. We start by using all utilities in the Danube region. The results shown in column 3 are very different from the before-after comparison in column 2 and suggest that utilities in countries that joined the EU or implemented a regulatory agency improved their performance substantially slower than other utilities. On average, utilities outside the EU and without a regulator increased their performance by roughly 5 and 6 WUPI points more than those affected by such reforms.

A further examination of the results shows that the results are very sensitive to the inclusion or exclusion of Moldova, which is in the control group because it is not in the EU and has no regulator in place. Particularly after 2000 the country experienced a consistent and strong growth trend. Since one might worry about the comparability of utilities in Moldova with the utilities being affected by the reforms, e.g. because it seems to be violating the common trend assumption, column 4 shows the results with Moldova excluded. The negative effect remains, but becomes statistically insignificant. This shows that utilities in countries with a regulator or in the EU improved, but certainly not faster than the average utility in the Danube region.

The overall appraisal of the effect of these events depends a lot on the chosen counterfactual Overall utility performance increased after joining the EU or with a regulator. But the pace of improvement was very modest and similar or even slower than in other countries. While some of these countries and the strong positive trends they experienced may not be comparable to those joining the EU or putting a regulator in place (e.g. Moldova), the results show that there is no guarantee that performance will improve by these events. Already the Figures 7 6 indicate that the country specific experiences were very heterogeneous. EU membership or the presence of a regulatory agency alone cannot explain much of the performance differences.

6.2 Utility level factors

Moving to the utility level factors, a few additional methodological problems need to be addressed. Most importantly, there are a number of factors on the utility level available in IBNET, but some of them are very highly correlated and measure similar things. The considered factors are:

- size (number of connections to water)
- type of provider (municipal, regional, corporatized)
- type of price oversight (municipal, regional, national)
- scope (water and sewerage, multi-utilities)
- private involvement (none, management contract, lease, concession)

The fact that several variables are highly correlated is called multicollinearity because they reflect similar underlying factors. For instance, the size of a utility is strongly negatively correlated with the probability that a utility is a municipal provider (in contrast to regional or a corporatized utility) because they are typically small utilities (see Figure 8). Multicollinearity can not only inflate standard errors but also bias the estimates in the sense that we cannot identify the driving factor. To avoid this, we calculate correlations between the variables. Variables that are correlated highly, should not enter the regression simultaneously. The identified cases are:

- size and type of provider
- regulator and the type of price oversight

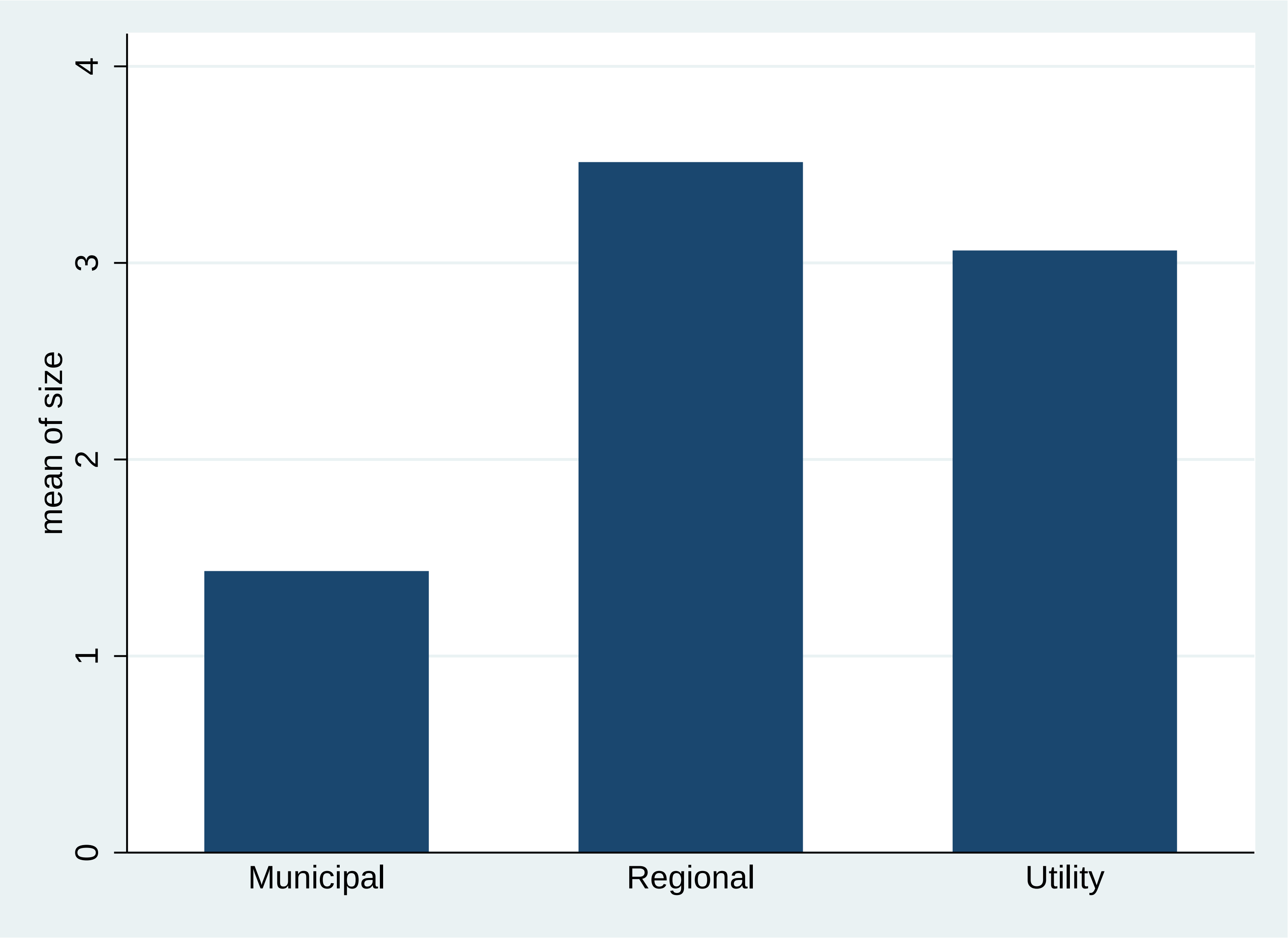

Figure 8: Average size (Number of customers) by type of water provider

As a result, and to deal with the fact that some variables are missing in a number of cases, we abstain from using the type of provider and the type of price oversight. It would also be questionable to add two variables measuring the same underlying features: what effect should be expected from changing a provider from a municipal utility to a regional provider, while keeping size constant?

As before, we start with a pooled regression before adding utility fixed effects. Given that on the utility level we have within country variation, we can even control for country-year effects to capture unobserved heterogeneity in the sense of country-year shocks. This latter specification leads to:

$$WUPI_{ict} = \beta_0 + \sum_{k=1}^{3} \beta_k * private_{k,ict} + \beta_4 * size_{ict} + \beta_5 * scope_{ict} + \gamma_i + \psi_{ct} + u_{ict} \tag{5}$$

with $WUPI_{ict}$ representing the WUPI score of utility i in country c in year t. $private_{k,ict}$ represents dummy variables indicating the type of private involvement, where no private involvement is the base category. $size_{ict}$ and $scope_{ict}$ represent utility size, measured by the number of connected customers, and utility scope, a dummy indicating whether the utility provides only water services or additional services. Here the base category is multiple services. The indicator variables ψ_{ct} represent the country-year fixed effects. In all regressions, we cluster errors at the utility level and make them heteroscedasticity robust (Huber-White).

The results of regression WUPI on the utility level factors are found in Table 14. As before in the initial specification in column 1 we do not control for country or year fixed effects. Coefficients in this specification there show the averages over the whole sample, across countries and years. On average utilities with private involvement, lease and concessions, have higher performance as measured by WUPI. Similarly, the first specification suggests that larger utilities and utilities with broader scope (multi utilities) exhibit higher performance.

However, as soon as utility fixed effects are taken into account (column 2), i.e. using only changes within utilities to identify the effects, only size remains statistically significant at the 5% level. The same is true for column 3, in which country-year fixed effects enter instead of the mere country fixed effects. Interestingly, private participation changes its sign, and leases and concessions seem to be related with lower performance, but not statistically significantly. To ensure that the results are not affected by utilities never changing from public to private, column 4 re-estimates the same model with only switching utilities. The results on private contracting remain the same, again with concession contracts being weakly negatively related with WUPI.

The bottom line is that when looking at performance evolution over time, columns 2 and 3, utility performance is not affected by switching from public to private or conversely. This result is in line with a strand of more recent literature concluding that public private differences are negligible for performance in network industries like water provision.

6.3 A digression on size

What remains strongly significant and positive, except in the last specification, is size. To zoom in on this discussion, it is first necessary to distinguish several aspects of size. Although the theoretical contributions relate to costs and cost functions (see Caves et al. (1980), Caves

Table 14: Utility level regressions

	(1) WUPIall	(2) WUPIall	(3) WUPIall	(4) WUPIall
private_manage	0.213	1.583	-0.764	-2.313
	(3.017)	(3.588)	(2.366)	(2.420)
private_lease	17.64***	-2.603*	-1.555	-0.994
	(2.043)	(1.494)	(1.643)	(2.252)
private_conces	11.34***	-1.681	-0.285	-2.033
	(2.207)	(2.142)	(1.708)	(1.547)
size	5.531***	10.64***	3.167***	2.561
	(0.447)	(1.437)	(1.124)	(4.310)
scope_double	-8.655***	-3.713*	0.782	-0.847
	(1.578)	(1.919)	(2.159)	(2.540)
_cons	60.02***	46.11***	58.29***	78.30***
	(1.774)	(3.654)	(3.027)	(13.93)
N	2449	2449	2449	263

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

et al. (1984), Garcia and Thomas (2001) for classical contributions or Nauges and van den Berg (2008) for a study using IBNET data), a similar logic can be applied regarding the effect of size on other quantities such as WUPI:

- Output density (Y): This captures the effect of size in the sense of volume of water.
- Consumer density (CD): This captures the effect of increasing the number of customers.
- Service area (SA): This captures the effect of increasing the number of serviced towns.

In reality, aggregations will affect possibly all of these quantities. Using a subsample of utilities in IBNET, where we believe data quality is highest, we find that the average aggregation increases volume by 5%, density by 5% and towns by 66% (7 cities on average).

To evaluate the effect of aggregations, we first focus on the average effect of aggregations. Through the data, we identify an aggregation if the number of serviced cities increases. The dummy variable 'after' is 1 after an aggregation and 0 otherwise. In the first step, we simply compare average performance before and after an aggregation. In the next step, we will compare it to utilities without an aggregation, similar to a difference-in-difference approach.[7] Some experimentation with the data has shown that the choice of the control group - e.g. the utilities without aggregation that is used as a comparison - is important for the obtained results. Since we are interested in the counter-factual scenario, how would the performance of a utility be in the absence of an aggregation, not all utilities are suitable for comparison. For this reason, we performed a propensity score matching, where we use pre-treatment characteristics to estimate the probability that a utility would experience an

---

[7] See Angrist and Pischke (2008) or Wooldridge (2010) for introductory texts on treatment effect evaluation.

aggregation (see Rosenbaum and Rubin (1985)). These variables $x_{k,ict}$ include important utility characteristics like the population in the service area, the number of towns already served, but also the volume of distributed water and treated sewerage. We therefore assume the following relationship in the decision to aggregate:

$$P(Aggregation_{ict} = 1|X) = F(\beta_0 + \sum_{k=1}^{k} \beta_k * x_{k,ict}) \tag{6}$$

where $P(Aggregation_{ict}$ is the conditional probability that municipality i in country c in year t will aggregate. While the variables in X are assumed to have a linear additive impact on the latent variable $Aggregation^{*}$[8], the response probability is actually a nonlinear function of the covariates.

While other approaches might be feasible, we follow a suggestion by Gelman and Hill (2006) to combine propensity score matching and regression by using the propensity score to identify utilities which are not comparable to the aggregated utilities. This means that utilities outside common support, in our case utilities whose probability to experience an aggregation is too small, are excluded from the analysis. Moreover, as before the results proved to be sensitive to including Moldova in the control group. As Moldova experiences a very strong positive trend in WUPI scores its inclusion leads to a higher performance of the control group and therefore the effect of aggregations turns insignificant (but still positive). Given that Moldova has had no aggregations and served thus enters the control group, the models and results without Moldova are preferred even if we present results with and without utilities from Moldova.

Using these subsample of utilities we first estimate a simple before-after model comparing the WUPI score before and after the aggregation while controlling for aggregate year shocks. In the next step, we compare the performance change in aggregating utilities to other non-treated utilities. This leads us to estimate the following generalized difference-in-difference specification with multiple time periods and arbitrary treatment patterns:

$$WUPI_{ict} = \beta_0 + \beta 1 * Aggregation_{ict} + \gamma_i + \eta_t + u_{ict} \tag{7}$$

To allow for the possibility that the effect of the aggregation is not immediate but distributed over time after the aggregation, we rerun the above model and replace the indicator variable $Aggregation_{ict}$ with dummy variables indicating the first year of the aggregation (1.*treattime*), the second year (2.*treattime*), the third and fourth (3.*treattime*), and the fifth and more years of the aggregation (4.*treattime*).

$$WUPI_{ict} = \beta_0 + \sum_{k=1}^{4} \beta_k * k.treattime_{ict} + \gamma_i + \eta_t + u_{ict} \tag{8}$$

As before, we cluster standard errors at the utility level and robustify for heteroscedasticity.

The results are shown in Table 15 and indicate that if we consider only utilities that underwent an aggregation, i.e. a before after comparison, performance was 1.8 points higher after the aggregations than before. However, in comparison to utilities without aggregations,

---

[8] The equivalent latent variable model is $Aggregation^{*}_{ict} = \beta_0 + \sum_{k=1}^{k} \beta_k * x_{k,ict} + u_{ict}$ with $Aggregation_{ict} = 1[Aggregation^{*}_{ict} > 0]$

Table 15: Treatment effect of aggregations on WUPI

	(1) WUPIall	(2) WUPIall	(3) WUPIall	(4) WUPIall	(5) WUPIall
after	1.950**	0.691		1.739*	
	(0.783)	(0.967)		(0.904)	
1.treattime			0.943		1.344*
			(0.775)		(0.783)
2.treattime			0.637		1.175
			(0.980)		(1.030)
3.treattime			0.654		1.392
			(1.103)		(1.096)
4.treattime			-1.611		-0.394
			(1.598)		(1.616)
_cons	70.50***	72.58***	58.59***	68.17***	59.51***
	(2.577)	(2.695)	(1.033)	(1.444)	(1.105)
N	220	629	741	590	701
Moldova	Yes	Yes	Yes	No	No

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the overall effects seemed slightly less beneficial (column 2 and column 4 without Moldova). Whether Moldova is included or not, the average effect of aggregations is estimated to be very small and not statistically significant, ranging from 0.4 to 1.4 WUPI points gain. Aggregations therefore had very little impact on the performance as measured by WUPI. To illustrate the over time effects of an aggregation, column 3 shows the change in performance in year 1 (treattime 1), in year 2 (treattime 2), in year 3 and 4 (treattime 3), and in 5 and more years (treattime 4) after the aggregation. As the table shows, if any positive effect, most changes happen in the first 2 years. Over time the effect dissipates.

This finding corresponds to the results for the WUPI subcomponents shown in tables 16, 17, and 18. The benefits from aggregation mainly stem from increases in management efficiency, which are statistically significant and large if Moldova is excluded. Conversely, the effect of aggregations is negative, particularly in the long term for the coverage indicators.

After this focus on the average effect of aggregations, the last step of this section is to separate the effect of aggregations in the changes in volume, customer density and the number of managed systems. This approach is closer to the typical approaches to measure economies of scale in the water industry (see Garcia and Thomas (2001) or Nauges and van den Berg (2008)). However, as we are not actually estimating cost or production functions, the standard approaches can not be applied directly. The reason is that theoretical results on cost minimization such as Shepard's lemma do not carry over to our case and we therefore cannot estimate a system including the cost share equations with the associated cross-equation restrictions. The consequence of this should, however, be limited to less precise estimates.

Table 16: Treatment effect of aggregations on coverage

	(1) coverageall	(2) coverageall	(3) coverageall	(4) coverageall	(5) coverageall
after	-0.309	-3.018		-0.639	
	(1.672)	(2.131)		(1.909)	
1.treattime			-1.431		-0.464
			(1.692)		(1.626)
2.treattime			-3.487		-2.380
			(2.111)		(2.131)
3.treattime			-5.833**		-4.194*
			(2.402)		(2.359)
4.treattime			-10.27***		-7.514**
			(3.403)		(3.506)
_cons	64.31***	62.35***	49.86***	54.99***	51.49***
	(4.657)	(2.997)	(3.318)	(2.465)	(3.639)
N	220	629	741	590	701
Moldova	Yes	Yes	Yes	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Treatment effect of aggregations on quality

	(1) qualityall	(2) qualityall	(3) qualityall	(4) qualityall	(5) qualityall
after	-0.961	1.758		1.282	
	(2.771)	(2.923)		(3.061)	
1.treattime			1.553		1.242
			(2.502)		(2.582)
2.treattime			2.469		2.111
			(2.884)		(3.001)
3.treattime			3.271		2.540
			(3.344)		(3.498)
4.treattime			6.289		4.982
			(4.001)		(4.379)
_cons	85.97***	87.51***	86.73***	89.77***	86.24***
	(0.965)	(3.091)	(2.132)	(3.447)	(2.018)
N	220	625	737	586	697
Moldova	Yes	Yes	Yes	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Treatment effect of aggregations on management efficiency

	(1) mgmtall	(2) mgmtall	(3) mgmtall	(4) mgmtall	(5) mgmtall
after	3.328^{**}	1.717		2.485^{*}	
	(1.327)	(1.294)		(1.322)	
1.treattime			1.776		2.141^{*}
			(1.113)		(1.148)
2.treattime			1.656		2.293^{*}
			(1.311)		(1.365)
3.treattime			3.119^{**}		3.958^{***}
			(1.347)		(1.366)
4.treattime			0.846		2.261
			(1.760)		(1.775)
_cons	67.87^{***}	74.23^{***}	55.48^{***}	69.66^{***}	56.71^{***}
	(4.273)	(3.026)	(1.141)	(2.044)	(1.007)
N	220	629	741	590	701
Moldova	Yes	Yes	Yes	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Overall, as shown in Figure 9 all three size-related variables have a positive association with WUPI, in the case of volume the relationship appears to be nonlinear.

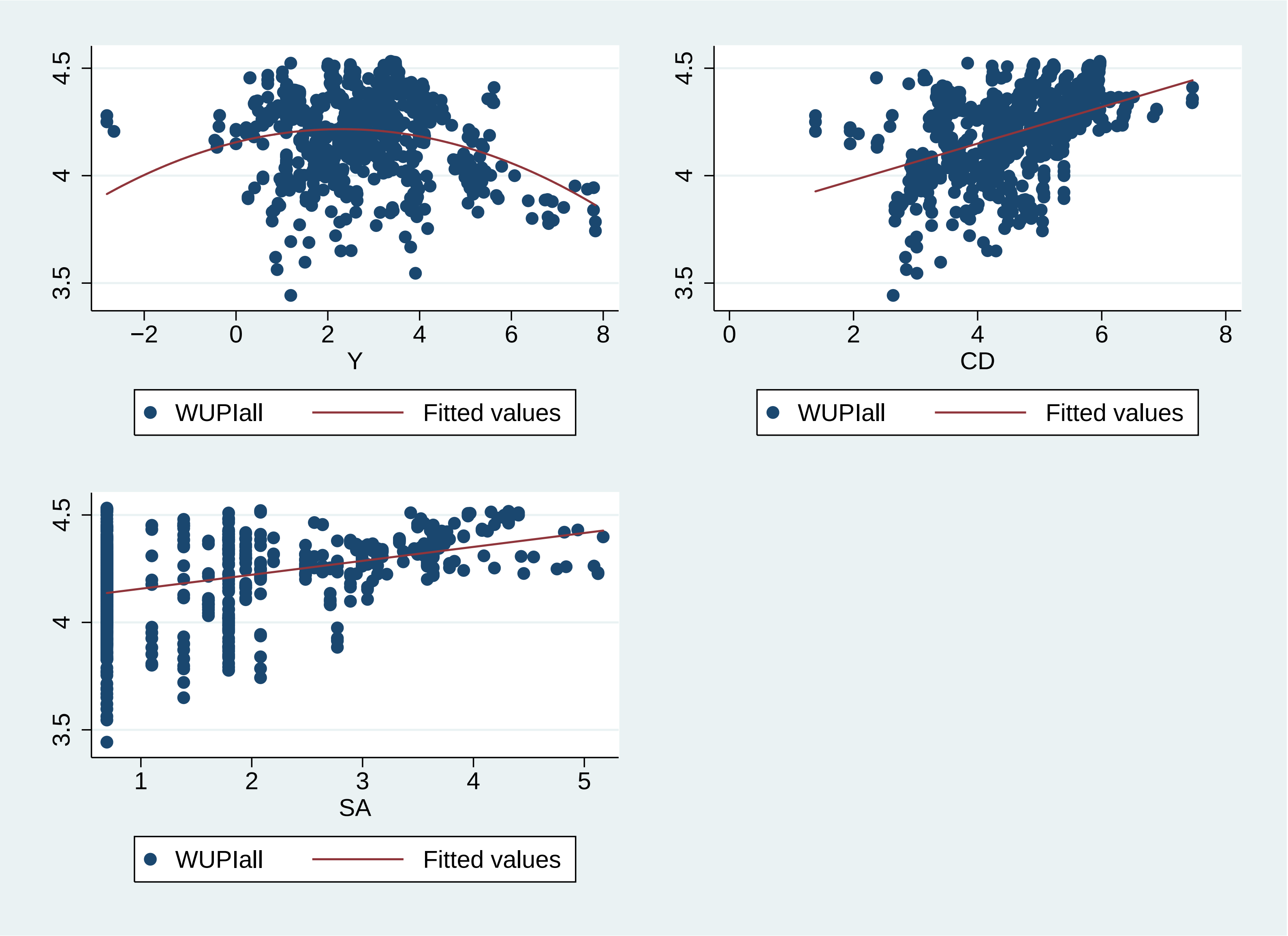

Figure 9: Scatter plots of volume (Y), customers (CD), and number of systems (SA) with WUPI

To evaluate the effect of aggregations in the sense of changes in these three aspects of size on WUPI, we are particularly interested in changes over time. While the cross-sectional relationship of size and WUPI may be positive, it is not evident that a utility that adds several additional towns to its service area will improve its performance. Using the same sample as in the previous regressions, we run a model regressing WUPI on these 3 factors as well as utility and year fixed effects:

$$ln(WUPI_{ict}) = \beta_0 + \beta_1 * ln(Y)_{ict} + \beta_2 * ln(CD)_{ict} + \beta_3 * ln(SA)_{ict} + \gamma_i + \eta_t + u_{ict} \tag{9}$$

where Y is the volume of water and wastewater by utility i in country c and year t. Similarly, CD captures the number of customers of a utility and SA the number of served towns. To allow for a more flexible data generating process, we also estimate a translog function of the form (see Christensen et al. (1973)):

$$\begin{aligned} ln(WUPI_{ict}) = \beta_0 + \beta_1 * ln(Y_{ict}) + \beta_2 * ln(CD_{ict}) + \beta_3 * ln(SA_{ict}) + \\ \beta_4 * ln(Y_{ict})ln(Y_{ict}) + \beta_5 * ln(CD_{ict})ln(CD_{ict}) + \beta_6 * ln(SA_{ict})ln(SA_{ict}) + \\ \beta_7 * ln(Y_{ict})ln(CD_{ict}) + \beta_8 * ln(Y_{ict})ln(SA_{ict}) + \beta_9 * ln(CD_{ict})ln(SA_{ict}) + \\ \gamma_i + \eta_t + u_{ict} \end{aligned} \tag{10}$$

This gives us the results presented in Table 19. As the variables are in natural logarithms, a percentage interpretation arises. For instance in the specification in column 1, we find that a 1% increase in customer density increases WUPI by 0.26%. In contrast, the negative sign for SA suggests that a 1% increase in the number of served towns decreases WUPI by 0.05%. A change in the amount of water produced does not seem affect WUPI. Given that the in most aggregations the service area increases much more than density and or volume, the composite effect may be positive in some cases and negative in others. The larger the number of additional systems relative to the number of additional customers, the more negative the effects of aggregations. Conversely, Aggregations with that add substantially more customers and only a limited number of additional towns are predicted to increase performance more strongly.

A more rich translog specification that allows for non-linear effects, column 2 of Table 19, shows similar results for the average utility. Consumer density is positively correlated with WUPI, the number of systems negatively, but not statistically significant. Given the high collinearity between the measures of size and their square and interaction terms, this is not surprising.

If we reestimate the translog specifications for different sizes of water utilities 20 and 21, we find that particularly large utilities may no longer benefit from aggregations because the positive effect of more consumers on WUPI disappears with increasing initial size. This is consistent for the regressions with Moldova and without. For the number of systems, the coefficient is negative in all specification but the pattern is less clear. The results do not show that the negative effect of adding systems vary with the size of the initial utility. Nevertheless, the effect is always negative and partially statistically significant.

Table 19: Size component regressions 1

	(1) WUPIall	(2) WUPIall	(3) WUPIall	(4) WUPIall
Y	-0.0271	-0.0304	-0.0499***	-0.0584***
	(0.0203)	(0.0219)	(0.0181)	(0.0181)
CD	0.256***	0.260***	0.247***	0.255***
	(0.0728)	(0.0734)	(0.0625)	(0.0629)
SA	-0.0505***	-0.0373***	-0.0269	-0.0123
	(0.0160)	(0.0139)	(0.0187)	(0.0192)
c.Y#c.Y			0.0185**	0.0221***
			(0.00748)	(0.00713)
c.SA#c.SA			-0.00863*	-0.00650
			(0.00506)	(0.00465)
c.CD#c.CD			-0.0526	-0.0354
			(0.0425)	(0.0403)
c.Y#c.CD			-0.0375**	-0.0438***
			(0.0181)	(0.0161)
c.Y#c.SA			0.0120	0.0130
			(0.0128)	(0.0122)
c.CD#c.SA			0.00329	-0.00539
			(0.0286)	(0.0271)
_cons	-0.0827***	-0.0682***	-0.0275	-0.0230
	(0.0161)	(0.0167)	(0.0341)	(0.0324)
N	738	698	738	698
Moldova	Yes	No	Yes	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Size component regressions for different utility sizes

	(1) WUPIall p10	(2) WUPIall p25	(3) WUPIall p50	(4) WUPIall mean	(5) WUPIall p75	(6) WUPIall p90
Y	-0.0705*	-0.0706***	-0.0634***	-0.0499***	-0.0414	-0.0180
	(0.0373)	(0.0274)	(0.0212)	(0.0181)	(0.0254)	(0.0412)
CD	0.441***	0.350***	0.247***	0.247***	0.145*	0.0712
	(0.105)	(0.0745)	(0.0750)	(0.0625)	(0.0783)	(0.0796)
SA	-0.0356	-0.0258	-0.0131	-0.0269	-0.0244**	-0.0390*
	(0.0474)	(0.0325)	(0.0179)	(0.0187)	(0.0110)	(0.0222)
N	738	738	738	738	738	738

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Size component regressions for different utility sizes without Moldova

	(1) WUPIall p10	(2) WUPIall p25	(3) WUPIall p50	(4) WUPIall mean	(5) WUPIall p75	(6) WUPIall p90
Y	-0.0784**	-0.0816***	-0.0758***	-0.0584***	-0.0513**	-0.0244
	(0.0309)	(0.0239)	(0.0212)	(0.0181)	(0.0254)	(0.0407)
CD	0.428***	0.345***	0.258***	0.255***	0.167**	0.0955
	(0.108)	(0.0745)	(0.0772)	(0.0629)	(0.0784)	(0.0772)
SA	-0.0136	-0.00883	-0.00176	-0.0123	-0.0129	-0.0256
	(0.0489)	(0.0319)	(0.0172)	(0.0192)	(0.00999)	(0.0200)
N	698	698	698	698	698	698

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.4 Aggregations and cost and tariffs

In addition to analyzing the effect of aggregations on WUPI, here we also consider how aggregations affect tariffs and costs. This is particularly important since WUPI is a purely outcome based indicator that holds no relation to the required inputs to achieve a certain WUPI level. In a previous section, we have already shown that a higher WUPI may require higher tariffs (higher costs). As the average effect of aggregation on WUPI was rather small, one might suspect that aggregations unfold more strongly on the input side, for a given level of WUPI. To address this concern, we repeat the above analysis for tariffs and costs.

As shown in Tables 22 and 23, overall there is little evidence that aggregations have an effect on tariffs or cost. Standard errors are very high, particularly regarding tariffs, which suggests that aggregations can go either way. A simple before-after comparison would suggest that aggregations could reduce cost and tariffs. Both coefficients are negative, but particularly in the case of tariffs the standard errors are very high, making the coefficient statistically insignificant. If we also consider other utilities that did not aggregate as comparisons, the effect of aggregations on tariffs switches to a positive sign, implying that aggregations lead to increasing tariffs. However, as before the results are clearly insignificant. In the case of costs, the sign remains negative, implying a negative effect of aggregations on tariffs but as for tariffs the coefficient is not statistically different from zero. As for WUPI, aggregations seem to vary gravely in their consequences for cost and tariffs such that on average no clear positive or negative effect can be established.

When looking at the different channels of aggregations - volume, consumer density, number of systems - we observe a pattern that is similar to what we found for WUPI: increasing consumer density through aggregations has a beneficial (negative) effect on tariffs and costs (see Tables 24 25). In contrast, for a given consumer density increasing the number of systems of produced volume increases costs. Although the negative effect of these variables is insignificant in many cases, the overall pattern mirrors the previous findings that aggregations which increase the number of customers are more likely beneficial for utility performance.

Finally, we again confirm that the benefits from aggregation are size dependent. Particularly small utilities may benefit from aggregations that lead to a significant increase

Table 22: Treatment effect of aggregations on tariffs

	(1) tariffs	(2) tariffs	(3) tariffs	(4) tariffs	(5) tariffs
after	-4.351	4.427		4.161	
	(7.182)	(7.038)		(7.708)	
1.treattime			1.883		2.237
			(4.894)		(5.236)
2.treattime			5.203		5.683
			(7.454)		(7.863)
3.treattime			6.467		7.218
			(9.337)		(10.48)
4.treattime			22.74**		24.06**
			(9.926)		(11.90)
_cons	57.92***	78.94***	77.47***	142.0***	80.70***
	(14.63)	(11.21)	(7.185)	(11.36)	(8.214)
N	225	651	768	609	725
Moldova	Yes	Yes	Yes	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Treatment effect of aggregations on cost

	(1) cost	(2) cost	(3) cost	(4) cost	(5) cost
after	-8.467*	-2.541		-4.554	
	(4.641)	(5.471)		(5.795)	
1.treattime			-1.558		-2.141
			(3.994)		(4.287)
2.treattime			0.353		-0.717
			(5.874)		(6.272)
3.treattime			-3.049		-4.359
			(6.266)		(6.894)
4.treattime			6.305		4.066
			(9.429)		(10.64)
_cons	58.05***	71.24***	73.69***	133.6***	74.53***
	(11.03)	(8.092)	(5.959)	(8.211)	(6.491)
N	222	645	760	603	717
Moldova	Yes	Yes	Yes	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Size component regressions 1 tariffs

	(1) tariffs	(2) tariffs	(3) tariffs	(4) tariffs
Y	0.0846**	0.0599*	0.109**	0.0789
	(0.0388)	(0.0360)	(0.0529)	(0.0551)
CD	-0.503***	-0.537***	-0.490***	-0.509***
	(0.162)	(0.176)	(0.157)	(0.166)
SA	0.0540	0.0659	0.0283	0.0303
	(0.0409)	(0.0450)	(0.0864)	(0.0874)
c.Y#c.Y			-0.0484***	-0.0388**
			(0.0171)	(0.0194)
c.SA#c.SA			-0.0405*	-0.0388*
			(0.0219)	(0.0217)
c.CD#c.CD			-0.0303	-0.0103
			(0.110)	(0.109)
c.Y#c.CD			0.119*	0.120*
			(0.0701)	(0.0690)
c.Y#c.SA			-0.121**	-0.117*
			(0.0588)	(0.0601)
c.CD#c.SA			0.151	0.138
			(0.0960)	(0.0947)
_cons	-0.499***	-0.492***	-0.483***	-0.524***
	(0.0435)	(0.0468)	(0.0921)	(0.0999)
N	765	722	765	722
Moldova	Yes	No	Yes	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 25: Size component regressions 1 cost

	(1) cost	(2) cost	(3) cost	(4) cost
Y	0.0469	0.0323	0.0953**	0.0775*
	(0.0303)	(0.0279)	(0.0390)	(0.0403)
CD	-0.328**	-0.319**	-0.350**	-0.342**
	(0.131)	(0.141)	(0.145)	(0.152)
SA	0.0516*	0.0522	0.0674	0.0706
	(0.0305)	(0.0344)	(0.0599)	(0.0628)
c.Y#c.Y			-0.0306**	-0.0233
			(0.0155)	(0.0163)
c.SA#c.SA			-0.00686	-0.00572
			(0.0105)	(0.0108)
c.CD#c.CD			0.0647	0.0708
			(0.0961)	(0.0929)
c.Y#c.CD			0.0404	0.0269
			(0.0395)	(0.0380)
c.Y#c.SA			0.0216	0.0311
			(0.0423)	(0.0434)
c.CD#c.SA			-0.0306	-0.0399
			(0.0608)	(0.0603)
_cons	-0.353***	-0.348***	-0.398***	-0.401***
	(0.0452)	(0.0469)	(0.0777)	(0.0838)
N	757	714	757	714
Moldova	Yes	No	Yes	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 26: Size component regressions for different utility sizes: tariffs

	(1) tariffs p10	(2) tariffs p25	(3) tariffs p50	(4) tariffs mean	(5) tariffs p75	(6) tariffs p90
Y	0.204**	0.221***	0.211***	0.109**	0.0507	-0.125
	(0.0951)	(0.0744)	(0.0737)	(0.0529)	(0.0808)	(0.137)
CD	-0.726**	-0.687***	-0.618***	-0.490***	-0.354**	-0.0904
	(0.309)	(0.208)	(0.167)	(0.157)	(0.168)	(0.225)
SA	0.0869	0.109	0.0982	0.0283	-0.00396	-0.123*
	(0.183)	(0.134)	(0.0920)	(0.0864)	(0.0489)	(0.0718)
N	765	765	765	765	765	765

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 27: Size component regressions for different utility sizes: cost

	(1) cost p10	(2) cost p25	(3) cost p50	(4) cost mean	(5) cost p75	(6) cost p90
Y	0.123	0.109*	0.0799*	0.0953**	0.0873	0.0968
	(0.0874)	(0.0638)	(0.0466)	(0.0390)	(0.0588)	(0.113)
CD	-0.558*	-0.447**	-0.325**	-0.350**	-0.240*	-0.189
	(0.295)	(0.198)	(0.137)	(0.145)	(0.137)	(0.174)
SA	0.0840	0.0783	0.0777	0.0674	0.0549	0.0344
	(0.124)	(0.0922)	(0.0635)	(0.0599)	(0.0352)	(0.0390)
N	757	757	757	757	757	757

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in consumer density. As shown in Tables 26 and 27, the negative coefficient for consumer density decreases with the size of the utility. The cost decrease from increasing consumer density by 1% is almost twice as large for utilities at the 25th percentile than utilities at the 75 percentile.[9]

---

[9] A utility at the 25th percentile delivery/treats a volume of 2.7 million m3 of water or sewerage per year, 25,000 consumers, and operates 2 systems. A utility at the 75th percentile delivers/treats a volume 30 million m3 of water or sewerage per year, 191,000 consumers, and operates 8 systems.

7 References

References

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Caves, D. W., Christensen, L. R., and Tretheway, M. W. (1980). Flexible cost functions for multiproduct firms. *The Review of Economics and Statistics*, pages 477–481.

Caves, D. W., Christensen, L. R., and Tretheway, M. W. (1984). Economies of density versus economies of scale: why trunk and local service airline costs differ. *The RAND Journal of Economics*, pages 471–489.

Christensen, L. R., Jorgenson, D. W., and Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The review of economics and statistics*, pages 28–45.

Garcia, S. and Thomas, A. (2001). The structure of municipal water supply costs: application to a panel of french local communities. *Journal of Productivity analysis*, 16(1):5–29.

Gelman, A. and Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

Li, G. (1985). Robust regression. *Exploring data tables, trends, and shapes*, 281:U340.

Nauges, C. and van den Berg, C. (2008). Economies of density, scale and scope in the water supply and sewerage sector: a study of four developing and transition economies. *Journal of Regulatory Economics*, 34(2):144–163.

Rosenbaum, P. R. and Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38.

Sala-i Martin, X. X. (1996). The Classical Approach to Convergence Analysis. *Economic Journal*, 106(437):1019–36.

van den Berg, C. and Danilenko, A. (2011). *The IBNET Water Supply and Sanitation Performance Blue Book: The International Benchmarking Network of Water and Sanitation Utilities Databook.* World Bank Publications.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.