Wastewater Botex Modelling for PR19: Developing a Robust Model for Regulatory Cost Assessment from an Appropriate Conceptual Model of How Cost Interactions Drive Integrated Expenditures

David Saal, Maria Nieswand, and Pablo Arocena
Loughborough University Centre for Performance and Productivity
Universidad Pública de Navarra, Spain

We know population density influences costs at firm level in most network industries

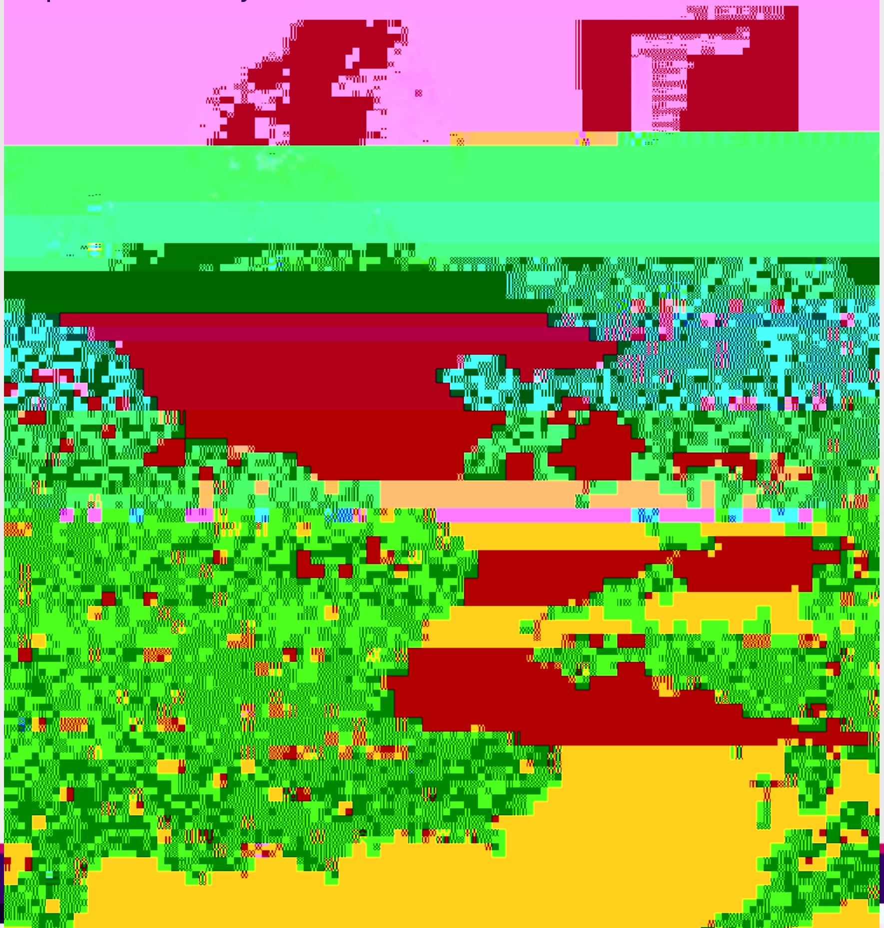
Well known that density has a nonlinear impact on costs of service (See Thames Water Consultation Response)

Academic models of network industries typically include connections, network length and squared terms and interactions with other variables to capture this impact on overall size economies

Ofwat developed measures of the proportion of served population in dense and sparse areas, but they were not been used by CEPA in its modelling

But accounting for population density is not enough to explain how the multiple optimal wastewater system designs that have been chosen by managers and engineers as the least cost solution to a given population settlement pattern resulting from demographic, economic, planning, and geographic factors influences costs.

Population density UK, 2011



Source: https://briskeat.com

Two Service Areas with alternative system configurations

Area A

4 distinct WW collection and WW treatment systems

Area B

1 Fully integrated WW collection and WW and Sludge treatment system

Configuration details of Service Areas A and B

- The size of each circle represents population.
- By construction, the sum of the areas of the circles is identical in both pictures, i.e. the population and sewage treatment load is the same in both Service areas .
- Same served area (represented by the two identical big squares) and thus same population density (population/ km^2).
- Both Service Areas generate the same volume of sludge, and each one has a single STC of equal size.
- The configuration of Service Area A:
  - Four disconnected wastewater collection and treatment systems (four WWTCs)
  - More dispersed population than Service Area B
- The configuration of Service Area B:
  - A single connected wastewater colection and treatment system (a single and larger WWTC)
  - Population is more concentrated

Key implications:

- In Service Area A most sludge is non-indigenously treated (60% by construction in picture A, 0% picture B)
- Higher transportation costs resulting from:
  - Wastewater collection (network transportation)
  - Non-indigenous sludge transport , dewatering, etc. (tankers and trucks)
- Suboptimal size of the smaller WWTCs in Service Area A, suggesting higher average costs of wastewater treatment (note scale economies are not fully represented in these Figures)

Implications for cost modelling

The indigenous versus non-indigenous treatment decision, is the result of rational cost minimizing managers' decisions with regard to

- Population size
- Population dispersion (Sparsity)
- Population density
- Sewage Collection and Sludge Transportation Costs

Need for a modelling approach capturing cost interactions between network and sewage and sludge treatment activities.

Our modelling approach is based on using indigenous and non-indigenous treatment of sludge as an effective proxy to capture key differences in systems configurations.

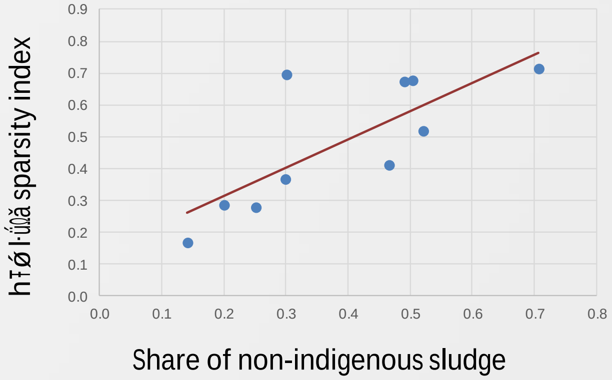
We cannot directly observe system level data, but is average Wastewater Collection and Treatment works data consistent with this conceptual model?

Company	Number of Systems (WWTCs)	Avg Pop Equiv Load per System	Connected Mains/ Pop Equiv Load	Share of Non Indig. Treated Sludge
ANH	1138	5.93	5.05	0.71
NES	412	7.35	3.98	0.47
NWT	567	15.34	3.59	0.25
SRN	365	13.17	3.50	0.52
SVT	1013	10.07	3.91	0.30
SWT	648	2.52	4.46	0.30
TMS	351	43.76	2.97	0.14
WSH	835	4.82	3.89	0.49
WSX	406	7.93	3.63	0.50
YKY	619	9.84	3.81	0.20
E. & W.	6354	12.07	3.88	0.39

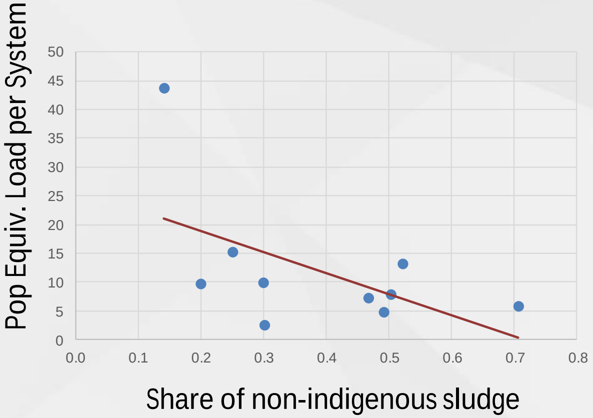
Connected Mains per Pop. Equiv. Load

Average Pop. Equiv Load per System

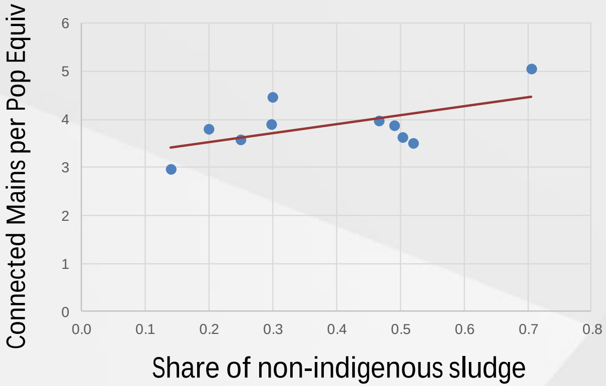
Indigenous treatment of sludge is a strong proxy of underlying demographic, geographic, planning, and economic conditions that inform decision making with regard to system design



Indigenous sludge treatment and hŦøl·ǘǚă sparsity threshold measures are positively correlated



Smaller systems with more non-indigenous treatment reflect the prohibitive cost of the networking needed to allow sufficient network scale to treat sludge indigenously



The higher the non-indigenous treatment, the larger the network length per population equivalent, reflecting the higher household passing distance required to connect households in sparsely populated areas.

The Average System Model (ASM)

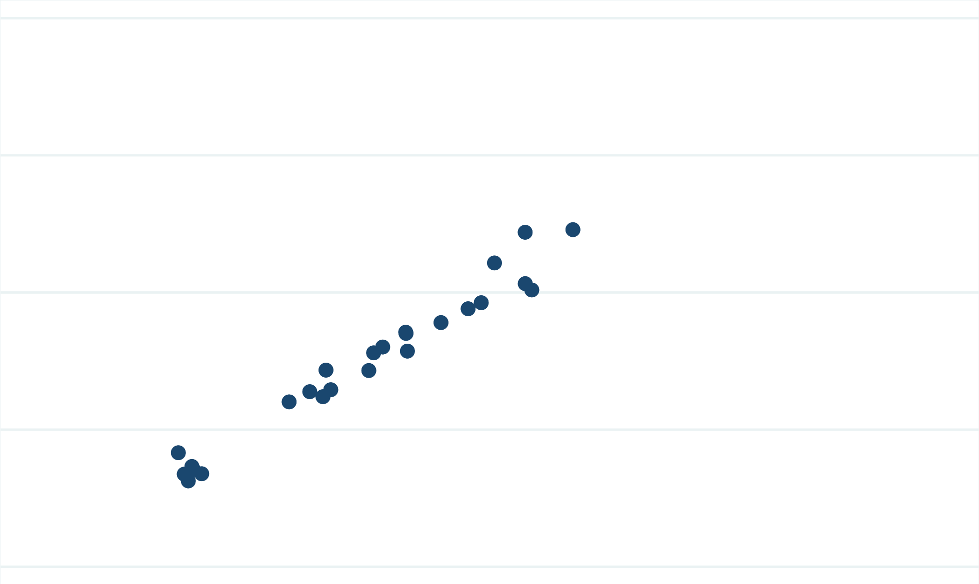
- Models conceptually average systems of a company and breaks out indigenous and non-indigenous treatment
- Average cost per system and average length and average population equivalent load for indigenous and non-indigenous treatment
- ĿƳǇƒƛŜă ǘƘĿǘ ǑƻƳǇĿyƛŜăǎ ǑƻǎǘăĿǋŜ ĿǑǘdzĿƭƭè ǘƘŜ ĿƎƎǋŜƎĿǘƛƻy ƻŦ ǘƘŜ ƻǇǘƛƳĿƭ ăèăǘŜƳ

ASM: Model specification test

Model Test Statistics	base model	general model	restricted model	restricted model with no time dummies
R-Squared	0.976	0.984	0.982	0.976
Adjusted R-Squared	0.971	0.978	0.978	[illegible]
Log likelihood	56.96	68.51	65.77	56.16
F-Test for the Overall Model	202.06	164.14	242.91	354.48
F-Test Significance	0.00	0.00	0.00	0.00
Ramsey Reset Test	1.65	0.35	0.17	1.16
Reset Test Signfiicance	0.1914	0.7875	[illegible]	0.3359
Akaike information criterion	-91.91	-103.02	-107.55	-98.32
Bayesian information criterion	-68.88	-67.41	-82.42	-83.66
root mean squared error	0.104	0.0913	0.090	0.101
VIF Max	98.68	14,261.8	29.96	29.78
VIF Mean	20.69	2,294.9	8.05	13.22
	60	60	60	60
Number of Parameters including Time Dummies	11	17	12	7
residual degrees of freedom	49	43	[illegible]	53
[illegible]	0.42	0.00	1.57	0.00
significance	0.26	1.00	0.11	1.00

Loughborough University

ASM also has very good predictive properties



Averages by Company

Comp	botex	prbotex	delta	deltashr
ANH	330.9	326.3	-4.6	-1.5%
NES	141.7	145.5		

What prevented these models in the CEPA report?

- A priori restrictions
  - No more than six variables
  - No translog, i.e. no squared and cross terms of variables
  - VIF < 5, i.e. no variables that are correlated > 90%
  - A modelling approach that defines cost groups (output, density, system [illegible] variable form each group, thereby fundamentally excluding the potential to model the complex systems and cost interactions that exist in the water and sewerage industry.
- We did not impose those constraints but our models pass all the appropriate tests.

Reproducing the OWWW3 Model

Reproduction of Representative Ofwat model OWWW3 with Botex definition agreed with Anglian

- Is representative and typical of models presented by Ofwat in the cost assessment consultation
- Conceptually models a company
- . I·ăŜŘ ƨуʃűKŜ I· ǇǊƛƨǊƛ I·уŘ Kƛ∃Kfè ǊŜăűǊƛŎűƛǾŜ ϓƨŘŜffƛу∃ I·ăădzϓǇűƛƨуã ƨdzűfƛуŜŘ ƛуʃűKŜ /9t!ℓă tвмф℮ă Ŏƨăű I·ăăŜăăϓŜуű ǊŜǇƨǊű
- A priori assumptions include: 6 groups of cost drivers, only one output variable max 6 variables, VIF < 5, no translog terms etc.

Representative Ofwat Model OWWW3 reproduced with BOTEX

[illegible] more flexible, specification in light of the failure of the Reset test. This search did not yield a better model

Appendixc1 Modelling Results,

OWWW3 model fails many specification tests and has very high range in delta share estimates (not resolved using RE and/or time dummies)

Averages by Company				
Comp	botex	prbotex	delta	deltashr
ANH	330.9	287.2	-43.7	-15.3%
NES	141.7	136.5	-5.2	-3.7%
NWT	437.4	415.9	-21.4	-5.2%
SRN	262.5	231.3	-31.2	-13.5%
SVT	388.0	407.9	19.9	4.8%
SWT	129.5	126.0	-3.4	-2.7%
TMS	575.0	604.0	29.0	5.1%
WSH	205.4	230.5	25.1	10.9%
WSX	126.2	140.7	14.5	10.4%
YKY	266.2	267.0	0.8	0.4%
Total	286.3	284.7	-1.6	-0.9%
			Range of Company Averages	26.2%

Averages by Time				
Yearend	botex	prbotex	delta	deltashr
2012	253.8	279.1	25.3	5.3%
2013	278.7	283.7	5.0	-1.0%
2014	283.5	283.7	0.2	-1.1%
2015	280.0	286.4	6.5	4.7%
2016	297.7	286.9	-10.7	-1.6%
2017	324.1	288.4	-35.7	-11.4%
Total	286.3	284.7	-1.6	-0.9%
			Range of Time Averages	16.7%

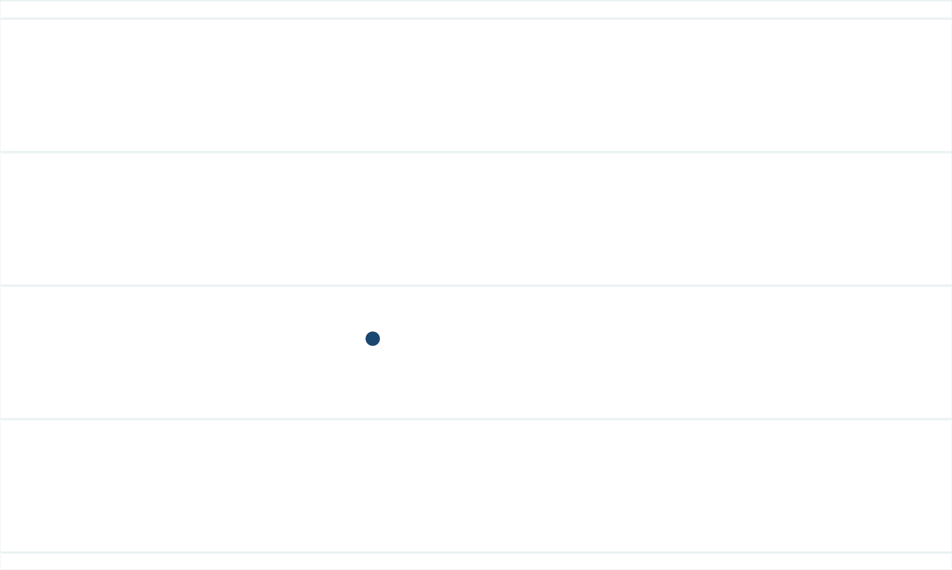
Conclusions

Appendix.

Alternative model: Extended Passing distance model (PDM)

- Developed with Anglian
- Models conceptually a company and breaks out indigenous and non-indigenous treatment thereby capturing differences in the combined botex elasticity and marginal costs of both overall wastewater and sludge treatment activities

PDM passes model specification tests, has relatively good fit, but has relatively high range in delta share estimates



Averages by Company				
Comp	botex	prbotex	delta	deltashr
ANH	330.9	328.4	-2.5	-0.8%
NES	141.7	132.1	-9.5	-7.4%
NWT	437.4	437.7	0.4	0.0%
SRN	262.5	242.6	-19.9	-8.5%
SVT	388.0	425.6	37.6	8.6%
SWT	129.5	128.2	-1.2	-0.8%
TMS	575.0	545.0	-30.0	-5.1%
WSH	205.4	239.2	33.7	13.9%
WSX	126.2	129.7	3.6	3.2%
YKY	266.2	244.4	-21.8	-8.6%
Total	286.3	285.3	-1.0	-0.6%
	Range of Company Averages			22.4%

Averages by Time				
Yearend	botex	prbotex	delta	deltashr
2012	253.8	258.2	4.4	-0.8%
2013	278.7	284.0	5.4	-0.3%
2014	283.5	284.8	1.3	-0.3%
2015	280.0	270.0	-9.9	-0.6%
2016	297.7	291.0	-6.7	-0.6%
2017	324.1	323.7	-0.4	-0.7%
Total	286.3	285.3	-1.0	-0.6%
		Range of Time Averages		0.5%