Utrecht University
School of Economics

The Optimum Size of Local Public Administration.

Jacob A. Bikker
Daan van der Linde

U.S.E. Discussion Paper Series nr: 15-05

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Kriekenpitplein 21-22
3584 EC Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the second author.

Jacob A. Bikker
Daan van der Linde
Utrecht University
Utrecht School of Economics
Kriekenpitplein 21-22
3584 TC Utrecht
The Netherlands.
E-mail: d.e.vanderlinde@uu.nl

This paper can be downloaded at: http:// www.uu.nl/rebo/economie/discussionpapers

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 15-05

The optimum size of local public administration

Jacob A. Bikker
Daan van der Linde

Utrecht School of Economics
Utrecht University

April 2015

Abstract

A renewed interest in decentralization has profoundly affected local public governance around the world. Faced with an increasing number of tasks, Dutch municipalities have recently sought physical centralization, merging into larger jurisdictions in order to target new policy areas more effectively and cost efficiently. Is such a policy of physical centralization wise? We study economies of scale in local public administration, and find – given transfer payments from central government and current cooperation between municipalities and after controlling for geographical, demographic and socio-economic variables – substantial unused scale economies of 20% for the average municipality. Between 2005 and 2014 the optimum size of municipalities increases from around 49,000 to 66,260 inhabitants, pointing at an increased importance of fixed costs relative to variable costs in local public administration. For other local government activities, we find either lower or no optimum scales.

Keywords: Fiscal federalism, Decentralization, Local public administration; Municipalities.

JEL classification: H71, H77

Acknowledgements
The authors are grateful to Mark Kattenburg and participants of the workshop on local government reform in the Netherlands and Spain (Universitat de Barcelona, October 2014) for useful comments.

I. INTRODUCTION

The analysis of the allocation of competencies and fiscal instruments across different levels of government dates back as far as Alexis de Tocqueville's (2000/1838) *Reflections on America*. Renewed interest in what is known as the research field of fiscal federalism has sprung from publications by Tiebout (1965) and Oates (1972), the latter in 1999 proclaiming decentralization to be 'back in vogue' (Oates 1999).

In what may be seen as a pendulum motion of decentralization policies around the world, many states have undergone policy reforms moving from decentralization to recentralization and back again (see *e.g.* De Vries 2000). At certain points in time, policy makers around the world valued the qualities of diversified local goods and services provision, better tailoring supply to local preferences, only later to focus on the efficiency gains from more uniform centralized policy, seeking to reap advantages of scale while reducing diversity across jurisdictions. The Netherlands has recently witnessed a paradoxical combination of policies aiming at increased decentralization while at the same time exerting pressure for *physical* centralization of local authorities to larger jurisdictions (Allers & De Kam 2010). Central government has delegated its Social Support and Provision Act (*Wet Maatschappelijke Ondersteuning; in Dutch*) to local government level, now proposing a near doubling of central government transfers to municipalities, delegating a wide array of new tasks including welfare and child welfare services to decentralized levels of government. In order to target these new policy fields effectively, central government asserts that jurisdictions must merge to 100,000 residents or more.

This paper focusses specifically on local public administration expenditure in the Netherlands, assessing at which number of inhabitants the optimum size for public administration is achieved. Section II gives an overview of the literature, while Section III presents our method, proposing a new set of functional forms in order to test for optimum size. Section IV includes economies of scale estimates in local public administration. In Section V we show our method can simultaneously distinguish between scale economies and managerial or X-inefficiency, which we show to be independent from scale. Section VI presents an overview of scale economies in other cost categories than local public administration. We will conclude with a reflection on our findings in Section VII.

II. THEORETICAL FRAMEWORK

Fundamental to economic thought about decentralization is the assumption of variance in demand for public goods. Heterogeneity in preferences can be utilized locally if the benefits of decentralization have a positive correlation with variance in demand (Panizza 1999). As the literature assumes, highly centralized systems are unable to respond to variance in preferences at the local level (Koethenbuerger 2008); in highly centralized systems with uniform levels of public goods and national taxes, individuals can only exert pressure on the allocation of their gross income through *voice* (Hirschman 1970). In his seminal paper, Tiebout (1956) introduces the *exit* option for individuals, claiming that demand and supply for local public goods can also be influenced by 'voting with one's feet'. A fully mobile consumer-voter can relocate to a district that satisfies his demand for local public goods at the cost of a matching local tax rate. While in centralized systems there is no incentive to reveal one's *true* preferences for public goods in order to free ride on others' demand for these goods, Tiebout (1956) claimed that when one 'votes with one's feet', moving to the region that better combines one's preferred level of local public goods and accompanying taxes, there is no longer an incentive to hide one's preferences. Using this information to coordinate local taxation and spending may then improve well-being throughout society.

But how much decentralization should we demand, and to *whom* we should decentralize? Oates' (1972) decentralization theorem seeks to answer this question, stating that from an efficiency perspective fiscal responsibilities should *always* be decentralized as long as there are (i) no cost savings to be gained from centralization and (ii) no interjurisdictional externalities.

Our paper focuses on the first argument, seeking to estimate the optimum size for municipalities in terms of local public administration (LPA). Our choice is guided by the fact that (i) costs of LPA generally does not seem to impose externalities on neighboring municipalities, (ii) LPA is a mainly homogenous service in which the *type* of tasks performed does not largely depend on the size of municipalities, as may be the case with other areas of spending and (iii) LPA pertains to a core task of municipalities. All local authorities face the costs of public administration, and larger municipalities can spread the costs across a larger number of inhabitants. This classic explanation of economies of scale however wears out as size increases further, as managerial problems, complexity, bureaucracy, declining motivation and commitment of staff, *et cetera*, may again induce rising costs of inefficiency from a certain point onward. The key question is whether such large local public administrative inefficiencies outweigh the basic effect of the monotonically declining fixed costs.

Two main strands of empirical literature have touched upon cost efficiency and scale at the local level. A first strand uses cross-sectional or panel data to study the general relationship between local government expenditure and population size, while a more recent strand of literature has provided quasi-natural experimental evidence for economies of scale using amalgamations of municipalities. Regressing local expenditure classes on a linear and quadratic population variable while controlling for a range of confounding variables, Drew *et al.* (2014) provide evidence of U-shaped cost curves in

Queensland, Australia with economies of scale up to 98,000 inhabitants. With Cobb-Douglas and translogarithmic cost functions, Geys, Heinemann & Kalb (2007) estimate the cost elasticity with respect to inhabitants for five different population classes of municipalities in Baden-Württemberg, Germany. The authors find substantial economies of scale that however wear out after approximately 10,000 inhabitants. Solé-Ollé & Bosch (2005) use a piecewise linear function postulating a nonlinear relationship between population size and costs in Spain, concluding that two local optimums at 5,000 and 50,000 inhabitants exist. Several studies have used a difference-in-difference framework to study the effects of municipal amalgamations. Reingewertz (2012) shows that amalgamations of small municipalities (around 10,000 inhabitants) lead to lower *per capita* expenditure levels, while not affecting quality. Using a government reform as an exogenous shock to the scale of Danish municipalities, Blom-Hansen, Houlberg & Serritzlew (2014) perceive scale effects with lower administration costs per inhabitant up to 100,000 inhabitants. While amalgamations were not exogenously determined as in the Danish case, Allers & Geertsema (2014) for the Netherlands show that voluntary amalgamations do not affect aggregate spending, while providing evidence for reduction in spending on local public administration.

III. RESEARCH DESIGN

1. Functional forms for local public administration

This section discusses the functional form of several familiar cost models and their underlying assumptions with respect to the shape of average costs per unit curve, which have a major impact on the existence and size of an optimum scale. Next, we develop empirical models for local public administration (LPA), or other local government activities. This section also explains the measurement of economies of scale.

The measurement and analysis of differences in LPA cost levels is based on the assumption that LPA production technology can be described by a production function that links the various types of LPA output to input factor prices, such as wages, office space rent, and so on. Under certain conditions, a dual cost function can be derived, using output levels and factor prices as arguments (Coelli *et al.*, 1998, pp. 43-49). In the literature, the translog cost function (TCF) to describe costs dominates other model specifications. Christensen *et al.* (1976) proposed the TCF as a second-order Taylor expansion, usually around the mean, of a generic function with all variables appearing as logarithms. This TCF is a flexible functional form that has proven to be an effective tool for the empirical assessment of efficiency, both in banking and elsewhere (Christensen *et al.*, 1976; Dietsch, 1993; Nauriyal, 1995; Edirisuriya *et al.*, 2001). It is an extension of the Cobb-Douglas function, which is capable of fitting

U-shaped average cost functions.[1] The TCF has also been applied for scale economies in local expenditure on waste, roads and parks by Drew *et al.* (2014) and Geys *et al.* (2007). A simple TCF reads as follows:

$$\ln LPAC\,(inh) = \alpha + \beta_1 (\ln inh) + \beta_2 (\ln inh - \overline{\ln \mathrm{inh}})^2 \qquad (1)$$

with *LPAC* for local public administration costs and '*inh*' or inhabitants for output volume, where we use the number of inhabitants. Note that in the squared term we take the logarithm of output in deviation from its mean (denoted by the bar above the variable), in line with the Taylor expansion.[2] In that case the model is extended with cross terms from both output measures. Since the Netherlands is a relatively small country, we expect little or no variation in input prices (Swank, 1996). Actually, LPAs do not report data on input prices, so that we are also unable to include municipality specific prices in the cost functions.[3] Unused scale economies exist where $\beta_1 < 0$, while concavity or a U-shaped average cost function requires $\beta_2 > 0$.

Shaffer (1998, page 94) proves that for a sample of monotonically declining average costs the TCF would estimate a concave function with an optimum scale, so that the existence of an optimum size and diseconomies of scale for larger firms is (incorrectly) imposed.[4] Indeed, the left leg of the TCF can be fitted to the hyperbolically declining average costs, with the optimum scale in the right-hand tail of the sample, or beyond the largest observation. Consequently, Shaffer (1998) suggests two additional cost functions to estimate scale economies which do not impose this U-shaped average-cost function.

The first alternative is the unrestricted Laurent function (ULF), which is similar to the TCF, but with two inverse terms added:

$$\ln LPAC\,(inh) = \alpha + \beta_1 (\ln inh) + \beta_2 (\ln inh - \overline{\ln \mathrm{inh}})^2 + \beta_3 /(\ln inh) + \beta_4 /(\ln inh)^2 \qquad (2)$$

The ULF can describe monotonically declining average cost, does not impose an optimum scale and allows different degrees of curvation for smaller and larger municipalities. For the concave properties to hold, the coefficients β_3 and β_4 should both be positive, next to β_2. According to Shaffer (1998), the

---

[1] For shortcomings of the TCF, see Shaffer (1998, p. 91).
[2] White (1980) and Shaffer (1998, p. 95) explain that this specification also helps to avoid multicollinearity. Note that $\overline{\ln \mathrm{inh}}$ is the arithmetic average of the *logarithms* of output measure inh_i.
[3] To simplify the presentation, we do not include one or more input prices in Eq. (1). Normally, a TCF would also include cross terms between output and input prices. Dropping input prices and other output measures implies that no cross terms remain. We do have data on real wages in the LPA that picks up input price effects over time, but at national level only, but that wage index is strongly correlated with other time trends. Hence, for practical reasons, we have not included that index in our models.
[4] Except possibly over limited ranges of scale with steeply declining marginal costs.

improvement of the ULF over the TLF may die down for skewed size as the squared nature of the cost-output relationship is built up by the relatively large share of observations in the smaller size region of the data sample. He consequently proposes a second alternative: the hyperbolically-adjusted Cobb Douglas (HACD) cost function, see also Adanu *et al.* (2009). Again ignoring input prices, this model reads as

$$\ln LPAC\,(inh) = \alpha + \beta_1 (\ln inh) + \beta_2 / inh \tag{3}$$

Thanks to the additional reciprocal term, this model can portray the U-shaped average cost function ($\beta_1 > \beta_2$), monotonically declining average costs ($\beta_1 > 1$ and $\beta_2 > 0$) and the L-shaped average cost function ($\beta_1 < 1$). Finally, we suggest an asymmetric TCF (ATCF), as the steepness of the left leg, reflecting scale economies, may differ from the mirrored steepness of the right leg, which describes diseconomies of scale, related to managerial problems, complexity, bureaucracy, *et cetera*. The squared term is split into positive and negative deviations from the mean, each with its own coefficient. To investigate which functional form best suits the sample data, we will apply Akaike's (1974) information criterion (AIC).

Cost elasticity (*CE*) is defined as the proportional increase in costs as a result of a proportional increase in output. In mathematical terms this results in the following formula for elasticity: $CE = \partial \ln OC / \partial \ln o$. Using Eq. (1-3), this results in for TCF, ATCF, ULF and HACD respectively:

$$CE^{tcf} = \beta_1 + 2\,\beta_2 (\ln inh - \overline{\ln \mathrm{inh}}) \tag{4}$$

$$CE^{atcf} = \beta_1 + 2\,\beta_2^{+}\,(\ln inh - \overline{\ln \mathrm{inh}})^{+} + 2\,\beta_2^{-}\,(\ln inh - \overline{\ln \mathrm{inh}})^{-} \tag{5}$$

$$CE^{\,ulf} = \beta_1 + 2\,\beta_2 (\ln inh - \overline{\ln \mathrm{inh}}) - \beta_3 /(\ln inh)^2 - 2\,\beta_4 /(\ln inh)^3 \tag{6}$$

$$CE^{\,hacd} = \beta_1 - \beta_2 / inh \tag{7}$$

The second term of the *CEs* in the TCF and the ULF becomes zero if the *CEs* are evaluated around the mean of the sampled logarithms of inhabitants inh_i, that is: $\overline{\ln \mathrm{inh}}$. The *CE* for the TCF (and the ATCF) is then equal to β_1, while for the ULF and the HACD, it depends on the sample observations.[5] We have also assumed a simplified ULF (SULF), that is (Eq. (2) with $\beta_4 = 0$ where the two inverse terms are too highly intercorrelated.

[5] For the reciprocal terms in the ULF we replace ln *inh* by the geometric mean of the output term. For the reciprocal term in the HACD model, arguments (to replace *inh*) exist in favour of both the geometric and arithmetic mean. This choice has no further consequences.

The scale economies (*SE*) can easily be distilled from the above by subtracting *CE* from unity: *SE* = 1 – *CE*. If the calculated *CE* has a value larger than 1, this indicates diseconomies of scale; a value smaller than 1 indicates economies of scale and a value of exactly 1 indicates constant returns to scale. To calculate a possible optimum size of a LPA, a value for *inh* has to be found to set *CE* equal to one (or to set *SE* to zero).

2. *Data*

We use annual data on budgeted municipal spending across nine expenditure items, collected by Statistics Netherlands (CBS 2014).[6] Table 1 shows costs per inhabitant for each of these items averaged, divided into five municipal size classes, over the 2005-2014 period. During these years, the size of Dutch municipalities increased markedly, while the number of municipalities fell to 403 from 467. This decline was most profound for municipalities with fewer than 20,000 inhabitants (declining in number to 137 from 223). Table 1 shows how costs for most categories initially decrease as size (measured in number of inhabitants) increases, hinting at economies of scale. However, as size increases further, costs rise again, suggesting larger-sized municipalities face increasing complexity and inefficiency. This pattern is clearly visible for the LPA cost categories, public order and safety and public health and environmental affairs, indicating U-shaped cost curves.

Table 1: Average budgeted expenses per inhabitant for each cost category (2005-2014, in 2005 EUR)

	Number of inhabitants (x 1000)				
	< 10	*10-20*	*20-50*	*50-100*	*> 100*
1. Local public administration (LPA)	270	163	142	**137**	170
2. Public order and safety	94	68	**66**	87	102
3. Infrastructure	171	155	**147**	195	279
4. Economic affairs	**13**	15	18	31	62
5. Education	157	**142**	150	203	245
6. Culture and recreation	222	**192**	214	259	314
7. Social services	**400**	476	551	814	1094
8. Public health and environmental affairs	284	236	**227**	243	292
9. Spatial planning and housing	**256**	290	348	449	597
10. Total expenses	1868	**1737**	1863	2419	3154
% of spending on public administration	14.4	9.4	7.6	5.7	5.4
Average number of municipalities	*48*	*128*	*188*	*43*	*26*

Note: The municipal classes with the lowest average costs per inhabitant are marked in bold.

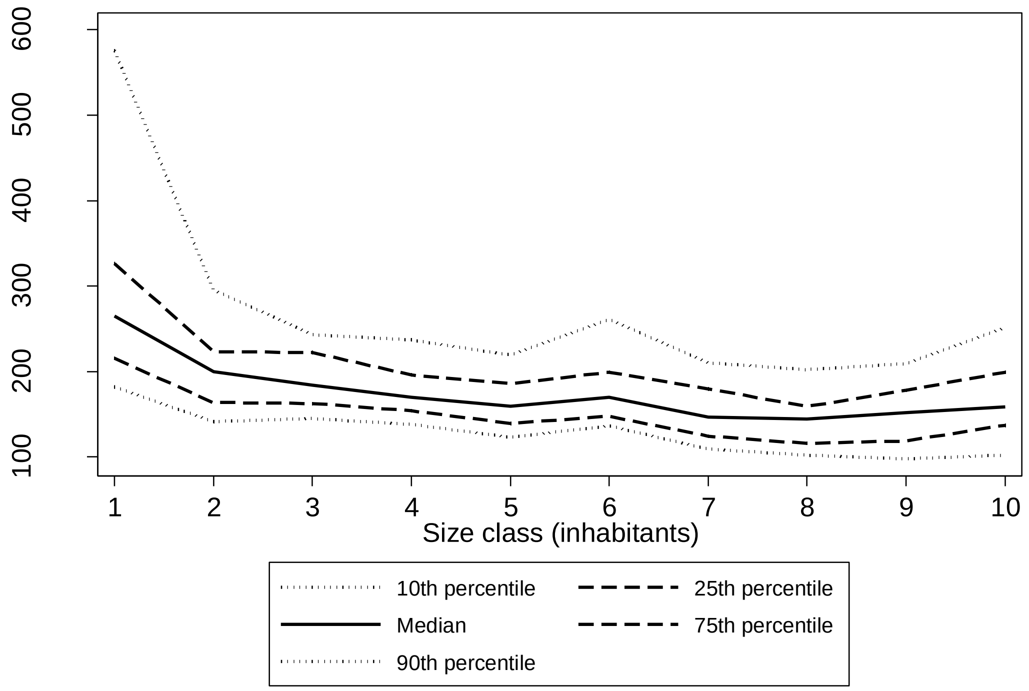
Social services are by far the largest expenditure item in municipal budgets – a cost category which in the light of future decentralization of tasks in the Netherlands is likely to keep increasing. For certain expenditure items municipalities act as mere intermediaries of national policies and have limited spending autonomy. More local autonomy is found in cost items such as spatial planning and housing,

[6] In the remainder of the paper we use budgeted, not actual expenditure due to data availability. When data overlap, we find both accounts are highly correlated and find estimates are robust to regressions on actual expenditure.

public transport, education – especially the housing thereof, waste management and the organization of bureaucracy. For some of the items we must assume cascading cost functions: the number of municipal tasks increases with size. For instance, small municipalities may suffice with just one grade school while larger municipalities serve a more regional function, also offering high schools or vocational education. This regional function entails higher costs, but at the same time this municipality also provides a higher level of services, possibly also to surrounding municipalities.

A final note should be made on central government transfers from the municipal fund (*Gemeentefonds*) and intermunicipal cooperation. A large share of local revenue in the Netherlands is made up of central government transfers, which among other factors in its distribution takes into account local expenditure needs, the capacity for local revenue collection, but also economies of scale in the provision of local goods and services. Consequently, our findings are to be interpreted *given* the current distribution of transfers. Similarly, as Bel, Fageda & Mur (2013) show, cooperation in the production of local services may lead to small municipalities in fact reaching their optimum scale, and results thus are inclusive of current intermunicipal cooperation.

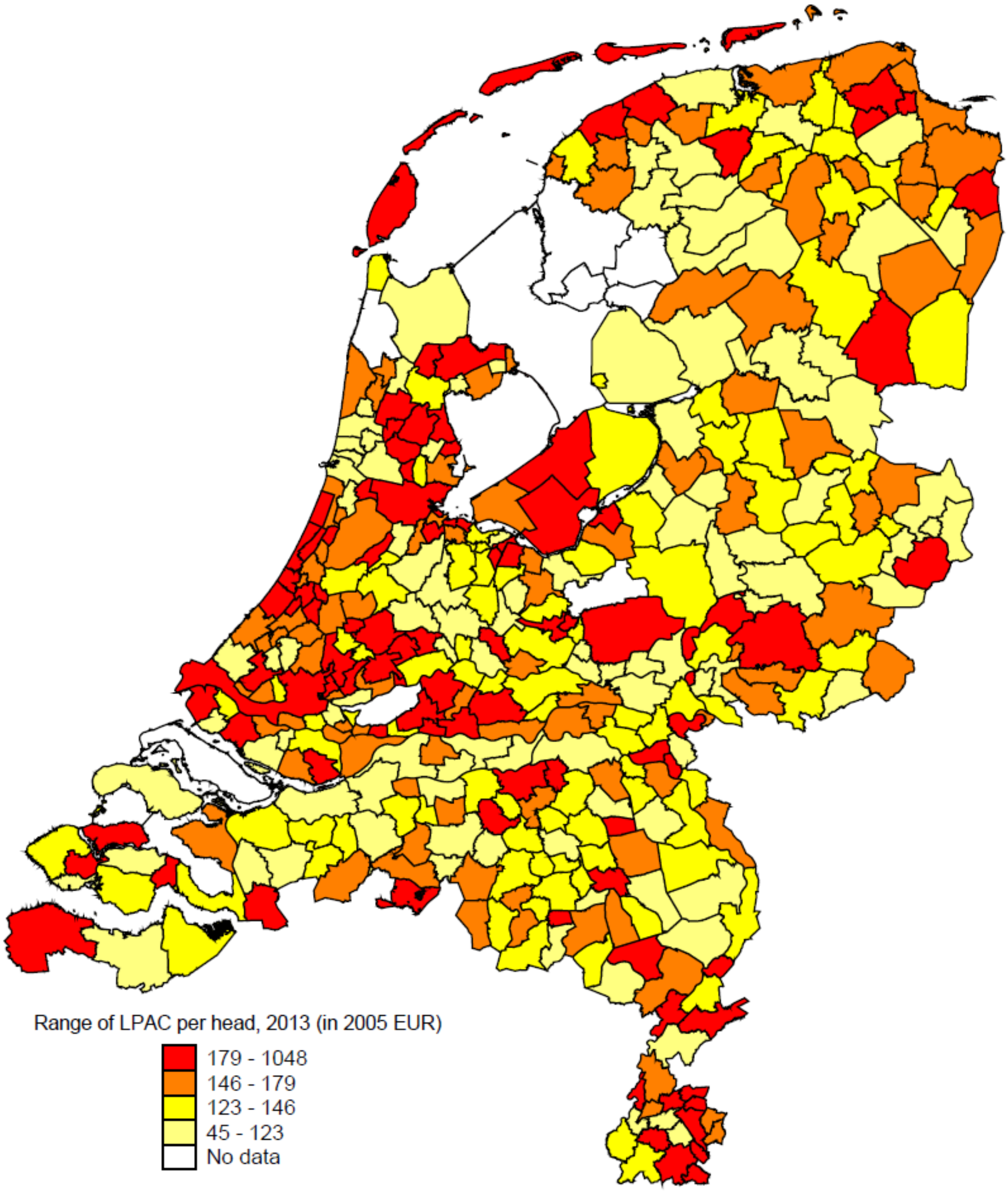
Figure 1: Costs of local public administration per inhabitant for ten size classes (for 2013 in current prices)



Data: CBS (2014)

Figure 1 shows the LPAC per inhabitant for the year 2013 in ten size classes. We find that average costs decline as the municipality grows (particularly for the smallest size classes), as well as strong variation within each size class. The former hints at economies of scale, while the latter points at differences in either efficiency across municipalities irrespective of their size, or heterogeneity where different municipalities have different tasks or service levels. Figure 2 gives the spatial distribution of LPAC per head for 2013. We find that relatively high per head LPAC costs are concentrated in the West Frisian or *Wadden* islands (averaging EUR 714 per head compared to the national average of EUR 163 per head), yet apart from this we have not observed a clear relationship between geographical location and costs. The three large cities (Amsterdam, Rotterdam and The Hague) rank relatively high, as do the Flevoland province and the Western part of the South Holland province.

Figure 2: Spatial distribution of the costs of local public administration per inhabitant (for 2013 in 2005 prices)



IV. ECONOMIES OF SCALE ESTIMATES IN LOCAL PUBLIC ADMINISTRATION

This section includes our results for a cross-sectional stochastic cost frontier (SCF) estimation (Coelli *et al*., 1998; Beloti *et al.*, 2012) for economies of scale in LPA activities using five functional forms. The SCF approach allows estimations using two error term components under the assumption of interindependency. The first error term v_{it} is normally distributed and represents errors in data and model specification, while the second error term u_{it} is a one-sided non-negative disturbance and approximates managerial or X-inefficiency. Throughout our estimations we have assumed u_{it} to be exponentially distributed with $u_{it} \sim \mathcal{E}(\sigma_u)$ (Meeusen & Van den Broeck 1977).[7] This estimation technique allows us to simultaneously study the existence of economies of scale in LPA work, and to estimate efficiency differences across similarly sized municipalities. In measuring scale economies, it is better to take inefficiencies into account and use SCF than to ignore inefficiencies and use OLS. Our model for LPA costs reads as follows:

$$ln(LPAC_{it}) = \beta_0 + \sum_j X_{ijt}\,\beta_j + \sum_k Z_{ikt}\,\gamma_k + u_{it} + v_{it} \quad (8)$$

where X_{ijt} represents variables (j = 1, …, N) relating to the number of inhabitants in municipality i in year t, and depends on our five functional specifications (TCF, ATCF, HACD, ULF or SULF). As exclusion of control variables may distort the estimation of the relationship between scale and LPAC, Z_{ikt} contains a number of control variables (k = 1, …, M): demographic pressure, defined as the sum of young (under 16 years of age) and old inhabitants (over 65 years of age) as a share of the working population (between 16 and 65), wealth, measured as the average home value (in current EUR 10,000), land surface (in 10 square kilometers), a dummy for the 'G4' (the four largest municipalities: Amsterdam, the Hague, Rotterdam and Utrecht), a dummy for the *Wadden* Islands (Texel, Vlieland, Terschelling, Ameland and Schiermonnikoog) and a time trend (*year*). Table 2 shows summary statistics for our dependent and control variables, including the within component of the standard deviations. For most variables, the table shows little variance within municipalities over time.

Table 2: Key features per municipality (2005-2014)

		Mean	**SD**	**Min**	**Max**	**Observations**
Local public administration costs, per capita (in 2005 EUR)	Overall	163.33	79.93	29.46	1138.95	N = 4327
	Within		27.59			
Inhabitants	Overall	38239	60701	932	810937	N = 4327
	Within		2286			
Demographic pressure	Overall	68.9	7.1	44.3	109.9	N = 4327
	Within		2.6			
Average home value (in current 10,000 EUR)	Overall	24.355	6.337	11.600	70.200	N = 4327
	Within		1.713			
Surface land (in 10 sq. km)	Overall	0.78	0.69	0.018	4.6	N = 4327
	Within		0.05			

Note: SD is standard deviation. 'Within' refers to 'within' municipalities.

[7] Our estimates are robust to Half-normal and Truncated normal distributions of the inefficiency term.

Table 3 presents the estimation results. According to the Wald test, all five models reject constant returns to scale. Based on the Akaike Information Criterion (AIC), the rather restricted HACD shows the worst performance,[8] while the other models' performance does not diverge much. The more flexible ATCF outperforms the TCF, but this model is still not sufficiently flexible, as it is

Table 3: SCF estimates of economies of scale in local public administration (2005-2014, in 2005 EUR)

	TCF	ATCF	ULF	Simplified ULF	HACD
Inhabitants (in log.)	0.823$^{***/000}$	0.768$^{***/000}$	3.115^{000}	-0.406^{000}	0.906$^{***/000}$
	(52.28)	(26.02)	(0.67)	(-0.95)	(36.76)
Inhabitants^2 (ln, mean dev.)	0.094***		0.035	0.225***	
	(8.27)		(0.14)	(4.36)	
1/inhabitants					1280.546***
					(4.14)
1/(ln inhabitants)			546.222	-123.085**	
			(0.62)	(-2.86)	
1/(ln inhabitants)^2			-1566.397		
			(-0.77)		
Inhabitants ^{2} (ln, mean dev., below mean)		0.139***			
		(5.00)			
Inhabitants ^{2} (ln, mean dev., above mean)		0.053**			
		(2.94)			
Average house value (in 10,000 EUR)	0.004*	0.004*	0.004*	0.004*	0.004*
	(2.05)	(2.07)	(2.05)	(2.11)	(2.10)
Surface land (in 10 sq. km)	-0.035*	-0.035*	-0.035*	-0.035*	-0.038*
	(-2.00)	(-2.02)	(-2.00)	(-2.03)	(-2.07)
Demographic pressure	-0.002	-0.001	-0.001	-0.001	-0.004*
	(-1.09)	(-0.64)	(-0.54)	(-0.55)	(-2.23)
G4 municipalities	0.193	-0.046	0.010	-0.082	0.779***
	(1.12)	(-0.21)	(0.04)	(-0.40)	(3.66)
Wadden Islands	0.535***	0.679***	0.727***	0.720***	0.574*
	(3.48)	(5.28)	(5.99)	(5.91)	(2.50)
Year	0.015***	0.015***	0.015***	0.015***	0.016***
	(7.55)	(7.44)	(7.43)	(7. 45)	(7.57)
Constant	-23.898***	-22.730***	-85.126	1.374	-26.273***
	(-6.07)	(-5.80)	(-0.75)	(0.15)	(-6.37)
σ_u (X-inefficiency)	0.117	0.115	0.114	0.115	0.124
σ_v (model and data errors)	0.227	0.227	0.227	0.227	0.234
λ	0.514	0.508	0.502	0.506	0.528
F-statistic	3,416.93***	2,737.69***	111.65***	3,410.98***	2,991.36***
Wald test [a]	227.32***	252.21***	263.65***	255.00***	112.57***
Cost elasticity at mean [b]	0.823	0.768	0.803	0.797	0.872
Scale economies at the mean	0.177	0.232	0.197	0.203	0.128
Ln (mean)	10.114	10.114	10.114	10.114	10.114
Optimum size (in inhabitants) [c]	63,400	221,100	55,900	57,600	-
Lower 95% bound	49,400	69,300	45,400	47,900	-
Upper 95% bound	92,200	-	76,300	77,300	-
Akaike's IC	455.041	426.38	412.77	413.39	734.63
Number of observations	4,111	4,111	4,111	4,111	4,111

Note: Standard errors clustered by municipality. *, ** and *** mean significantly different from zero at, respectively, the 95%, 99%, and 99.9% confidence level, while ^{000} indicates that the Wald test on constant returns to scale (see footnote *a*) is rejected at the 99.9% confidence level.

[a] The Wald test regards the constant returns to scale hypothesis: the coefficient of the linear term ln (number of participants) is equal to 1, and the coefficient(s) of the non-linear term(s) is (are) equal to 0; [b] With mean referring to the mean value of the output measure, the number of inhabitants, see Section III.1; [c] The optimum scale is calculated by setting the first derivative equal to one, see Section III.1 for details.

[8] This contrasts with a simulation test by Shaffer (1998) on generated data, where the HACD model performed best.

outperformed by the ULF models. The AIC values of the ULF and SULF are roughly equal,[9] so that we prefer the most parsimonious model, the SULF. For the SULF, we find a cost elasticity of 0.797 at the geometric mean of 24,686 inhabitants.[10] Hence, keeping our control variables constant, we find that a 1 percent increase in the number of inhabitants pushes up costs for public administration by 0.797%. This implies average unused scale economies of 20.3%. Note that the results across the models, also with regard to scale economies, are quite similar, except for the poorly performing HACD.

Turning to the SULF model, a number of control variables significantly affect the costs of local public administration. An EUR 10,000 increase in the average home value, increases spending on public administration by 0.4%, holding all other variables constant. Apparently, wealthier municipalities can afford and demand higher service levels. We cannot exclude the possibility of reverse causality in this variable, as we may to a limited extent expect the quality of local public administration to in turn affect house prices. On average, municipalities covering larger areas have lower costs, as a 10 square kilometer increase in surface depresses costs by 3.6%, holding other factors constant. Possibly, more agricultural municipalities have more austere preferences. All other things being equal, the isolated *Wadden* islands on average spend 105% more on local public administration than other municipalities, while we find no discernible differences between the 'G4' municipalities and others, indicating that these large municipalities are a good fit to our model. In addition , we observe that spending on local public administration (in constant prices) increases over time by roughly 1.4% annually. λ refers to the (estimated) signal-to-noise ratio (σ_u/σ_v), which lies around 50% for the SULF. Low values for λ may infer that it is difficult to differentiate between efficiency and noise from the model, which does not seem to be the case for our model specifications.

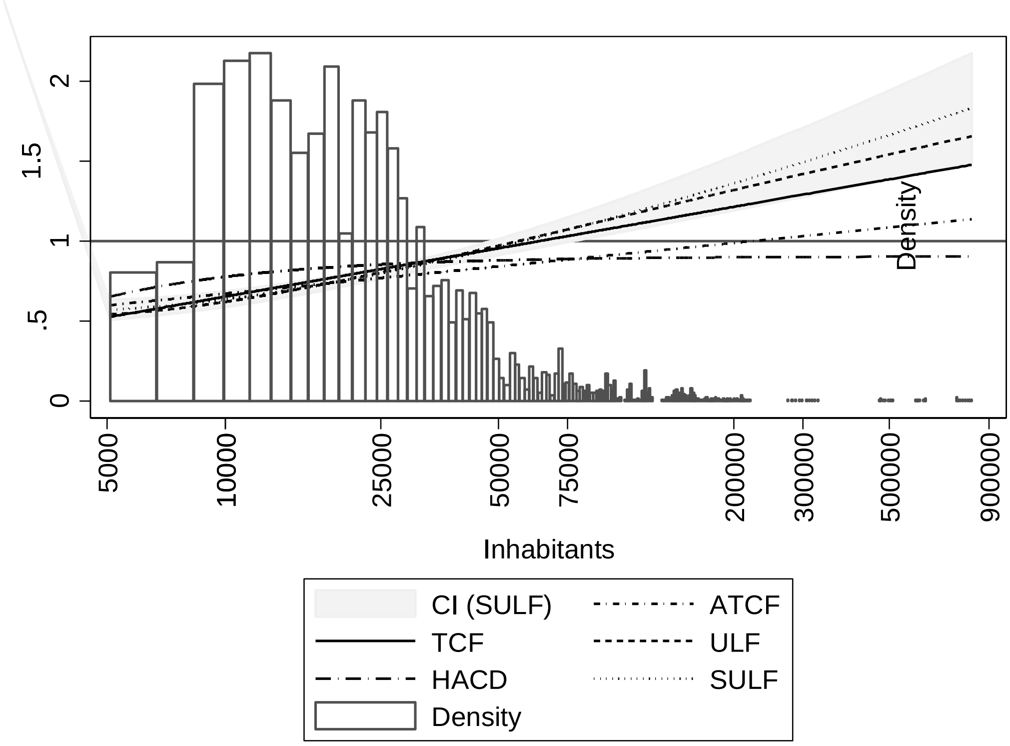
The coefficients of the squared and inverse terms of the SULF are quite large and highly significant, implying substantial curvature in scale economies, see Figure 3. This graph reflects the predicted cost elasticity for our four estimations over the 2005-2014 period as a function of the number of inhabitants, including 95% confidence bands for the SULF. The confidence bands for the ULF are wider, as this model is less parsimonious, while the bands for the TCF and the HACD are much smaller. This graph shows the effect of the non-linear terms of the output measure inhabitants. TCF is a straight line, while the ULF and SULF results in a convex curve. By nature the HACD is concave. The background histogram gives the size distribution for inhabitants. All models except the HACD are close to each other for the most frequent municipal sizes. Small changes in the curvature result in more

[9] As a rough rule of thumb, Bunham & Anderson (2002) argue AIC differences between zero and two imply substantial empirical support for the argument that a model is considered competitive with the selected best model.

[10] The logarithm of the geometric mean 24,686 is 10.114, see Table 3.

substantial changes in the optimum scale. The intersection with the horizontal line y=1 (reflecting constant returns to scale) gives the optimum sizes of the local public administration over our sample. The optimum size according to SULF is 55,950 inhabitants, with a 95% confidence interval ranging between 45,420 and 76,410.

Figure 3: Cost elasticity of local public administration according to five models (2005-2014, in 2005 EUR)



In order to investigate developments over time, we disaggregate the data into annual samples and estimate the SULF model for each year separately, see Table 4. The cost elasticity at the mean is virtually stable over time, showing a slightly decreasing tendency, which points to somewhat higher scale economies in later years. The smaller annual sample lowers the significance level of the control variable coefficients, so that only the Island dummy remains statistically significant. Our analysis reveals that the optimum size has grown over the 2005-2014 period, as in 2005 the optimum size is 49,000 (with 95% confidence interval 39,910-68,420 inhabitants), while in 2014 the optimum size is 66,260 (with 95% confidence 52,690-102,310 inhabitants). Figure 4 shows this development of optimum scale over all our sample years for the ULF and SULF models (with SULF confidence bounds). Given the wide confidence bands, this increase is not statistically significant, however. The estimates in the graph for ULF show that both models have comparable estimates of the optimum size over time.

Table 4: Annual SCF estimation of economies of scale in local public administration for 2005, 2010 and 2014 using the SULF model.

	2005	2010	2014
Inhabitants (in log.)	0.311^{000}	-0.144^{000}	-1.088$^{*/000}$
	(0.67)	(-0.27)	(-1.99)
Inhabitants ^{2} (logs, mean dev.)	0.159**	0.194**	0.295***
	(2.88)	(3.07)	(4.18)
1/(ln inhabitants)	-49.331	-97.329	-192.588***
	(-1.10)	(-1.80)	(-3.45)
Demographic pressure	0.002	-0.001	-0.000
	(1.04)	(-0.56)	(-0.09)
Average home value (in 10,000 EUR)	0.002	0.002	0.003
	(0.57)	(1.10)	(1.04)
Surface land (in 10 sq. km)	-0.034	-0.039*	-0.037
	(-1.47)	(-2.00)	(-1.69)
G4 municipalities	-0.101	0.084	-0.263
	(-0.36)	(0.35)	(-0.96)
Wadden Islands	0.487***	0.738***	0.869***
	(5.06)	(3.69)	(9.01)
Constant	16.397	26.152*	44.994***
	(1.80)	(2.43)	(4.09)
σ_u (X-inefficiency)	0.145	0.117	0.128
σ_v (model and data errors)	0.205	0.228	0.233
λ	0.706	0.513	0.548
F-statistic	1,989.25***	1,697.50***	1,450.96***
Wald test	236.54***	142.53***	182.02***
Cost elasticity at mean	0.808	0.805	0.759
Scale economies at mean	0.192	0.195	0.241
Ln (mean)	9.965	10.128	10.210
Optimum size (in inhabitants)	49,000	60,100	66,300
Lower bound	39,900	47,000	54,000
Upper bound	68,400	96,300	102,300
Akaike's IC	45.34	70.23	94.39
Number of observations	427	425	396

Huber–White standard errors. For all other items, see the footnotes below Table 3

V. INEFFICIENCY

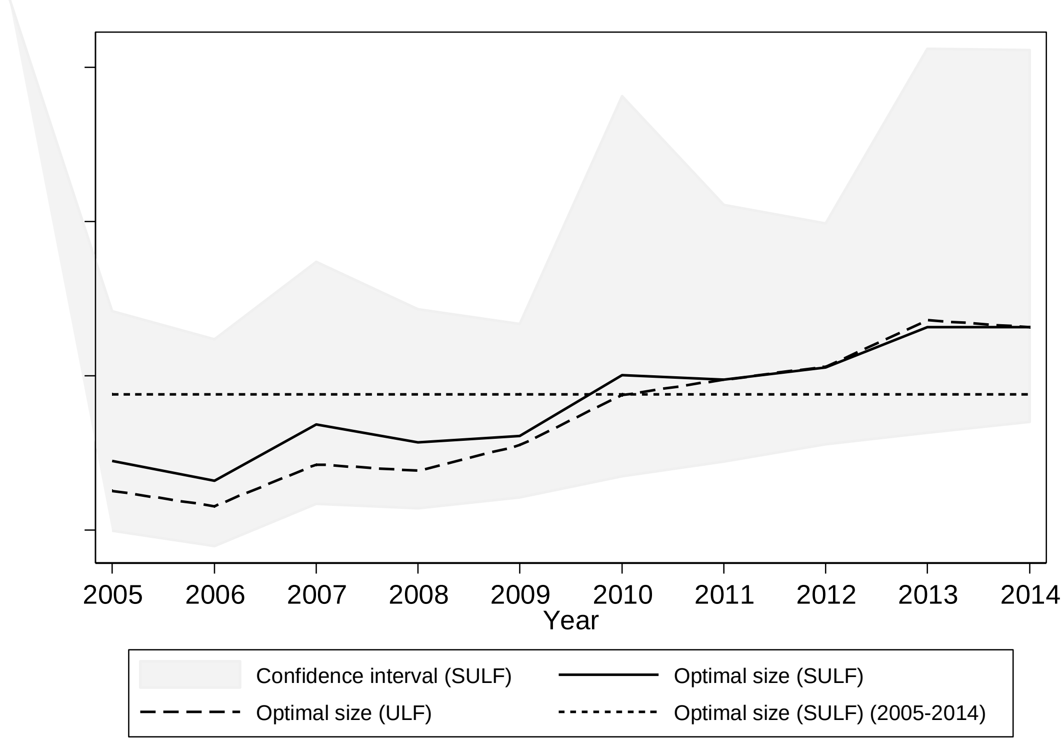
This section presents the managerial or X-inefficiency terms u_{it} from the SULF stochastic cost frontier as presented in Table 3 for the entire 2005-2014 sample. Similarly, we show the X-inefficiency terms u_i for the series of single-year observations from the model in Table 4. Table 5 gives the descriptive statistics for these two groups of predicted inefficiency terms. We find that around half of the municipality observations have a predicted inefficiency score of 10% or above, while 16% have an inefficiency of 20% or above. Both methods of estimation show a strong correlation (ρ = 0.893).

Table 5: Summary statistics for X-inefficiency (estimates from pooled model and annual regressions)

		Mean	SD	Min	Max	Observations
X-inefficiency (Pooled OLS model)	Overall	0.1147	0.061	0.0098	0.6915	N = 4,111
	Within		0.048			
X-inefficiency (Aggregated yearly regressions)	Overall	0.1162	0.076	0.0310	0.8580	N = 3,286
	Within		0.056			

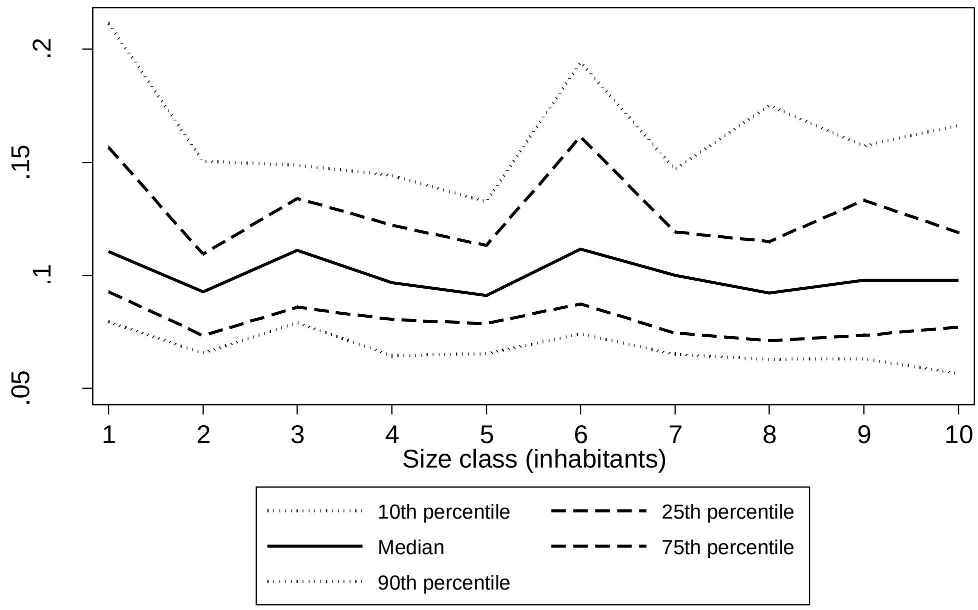
Note: SD is standard deviation. Within refers to within'' municipalities.

Figure 4: Annual estimates of the optimum size for local public administration according to the ULF and SULF models (2005-2014)



Such levels of X-inefficiency are not uncommon in these kinds of analyses. For instance, banks and insurers have average X-inefficiencies of 20% and 30% (Bikker 2010), typically much higher than unused scale economies at 5% and 10%. For LPA the reverse holds true: the levels of inefficiencies are around half that of scale economies. When interpreting these figures, we should take into account that our X-inefficiency estimates pick up possible managerial inefficiency, which may also stem from costs due to heterogeneity in terms of differences in service levels, *e.g.* costs related to administrating high schools and professional education in larger cities. Such differences in service levels occur in many spending areas and are consequently spread across a wide range of sizes. Figure 5 presents the distribution of X-inefficiencies across ten size classes. The pattern is quite similar along these size classes, yet shows that X-inefficiencies peak at the smallest class (between 900 and 9,800 inhabitants) and the sixth largest size class (between 24,000 and 27,500), largely due to one outlier municipality with special circumstances (with a mean X-inefficiency term of 0.50).

Figure 5: X-inefficiencies across ten classes of municipality size (2005-2014)



Finally, we can compare the inefficiency terms over time. Figure 6 presents rather constant inefficiency over time.

Figure 6. X-inefficiency term by year

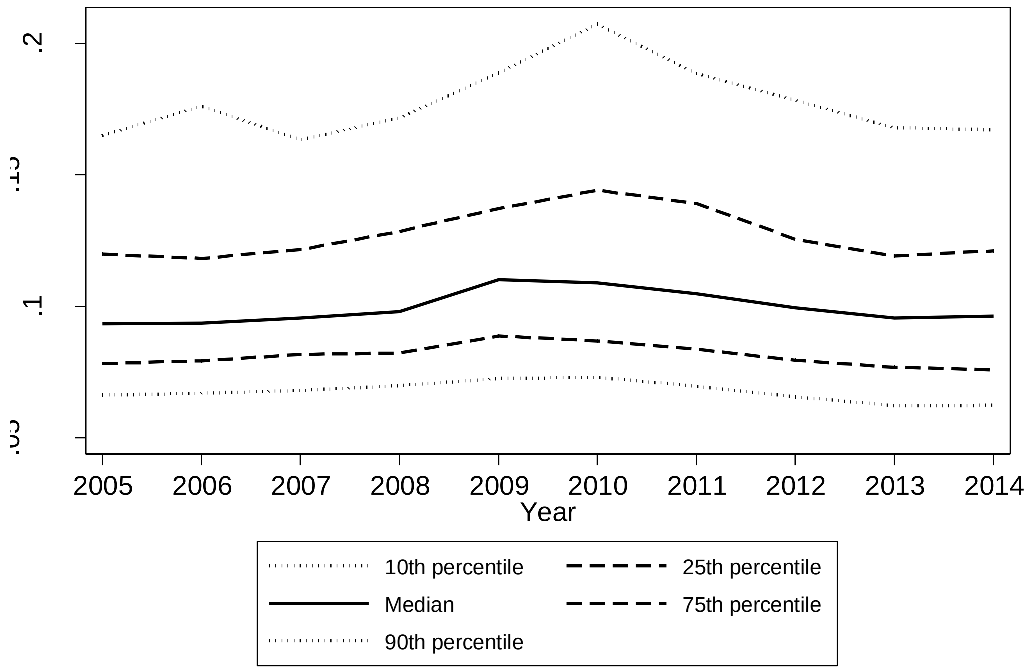
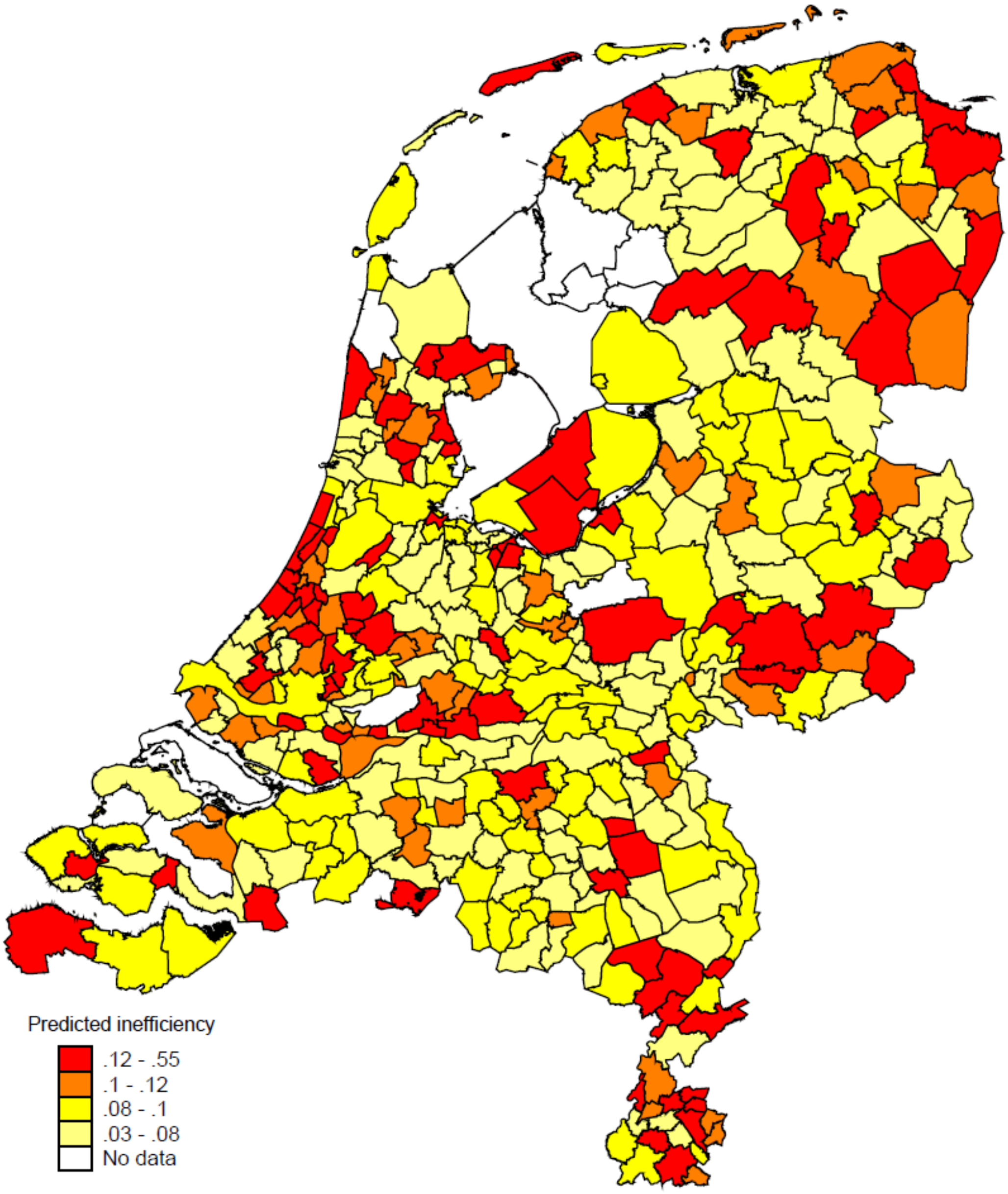


Figure 7 presents the geographical distribution of the efficiency term for 2013, based on pooled regression. We find evidence of regional clustering as the test of Moran's I reveals significant spatial autocorrelation in the inefficiency term at the 1-percent level (*I*=0.071). Results are insensitive to controlling for the number of neighboring municipalities, which would indicate complexity of coordination.

Figure 7: Geographical distribution of the efficiency terms for 2013



VI. ECONOMIES OF SCALE ESTIMATES IN OTHER COST CATEGORIES

We also applied our approach to the other cost categories listed in Table 1. As said, the results for these categories should be interpreted with caution in the cases where municipalities with limited spending autonomy act as mere intermediaries for national policies, or where the number of municipal tasks increases with size, so that unit costs increase rather than decrease with size. The former can especially be thought the case for costs related to social services, the latter especially for costs related to education or culture and recreation.

Table 6: Optimum scale for ten categories of municipal activities (2005-2014)

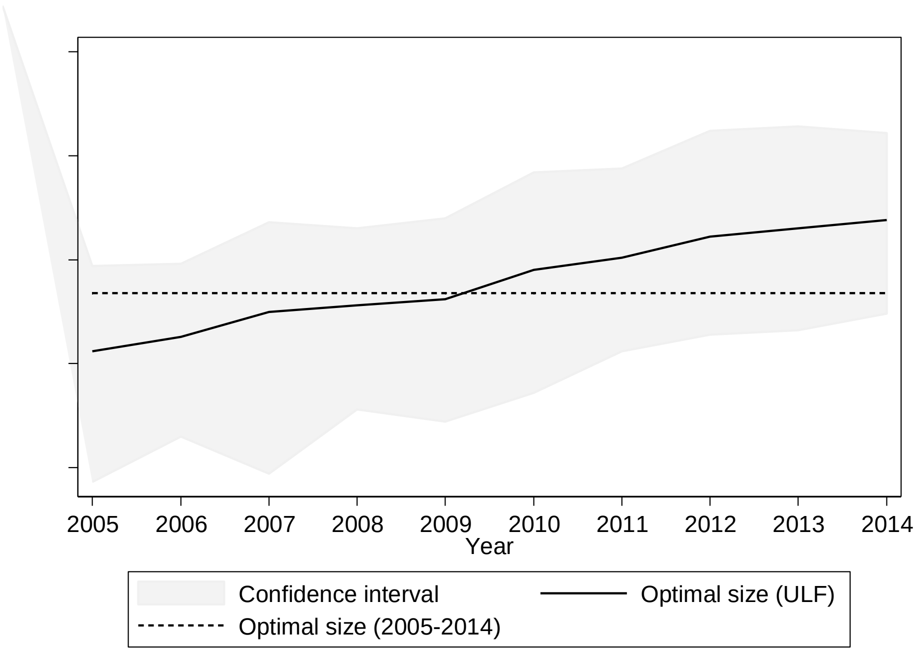
Cost category	Model [a]	Optimum scale (in inhabitants)
1. Local public administration (LPA)	ULF	55,900 (45,400-76,400)
	SULF (ΔAIC=0.6)	57,700 (47,900-77,400)
2. Public order and safety	ULF	18,400 (14,100-22,100)
3. Infrastructure	ULF	18,300 (12,000-26,700)
4. Economic affairs	ULF	-
5. Education	ULF	-
6. Culture and recreation	ULF	-
7. Social services	SULF	-
	ULF (ΔAIC=0.05)	-
8. Public health and environmental affairs	ULF	36,100 (27,400-47,800)
9. Spatial planning and housing	ULF	7,200 (-11,500) 372,600 (165,300-)[b]
10. Total expenses	SULF	8,400 (6,900-10,300)
	ULF (ΔAIC=1.6)	8,800 (6,400-11,000)

[a] If applicable, we present both the best fitting specification according to AIC and any arguably non-competing specification (*i.e.*, $\Delta AIC < 2$). [b] Two optimums at the tails of the size distribution so that the confidence intervals are in part out-of-sample.

Table 6 shows for four other cost categories as well as for total expenses a cost elasticity curve that crosses the y=1 axis, reflecting an optimum scale, similar to Figure 2. For the other cost categories, the cost elasticity curve remains either below the y=1 axis, so that unused scale economies exist across all sizes, or above that line so that diseconomies of scale exist for all sizes. Generally, the unrestricted Laurent function (in certain cases simplified) best fits the data.

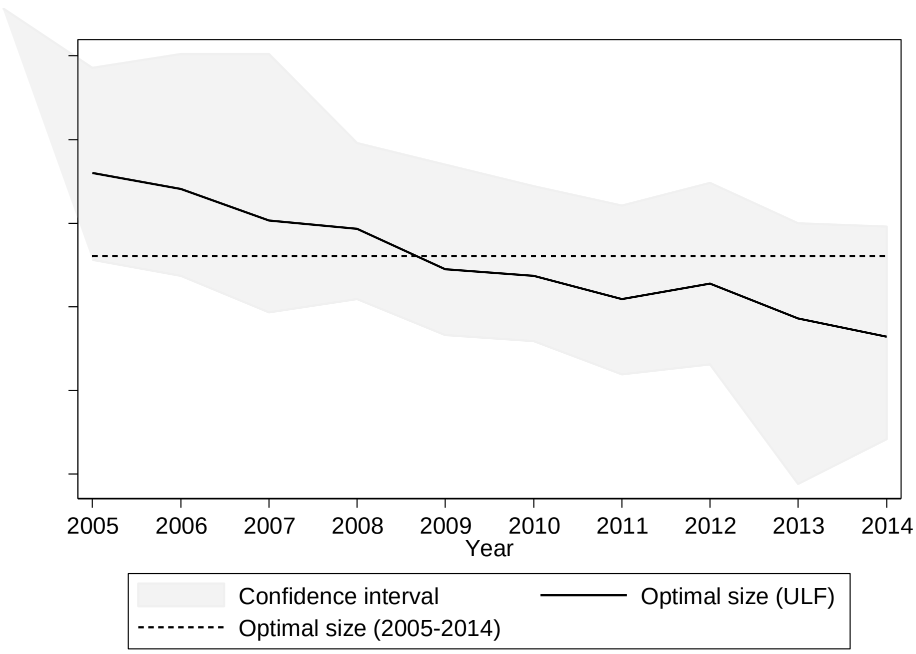
Figure 8 presents annual estimates of the optimum scale for costs related to public order and safety (category 2), based on the ULF model, with the corresponding 95% confidence bands. We see an optimum scale that continuously increases over time, though the change is not statistically significant. The average or full-sample optimum size is given by the horizontal line at 18,400 inhabitants (between 95% confidence bands of 14,100 and 22,100).

Figure 8: Annual estimates of the optimum scale for public order and safety over 2005-2014



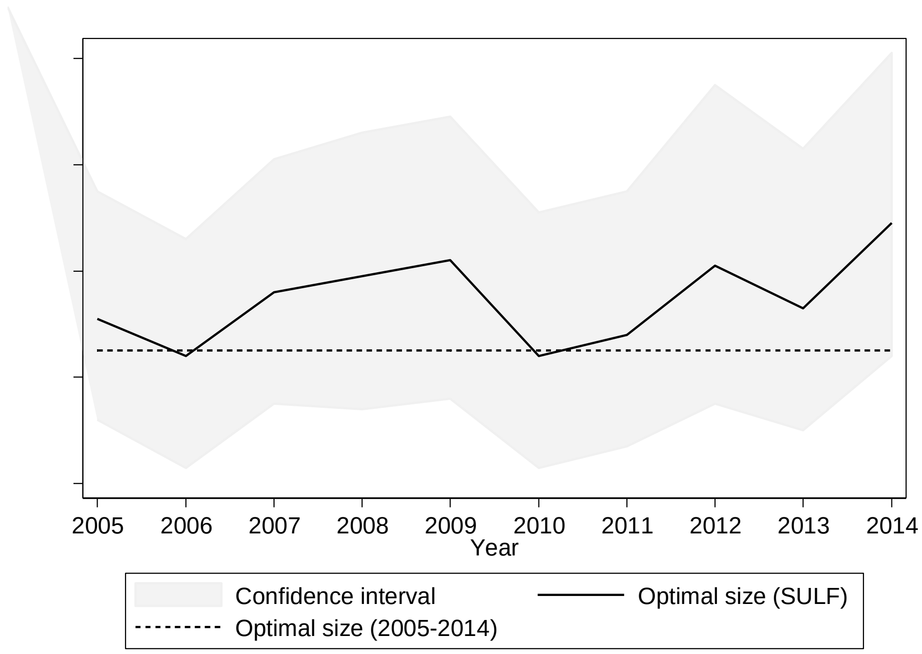
Based on the SULF model, Figure 9 shows annual estimates of the optimum scale for costs related to public health and environmental affairs (category 8, including all costs associated with municipal healthcare policies, costs associated with sewage and garbage collection and costs associated with maintaining air and soil quality), which gradually declines over time, albeit not significantly. Here, the average or full-sample optimum size is 31,600 inhabitants (between 95% confidence bands of 27,400 and 47,800).

Figure 9: Annual estimates of the optimum scale for public health and environmental affairs over 2005-2014



Finally, Figure 10 reflects the annual estimates of the optimum scale for total municipal expenditures for 2005-2014. We find that the optimum size roughly ranges between 6,000 and 14,000 inhabitants. As said, any meaningful interpretation of these figures is limited as we must realize that costs in certain cost categories increase with size due to additional tasks and service levels, arguably leading to a downward bias in a possible optimum scale relating to total expenditure.

Figure 10: Annual estimates of the optimum scale for total expenditure over 2005-2014



VII. CONCLUSIONS

In order to answer whether and how activities should be allocated to lower levels of government, the decentralization theorem claims in the absence of externalities and economies of scale that the provision of all public goods and services should be decentralized. The aim of this paper is to discover the possible existence of economies of scale in local public administration (LPA) in the Netherlands , a rather homogenous good which pertains to the core of municipal tasks.

We test for the existence of economies of scale using a number of non-linear functional forms in a stochastic cost frontier estimation over the 2005-2014 period and find that traditional functions, such as the Translog Cost Function, are not sufficiently flexible to describe the production of LPA properly. A simplified unrestricted Laurent function appears to provide the optimum model. This model indicates that scale economies for LPA exist at 20% around the mean – higher for smaller and lower for larger municipalities. The optimum size for municipalities hovers around 57,500 inhabitants. Disaggregated analyses on annual data show that the optimum size increased over our sample period to 66,260 inhabitants in 2014 from around 49,000 inhabitants in 2005. For LPA, this points at increasing importance of fixed costs over variable costs over time.

Beyond the number of inhabitants as a measure of size or output, LPA costs are to a large extent determined by the wealth of the inhabitants and the surface area. While the model can accommodate the LPA costs of the four largest municipalities, the specificity of its geographies cause the *Wadden* island municipalities in the Netherlands to be generally more expensive.

Applying stochastic cost frontier estimation allows us to differentiate between managerial or X-efficiency and noise. We find that roughly half of the municipality observations have a predicted inefficiency score of 10% of costs or above, while 16% have inefficiencies of 20% of costs or more. These inefficiencies appear spatially correlated, as we find evidence of geographic clustering. We should be cautious in interpreting these figures, as inefficiency indicates possible managerial inefficiency, but may also represent heterogeneity, *i.e*. costs related to differences in service levels across municipality sizes.

Similar interpretation problems occur in the analysis of other cost categories, where we again find (simplified) unrestricted Laurent functions best describe cost data. Economies of scale are found for costs of public order and safety (optimum size: 18,400), infrastructure (optimum size: 18,300), public health and environmental affairs (optimum size: 36,100) and for total expenditure (optimum size: 8,400). As we assume that in certain cost categories service levels change with size, we expect a potential downward bias in economies of scale, which only further disaggregation can overcome, but at the cost of losing generality of findings.

In the Netherlands where central government recently delegated a range of new tasks to local authorities, we expect increased complexity to lead to further rises in fixed costs for LPA relative to its variable cost components, moving the optimum scale further upwards. Whether the optimum scale will rise to 100,000 residents, as put forth by central government, remains to be seen.

References

Adanu, K., Hoehn, J. & Iglesias, E. (2009). *Cost function estimation in the water industry – functional forms and efficiency measures*, essay 3, Department of Agricultural Economics, presented at the 2009 USDA-CSREES National Water Conference; St. Louis, MO (February), and at the First Annual Graduate Academic Conference Michigan State University, East Lansing (March).

Akaike, H. (1974). A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, 19, 716–723.

Allers, M.A. & De Kam, C.A. (2010). Renovatie van het huis van Thorbecke (In Dutch). In: C.A. de Kam, J. Donders en A. Ros (Eds.). *Miljardendans in Den Haag: over bezuinigingen en belastingen*. Den Haag: Sdu Uitgevers.

Allers, M.A. & Geertsema, J.B. (2014). The effects of local government amalgamation on public spending and service levels. Evidence from 15 years of municipal boundary reform. *COELO*, working Paper 14019-EEF.

Bel, G., Fageda, X. & Mur, M. (2013) Why Do Municipalities Cooperate to Provide Local Public Services? An Empirical Analysis, *Local Government Studies*, 39(3), 435-454, DOI: 10.1080/03003930.2013.781024.

Belotti, F., S. Daidone., G. Ilardi, & Atella, V. (2012). Stochastic frontier analysis using Stata. *CEIS Tor Vergata Research Paper Series* 251.

Bikker, J.A., 2010, Measuring performance of banks: an assessment, *Journal of Applied Business and Economics* 11 (4), 141-159.

Bikker, J.A. (2013). Is there an optimal pension fund size? A scale-economy analysis of administrative and investment costs, *DNB Working Paper* No. 376.

Bish, R. L. (2001). Local government amalgamations: discredited nineteenth-century ideals alive in the twenty-first. *The Urban Papers*, 150, 1-35.

Blom-Hansen, J., Houlberg, K., & Serritzlew, S. (2014). Size, Democracy, and the Economic Costs of Running the Political System. *American Journal of Political Science*, 58(4), 790-803.

Burnham, K. & Anderson, D. (2002). *Model selection and multimodel inference a practical information-theoretic approach.* New York: Springer.

CBS (2014). Gemeentebegrotingen. CBS: STATLINE.

Christensen, L.R. & Greene, W.H. (1976). Economies of Scale in US Electric Power Generation, *Journal of Political Economy*, 84, 655-676.

Coelli, T.J., R. Prasada & Battese, G.E. (1998). *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers, Boston/Dordrecht/London.

De Vries, M.S. (2000). The Rise and Fall of Decentralization: A comparative analysis of arguments and practices in European Countries. *European Journal of Political Research*, 38, 193-224.

Dietsch, M. (1993). Economies of scale and scope in French commercial banking industry, *The Journal of Productivity Analysis*, 4, 35-50.

Drew, J., Kortt, M.A. & Dollery, B. (2014). Did the Big Stick Work? An Empirical Assessment of Scale Economies and the Queensland Forced Amalgamation Program, *Local Government Studies*, DOI: 10.1080/03003930.2013.874341

Edirisuriya, P. & O'Brien, G. (2001). Financial deregulation and economies of scale and scope: evidence from the major Australian banks, *Asia-Pacific Financial Markets* 8, 197–214.

Geys, B., F. Heinemann & Kalb, A. (2007). Local governments in the wake of demographic change: Efficiency and economies of scale in German municipalities, *ZEW Discussion Papers* 07-036, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Hirschman, A.O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

Koethenbuerger, M. (2008). Revisiting the "Decentralization Theorem"—On the role of externalities, *Journal of Urban Economics*, 64(1), July 2008, 116-122.

Meeusen, W. & Van den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review*, 18(2), 435-444.

Nauriyal, B.B. (1995). Measures of cost economies in Chilean banking: 1984-1991, *Revista de analisis economico*, 10, 71-99.

Oates, W.E. (1972) *Fiscal Federalism*. Harcourt, Brace and Jovanovich.

Oates, W.E. (1999). An essay on Fiscal Federalism. *Journal of Economic Literature*, 37(3), 1120-1149.

Panizza, U. (1999). On the determinants of fiscal centralization: Theory and evidence. *Journal of Public Economics*, 74(1), 97-139.

Reingewertz, Y. (2012). Do municipal amalgamations work? Evidence from municipalities in Israel. *Journal of Urban Economics*, 72(2), 240-251.

Shaffer, S., 1998, Functional forms and declining average costs, *Journal of Financial Services Research* 14, 91-115.

Solé-Ollé, A., & Bosch, N. (2005). On the relationship between authority size and the costs of providing local services: lessons for the design of intergovernmental transfers in Spain. *Public*

Finance Review, 33(3), 343-384.
Swank, J. (1996). How stable is the multiproduct translog cost function? Evidence from the Dutch banking industry, *Kredit und Kapital*, 29, 153-172.
Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *The Journal of Political Economy*, 64(5), 416-424.
Tocqueville, A. de (2000/1838). *Democracy in America*. Chicago: University of Chicago Press.
White, H. (1980). Using least squares to approximate unknown regression functions, *International Economic Review*, 21, 149-170.