Productivity of public capital in Finnish regions: What do the new panel econometric techniques tell us?*

Pellervo Hämäläinen[†] Tuomas Malinen[‡]

October 17, 2008
(Incomplete)

Abstract

This paper seeks to estimate the productivity effect of public capital in the Finnish regions in the period of 1975-2004. The focus is twofold. On the one hand we are interested on the size of the productivity effect of public capital in one of the Nordic countries. On the other hand we are looking for better ways than fixed effects OLS to estimate this effect in the production function framework. Attention is paid especially to panel unit root tests that take possible spatial dependence into account and panel data estimation techniques that solve some endogeneity problems.

JEL Codes: C33, H11, H54
Keywords: Public capital, productivity, panel unit root tests, panel estimation

1 Introduction

The productivity of public investments has been in the research agenda for 20 years. Since Aschauer's (1989) article number of studies using different approaches and data sets have been made.[1] Researchers using aggregate level data have generally ended up with the conclusion that the impact of public capital on the private sector is positive, but smaller than Aschauer's original estimate. However, there are still some open questions such as the size of the effect, for instance.

At the aggregate level studies, the two main motivations for this kind of productivity analyses have usually been a diminishing trend in public investments and a similar trend in overall economic growth. Some of the researchers have argued that these are cause and effect. However, the direction of the causality is not self-evident as it will be discussed later on.

*TBA

[†]Department of Economics, University of Turku. email: pellervo.hamalainen@utu.fi.
[‡]Department of Economics, University of Helsinki. email: tuomas.malinen@helsinki.fi. Wishes to thank Henri Nyberg for comments.

[1]For theoretical background look e.g. Arrow & Kurz (1970) and for older empirical paper e.g. Ratner (1983).

Due to some common problems in aggregate time series analyses or cross-country panels[2], the direction of the research on the productivity of public capital has recently changed toward regional analyses. The tightest restriction in regional analyses has been and still is the availability of suitable capital stock data. Therefore, only a few studies have been made in outside the US.

In the regional setup public capital is often seen as an instrument of regional policies. Central government can support poorer areas by financing large infrastructure projects, which hopefully will have also long term impacts on the local economic activity. One purpose of the research on the regional productivity studies is to find out, if public capital can be used effectively in the regional policy.

In addition to the division between country and regional level studies, previous research can be divided based on approach or level of aggregation of the data. The main approaches have been production functions, cost functions and VAR/VECM models. Some studies have used data over all sectors, while others have focused only to some specific sectors such as manufacturing or agriculture. There are also differences in the definition and scope of the public capital variable.

More complete view about the whole branch of research is given in literature reviews. Romp and de Haan (2007) have recently updated their summary of the extensive literature. Older reviews have been done, for example, by Gramlich (1994) and Strum, Kuper and deHaan (1996).

The purpose of this paper is to apply the production function approach to Finnish regional data and analyze how productive public capital is. In addition, we apply new methods of panel data econometrics to test and analyze the data. In that respect, we are taking into account some econometric issues that have been neglected thus far.

The history of regional productivity analyses focusing especially on the productivity of public capital and applying the production function approach could be thought to begin from Mera's (1973) research with the regions of Japan. Most of the regional studies in the 90's used the data from the US states (e.g. Munnell (1990), Garcia-Mila & McGuire (1992), Evans & Karras (1994), Holtz-Eakin (1994) and Garcia-Mila, McGuire & Porter (1996)). The results from these first regional analyses were contradictory. Some researchers found positive and some insignificant effect of public capital or infrastructure on the private sector productivity.[3]

More recently, regional capital stock data have been available increasingly also in European countries, which has lead to an increasing extent of regional studies. Stephan (2003) has studied a panel of 11 West German 'Bundesländer' and found that public capital is productive for manufacturing sector. Cadot, Röller and Stephan (2006) have studied 21 French regions in 1985-1992 using simultaneous equation model. They found that the statistically significant estimate for elasticity of private production with respect to infrastructure is 0,08. Moreno and López-Bazo (2007) have applied production function approach to 50 Spanish regions in 1965-1997. They found that local public capital is more important than transportation infrastructure. This result holds, even, if spill-over effects are taken into account. Interestingly, Moreno and López-Bazo also argue that the public capital is more productive in regions where the ratio of public capital stock to private capital stock is low. This result gives some guidance

[2]High multicollinearity, lack of co-integration and economically unreasonable size of elasticity estimates have been common in country-specific aggregate time series analyses. In cross-country panels problems arise, for example, from different definitions of data and various economic environments among included countries.

[3]In the 21th century US state data has been used mainly in cost function studies (e.g. Cohen & Morrison (2004).

for the effectiveness of public capital as a tool for regional policy. Spanish case has been studied also by Salinas-Jimenez (2004), who got positive effect only if spatial (or spill-over) effects were taken into account. Italian case has been studied for instance by Destefanis and Sena (2005), who ended up with the conclusion that public capital has positive effect at least in some regions.

Usually results from regional studies have ended up with lower estimates for the impact of public capital than studies using aggregate level data. A natural explanation to this is that the aggregate analysis takes externalities into account. Especially with transportation infrastructure, these externalities can be remarkable. Recently, many studies have been focusing on modeling and estimating these spill-over effects (see e.g. Pereira & Roca-Sagalés (2003))

Econometric methodology has varied a lot in previous studies although basic fixed effects OLS is the most commonly used especially in the older literature. Stephan (2003) estimates feasible generalized least squares (FGLS) both in levels and in first differences to correct problems with serial correlation, group wise heteroscedasticity and cross-sectional correlation, which were present in basic fixed effects OLS. However, he also presents results estimated by OLS with PCSEs. Also AR(1)-term, regional dummies (i.e. fixed effects) and linear time trend is added to equations estimated in levels.[4] Stephan argues that first differences gives the most reliable estimates.

Moreno and López-Bazo (2007) bring out many possible econometric problems, but handling of these problems remain quite superficial. Fixed effects OLS with time dummies has been used in estimation and the robustness of results has not been tested using some other estimator. Destefanis and Sena (2005) use t-bar test to examine possible unit roots. T-bar test allows individual constant and time trend to each region and, in addition, allows different autoregressive parameters and short run dynamics. Test results indicate that TFP would be trend stationary, but authors take it as I(1)-series and proceed to cointegration analyses and testing the possible long-run relationships. Destefanis and Sena (2005) apply also free disposal hull to get non-parametric estimates. Salinas-Jiminez (2004) uses fixed effects instrumental variable -estimator using lagged values as instruments.

For example, among these four studies results from unit root studies have been reported only in Destefanis and Sena (2005) although cointegration has been tested in three of them. Especially in some older studies, econometric specification does not get much attention. In the matter of fact, studies that mainly focus on the econometric problems are clearly in a minority. This is one motivation to the point of view of our article.

Kelejian and Robinson (1997) are one exception by focusing on empirical problems in regional panel data analyses. However, econometric methods have developed a lot especially in the field of panel data econometrics. According to our knowledge, there is no recently published articles applying the newest econometric methods to analyze the productivity of regional public capital. Kelejian and Robinson (1997) noted that if all econometric problems are ignored and production function is estimated with basic OLS, the results are in line with previous studies. However, the picture is quite different, when they correct, for example, autocorrelation, variable endogeneity and heteroscdasticity. They conclude, based on estimation results for the US data ,that public capital is productive only in the specifications that ignore econometric problems.[5]

[4]Capacity utilization rate is added as an input, which is critised, for example, in Duggal, Saltzman & Klein (1999).

[5]However, they point out that results hold only for marginal productivity of public capital.

In this paper, we are using, in addition to traditional panel unit root tests, also tests that allow for spatial dependence across regions and breaks in the tested series. To our knowledge, these test have not been used before in this branch of literature. We found that value added is a trend-stationary process in the period tested. This strongly contradicts the results obtained in previous studies, where value added series is usually found to be a I(1)-process, and cointegration analyses has therefore been a common part of previous empirical analyses. In addition, we are not satisfied with the basic fixed effects OLS-estimator, but use more advanced 2SLS, GMM and LIML -estimators. We also examine the exogeneity of each variable using Wu-Hausman test. The estimation section is, however, still incomplete and we cannot present any final conclusions at the moment.

One caution about the terminology is in order. Public capital and infrastructure capital should not be confused or used as synonyms. Public capital includes part of the infrastructure, but there are also a great many other items included. It should also be noticed that in the National Accounts part of the infrastructure capital is included in the private sector's accounts. This research focuses on public capital, which is quite broad concept compared to infrastructure capital. This choice is rationalized by the lack of proper data of the whole infrastructure capital in Finnish regions.

Focusing on public capital means that a part of the infrastructure capital is not included in the analysis of this study. Actually, nowadays in Finland only roads and railroads[6] are included in public capital and other parts of infrastructure (airports, energy and telecommunication networks, water supply and sewer systems, for instance) are included in the accounts of the private sector.[7]

2 Theoretical framework

We assume that every region has a following Cobb-Douglas type aggregate production function

$$Y_t = A \cdot L_t^{\alpha} \cdot K1_t^{1-\alpha}, \tag{1}$$

where Y is private output, L is private labor and $K1$ is private capital stock of the region. Parameter α measures the elasticity of private output with respect to private labor.

In addition, Hicks-neutral technological progress A of the region is assumed to depend on public capital stock $K2$ and a linear time trend t.

$$A = Ce^{\delta \cdot t} \cdot K2_t^{\gamma}, \tag{2}$$

where C is a constant describing initial level of productivity in the region. However, the rate of technological growth δ is assumed to be the same for all regions.

Combining equations (1) and (2) yields

$$Y_t = Ce^{\delta \cdot t} \cdot K2_t^{\gamma} \cdot L_t^{\alpha} \cdot K1_t^{1-\alpha}. \tag{3}$$

Here, it is assumed that $0 < \alpha < 1$ and $\gamma > 0$, which means that the inclusion of public capital yields increasing returns to scale over the economy as a whole. It would also

[6]However, these comprised about 40 % of the total fixed public capital in 2005

[7]I have not found any publication or research, where these movements between public and private sector would have been summarized. For example, airports moved to private sector's accounts in 1989 and over half of the public buildings switched from public to private in 1999.

be possible to induce constant returns to scale on the whole economy. Equation (3) can be easily estimated and all assumptions are testable. A similar specification is used for instance in Stephan (2003).

Now, if γ is positive, an increase in the public capital stock leads to increase in the private production.[8] This is the direct effect of public capital. Indirectly public capital can affect via marginal productivities of private capital and labor. Here, the focus is on the direct effect, because equation (3) cannot capture the indirect effect of public capital.

In estimation, also spatial effects are taken into account by adding spill-over variable in the estimated regression. This variable is constructed similarly as in Moreno and López-Bazo (2007) i.e. by calculating average public capital in neighboring regions[9].

3 Data and tests

3.1 Data

The data consists of yearly observations from 77 Finnish sub-regional units in the period of 1975-2004. Private sector's regional production is measured as value-added at factor prices. Private labor consists of number of workers in each region. Regional net capital stocks are taken from Salmela (2008), who has constructed those series using the current National Account standards. Variables are measured as constant prices at 2000 and the regional division corresponds to the situation in the year 2005[10]. The data is described more detailed in the data appendix.

3.2 Unit root tests

Some of the previous studies made on the topic have proceeded to cointegration analyses without testing the unit roots at all (e.g. Moreno & López-Bazo (2007))and many have relied on so called traditional panel unit root tests that assume independence of cross-sections (e.g. Destefanis & Sena (2005)). This is a very restrictive assumption when testing includes regions within a country. When the data consist on regions of a country, some of the different regional series are very likely to be correlated and/or cointegrated with each other. In a testing setup, where different sub-regions' series of value added are tested, majority of the different series are surely correlated and/or cointegrated with each other.

The traditional panel unit root tests are usually based on the following regression:

$$y_{it} = \rho_i y_{i,t-1} + \delta_i + \eta_i t + \theta_t + \varepsilon_{it}, \tag{4}$$

where δ_i are the individual constants, $\eta_i t$ are the individual time trends, and θ_t are the common time effects. Tests rely on the assumption that $E[\varepsilon_{it}\varepsilon_{js}] = 0 \; \forall \; t, s$ and $i \neq j$, which is required for the calculation of common time effects. Thus, if the different series are correlated and/or cointegrated the last assumption is violated. Despite of this restriction some tests are proved to be consistent and unbiased under cross-sectional correlation or even cointegration (Banerjee et. al 2005).

[8] Given the inputs of the private sector, private productivity could be used as well.

[9] Neighbor region is defined to have common boarder. However, Åland's three sub-regional units are assumed to be neighbors only to each other not to regions located in the coast, which they have common (sea) boarder.

[10] The names and the locations of sub-regional units can be found from appendix

The traditional panel unit root tests used in this study are based on the regression presented in equation (4). The null hypothesis is that $H_0 : \rho_i = 1 \ \forall \ i$. Tests have different assumptions about the heterogeneity of the unit root process. Levin, Lin and Chu (2002) (LLC), and Breitung tests assume that the unit root process is common to all cross-sections and Im, Pesaran, and Shin (2003) test (IPS) and Fisher type ADF and PP tests, presented by Maddala and Wu (1999), allow for a individual unit root processes. The inclusion of individual constants and time trends is optional, although Breitung's (2000) test requires that individual trends are included.

Karlsson and Löthgren (2000) have studied how does including some stationary series in the dataset effect on the results of panel unit root tests. They found that when the T dimension of a dataset is large (50 or over), small fraction of stationary series in the dataset results to high power and *vice versa*. Therefore, there is a risk that panels with large T would erroneously be modeled as stationary and that panels with small T be modelled as non-stationary. To eliminate the possible effect that some stationary series might have on our tests we run the individual ADF-tests and remove the sub-regional units whose value added series was stationary from the panel test.[11] First test includes individual constants and the second test includes individual constants and deterministic trends. Table 1 presents the results of five traditional panel unit root tests.

Table 1: Traditional panel unit root tests

variable	LLC	Breitung	IPS	ADF	PP
value added (c)	1.570	-	5.440	93.033	84.646
	(0.942)		(1.000)	(0.991)	(0.999)
value added (c&t)	4.459*	0.220	2.912*	171.91*	155.84*
	(<.0001)	(0.587)	(0.0018)	(0.0058)	(0.0476)
labor (c)	-5.414	-	-0.012	110.20	87.655
	(<.0001)		(0.495)	(0.870)	(0.998)
labor (c&t)	2.227	-0.319	2.133	86.022	50.213
	(0.987)	(0.3750)	(0.9835)	(0.9984)	(1.000)
private capital (c)	-7.361	-	-2.433	187.99	305.64
	(<.0001)		(0.008)	(0.0004)	(<.0001)
private capital (c&t)	-6.133*	0.468	0.419	116.06	134.98
	(<.0001)	(0.6579)	(0.6623)	(0.3192)	(0.3192)
public capital (c)	-12.890	-	-4.874	218.34	229.20
	(<.0001)		(<.0001)	(<.0001)	(<.0001)
public capital (c&t)	5.117	8.577	14.130	22.833	12.758
	(1.0000)	(1.0000)	(1.0000)	(1.0000)	(1.0000)

All variables are tested in logarithms. (c) denotes that individual constants and (c&t) that individual constants and trends have been included in the test. Probabilities of the test statistics are presented in parentheses. * denotes the rejection of unit root hypothesis at 5 percent or higher level.

One thing is clearly visible in the results of table 1. The result of the tests crucially depends on the inclusion of individual trends. If there is no deterministic trends included in the test, they may give flawed results. This is because the inclusion of individual deterministic trend does not alter the test. It just removes a trend in the series *if*

[11] All the tests have been done with Eviews 6. Lag lengths have been determined using Schwarts information criterion, spectral estimation has been conducted with Bartlett kernel and bandwidth has been selected using Newey-West method.

there is a trend. So according to results presented in table (1), individual trends should always be included in panel unit root tests. [So, we concentrate only on the results of those tests which include both the individual constants and trends.]

According to LLC, IPS, ADF and PP test the series of value added is stationary and according to Breitung's test it is non-stationary. All the other variables are non-stationary according to all tests except private capital which is stationary according to LLC. However, it is likely that most of the tested series are correlated across sub-regional units. This would violate the assumption of uncorrelated residuals among cross-sections, i.e. $E[\varepsilon_{it}\varepsilon_{js}] = 0 \; \forall \; t,s$ and $i \neq j$. Banerjee et al. (2005) have studied the effect of the violation of the assumption of no cross-unit cointegration to rejection frequencies of the null hypothesis. Their results show that in the presence of cross-unit cointegration ADF, PP, and IPS tests grossly overreject the null hypothesis of unit root and that LLC test keeps its nominal size and power with small T and relatively large N dimension of data. As all test accept the null hypothesis of unit root in labor and public capital series, they seem to be unit root processes whereas the series of valued added seems to be stationary. Results for private capital are inconclusive. Because the time dimension of our data is reasonably small (30 years), the bias presented by Karlsson and Löthgren (2000) should be towards non-stationary. This enforces the view that value added series is stationary.

However, the dependence of series may go further than the one studied by Banerjee et al. (2005) . Different regions of a country are likely to be *spatially dependent* as they (usually) lie in the same geographical area. This would violate the assumption of independence of error processes, but the different spatially dependent statistical units *need not to be* statistically correlated or integrated (Baltagi et al. 2007). Baltagi et al. (2007) have tested the performance of different panel unit root tests and found that tests allowing for cross-sectional dependence perform better than those assuming cross-sectional independence when statistical units are spatially dependent.

To account for spatial dependence in the series we conduct two tests. Baltagi et al. (2007) found that Phillips and Sul's (2003) (PS) and Pesaran's (2007) panel unit root tests allowing for cross-sectional dependence performs robustly in the presence of spatial dependence compared to many traditional panel unit root tests and to other panel unit root tests allowing for cross-sectional dependence. Phillips and Sul's and Pesaran's tests are based on a regression

$$y_{it} = \rho y_{i,t-1} + \eta_i t + \alpha_i + \delta_i \theta_t + \varepsilon_{it}, \tag{5}$$

where α_is are the individual constants, $\eta_i t$ are the individual time trends, θ_t is the common time effect whose coefficients δ_i, assumed to be non-stochastic, measure the impact of the common time effects of series i, $\varepsilon_{it} \sim i.i.d.N(0,\sigma^2)$ over t, and ε_{it} is independent of ε_{js} and θ_s for all $i \neq j$ and s,t. Cross-sectional dependence is allowed through the non-stochastic measure of the common time effects ($\delta_i\theta_t$).

We run two versions of each test. First we conduct tests including only individual constants and then a test that includes both individual constants and deterministic trends. Table 2 presents the results.[12]

Results of the PS test support the findings of traditional tests, i.e. the inclusion of individual trends in the test alter the results significantly. Results of Pesaran's test are more stable between the assumption of individual constants and individual constants and trends. According to it value added series would be a stationary $AR(p)$ process.

[12] Tests have been conducted with Gauss and Stata. The code for PS has been provided by the author of the test.

Many Nordic countries experienced a severe economic downturn in the beginning of 1990s. in Finland, one of the most important factors that contributed to this rapid downturn was financial crisis that stemmed from reckless lending by banks after credit restrictions were eased in the late 1980s. In the aftermath one of the major banks in Finland went bust and Finland was driven on the verge of bankruptcy. Cause of bursting property and equity bubbles and aggressive cutbacks in lending, the downturn was very rapid (GDP growth was +5,4% in 1989, +0,1 in 1990 and -6,2% in 1991 followed by 2 years of contraction). This *structural shift* is clearly visible in the Finnish GDP series. It also likely that in the span of 30 years almost all countries in the world have experienced a somewhat severe recession. That's why the possibility of structural breaks need to be taken into account in unit root tests.

Im, Lee and Teislau (2005) (ILT) have developed a unit root test that allows for structural shifts in the series. Im, Lee and Teislau's (2005) test assumes the following data generating process:

$$\begin{aligned} y_{it} &= z_{it} + x_{it} \\ z_{it} &= \gamma_{1i} + \gamma_{2i} t + \delta_i D_{it}, \\ x_{it} &= \phi_i x_{i,t-1} + \varepsilon_{it} \end{aligned} \tag{6}$$

where

$$D_{it} = \begin{cases} 0 & t \leq T_{B,i} \\ 1 & t \geq T_{B,i} + 1 \end{cases},$$

where $T_{B,i}$ is the time period of structural shift in the ith series and $\varepsilon_{it} \sim i.i.d.N(0, \sigma^2)$.

Once again, we conduct two tests. First ILT test allows for no breaks in the tested series and the second one allows for one common break in the series. Im, Lee and Teislau's test estimates the time of the break in the different series and then uses a common time dummy to control for the break. Table 3 gives the results of ILT panel unit root tests.[13]

In the combined individual time series of value added, there is a clear break point visible in the value added series in the year 1990. This is also the same year that the ILT test estimates as a break point. National labor series seems to have two break points: in

[13]Test have been done with Gauss. Estimation code have been provided by author's of the test.

Table 2: Panel unit root tests allowing cross-sectional dependence

variable	PS (c)	PS (c&t)	PE (c)	PE (c&t)
log(value added)	85.646	244.64*	-7.384*	-6.311*
	(0.998)	(<.0001)	(<.0001)	(<.0001)
log(labor)	45.937	202.31*	-1.836*	-6.063*
	(1.000)	(<.0001)	(0.033)	(<.0001)
log(private capital)	136.08	206.62*	5.456	5.456*
	(0.255)	(<.0001)	(1.000)	(1.000)
log(public capital)	154.37*	78.541	4.031	2.510
	(0.044)	(0.9997)	(1.000)	(0.994)

Probabilities of the test statistics are presented in parentheses. The values presented without brackets denotes the value of z-statistics. * denotes the rejection of unit root hypothesis at 5 percent or higher level.

the late 1970s and around 1990. ILT estimates the break point to be in the year 2000, which is clearly off. In the national private capital series there is a clear break visible in 1990-1991. ILT estimates that break point is in 1991.In the case of public capital, ILT test estimates the break point to the year 1992. However, there is no break visible in the national public capital series during that era, although there is some kink in the national series in the late 1990s. There also seems to be no common break point in the series in different sub-regions So, there may not be such common break in the series that would be identifiable to the ILT test, and the result obtained assuming no break in the series are probably more reliable. Results of the ILT tests may also have been affected by the correlation between cross-sectional units as it assumes independently distributed errors.

Traditional panel unit root tests and tests allowing for cross-sectional correlation and structural shifts give an unanimous conclusion: Value added series is trend-stationary in the Finnish data in the period tested. E.g. Destefanis and Sena (2005) have found this same result, but *assumed* that this was due to some errors and/or deficiencies in the panel unit root tests. However, many of arguments made in previous studies supporting the idea on non-stationary value added series are not supported by results. Thus the value added series may, at least in some cases, actually be trend-stationary.

This allows us to use traditional panel estimators on the Finnish data. Some previous studies have used standard panel estimators to estimate variables that are cointegrated. However, many standard panel estimators are *inconsistent* in panel cointegrated data. For example, GMM estimator is, by definition, inconsistent in panel cointegrated data. OLS and GLS are also inconsistent estimators if panel data includes cointegrating relations between the dependent and explanatory variables, a result that strongly contradicts the consistency of OLS in cointegrated time series data (Choi 2002, Kao & Chiang 2000).

4 Estimation

We estimate a model:

$$\begin{aligned} log(valueadded)_{it} = \alpha + \beta_1 log(labor)_{it} + \beta_2 log(publiccapital)_{it} \\ + \beta_3 log(privatecapital)_{it} + \beta_4 log(spill-over)_{it} \\ + dummy(1991-1993) + \varepsilon_{it}. \end{aligned} \tag{7}$$

Table 3: Panel unit root test allowing for structural breaks

variable	ILT (no break)	ILT (1 break)
log(value added)	-8.384*	-32.481*
	(<.0001)	(<.0001)
log(labor)	-3.613*	-28.864*
	(0.0387)	(<.0001)
log(private capital)	-0.145	-29.087*
	(0.934)	(<.0001)
log(public capital)	-0.115	-27.454*
	(0.948)	(<.0001)

Probabilities of the test statistics are presented in parentheses. The values presented without brackets denote *t*-statistics. * denotes the rejection of unit root hypothesis at 5 percent or higher level.

On ε_{it} we assume two different error structures:

$$\varepsilon_{it} = \mu_i + v_{it} \tag{8}$$

and

$$\varepsilon_{it} = \mu_i + \lambda_t + v_{it} \tag{9}$$

i.e. we assume that error process is either one-way (8) or two-way (9).

For one-way error process model we assert a dummy variable to account for the severe economic downturn in 1991-1993. In both two-way error models we assume that the unobservable time effect captures the effect of the downturn. Table 4 reports the result of estimation of equation 7. In addition, the equation is estimated only for the private sector and then with public capital, but without spill-over variable. In a way, three first equations in Table 4 serve as a benchmark cases estimated in the most of previous studies.

Table 4: Production function estimates

variable	One-way FE	One-way FE	One-way FE	Two-way, RE	Two-way, FE
constant	-3.323***	-3.619***	-4.017***	-4.061***	-4.327***
	(0.220)	(0.227)	(0.206)	(0.1349)	(0.2465)
log(labor)	0.7312***	0.7458***	0.7598***	0.7151***	0.8952***
	(0.0211)	(0.02058)	(0.0197)	(0.0210)	(0.0236)
log(private capital)	0.2788***	0.2319***	0.2058***	0.3415***	0.1838***
	(0.0244)	(0.0260)	(0.028155)	(0.0227)	(0.0260)
log(public capital)	-	0.0900***	0.0598**	0.0540***	0.0306
		(0.0175)	(0.0192)	(0.0183)	(0.0153)
log(spill-over)	-	-	0.1046***	0.0540***	0.0583*
			(0.0231)	(0.0147)	(0.0251)
dummy	-0.021***	-0.027***	-0.032***	-	-
(1991-1993)	(0.0053)	(0.0054)	(0.0054)		
trend	0.0260***	0.0252***	0.0243***	-	-
	(0.0006)	(0.0006)	(0.0007)		
joint	41.05	41.66	41.36	-	-
significance	(<0.0001)	(<0.0001)	(<0.0001)		
sub-regions	76	76	76	76	76
years	30	30	30	30	30
observations	2280	2280	2280	2280	2280

* = p<.05, ** = p<.01, *** = p<.001. Dependent variable: log(value added). Standard errors are presented in parentheses except in the test for joint significance where it denotes the *p*-value of rejecting H_o. Joint significance gives the value of F -statistics for the test of equal sub-regional dummy coefficients.

The value of the F -test for joint significance of sub-region dummy coefficient indicates that sub-regions dummies are jointly significant. Thus, normal cross-sectional estimation would suffer from omitted-variables bias. The difference in the estimates of fixed- and random-effect models implies that some of the explanatory variables may be correlated with the unobserved individual effect.

Basic production function estimates seems economically reasonable in sign and size. Estimates are also highly statistically significant also, when both capital stocks and spill-over variable are included. With one-way fixed effects, public capital gets estimate of 0.09, which is well in line with previous results from regional studies done for European countries. Quite logically, estimate is lower when spill-over variable is added. Only two-way fixed effects produce smaller and statistically insignificant estimate for public capital. However, spill-over variable stays significant. All in all, results indicate that public capital stock in neighborhood regions is more important than the capital stock in the region i.e. spill-over effects of public capital are important. So far, this looks quite similar as in Kelejian and Robinson (1997).

The general problem in production function estimation, and econometrics in general, is the possible endogeneity of regressors. We assert this question by assuming that the instrumental variables may be *predetermined*, but not *strictly exogenous*.[14] If this assumption holds, instrumental variables estimation can be used to obtain consistent and efficient estimates on otherwise endogenous explanatory variables given that instruments are efficient. Efficiency on instruments means that their correlation with instrumented variable is strong enough.

In recent years generalized method of moments has become one of the most used instrumental variables estimation methods in econometrics. The common reason for this has been the fact that, unlike GMM, many other instrumental variable estimators suffer from finite- or small-sample bias. However, when the time dimension and number of observations increase, GMM becomes biased toward OLS as the moments conditions grow over the number used by standard FE or RE estimators (Ziliak 1997). This can be controlled by controlling the number of instruments used in estimation. Ziliak (1997) studied the performance of several instrumental variable estimators in the setup where instruments were predetermined and found that as moment conditions increased, there are clear differences in the performance of GMM, 2SLS, and Forward-Filtering Keane and Runkle (1992) (FF) estimators. As the number of moment conditions increase the GMM parameter estimates become biased downwards relative to estimates of 2SLS and FF, although there was no differences in standard errors between estimators. However, this problem can be solved by restricting the number of moment conditions included in estimation.

As we have relatively large sample (over 2000 observations) we use also the 2SLS estimator. We use the of first differences as instruments for variables in levels. The problem with differenced instruments is that they may be "weak" instruments. Stock and Yogo (2002) have tested the effect that weak instruments may have on different instrumental variables estimators. They found that LIML is less biased than 2SLS when instruments are weak. First differenced instruments may also perform poorly when some of the explanatory variables are non-stationary variables (Blundell & Bond 1998, Bond et al. 2001). This problem can be dealt with System-GMM estimator. GMM estimation is also found to be consistent and asymptotically normal in the presence of spatial dependence (Conley 1999).

First we estimate equation 7 by 2SLS and LIML and test for overidentifying and underidentifying restrictions. Exogeneity of regressors will be tested later on using Wu-Hausman test . We test for autocorrelation using Arellano and Bond's (1991) test.[15]

Anderson LR-test does not reject the null of under-identification at 5% level. Ac-

[14]Strict exogeneity means that instrumental variables are independent from all future and past values of dependent variable. We thus assume that instruments are independent only on all *future* values.

[15]The drawback in Arellano and Bond's test is that it assumes uncorrelated error terms between cross-sections. As mentioned before this assumption is highly questionable in our setup. We are still forced to rely on these results, because, at least in our knowledge, there is no other testing method for autocorrelation in residuals in panel data.

Table 5: Production function estimates, instrumental variables tests

Dependent variable: log(value added)		
variable	2SLS	LIML
constant	-42.579****	-48.018****
	(4.7981)	(12.044)
log(labor)	0.5045****	0.6560*
	(0.0933)	(0.2825)
log(private capital)	0.5150****	0.1795
	(0.1334)	(0.5226)
log(public capital)	0.0596	0.2064
	(0.0536)	(0.2131)
log(spill-over)	0.0090	0.0391
	(0.0211)	(0.0605)
trend	0.0199****	0.0227****
	(0.0024)	(0.0060)
dummy (1991-1993)	-0.0588****	-0.0430
	(0.0159)	(0.0361)
Anderson LR-test	26.296	26.296
	(0.0692)	(0.0692)
Sargan-Hansen test	15.188	8.600
	(0.5109)	(0.9290)
Wu-Hausman test:	TBA	
log(labor)	-	
log(private capital)	-	More LIML estimations will be added later
log(public capital)	-	
log(spill-over)	-	
sub-regions	76	76
years	30	30
observations	2277	1824

* = p<.05, ** = p<.01, *** = p<.001, **** = p<.0001. Standard errors are robust to arbitrary heteroscedasticity. First, second, third, fourth and fifth lags of first differences are used as instruments for explanatory variables. Anderson LR-statistics measures underidentification of the estimated equation and the null hypothesis is that the matrix of reduced form has a rank smaller than the number of regressors. Sargan-Hansen test measures the overindentification of estimated equation and the correlation of instruments with error term. Under the null instruments are uncorrelated with the error term and the estimated equation is not overidentified. Wu-Hausman tests the exogeneity of regressors. The null hypothesis is that regressors are exogenous. Standard errors are presented in parentheses, except the case of Anderson, Sargan and Arellano-Bond test statistics where it presents the p-value of the rejection of H_0.

cording to Sargan-Hansen test instruments are not correlated with error terms and there is no overidentification present in the model. As 2SLS estimator will be more affected in the presence of possible under-identification, the results obtained using LIML can be viewed as more reliable

Table 6 presents preliminary results of estimation of equation 7 assuming one-way error process (8) using System-GMM. We use the lags from the third lag onwards for instruments in equations in differences as Arellano and Bond's (1991) autocorrelation test shows that first differences are autocorrelated in the second order. To control for too much moment conditions we restrict the number of instruments by allowing for one instrument for each variable and lag distance.[16]

Table 6: Production function estimates, System-GMM

Dependent variable: log(value added)		
variable		
constant	-50.904****	
	(3.4000)	
log(labor)	0.7426****	
	(0.0822)	
log(private capital)	0.0656	
	(0.1603)	
log(public capital)	0.2381*	
	(0.1006)	
log(spill-over)	0.0162	
	(0.0750)	
trend	0.0241****	
	(0.0017)	
dummy (1991-1993)	-0.0329****	
	(0.0069)	
endogenous	labor	Results from other combinations
	public capital	will be added later on..
	private capital	
	spill-over	
Arellano-Bond $AR(1)$	-5.07 (<.0001)	
Arellano-Bond $AR(2)$	-4.05 (<.0001)	
Arellano-Bond $AR(3)$	-0.70 (0.485)	
Arellano-Bond $AR(4)$	0.35 (0.725)	
Arellano-Bond $AR(5)$	1.07 (0.283)	
sub-regions	76	
years	30	
observations	2280 (total)	

* = p<.05, ** = p<.01, *** = p<.001, **** = p<.0001. Lags from third level and second lag of first difference onwards are used as instruments for explanatory variables. Standard errors are presented in parentheses except in the case of Arellano and Bond's autocorrelation test, where the p-value of rejecting null hypothesis is presented in parentheses.

Based on the results presented in Table 5, the statistically significant effects found with basic fixed effects OLS have disappeared. System-GMM (Table 6) results to significant estimate for public capital, but the size of the estimate is rather large. Moreover,

[16]Estimation is conducted with Stata.

private capital gets low and insignificant estimate. However, in Tables 5 and 6 we have assumed that all the explanatory variables are endogenous. Later on, we will change the assumption of endogenous variables to correspond the results from Wu-Hausman exogeneity test and perform more system-GMM and LIML estimations. Therefore, it is not possible to draw final conclusions before more careful estimations.

5 Conclusions

TBA

References

References will be completed later on...

ALVAREZ, JAVIER & ARELLANO, MANUEL (2003):

ARELLANO, MANUEL & (1991):

ARROW, KENNETH J. & KURZ, MORDECAI (1970): *Public Investment, the Rate of Return, and Optimal Fiscal Policy.* Baltimore: Johns Hopkins.

ASCHAUER, DAVID A. (1989a): Is Public Expenditure Productive? *Journal of Monetary Economics, 23*, 177–200.

BALTAGI, BADI H. & (2007):

BALTAGI, BADI H. & PINNOI, NAT (1995): Public Capital Stock and State Productivity Growth: Further Evidence from an Error Components Model. *Empirical Economics, 20*, 351–359.

BANERJEE, & (2005):

BREITUNG, & (2000):

CADOT, OLIVIER, RÖLLER, LARS-HENDRIK & STEPHAN, ANDREAS (2006): Contribution to productivity or pork barrel? The two faces of infrastructure investment. *Journal of Public Economics, 90*, 1133–1153.

CHOI, IN (2002): Instrumental variables estimation of a nearly nonstationary, heterogenous error component model. *Journal of Econometrics, 109*, 1-32.

CONLEY, TIMOTHY G. (1999): GMM estimation with cross sectional dependence. *Journal of Econometrics, 92*, 1-45.

COSTA, JOSE DA SILVA, ELLSON, RICHARD W. & MARTIN, RANDOLPH C. (1987): Public Capital, Regional Output and Development: Some Empirical Evidence. *Journal of Regional Science, 27(3)*, 419–437.

DESTEFANIS, SERGIO & SENA, VANIA (2005): Public Capital and Total Factor Productivity: New Evidence from the Italian Regions 1970-98. *Regional Studies, 39(5)*, 603–617.

DUGGAL, VIJAYA G., SALTZMAN CYNTHIA & KLEIN, LAWRENCE R. (1999): Infrastructure and Productivity: A Nonlinear Approach. *Journal of Econometrics, 92*, 47–74.

EUROSTAT (1996): *Euroopan kansantalouden tlinpitojärjestelmä – EKT 1995.* Luxemburg: Euroopan yhteisöjen virallisten julkaisujen toimisto.

EVANS, PAUL & KARRAS, GEORGIOS (1994): Are Government Activities Productive? Evidence from a Panel of U.S. States. *The Review of Economics and Statistics, 76(1)*, 1–11.

EVERAERT, GERDIE & HEYLEN, FREDDY (2001): Public Capital and Productivity Growth: Evidence for Belgium 1953–1996. *Economic Modelling, 18*, 97–116.

GARCIA-MILA, TERESA & MCGUIRE, THERESE J. (1992): The Contribtion of Publicly Provided Inputs to States' Economies. *Regional Science and Urban Economics, 22*, 229-241.

GARCIA-MILA, TERESA, MCGUIRE, THERESE J. & PORTER, ROBERT H. (1996): The Effect of Public Capital in State-Level Production Functions Reconsidered. *The Review of Economics and Statistics, 78(1)*, 177-180.

HOLTZ-EAKIN, DOUGLAS (1994): Public-Sector Capital and the Productivity Puzzle. *The Review of Economics and Statistics, 76(1)*, 12–21.

IM, KYUNG-SO, LEE, JUNSOO & TIESLAU, MARGIE (2005): Panel LM Unit-root Tests with Level Shifts. *Oxford Bulletin of Economics and Statistics, 67(3)*, 393–419.

KAO, CHIHWA & CHIANG, MIN-HSIEN (2000): On the estimation and inference of a cointegrated regression in panel data. *Advances in Econometrics, 15*, 179-222.

KARLSSON, & (2000):

KELEIJAN, HARRY & ROBINSON, DENNIS (1997): Infrastructure Productivity Estimation and its Underlying Econometric Specifications: A Sensitivity Analysis. *Papers in Regional Science, 76(1)*, 115–131.

LA FERRAR, ELENA & MARCELLINO, MASSIMILIANO (2005): TFP, costs and public infrastructure: an equivoical relationship. In Artis, M., Banerjee, A. & Marcellino, M. (ed.) (2005): *The Central and Eastern European Countries and the European Union*. Cambridge University Press.

LEVIN, & (2002):

LIGTHART, JENNY E. & SUAREZ, ROSA M. (2005): The Productivity of Public Capital: A Meta Analysis. URL: http://center.uvt.nl/staff/ligthart/research.html

MADDALA, & (1997):

MANKIW, GREGORY N., ROMER, DAVID & WEIL, DAVID N. (1992): A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics, 107(2)*, 407–437.

MEHROTRA, AARON & VÄLILÄ, TIMO (2006): Public Investment in Europe: Determinants in Perspective. *Fiscal Studies, 27(4)*, 443–471.

MERA, KOICHI (1973): Regional Production Functions and Social Overhead Capital: An Analysis of the Japanese Case. *Regional and Urban Economics, 3(2)*, 157–186.

MUNNELL, ALICIA H. (1990): How Does Public Infrastructure Affect Regional Economic Performance? In Alicia Munnell (ed.) Is There a Shortfall in Public Capital Investment?, *Conference Series No. 34*, 69–103, Federal Reserve Bank of Boston.

MORENO, ROSINA & LÓPEZ-BAZO, ENRIQUE (2007): Returns to Local and Transport Infrastructure under Regional Spillovers. *International Regional Science Review, 30(1)*, 47–71.

PESARAN, HASHEM M. (2007): A Simple Panel Unit Root Test in the Presence of Cross-section Dependence. *Journal of Applied Econometrics, 22*, 265–312.

PHILLIPS, PETER C.B. & SUL, DONGGYU(2003): Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence. *Econometrics Journal, 6*, 217–259.

RATNER, JONATHAN B. (1983): Government Capital and the Production Function for U.S. Private Output. *Economics Letters, 13*, 213–217, North-Holland.

ROMP, WARD & DE HAAN, JAKOB (2007): Public Capital and Economic Growth: A Critical Survey. *Perspektiven der Wirtschaftspolitik 2007, Vol 8 (Special Issue)*, 6–52.

SALINAS-JIMENEZ, DEL MAR, M. (2004): Public infrastructure and private productivity in the Spanish regions. *Journal of Policy Modeling, 26*, 47–64.

SALMELA, JAAKKO (2008): Teollisuuden alueelliset pääomakannat Suomessa 1975-2004. [Regional capital stocks for Finnish industry 1975-2004] Pro gradu -työ. Taloustieteiden tiedekunta, Jyväskylän yliopisto (Master Thesis, Faculty of Economics, University of Jyväskylä).

STATISTICS FINLAND (2005): *Municipalities and Regional Division Based on Municipalities 2005.* Handbooks 28.

STEPHAN, ANDREAS (2003): Assessing the contribution of public capital to private production: Evidence from German manufacturing sector. *International Review of Applied Economics, 17(4)*, 399–417.

STRUM, JAN-EGBERT, KUPER, GERARD H. & DE HAAN, JAKOB (1996): Modelling government investment and economic growth on a macro level: a review. *CCSO Working Paper Series No. 29*.

UIMONEN, SAKARI (2007): Suomen infrastruktuuripääoma: Tiet [Finnish infrastructure capital: Roads]*VATT-keskustelualoitteita, 436* (Discussion Papers, 436. Government Institute for Economic Research).

UIMONEN, SAKARI (2008): Suomen infrastruktuuripääoma: Rautatiet [Finnish infrastructure capital: Railroads] *VATT-keskustelualoitteita, 439* (Discussion Papers, 439. Government Institute for Economic Research).

UIMONEN, SAKARI & TUOVINEN, TARJA (2008): Tieinfrastruktuurin vaikutukset teollisuuden tuottavuuteen Suomessa. [The effects of road infrastructure on the productivity of Finnish industry] Julkaisematon käsikirjoitus (Unpublished manuscript).

ZILIAK, & (1997):

Data Appendix

General notices: The data consists of yearly observations from 77 Finnish sub-regional units (seutukunta in finnish)[17] in the period of 1975-2004.[18]

The division to private and public sector is done as follows. Sectors classified under S.11 (non-financial corporations and housing corporations), S.12 (financial and insurance corporations), S.14 (households) and S.15 (non-profit instutions serving households) are included in the private sector, while sector S.13 (general government) comprises the public sector.[19] All industries are included. Variables are measured as constant prices at 2000 and the regional division corresponds to the situation in the year 2005 (see appendix).

Regional output Y: Private sector's regional production is measured as value-added at factor prices. The data is taken from official statistics complied by Statistics Finland after wide revision of National Account statistics finalized in the spring 2006.

Labor L : Private labor consists of number of workers. The number of hours would be better variable, but it is not available at sub-regional level prior to 1995. The data is taken from official statistics complied by Statistics Finland after wide revision of National Account statistics finalized in the spring 2006.

Private capital $K1$: Private capital is measured as private net capital stock. Net capital stocks are taken from Salmela (2008) and they are constructed using the current National Account standards. Here, residential buildings are included and this choice is justified by the fact that in the aggregate level the correlation between non-residential and total capital stock series are 0,994 and 0,999 in the private and public sector, respectively. Omission of residential buildings would be troublesome. In addition, then also the impact of residential buildings on the value-added and employment should be removed, although the impact maybe quite small.

Public capital $K2$: Public capital is measured as public net capital stock, which includes both central and local governments capital stocks. Net capital stocks are taken from Salmela (2008) and they are constructed using the current National Account standards. Public capital stock is used instead of some infrastructure capital measurement. The use of the whole public capital stock can be justified by the restrictions in the availability of more proper data. Recently, also in Finland has been attempts to construct variables for infrastructure capital (see Uimonen 2007,2008). Until now, this is done only for roads and railroads. The division of capital stock to private and public is artificial, especially, when the productivity of infrastructure is examined. In the National Accounts infrastructure capital is divided in both sectors' accounts. Therefore, infrastructure variable that is independent from origin of funding would be preferable.

However, the use of the whole public capital stock most probably leads to underestimation of the effects of sole infrastructure and thus this choice does not induce upward

[17] Sub-regional units do not enter in the NUTS (Nomenclature des Unités Territoriales Statistiques) classification, which became effective in 2003 as a European Union's regulation. They are one step lower than the NUTS level 3, which would be regions in the Finnish case.

[18] However, Porvoo is excluded due to data problems. For instance, value added drops 75 % from 1985 to 1986.

[19] The sector classification codes are according to ESA95.

bias in the elasticity estimate. Actually, OECD's current recommendation considering the productivity studies is to use effective capital, which is a flow variable measured as a volume index of serviced provided by capital. These kind of variables are available in official statistics only in three countries (the United States, Canada and Australia).

Appendix: Sub regional units

Appendix: Sub regional units

Finnish sub-regional units in 2005. Source: Statistics Finland (2005).

011	Helsinki	093	Imatra	152	Vaasa
012	Lohja	101	Mikkeli	153	Sydösterbottens region
013	Tammisaari	102	Juva	154	Jakobstadsregionen
021	Turunmaa	103	Savonlinna	161	Kaustinen
022	Salo	105	Pieksämaäki	162	Kokkola
023	Turku	111	Ylä-Savo	171	Oulu
024	Vakka-Suomi	112	Kuopio	173	Oulunkaari
025	Loimaa	113	Koillis-Savo	174	Raahe
041	Rauma	114	Varkaus	175	Siikalatva
043	Pori	115	Sisä-Savo	176	Nivala-Haapajärvi
044	Pohjois-Satakunta	122	Joensuu	177	Ylivieska
051	Hämeenlinna	124	Keski-Karjala	178	Koillismaa
052	Riihimäki	125	Pielisen Karjala	181	Kehys-Kainuu
053	Forssa	131	Jyväskylä	182	Kajaani
061	Luoteis-Pirkanmaa	132	Joutsa	191	Rovaniemi
062	Kaakkois-Pirkanmaa	133	Keuruu	192	Kemi-Tornio
063	Etelä-Pirkanmaa	134	Jämsä	193	Torniolaakso
064	Tampere	135	Äänekoski	194	Itä-Lappi
068	Lounais-Pirkanmaa	138	Saarijärvi-Viitasaari	196	Tunturi-Lappi
069	Ylä-Pirkanmaa	141	Suupohja	197	Pohjois-Lappi
071	Lahti	142	Seinäjoki	201	Porvoo
072	Heinola	143	Eteläiset seinä-naapurit	202	Loviisa
081	Kouvola	144	Kuusiokunnat	211	Mariehamns stad
082	Kotka-Hamina	145	Härmänmaa	212	Ålands landsbygd
091	Lappeenranta	146	Järviseutu	213	Ålands skärgård
092	Länsi-Saimaa	151	Kyrönmaa		

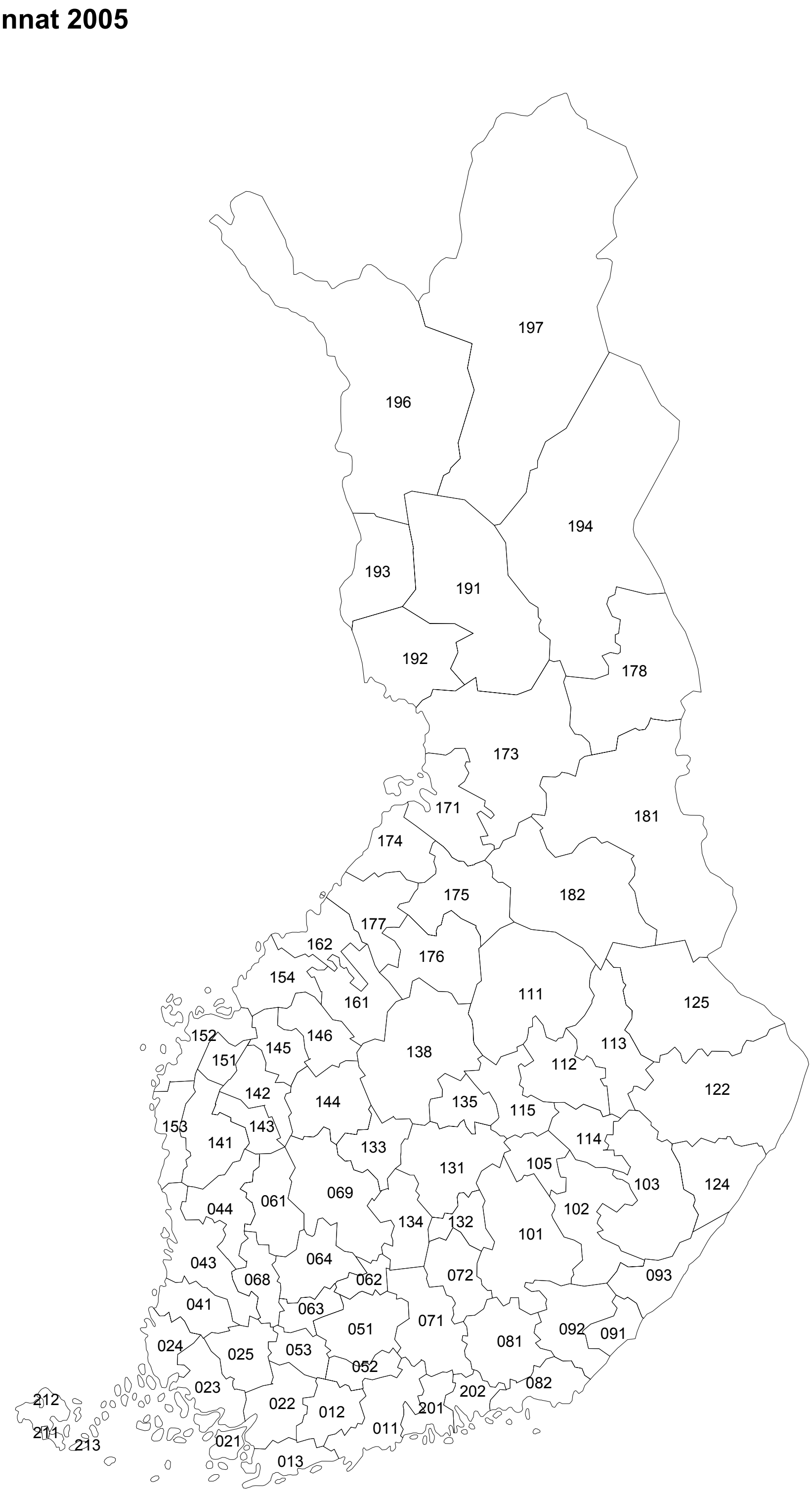

Figure 1: Map of Finnish sub-regional units in 2005. Source: Statistics Finland.