

Policy evaluation using multi-level panel designs:

Institutional capacity building among Polish regional zones administrations

by Camilla Jensen

Institute for Marketing and Management, University of Southern Denmark

Center for Social and Economic Research (CASE), Warsaw, Poland

First draft working paper submitted to the 21st Panel Data Conference, Central European University,
Budapest, Hungary, 29-30 June, 2015

Abstract

The research investigates the problem of designing policies to alleviate emerging employment differences across Poland's regions during transition to a market economic system. Poland has relied on alleviating problems of lacking local 'captains of regional development' by a design of zones administrations that cut across traditional regional boundaries and governance systems.

The questions addressed here are whether the individual zones administrations matter and if so, how exactly do they matter? Do they matter because of the tiered tax incentive structure offered or due to the services and institutional building efforts of the new alternative and footloose local zones administrations?

Overall it is concluded that zones administrations matter both because of the impact they have on the intensive (reeking in large projects) and extensive margins (attracting new investors) of employment creation. There are great differences in these capacities across the zones. The intensive margin is found to be declining over time across almost all the special economic zones. However, this is not necessarily a bad sign as the economic model behind the estimations suggest as it may be a sign of division of labour or specialization in the zones.

What seems to matter most in the dynamic perspective is whether the zones are able to increase their extensive margin over time, e.g. did any of the zones generate a snow balling type of effect in their ability to attract investors? Again are there differences across zones in their ability to attract investors. Only two zones were found capable of lifting their extensive margins over time.

Keywords: Panel data designs, multi-level data, policy evaluation, regional policy, institutions, evaluation of substitutions by design

JEL Codes: C19, C23, H25, L53, O25, O35, P25, P48, R58

1. Introduction

Evaluating policy experiments with panel data can be challenging for several reasons. The problem of counterfactuals in treatment effect models is the most commonly discussed problem in the literature on policy evaluation that uses econometrics and panel data designs as the primary tools of investigation (Heckman, 2001). Less attention has been paid to tackling and exploring heterogeneity in panels including heterogeneity in the responses of treated subjects to the policy experiment itself (Bondonio and Greenbaum, 2007).

The objective in this paper using the special economic zones in Poland as an example is to investigate and explore heterogeneity in responses to policy treatment. The counterfactual research question – e.g. of whether the policy had the intended impact on raising regional employment levels over and above similar areas that did not receive the same policy treatment (see e.g. Jensen and Winiarczyk, 2014), is complemented here with a top-down research design exclusively focusing on what happens in the areas affected by the policy and why there are such large differences in the responses of treated areas to the policy. Hence the research is grounded in the structural tradition with roots in economic theory (Heckman, 2001).

A short background introduction to the policy experiment is necessary to understand the context of the specific research questions.

In 1994 Poland designed a regional policy experiment to combat increasing regional unemployment in areas that were adversely affected by the economic transition. This started an unusual policy experiment from a regional perspective. A centrally controlled regional system (U form) existed in a vacuum of weakly equipped regional administrations to provide for the institutions that were needed for capitalist enterprises to work well.

In place was therefore put Special Economic Zones (SEZ) administrations to fill the institutional gap quicker than what could have been achieved from central or local hand. Unexpectedly these zones soon started to take on lives of their own- They spread beyond the borders of ordinary territorial administrations competing for investors as new administrations in their own right.

It is in particular this multi-level break in the panels of polish regions that the paper seeks to explore towards answering specific research questions about why or why not the policy experiment was successful at the level of individual zones administrations.

Exclusively focusing on the treated areas of the policy (being 334 localities, involving all of Poland's 44 regions, carried out by 14 zones administrations over the period 1997-2011), the research questions addressed are the following.

Do the zones administrations make a difference in this policy design? This is the first research question addressed. If it can be shown that the policy administrations matter (as they break traditional regional and local administrative boundaries and hence are not collinear in traditional regional fixed effects), it is

a strong indication that they have been partly successful in providing the institutions that the policy experiment is a promise of.

The second research question is, if the zones administrations matter, how exactly do they matter?

This is much more difficult to investigate using purely quantitative tools such as econometrics and panel data designs. The hypothesis is that the zones matter because they provide for institutions that businesses need to function and operate in a transition country setting. Hence they could be providing for a mechanism of substitution when new institutions are slow to emerge as long ago proposed by Gershenkron (1967).

Many similar policy experiments of using substitutions in the same or similar contexts have proven to be difficult to carry through according to the intend behind it or following the hands or logic of the master minds that designed them. Examples abound especially with failed designs of mass privatization (Birdsall and Nellis, 2003). Many well intended designs fell to pieces or led to adverse outcomes for the economic transition across Eastern Europe. Similar examples abound from developing countries (Parker and Kirkpatrick, 2005).

This is why the second research question is important, e.g. how can we investigate this using panel data, adding content to policy evaluations whereby more can be understood about the circumstances under which policies sometimes are successful, sometimes not and why it is so?

It is argued that this kind of problem poses completely different challenges for research with panel data in multi-level designs than merely a problem of counterfactuals and yes/no type of answers. A full design or research cycle could then include an initial framing with a counterfactual followed up by an in-depth investigation of factors behind relative policy success as the one conducted here and concluded with some in-depth case studies or check-list interviews about the remaining questions that it is not possible to fully answer with quantitative dataset readily available to the researcher.

2. Data sources and background.

The policy experiment of establishing special economic zones in Poland after the transition had started in 1994 is a good example of a policy intervention to study addressing the statistical issues of central concern in this paper. It has been at least hypothesised that the zones administrations in a certain sense of local political economy have become spatially wandering administrations that criss-cross traditional boundaries of division of labour among institution-making entities under the traditional regional set-ups of countries (e.g. federal vs. unitary systems). There are only a total of 14 zones administration, but the subzone areas they administer are manifold (typically 10 or more for each zone) and it is these subzones that cut across traditional regional administrative layers in Poland.

Exactly this aspect of the policy renders it amenable to conduct the policy evaluation in question. The main research question for this particular example is therefore whether those new administrative institutions matter for the particular policy outcomes in question. In other words does the institutional capacity building effort at the level of the zones have an impact on the relative success of the policy?

The general policy outcomes are treated in a number of papers and analysis (see e.g. Jensen and Winiarczyk, 2014). The main objective with the policy at least at the onset was to combat rising rural or regional unemployment in Poland because of the strong move towards urbanisation; where post-transition urbanisation was a natural geographical or spatial consequence of the economic transition itself (Jensen, 2004).

Hence the main policy outcome to be determined is how it affected local employment. That in itself is a relatively difficult thing to do as the special economic zones from a development economic viewpoint are really artificial areas e.g. they are not typical bottom-up economic areas where activity starts to develop because of resource availability, location or that simply some people live there in the first place. This problem is sought tackled by using a complementary dependent variable of employment that measures the share in the localities employment in the context of the greater region that it belongs to. This can to some extent solves for the counterfactual problem and also help to eradicate the influence of the business cycle on employment outcomes.

The data on localities and regions are taken from the Polish local databank published by the Central Statistical Office in Warsaw. The exact variables drawn from the local databank are shown in the upper part of Table 1. The data on SEZ administrations is taken from KPMG (2014). These variables are explained in the lower part of Table 1 and also summarized in Table 2.

The data variables are drawn from these sources in a compromise between data availability and the specific structural equation to be estimated. This is explained further in Section 3 of the paper.

INSERT TABLE 1 AND 2 HERE

3. Data and econometric strategy

Towards modelling employment at the regional level is adopted the regional labour demand model from Puga (1999):

$$L_i^T = (1 - \mu)n_i \frac{c_i}{w_i} - K_i r_w(w_i) \quad (1.1)$$

In this model capital is represented with intermediate goods, where μ is the share of intermediates in final output. A higher share of own value added ($1 - \mu$) increases labour demand locally. Similarly a higher number of firms n increases labor demand. The opportunity cost of sourcing from other firms (affected by producer prices and trade cost among other) c will lead the firm to internalize and produce more on its own. Oppositely will cost towards own value added or higher wages w decrease labor demand. Total labour demand of the region or city i may also be affected by the conditions in agriculture and services (however services are not considered separately in Puga, 1999). In Puga's model total labour demand is affected negatively by land endowment K devoted to agricultural production and the returns to agriculture r . See further below about some consideration made in the application to the data of the potential influences on industrial employment due to outside conditions for manufacturing workers in other sectors such as agriculture and services.

With regard to equilibrium considerations it is assumed that labour supply is perfectly elastic at the prevailing regional wage rate. This assumption makes it possible to identify the relationship between the explanatory factors and employment. The equation for labour demand (1.1) can now be estimated using logarithmic transformation:

$$\log L = \log(1 - \mu) + \log n + \log c - \log w - \log K - \log r \quad (1.2)$$

Some adaptation is necessary in view to data availability as introduced in Section 2. Employment (*Log Employment*), number of firms (*Log Firms* and *Log Foreign Firms*) and wages (*Wages*) are variables readily available from the statistical offices' local databank for Polish regions.

It is important to note that the relative size of firms is here modelled with the intercept α_0 where a decrease in μ in practice could be reflected by a variety of factors pertaining to industry, whether the firm is a start-up or part of a larger business group etc. Variations in the intercept across zones administrations could therefore reflect the economic characteristics of the zones. Changes in the intercept over time could be interpreted as an improvement in the intensive margin (or specialization if negatively signed) of employment with individual firms.

$$\log Employment_{it} = \alpha_0 + \beta_1 \log Firms_{it} + \beta_2 \log Foreign\ firms_{it} + \beta_3 Wages_{it} + \beta_4 Urban_i - \beta_5 \log Subsidies_{it} + \beta_6 \log EU\ Subsidies_{it} + \beta_7 EU\ Member_{it} + \epsilon_{it} \quad (1.3)$$

The extensive margin of employment is observed using the coefficient that measures the response in employment relative to the number of firms e.g. with β_1 and β_2 . Zones with a higher response in the number of firms (number of foreign firms) are zones that perform better in terms of attracting new investors (foreign investors that generate employment in particular e.g. the effect for foreign firms is the product of β_1 and β_2).

The producer prices and trade cost by region expected to affect c cannot be directly observed with the available data. However, some of the variables at the zones level could be used to control for this factor (such as in particular location and the relative size of land that the zones administrations govern).

To the base empirical model (equation 1.3) are added control variables for the influence of other policy instruments besides the main policy experiment of interest being the zones administrations. One control is for central government subsidies with the variable *log Subsidies*. Another control is for EU subsidies (*log EU Subsidies*) flowing from the EU's structural funds since 2006 onwards to each locality.

Land endowment is assumed to be constant over time and will in the empirical model be absorbed by the unobserved cluster effects (modelled as random) at the level of traditional territorial boundaries (e.g. localities at the lau2 level). However, the influence of the common agricultural policy as a fundamental structural break in terms of the rent structure of Polish agriculture is modelled by inserting a dummy for EU membership (*EU Member*).

Finally is there also made control for the influence of urbanization on employment with the dummy *Urban*, where it is assumed that urban localities relative to rural localities will generate more

employment opportunities as the availability of services allow households to free up more time for professional employment in the urban context.

The possible influence of the zones administrations on employment outcomes is modelled in a top-down design. How the effect of the zones administrations is to be captured is entirely an open and relatively difficult question and open to experiment and design.

The first model (investigating zones administrations using fixed effects) assumes that zones administrations increase employment due to unobserved factors that can be modelled as fixed effects. In the context of the structural model estimated and the underlying economic theory the indirect assumption behind this approach is that the zones administrations mainly work to increase the intensive margin e.g. by attracting firms that will employ more workers. Typically the tax incentives offered come with some size requirements in terms of minimum investment and minimum places of work (KPMG, 2009), therefore it is not unlikely that the zones policy will work through the intensive margin. Hence the influence is captured by including dummies or fixed effects for individual zones, e.g. α_0 is decomposed at the zones level. However, note that the basic panel data model is cast as random effects – e.g. these effects follow the ordinary territorial and time structure in the data.

The next model (investigating zones administrations using explanatory factors) assumes that the influence of zones administrations work through real variables that can be observed and related with the administrator's economic, political and management strategies. In practice explanatory factors are added to the model but at the level of the zones administrations. The potential variables to be added to Equation 1.3 are those listed in Table 1. One potential problem with including these variables added in an ad hoc manner is that they have little theoretical content in the context of the model and may therefore be difficult to interpret.

The final case-by-case model investigates in complement hereto whether all the structural parameter estimates in the fundamental model of labour demand (e.g. all the β) are affected by the zones administrations. For example is it possible that zones administrations work through both the intensive and extensive margins, the latter in particular is highly plausible as the zones administrations' foremost task is to attract new investors to the land under their governance. Importantly is it modelled with the case-by-case approach whether there are any changes at the zones level in these parameters over time. This is really one of the most salient questions to answer from the perspective of the designers behind the policy as the intention is eventually to dismantle the tax incentives structure and leave behind local administrators with viable economic areas that will sustain continuous employment.

4. Results

4.1. Investigating zones administrations using fixed effects

Table 3 compares results from using a random effect and between effect model on the aggregate data where zones are purely modelled as fixed effects. Note that the between effect model is not aggregated up by the ordinary panel dimension but rather by the upper level of the panel which is the zones administrations. This was already explained in Section 3. The between effect model serves as a robustness check on the variables (e.g. zones dummies or fixed effects in Table 3) cast at this level – e.g. at the level of the zones administrations - due to Wooldridge (2003).

INSERT TABLE 3 HERE

Two results come out of Table 3 with respect to the first research question whether the zones administrations matter in their own right (e.g. beyond affecting the ordinary explanatory variables such as attracting firms, investors from outside Poland, whether the area is rural or urban, their ability to bargain for subsidies etc.).

The zones administrations do matter since on average it is found that these administrations individually have a systematic and differential impact on employment. There are considerable differences in the performance of the zones administrations towards generating employment.

There are some differences across in particular dependent variables and also due to panel heterogeneity which renders different ranks in the model that uses employment as dependent variable. Specific results for the coefficient estimates must be read in the context of the specific model. For example, the firm level variable changes sign moving from the random effect to between effect model. This could in principle be due to different size strategies of individual zones when they target particular investors. It could also be due to misspecification of the model in view to underlying heterogeneity in individual panels.

Using the share in region as dependent variable has the advantage that it draws into the analysis in a very simple way a solution to the counterfactual problem. Here it is not employment generation *per se* that is evaluated with the share in region as dependent variable but rather employment generation in the specific regional context of the locality. The disadvantage of using this variable is that it suffers from non-stationarity. Only the case-by-case approach (see Table 5) involves strategies to address this particular problem since it is cast with a dynamic or time-varying dimension.

According to the results in Table 3 some zones administrations perform much better than others at the moment noting that Suwalki, Krakow, Katowice and Mielec persistently outperform across models other zones administrations in terms of generating absolute and relative employment. Oppositely do other zones such as Slupsk and Kamienna Gora perform poorly again both in absolute and relative terms across the two models.

This is only captured in Table 3 due to the varying ranks and significances of the zones administrations. The models in Table 3 are not useful towards understanding why and how the zones administrations

matter. For example, it could be the case that some zones administrators are better managers and hence more capable at identifying the right areas for industrial development. It could also be the case that some zones administrators have better political clout with central government. But it could also be the case that some zones are better at offering the type of institutions and economic governance that will attract and tie down investors to particular locations. To investigate this it is necessary to add data variables with real content that seek to capture some of the qualities among zones administrations. It is to these questions that the research turns next further exploring the multi-level character of the panel.

4.2. Investigating zones administrations using explanatory variables

Table 4 reports the results of including these explanatory variables. Here focus is on particular variables that are readily available and could be used as proxies for institutional capacity and other economic characteristics that relate with such strategies and management practices of individual zones administrations.

INSERT TABLE 4 HERE

The total area of each zone is adopted as an important control variable. This variable in particular would capture whether there are spillovers or increasing returns from running larger zones compared to smaller zones. In terms of hectares the zones size varies quite a lot from the smallest to the largest. Two of the variables adopted at the level of zones administrations are expected to be directly related with their institutional capacity building efforts. The most reliable comes from a survey conducted by KPMG among individual investors located with the zones where they are asked to rate the zone administration on a scale from 1 to 5. The main problem with this measure is that there is low variation in the average score obtained by each zone administration in part owing to the methodology of the study. The last variable measures the capital assets of the zones administrations relative to the area in hectares they have to administer and is therefore a direct measure of the administration's capital to land ratio. This measure is less valid as there need not be a direct correlation between the ability of the zones administration to accumulate capital and their provision of investor friendly institutions.

The results obtained as to why the zones matter are relatively weak when adopting these explanatory variables. The zone level variables are only significant in the between effect model. This result could owe to the fact the effect is absorbed by the individual localities effects in the random effect model or due to the necessary omission of the urban factor in the between effect model. The between effect model suggests that there may be some effect of the zones administrations due to governance quality. However, this effect is much weaker for the share in region results. The land variable in the latter is highly negative which may again be due to the relative ratio of urban to rural localities in each zones administration's portfolio which is not controlled for in the between effect model.

An important concern with the above results is in regard to the problem of inference with multi-level data when there are few observations available at the upper level. Even though the analysis is cast in a panel setting that has the advantage of producing relatively many observations it does hide over the fact that the real number of observations at the upper level here is only 14. This makes it difficult to use statistical inference relying on this approach in the present context. However, one econometric strategy

to alleviate this problem is to complement the previous two levels of analysis with another at the level of the zones administrations which is here called a case-by-case approach. Perhaps more can be discovered about the zones administrations at this level and in particular exploring the dynamic or time dimension in the dataset.

4.3. Investigating zones administrations case-by-case

In Table 5 (here only reporting results for employment as dependent variable) results are reported when estimating the original model individually for each zone and including a time trend that is interacted with all the explanatory variables. Note that the wage variable was dropped in these estimations due to its lower availability for the early period. This renders it possible to increase the number of observations for the more detailed analysis with 6 years back to 1995.

This approach appears to be the most fruitful with the present dataset and research questions. Zones come out as highly heterogeneous both in terms of their intensive and extensive margins but perhaps less so in the trends in these over time. Many zones follow a particular pattern of a decreasing trend in the intensive margin over time (Suwalki and Krakow are the only exceptions to this pattern). This could be interpreted so that individual firms loose places of work or generate less employment over time. However, it could also be seen as a sign of specialization at the zone level as more and more firms are added to the economic area.

With respect to the extensive margin few zones are able to significantly improve their extensive margin over time whereas most zones perform quite stable. One zone is an outlier and able to increase its extensive margin over time (Kostrzyn Slubice). Two zones (Katowice and Walbrzych) loose on their extensive margin up to EU accession in 2003-2004.

INSERT TABLE 5 HERE

Most of the time trends are not significant (e.g. the estimates are small and insignificant) suggesting that most zones do less well from a dynamic perspective. However, a few surprises did arise from this exercise. Notably the results revealed the potential sensitivity of the zones policy's employment generation ability to political developments surrounding this policy.

Politically there have been significant uncertainties surrounding the future fate of the zones three times during the last 20 years. The first time was in 2003-04 during the period where Poland negotiated the final conditions of its accession to the European Union in 2004. At this time there were significant rumors that the zones policy would be phased out. At that time it was agreed between the Polish government and the Commission that the zones policy experiment would be phased out by 2016. However, owing to subsequent developments across Europe and arguably a relative increase in the bargaining power of the Polish government it was decided finally in 2014 to relay the phasing out of the zones to 2027. Besides this there may also have been side effects on manufacturing employment in general due to the change in rent structures across agricultural and manufacturing when Poland also simultaneously joined the Common Agricultural Policy in 2004. Finally, Poland did go through a

recession in the period 2002-2003 which may also in part explain the general decline in employment in the same period.

9 out of the 14 zones analyzed exhibited surprisingly very high losses in employment levels at the time of EU accession and it is concluded that this is more likely due to the rumor effect that EU accession would lead to a phasing out of the zones policy in a near future during the accession negotiations in the years 9 and 10 of the policy (2003 and 2004). This effect according to the results reported in Table 5 came through mostly as a decline in the intensive margin and only as a decline in the extensive margin for a few zones.

From the detailed case-by-case results it would also appear that ordinary policies of central government subsidies to combat employment losses were reinstated in this period hence most of the zones administrations that experienced such adverse effects of the accession proved capable of bargaining for subsidies as an alternative shelter to employment losses due to the declining effect of the tax incentives in terms of securing places of employment in this period.

5. Preliminary conclusion and perspectives

The research addressed the problem of designing policies to alleviate emerging employment differences across Poland's regions during transition to a market and capitalist economic system. Hence the research sets out to investigate whether zones administrations matter and if so, how exactly they matter through a policy design that ties down investors due to a tiered tax incentive structure but also due to the services and institutional building efforts of the new alternative local administrators embedded with the footloose zones administrations. Poland has in this way relied on alleviating the problem of lacking local 'captains of regional development' by a design of zones administrations that cut across traditional regional boundaries and governance systems.

As is shown and discussed in the remainder of the paper a fixed effect model is less than satisfactory as it cannot answer both research questions (e.g. in particular the second which is why the zones administrations matter). A multi-level approach e.g. by adding explanatory power to the model through the variables gathered at the level of zones administrations also has severe drawbacks in the present case where the full population at the upper level is the 14 zones administrations. The number of observations available at this level renders significant drawbacks for statistical inference (Donald and Lang, 2007, Wooldridge, 2003).

Hence the last approach is the preferred where the zones administrations are seen to matter via several of the parameters in the original structural model. This is called a case-by-case approach even though the results can be obtained estimating one long or a set of zone-by-zone equations. The economic content of the model (e.g. making possible to see and interpret the distinction between the impact that the zones administrations have on the intensive and extensive margins of employment creation e.g. with the expansion of existing firms and the entry of new firms) is paramount towards leading us to these particular results and conclusions for the policy experiment.

Overall it is concluded that zones administrations matter both because of the impact they have on the intensive and extensive margins of employment creation. There are great differences in these capacities across the zones. The intensive margin is found to be declining over time across almost all the special economic zones. However, this is not necessarily a bad sign as the economic model behind the estimations suggest. In fact these results should be read so that in most periods the negative trend in the intensive margin may be due to the specialization impact that adding more and more investors to the zones have on the decision of firms to internalize and produce their own value added or source it from other firms in the zones. In particular periods however the decline in the intensive margin is cyclical and possibly also due to the rumor effect of the policy potentially under threat to be phased out at the time of Poland's accession to the EU.

What seems to matter more in the dynamic perspective is whether the zones are able to increase their extensive margin over time, e.g. did any of the zones generate a snow balling type of effect in their ability to attract investors? Again are there differences across zones in their ability to attract investors. Only two zones were found capable of lifting their extensive margins over time.

A very similar substitution as the one investigated here was used in Poland's mass privatization program where management companies were set in to substitute for the lack of 'captains of industry' at the time the program was finally launched in 1997. However, available evaluations question whether this was successful or whether in the end the management companies mainly worked to enrich themselves and poorly fulfilled their original mandate of turning around ailing socialist enterprises and making them viable capitalist firms. At the same time it is clear that the mass privatization program was much more successful at privatizing enterprises than any other approach that was used towards solving the ownership and management problem of these large Polish enterprises (Błaszczuk, 2001). Hence the answer here is probably that the management companies did exactly both and that is how capitalism works since no outcomes are achieved without designs that ensure significant benefits for those that carry out the actual work.

In parallel to secure a better grounding of the zones policy for future regional employment figures, policy makers need perhaps to rethink how they can create more long-term incentives among zones administrators to create the type of institutions that will make investors commit to the land in more permanent ways. Hence a successful design of substitutions requires not only getting incentives right but getting incentives and the concerted direction of institution capacity building efforts towards sustainable societal goals resulting from these incentives right. It is easy to think about but extremely hard to carry out in practice as many failed experiments in this area have demonstrated.

REFERENCES

- Birdsall, N., & Nellis, J. (2003). Winners and Losers: Assessing the distributional impact of privatization. *World Development*, 31(10), 1617-1633.
- Blaszczyk, B. (2001). Secondary privatization in Poland (part II): Evolution of ownership structure and performance in national investment funds and their portfolio companies. *CASE Network Reports*, (48).
- Bondonio, D., & Greenbaum, R. T. (2007). Do local tax incentives affect economic growth? What mean impacts miss in the analysis of enterprise zone policies. *Regional Science and Urban Economics*, 37(1), 121-136.
- Donald, S. G., & Lang, K. (2007). Inference with difference-in-differences and other panel data. *The review of Economics and Statistics*, 89(2), 221-233.
- Gerschenkron, A. (1967). The discipline and I. *The Journal of Economic History*, 27(04), 443-459.
- Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of political Economy*, 109(4), 673-748.
- Jensen, C. (2004). Localized spillovers in the Polish food industry: the role of FDI in the development process?. *Regional Studies*, 38(5), 533-548.
- Jensen, C., & Winiarczyk, M. (2014). Special Economic Zones—20 Years Later. *CASE Research Paper*, (467).
- Parker, D., & Kirkpatrick, C. (2005). Privatisation in developing countries: A review of the evidence and the policy lessons. *Journal of Development Studies*, 41(4), 513-541.
- Puga, D. (1999). The rise and fall of regional inequalities. *European economic review*, 43(2), 303-334.
- Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review*, 133-138.

Table 1: The study variables (descriptive statistics are reported in Appendix Table A1)

Variables - localities	Explanation
Employment _{it}	Number of employed persons (lau2).
Share in region _{it}	Employment at the lau2 level divided with total regional employment at nuts3 level.
All firms _{it}	Number of firms (lau2).
Foreign firms _{it}	Number of foreign owned firms (lau2).
Wages _{it}	Average wage rate index, Poland=100 (nuts3).
Nuts51 _{it}	A dummy that takes the value of 1 when the area is classified as urban (lau2).
Subsidies _{it}	Total subsidies paid out from the central government budget (lau2).
EU subsidies _{it}	Total subsidies paid out from the EU structural funds budget (lau2).
EU member _t	A dummy that take the value of 1 when Poland is a full member of the EU, e.g. from 2004 onwards.
Code (i)	Territorial code of the observation at the lau2 level (cross section panel data identifier).
Year (t)	Year of the observation (time series panel data identifier).
Variables - SEZ administrations	
Land in hectares _a	Total area of land governed by the zones administration.
Capital assets _a	Total capital assets accumulated by the zones administration.
Capital-land ratio _a	Total capital assets to land area.
Governance quality _a	The quality of the zones management as evaluated by investors in each zone on a scale from 1 to 5.
Project size _a	Average project size, where small zones have an average project size below 100 jobs, medium zones have an average project size above 100 but below 175 jobs and finally large zones have an average project size above 175.
No. of subzones _a	The total number of subzones run by each administration.
Tax rebate tier _a	Tax rebates are rated in terms of three tiers, where administrations operating out of tier 1 regions can give the least rebates and administrations operating out of tier 3 regions can give the highest rebates.

Data sources: Local Data Bank, Central Statistical Office of Poland (http://stat.gov.pl/bdlen/app/strona.html?p_name=indeks), KPMG (2014): '20 years of Special Economic Zones in Poland – A Guide to SEZs', KMPG, Poland.

Table 2: Economic characteristics of the zones administrations

	KMG Kamienna Gora	KTW Katowice	KZS Kostrzyn Slubice	KRW Krakow	LGA Legnica	LDZ Lodz	MLC Mielec
Land in hectares	413	2347	1747	708	1212	1302	1363
Capital assets, mio PLN	11.5	9.2	27	17.6	31.6	24.9	5040
Governance quality	4.1	4.46	4.41	3.73	4.38	4.18	4.29
Average project size in jobs created	4864/59=82 small	52575/250=210 large	22630/140=162 medium	16779/101=166 medium	10237/65=157 medium	28882/173=167 medium	23562/183=129 medium
Location	West	South	West	South	West	Center	East
No. of subzones	18	35	26	14	13	51	20
Tax rebate tier	1	2	2	2	2	2	3
	PMR Pomeranian	SLP Slupsk	STW Starachowice	SWK Suwalki	TBZ Tarnobrzeg	WBZ Walbrzych	WMZ Warmia Mazury
Land in hectares	1863	817	644	376	1677	2649	1615
Capital assets	255	28.8	16.7	20	5040	224.7	32.2
Governance quality	4.24	3.58	4.31	3.92	4.48	4.27	3.78
Average project size in jobs created	15394/111=139 medium	3515/58=61 small	6380/71=90 small	5425/71=76 small	27230/151=180 large	36164/204=177 large	13063/72=181 large
Location	North	North	East	East	East	West	North
No. of subzones	22	12	16	8	36	41	27
Tax rebate tier	2	2	2	3	3	1	3

Source: KPMG (2014): '20 years of Special Economic Zones in Poland – A Guide to SEZs', KPMG, Poland.

Table 3 – Modelling zones administrations as fixed effects

	Log Employment – RE		Log Employment – BE		Share in region – RE		Share in region – BE	
	<i>Coefficient</i>	<i>s.e.</i>	<i>Coefficient</i>	<i>s.e.</i>	<i>Coefficient</i>	<i>s.e.</i>	<i>Coefficient</i>	<i>s.e.</i>
Log All firms	.765***	.047	-.331	.679	2.612***	.230	-.885*	.477
Log Foreign firms	.200***	.038	1.834**	.795	.780***	.099	1.318**	.541
Wages (indexed)	-.005**	.002	-.007	.006	.007	.008	.007	.019
Urban dummy	.228***	.067	NA	NA	10.734***	1.160	NA	NA
Log Subsidies	.074***	.005	-.359**	.125	.055	.091	-.237**	.080
Log EU subsidies	.009***	.003	.0344***	.004	-.008***	.003	.021***	.003
EU member dummy	-.252***	.052	-.637***	.089	-.281***	.049	-.402***	.079
<u>SEZ Administrations</u>	<i>Rank</i>	<i>Coefficient</i> ✱	<i>Rank</i>	<i>Coefficient</i> ✱	<i>Rank</i>	<i>Coefficient</i> ✱	<i>Rank</i>	<i>Coefficient</i> ✱
Katowice: KTW	KTW	.581***	SWK	1.029*	SWK	14.601***	SWK	17.260***
Tarnobrzeg: TBZ	TBZ	.566***	MLC	.822*	KRW	9.241***	KRW	11.019***
Legnica: LGA	LGA	.553***	WMZ	.462	KTW	4.593**	KTW	6.366***
Warmia Mazury: WMZ	WMZ	.484***	KRW	.453▷	KZS	4.215*	KZS	2.968***
Mielec: MLC	MLC	.479***	STW	.240	MLC	2.869	MLC	2.258***
Lodz: LDZ	LDZ	.471***	TBZ▷	12.342***	TBZ	2.822	PMR	1.307***
Krakow: KRW	KRW	.463***	PMR	-.240	LDZ	1.828	WBZ	.3238
Starachowice: STW	STW	.425***	LDZ	-.673*	LGA	1.734	TBZ▷	13.619
Suwalki: SWK	SWK	.424**	LGA	-.750	WBZ	1.117	LDZ	-.0305
Pomeranian: PMR	PMR	.393***	KTW	-1.121*	PMR	.780	WMZ	-.283
Walbrzych: WBZ	WBZ	.211	KMG	-1.241*	SLP	.632	SLP	-.8679*
Kostrzyn Slubice: KZS	KZS	.210	WBZ	-1.370**	STW	.074	STW	-2.141***
Kamienna Gora: KMG	KMG	.120	SLP	-1.789**	KMG	-2.656	LGA	-2.144***
Slupsk: SLP	SLP▷	.871**	KZS	-2.032**	WMZ▷	-19.951***	KMG	-4.114***
Local dummies (Lau2)	Yes (random)		No		Yes (random)		No	
Year dummies	Yes (random)		Yes (random)		Yes (random)		Yes (random)	
No. of observations		3320		140		3320		140
No. of localities		334		14		334		14
R ² within/between/tot		0.02/0.93/0.64		0.87		0.03/0.53/0.53		0.99

***/**/* the coefficient estimate is significant at the 1 percent/5 percent/ 10 percent level. ✱ Coefficients should be read as deviations from the omitted dummy (marked with ▷) in the column.

Table 4. – Modelling zones administrations using descriptors and unobserved cluster effects

	Employment RE model		Employment BE model		Share in region RE model		Share in region BE model	
	<i>Coefficient</i>	<i>s.e.</i>	<i>Coefficient</i>	<i>s.e.</i>	<i>Coefficient</i>	<i>s.e.</i>	<i>Coefficient</i>	<i>s.e.</i>
Log All firms	.844***	.079	.998***	.090	2.613***	.229	10.530***	.686
Log Foreign firms	.102*	.058	-.155***	.037	.764***	.099	-1.439***	.708
Wages (indexed)	-.002	.001	.004***	.001	.007	.008	.203***	.040
Urban dummy	.274***	.083	NA	NA	10.940***	1.178	NA	NA
Log Subsidies	.073***	.003	.112	.065	.060	.092	.691***	.125
Log EU subsidies	.010***	.002	.042***	.002	-.008***	.003	.082	.083
EU member	-.242***	.063	-.458***	.041	-.279***	.049	-.479	.273
<u>SEZ Administrations:</u>								
Log Land	.027	.098	.152***	.019	.251	1.274	-3.074***	.853
Log KL ratio	-.006	.013	-.028***	.005	-.245	.2356	.178	.234
Governance quality (scaled)	.155	.201	.080***	.013	.164	2.593	.645*	.370
Local dummies (Lau2)	Yes (random)		No		Yes (random)		No	
Year dummies	Yes (random)		Yes (random)		Yes (random)		Yes (random)	
Number of observations	3320		140		3320		140	
Number of localities	334		14		334		14	
R ² within/between/total	0.02/0.92/0.64		0.84		0.03/0.51/0.50		0.76	

***/**/* the coefficient estimate is significant at the 1 percent/5 percent/ 10 percent level.

Table 5. – Modelling zones administrations case-by-case – Dependent variable is Log Employment (continued on next page...)

	KMG Kamienna Gora	KTW Katowice	KZS Kostrzyn Slubice	KRW Krakow	LGA Legnica	LDZ Lodz	MLC Mielec
Intercept	1.585**	1.936***	3.241***	-0.860	1.483	3.192***	1.450**
Trend	-(9, 10 -***)	-(9,10 -***)	-*** (9, 10 >)	++	-(9,10 -***)	-(9,10 -***)+	/
Log All firms	0.950***	0.982***	0.650***	1.01	1.04***	0.705***	1.047***
• Trend	+	+(9,10 -**)	+***	-	-	-	-
Log Foreign firms	0.047	-0.035	0.275***	-0.276	0.309	0.137	-0.153
• Trend	-	-	-**	+*	+	+	+
Urban dummy	0.336	0.206	0.125	1.010***	-0.720	0.683***	0.844***
• Trend	-	+	-	-	+	-	-
Log Subsidies	-0.011	-0.009	-0.015	0.161	-0.035	-0.027	-0.035
• Trend	-(9, 10 +***)	-(9, 10 +***)	-(9, 10 +***)	-	-(9, 10 +***)	+(9, 10 +***)	-
Log EU subsidies	-0.287***	-0.047	-0.052	-0.158	0.151	-0.011	0.044
• Trend	+***	+	+	-	-	+	-
Number of observations	289	577	442	238	221	848	340
Number of localities	18	35	26	14	13	51	20
R ² within/ between/ total	0.94/ 0.98/ 0.97	0.98/ 0.96/ 0.97	0.92/ 0.97/ 0.96	0.51/ 0.97/ 0.97	0.96/ 0.92/ 0.93	0.95/ 0.83/ 0.87	0.23/ 0.96/ 0.96

***/**/* the coefficient estimate is significant at the 1 percent/5 percent/ 10 percent level.-/+// the coefficient estimate is negative, positive or goes in both directions. > the coefficient estimate is larger in particular years.

Table 5. – Concluded – Dependent variable is Log Employment.

	PMR Pomeranian	SLP Slupsk	STW Starachowice	SWK Suwalki	TBZ Tarnobrzeg	WBZ Walbrzych	WMZ Warmia Mazury
Intercept	1.570	2.062**	1.511	2.417***	3.540***	2.551***	1.001
Trend	/	-**(9, 10 >)	-**(9, 10 >)	+**	-(9,10 -***)	-*** (9, 10 >)	-
Log All firms	0.870***	0.819***	1.056***	1.083***	0.753***	0.767***	0.955***
• Trend	/	+	+	+	+	+(9,10 -***)	+
Log Foreign firms	0.030	0.106	-0.253	0.024	0.014	0.208***	0.049
• Trend	+	/	+	+	+	/	-
Urban dummy	0.220	0.420	0.576*	-0.014	0.711***	-0.005	-0.067
• Trend	/	-	-	+	-	-	+
Log Subsidies	0.056	0.003	-0.001	-0.101	-0.035**	0.013	0.049
• Trend	-	-	-(9, 10 +***)	-	+(9, 10 +***)	-	-
Log EU subsidies	0.013	-0.050	0.114	0.070	0.190	-0.142	0.111
• Trend	-	+	/	-	-	+**	-
Number of observations	364	204	272	126	612	673	459
Number of localities	22	12	16	8	36	41	27
R ² within/ between/ total	0.36/ 0.96/ 0.93	0.98/ 0.98/ 0.98	0.96/ 0.96/ 0.96	0.73/ 0.99/ 0.99	0.97/ 0.89/ 0.93	0.93/ 0.96/ 0.96	0.16/ 0.97/ 0.95

***/**/* the coefficient estimate is significant at the 1 percent/5 percent/ 10 percent level.-/+// the coefficient estimate is negative, positive or goes in both directions. > the coefficient estimate is larger in particular years.