II. **ABORDĂRI ECONOMETRICE ALE UNOR PROCESE ECONOMICE COMPLEXE**

LEAST DEVIANCE ESTIMATION BOOTSTRAP TECHNIQUES APPLIED TO AGGREGATED PRODUCTION ELASTICITY COEFFICIENTS. EMPIRICAL EVIDENCE FROM THE PALESTINIAN INDUSTRY

Alexandrina Ioana SCORBUREANU
Doctorand and Teaching Assistant, University of Verona
and
Doctorand, Ecole Normale Supérieure du Cachan
emails: alexandrina.scorbureanu@univr.it
alexandrina_scorbureanu@yahoo.com

Abstract

The aim of this paper is to provide production elasticity estimates for the aggregate production functions of developing countries. We use aggregate data concerning the production sectors from two Middle Eastern countries. Unfortunately, the available data are quite of bad quality (small samples with high variability and time inconsistency), implying that the traditional OLS-estimates are biased. We propose an estimation procedure based on the bootstrap least deviance technique and find that the estimated elasticity is both significant and robust. For time-saving purposes, we repeat estimates for three available cross-sections of 71 manufacturing aggregates, and obtain increasing returns to scale for the manufacturing sector, which are supposed to reflect the imperfect competition of the market and/or the existence of high set-up or sunk costs, that are mandatory in order to produce at all.

Key-words: *aggregate elasticity estimation, bootstrap LAD estimator, production elasticity, developing countries*

JEL Classification: C32, C10

Literature review

The Semiparametric approach of Least Deviance Estimator is already found in the studies of Butler, McDonald, Nelson and White (1990) and McDonald and White (1993)[1]. This technique historically precedes the Ordinary Least Squares family of estimators (OLS) and was successfully applied to the estimation of production functions in small samples with high variability for the transportation industry in US, by Eellner & Revankar (1970), to show that economies of scale vary with output. They demonstrated therefore that bootstrap LAD estimates were unbiased.

In this paper, we found that the same conclusion applies to production estimates from the Palestinian manufacturing sector. We inspect on simple Cobb-Douglas production aggregates and obtain both statistically significant and

[1] For a technical implementation of the LAD estimator, see Hardle (1970).

robust elasticity for output with respect to labor and respectively intermediary consumption in the manufacturing sector.

Production elasticity for the Palestinian *stone* industry are available for the year 2003 and they are OLS-estimates of a CES-production function proposed by B. Makhool[2], whereas Cobb-Douglas production functions of the Palestinian *stone cutting* industry are dated back in 1997 (the same author). Results of the study revealed that the stone industry, in general, was characterized by decreasing returns to scale, while small firms enjoyed constant returns to scale. Also, it was found that the output elasticity with respect to labor was greater than the output elasticity with respect to capital. In addition, a significant statistical difference at 1% significance level was found between large and small firms in the sense that large firms faced a low elasticity of substitution between labor and capital, while small firms had higher possibilities of substituting labor for capital.

Data description

The economic background

The manufacturing sector in WB&G has constantly decreased its contribution to GDP since 1994 (from 22% in 1994 to around 12% in 2004) and gave more and more space to a service-based economy (which on the contrary to manufacturing, increased its contribution from 53% to 72% in the total GDP[3]). The failure in establishing growth patterns for the Palestinian private sector, in particularly the manufacturing industry, are also caused by fundamental changes in the economy. The local industry developed to produce low value-labor intensive goods for the Palestinian and Israeli domestic markets. Often, this was done by collections of small Palestinian enterprises serving as sub-contractors for larger Israeli firms who designed and marketed the goods[4]. Also, between 1994 and 2004 the manufacturing sector's share in total employment fell from an estimated 14% to 12%.

From the microeconomic point of view, we expect to obtain increasing production returns to scale which may find the explanation in two main causes:

- they reflect the imperfect competition on the market (typically oligopolistic) and/or
- the fact that any feasible input-output vector may be scaled-up (or in other words, units of a good can be produced at a constant cost of input, given that fixed set-up costs are required in order to produce at all).

Datasets

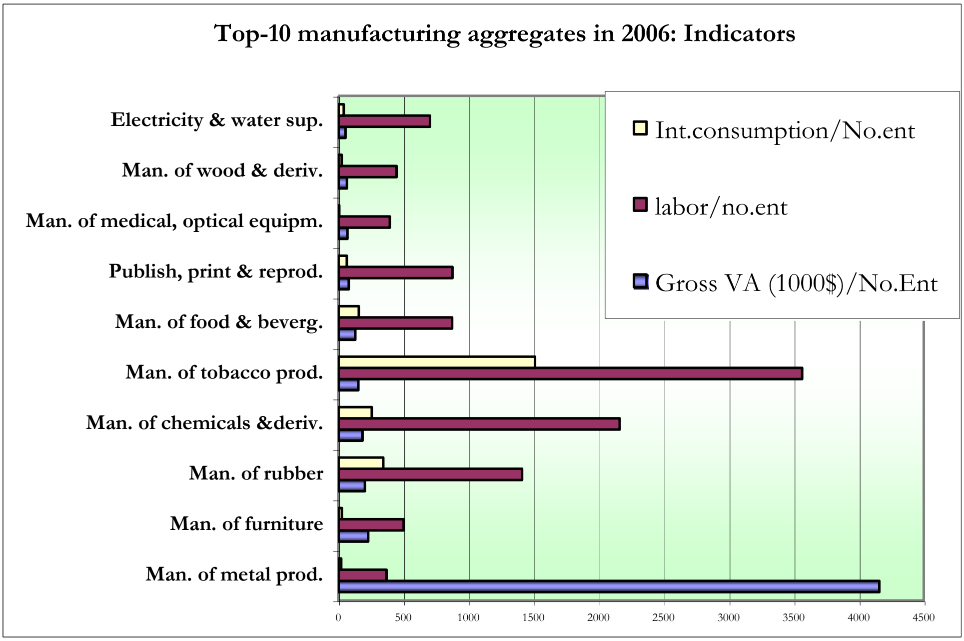
We use three cross section datasets for the years 2000, 2002 and 2006 containing 71 aggregates at the subsector level of the manufacturing in West Bank and Gaza (source: Paltrade & PCBS). We present a summarizing distribution of Gross Value Added (output) over these industries grouped in 23 aggregates (of which we present

[2] Basim Makhool, 2003 (see the References part for a detailed citation).

[3] Source*:* PCBS, 2006 National Accounts.

[4] According to the "Investment climate assessment 2007 Report No. 39109 – GZ" – World Bank Organization.

the 10 most relevant ones – situation in 2006). Remark that not necessarily the most productive sectors are the ones that absorb most resources (labor/intermediary goods): the two extreme cases are the *manufacturing of tobacco* (resource-intensive and less productive) and *manufacturing of metal products* (less resource intensive and highly productive).



Since the data available are often declared inconsistently across the years, for time-saving purposes, we do not use a panel dataset, but rather select three cross-sections: one for 2000, one for 2002 and a third one for 2006, and compare the results. Therefore, the main issue that arises when it comes to estimate elasticity is the small sample problem – there are on average 71 industries by period – characterized by a high variability of data between aggregates and across time. In this case, asymptotic approximations need not be very good, especially with small sample sizes and unusual features of the population distribution (i.e. thick-tailed distribution of dependent variable across data). Therefore, simulation methods, while always special, can help determine how well the asymptotic approximations work, whereas resembling methods can allow us to improve on the asymptotic distribution approximations. They also may simplify the calculation of standard errors, confidence intervals, and p-values for test statistics, and we can get a good idea of the amount of finite-sample bias in the estimation method. In addition, is well known from the literature that under certain assumptions and for certain statistics, resembling methods can provide quantifiable improvements to the usual asymptotics.

In the following tables, we present summarizing statistics of the three datasets, in which variables of the form *l_variableN* stand for the logarithm transform of the value *variable/number of enterprises* and we use them for estimation purposes.

Year 2000

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
year	64	2000	0	2000	2000
id	64	46.09375	28.4949	1	98
vad	38	13012.63	21766.66	82.73149	83799.2
icons	38	16206.55	26569.38	116.0025	128807.9
output	38	29219.18	44889.19	240.3014	179217.2
wages	38	4732.529	9976.636	25.00681	53700.54
noempl	38	1529	1529	3778.211	22050
noent	38	227.9763	480.6635	11.65098	2078.409
exp	38	4965.079	10923.88	0	62476.32
local	38	21032.89	35605.02	29.01351	169978.9
finprod	38	26280.39	41472.91	29.01351	178281
l_vadn	38	3.910689	1.406147	1.393999	7.938782
l_laborn	38	1.791587	.8962016	1.803845	3.977748
l_iconsn	38	4.142978	1.603312	1.423005	7.22785

Year 2002

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
year	89	2002	0	2002	2002
id	89	47.52809	27.51986	1	94
vad	51	8889.166	12591.48	-3699.396	57600.46
icons	51	15660.98	25748.49	52.21857	128807.9
output	51	24550.15	35308.46	237.5958	147780
wages	51	3382.55	5885.408	10.21638	33978.19
noempl	51	1340.238	2692.835	30.72759	16837.19
noent	51	282.0943	563.8627	10.61408	2886.372
exp	51	2776.191	6080.892	0	33797.75

local	51	19881.51	30240.98	68.10924	147739.8
finprod	51	22875.8	33918.38	69.30115	147770.9
l_vadn	49	3.605849	1.337986	1.393999	7.938782
l_laborn	51	1.693411	.7268931	.6546682	4.300247
l_iconsn	51	3.888265	1.529946	.9065158	8.283315

Year 2006

Variable	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
id	55	45.45455	27.08069	1	92
vad	47	9211.788	24010.62	27.50754	139969.7
icons	47	11195.92	24101.78	28.05374	113765
output	47	20407.71	45796.55	81.86168	239352
wages	47	2002.3	4674.008	6.4207	22944.48
noempl	47	665.45	1630.676	11	8802.271
noent	47	109.0168	251.2563	4	1016.277
exp	47	1665.41	5891.217	0	39193.18
local	47	17531.98	41812.1	3.3675	233788.2
finprod	47	19246.94	44987.16	3.2777	237940.1
year	55	2006	0	2006	2006
l_vadn	47	4.049711	1.365381	1.928166	8.66451
l_iconsn	47	4.352924	1.56788	1.542363	7.524533
l_laborn	47	1.860212	.7530818	.4519851	4.318272

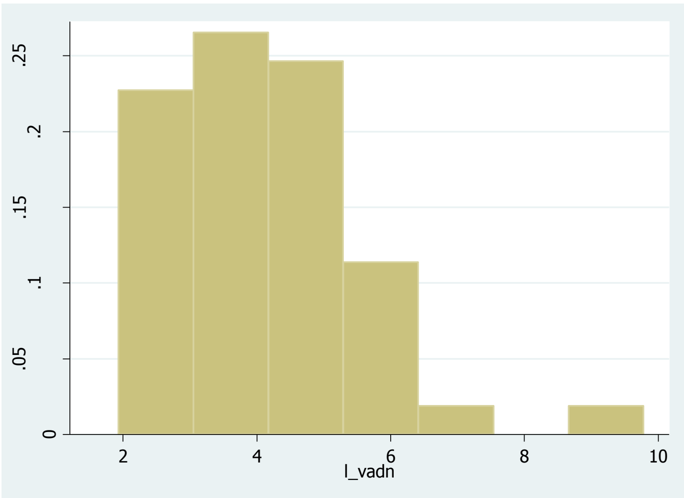
Remark that the average number of firms by industry decreases by more than a half between 2000 and 2006 (from 227 in 2000 to 109 in 2006). This could be an effect of some administrative barriers (i.e. more rigid regulations for firm-creating bureaucracy), which may lead to an oligopolistic market structure particularly accentuated, given that the market power is already concentrated in the hands of few powerful and rich owners. Another effect may be the high barriers to enter on the manufacturing industry market, due to existing high levels of fixed set-up costs. Also, the average number of employees by industry decreased dramatically, from

1529/sector in 2000 to an average of 665 in 2006. This value may represent the cause of a twofold effect:

- the *follow-up effect*, which is due to the presence of a smaller number of companies on the market in 2006 with respect to 2000, which are not able to absorb as much resources as before; and
- the *migration* effect, which is due to the fact that the non-tradable sectors are the principal labor-donors in the West Bank (Ramallah, in particular) and they absorbed resources, which initially were employed in manufacturing sectors.

We consider that two additional remarks are important at this point:

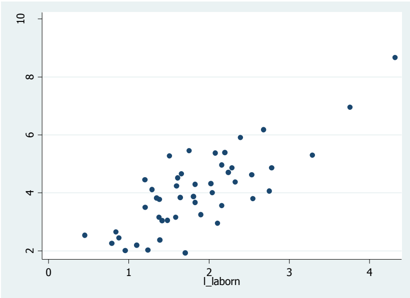
- there is a relatively important variability of data across periods (*time variability*);
- the high variability in productivity among industrial sectors (*sector-variability*) announces a thick-tailed distribution of prediction error terms (see the graphic below for 2006 distribution of log(value-added) across industries).

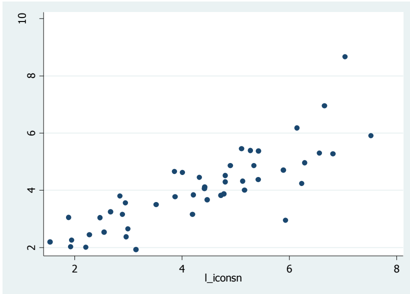


Log-Value Added distribution across industries in 2006

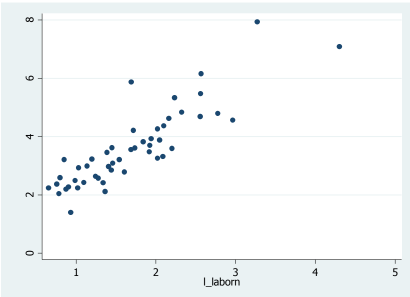
We also expect a linearly positive effect of labor and intermediary consumption on the industry output in each period, as Value added distributions in the three samples suggest (see the following graphics).

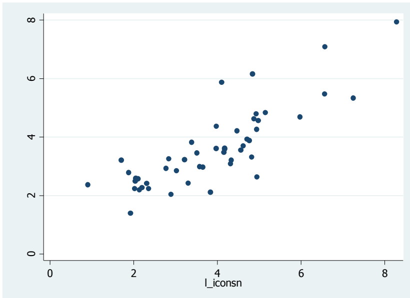
Year 2006



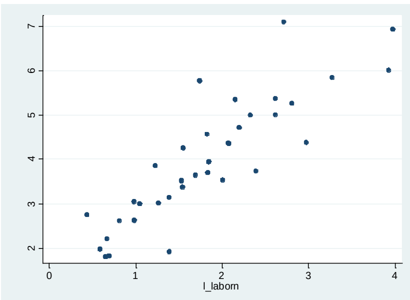


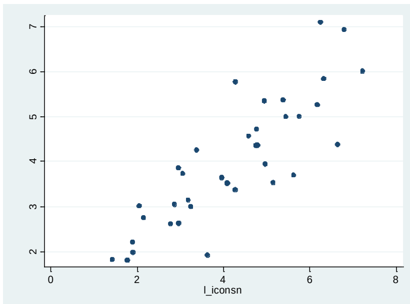
Year 2002





Year 2000





Panel 1. Value-added distribution with respect to Labor employment and Intermediary consumption in WB&G

The model

Notation

The model to be estimated is the following log-log Cobb-Douglas production aggregate:

$$\ln\left(\frac{y_i}{N_i}\right) = \beta_0 + \beta_1 \ln\left(\frac{C_i}{N_i}\right) + \beta_2 \ln\left(\frac{L_i}{N_i}\right) + \varepsilon_i$$

$$\Leftrightarrow \tag{0.1}$$

$$l\_vadn = \beta_0 + \beta_1 \cdot l\_iconsn + \beta_2 \cdot l\_laborn + \varepsilon_i$$

in which *i=1...n* are the industries (corresponding to *n* observations in each cross section dataset), y_i is the output of each industry (value added), N_i is the number of firms for each industry *i*, C_i is the intermediate consumption and L_i is the value of labor employed in industry i (equal to the number of employees multiplied by the total number of hours worked in a month and normalized by the number of firms).

Some remarks on LDA and Bootstrap estimations

In our case, we found bootstrap technique particularly useful in obtaining estimates of the standard errors of quintile-regression coefficients. *Stata* software performs quintile and obtains the standard errors using the method suggested by Koenker and Bassett (1978,1982). Rogers (1992) reports that these standard errors are satisfactory in the *homoscedastic* case but that they appear to *be understated* in the presence of *heteroscedastic* errors. We follow the traditional notation used in the econometric theory[5], therefore the OLS estimates are as usual, while we provide a refreshment for LAD estimates, which are the solution to the minimization problem:

$$\min_{b_0} \sum_{i=1}^{n} | y_i - x_i b_0 | \tag{0.2}$$

which is a special case of the quintile regression:

$$\Pr ob\left[y_i \le x_i{}'\beta\right] = q \tag{0.3}$$

In particular, LAD estimation corresponds to the median regression (i.e. *q=0.5*). Results suggest an estimation for the asymptotic covariance's matrix of the quintile regression:

$$Est.Asy.Var\left[b_q\right] = \left(X'X\right)^{-1} XDX \left(X'X\right)^{-1} \quad ({}^6)$$

in which D is the diagonal matrix containing the weights associated to different variances d_i defined as following:

[5] See Greene, *Econometric Analysis for a summary discussion on LDA method*.

[6] See Koenker and Bassett (1978,1982), Huber and Rogers (1993) that have analysed this regression and found the estimator for the asymptotic covariance matrix of the quintile regression estimator.

$$\left\{ \begin{aligned} d_i &= \left[\frac{q}{f(0)} \right]^2 if: y_i - x_i\beta > 0, and \\ d_i &= \left[\frac{1-q}{f(0)} \right]^2, otherwise. \end{aligned} \right\} \tag{0.4}$$

in which f(0) is the true distribution of disturbances. Now we obtain an estimate for f(0), supposing that it is normally distributed with variance σ^2:

$$d_i = \sigma^2 \frac{\pi}{2} \left(X'X \right)^{-1} \tag{0.5}$$

For small sample estimates, which is in our case, estimation of f(0) is computed as:

$$f(0) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K \left[\frac{e_i}{h} \right] \tag{0.6}$$

in which *h* is the bandwidth, $e_i = x_i - x$ represents the set of residuals and *K[.]* is a weighting, or the *kernel function*. We used the software *Stata Version 10.0*, which assumes the following forms for *h* and *K*:

$$\begin{aligned} h &= \frac{0.9s}{n^{1/5}} \\ K[.] &= Logit \end{aligned} \tag{0.7}$$

Bootstrap estimator for the asymptotic covariance matrix is known as:

$$Est.Var\left[b_{LAD} \right] = \frac{1}{R} \sum_{r=1}^{R} \left(b_{LAD}{}^{r} - b_{LAD} \right) \left(b_{LAD}{}^{r} - b_{LAD} \right)' \tag{0.8}$$

where *r=1...R* are the number of replications chosen, b_{LAD} is the LAD estimator of β based on a sample of n observations drawn from the original dataset. This estimator is robust to the fact that some marginal observations may exert a high influence on sample's estimates, due to the fact that b_{LAD} *penalizes* those observations, which tend to *matter mostly* in the sample, by the means of their variance's weighting. The standardized LS residuals would otherwise suggest different results, according to the exclusion or not of the *distorting observations* from the sample.

Results

We first present OLS results and compare them with LSD results for the Palestinian economy in the three periods. In the appendix, we also include auxiliary OLS estimates for the Israeli manufacturing sector. Unfortunately, these results have a low degree of comparability due to incongruencies in data registration and industry nomenclature, which is changing from one economy to another. We also provide (in the appendix) OLS estimates for 5 periods (years).

Furthermore, OLS estimates on log-technology intercept (the constant term), intermediary consumption and labor coefficients are provided for the complete datasets as well as for the "corrected" datasets for the three periods 2000, 2002 and 2006 (from which we excluded the industries which caused distortions in the results[7]). Most coefficients are significant for all three periods (except for the intermediary). Consequently, we estimate a median regression of production (value added) on intermediary consumption and labor input for the same period. We obtain LAD estimates along with Koenker–Bassett standard errors, which are invariant for the two types of dataset (complete vs. corrected).

Table 1

OLS vs. LAD estimates for output in 2006 manufacturing industries of WB&G

Dependent variable: *l_vadn* Independent variables:	**OLS *Complete dataset***	**ESTIMATES *Dataset without 2[a'] observations:***	**LAD ESTIMATES *Bootstrap (500 replications)***	**COMPLETE DATASET *Conf. Interval (95% Norm. based)***
l_iconsn	0.47*** *(0.07)[a)]*	0.49*** *(0.06)[a)]*	0.47*** *(0.06)*	[0.351; 0.607]
l_laborn	0.78*** *(0.16)*	0.77*** *(0.13)*	0.64*** *(0.18)*	[0.269; 1.011]
Const	0.52* *(0.28)*	0.54** *(0.23)*	0.76 *(0.56)*	-
R-squared	0.79	0.85	MaxLhood: -33258.41	-
***significant at 1% level	**significant at 5% level	*significant at 10% level	a) in parenthesis: Standard Errors	a') missing obs.: Man. of diary prod; Man. of rubber; Man. other texts

Table 2

OLS vs. LAD estimates for output in 2002 manufacturing industries of WB&G

Dependent variable: *l_vadn* Independent variables:	**OLS *Complete dataset***	**ESTIMATES *Dataset without 2[b'] observations***	**LAD ESTIMATES *Bootstrap (500 replications)***	***COMPLETE DATAS Conf. Interval (95% Norm. based)***
l_iconsn	0.30*** *(0.10)[b)]*	0.33*** *(0.06)[a)]*	0.35* *(0.13)*	[0.1014; 0.6117]
l_laborn	1.05*** *(0.21)*	1.04*** *(0.13)*	0.92*** *(0.26)*	[0.4105; 1.4446]
Const	0.62*** *(0.24)*	0.54** *(0.23)*	-	-
R-squared	0.79	0.85	-	-
***significant at 1% level	**significant at 5% level	*significant at 10% level	b) in parenthesis: Standard Errors	b') missing obs.: Man. of vegetable & animal oil; Man. of grain mill prod.; Man. of soap and detergents

[7] See the appendix for a scatter-plot representation of the prediction errors by industry.

Table 3

OLS vs. LAD estimates for output in 2000 manufacturing industries of WB&G

Dependent variable: ***l_vadn*** **Independent variables:**	**OLS** ***Complete dataset***	**ESTIMATES:** ***Dataset without 4c) observations***	**LAD ESTIMATES** ***Bootstrap (500 replications)***	***COMPLETE DATAS Conf. Interval (95% Norm. based)***
l_iconsn	0.32** *(0.16)c)*	0.27** *(0.12)*	0.35*(0.25)	[-0.1460; 0.8573]
l_laborn	0.79*** *(0.28)*	0.87****(0.21)*	0.79*** (0.36)	[0.0732; 1.5219]
Const	1.29*** *(0.29)*	1.31****(0.22)*	-	-
R-squared	0.79	0.85	-	-
***significant at 1% level	**significant at 5% level	*significant at 10% level	c) in parenthesis: Standard Errors	c') missing obs.: Man. of soft drink & mineral water; Manufacturing of articles of paper;

First of all, remark that distortion of OLS coefficients is worst in the case of the year 2000 (variations between 8%-20% for OLS estimates when we rely on the whole dataset compared with estimates done on the dataset without the two observations: manufacturing of soft drink and mineral water and manufacturing of articles of paper). By difference, the bootstrapped LAD coefficients are invariant from one dataset to another (therefore we only present estimations for the complete dataset). We also remark that throughout estimated coefficients suggest a strong reliability of industry value added on labor resources (coefficients associated to the normalized log-labor are 0.79 in 2000 and 0.92 in 2002).

Conclusions

We presented a method of estimating robust coefficients in a context of small sample size, with high variance in data, as is the case of the uncertain situation on the Palestinian manufacturing market. Nevertheless, this technique may be improved once we will have the appropriate data to test it: for instance, a more complete pooled dataset and eventually microdata tests must be taken into account. Also, explaining the impact of fixed capital on value added and eventually estimating cost functions in the future may be revealing.

Precedent studies on the Palestinian market identified some factors that are responsible for low rates of increase in productivity for the Palestinian manufacturing sector. The first is related to low rate of embodied technical progress resulting from negative rate of growth in physical capital. The others are related to factors causing inefficiency and they are: the misallocation of factors of production among sectors and firms caused by various impediments to free mobility of persons and goods, and finally the inefficiency resulting from the existence of idle resources (both labor and capital). Also, the cost structure of the sector reveals that wages account for one fourth (25.3%) of the total cost. It concentrates on three types of cost constituting the remaining three fourth of total cost and calculates their growth rates, and a weighted average of which is usually found to be negative. It observes that, despite the negative rate, the level of certain non wage costs are relatively high (cost of utilities – electricity and water, the cost of transportation and the cost of clearing imported goods through Israeli customs).

Acknowledgements

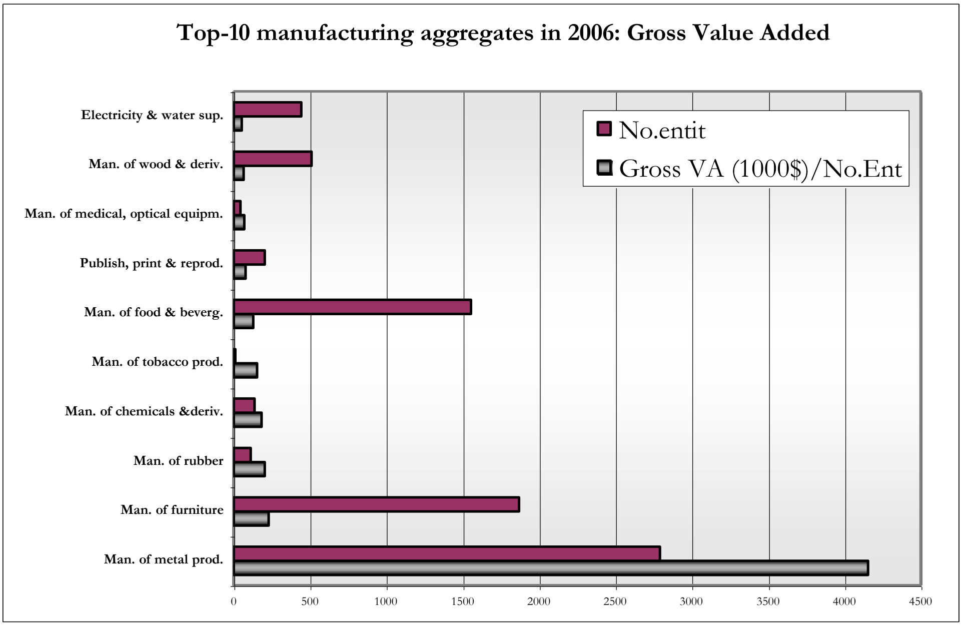
The research was realized in the Department of Economic Sciences at the University of Verona under the project EuroMidBridge and it is part of the feasibility study for the realization of an integrated freight village at Jenin (West Bank), project which is financed by the Italian Minister of Foreign Affairs and the Italian Cooperation for Palestine Office in Jerusalem. Nevertheless, this paper does not represent any political orientation, nor the Institutions' standpoint with respect to the subject, it objectively presents results obtained by using available data (see references). We also assume the responsibility of any errors, which despite our efforts, may have remained unobserved.

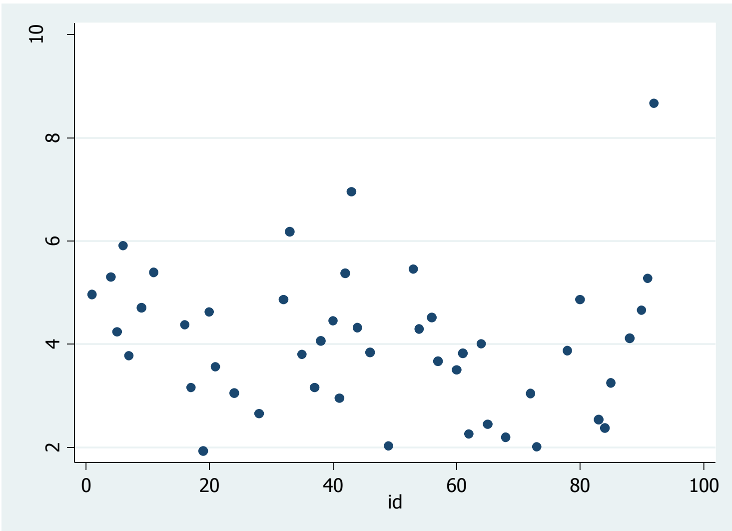
REFERENCES

- Datasets: Paltrade and PCBS.
- Basim Makhool, *Function Analysis of the Stone Industry in the West Bank and Gaza.* An-Najah University Journal for Research – Humanities (B), 2003, ISSN: 1727-8449.
- World Bank Organization document, *Investment Climate Assessment,* March 2007 Report No. 39109 – GZ.
- William Greene, *Econometric Analysis*, MIT Press, 1995.

APPENDICES

A. Summary statistics

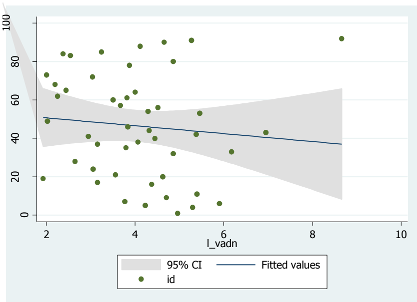




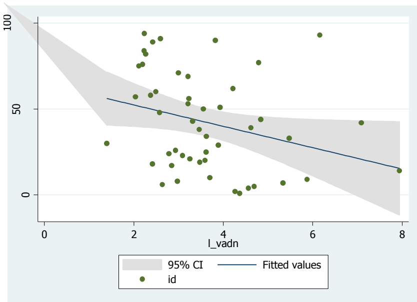
Output distribution on industrial sector 2006 (by industry id)

B. **Estimates and predictions**

Year 2006

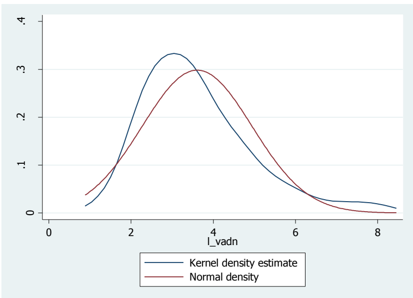


Year 2002

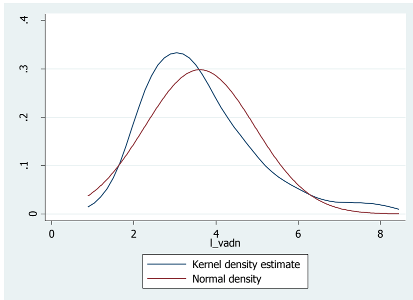


OLS Linear prediction of log value added (normalized by *Ni*)

Year 2002

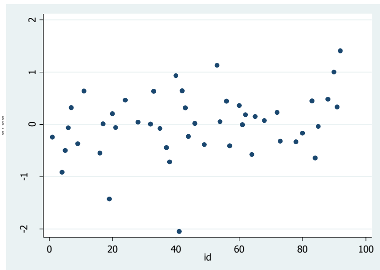


Year 2006

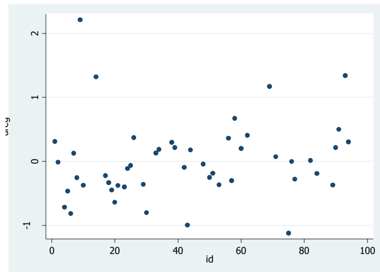


Linear kernel density estimation of *l_vadn* (vs. Normal density)

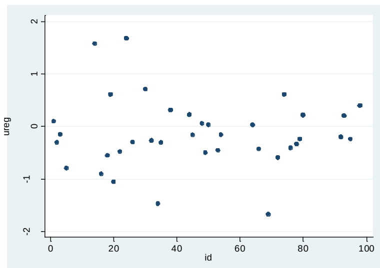
Year 2006



Year 2002



Year 2000



Linear prediction Errors after OLS, distribution by industry sectors (id)

C. **Auxiliary estimates**

In this section, we report OLS estimates for 2000 and 2002 from the Israel manufacturing account[8]. In this case there is no reason to apply bootstrap techniques, given that linear prediction errors are normally distributed (there are no distortionary observations in the sample and coefficients are both unbiased and efficient). Remark that we obtain even in this case increasing returns to scale, which causes will be furthermore investigated in more detail (we do not have data on the sector firms number).

The estimated model is:

$$Y_i = A \cdot C_i^{\alpha 1} \cdot L_i^{\alpha 2}$$

$$\Leftrightarrow$$

$$\ln(Y) = \alpha_0 + \alpha_1 \ln(C_i) + \alpha_2 \ln(L_i)$$

$$\Leftrightarrow$$

$$l\_output = \alpha_0 + \alpha_1 \cdot l\_totin + \alpha_2 \cdot l\_totlab + \upsilon_i$$

where, as before, *l_totin* is the log-transform of the total input, *l_totlab* is the log-transform of total labor input in the industry and α_0 is the technology constant term.

Table 4

OLS estimates for output in 2000 manufacturing industries of Israel

Dependent variable: *l_out* **Independent variables:**	**OLS** ***With constant term***	**ESTIMATES** ***Without the constant term***
L_totin	0.30*** *(0.10)*$^{b)}$	0.33*** *(0.06)*$^{a)}$
L_totlab	1.05*** *(0.21)*	1.04*** *(0.13)*
Const	0.62*** *(0.24)*	0.54** *(0.23)*
R-squared	0.79	-
***significant at 1% level	**significant at 5% level	*significant at 10% level

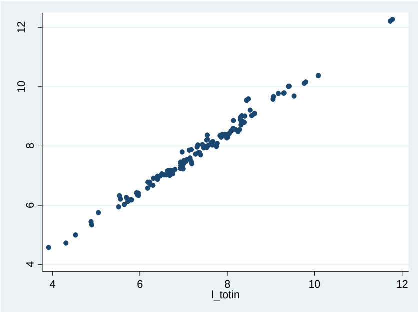
Table 5

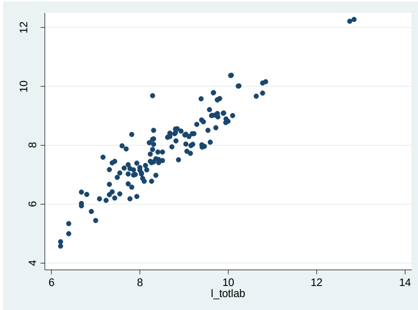
OLS estimates for output in 2002 manufacturing industries of Israel

Dependent variable: *l_out* **Independent variables:**	**OLS** ***With constant term***	**ESTIMATES** ***Without the constant term***
l_totin	0.30*** *(0.10)*$^{b)}$	0.33*** *(0.06)*$^{a)}$
l_totlab	1.05*** *(0.21)*	1.04*** *(0.13)*
Const	0.62*** *(0.24)*	0.54** *(0.23)*
R-squared	0.79	-
***significant at 1% level	**significant at 5% level	*significant at 10% level

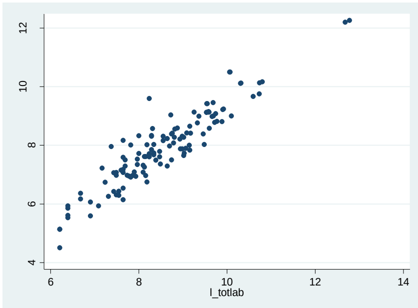
[8] *Source*: Central Bureau of Statistics of Israel.

Year 2000





Year 2003



Labor use/Intermediate consumption and Value Added in Israel