Analyzing the Impact of the Sample Renewing Effect on Aggregate Productivity: Evidence from the Greek Olive Oil FADN Data

Giannis Karagiannis
University of Macedonia, Thessaloniki, Greece

karagian@uom.gr

Slides prepared for presentation at the International Agricultural Trade Research Consortium’s (IATRC’s) 2013 Symposium: Productivity and Its Impacts on Global Trade, June 2-4, 2013, Seville, Spain
Analyzing the Impact of the Sample Renewing Effect on Aggregate Productivity: Evidence from the Greek Olive Oil FADN data

Giannis Karagiannis

Professor, Department of Economics, University of Macedonia, Thessaloniki, GREECE
Introduction

The Farm Accounting Data Network (FADN) data set has been extensively (48 journal papers during the period 1996-2011) for performance evaluation of agricultural holdings in EU-27

Previous studies analyze efficiency and productivity differences and changes over a given period of time

Also infer aggregate (industry-level) efficiency and productivity by usually taking *simple arithmetic averages* of the farm-level estimates
Motivation

In rotated panels, as the EU FADN, farms are replaced in the sample after a period of five years by seemingly similar farms in terms of some observable characteristics such as location, size, and type of farming to maintain the stratified nature of the data.

However, farms may have different efficiency and productivity levels that are not directly observable by the designers of the replacement process and this may affect our estimates of aggregate productivity.
Research Question & Hypothesis

examine the impact that the replacement process in rotated panels may have on industry-level productivity changes

If the unobserved performance heterogeneity is perfectly or highly correlated with the observable characteristics used for the replacement process then the impact of entering and dropped from the sample farms on aggregate productivity changes will be near zero or negligible

Otherwise, the replacement process will affect changes in aggregate productivity
Sample Renewing Effect

that part of aggregate productivity change that is related to the replacement process (i.e., stratification and rotation)

It contains the effects of
1. Entering to the sample (selected and voluntarily participate)
2. Exit the market (self selection)
3. Voluntarily withdraw from the sample
4. Replaced
The FADN data set

it is a harmonized, stratified, rotated and unbalanced data set

It is drawn from the population of commercial farms, defined as those that are large enough to provide a main activity to the farmer and a level of income sufficient to support his/her family. This is determined by the imputation of a economic size threshold $= \{ \text{sum of scale} \times \text{SGM} \} / \text{ESU}$, which is different for each member state.

the coverage range from a low of 5% in Slovakia to a high of 83% in Ireland, with the average being around 45%
The FADN data set

**Stratification Criteria:** location of the farm
economic size
type of farming

A farm is classified into a particular type of farming when at least \( \frac{2}{3} \) of its total SGM is coming from a particular activity

Stratification matrix = 140 regions x 72 types of farming x 9 economic classes

For each cell a representative number of farms are included in the FADN sample
The FADN data set

Sample fractions may differ depending on the number of available observations as sometimes it is difficult to find farms for particular cells.

The agricultural holdings included in the FADN sample are selected at random conditional upon availability of farm accounts in order to be able to complete the EU FADN Farm Return and on that participation is voluntary.

Farms usually stay for a period of five years in the sample and then are replaced with similar ones in terms of observable characteristics (location, size and type of farming).
Our Sub-sample: Olive producers in Greece

6,004 observations available for olive producers during the period 1991-2002, corresponding to 1,281 farms

around 14% of farms are dropped from the sample and replaced every year

Variables

Output = deflated revenue by an industry price index
Land = hectares
Labour = annual working hours of family and paid labor
Intermediate inputs = deflated expenditures by a sector price index
Capital = deflated end-of-the-year booking value
<table>
<thead>
<tr>
<th>Year</th>
<th>Staying in the sample</th>
<th></th>
<th>Entering to the sample</th>
<th></th>
<th>Dropped from the sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
</tr>
<tr>
<td>1991-1992</td>
<td>444</td>
<td>86.2</td>
<td>71</td>
<td>13.8</td>
<td>49</td>
<td>9.9</td>
</tr>
<tr>
<td>1992-1993</td>
<td>436</td>
<td>85.4</td>
<td>87</td>
<td>14.6</td>
<td>65</td>
<td>12.6</td>
</tr>
<tr>
<td>1993-1994</td>
<td>402</td>
<td>79.9</td>
<td>101</td>
<td>20.1</td>
<td>125</td>
<td>23.7</td>
</tr>
<tr>
<td>1994-1995</td>
<td>361</td>
<td>68.9</td>
<td>163</td>
<td>31.1</td>
<td>142</td>
<td>28.2</td>
</tr>
<tr>
<td>1995-1996</td>
<td>404</td>
<td>81.8</td>
<td>90</td>
<td>18.2</td>
<td>120</td>
<td>22.9</td>
</tr>
<tr>
<td>1996-1997</td>
<td>447</td>
<td>91.6</td>
<td>41</td>
<td>8.4</td>
<td>47</td>
<td>9.5</td>
</tr>
<tr>
<td>1997-1998</td>
<td>458</td>
<td>91.8</td>
<td>41</td>
<td>8.2</td>
<td>30</td>
<td>6.1</td>
</tr>
<tr>
<td>1998-1999</td>
<td>464</td>
<td>92.2</td>
<td>39</td>
<td>7.8</td>
<td>35</td>
<td>7.0</td>
</tr>
<tr>
<td>1999-2000</td>
<td>484</td>
<td>92.0</td>
<td>42</td>
<td>8.0</td>
<td>19</td>
<td>3.8</td>
</tr>
<tr>
<td>2000-2001</td>
<td>398</td>
<td>85.4</td>
<td>68</td>
<td>14.6</td>
<td>128</td>
<td>24.3</td>
</tr>
<tr>
<td>2001-2002</td>
<td>411</td>
<td>88.2</td>
<td>55</td>
<td>11.8</td>
<td>55</td>
<td>11.8</td>
</tr>
<tr>
<td>Average</td>
<td>85.8</td>
<td>14.2</td>
<td></td>
<td></td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staying in the sample</td>
<td></td>
<td>Entering to the sample</td>
<td></td>
<td>Dropped from the sample</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------</td>
<td>-------</td>
<td>------------------------</td>
<td>-------</td>
<td>-------------------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
</tr>
<tr>
<td>1991-1995</td>
<td>199</td>
<td>24.3</td>
<td>325</td>
<td>39.7</td>
<td>294</td>
<td>35.9</td>
</tr>
<tr>
<td>1992-1996</td>
<td>172</td>
<td>20.5</td>
<td>322</td>
<td>38.5</td>
<td>343</td>
<td>41.0</td>
</tr>
<tr>
<td>1993-1997</td>
<td>184</td>
<td>22.1</td>
<td>304</td>
<td>36.6</td>
<td>343</td>
<td>41.3</td>
</tr>
<tr>
<td>1994-1998</td>
<td>236</td>
<td>30.8</td>
<td>263</td>
<td>34.3</td>
<td>267</td>
<td>34.9</td>
</tr>
<tr>
<td>1995-1999</td>
<td>328</td>
<td>46.9</td>
<td>175</td>
<td>25.0</td>
<td>196</td>
<td>28.1</td>
</tr>
<tr>
<td>1996-2000</td>
<td>380</td>
<td>59.4</td>
<td>146</td>
<td>22.8</td>
<td>114</td>
<td>17.8</td>
</tr>
<tr>
<td>1997-2001</td>
<td>296</td>
<td>45.0</td>
<td>170</td>
<td>25.8</td>
<td>192</td>
<td>29.2</td>
</tr>
<tr>
<td>1998-2002</td>
<td>310</td>
<td>47.3</td>
<td>156</td>
<td>23.8</td>
<td>189</td>
<td>28.9</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>37.0</strong></td>
<td></td>
<td><strong>30.8</strong></td>
<td></td>
<td><strong>32.2</strong></td>
<td></td>
</tr>
</tbody>
</table>
Analytical Framework

\[ A_t = \sum_{i=1}^{N} \theta_{it} A_{it} \]

\( \theta \) is specified according to the denominator rule to guarantee consistency in aggregation (van Biesebroeck, 2008; Fare and Karagiannis, 2013)

partial factor productivity measures to be aggregated using the shares of the reference input and TFP using the share of aggregate (real) input

the resulting aggregate measure has exactly the same interpretation as the farm-level measures and there are no monotonicity violations
Analytical Framework

\[
A = A_t - A_{t-1} = \sum_{i=1}^{N} \theta_{it}A_{it} - \sum_{i=1}^{M} \theta_{it-1}A_{it-1}
\]

\[
= \left( \sum_{s} \theta_{it}A_{it} - \sum_{s} \theta_{it-1}A_{it-1} \right)
+ \left( \sum_{n} \theta_{it}A_{it} - \sum_{d} \theta_{it-1}A_{it-1} \right)
\]

= contribution of remaining in the sample farms +
+ the sample renewing effect
Analytical Framework

we use the symmetric weighting scheme suggested by Griliches and Regev (1995) to decompose the contribution of the remaining in the sample farms

\[
\Delta A = \sum_s \bar{\theta}_i \Delta A_i + \sum_s \bar{A}_i \Delta \theta_i \\
+ \left( \sum_n \theta_{it} A_{it} - \sum_d \theta_{it-1} A_{it-1} \right)
\]

(a) productivity changes of farms being in the sample for both periods
(b) share weights shift among staying in the sample farms
Analytical Framework

Since entering farms essentially replace dropped farms in order to guarantee the representativeness of the sample, to properly account for the contribution of this replacement effect we should compare productivity between entering and exiting units.

Moreover, as long as farms that dropped from the sample look very much like the new entrants, the share of entrants will in general be quite similar to the market share of drops.

Aggregate all farms that exit into a single unit and all farms that enter into another single unit and measure their productivity using share within each group.
Analytical Framework

$$
\Delta A = \sum_s \bar{\theta}_i \Delta A_i + \sum_s \bar{A}_i \Delta \theta_i + \frac{A_{at-1} + A_{nt}}{2} \left( \sum_n \theta_{it} - \sum_d \theta_{it-1} \right) + \frac{\sum_n \theta_{it} + \sum_d \theta_{it-1}}{2} (A_{nt} - A_{at-1})
$$

(a) reallocation of market share between the entering to and dropped from the sample farms

If stratification by economic size is reflected in their market share, then this term will be small in magnitude and will reflect the changes (if any) in the sample size

(a) Productivity difference between entering and dropped farms
Farm-level Productivity Measurement

\[ \Delta FP_{it}^j = (lny_{it} - \bar{lny}_t) - (lnx_{jit} - \bar{lnx}_jt) \]

\[ + \sum_{r=2} (lny_r - lny_{r-1}) - \sum_{r=2} (lnx_{jr} - lnx_{jr-1}) \]

the first two terms of this index measure cross-sectional productivity differences (i.e., a catching-up effect) and the other two times-series productivity changes that result from the shift in the productivity distribution over time.

the resulting productivity index measure the proportional difference in productivity between for a particular farm at year t relative to the hypothetical unit in the base period.
## Empirical Results

For 1991-1992

<table>
<thead>
<tr>
<th>Change in Aggregate Productivity</th>
<th>Change in the Productivity of Staying in the Sample Farms</th>
<th>Reallocation among Staying in the Sample Farms</th>
<th>Productivity Differences between Entering to and Dropped from the Sample Farms</th>
<th>Reallocation between Entering to and Dropped from the Sample Farms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land productivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.590</td>
<td>-0.532</td>
<td>0.013</td>
<td>-0.061</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.3%</td>
</tr>
<tr>
<td>Labor productivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.564</td>
<td>-0.505</td>
<td>0.043</td>
<td>-0.089</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.8%</td>
</tr>
</tbody>
</table>
Further Work….

Analysis with 5-years rolling windows

Analysis by region

TFP estimates based on econometric estimation

Other products, countries