Monitoring of Markets and Sectors

MMS Project

Final Report

Appendices

December 05, 2011

Edited by:

Johan Eyckmans, Stijn Kelchtermans, Stijn Vanormelingen, Kristien Coucke, Annabel Sels, Cherry Cheung, Daniel Neicu (*Hogeschool-Universiteit Brussel, HUB*), Frank Verboven (*Katholieke Universiteit Leuven*), Catherine Schaumans (*Universiteit Tilburg*), Luc Mariën (*Federal Public Service Economy, FOD Economie*)

Financial support by the Belgian Federal Science Policy Office in the framework of the AGORA program (AG/00/151) is gratefully acknowledged.

Appendix 1: Sector aggregation and selection

AGRI (agriculture):

NACE 2 sectors 01 to 03

EXTRA (extraction):

NACE 2 sectors 05 to 09

MANUF (manufacturing):

NACE 2 sectors 10 to 33

UTILE (utilities and electricity):

NACE 2 sectors 35 to 39

CONST (construction):

NACE 2 sectors 41 to 43

TRADE (trading):

NACE 2 sectors 45 to 47

TRANS (transport):

NACE 2 sectors 49 to 53

PSERV (personal services):

NACE 2 sectors 55, 56 and 95, 96

BSERV (business services):

NACE 2 sectors 58 to 82

OTHER (other): everything else

NACE 2 sectors 84 to 99, except for 95 and 96

For the COMPOSITE INDICATOR, we have used all sectors except AGRI and OTHER.

For the BENEFIT OF THE DOUBT, we have used two groups of sectors:

GROUP 1: EXTRA + MANUF + UTILE

GROUP 2: CONST + TRADE + TRANS + PSERV + BSERV

Appendix 2: File Locations at FPS Economy

	File Location on Windows Network			SAS Library	
Indicators	*X:\Indicators*			*\SHARE*	
Capital Intensity	**Documentation**		**File Name Format**		
Churn	Factsheet	*"Indicators"\Documentation\Factsheet*	Indicators_Form		
Concentration	Literatures	*"Indicators"\Documentation\Literatures*	Author-Year		
Import Penetration	**Results**			**Results**	**File Name Format**
Labour Productivity	Excel	*"Indicators"\Results*	"Indicators"_Results	SAS	ID_"Indicators"_NACE"X"
Price-cost Margin	**Projects**				
Volatility of Market Share	SAS	*"Indicators"\SAS Projects*	"Indicators"		
R&D					
Composite Indicator	*X:\Composite Indicator*				
Traditional Composite Indicator	**Documentation**		**File Name Format**		
Benefit of the Doubt	Literatures	*"Composite Indicator"\Documentation\Literatures*	Author-Year		
	Results			**Results**	**File Name Format**
	Excel	*"Composite Indicator"\Results*	"Composite Indicator"_Results	SAS	ID_"Composite Indicator"_NACE"X"
	Projects				
	SAS	*"Composite Indicator"\SAS Projects*	"Composite Indicator"		
Case Studies	*X:\Case Studies*				
Entry Threshold Ratios	**Documentation**		**File Name Format**		
Quick Scan	Paper	*"Case Studies"\Documentation\Paper*	"Case Studies"-Authors-Year		
Persistence of Profits	Literatures	*"Case Studies"\Documentation\Literatures*	Author-Year		
	Results			**Results**	**File Name Format**
	Excel	*"Case Studies"\Results*	"Case Studies"_Results	SAS	ID_"Case Studies"_NACE"X"
	Projects				
	SAS	*"Case Studies"\SAS Projects*	"Case Studies"		
General	*X:\General*				
Final Report	\Final Report\				
Expert Workshops	\Expert Workshops\				
Sources	\Sources\				
Meetings	\Meetings\				

Appendix 3: Additional Documents

- Technical notes
    - Luc Mariën on the selected turnover
    - Stijn Kelchtermans on R&D data
    - Johan Eyckmans on technical implementation of Benefit of the Doubt
- Referee comments by Marcel Cannoy (Ecorys Nederland)
- Referee comments by Jan Bouckaert (Universiteit Antwerpen)
- Expert meeting 2010
    - Program
    - Conclusions
- Expert meeting 2011
    - Program
    - Conclusions
- Technical forms on the indicators
    - Capital Intensity
    - Churn
    - Concentration
    - Import Penetration
    - Volatility of Market Shares (1 and 2)
    - Price-Cost Margin
    - Labor Productivity
- Papers:
    - Cheung, C. , Coucke, K. and Neicu, D. (2011). Decision tree structure as screening tool for market malfunctioning
    - Schaumans and Verboven (2011). Entry and Competition in Differentiated Products Markets
    - Cheung, C. and Vanormelingen, S. (2011), Persistence of profits

Selected Turnover: Technical Note

Luc Mariën (FOD Economie)

1. In general

Yearly tables TU_SEL_AGGREGATES_YEAR (from 2000 to 2009) are created with the objective to include calculated variables per company (and also the background variables used for the calculation) that allow the production of values, aggregated at Nace 2, 3 or 4-digit-level, that allow a maximal consistency with aggregated values produced by the National Accounts, that can be considered as an essential reference. The objective is to strengthen the consistency and complementarity between National Accounts data (=aggregates) and the company level data in the sectoral database.

For technical elements on the National Accounts, the NBB publication "De berekeningsmethode voor het Bruto Binnenlands Product en het Bruto Nationaal Inkomen volgens het ESR 1995" is used.

The first variable SELECTED_TRNOV = selected turnover or operating income. This variable is related to the national accounts variable P.1 (Output). This note gives technical elements on its calculation.

2. Important recent elements on Company Accounts data (tables TU_NBB_YEAR)

The actual tables adopt a ventilation of accounting periods data to calendar years data similar to that applied by the national accounts. It takes into account that the big majority of the accounting periodes cover more or less 12 months, but that there are also exceptions (varying between 1 and 64 months).

The ventilation is done as follows:

1) If the start date and the end date of the accounting period fall in the same year, the accounting periods data are ventilated to that calendar year.

2) If the start and the end data belong to 2 consecutive years (say year 1 and year 2):
   a) Either the accounting period covers between 10 and 15 months:
      - if the accounting period covers at least 74% of year 2, the accounting periods data are entirely ventilated to year 2
      - if the accounting period covers at least 74% of year 1, the accounting periods data are entirely ventilated to year 1
      - if neither of the two cases is fulfilled, the accounting periods data are pro rata ventilated over year 1 and year 2 according to the proportion (weight) of each calendar year
   b) Either the accounting period covers 9 months or less: the accounting periods data are entirely ventilated to either year 2 or year 1, depending on the which of the two coincides most with the accounting period

c) Either the accounting period covers between 16 and 24 months: then the accounting periods data are pro rata ventilated over year 1 and year 2 according to the proportion (weight) of each calendar year

3) If the period from the start date to the end date covers 3 consecutive years (say year 1, year 2 and year 3), the ventilation depends on the weight of respectively year 1 and year 3 in the total accounting period.
   a) If as well year 1 as year 3 have both a weight of at least 20%, the accounting data are pro rata ventilated of the three years (according to the respective weights of each year). . If only year 1 and not year 3 has a weight of 20% or more, the ventilation goes to year 1 and 2. In only year 3 and not year 1 has a weight of 20% or more, the ventilation goes to year 2 and 3
   b) If the weight of neither year 1 neither year 3 reaches 20%, the accounting data are entirely attributed to year 2.

4) In the other cases (almost not existant), the accounting data are entirely attributed to the calendar year of the stop date.

Each yearly table has the following three new variables:
- NR_ACCPER : the total number of accounting periods incorporated in the data: in the most of the cases this is 1, in some cases it is two (= the maximum).
- NR_PRORATA: the total number of "pro-rata-calculated" amounts incorporated in the data. In most of the cases this variable is 0. The maximum for this variable = the previous variable (NR_ACCPER). Both variables allow to make the link, if necessary, to the original accounting data as produced by the company.
- CD_SCHM_TYPE (this variable existed before in the TU_BR_ACTIVE_YEAR tables, where it will be omitted):

Values	Signification
1	Abbreviated accounting scheme for companies
2	Complete accounting scheme for companies
4	Abbreviated accounting scheme for associations
5	Complete accounting scheme for associations

3. Selected Turnover

- The selected turnover is calculated by selecting one of four sources, having priority 1 to 4: this means:
  > if source 1 is available, selected turnover equals this one,
  > if source 1 is not available and source 2 is available, selected turnover equals this one,
  > if neither source 1 or 2 are available and source 3 is available, selected turnover equals this one,
  > selected turnover equals source 4 if it's available and if sources 1 to 3 are not available

- The four sources are the following:

  1) COMPACC_TRNOV_TOT = the total operating income based on the yearly company accounts, more precisely the accounts 70 (Turnover) + 71 (Stocks of finished goods and work in progress: increase (decrease) + 72 (Own work capitalised) + 74 (Other operating

income) - 740 (Operating subsidies and compensatory amounts received from public authorities)

For companies with a complete schema, this variables are mandatory, for companies with an abbreviated scheme, they're facultative.

2) SBS_TRNOV_TOT = the operating income based on the yearly SBS-survey (=Structural Business Survey). SBS are available from 2000 to 2008.

3) <u>EXTRAPOL_TOT = the operating income obtained from an extrapolation based on the gross operating income</u>

This turnover is calculated as follows:

a) For each year and each Nace-3-digit-sector, a population of companies is composed with the following characteristics:
   - either it has an abbreviated scheme and its has a turnover figure and a positive gross margin (=account 9900). These companies get the code B1 (in CD_COMP) identical to the national accounts scheme
   - either it has a complete scheme (and registers automatically a turnover) and it is "small" (its yearly turnover doesn't exceed 3 mio euro): these companies receive the code

b) A coefficient (see variable MS_COEFF) is calculated as the total operating income divided by the total gross operating income. This coefficient is calculated for each year and for each Nace-3-digit-sector, except for a number of sectors excluded because of the limited number of companies (generally less than 10) and, related to that, the unreliability (unstability) of the results. The excluded sectors are 017, 089, 091, 099, 104, 120, 143, 142, 192, 202, 206, 211, 235, 241, 244, 254, 264, 266, 267, 268, 272, 301, 302, 03, 352, 353, 390, 492, 495, 501, 512, 531, 643, 652, 653, 68, 783, 799, 803, 822, 841, 842, 854, 871, 872, 881, 970 (for all the years) and 243 (for 2008 and 2009) and 852 and 853 (for 2000 to 2005).

c) A code B2 is given to those companies having an abbreviated scheme, that do not report a turnover but that report a positive gross operating margin. The "extrapolated turnover" of the company is calculated as the gross operating margin multiplied by the MS_COEFF of the Nace-3-digits-sector to which the company belongs.

4) VAT_TRNOV_TOT = turnover based on VAT data

- VAT-units

Data for all the companies called "VAT-Units" has been omitted from the calculation of the selected turnover. "VAT-units" are companies (about 1000 now), started up since 2007 and, still more active in 2008 and 2009, that are created by groups of related companies (their "affiliates") and that are charged with the relationships, for all their affiliates, with the VAT-administration. Examples are "BTW-eenheid Colruyt" or "Procter and Gamble Belgium".

Data on the VAT-turnover from these companies are not taken into account in order to avoid double counting and inconsistencies in the calculation of the selected turnover: indeed, some or all of the affiliates, register already a turnover from other possible sources (company accounts, SBS and/or extrapolation).

- Marketable goods (handelsgoederen / marchandises):

  Like explained on page 142 of the Manual of the SDB, in the National Accounts, the costs related to the purchases and the stock changes of marketable goods are subtracted from total output.

  This is particularly important for the the trade sector in the economy (= trade in cars, wholesale, retail, reparation cars, etc.) (= Nace 50, 51 and 52 (Nace-2003) and 45, 46 and 47 (Nace-2008)). The total output of these sectors, after subtraction of the costs of marketable goods, correspond to their commercial (trade) margins. Also in other sectors, these costs are subtracted from total output, but there it's less important.

  For our comparison between the SDB and NA, we estimated, using SBS figures, these costs for the sectors 45-47 and subtracted it from total output.

- Final results of the comparison SDB - NA: the differences SDB-NA seem reasonable: in general: they turn around 10%. For the years 2006, 2007 and 2008 they are higher (respectively 15,5%, 16,3% and 17%). The yearly growth figures are highly parallel (except for 2002 and 2003).
  (P.S.: version of 16/5/2011 of this paragraph: the differences SDB-NA are remarkably low (generally less than 1% of NA figures). Also the yearly growth figures are highly parallel).

Combination of innovation data with sectoral database: methodological note

For analytical purposes, it is important that all sectoral indicators are based on a sector definition that is consistent over time. Since there is no simple 1-to-1 mapping between NACE rev1.1 and NACE rev.2 (with the latter used from 2008 onwards), the Directorate General Statistics and Economic Information[1] carried out a NACE 'backcasting' exercise in which multiple information sources (Structured Business Survey, PRODCOM, ONSS) are used to assign firms to a NACE rev.2 sector based on a propensity score. This assignment of firms to sectors was done on a yearly basis and resulted in yearly tables of firm-level identifiers linked to the NACE rev.2 code for the firm in that year. These firm-level mapping tables can in principle be used to integrate external firm-level data sources into the sectoral database, ensuring that firms are linked to sectors in the same way as for other data sources.

Also for the Community Innovation Survey data, this approach was used since the CIS4 and CIS2006 surveys use the NACE rev.1.1 classification to designate firms' sector membership while CIS2008 is based on NACE rev.2. Using the conversion tables, the firms in the CIS-data were linked to their NACE rev.2 code as defined by the NACE backcasting exercise. This results in a linkage of firms to sectors using a common classification across the CIS waves, which is also consistent with the other indicators in the sectoral database.

However, the following issues arise with re-assigning firms surveyed in the CIS to NACE rev.2 sectors using the NACE backcasting approach.

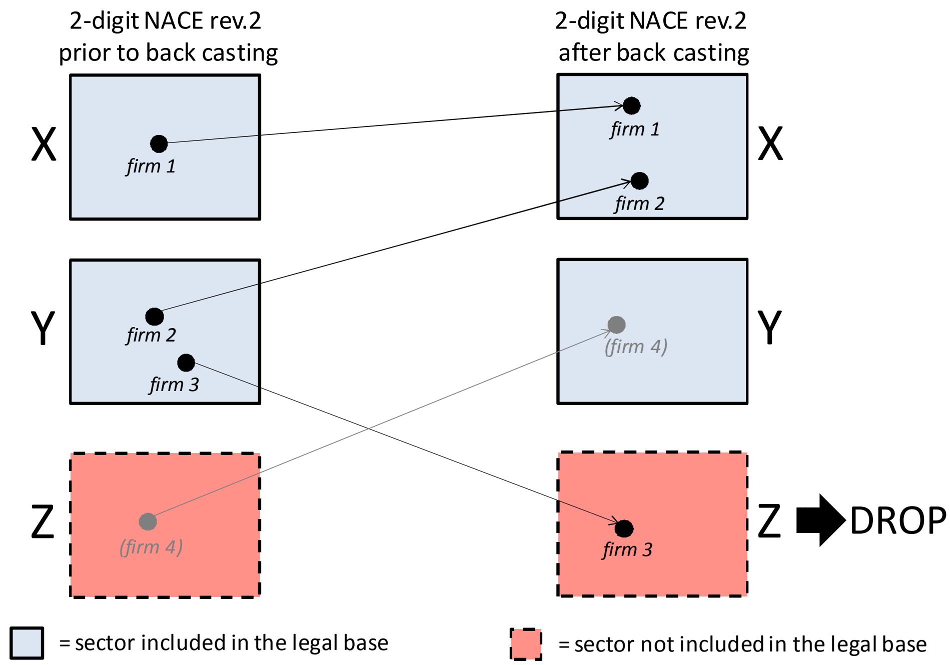
First, ***the NACE backcasting gives rise to a non-representative coverage of sectors given that the CIS survey does not cover the entire economy.*** The set of NACE sectors surveyed for the Community Innovation Survey is based on the Eurostat legal base, which is a subset of the entire economy. The current legal base is defined at the 2-digit NACE rev.2 level and covers the sectors 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 36, 37, 38, 39, 46,49, 50, 51, 52, 53, 58, 61, 62, 63, 64, 65, 66, 71, and 72. These sectors are surveyed in the CIS 2008 survey. A problem may arise with the representativeness of the data for some sectors since our sector-level data is based on firms' NACE rev.2 sector membership according to the NACE backcasting exercise, which may reclassify firms across the boundaries of the legal base. Figure 1 gives an overview of the possible cases for firms in CIS2008.

- The firms in a sector within the legal base (=surveyed in CIS2008) that are reclassified to a sector within the legal base (firm 2 in Figure 1), represent no immediate problem: this is essentially a regrouping of firms in sectors according to NACE rev.2 that is considered to be a more sensible grouping of firms than the previous NACE rev.1.1 classification.[2]

[1] DGSEI, part of the Federal Public Service Economy and in charge of the national statistics in Belgium.
[2] An example of this situation are firms that are reclassified from NACE rev.2 sector 28 (*Manufacture of machinery and equipment*) to NACE rev.2 sector 33 (*Repair and installation of machinery and equipment*).

- The firms in a sector <u>within</u> the legal base that are reclassified to a sector <u>outside</u> of the legal base (firm 3 in Figure 1), lead to a problem of representativeness: since these sectors Z were not surveyed in the CIS, there is no guarantee that the group of firms that is reclassified to such a sector yields a representative picture of the sector composition.[3] *The sectors Z should be excluded from any analysis.*[4]
- The firms in a sector <u>outside</u> of the legal base that *would be* reclassified (if they had been surveyed!) to a sector <u>inside</u> the legal base (firm 4 in Figure 1), also give rise to incomplete coverage of sectors. Since these firms are per definition not observed, the magnitude of the problem cannot be assessed directly although one could assume that it is similar in size to the previous case.

Figure 1: Reclassification of firms (NACE backcasting) in CIS2008



Second, ***the change of the legal base from NACE rev1.1*** (used for CIS4 & CIS2006) ***to NACE rev.2*** (used for CIS2008) ***gives rise to a non-representative coverage of sectors***. The legal base defined in terms of NACE rev.1.1 covers the sectors 10, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 40, 41, 51, 60, 61, 62, 63, 64, 65, 66, 67, 72, 74.2, and 74.3. Figure 2 gives an overview of the possible cases for firms in CIS4 and CIS2006. The first three cases are analogous to the ones for the CIS2008 data.

[3] An example of this situation are firms that are reclassified from NACE rev.2 sector 30 (*Manufacture of other transport equipment*) to NACE rev.2 sector 42 (*Civil engineering*).
[4] In total, 18 sectors at the NACE 4-digit level outside of the NACE rev.2 legal base have a positive firm count after the NACE backcasting exercise for the firms in CIS2008, accounting for 13.5% of all observations at the 4-digit sector level. It concerns NACE rev.2 sectors 41, 42, 43, 45, 47, 60, 68, 70, 73, 74, 77, 78, 79, 80, 81, 82, 92 and 95.

- The firms in a sector within the legal base (=surveyed in CIS2008) that are reclassified to a sector within the legal base (firm 2 in Figure 2), represent no immediate problem: this is essentially a regrouping of firms in sectors according to NACE rev.2 that is considered to be a more sensible grouping of firms than the previous NACE rev.1.1 classification.[5]
- The firms in a sector within the legal base that are reclassified to a sector outside of the legal base (firm 3 in Figure 2), lead to a problem of representativeness: since these sectors Z were not surveyed in the CIS, there is no guarantee that the group of firms that is reclassified to such a sector yields a representative picture of the sector composition.[6] ***The sectors Z should be excluded from any analysis.***
- The firms in a sector outside of the legal base that *would be* reclassified (if they had been surveyed!) to a sector inside the legal base (firm 4 in Figure 2), also give rise to incomplete coverage of sectors. Since these firms are per definition not observed, the magnitude of the problem cannot be assessed directly although one could assume that it is similar in size to the previous case.
- The change of the NACE system implies regroupings of sectors, which combined with the change in the legal base leads to incomplete coverage of certain sectors. More specifically, a certain NACE rev.2 sector may be linked[7] to multiple NACE rev1.1 sectors where at least one of the NACE rev1.1 sectors was not within the legal base i.e. it was not surveyed in CIS4 or CIS2006.[8] ***The sectors Y that are linked to multiple NACE rev1.1 sectors where at least one of the NACE rev1.1 sectors was outside of the legal base should be excluded from analysis***. It concerns 11 NACE rev.2 sectors: 9, 10, 11, 16, 37, 38, 39, 52, 63, 64, 71.

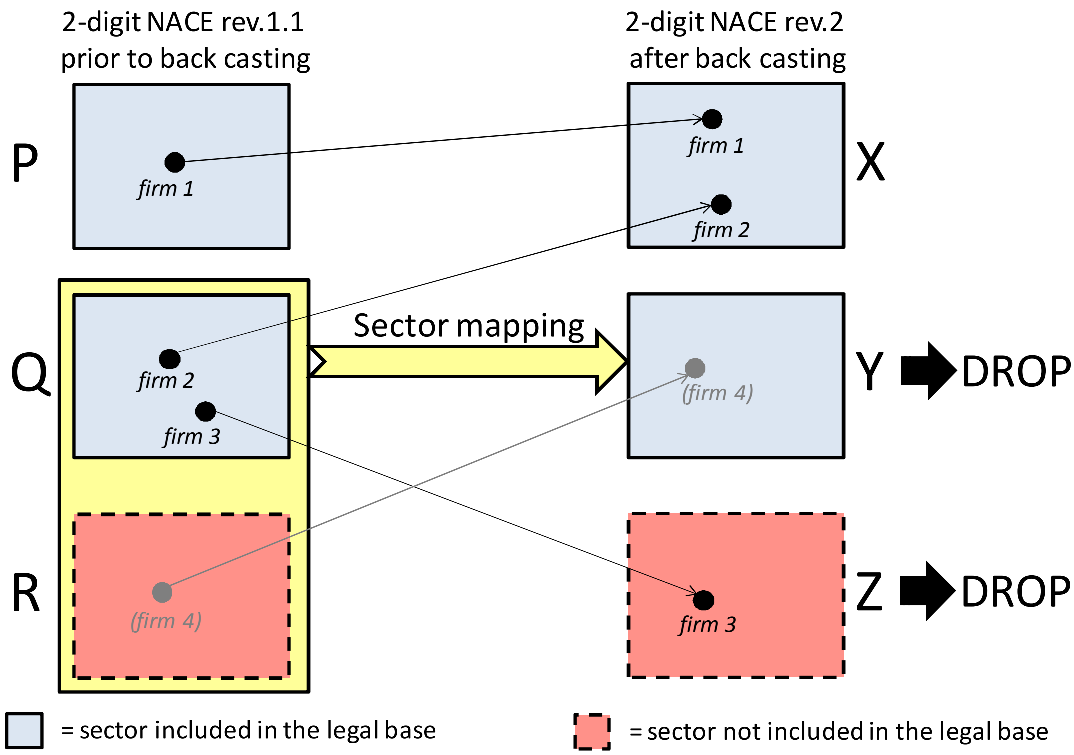
[5] An example of this situation are firms that are reclassified from NACE rev.1.1 sector 22 (*Publishing, printing, and reproduction of recorded media*) to NACE rev.2 sector 17 (*Manufacture of paper and paper products*).

[6] An example of this situation are firms that are reclassified from NACE rev.1.1 sector 15 (*Manufacture of food products and beverages*) to NACE rev.2 sector 47 (*Retail trade, except of motor vehicles and motorcycles*). Another example is NACE rev.1.1 sector 63.3 (Travel Agencies) that were included in the legal base of CIS4/CIS2006 as part of 'support and auxiliary transport activities' within Section I (Transport, storage and communication). The NACE backcasting exercise classifies these firms in NACE rev.2 sector 79, which is outside of the NACE rev.2 legal base.

[7] By 'linked' we mean that the sector-level conversion tables for NACE rev 1.1 and NACE rev.2 contain a mapping between the sectors. This is illustrated by the 'sector mapping' arrow in Figure 2.

[8] An example is NACE rev.2 sector 38 (*Waste collection, treatment and disposal activities; materials recovery*). This sector was surveyed in CIS2008 since it is part of the legal base. The NACE conversion tables indicate that NACE rev1.1 sector 90 (*Sewage and refuse disposal, sanitation and similar activities*) is a related sector in the previous NACE classification. However, NACE rev1.1 sector 90 was not part of the legal base and was therefore not surveyed in CIS4 or CIS2008.

Figure 2: Reclassification of firms (NACE backcasting) in CIS4 & CIS2006



Technical note on the calculation of Benefit of the Doubt scores

using linear programming techniques

Johan Eyckmans (HUBrussel)

June 28, 2011

This technical note describes how we implemented the Benefit of the Doubt composite indicator for the AGORA-MMS project. More information on composite indicators, literature references and so on can be found in the final report of the project. This note is only a complement to the report and is not intended as a standalone or self contained document.

The Benefit of the Doubt (BoD in the sequel) technique is a composite indicator methodology. This means that it is a technique to aggregate the information of several indicators into one single number, a composite indicator score. For the AGORA-MMS project, this means that we have information for sectors (at NACE 2, 3 or 4 level) of structural indicators like for instance concentration, volatility of market shares, price cost margins, … and that we want to aggregate the scores of a particular sector on each of the structural indicators into one single composite indicator score.

Many traditional composite indicators aggregate the information by computing a weighted average of the (normalized) indicator values. It is very common to use the same set of weights for all sectors and to give equal weight to each dimension. Assume that y_s^i denotes the value of indicator i for sector s. The traditional composite indicator score of sector s is given by:

$$CI_s = \sum_i \omega^i \cdot y_s^i$$

Note that the weights of the different indicators ω^i are not indexed on the sectors, hence they are assumed to be the same for all sectors s.

The innovative idea of the BoD aggregation methodology is to allow for more flexibility in the weights. Different indicators can have different weights and the set of weights can be different for different sectors. Hence, the BoD approach relaxes in two important ways the usual restrictions on traditional composite indicators (equal weights for all indicators and equal sets of weights for all sectors). In terms of aggregation, the idea behind the BoD methodology is to give sectors more credit for dimensions they are good in, compared to dimensions they are lagging behind compared to other sectors.

Technically, the calculation of the BoD score for a sector requires solving a linear programming problem. Consider a set of sectors $S=\{1,2,\ldots,\#S\}$ indexed by s or r and a set of indicators $I=\{1,2,\ldots,\#I\}$ indexed by i or j. For practical purposes it is often convenient to consider only a subset of sectors $SS(S)\subseteq S$ and a subsets of indictors $II(I)\subseteq I$ for calculating the BoD scores. For instance, we want to limit the set of peers for a manufacturing sector to the set of manufacturing sectors only (we do not want to compare the steel sector to the sector of hairdressers). Or we want to include only a subset of all possible indicators, for instance because we have no full coverage of the data for some indicators for all sectors.

The BoD score for a particular sector $s\in SS(S)$ is the given by the optimal objective value of the following linear programming problem:

$$\mathrm{BoD}_s = \max_{\{\omega_s^i\}_{i\in II(I)}} \sum_{i\in II(I)} \omega_s^i \cdot y_s^i$$

$$\text{s.t.}\begin{cases} \sum_{i\in II(I)} \omega_s^i \cdot y_r^i \le 1 & r\in SS(S) \qquad [\lambda_r] \\ \underline{w}_s^i \le \omega_s^i \cdot y_r^i \le \overline{w}_s^i & i\in II(I) \\ \omega_s^i \ge 0 & i\in II(I) \end{cases}$$

Note that, compared to traditional composite indicators, the weights ω_s^i are indexed on sectors and hence they can differ across sectors. The linear program seeks a set of weights for the different indicators such that the weighted average for sector s of its indicators' values is maximal, under the constraint that no sector has a score higher than one using the same set of weights (i.e. the first constraint which is a normalization constraint). In addition, it is required that all weights are non-negative (cfr. third constraint) and often it is imposed that the share of a particular indicator in the overall BoD score lies in an interval $\left[\underline{w}_s^i,\overline{w}_s^i\right]$ (cfr. second constraint which is often based on expert opinion or theoretical indications).

In terms of dimension, the typical linear program to be solved has as many decision variables as there are indicators (for instance 6 to 8 in the AGORA-MMS project) and as many normalization restrictions as there are sectors (for instance 100 at NACE 3 level or 200 at NACE 4 level). As such, these are relatively small linear programming problems without too many complications (for instance there are no integer decision variables) which can be solved by standard optimization algorithms (for instance variations on the original simplex algorithm by Dantzig for linear program problems, or more sophisticated modern linear programming solvers like CPLEX). The real technical challenge for the implementation of these problems in SAS is therefore not the solution of the linear programs itself, but more the set up of the different LP problems and the management of the data and results. It requires flexible routines to set up efficiently many different LP problems (one for each sector) with different sets of constraints.

As of today, the BoD implementation used for the AGORA-MMS project is written in GAMS (General Algebraic Modeling System, see www.gams.com), a generic programming language dedicated to solving numerical optimization problems. We included some crucial elements of the GAMS code to illustrate how the LP problems are set up and solved.

Excerpts of GAMS code:

```
xxx
reading and preparing data
defining parameters
xxx
SETS
set I indicators /HHI, CAPINT, CHURN, VOLAT, LPG, PCM, IMPENE, RDINT/ ;
set S sectors / "0111","0112","0113",...,"3900"/ ;

SETS
II(I) subset of active indicators
SS(S) subset of active sectors
;

ALIAS S, S1, S2, S3 ;

VARIABLES
w(s,i)   weight of indicator i for sector s
obj      objective value
;

POSITIVE VARIABLE w(s,i) ;

*** equations
EQUATIONS
E_OBJ              objective equation
```

```
E_CONSTRAINT(s,s1) benchmarking constraints
E_BOUND_lo(i,s)    lower bound on individual indicator
E_BOUND_up(i,s)    upper bound on individual indicator
;

E_OBJ..
      OBJ =E= sum((s,i)$(ss(s) AND ii(i)), d(s)*w(s,i)*y(s,i)) ;
E_CONSTRAINT(s,s1)$(ss(s) AND ss(s1))..
      sum(i$ii(i), d(s)*w(s,i)*y(s1,i)) =l= 1 ;
E_BOUND_lo(i,s)$(ss(s) AND d(s) AND ii(i))..
      w(s,i)*y(s,i) =L= 0.50 ;
E_BOUND_up(i,s)$(ss(s) AND d(s) AND ii(i))..
      w(s,i)*y(s,i) =G= 0.000001 ;

*** models
MODEL BOD /all/ ;

*** begin loop over SECTORS
loop(s3$ss(s3),
    {* initialize membership dummies *}
    d(s2) = 0 ;
    d(s3) = 1 ;
    {* solving model BOD *}
    w.L(s,i) = 0.1 ;
    SOLVE BOD using LP Maximizing OBJ ;
    {* writing output *}
    score(s3,i) = w.L(s3,i)*y(s3,i) ;
    outw(s3,i)  = w.L(s3,i) ;
    outobj(s3)  = obj.L ;
    outpeer(s3,s2) = E_CONSTRAINT.M(s3,s2) ;
    bodstat(s3) = BOD.modelstat ;
) ;
*** end loop over SECTORS
```

```
xxx
writing output
xxx
```

The following remarks should be made.

- In order to construct a general algorithm that can be applied automatically to the full set of sectors and indicators under consideration, the objective value and constraints have been defined using a “membership dummy vector”. For instance `d(s) = (0,0,0,1,0,…,0)` if we want to solve the LP problem for sector 4. The `d(s)` picks the relevant part of the more general objective function and set of constraints (only those constraints referring to sector s that we want to evaluate).
- The actual BoD score of the sectors are given by the value of the objective variable `OBJ` and are recorded for output reporting outside the loop over sectors. The optimal value of the solution is

given in GAMS by the ".L" (L of "level") suffix:

```
outobj(s3) = obj.L ;
```

- The set of peers, i.e. the sectors for which the normalization constraint is binding (i.e. holds with equality) is constructed by using information on the marginal value of the constraint in the optimum. If a sector r is a peer for sector s, it will show up in the solution because the shadow price or multiplier of that particular constraint is nonzero. Hence, the set of peers for sector s is the set of sectors for which the marginal value of the corresponding normalization constraint is nonzero: $P_s(S) = \left\{ r \in SS(S) \middle| \sum_{i \in II(I)} \omega_s^i \cdot y_r^i = 1 \right\} = \left\{ r \in SS(S) \middle| \lambda_r > 0 \right\}$. In the GAMS program we therefore record the value of the slack variables associated with the normalization constraints, i.e. the marginal values (".M" suffix in GAMS).

```
outpeer(s3,s2) = E_CONSTRAINT.M(s3,s2))i
```

- The actual implementation in GAMS is more complicated because we solve BoD problems for every year between 2001 and 2009. Hence, the excerpt of the GAMS code above is embedded in an additional loop over at set of years:

```
SET YEAR years /2001,2002,2003,2004,2005,2006,2007,2008,2009/ ;
```

- It is important to keep track of the status of the solution (infeasible, optimal solution found, …) in order to check whether the problems have been solved correctly. This information is recorded in GAMS in the "modelstat" (model status) variable. A value of "1" for modelstat means that the LP program has been solved correctly (no infeasibilities, no convergence problems and so on). Other values than "1" are indications of non-optimal solutions.

```
bodstat(s3) = BOD.modelstat ;
```

- The typical solution time for 9 years of data, 100 NACE 3 manufacturing sectors and 8 indicators (i.e. 900 LP problems of 8 decision variables and 125 constraints each) is about 15 minutes on a standard PC.
- It is important to warn against "mechanical" implementation of the BoD methodology. In the process of solving the LP problems, many things can go wrong (for instance, the lower bound constraints $\underline{w}_s^i \leq \omega_s^i \cdot y_s^i$ become infeasible when $y_s^i = 0$ and $\underline{w}_s^i > 0$, hence indicators with zero values are problematic when combined with lower bound constraints). The analysist should always carefully check the detailed output of the optimization software in order to detect possible anomalies. We therefore have to warn against "push the button" implementations of the BoD methodology.

- For completeness, we have included all GAMS programs for a typical BoD problem in the AGORA-MMS project, in particular for the manufacturing sectors (95 at NACE 3) for all 8 indicators and all 9 years for which data is available:
    - `COMPIND.GMS`:
      main GAMS program (DOS command line "GAMS COMPIND.GMS PS=9999")
    - `DATA3.INC`:
      include file for including and preparing data in which the set of indicators, sectors and years has to be chosen by the user
    - `data_NACE3_2001.TXT` to `data_NACE3_2009.TXT`:
      text files containing data for all sectors and indicators for years 2001 to 2009
- Output is gathered in different text files that can easily be imported in Excel for editing and reporting.
    - `EXCEL_BOD.TXT`:
      output of the different BOD scores for all sectors (rows) and years (columns)
    - `DETAIL_BOD.TXT`:
      weight or load of every indicator (columns) for every sector (rows) and every year (tables are appended from 2001 to 2009)
    - `PEERS.TXT`:
      overview of all peers (columns) for all sectors (rows) and years tables are appended from 2001 to 2009)

```
$ontext
===========================================================================
Benefit of the Doubt composite indicator
input: indicator data
output: weights and composite indicators and rankings
===========================================================================
(c) 2011 Johan Eyckmans
version 25062011
===========================================================================
$offtext

$TITLE MARKET FUNCTIONING MONITORING TOOL

$inlinecom {* *}
$offupper
$offsymxref offsymlist offuellist offuelxref



**************************************
*** set definitions and data input ***
**************************************
* set definitions and raw data input
$batinclude data3.inc ;


******************
*** parameters ***
******************

PARAMETERS
d(s)             membership dummy sectors
y(s,i)           value for sector s of indicator i
score(s,i)       output score of sector s for indicator i
outw(s,i)        output weight of sector s for indicator i
outobj(s)        output objective function sector s
outpeer(s,s)     output peers sector s
bodstat(s)       model status for BOD
data(s,i,*)      data table
restriction      restrictions dummy
peernum(s)       number of peers
CI(s,*)          composite indicator score for sector s
CIR(s,*)         composite indicator rank of sector s
we(s,i)          weight of sector s for indicator i
yn(s,i)          normalized indicator of sector s for indicator i
ymin(i)          minimum indicator value
ymax(i)          maximum indicator value
ys(s)            sorted indicator
rank(s)          rank
order(s)         order
tel              teller
missing(S,I)     dummy missing value for indicator i
yaver(i)         average indicator value
ystdev(i)        standard deviation indicator value
xCI(s,*,year)    composite indicator score for sector s in year
xCIR(s,*,year)   composite indicator rank of sector s in year
perc             percentage
ytemp(s,i)       temporary variable
cow              column wide
dec              decimals
xpeer(s,s1,year) peers
xpeernum(s,year) number of peers
xmis(s,year)     missing observations
;

perc = 0.25 ;
cow = 10 ;
dec = 4 ;
xmis(s,year) = 0 ;
xmis(s,year)$(not ss(s)) = 1 ;


*********************
*** raw data input ***
*********************
```

```
* data tables per year
$batinclude data_NACE3_2001.txt ;
$batinclude data_NACE3_2002.txt ;
$batinclude data_NACE3_2003.txt ;
$batinclude data_NACE3_2004.txt ;
$batinclude data_NACE3_2005.txt ;
$batinclude data_NACE3_2006.txt ;
$batinclude data_NACE3_2007.txt ;
$batinclude data_NACE3_2008.txt ;
$batinclude data_NACE3_2009.txt ;


* choose one year
y(S,I) = 999999 ;
y(S,I) = indicators2001(S,I) ;


******************
*** variables ****
******************

VARIABLES
w(s,i)   weight of indicator i for sector s
obj      objective value
;

POSITIVE VARIABLES
w(s,i)
;

*** equations
EQUATIONS
E_OBJ               objective equation
E_CONSTRAINT(s,s1) benchmarking constraints
E_BOUND_lo(i,s)    lower bound on individual indicator
E_BOUND_up(i,s)    upper bound on individual indicator
E_BOUND_STRU(s)    relative bound on weight for STRUCTURE dimension
E_BOUND_COND(s)    relative bound on weight for CONDUCT dimension
E_BOUND_PERF(s)    relative bound on weight for PERFORMANCE dimension
;

E_OBJ..               OBJ =E= sum((s,i)$(ss(s) AND ii(i)), d(s)*w(s,i)*y(s,i)) ;
E_CONSTRAINT(s,s1)$(ss(s) AND ss(s1))..
                      sum(i$ii(i), d(s)*w(s,i)*y(s1,i)) =l= 1 ;
E_BOUND_lo(i,s)$(ss(s) AND d(s) AND ii(i))..
                      w(s,i)*y(s,i) =L= 0.50 ;
E_BOUND_up(i,s)$(ss(s) AND d(s) AND ii(i))..
                      w(s,i)*y(s,i) =G= 0.000001 ;
E_BOUND_STRU(s)$(ss(s) AND d(s))..
                      sum(i$STRU(i), w(s,i)*y(s,i)) =G= (1/5)*sum(i, w(s,i)*y(s,i)) ;
E_BOUND_COND(s)$(ss(s) AND d(s))..
                      sum(i$COND(i), w(s,i)*y(s,i)) =G= (1/5)*sum(i, w(s,i)*y(s,i)) ;
E_BOUND_PERF(s)$(ss(s) AND d(s))..
                      sum(i$PERF(i), w(s,i)*y(s,i)) =G= (1/5)*sum(i, w(s,i)*y(s,i)) ;



**************
*** models ***
**************

*MODEL BOD /all/ ;
*MODEL BOD /E_OBJ, E_CONSTRAINT/ ;
MODEL BOD /E_OBJ, E_CONSTRAINT, E_BOUND_lo/ ;
*MODEL BOD /E_OBJ, E_CONSTRAINT, E_BOUND_lo, E_BOUND_up/ ;
*MODEL BOD /E_OBJ, E_CONSTRAINT, E_BOUND_is, E_BOUND_STRU, E_BOUND_COND, E_BOUND_PERF/ ;



**********************
*** solver options ***
**********************

OPTION optcr    = 0 ;
OPTION iterlim  = 1000000 ;
OPTION reslim   = 1000000 ;
```

```
OPTION LIMROW   = 5 ;
OPTION LIMCOL   = 5 ;
OPTION SOLPRINT = OFF ;
option decimals = 6 ;

* output BOD
file detail_BOD /detail_BOD.txt/ ;
detail_BOD.PW = 150 ;
*detail_BOD.ap = 1 ;


************************
*** begin loop YEARS ***
************************
loop(year$yy(year),

* reconstruct base set of sectors
ss(s) = NO ;
ss(s)$show(s) = YES ;

*** loading data
y(S,I) = 999999 ;
if(ord(YEAR) EQ 1, y(S,I) = indicators2001(S,I)) ;
if(ord(YEAR) EQ 2, y(S,I) = indicators2002(S,I)) ;
if(ord(YEAR) EQ 3, y(S,I) = indicators2003(S,I)) ;
if(ord(YEAR) EQ 4, y(S,I) = indicators2004(S,I)) ;
if(ord(YEAR) EQ 5, y(S,I) = indicators2005(S,I)) ;
if(ord(YEAR) EQ 6, y(S,I) = indicators2006(S,I)) ;
if(ord(YEAR) EQ 7, y(S,I) = indicators2007(S,I)) ;
if(ord(YEAR) EQ 8, y(S,I) = indicators2008(S,I)) ;
if(ord(YEAR) EQ 9, y(S,I) = indicators2009(S,I)) ;

*** data manipulation
* detecting missing values
missing(S,I) = 0 ;
loop(I$ii(i),
   loop(S$ss(s),
       if(y(S,I) GE 9999998,
           missing(S,I) = 1 ;
       else
           missing(S,I) = 0 ;
       ) ;
   ) ;
) ;

* drop sectors for which there are missing values
ss(s)$(sum(i$ii(i), missing(s,i)) GE 1) = NO ;
xmis(s,year) = 1 ;
xmis(s,year)$ss(s) = 0 ;

* all indicators should be "goods", not "bads"

* high concentration is bad: inverse transformation
*y(s,"c4")$ss(s)   = 1 / y(s,"c4") ;
*y(s,"c8")$ss(s)   = 1 / y(s,"c8") ;
* inverse transformation IS NOT NEUTRAL for BOD
*y(s,"hhin")$ss(s) = 1 / y(s,"hhin") ;
* linear transformation
y(s,"hhin")$ss(s) = smax(s1$ss(s1), y(s1,"hhin")) - y(s,"hhin") + 1;

* high CAPINT is bad:
* inverse transformation
*y(s,"capint")$ss(s) = 1 / y(s,"capint") ;
* linear transformation
y(s,"capint")$ss(s) = smax(s1$ss(s1), y(s1,"capint")) - y(s,"capint") + 1 ;

* high MES is bad:
* inverse transformation
*y(s,"MES")$ss(s) = 1 / y(s,"MES") ;
* linear transformation
y(s,"MES")$ss(s) = smax(s1$ss(s1), y(s1,"MES")) - y(s,"MES") + 1 ;

* high DLP is good
* but deduct minimum to convert to positive numbers
y(s,"dlp")$ss(s) = y(s,"dlp") - smin(s1$ss(s1), y(s1,"dlp")) + 1 ;
```

```
* high PCM is bad
* deduct minimum to convert to positive numbers
y(s,"pcm")$ss(s) = y(s,"pcm") - smin(s1$ss(s1), y(s1,"pcm")) + 1 ;
* inverse transformation
*y(s,"pcm")$ss(s) = 1 / (y(s,"pcm")+1) ;
* linear transformation
*y(s,"pcm")$ss(s) = smax(s1$ss(s1), y(s1,"pcm")) - y(s,"pcm") ;

* high RD is good

*** Benefit of the doubt LP programs

* for loglinear specication
if(LOGLINEAR, y(s,i)$(ii(i) AND ss(s)) = log(y(s,i))) ;

*******************************
*** begin loop over SECTORS ***
*******************************
loop(s3$ss(s3),
    {* initialize membership dummies *}
    d(s2) = 0 ;
    d(s3) = 1 ;
    {* solving model BOD *}
    w.L(s,i) = 0.1 ;
    SOLVE BOD using LP Maximizing OBJ ;
    {* writing output *}
    score(s3,i) = w.L(s3,i)*y(s3,i) ;
    outw(s3,i)  = w.L(s3,i) ;
    outobj(s3)  = obj.L ;
    outpeer(s3,s2) = E_CONSTRAINT.M(s3,s2) ;
    bodstat(s3) = BOD.modelstat ;
) ;
*****************************
*** end loop over SECTORS ***
*****************************

*** anti log
if(LOGLINEAR, y(s,i)$(ii(i)  AND ss(s)) = exp(y(s,i))) ;

* display solution in listing file
display outw, score, outobj, outpeer ;
outpeer(s,s1)$(outpeer(s,s1) GT EPS) = 1 ;
outpeer(s,s1)$(outpeer(s,s1) LE EPS) = 0 ;
display outpeer ;
peernum(s) = sum(s1, outpeer(s1,s)) ;
display peernum ;

* store BOD
xCI(s,"BOD",year) = round(sum(i$ii(i), score(s,i)),6) ;
xCI(s,"BOD",year)$(NOT ss(s)) = 999999 ;
xpeer(s,s1,year) = outpeer(s,s1) ;
xpeernum(s,year) = peernum(s) ;

* output in detailed filed
put detail_BOD ;
put year.TL:>cow / ;
put "sector":<cow ;
loop(i$ii(i), put i.TL:>cow ) ;
put "BOD":>cow ;
put "peer":>cow ;
put "test":>cow ;
put / ;
loop(s$ss(s),
    put s.TL:<cow ;
    loop(i$ii(i),
        put score(s,i):cow:dec ;
    ) ;
    put xCI(s,"BOD",year):cow:dec ;
    put xpeernum(s,year):cow:0 ;
    put BODstat(s):cow:0 ;
    put / ;
) ;
put  //  ;

* ordinary arithmetic average z-score normalized data
yaver(i) = sum(s$ss(s), y(s,i)) / card(ss) ;
```

```
ystdev(i) = sqrt((1/card(ss))*sum(s$ss(s), (y(s,i)-yaver(i))*(y(s,i)-yaver(i)) )) ;
yn(s,i) = (y(s,i) - yaver(i)) / ystdev(i) ;
we(s,i) = 1 / card(ii) ;
xCI(s,"STDEV",year) = sum(i$ii(i), we(s,i)*yn(s,i)) ;
xCI(s,"STDEV",year)$(NOT ss(s)) = 999999 ;

* ordinary arithmetic average minmax
ymin(i) = smin(s$ss(s), y(s,i)) ;
ymax(i) = smax(s$ss(s), y(s,i)) ;
yn(s,i) = (y(s,i) - ymin(i)) / (ymax(i) - ymin(i)) ;
we(s,i) = 1 / card(ii) ;
xCI(s,"MINMAX",year) = sum(i$ii(i), we(s,i)*yn(s,i)) ;
xCI(s,"MINMAX",year)$(NOT ss(s)) = 999999 ;

) ;
**********************
*** end loop YEARS ***
**********************

* display
display xCI ;

* write to txt files
* output BOD
file excel_BOD /excel_BOD.txt/ ;
excel_BOD.PW = 150 ;
put excel_BOD ;
*excel_BOD.ap = 1 ;

put / ;
put @(cow+1);
loop(year$yy(year),
    put year.TL:>cow ;
) ;
put / ;
loop(s$(prod(year$yy(year), xmis(s,year)) EQ 0),
    put s.TL:<cow ;
    loop(year$yy(year),
        if(xCI(s,"BOD",year) NE 999999,
            put xCI(s,"BOD",year):cow:dec ;
        else
            put "n.a.":>cow ;
        ) ;
    ) ;
    put / ;
) ;
put / ;

* output STDEV
file excel_STDEV /excel_STDEV.txt/ ;
excel_STDEV.PW = 150 ;
put excel_STDEV ;
*excel_STDEV.ap = 1 ;

put / ;
put @(cow+1);
loop(year$yy(year),
    put year.TL:>cow ;
) ;
put / ;
loop(s$(prod(year$yy(year), xmis(s,year)) EQ 0),
    put s.TL:<cow ;
    loop(year$yy(year),
        if(xCI(s,"STDEV",year) NE 999999,
            put xCI(s,"STDEV",year):cow:dec ;
        else
            put "n.a.":>cow ;
        ) ;
    ) ;
    put / ;
) ;
put / ;

* output MINMAX
file excel_MINMAX /excel_MINMAX.txt/ ;
excel_MINMAX.PW = 150 ;
```

```
put excel_MINMAX ;
*excel_MINMAX.ap = 1 ;

put / ;
put @(cow+1);
loop(year$yy(year),
    put year.TL:>cow ;
) ;
put / ;
loop(s$(prod(year$yy(year), xmis(s,year)) EQ 0),
    put s.TL:<cow ;
    loop(year$yy(year),
        if(xCI(s,"MINMAX",year) NE 999999,
            put xCI(s,"MINMAX",year):cow:dec ;
        else
            put "n.a.":>cow ;
        ) ;
    ) ;
    put / ;
) ;
put / ;

* output peers
file peers /peers.txt/ ;
peers.PW=150 ;
put peers ;
*peers.ap = 1 ;

put / ;
loop(year$yy(year),
    put year.TL:>cow ;
    put / ;
    put @6 ;
    loop(s$ss(s), put$(xpeernum(s,year) GT 0) s.TL:>5) ;
    put / ;
    loop(s$ss(s),
        put s.TL:<5 ;
        loop(s1$ss(s1),
            if(xpeernum(s1,year) GT 0,
                put$(not xpeer(s,s1,year)) "     " ;
                put$xpeer(s,s1,year) 1:5:0 ;
            ) ;
        ) ;
        put / ;
    ) ;
    put @6 ;
    loop(s$ss(s), put$(xpeernum(s,year) GT 0) xpeernum(s,year):5:0) ;
    put /// ;
) ;
put  //  ;


***************************************************************************

$label END

***************************************************************************
```

```
$ontext
================================================================================
DATA3.TXT
set definitions and data input
=> choose data file and year
=> include import penetration (impene) or not
================================================================================
(c) 2011 Johan Eyckmans
version 13062011
================================================================================
$offtext


******************
*** parameters ***
******************

parameter LOGLINEAR ;
LOGLINEAR = 0 ;


************
*** sets ***
************

*** set of YEARS
set YEAR years
/2001,2002,2003,2004,2005,2006,2007,2008,2009/ ;

* alias
alias(year,year1,year2) ;

* subsets of years
set yy(year) active years ;
yy(year) = YES ;
*yy("2007") = YES ;
*yy("2008") = YES ;
*yy("2009") = YES ;

*** set of INDICATORS
set I indicators
/c4,
 c8,
 hhin,
 capint,
 mes,
 churn,
 volat,
 dlp,
 pcm,
 impene,
 rd
/ ;

* alias
alias(I,I1,I2) ;

* subsets of indicators
sets
STRU(I) STRUcture subset of indicators
COND(I) CONDuct subset of indicators
PERF(I) PERFormance subset of indicators
;
STRU(I)          = NO ;
STRU("hhin")     = YES ;
STRU("capint")   = YES ;
STRU("mes")      = YES ;
STRU("churn")    = YES ;
STRU("impene")   = YES ;
STRU("rd")       = YES ;
COND(I)          = NO ;
COND("volat")    = YES ;
PERF(I)          = NO ;
PERF("pcm")      = YES ;
PERF("dlp")      = YES ;
```

```
*** set of SECTORS
* this set contains all sectors that are in the raw data file
set S sectors
/
"000",
"011",
"012",
"013",
"014",
"015",
"016",
"017",
"021",
"022",
"023",
"024",
"031",
"032",
"051",
"071",
"072",
"081",
"089",
"091",
"099",
"100",
"101",
"102",
"103",
"104",
"105",
"106",
"107",
"108",
"109",
"110",
"120",
"131",
"132",
"133",
"139",
"141",
"142",
"143",
"151",
"152",
"157",
"161",
"162",
"171",
"172",
"173",
"181",
"182",
"191",
"192",
"201",
"202",
"203",
"204",
"205",
"206",
"211",
"212",
"221",
"222",
"231",
"232",
"233",
"234",
"235",
"236",
"237",
"239",
"241",
```

```
"242",
"243",
"244",
"245",
"251",
"252",
"253",
"254",
"255",
"256",
"257",
"259",
"261",
"262",
"263",
"264",
"265",
"266",
"267",
"268",
"269",
"271",
"272",
"273",
"274",
"275",
"279",
"281",
"282",
"283",
"284",
"289",
"291",
"292",
"293",
"299",
"301",
"302",
"303",
"304",
"309",
"310",
"321",
"322",
"323",
"324",
"325",
"329",
"331",
"332",
"351",
"352",
"353",
"360",
"370",
"381",
"382",
"383",
"390",
"399",
"411",
"412",
"421",
"422",
"429",
"431",
"432",
"433",
"439",
"451",
"452",
"453",
"454",
"460",
"461",
"462",
```

```
"463",
"464",
"465",
"466",
"467",
"469",
"471",
"472",
"473",
"474",
"475",
"476",
"477",
"478",
"479",
"492",
"493",
"494",
"495",
"501",
"502",
"503",
"504",
"511",
"512",
"521",
"522",
"531",
"532",
"551",
"552",
"553",
"559",
"561",
"562",
"563",
"581",
"582",
"591",
"592",
"601",
"602",
"611",
"612",
"613",
"619",
"620",
"631",
"639",
"641",
"642",
"643",
"649",
"651",
"652",
"653",
"660",
"661",
"662",
"663",
"681",
"682",
"683",
"691",
"692",
"701",
"702",
"711",
"712",
"721",
"722",
"731",
"732",
"741",
"742",
"743",
```

```
"749",
"750",
"771",
"772",
"773",
"774",
"781",
"782",
"783",
"791",
"799",
"801",
"802",
"803",
"811",
"812",
"813",
"821",
"822",
"823",
"829",
"841",
"842",
"843",
"851",
"852",
"853",
"854",
"855",
"856",
"861",
"862",
"869",
"871",
"872",
"873",
"879",
"881",
"889",
"900",
"910",
"920",
"931",
"932",
"941",
"942",
"949",
"951",
"952",
"960",
"970",
"981",
"982",
"990",
"999"
/ ;

* alias
alias(S,S1,S2,S3) ;

* subsets of sectors
set AGRIC(S) AGRICulture ;
AGRIC(S) = NO ;
*AGRIC("000") = YES ;
AGRIC("011") = YES ;
AGRIC("012") = YES ;
AGRIC("013") = YES ;
AGRIC("014") = YES ;
AGRIC("015") = YES ;
AGRIC("016") = YES ;
AGRIC("017") = YES ;
AGRIC("021") = YES ;
AGRIC("022") = YES ;
AGRIC("023") = YES ;
AGRIC("024") = YES ;
AGRIC("031") = YES ;
```

```
AGRIC("032") = YES ;
set EXTRA(S) EXTRAction ;
EXTRA(S) = NO ;
EXTRA("051") = YES ;
EXTRA("071") = YES ;
EXTRA("072") = YES ;
EXTRA("081") = YES ;
EXTRA("089") = YES ;
EXTRA("091") = YES ;
EXTRA("099") = YES ;
set MANUF(S) MANUFacturing ;
MANUF(S) = NO ;
*MANUF("100") = YES ;
MANUF("101") = YES ;
MANUF("102") = YES ;
MANUF("103") = YES ;
MANUF("104") = YES ;
MANUF("105") = YES ;
MANUF("106") = YES ;
MANUF("107") = YES ;
MANUF("108") = YES ;
MANUF("109") = YES ;
MANUF("110") = YES ;
MANUF("120") = YES ;
MANUF("131") = YES ;
MANUF("132") = YES ;
MANUF("133") = YES ;
MANUF("139") = YES ;
MANUF("141") = YES ;
MANUF("142") = YES ;
MANUF("143") = YES ;
MANUF("151") = YES ;
MANUF("152") = YES ;
*MANUF("157") = YES ;
MANUF("161") = YES ;
MANUF("162") = YES ;
MANUF("171") = YES ;
MANUF("172") = YES ;
*MANUF("173") = YES ;
MANUF("181") = YES ;
MANUF("182") = YES ;
MANUF("191") = YES ;
MANUF("192") = YES ;
MANUF("201") = YES ;
MANUF("202") = YES ;
MANUF("203") = YES ;
MANUF("204") = YES ;
MANUF("205") = YES ;
MANUF("206") = YES ;
MANUF("211") = YES ;
MANUF("212") = YES ;
MANUF("221") = YES ;
MANUF("222") = YES ;
MANUF("231") = YES ;
MANUF("232") = YES ;
MANUF("233") = YES ;
MANUF("234") = YES ;
MANUF("235") = YES ;
MANUF("236") = YES ;
MANUF("237") = YES ;
MANUF("239") = YES ;
MANUF("241") = YES ;
MANUF("242") = YES ;
MANUF("243") = YES ;
MANUF("244") = YES ;
MANUF("245") = YES ;
MANUF("251") = YES ;
MANUF("252") = YES ;
MANUF("253") = YES ;
MANUF("254") = YES ;
MANUF("255") = YES ;
MANUF("256") = YES ;
MANUF("257") = YES ;
MANUF("259") = YES ;
MANUF("261") = YES ;
MANUF("262") = YES ;
```

```
MANUF("263") = YES ;
MANUF("264") = YES ;
MANUF("265") = YES ;
MANUF("266") = YES ;
MANUF("267") = YES ;
MANUF("268") = YES ;
*MANUF("269") = YES ;
MANUF("271") = YES ;
MANUF("272") = YES ;
MANUF("273") = YES ;
MANUF("274") = YES ;
MANUF("275") = YES ;
MANUF("279") = YES ;
MANUF("281") = YES ;
MANUF("282") = YES ;
MANUF("283") = YES ;
MANUF("284") = YES ;
MANUF("289") = YES ;
MANUF("291") = YES ;
MANUF("292") = YES ;
MANUF("293") = YES ;
*MANUF("299") = YES ;
MANUF("301") = YES ;
MANUF("302") = YES ;
MANUF("303") = YES ;
MANUF("304") = YES ;
MANUF("309") = YES ;
MANUF("310") = YES ;
MANUF("321") = YES ;
MANUF("322") = YES ;
MANUF("323") = YES ;
MANUF("324") = YES ;
MANUF("325") = YES ;
MANUF("329") = YES ;
MANUF("331") = YES ;
MANUF("332") = YES ;
set UTILE(S) UTILities and Energy ;
UTILE(S) = NO ;
UTILE("351") = YES ;
UTILE("352") = YES ;
UTILE("353") = YES ;
UTILE("360") = YES ;
UTILE("370") = YES ;
UTILE("381") = YES ;
UTILE("382") = YES ;
UTILE("383") = YES ;
UTILE("390") = YES ;
*UTILE("399") = YES ;
set CONST(S) CONSTruction ;
CONST(S) = NO ;
CONST("411") = YES ;
CONST("412") = YES ;
CONST("421") = YES ;
CONST("422") = YES ;
CONST("429") = YES ;
CONST("431") = YES ;
CONST("432") = YES ;
CONST("433") = YES ;
CONST("439") = YES ;
set TRADE(S) TRADE ;
TRADE(S) = NO ;
TRADE("451") = YES ;
TRADE("452") = YES ;
TRADE("453") = YES ;
TRADE("454") = YES ;
TRADE("460") = YES ;
TRADE("461") = YES ;
TRADE("462") = YES ;
TRADE("463") = YES ;
TRADE("464") = YES ;
TRADE("465") = YES ;
TRADE("466") = YES ;
TRADE("467") = YES ;
TRADE("469") = YES ;
TRADE("471") = YES ;
TRADE("472") = YES ;
```

```
TRADE("473") = YES ;
TRADE("474") = YES ;
TRADE("475") = YES ;
TRADE("476") = YES ;
TRADE("477") = YES ;
TRADE("478") = YES ;
TRADE("479") = YES ;
set TRANS(S) TRANSport ;
TRANS(S) = NO ;
TRANS("492") = YES ;
TRANS("493") = YES ;
TRANS("494") = YES ;
TRANS("495") = YES ;
TRANS("501") = YES ;
TRANS("502") = YES ;
TRANS("503") = YES ;
TRANS("504") = YES ;
TRANS("511") = YES ;
TRANS("512") = YES ;
TRANS("521") = YES ;
TRANS("522") = YES ;
TRANS("531") = YES ;
TRANS("532") = YES ;
set PSERV(S) Personal SERVices ;
PSERV(S) = NO ;
PSERV("551") = YES ;
PSERV("552") = YES ;
PSERV("553") = YES ;
PSERV("559") = YES ;
PSERV("561") = YES ;
PSERV("562") = YES ;
PSERV("563") = YES ;
PSERV("951") = YES ;
PSERV("952") = YES ;
PSERV("960") = YES ;
set BSERV(S) Business SERVices ;
BSERV(S) = NO ;
BSERV("581") = YES ;
BSERV("582") = YES ;
BSERV("591") = YES ;
BSERV("592") = YES ;
BSERV("601") = YES ;
BSERV("602") = YES ;
BSERV("611") = YES ;
BSERV("612") = YES ;
BSERV("613") = YES ;
BSERV("619") = YES ;
BSERV("620") = YES ;
BSERV("631") = YES ;
BSERV("639") = YES ;
BSERV("641") = YES ;
BSERV("642") = YES ;
BSERV("643") = YES ;
BSERV("649") = YES ;
BSERV("651") = YES ;
BSERV("652") = YES ;
BSERV("653") = YES ;
BSERV("660") = YES ;
BSERV("661") = YES ;
BSERV("662") = YES ;
BSERV("663") = YES ;
BSERV("681") = YES ;
BSERV("682") = YES ;
BSERV("683") = YES ;
BSERV("691") = YES ;
BSERV("692") = YES ;
BSERV("701") = YES ;
BSERV("702") = YES ;
BSERV("711") = YES ;
BSERV("712") = YES ;
BSERV("721") = YES ;
BSERV("722") = YES ;
BSERV("731") = YES ;
BSERV("732") = YES ;
BSERV("741") = YES ;
BSERV("742") = YES ;
```

```
BSERV("743") = YES ;
BSERV("749") = YES ;
BSERV("750") = YES ;
BSERV("771") = YES ;
BSERV("772") = YES ;
BSERV("773") = YES ;
BSERV("774") = YES ;
BSERV("781") = YES ;
BSERV("782") = YES ;
BSERV("783") = YES ;
BSERV("791") = YES ;
BSERV("799") = YES ;
BSERV("801") = YES ;
BSERV("802") = YES ;
BSERV("803") = YES ;
BSERV("811") = YES ;
BSERV("812") = YES ;
BSERV("813") = YES ;
BSERV("821") = YES ;
BSERV("822") = YES ;
BSERV("823") = YES ;
BSERV("829") = YES ;
set OTHER(S) other sectors ;
OTHER(S) = YES ;
OTHER(S) = OTHER(S) - AGRIC(S) - EXTRA(S) - MANUF(S) - UTILE(S) -
           CONST(S) - TRADE(S) - TRANS(S) - PSERV(S) - BSERV(S) ;

display AGRIC, EXTRA, MANUF, UTILE, CONST, TRADE, TRANS, PSERV,
        BSERV, OTHER ;


* subset of sectors
set ss(s)   subsample of sectors ;
set show(s) sectors to be displayed ;
alias (ss,ss1,ss2,ss3) ;

* subset of indicators
set ii(i) subsample of indicators ;
alias (ii,ii1,ii2,ii3) ;

*** choosing indicators
ii(i) = NO ;
*ii("C4")    = YES ;
*ii("C8")    = YES ;
ii("HHIN")  = YES ;
ii("CAPINT")= YES ;
*ii("MES")   = YES ;
ii("CHURN") = YES ;
ii("VOLAT") = YES ;
ii("DLP")   = YES ;
ii("PCM")   = YES ;
ii("IMPENE")= YES ;
ii("RD")    = YES ;
display ii ;

*** choosing sectors
ss(s) = NO ;
*ss(s)$AGRIC(s) = YES ;
ss(s)$EXTRA(s) = YES ;
ss(s)$MANUF(s) = YES ;
ss(s)$UTILE(s) = YES ;
*ss(s)$CONST(s) = YES ;
*ss(s)$TRADE(s) = YES ;
*ss(s)$TRANS(s) = YES ;
*ss(s)$PSERV(s) = YES ;
*ss(s)$BSERV(s) = YES ;
*ss(s)$OTHER(s) = YES ;
display ss ;
show(s) = NO ;
show(s)$ss(s) = YES ;
display show ;


***************************************************************************
*** end of data3.inc
***************************************************************************
```

```
$ontext
"CD_NACE3"
"MS_C4_DT_2009"
"MS_C8_DT_2009"
"MS_HHI_NORM_DT_2009"
"MS_W_CAPINT_TRN_2009"
"MES_2009"
"MS_CHURN_2009_WG"
"MS_VOLAT_IDX_2009"
"MS_W_LP_CH_2009"
"MS_W_PCM_2009"
"MS_IMPEN_2009"
"RD_INT_2009"
$offtext
TABLE indicators2009(s,i)
            C4      C8    HHIN  CAPINT     MES   CHURN   VOLAT     DLP     PCM  IMPENE      RD
"011"  0.06330 0.09009 0.00205 0.33589 0.00045 0.04587 0.26437 0.09228 0.10550 0.71063 9999999
"012"  0.05691 0.09306 0.00236 0.68201 0.00187 0.08307 0.62563 -.01667 0.11454 0.90669 9999999
"013"  0.16191 0.21725 0.01042 0.18814 0.00255 0.03573 0.13940 0.00191 0.05292 0.51698 9999999
"014"  0.17380 0.23831 0.01085 0.26066 0.00045 0.03808 0.21150 0.14031 0.02540 0.26377 9999999
"015"  0.01434 0.02572 0.00030 0.96864 0.00022 0.06673 0.30009 0.05576 0.16987 9999999 9999999
"016"  0.15322 0.20678 0.00900 0.44299 0.00084 0.08505 0.04931 0.59336 0.01425 9999999 9999999
"017"  0.36963 0.59513 0.03454 1.63115 0.04272 0.08021 0.18119 0.60987 0.33001 9999999 9999999
"021"  0.30822 0.40046 0.02960 0.25785 0.00676 0.05183 0.16051 0.10708 -.02115 0.16095 9999999
"022"  0.16860 0.22597 0.01027 0.25224 0.00136 0.04547 0.22736 0.07187 0.10114 0.38228 9999999
"023"  1.00000 1.00000 9999999 9999999 1.00000 1.00000 9999999 9999999 9999999 0.99963 9999999
"024"  0.21826 0.32120 0.01661 0.83407 0.00666 0.05852 0.43930 -.22195 -.18005 9999999 9999999
"031"  0.34231 0.41138 0.07843 0.28624 0.01403 0.02238 0.03605 3.34412 0.01188 9999999 9999999
"032"  0.39393 0.58482 0.04780 0.14954 0.01971 0.02916 0.18644 -.15557 0.07691 9999999 9999999
"051"  1.00000 1.00000 9999999 9999999 1.00000 0.00000 1.95016 9999999 9999999 0.99937 9999999
"072"  1.00000 1.00000 9999999 9999999 1.00000 0.00000 0.00000 0.00723 9999999 0.99803 9999999
"081"  0.32719 0.55302 0.03913 0.44345 0.01013 0.27602 0.29637 -.01716 0.12363 0.38012 0.01513
"089"  0.82492 0.98498 0.14250 0.13961 0.12411 0.00655 0.70056 1.24976 0.07309 0.98493 9999999
"091"  0.97617 0.99583 0.28547 2.40668 0.11075 0.00778 0.69420 9999999 -.32706 9999999 9999999
"099"  0.74868 0.93561 0.12686 5.04487 0.10952 0.00191 0.48380 -.13487 0.27607 9999999 9999999
"101"  0.30251 0.38728 0.03946 0.12543 0.00230 0.01561 0.55219 -.11214 0.04205 0.29337 0.00172
"102"  0.33374 0.57944 0.03492 0.16579 0.04676 0.19665 0.36365 -.11336 0.05771 0.88934 0.00140
"103"  0.22988 0.41969 0.02816 0.21295 0.01201 0.09193 0.18098 -.06548 0.06326 0.62727 0.00310
"104"  0.95863 0.98516 0.52369 0.06314 0.08287 0.00201 0.04018 1.19494 0.01927 0.50883 0.00227
"105"  0.42515 0.62463 0.06070 0.09441 0.00438 0.00143 0.07835 0.09098 0.05366 0.47135 0.00122
"106"  0.67892 0.83246 0.21613 0.09062 0.01811 0.00008 0.44208 0.36810 0.07920 0.36475 0.00427
"107"  0.18898 0.27924 0.01333 0.17633 0.00039 0.02750 0.16984 0.10769 0.06893 0.20792 0.00391
"108"  0.46920 0.58360 0.09895 0.13434 0.00219 0.03181 0.10808 0.58656 0.06362 0.31354 0.01040
"109"  0.38188 0.51971 0.05178 0.07475 0.00881 0.00576 0.06656 -.02348 0.02211 0.20935 0.00516
"110"  0.64466 0.73125 0.16258 0.30594 0.00679 0.00779 0.03982 0.14678 0.16939 0.37482 0.00138
"120"  0.88525 0.94865 0.58302 0.26107 0.06574 0.00001 0.37661 0.13268 0.10351 0.26201 0.00165
"131"  0.26505 0.42210 0.02669 0.21131 0.00963 0.14273 0.72174 0.21302 0.02756 0.69157 0.00786
"132"  0.33343 0.49775 0.03830 0.14430 0.01237 0.06533 0.26641 0.00197 -.00778 0.67861 0.01208
"133"  0.26243 0.42845 0.02959 0.28755 0.00990 0.03927 0.23368 0.10694 0.00942 9999999 0.00523
"139"  0.19834 0.27017 0.01584 0.12342 0.00221 0.17440 0.20637 0.12020 0.07912 0.58592 0.01303
"141"  0.33745 0.40052 0.05884 0.05000 0.00172 0.02632 0.51810 -.08351 0.05545 0.85028 0.00133
"142"  0.66264 0.83122 0.21723 0.19943 0.06518 0.00050 0.43859 -.01130 -.13251 0.26017 9999999
"143"  0.42429 0.60934 0.04212 0.15429 0.04498 0.01297 0.20211 0.17846 0.11475 0.97132 9999999
"151"  0.60389 0.76338 0.19262 0.05517 0.01955 0.04044 0.69492 0.09223 0.06797 0.87168 0.01237
"152"  0.71287 0.86881 0.16611 0.07837 0.03823 0.06488 0.12267 0.00599 0.07127 0.96934 0.00988
"161"  0.21665 0.32238 0.02044 0.52817 0.00528 0.00448 0.17698 0.06276 0.11531 0.50799 0.00113
"162"  0.15347 0.20324 0.01412 0.21276 0.00115 0.32378 0.28753 0.00477 0.12523 0.35777 0.00132
"171"  0.69408 0.90334 0.15400 0.55390 0.02855 0.01690 0.46207 0.12095 0.04566 0.79289 0.00804
"172"  0.50838 0.61462 0.13742 0.15981 0.00719 0.00605 0.22717 0.09608 0.09161 0.41567 0.00227
"181"  0.13086 0.19545 0.00890 0.34236 0.00037 0.02895 0.07622 -.04123 0.03023 0.05852 0.00539
"182"  0.50724 0.57509 0.16148 1.35586 0.00862 0.05808 0.43549 0.36398 -.21913 9999999 9999999
"191"  1.00000 1.00000 0.99974 0.07296 0.99999 0.00000 0.03067 -.18235 0.16760 0.86240 9999999
"192"  0.97640 0.98859 0.45032 0.08330 0.05544 0.00000 0.15550 -.38796 -.00839 0.27624 0.00000
"201"  0.35701 0.49040 0.05358 0.23115 0.00686 0.01899 0.16134 -.05055 0.08002 0.73949 0.01037
"202"  0.89527 0.99149 0.31715 0.89547 0.11049 0.00044 0.30279 -.24874 0.03251 0.70678 0.48828
"203"  0.65967 0.77019 0.11064 0.09223 0.01686 0.00017 0.40289 -.07160 0.00958 0.48674 0.01398
"204"  0.61959 0.77924 0.10155 0.19871 0.01085 0.03955 0.05365 0.05851 0.04884 0.66644 0.00685
"205"  0.55579 0.69959 0.10128 0.23358 0.01270 0.00406 0.24549 0.10052 0.03477 0.66805 0.01162
"206"  0.81903 0.96478 0.25494 0.06253 0.08978 0.00698 0.11132 0.67904 0.03866 0.57946 0.00040
"211"  0.81091 0.92371 0.22648 0.17702 0.04339 0.12372 0.37832 0.41013 0.14487 0.95325 0.12744
"212"  0.70741 0.85605 0.27378 0.41687 0.01666 0.20848 0.15324 0.17043 -.08537 0.79775 0.38947
"221"  0.64250 0.74837 0.16691 0.17259 0.01831 0.05162 0.44215 -.30132 0.01229 0.80779 0.00352
"222"  0.20764 0.29041 0.01693 0.17208 0.00257 0.05066 0.08853 0.00667 0.05944 0.59331 0.01229
"231"  0.59489 0.71984 0.14845 0.27775 0.00924 0.04922 0.19026 -.13583 -.05789 0.44174 0.02507
"232"  0.56544 0.84614 0.07759 0.30228 0.05682 0.13519 0.46129 -.25133 0.01384 0.63165 0.00125
"233"  0.64831 0.84634 0.12419 0.40356 0.02123 0.16708 0.21088 0.01624 0.08521 0.57984 0.00013
```

"234"	0.77638	0.89984	0.25497	0.57898	0.01913	0.00919	0.42445	0.07605	0.09855	0.77749	0.00000
"235"	0.74267	0.99980	0.09283	0.39425	0.16634	0.00044	0.62701	-.01203	0.10051	0.14923	0.01201
"236"	0.17791	0.26528	0.01191	0.24639	0.00356	0.02625	0.14678	0.03987	0.07593	0.05879	0.00198
"237"	0.11283	0.17785	0.00641	0.21756	0.00265	0.01986	0.08447	-.18813	0.07511	0.24594	0.00004
"239"	0.43090	0.66156	0.05501	0.15879	0.03764	0.00388	0.13198	0.03012	0.06504	0.64277	0.00461
"241"	0.69683	0.84538	0.17183	0.21502	0.01723	0.00229	0.41931	-.65556	0.00765	0.73099	0.00391
"242"	0.64070	0.79367	0.11142	0.27544	0.03970	0.12886	0.25889	-.33027	0.02504	0.91613	0.00135
"243"	0.74079	0.86863	0.21789	0.15864	0.02930	0.05191	0.78914	-.11545	-.16233	0.52612	0.03507
"244"	0.78863	0.90059	0.22057	0.12119	0.02082	0.00470	0.74400	-.19096	-.00436	0.51888	0.00498
"245"	0.53296	0.67895	0.09536	0.21365	0.01770	0.00421	0.17588	0.23346	0.03794	0.17414	0.03923
"251"	0.22216	0.27289	0.02506	0.12186	0.00079	0.02406	0.16028	-.01243	0.07321	0.10358	0.00468
"252"	0.42775	0.61866	0.05662	0.11157	0.01367	0.08386	0.25956	-.29216	0.11576	0.39477	0.00251
"253"	0.45569	0.56823	0.11774	0.05836	0.01490	0.07784	0.19699	-.06052	0.04384	0.29686	0.00000
"254"	0.85686	0.93649	0.29413	0.18814	0.05548	0.00063	0.45111	-.03040	0.11335	0.66087	9999999
"255"	0.28549	0.37837	0.03187	0.20525	0.00244	0.00890	0.36638	-.07726	0.01341	9999999	0.00030
"256"	0.08243	0.13458	0.00378	0.26627	0.00063	0.00986	0.18784	0.04955	0.04193	9999999	0.00867
"257"	0.29779	0.40014	0.02873	0.19005	0.00520	0.02032	0.08858	0.11619	0.02484	0.77487	0.04065
"259"	0.29950	0.43346	0.03247	0.13350	0.00395	0.07767	0.18021	-.06895	0.07225	0.73605	0.00709
"261"	0.53619	0.68672	0.09492	0.19118	0.01881	0.01763	0.14825	-.10342	-.04411	0.84032	0.12224
"262"	0.49211	0.69403	0.07690	0.11951	0.01189	0.09802	0.27281	-.01948	0.01481	0.96247	0.08760
"263"	0.85127	0.91526	0.29656	0.05996	0.01264	0.00097	0.08132	-.13897	-.01915	0.48469	0.08273
"264"	0.91098	0.95332	0.31654	0.05339	0.02269	0.00489	0.48090	0.06611	-.05096	0.91032	0.11041
"265"	0.66313	0.80369	0.15811	0.07962	0.01485	0.00837	0.37264	0.22879	-.01287	0.80491	0.05527
"266"	0.92355	0.99538	0.24300	9999999	0.19005	0.56760	0.93057	9999999	9999999	0.99316	0.17999
"267"	0.90283	0.95300	0.21137	0.14654	0.06196	0.00929	0.28736	-.02005	0.00258	0.89059	0.26336
"268"	0.98772	0.99912	0.91635	9999999	0.16603	0.00368	0.51085	0.70414	9999999	0.85935	9999999
"271"	0.60307	0.77168	0.10648	0.10838	0.00810	0.01749	0.14307	0.02805	0.05377	0.60133	0.02839
"272"	0.99426	0.99981	0.43195	0.29544	0.19946	0.00387	0.38048	-.31556	0.09591	0.73468	0.00000
"273"	0.78669	0.87692	0.22632	0.09076	0.03551	0.00479	0.08014	0.15439	0.02853	0.67856	0.04472
"274"	0.57316	0.66359	0.13760	0.11582	0.00733	0.00412	0.26723	-.07735	0.07450	0.56689	0.02284
"275"	0.71372	0.79813	0.33891	0.10132	0.01946	0.00681	0.05445	-.00346	0.05931	0.79416	0.02521
"279"	0.89600	0.94752	0.36178	0.24499	0.02622	0.01245	0.43853	-.23845	0.17560	0.78484	0.06017
"281"	0.73583	0.82697	0.24722	0.14579	0.01365	0.07088	0.24494	-.10939	0.13661	0.76015	0.03180
"282"	0.36150	0.46999	0.05616	0.09707	0.00271	0.01816	0.09712	-.10270	0.07561	0.70568	0.01120
"283"	0.53156	0.62000	0.12589	0.06033	0.01480	0.00169	0.52310	-.16074	0.04022	0.72339	0.02482
"284"	0.47477	0.63677	0.06489	0.12410	0.01602	0.02427	0.34753	-.16671	0.09572	0.80740	0.01525
"289"	0.30865	0.45130	0.03180	0.09529	0.00616	0.01988	0.16261	-.02657	0.03179	0.77167	0.02102
"291"	0.83374	0.93763	0.28431	0.09959	0.04165	0.00232	0.07366	0.10234	0.01100	0.78472	0.00168
"292"	0.20265	0.32740	0.01727	0.14623	0.00671	0.03118	0.21319	0.05830	0.05239	0.51889	0.00768
"293"	0.62545	0.77156	0.17336	0.10266	0.01120	0.02219	0.21512	0.02249	0.01080	0.66417	0.01874
"301"	0.63545	0.77802	0.13788	0.12710	0.04399	0.09846	0.31443	-.11312	0.42897	0.85181	0.00000
"302"	0.88059	0.99118	0.50441	0.29011	0.15956	0.00061	1.08899	0.44719	0.13511	0.61009	0.00000
"303"	0.89437	0.97610	0.26773	0.12190	0.04994	0.00234	0.96292	0.04894	0.04749	0.82345	0.01139
"304"	1.00000	1.00000	9999999	9999999	1.00000	0.00000	0.00000	9999999	9999999	0.63092	9999999
"309"	0.38911	0.57352	0.04867	0.10988	0.01594	0.00390	0.35578	-.27047	0.00402	0.85536	0.09021
"310"	0.06327	0.11150	0.00389	0.18566	0.00077	0.01447	0.08455	0.00344	0.07432	0.51141	0.00463
"321"	0.70089	0.78548	0.25411	0.05432	0.00201	0.00583	0.11366	-.29312	0.01108	0.49112	0.00037
"322"	0.34837	0.48028	0.03779	0.35996	0.01271	0.02128	0.36963	-.09428	0.12659	0.94068	0.00000
"323"	0.62823	0.78317	0.11949	0.05239	0.03476	0.22548	0.28799	0.21165	0.04568	0.94390	9999999
"324"	0.49817	0.71205	0.07184	0.11080	0.01633	0.25418	0.70469	0.12925	0.10019	0.92080	0.01866
"325"	0.32731	0.42037	0.03505	0.12352	0.00153	0.08639	0.49223	0.09926	0.10703	0.88755	0.07508
"329"	0.29518	0.41950	0.03709	0.13657	0.00613	0.02225	0.21537	0.00510	0.03975	0.73540	0.00373
"331"	0.38991	0.48782	0.04746	0.22267	0.00144	0.03216	0.24300	0.10705	0.05583	9999999	0.02713
"332"	0.52179	0.65139	0.11002	0.07029	0.01030	0.07826	0.15770	0.01412	0.03147	9999999	0.10159
"351"	0.78172	0.84937	0.26545	0.83705	0.00934	0.00166	0.04534	0.16622	0.08331	0.01191	0.00430
"352"	0.95172	0.98905	0.68775	3.13849	0.05259	0.00844	0.20462	-.18835	0.10006	0.00000	0.00153
"353"	0.72915	0.86641	0.13572	0.21109	0.07789	0.00608	0.15963	0.08674	0.02609	9999999	9999999
"360"	0.60063	0.87730	0.10347	2.72662	0.02699	0.00273	0.05487	0.02343	0.01552	9999999	0.00346
"370"	0.81511	0.86852	0.37935	2.41564	0.00961	0.01060	0.12765	0.04410	0.36928	0.00080	0.00875
"381"	0.39615	0.58925	0.05486	1.01497	0.01413	0.00856	0.07302	-.02672	0.06812	0.71602	0.01851
"382"	0.42294	0.56258	0.06458	0.48984	0.00903	0.00732	0.07838	-.00106	0.06017	0.00206	0.01639
"383"	0.18868	0.29366	0.01482	0.22913	0.00377	0.08408	0.33872	-.14968	0.06778	0.00040	0.00309
"390"	0.80683	0.88858	0.25454	0.26392	0.03961	0.07834	0.05713	-.07245	0.15144	9999999	0.00468
"411"	0.08679	0.13042	0.00408	0.91003	0.00047	0.07967	0.93853	0.11028	0.12926	9999999	9999999
"412"	0.10726	0.14872	0.00562	0.08331	0.00015	0.03944	0.05506	-.09799	0.03846	9999999	9999999
"421"	0.34178	0.41215	0.07516	0.14247	0.00094	0.05601	0.10238	0.00883	-.05671	9999999	9999999
"422"	0.22458	0.29951	0.02195	0.13223	0.00141	0.08328	0.10995	-.00987	0.06460	9999999	9999999
"429"	0.54383	0.64646	0.09862	0.30569	0.00541	0.02935	0.20787	-.04365	0.02161	9999999	9999999
"431"	0.09164	0.12698	0.00391	0.22770	0.00045	0.04379	0.57023	-.00489	0.08504	9999999	9999999
"432"	0.08229	0.11081	0.00263	0.09910	0.00008	0.05330	0.11353	0.05848	0.04289	9999999	9999999
"433"	0.02331	0.03611	0.00048	0.15840	0.00007	0.04362	0.35005	-.07293	0.07677	9999999	9999999
"439"	0.06137	0.09321	0.00207	0.18935	0.00011	0.06392	0.17598	0.05429	0.07297	9999999	9999999
"451"	0.25366	0.38762	0.02382	0.04385	0.00025	0.01214	0.21268	-.12601	-.01096	9999999	9999999
"452"	0.05322	0.07470	0.00166	0.10198	0.00020	0.02183	0.19366	-.06519	0.01042	9999999	9999999
"453"	0.29901	0.40632	0.03632	0.06936	0.00098	0.01797	0.07291	-.04479	0.02101	9999999	9999999
"454"	0.27839	0.34540	0.03035	0.10203	0.00152	0.02461	0.13816	0.01385	-.00350	9999999	9999999

"460"	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999
"461"	0.30066	0.36453	0.04165	0.05456	0.00013	0.04474	0.19954	-.04864	-.02043	9999999	0.04263
"462"	0.18943	0.24914	0.01428	0.09033	0.00062	0.03331	0.05157	-.40385	0.02784	9999999	0.00143
"463"	0.18652	0.25301	0.01151	0.05937	0.00030	0.03484	0.12494	0.01113	-.00064	9999999	0.00216
"464"	0.21239	0.29029	0.01683	0.05439	0.00017	0.02502	0.56051	-.09524	0.02630	9999999	0.00521
"465"	0.29496	0.39094	0.03146	0.04389	0.00125	0.03857	0.10414	0.14354	-.03027	9999999	0.03020
"466"	0.19688	0.25358	0.01451	0.07557	0.00031	0.01819	0.43804	-.11399	0.03905	9999999	0.00190
"467"	0.14013	0.20110	0.01028	0.06963	0.00026	0.05611	0.18056	-.07712	0.05980	9999999	0.00230
"469"	0.41157	0.52652	0.04753	0.09529	0.00144	0.06210	0.11997	-.29265	0.03728	9999999	0.00000
"471"	0.49955	0.58976	0.07511	0.08003	0.00023	0.01313	0.03486	0.05423	0.00348	9999999	9999999
"472"	0.09637	0.11704	0.00447	0.14418	0.00015	0.04408	0.02506	-.00396	0.04919	9999999	9999999
"473"	0.28796	0.39096	0.03088	0.05952	0.00109	0.02625	0.10133	-.09275	0.01968	9999999	9999999
"474"	0.21591	0.28699	0.01837	0.08511	0.00040	0.02447	0.10596	0.14619	0.01483	9999999	9999999
"475"	0.17484	0.21719	0.01077	0.11477	0.00017	0.02047	0.02408	0.11110	0.05671	9999999	9999999
"476"	0.24178	0.33169	0.01822	0.10106	0.00029	0.03553	0.07700	0.01419	0.01970	9999999	9999999
"477"	0.06851	0.11395	0.00241	0.14376	0.00006	0.03323	0.09648	0.06091	0.05052	9999999	9999999
"478"	0.02901	0.05025	0.00089	0.17010	0.00045	0.07759	0.48821	0.06807	0.09967	9999999	9999999
"479"	0.28307	0.40676	0.03281	0.04974	0.00093	0.07045	0.05372	0.17511	-.04073	9999999	9999999
"492"	0.99591	0.99872	0.45761	0.99363	0.09997	0.00275	0.04456	0.04929	-.07062	9999999	0.01871
"493"	0.53813	0.61156	0.13334	0.75427	0.00073	0.00766	0.04843	0.01319	0.03730	9999999	0.00569
"494"	0.07784	0.11155	0.00317	0.23644	0.00023	0.02610	0.14888	-.02427	0.05507	9999999	0.00121
"495"	0.67336	0.86252	0.11722	0.74224	0.04864	0.06892	0.12023	-.25589	0.13383	9999999	9999999
"501"	0.97057	0.99787	0.63991	9999999	0.14223	1.38773	0.59345	-.61135	-.10134	9999999	9999999
"502"	0.75122	0.91425	0.29339	1.08790	0.01264	0.00089	0.21891	-.69110	0.05639	9999999	0.00000
"503"	0.39030	0.53945	0.04286	1.89502	0.02379	0.01290	0.57685	0.32615	0.12337	9999999	9999999
"504"	0.34316	0.47561	0.03566	0.97285	0.00994	0.06194	0.39498	-.08836	0.11646	9999999	0.00000
"511"	0.59084	0.72806	0.12617	0.06991	0.00847	0.01730	1.14116	-.01904	0.01693	9999999	0.00285
"512"	0.94963	0.98328	0.41645	0.21233	0.02380	0.00015	0.21622	-.06605	-.08178	9999999	0.00255
"521"	0.29495	0.40780	0.03044	0.69920	0.00346	0.03139	0.25515	0.05724	0.11735	9999999	0.00112
"522"	0.16755	0.24078	0.01356	0.69999	0.00074	0.06306	0.22686	-.17119	0.05256	9999999	0.00292
"531"	0.99587	0.99661	0.98839	0.22128	0.00909	0.00140	0.63105	0.07743	0.11237	9999999	9999999
"532"	0.51960	0.63799	0.07322	0.05952	0.00070	0.02825	0.07703	-.03763	-.01145	9999999	0.00194
"551"	0.14219	0.19133	0.01052	1.32734	0.00109	0.02941	0.15780	-.13898	0.09420	9999999	9999999
"552"	0.32154	0.43708	0.03479	1.27959	0.00291	0.04737	0.63339	0.08389	-.00771	9999999	9999999
"553"	0.12922	0.21293	0.00840	1.00763	0.00431	0.01683	0.27870	-.01122	0.14819	9999999	9999999
"559"	0.85796	0.89521	0.59701	0.21209	0.01067	0.00128	0.07457	-.15461	0.01515	9999999	9999999
"561"	0.04494	0.06223	0.00077	0.30099	0.00006	0.06649	0.02672	-.01666	0.03084	9999999	9999999
"562"	0.31739	0.38308	0.04098	0.07064	0.00049	0.05096	0.03779	0.15875	-.00246	9999999	9999999
"563"	0.01675	0.02853	0.00031	0.28458	0.00010	0.10475	0.18163	0.08619	0.07013	9999999	9999999
"581"	0.33394	0.45457	0.03564	0.10351	0.00124	0.01286	0.05154	-.05603	0.05357	0.21549	0.01553
"582"	0.17943	0.28334	0.01334	0.09549	0.00373	0.05887	0.31733	-.02168	0.02722	0.80263	0.00426
"591"	0.15139	0.25779	0.01314	0.30111	0.00083	0.02762	0.10943	0.01221	0.16644	0.10578	9999999
"592"	0.32613	0.48294	0.03791	0.19861	0.00289	0.02530	0.30594	-.09722	0.08951	0.31532	9999999
"601"	0.75067	0.85186	0.30809	0.05794	0.00833	0.03115	0.19673	0.47018	0.09845	9999999	9999999
"602"	0.89948	0.96652	0.42034	0.15157	0.01149	0.00078	0.01154	-.02158	0.25085	9999999	9999999
"611"	0.83914	0.94069	0.35084	0.74319	0.00628	0.00291	0.12852	0.09539	0.39373	9999999	0.00000
"612"	0.87189	0.92431	0.39453	0.55317	0.00383	0.00390	0.03603	0.04526	0.23045	9999999	0.00133
"613"	0.97277	0.98757	0.83837	0.33751	0.02082	0.00267	0.22281	0.07430	0.35235	9999999	0.00005
"619"	0.77244	0.85679	0.23763	0.19598	0.00322	0.02641	0.13532	0.16979	0.08082	9999999	0.00475
"620"	0.18525	0.23984	0.01537	0.08660	0.00011	0.06648	0.09436	0.00640	0.04406	9999999	0.08427
"631"	0.34872	0.47445	0.04228	1.06500	0.00139	0.05817	0.07511	0.20346	0.15420	9999999	0.03243
"639"	0.64400	0.69410	0.15298	0.30466	0.00174	0.03841	0.10003	0.24342	0.02689	9999999	0.02670
"641"	0.72851	0.80459	0.19015	0.14564	0.00132	0.04343	0.19883	0.11881	0.08751	9999999	0.00222
"642"	0.40057	0.47487	0.05798	0.33466	0.00036	0.04452	0.48903	0.07260	0.21602	9999999	0.10918
"643"	0.93301	0.98384	0.38230	0.03637	0.04536	0.00322	0.27049	-.26123	0.02692	9999999	0.00000
"649"	0.29878	0.44140	0.03614	1.04513	0.00332	0.01768	0.79129	-.12586	0.12450	9999999	0.00919
"651"	0.66507	0.79206	0.18155	0.17245	0.02381	0.03331	0.24659	0.27435	0.00801	9999999	0.00199
"652"	1.00000	1.00000	0.49322	9999999	0.49945	0.00410	0.37392	0.05392	0.11913	9999999	0.00000
"653"	1.00000	1.00000	0.99935	9999999	0.99984	0.13867	1.07397	-.77877	0.01680	9999999	9999999
"660"	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999
"661"	0.77045	0.82499	0.19346	0.06578	0.00229	0.02245	0.23296	0.16019	0.19076	9999999	0.01151
"662"	0.13479	0.18602	0.00627	0.12666	0.00049	0.06522	0.05395	0.02235	0.12683	9999999	0.00261
"663"	0.79038	0.84510	0.35911	0.28978	0.00655	0.01848	0.22459	-.30289	0.08249	9999999	0.01197
"681"	0.13015	0.18184	0.00715	1.33961	0.00083	0.14218	0.16369	-.01828	0.24041	9999999	9999999
"682"	0.08871	0.13542	0.00413	2.16367	0.00043	0.05281	0.13123	-.02155	0.32113	9999999	9999999
"683"	0.04606	0.08278	0.00184	1.15465	0.00021	0.06495	0.68735	-.03716	0.11731	9999999	9999999
"691"	0.11460	0.17164	0.00540	0.08113	0.00064	0.12399	0.17682	0.00966	0.17407	9999999	9999999
"692"	0.17475	0.24867	0.01281	0.23558	0.00014	0.02685	0.56469	-.06429	0.05742	9999999	9999999
"701"	0.42453	0.53030	0.07800	0.04797	0.00108	0.03507	0.18369	-.02325	0.00792	9999999	9999999
"702"	0.30420	0.35778	0.03351	0.16387	0.00006	0.04517	0.08836	-.00765	0.04023	9999999	9999999
"711"	0.12778	0.18723	0.00668	0.12153	0.00008	0.08336	1.20364	0.06389	0.05917	0.00017	0.09951
"712"	0.35398	0.44988	0.04319	0.21696	0.00237	0.03238	0.69655	0.00302	0.05232	9999999	0.07651
"721"	0.37735	0.53153	0.04891	0.40656	0.00450	0.02750	0.06359	-.04355	-.01971	9999999	0.47351
"722"	0.65615	0.79596	0.28010	0.40697	0.01128	0.01829	0.47864	0.30386	0.24495	9999999	0.73283
"731"	0.16500	0.29079	0.01415	0.09713	0.00029	0.01575	0.12682	-.10642	0.02173	9999999	9999999
"732"	0.88870	0.89955	0.75653	0.04800	0.00046	0.00330	0.16695	0.87516	-.02508	9999999	9999999

"741"	0.07921	0.14002	0.00481	0.41543	0.00086	0.12812	0.99928	0.19700	0.09941	9999999	9999999
"742"	0.15628	0.24619	0.01044	0.15759	0.00063	0.03684	0.38467	0.27615	0.07177	0.02235	9999999
"743"	0.18108	0.26766	0.01193	0.12804	0.00070	0.04220	0.11774	0.00436	0.06298	9999999	9999999
"749"	0.31991	0.52750	0.04044	0.54572	0.00153	0.02383	0.12722	0.05815	0.07111	9999999	9999999
"750"	0.06158	0.08541	0.00163	0.49078	0.00043	0.05247	0.05587	0.00370	-.09849	9999999	9999999
"771"	0.32434	0.51809	0.04263	1.56291	0.00258	0.00394	0.04500	0.03474	0.39049	9999999	9999999
"772"	0.31984	0.41888	0.04101	0.43670	0.00098	0.02875	0.40748	0.01283	0.16177	9999999	9999999
"773"	0.32291	0.44771	0.04403	1.03771	0.00114	0.02215	0.83842	-.08107	0.10356	9999999	9999999
"774"	0.82084	0.93688	0.14473	9999999	0.10576	0.60454	0.67443	9999999	9999999	9999999	9999999
"781"	0.16416	0.24338	0.01169	0.07744	0.00163	0.04512	0.32523	-.00378	0.02095	9999999	9999999
"782"	0.39527	0.59080	0.05677	0.01593	0.00446	0.01562	0.05053	0.05078	-.00033	9999999	9999999
"783"	0.62968	0.82331	0.12130	0.04626	0.04038	0.27764	0.32722	0.54063	0.02279	9999999	9999999
"791"	0.35991	0.51454	0.04992	0.05472	0.00153	0.03813	0.15062	0.08943	-.01648	9999999	9999999
"799"	0.44160	0.61904	0.05707	0.27093	0.00650	0.06726	0.60289	0.14955	0.01776	9999999	9999999
"801"	0.58262	0.74367	0.11345	0.06242	0.00295	0.01422	0.03943	-.01161	0.00239	9999999	9999999
"802"	0.59252	0.68380	0.14070	0.05782	0.00615	0.02682	0.15476	0.09285	0.00791	9999999	9999999
"803"	0.38262	0.50770	0.04529	0.10720	0.01961	0.03066	0.44838	-.29426	-.05571	9999999	9999999
"811"	0.35850	0.57423	0.04492	0.93160	0.00807	0.05261	0.39770	-.21616	0.17722	9999999	9999999
"812"	0.19560	0.28332	0.01760	0.10844	0.00039	0.04477	0.02575	0.04978	0.03809	9999999	9999999
"813"	0.04646	0.06674	0.00135	0.46930	0.00016	0.05944	0.25537	-.03064	0.14787	9999999	9999999
"821"	0.27804	0.43985	0.03331	0.33653	0.00063	0.02999	0.07616	-.07453	-.05450	9999999	9999999
"822"	0.59347	0.72805	0.13360	0.18708	0.02293	0.08932	0.11053	0.13161	0.04987	9999999	9999999
"823"	0.40141	0.49769	0.05589	0.20141	0.00131	0.03029	0.15122	0.03474	-.01895	9999999	9999999
"829"	0.20582	0.32172	0.01923	0.18696	0.00043	0.07152	0.21776	0.02189	0.01343	9999999	9999999
"841"	0.52060	0.64460	0.09732	2.99532	0.00471	0.00523	0.65957	0.04495	0.19724	9999999	9999999
"842"	0.98775	0.99176	0.85052	0.02467	0.03443	0.00036	0.13899	0.06742	-.09606	9999999	9999999
"843"	0.44360	0.63531	0.06232	0.05397	0.03541	0.00918	0.34883	-.21775	0.01735	9999999	9999999
"851"	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999
"852"	0.33819	0.50510	0.03391	2.58082	0.02505	0.14699	0.27524	-.13848	0.09579	9999999	9999999
"853"	0.32475	0.44890	0.03337	1.39353	0.00577	0.06589	0.49763	0.00438	-.01948	9999999	9999999
"854"	0.43864	0.61431	0.06977	1.09618	0.03066	0.03351	0.58915	0.36075	-.01393	9999999	9999999
"855"	0.15661	0.22652	0.01145	0.68252	0.00048	0.12949	0.11670	-.01650	0.01707	9999999	9999999
"856"	0.32114	0.43281	0.03197	0.68674	0.00690	0.02798	0.55717	0.16350	0.27875	9999999	9999999
"861"	0.30777	0.52508	0.03774	0.65564	0.01629	0.29221	0.03378	0.19664	-.00507	9999999	9999999
"862"	0.12339	0.18131	0.00558	0.30481	0.00042	0.14863	0.18567	0.08590	0.03323	9999999	9999999
"869"	0.20897	0.33751	0.02029	0.17112	0.00073	0.05758	0.10553	0.07894	0.09547	9999999	9999999
"871"	0.90837	0.99992	0.49241	0.08459	0.18937	0.19786	0.20681	9999999	-.02240	9999999	9999999
"872"	0.54119	0.63099	0.19643	0.79375	0.01106	0.05199	0.02010	0.06385	0.02143	9999999	9999999
"873"	0.09111	0.11968	0.00355	1.03593	0.00168	0.20037	0.23348	0.02242	-.05728	9999999	9999999
"879"	0.16977	0.23104	0.01281	1.14487	0.00593	0.05081	0.04271	0.03536	-.02969	9999999	9999999
"881"	0.18749	0.30595	0.01629	0.80353	0.00830	0.15677	0.11221	0.01776	-.17575	9999999	9999999
"889"	0.12371	0.19098	0.00747	0.67775	0.00144	0.12891	0.12547	0.03587	-.40457	9999999	9999999
"900"	0.29160	0.36445	0.03649	0.33585	0.00027	0.10924	0.19068	0.00288	0.09567	0.03364	9999999
"910"	0.31362	0.39184	0.04237	1.12962	0.00567	0.03922	0.17749	0.02538	0.09750	0.11984	9999999
"920"	0.88147	0.91010	0.56685	0.03685	0.00674	0.01443	0.19978	0.06212	0.24797	9999999	9999999
"931"	0.09626	0.16550	0.00522	0.55729	0.00031	0.07421	0.19166	-.14382	0.02651	9999999	9999999
"932"	0.24427	0.33289	0.02159	0.54758	0.00073	0.03908	0.04817	0.01511	0.27685	9999999	9999999
"941"	0.08905	0.14480	0.00491	0.47002	0.00193	0.08012	0.45738	-.05436	-.11680	9999999	9999999
"942"	0.80278	0.93047	0.21783	0.06449	0.08145	0.23513	0.20912	-.44214	0.29091	9999999	9999999
"949"	0.21616	0.29653	0.01808	0.46737	0.00137	0.07900	0.36473	-.01271	-.05884	9999999	9999999
"951"	0.64680	0.80776	0.14537	0.03465	0.00271	0.00477	0.08901	-.02314	-.03937	9999999	9999999
"952"	0.32671	0.39304	0.05247	0.07966	0.00060	0.03048	0.25578	-.03070	-.00598	9999999	9999999
"960"	0.04264	0.06899	0.00107	0.54373	0.00006	0.04294	0.07290	0.03203	0.14724	0.00003	9999999
"970"	0.27648	0.40130	0.02212	0.16242	0.01886	0.06177	0.16104	-.24611	0.13766	9999999	9999999
"981"	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999	9999999
"982"	1.00000	1.00000	0.84786	9999999	0.96040	0.84159	1.71682	1.44243	9999999	9999999	9999999
"990"	1.00000	1.00000	0.22538	9999999	0.37975	0.04199	0.14923	0.01938	9999999	9999999	9999999

;

Referee report on the AGORA-MMS project

Marcel Canoy
Chief Economist Ecorys
Marcel.canoy@ecorys.com

This report reflects on different contributions that measure competition in the context of the AGORA-MMS project executed by a team headed by Professor Johan Eyckmans (HUBrussel) commissioned by the FOD Economie.

The aim of the whole project is to provide methodologies that serve as ‘early warnings’ or screening devices that in some sectors there could be a problem. The exercise is similar to the Market monitoring exercise of the European Commission. Thus, the focus is on methodologies that serve that purpose (unlike methods that directly try to measure competition or abuse of a dominant position in a legal context).

One important consequence is that the methods are not geared to measure competition at the aggregation level of a relevant market (in its legal definition) but at a sectoral level.

For most indicators it holds that their a priori theoretical basis is vulnerable. Often there is a ‘correlation’ between competition and the indicator. This holds e.g. for Lerner index, price cost margins, Herfindahl indexes, churn etc. What this means is that often if there is a problem with the indicator (high or low compared to some appropriate benchmark) that there could be a competition problem, but there need not to be. This does not disqualify the usage of the indicators at all (there is no perfect measure that is both theoretically sounds and empirically useful for this purpose) but it is important to keep this in mind.

This less than perfect correlation between the used indicators and theoretical notions of competition has several consequences. First of all, one has to use more than one indicator. Second, one should interpret the conclusions with some care. Since the purpose is screening, a conclusion of the type: ‘this sector needs further scrutiny’ is often the appropriate conclusion. Third, there is merit in looking at composite indicators that try to use information from various sources.

Contribution 1: Entry and Competition in differentiated products markets

This contribution looks at a very specific type of market (sector), namely a sector that is characterized by local competition. The aim is again to check which sectors that are characterized by local competition seem to call for a closer scrutiny, i.e. the method is not geared towards accurate measures of competition in a relevant market. For local sectors traditional methods are indeed not very insightful, as is mentioned on slide 2. The advantage of the method suggested is that it is well tested and is low on data requirements (often an important bottleneck as this project also demonstrated). The contribution of the researchers is that they want to use the Bresnahan and Reiss method also for heterogeneous goods (whereas it was designed originally - at least implicitly - for homogeneous goods).

The way the researchers want to use the method also for heterogeneous goods is to separate the business creation effect from the competitive conduct. I quite like the basic idea from this. The only real drawback I see is that the method is very blackbox natured, in the sense that one often has considerable difficulties interpreting the results. The examples of 7 sectors on slide 12 prove the point. In many cases it is not clear why certain sectors score in certain ways. Are these data anomalies, technical issues or real economic effects?

Going into the detail of bakeries and real estate agencies (slide 16): the slide concludes that ‘this is a clear signal of a problem concerning competition in the bakery market’, but this seems extremely unlikely since that sector is likely to be very competitive (unlike the real estate agency market). So I conclude that while the method looks promising, it needs detailed institutional knowledge of the sector or the local differences within the sector to become of real value.

Contribution 2: Persistence of profits

The basic idea of looking at persistence of profits is that the measure is first of all more dynamic in nature than traditional static measures and second that whilst profits themselves say preciously little about competition, persistence of profits hints at a lack of entry or other disciplining devices.

I have two questions in relation to this measure. The first one is that profits are notoriously difficult to measure. Reported profits rarely say much about real economic profits, inter alia because of accounting and tax rules. The slides do not address this issue. I am not sure therefore how to interpret the results form slides 13-15.

A second question is whether (in the light of the first problem) other measures of capturing dynamics are not able to produce similar results without the data problems associated with profits. One can measure entry exit in a dynamic way.

Contribution 3: Composite Indicators

Whilst there is a comprehensive literature on composite indicators in general (e.g. the OECD JRC Handbook) the application to competition has been fairly limited so far. The most important thing with composite indicators is that the results can be traced back to their origin. I.e., if a certain sector shows a problematic number, one should be able to trace back why this number has been high (or low). Otherwise it becomes a black box again. Slides 15-17 shows that the authors are aware of this issue.

The contribution of the authors is that they suggest a solution for the black box issue sketched above by the benefit of doubt (BoD) approach, where weights are endogenously determined by the data using linear programming techniques. Whilst I see merit in the approach I cant make much of the conclusions on the basis of the slides alone. It would have helped the reader if more efforts are put in explaining the results. Which sectors are picked up by this method that would not have been picked up by traditional methods? For me it remains high brow technique the merits of which I cannot judge at this stage.

Contribution 4: Decision Tree

This contribution aims to come up with a decision tree based on well-known indicators of competition such as entry rate or HHI. On the basis of a limited set of questions it aims to point at high risk or low risk sectors. The first question is whether entry barriers exist or not. If entry barriers are deemed high (the slides do not report how exactly this is measured) the second question is whether the sector faces international competition. If the answer is no there a third question is whether the market is concentrated or not. If yes the sector is deemed to be a high risk sector.

Of all the methods employed I am least convinced by this one. International competition is not a great measure by itself, and I am not sure how this improves over simple composite indicator methods. Also some measures (HHI) are better calculated at the relevant market level rather than sectoral level. Also reading the draft paper, it becomes apparent that one of the merits of the approach could be to group sector into four different groups (i) potential and internal competition; (ii) potential but no internal competition; (iii) no potential but internal competition; (iv) no potential and no internal competition. Assumption is that if sectors are grouped in this way it will provide information on the risk of competitive problems. I am not convinced yet that this method will yield better results than other methods.

Conclusions

The most important thing still to do for the research team is to see how the different contributions add up. It would e.g. be highly interesting to see and compare which sectors were chosen by one or the other methodology as high risk sectors and then to add some institutional knowledge on the sector, so to conclude what this says about the methodologies employed and their potential advantages and disadvantages. The overall conclusion can then be: in this or that situation use method A, in other employ method B, in others C and D together. The researchers mention the following priorities for the FOD Economie: Priorities for further research at FOD Economie:

- Data work (Import penetration: scale up sample to Belgian economy, Labor productivity: real instead of nominal terms, R&D data integration)
- Other synthetic indicators (Boone’s profit elasticities)
- Econometric estimation of PCM
- Future data access for researcher (Data safe center project)

I don’t deny that these issues are important, but I would like to add a priority, perhaps even suggesting this to be more important than the ones mentioned above. In my view an approach where existing indicators and methods are grouped according to their usefulness in particular situations with particular sector and data characteristics is vital and is likely to yield more than ‘never ending’ data and technique improvements.

Referee Commentaren op MMS-AGORA project

Jan Bouckaert (Universiteit Antwerpen)

Ik vind deze oefeningen/analyses heel waardevol voor het beleid. In elk geval is duidelijk, en dit staat ook in de Intro, er is geen “one size fits all”. Elke methode heeft voor- en nadelen.

Mijn indruk is dat inzicht in de werking van lokale markten specifieke inzichten geeft die maximaal rekening houden met de lokale marktcondities. Vanuit beleidsoogpunt is dit interessant, denk ik.

“Quickscan”

Deze studie gaat uit van nationale of internationale sectoren. Je zou kunnen zeggen dat dit een arbitrair uitgangspunt is. Ik verwijs hierbij naar de Schaumans/Verboven analyse die kijkt naar lokale markten, maar ook naar het algemeen concept van relevante markten die (inter)nationale grenzen niet noodzakelijk als enige criterium neemt. In de presentatie zie ik weliswaar een verwijzing naar HHI en MS maar geen vermelding naar de manier waarop een relevante antitrust markt bepaald wordt (bv. via SSNIP test, …). De vraag is hoe dus de relevantie van de markt bepaald wordt. Slide 12 vermeldt bijvoorbeeld “electric generation, transmission and distribution” in één adem terwijl dit drie verschillende relevante markten zouden kunnen zijn.

“CASE_POP”

Slide 4: het is voor mij niet duidelijk hoe “winst” gedefinieerd/gemeten wordt (zie ook slide 7: is “total assets” de boekwaarde of marktwaarde), en als er winst is waarom die zou moeten geïnterpreteerd worden als abnormaal. De interpretatie kan wel iets zeggen over persistentie van winst over de tijd, maar de hoogte van de winst is niet noodzakelijk “abnormaal” te noemen. Misschien is er wel een grotere persistentie over de tijd wanneer de winsten niet supranormaal maar economisch zijn. Er wordt ook impliciet verondersteld dat alle bedrijven op basis van zelfde classificatie met elkaar concurreren; competitie kan lokaal zijn of breder/smaller dan de classificatie. Ik vind dit wel een belangrijke oefening maar de vraag is ook hoe interpreteer je de geschatte parameters: welke theorie of harm heb je onderliggend. Een lage persistentie kan het gevolg zijn van roterende winsten in een collusieve omgeving, maar ook van echte concurrentie. Hoe kan je dit identificeren van elkaar?

“entry_tresholds”

Zeer gefundeerde analyse (heb de paper ook gelezen) maar wel een (te?) voorzichtige conclusie.

De presentaties over “indicators” en “composite indicators” zijn voor mij moeilijker om commentaar op te geven.

International Expert Workshop
Market Monitoring Indicators

Friday, March 26, 2010

Hogeschool-Universiteit Brussel

	Program
9:15-9:45h	Welcome coffee
9:45-10:00h	*The AGORA program and MMS project* Aziz Naji, Federal Public Service Science Policy *Introduction: goals and set-up of the MMS project* Marie-Thérèse Peeters, Federal Public Service Economy
10:00-11:00h	*Revised methodology of the screening stage of the Market Monitoring* Dominique Simonis, Head of Sector, DG ECFIN, European Commission
11:00-11:15h	Coffee break
11:15-12:00h	*Experiences of the Office of Fair Trading in using empirical indicators for market investigations* John Gibson, Deputy Director Strategy and Planning, Office of Fair Trading, UK
12:00-13:15h	Lunch
13:15-14:00h	*Composite Indicators: Methodology & Guidelines* Tom Van Puyenbroeck, HUBrussel
14:00-15:00h	MMS Project: *Preliminary Findings for the composite Market Functioning Indicator* *Choice of Indicators* Stijn Kelchtermans, HUBrussel
15:00-15:15	Coffee break
15:15-16:15h	MMS Project: *Preliminary Findings for the composite Market Functioning Indicator* *Aggregation of Indicators* Johan Eyckmans, HUBrussel
16:15-17:00h	MMS Project: *Preliminary Findings of An Entry Threshold Ratio Approach for Competition in Local Markets* Frank Verboven, K.U.Leuven
17:00	Closing workshop

economie
FPS Economy, S.M.E.s, Self-employed and Energy

KATHOLIEKE UNIVERSITEIT BRUSSEL — EHSAL — VLEKHO - HONIM
partners in HUB Hogeschool-Universiteit Brussel

FEDERAAL WETENSCHAPSBELEID — POLITIQUE SCIENTIFIQUE FEDERALE

Venue

Hogeschool-Universiteit Brussel HUB
Stormstraat / Rue d'Assaut 2, B-1000 Brussels
room: 6306 (multimedia aula in EHSAL 3, 6th floor)

Travel Directions

It takes 5 minutes walking from Brussels Central Station to the HUB Stormstraat campus, see http://www.hubrussel.be/eCache/IEE/13/250.html for more information how to reach us.

If you want to come by car, please let us know in advance by email so that we can make reservations. (we need your licence plate number and car brand / color)

Registration

Participation is free but please confirm your participation by email to yolande.degroote@hubrussel.be

This workshop is organised by Hogeschool-Universiteit Brussel in collaboration with the Federal Public Service Economy. Financial support by the AGORA program of Belgian Federal Science Policy Office is gratefully acknowledged.

AGORA MMS project

Debriefing International Expert Workshop, March 26, 2010

Main comments and conclusions to be incorporated in our future work

- AIM OF THE SCREENING TOOL
  After the presentation of J. Gibson (OFT), it is clear that the aim of our market screening tool is not the detection of abuses of market power by individual firms or cartels. The aim of the tool is rather to foster understanding of the importance and specific nature of different sectors (not markets!). The monitoring tool should be kept simple and transparent and always, we should go back to the raw data, i.e. the values of the original underlying indicators. The monitoring tool is rather an information transmission device than a surveillance and detection system.
- DYNAMICS
  Several participants stressed the importance of looking at evolution over time of indicators. This should be an important priority in our future data work. For some indicators, we can consider taking up both the absolute level and the rate of change of the indicator. The Persistence of Profits approach that we proposed earlier is a good way to incorporate dynamics in our screening tool.
- CHOICE OF INDICATORS
  We have heart little negative comments on the set of indicators that we selected. There were however detailed comments on the computation of particular indicators (for instance concentration should account for imports/exports and churn should be take into account mergers & acquisitions). No suggestions were made to include additional indicators compared to the set of indicators that we proposed earlier.
- AGGREGATION OF INDICATORS
  Workshop participants were interested in the results of the alternative aggregation method of Benefit of the Doubt. For the MMS project, we will do both types of aggregations: classical linear aggregation with fixed (and mostly equal) weights AND more sophisticated benefit of the doubt approach (but using different subgroups of sectors as peers: manufacturing and services separately for instance).

Detailed comments by participants

SIMONIS (DG ECFIN):

- DG ECFIN's new methodology for Market Monitoring (part of assessing the Single Market) has two dimensions: economic importance and market performance.
- This approach doesn't rank sectors anymore, but plots them onto a 2-way axis with 4 "importance" zones – A, B, C, D. Automatically screened sectors are in A, and some in B.
- Services and manufacturing are analyzed separately and have different benchmarks (Construction is under services, as it was an outlier under manufacturing).
- DG checked for correlations between different indicators in order to keep the most relevant ones.
- Regulation on services could be used to choose sectors in quadrant B that should make the subject of further investigation.
- OECD has just revised their product-market regulation index in order to make use of a better weighting scheme.

GRILO (DG ECFIN):

- Dynamics could be used for some indicators within the CI or for the CI itself; this has not been done so far by the DG.
- The indicator "investment share" measures the share in total investments that the sector supplies to other sectors.
- DG ECFIN's tool is not for competition analysis, but for market monitoring.

MOLLEN (DG ECFIN):

- Regarding Johan's question about looking at both dynamic and static levels for the composite indicator, she suggested we could combine both, if relevant.
- OFT 's study is similar to their DG study but there are some differences. The DG study has two stages to screening of sectors.

GIBSON (OFT):

- OFT's scope was different than the Commission's – the office's role is to take into court cases of abuse of market power. Therefore, market definition is very important.
- The 2004 exercise tried to combine indicators into a CI, but the OFT dropped this approach due to the very different results they got when changing the weights used.
- Weighting should be aligned to economic policies.
- Comparison between sectors could be redundant.
- SIC4 data was too heterogeneous to correspond to actual markets.
- Issues with large firms having only one SIC (NACE) code and many secondary activities (issue gets worse at SIC 3-4-5 digits).
- The 2006 exercise used only two dimensions – competition and productivity.
- Churn was measured among the bottom firms (by market share).
- The benchmark used was EU15 average, not cross-sectoral.
- Sectors that comprised too many markets have been filtered out.

- Different database was used to check the robustness of their analysis.
- The 2004 study's conclusion was that a bottom-up or case-by-case approach would have worked better for the OFT's goals in order to understand the sector better.
- The top-down approach is useful as an additional source, when other sources signal problems on some markets such as consumer complaints.

CANOY (ECORYS):

- The CI should only be used to send a simpler message, not as an analysis tool, so we should always refer to the raw data as well.
- When computing churn, mergers should also be taken into account.
- Regarding our study, we should exclude non-business sectors.
- Using the composite indicator to see the sector performance, we can for example use a 10 point scale for each S-C-P and see how each sector scores.
- For PCM, we should look at the dynamic level and see how it influences competition in the sector.

DRESSE (NATIONAL BANK):

- Before aggregating the firm level data to NACE 2 level, we should kick out the outliers first.

BOUCKAERT (UNIVERSITEIT ANTWERPEN):

- Using HHI based on market share as an indicator itself would be misleading; we should take the openness into account and look into whether it is local or international competition.
- HHI is not based on actual market shares, as it does not capture the results of foreign firms.

BRAMATI (FOD):

- When computing the import penetration using the PRODCOM database, what do we do with the service sector?
- Regarding to our study, how do we put weight with the negative PCM, do we put positive weight or not. Johan answered that before weighting, we adjust the values so that each indicator would point in the same direction and all numbers are positive.

OTHERS:

- CI's are also used beyond communication purposes (e.g. as budgeting tools).
- Theoretical benchmarks could be used on some indicators instead of empirical benchmarks.

[thanks to Daniel Neicy and Cherry Cheung for taking note of these detailed comments by workshop participants]

Expert Workshop
Market Monitoring Indicators

Friday, May 20, 2011

Hogeschool-Universiteit Brussel

	Program
9:15-9:45h	Welcome coffee
9:45-10:00h	*The AGORA program and MMS project* Aziz Naji, Federal Science Policy *Introduction: goals and set-up of the MMS project* Luc Mariën, Federal Public Service Economy
10:00-11:00h	*Sectoral indicators for market functioning* HUBrussel team
11:00-11:15h	Coffee break
11:15-12:30h	*Results for the composite Market Functioning Indicator:* *aggregated indicators and benefit of the doubt* HUBrussel team
12:30-13:30h	Sandwich lunch
13:30-14:15h	*Results for the composite Market Functioning Indicator:* *a quick scan* HUBrussel team
14:15-15:00h	*Case study:* *Entry and Competition in Differentiated Products Markets* Catherine Schaumans (UTilburg) and Frank Verboven (K.U.Leuven)
15:00-15:15	Coffee break
15:15-16:00h	*Case study:* *Persistence of Profits: sectoral approach for the Belgian Economy 2000-2009* HUBrussel team
16:00-17:00h	Feedback and discussion
17:00	Closing

economie
FPS Economy, S.M.E.s, Self-employed and Energy

KU Brussel EHSAL VLEKHO - HONIM
partners in HUB
Hogeschool-Universiteit Brussel

FEDERAAL WETENSCHAPSBELEID POLITIQUE SCIENTIFIQUE FEDERALE

Venue

Hogeschool-Universiteit Brussel HUB
Stormstraat / Rue d'Assaut 2, B-1000 Brussels
room: 6306 (multimedia aula in Hermes building, 6th floor)

Travel Directions

It takes 5 minutes walking from Brussels Central Station to the HUB Stormstraat campus, see http://www.hubrussel.be/eCache/IEE/13/250.html for more information how to reach us.

If you want to come by car, please let us know in advance by email so that we can make reservations. (we need your license plate number and car brand / color)

Registration

Participation is free but please confirm your participation by email to yolande.degroote@hubrussel.be

This workshop is organised by Hogeschool-Universiteit Brussel in collaboration with the Federal Public Service Economy. Financial support by the AGORA program of the Belgian Federal Science Policy Office is gratefully acknowledged.

AGORA MMS project

Debriefing Expert Workshop,
May 20, 2011

Comments by participants[1], ordered by the workshop agenda

Introduction

NAJI (Federal Science Policy)

- It is important that the final data set-up is easily accessible and that a guide/manual for accessing the data is in place.

Indicators

VAN DER LINDEN (FEDERAL PLANNING BUREAU)

- Bear in mind the ambiguity in interpretation of the indicators
    - *Multiple indicators are used; each has been carefully defined in terms of 'good' vs 'bad'*

VERMEULEN (EUROPEAN CENTRAL BANK)

- Selected turnover: why not combine method 3 & 4?
- pcm: should in principle also subtract capital return to obtain the true profit rate. See for example pharma: high pcm (in our definition) but high capital expenses.
    - *The FPB has been working on this (see report by Glenn Rayp), but no data for recent years[2]. More sophisticated methods of pcm estimation (Hall, Roeger) would also be an option.*

---

[1] Replies and clarifications to participants' comments that were already given during the workshop are printed in italics.

[2] Federal Planning Bureau (2010). Competition and regulation in Belgium, 1997-2004, Working paper 3-10. -> see section 2.2 (average profitability), p6. Data sources were EUKLEMS and (for the cost of capital since for Belgium this information is not in EUKLEMS) the FPB.

Composite indicator

MALEK MANSOUR (FEDERAL SCIENCE POLICY)

- A currently ongoing composite indicator exercise at the OECD involves a principal component (PC) analysis. Such a PC-analysis could be considered in the MMS-project as a robustness check.

VAN DEN CRUYCE (FEDERAL PLANNING BUREAU)

- Bootstrapping method: careful with the interpretation of a wide confidence interval of the composite indicator score. A sector may be particularly 'unlucky' with respect to its position in the data cloud. One should be careful not to suggest that a wide confidence interval equals a strangely behaving and thus malfunctioning sector.
- Note that even within sectors, the included firms may offer very heterogeneous products.

SIMONIS (EUROPEAN COMMISSION)

- The analysis has been split into manufacturing vs services. Another way to make that split is based on factor intensity i.e. consider the sectors with low capital intensity separately from the sectors with high capital intensity.

VERMEULEN (EUROPEAN CENTRAL BANK)

- BoD seems to smoothen out the outliers while these could be the most interesting data points. Therefore, an alternative (or additional) approach could be to pay special attention to those sectors that behave as an outlier for one or more indicators.

Quick scan

VERMEULEN (EUROPEAN CENTRAL BANK)

- Clarify whether import data is really available at NACE 3-digit level.
- Check how 'high-risk sectors' perform on other indicators.

PEPERMANS (HUBRUSSEL)

- Heterogeneity within sectors is still a major issue even at the 3-digit level. E.g. sector 351 covers generation, transmission and distribution of electricity.
  - *Whenever possible, an analysis at the 4-digit level is preferred.*

<u>Entry threshold ratios</u>

VERMEULEN (EUROPEAN CENTRAL BANK)

- The identification in your model is based on cross-sectional variation. Could variation over time also be exploited?
    - *In principle yes (see dynamic entry models), but in practice this is a major challenge due to the occurrence of multiple equilibria i.e. it is very hard to make these models converge.*
- Does this approach assume constant returns to scale?
    - No explicit assumption is made, although it should indeed be clarified how returns to scale are accommodated in the model. In particular, whether these are picked up by the revenue equation or the entry equation.
- What is the data source for establishment data?
    - *The KBO-data lists the number of establishments of firms. This was merely used for the <u>selection</u> of sectors to analyze since revenues are not split out per establishment. In the future, revenues per establishment would be available.*

VAN DER LINDEN (FEDERAL PLANNING BUREAU)

- Analyzing retail trade is tricky. E.g. bakeries: should also include the supermarket establishments.
    - In principle this would require separate sales data on the bakery departments within supermarket establishments, which is infeasible in practice. Alternatively, a dummy could be added to indicate the presence of a supermarket establishment in the zip code, but this would result in very little variation in the data since many zip codes will have a supermarket.

CORNILLE (NBB)

- Watch out which NACE codes to include when, for example, analyzing bakeries. There are “bakery-shops” and “bakery-manufacturing units”, which are in different NACE codes.
    - *This was verified and bakeries are consistently classified into one NACE code only (both shops and manufacturers), which is the one used in the analysis.*

VAN DEN CRUYCE (FEDERAL PLANNING BUREAU)

- What’s the explanation for the real estate agents (where no competitive effect is found)?
    - *The model does not provide a final judgment; it is merely a first step i.e. a signal for further investigation involving detailed sector knowledge.*

Persistence of profits

MALEK MANSOUR (FEDERAL SCIENCE POLICY)

- Is the measure used normal or supranormal profits?
    - *Supranormal profits i.e. after deduction of labor, materials…*

VERMEULEN (EUROPEAN CENTRAL BANK)

- An extension of the analysis could be to run regressions at NACE 3-digit level rather than at the firm level. This would make it more robust, e.g. you would always have the full 10 years of observations. Could then also add the other sector indicators so the analysis would become more informative.
- Clarify whether the analysis controls for sector-level business cycle effects.

VAN DEN CRUYCE (FEDERAL PLANNING BUREAU)

- Note that dropping firms that exit the market may bias the results.

SECTOR / MARKET INDICATOR FORM		
1. Name		Capital Intensity
2. Description		Capital intensity has an impact on industry profitability (Schepherd (1972), Schmalensee, Willig,(1989),Tirole (1988)). Capital requirements are identified by Bain (1956) as an element of market structure that affects the ability of established firms to prevent supra-normal profits from being eroded by entry. The intuition is that entrants may have trouble finding financing for their investments because of the risk to the creditors. One argument is that banks are less eager to lend to entrants because they are less well known than incumbents. Besides, entrants may be prevented from growing as existing players inflict losses on them in the product market in order to reduce their ability to find financing for new investments (Tirole, 1988). $CAPINT_i^t = \sum_{i \in S} m_i^t \frac{K_i^t}{y_i^t}$ where K_i^t stands for firm i's capital stock value in period t, y_i^t for its turnover and $m_i^t = y_i^t / y_s^t$ for its share in total sector turnover (i.e. its market share). The capital intensity for sector s is defined as the weighted sum of the ratio of individual firms' capital stock value over turnover. The weights are typically based on firm's share in the sector total turnover or value added of the sector.
3. Result tables in sectoral database		ID_CAPINT_NACE&NACE (NACE2, NACE3 and NACE4)
4. Source data used		TU_NBB_&YEAR (2000 - 2009), containing data on: a) Tangible Fixed Assets (code 27) = raw material, consumables, services and other goods b) Total Assets (code 20/58) c) Turnover (code 70)
5. Availability		2000-2009
6. 1 Variable1	**Name**	CD_NACE&NACE
	Label	
	Formula	
	Comments	NACE 2, NACE3 or NACE4
6. 3 Variable3	**Name**	MS_W_CAPINT_TOTASS_&YEAR
	Label	
	Formula	W_CAPINT (weighted average capital intensity in the sector using total assets) 1) W_CAPINT=sum(each firm's tangible fixed assets in the sector)/sum(each firm's total assets in the sector)
	Comments	
6. 4 Variable4	**Name**	MS_NO_OF_FIRMS_TOTASS_&YEAR
	Label	
	Formula	No of firms= counting the number of firms in corresponding sector based on firms which have tangible fixed assets and total assets data.
	Comments	
6. 6 Variable6	**Name**	MS_W_CAPINT_TRN_&YEAR
	Label	
	Formula	W_CAPINT (weighted average capital intensity in the sector using turnover) 1) W_PCM=sum(each firm's tangible fixed assets in the sector)/sum(each firm's turnover in the sector)
	Comments	
6. 7 Variable7	**Name**	MS_NO_OF_FIRMS_TRN_&YEAR

	Label	
	Formula	No of firms= counting the number of firms in corresponding sector based on firms which have tangible fixed assets and turnover data
	Comments	
7. Methodology		See the Final Report for details
8. Literature		Refer to the Final Report
9. Last exercise		June 20, 2011
10. Responsible		Cherry Cheung
11. Coverage		All NACE 2-3-4 digit sectors over 2000-2009
12. Reliability		
13. Annexe(s)		
14. Remarks(s)		Only those companies are included that have a NACE code of at least 4 digits (>=4). All other firms (including those with missing NACE code) are dropped. Further analysis on the NACE can be made, e.g. which sector has the most missing NACE.

<table>
<tr><th colspan="3">SECTOR / MARKET INDICATOR FORM</th></tr>
<tr><td colspan="2">1. Name</td><td>Economic Churn Rate</td></tr>
<tr><td colspan="2">2. Description</td><td>The churn rate is an indicator that reflects the presence of entry and exit barriers in a non-extensive way. We define the churn ratio in year y and on sector s as the ratio of the number of firms that enter or exit the industry to the number of active firms.
$$CHURN_s^t = \frac{\sum_{i \in s} \left[EN_i^t + EX_i^t \right]}{\sum_{i \in s} AF_i^t}$$
The variables EN_i^t and EX_i^t are dummy variables taking value one if firm i was entering or exiting the industry respectively. AF_i^t takes value one for firms that can be considered active in the industry during the time frame considered. Gross entry and exit rates are defined by the ratio’s EN_i^t / AF_i^t and EX_i^t / AF_i^t .
A second definition, taking into account the relative importance of each company, weights entries, exits and active firms by their respective market shares, which is the preferred choice, as it allows to measure the importance of entry and exit relative to the active companies. The formula for this (preferred) method is as follows:
$$WCHURN_s^t = \frac{\sum_{i \in s} \left[EN_i^t \cdot m_i^t + EX_i^t \cdot m_i^t \right]}{\sum_{i \in s} AF_i^t \cdot m_i^t}$$
where m_i^t denotes the market share of company i in year t.
A company is considered an entry only once in the selected period (2001-2009); this is so for the first year it recorded positive turnover. Also, a company is considered an exit only once in the selected period (2001-2009); this is so for the first year after the last year it recorded positive turnover.
Firms with positive turnover in the selected year are defined as active firms.
Due to this dynamic definition, churn rates for the first year for which there is data available (here 2000) cannot be calculated (as we need one period before to determine exits).
Furthermore, due to the way we define active firms (see sections 6.4 and 14 below), we define “sleeping firms” as those that are inactive in year y, but have been active before and after year y (in sector s). Thus, the relationship between the different variables is as follows:
$$active_s^y = active_s^{y-1} + sleeping_s^{y-1} - sleeping_s^y + entries_s^y - exits_s^y$$</td></tr>
<tr><td colspan="2">3. Result tables in sectoral database</td><td>ID_CHURN_NACE2, ID_CHURN_NACE3, ID_CHURN_NACE4</td></tr>
<tr><td colspan="2">4. Source data used</td><td>TU_SEL_TRNOV_&YEAR (2000 - 2009), containing data on domestic Turnover (DT) = an estimation of the turnover in Belgium, calculated as the difference between Selected Turnover and Total Exports (based on data from the NBB).
TU_BR_ACTIVE_&YEAR (2000-2009) from the Sectoral DB, containing active companies within the selected period.</td></tr>
<tr><td colspan="2">5. Availability</td><td>2001-2009</td></tr>
<tr><td rowspan="2">6. 1 Variable1</td><td>Name</td><td>CD_NACE&nace</td></tr>
<tr><td>Label</td><td>x-digit sector of activity</td></tr>
</table>

	Formula	First x digits of the NACE sector
	Comments	The number of digits of the NACE code can be defined via a parameter &nace
6. 2 Variable2	**Name**	MS_ENTRIES_&year
	Label	Number of entries in selected year
	Formula	Count of firms switching from inactivity to activity in the selected year
	Comments	A company is considered an entry only once in the selected period (2001-2009); this is so for the first year it recorded positive turnover. See general remarks further down
6. 3 Variable3	**Name**	MS_EXITS_&year
	Label	Number of exits in selected year
	Formula	Count of firms switching from activity to inactivity in the selected year
	Comments	A company is considered an exit only once in the selected period (2001-2009); this is so for the first year it starts recording zero or negative turnover (or is deregistered) after the last year it recorded positive turnover. See general remarks further down
6. 4 Variable4	**Name**	MS_ACTIVE_&year
	Label	Number of active companies in selected year
	Formula	Count of firms with positive turnover in selected year
	Comments	See comments in section 14
6. 5 Variable5	**Name**	MS_ENT_RT_&year
	Label	Entry rate in selected year
	Formula	Number of entries divided by number of active firms for selected year
	Comments	The entry rate (MS_ENT_RT) is to be used in the **quick scan** method.
6. 6 Variable6	**Name**	MS_CHURN_&year
	Label	Churn rate for selected year
	Formula	Churn rate = (Number of entries + Number of exits) / Number of active co.
	Comments	
6. 7 Variable7	**Name**	MS_WG_ENTRIES_&year
	Label	Weighted entries in selected year
	Formula	Sum of market shares of entrants in selected year
	Comments	

6. 8 Variable8	**Name**	MS_WG_EXITS_&year
	Label	Weighted exits in selected year
	Formula	Sum of market shares of exiting firms in selected year
	Comments	
6. 9 Variable9	**Name**	MS_CHURN_&year_WG
	Label	Weighted churn rate
	Formula	Wg churn rate = Wg entries + Wg exits
	Comments	Captures the relative importance of exits and entries
7. Methodology		Please refer to the final report
8. Literature		Please refer to the final report
9. Last exercise		December 10 2010
10. Responsible		Daniel Neicu
11. Coverage		All NACE 2-3-4 digit sectors over 2001-2009 (results for 2001 are based on data for 2000)
12. Reliability		See comments under section 14
13. Annexe(s)		• Stats_nr_companies.xls – summary of total number of companies in sample across years, including statistics on negative turnover and missing NACE codes.
14. Remarks(s)		• Our definition of churn does not directly capture the fact that companies change their sector of activity. Thus, if a company is active in year y in sector s and in year y+1 in sector t, it will not be counted as an exit from sector s and an entry in sector t. However, it will be counted as an active firm in sector s in year y and in sector t in year y+1. We argue that capturing this type of event is not possible given the current data, which are unreliable insofar as some companies seem to switch back and forth between 2 or 3 different NACE sectors over longer periods of time. The code to calculate churn rates can be adjusted to capture this issue if the data will become more reliable. • Companies with NACE codes composed of less than 4 digits are recoded as having missing NACE codes. • We do not take into account mergers & acquisitions for our calculation of churn because of data constraints. Extensive literature suggests that these events are important for churn rates and relate strongly to competition within a sector (horizontal mergers). • Firms with negative, missing or zero turnover are considered inactive (FL_ACT=0), but not necessarily exits from the market (see conditions for exit dummies above). Therefore, the sum of active firms in sector N in year Y is not equal to the sum of active firms in sector N in year Y-1 plus entries minus exits, because we define “sleeping firms” in the manner described above.

<table>
<tr><th colspan="3">SECTOR / MARKET INDICATOR FORM</th></tr>
<tr><td colspan="2">1. Name</td><td>Market Concentration</td></tr>
<tr><td colspan="2">2. Description</td><td>The Herfindahl-Hirschman Index (HHI) is a traditional indicator for measuring market concentration. The HHI is calculated as the sum of squared market shares of all firms in the sector or market.
$HHI_i^t = \sum_{i \in s} \left[m_i^t \right]^2$
Non-aggregated data on a measure of economic activity, for instance production in physical units or turnover, of all firms in the sector is needed to compute the market shares. One of the traditional indicators for measuring market concentration is the Herfindahl Index, which is widely used both by policy makers, as well as policy analysts or courts of law. C4 and C8 are calculated for robustness checking purpose.</td></tr>
<tr><td colspan="2">3. Result tables in sectoral database</td><td>ID_CONCRT_NACE&NACE (NACE2, NACE3 and NACE4)</td></tr>
<tr><td colspan="2">4. Source data used</td><td>TU_SEL_TRNOV_&YEAR (2000 - 2009), containing data on:
a) Selected Turnover (ST) = an estimation of the total turnover, based on three sources with their respective priorities: 1° Company Accounts, 2° SBS (Structural Business Survey) and 3°VAT
b) Domestic Turnover (DT) = an estimation of the turnover in Belgium, calculated as the difference between Selected Turnover and Total Exports (based on data from the NBB).</td></tr>
<tr><td colspan="2">5. Availability</td><td>2000-2009</td></tr>
<tr><td rowspan="4">6. 1 Variable1</td><td>Name</td><td>CD_NACE&NACE</td></tr>
<tr><td>Label</td><td></td></tr>
<tr><td>Formula</td><td></td></tr>
<tr><td>Comments</td><td>NACE 2, NACE3 or NACE4</td></tr>
<tr><td rowspan="4">6. 2 Variable2</td><td>Name</td><td>MS_C4_ST_&YEAR</td></tr>
<tr><td>Label</td><td></td></tr>
<tr><td>Formula</td><td>C4 (Concentration Ratio for top 4 firms in the sector)
1) Rank each firm in each NACE sector according to its market share, based on Selected Turnover
2) Pick the top 4 firm with highest market share in each NACE.
3) C4 is the total market share of the 4 largest firms in the sector.</td></tr>
<tr><td>Comments</td><td></td></tr>
<tr><td rowspan="4">6. 3 Variable3</td><td>Name</td><td>MS_C4_DT_&YEAR</td></tr>
<tr><td>Label</td><td></td></tr>
<tr><td>Formula</td><td>C4 (Concentration Ratio for top 4 firms in the sector)
1) Rank each firm in each NACE sector according to its market share, based on Domestic Turnover
2) Pick the top 4 firm with highest market share in each NACE.
3) C4 is the total market share of the 4 largest firms in the sector.</td></tr>
<tr><td>Comments</td><td></td></tr>
<tr><td rowspan="4">6. 4 Variable4</td><td>Name</td><td>MS_C8_ST_&YEAR</td></tr>
<tr><td>Label</td><td></td></tr>
<tr><td>Formula</td><td>C8 (Concentration Ratio for top 8 firms in the sector)
1) Rank each firm in each NACE sector according to its market share, based on Selected Turnover
2) Pick the top 8 firm with highest market share in each NACE.
3) C8 is the total market share of the 4 largest firms in the sector.</td></tr>
<tr><td>Comments</td><td></td></tr>
</table>

6. 5 Variable5	**Name**	MS_C8_DT_&YEAR
	Label	
	Formula	C8 (Concentration Ratio for top 8 firms in the sector) 1) Rank each firm in each NACE sector according to its market share, based on ***Domestic Turnover*** 2) Pick the **top 4** firm with highest market share in each NACE. 3) C4 is the total market share of the 4 largest firms in the sector.
	Comments	
6. 6 Variable6	**Name**	MS_HHI_ST_&YEAR
	Label	
	Formula	HHI (Herfindahl-Hirschman Index) based on ***Selected Turnover*** 1) Take out the observations with negative or zero turnover 2) Calculate the market share of each firm in the particular sector =turnover of the firm / total turnover of the particular sector (tot/sum tot) 3) HHI is calculated by squaring the market share of each firm in each NACE, and then summing the resulting numbers by NACE.
	Comments	
6. 7 Variable7	**Name**	MS_HHI_DT_&YEAR
	Label	
	Formula	HHI (Herfindahl-Hirschman Index) based on ***Domestic Turnover*** 1) Take out the observations with negative or zero turnover 2) Calculate the market share of each firm in the particular sector =turnover of the firm / total turnover of the particular sector (tot/sum tot) 3) HHI is calculated by squaring the market share of each firm in each NACE, and then summing the resulting numbers by NACE.
	Comments	
6. 8 Variable8	**Name**	MS_HHI_NORM_ST_&YEAR
	Label	
	Formula	HHI NORM=(HHI-1/N)/(1-1/N)
	Comments	HHI Normalization calculation is based on ***Selected Turnover***
6. 9 Variable9	**Name**	MS_HHI_NORM_DT_&YEAR
	Label	
	Formula	HHI NORM=(HHI-1/N)/(1-1/N)
	Comments	HHI Normalization calculation is based on ***Domestic Turnover***
6. 10 Variable 10	**Name**	MS_NO_OF_FIRMS_ST_&YEAR
	Label	
	Formula	
	Comments	No. of Firms in corresponding sector (counting is based on Firms which have “***selected turnover***” Data
6. 11 Variable11	**Name**	MS_NO_OF_FIRMS_DT_&YEAR
	Label	
	Formula	No. of Firms in corresponding sector (counting is based on Firms which have “***domestic turnover***” Data
	Comments	Based on Domestic Turnover
7. Methodology		See the Final Report for details
8. Literature		Refer to the Final Report
9. Last exercise		June 20, 2011
10. Responsible		Cherry Cheung, validated by Jean-Yves Jaucot and Luc Mariën
11. Coverage		All NACE 2-3-4 digit sectors over 2000-2009

12. Reliability	
13. Annexe(s)	
14. Remarks(s)	1) Only those companies are included that have a NACE code of at least 4 digits (>=4). All other firms (including those with missing NACE code) are dropped. Further analysis on the NACE can be made, e.g. which sector has the most missing NACE… 2) Alternative Calculation of HHI based on other literatures can be done in the future. 3) Firms with positive turnover are included in the calculations; further adjustment will be taken into account.

SECTOR / MARKET INDICATOR FORM		
1. Name		Import penetration
2. Description		The indicator import penetration IP for a given sector in a selected period is computed by dividing the total imports of products included in that sector (CN8 – CPA codes) by the sum of total turnover of companies included in the sector (NACE codes) plus the total imports of products in that sector (CN8 – CPA codes) in a given year. Formula: $IP_s^t = \frac{\sum_{p \in s} IMP_p^t}{\sum_{i \in s}\left[y_i^t - EXP_i^t\right] + \sum_{p \in s} IMP_p^t}$ where i denotes a firm in sector s, p the product(s) in the corresponding sector and t the time period, IMPpt denote imports of product p in year t, EXPit denotes exports of firm i in year t, and yit denotes the total turnover of firm i in year t.
3. Result tables in sectoral database		ID_IMPEN_NACE&nace
4. Source data used		TU_NBB_IMPEXP_&year. (2000-2010) from the Sectoral DB, containing data on imports and exports by product type from the NBB. TU_CNVN_CN_CPA2008 from the Sectoral DB: conversion table between yearly CN8 codes and CPA 2008 codes.
5. Availability		2000-2010
6. 1 Variable1	**Name**	CD_NACE&nace
	Label	x-digit sector of activity
	Formula	First x digits of the NACE sector
	Comments	The number of digits of the NACE code can be defined via a parameter &nace. The NACE code has been substracted as the first x digits of the NACE codes (for turnover) and the first x digits of CPA 2008 codes (for imports).
6. 2 Variable2	**Name**	MS_IMPEN_&year
	Label	Import penetration
	Formula	The import penetration indicator, calculated with the formula above
	Comments	
7. Methodology		Please refer to the final report
8. Literature		Please refer to the final report
9. Last exercise		June 2011
10. Responsible		Daniel Neicu
11. Coverage		All NACE 2-3-4 digit sectors over 2000-2010
12. Reliability		See comments under section 14

13. Annexe(s)	
14. Remarks(s)	

SECTOR / MARKET INDICATOR FORM		
1. Name		Volatility of market shares
2. Description		A company's individual volatility index for a given year is the difference in market shares of that company from the year before, divided by the average market share of the company over the two years (year of analysis and the year before). An individual company's market share in a given year is its domestic turnover for that year, divided by the total domestic turnover for that year of all the companies in that sector. The sectoral indicator *VI* for a given sector in a selected year is computed by summing, for those companies that have been in the top4 (by market shares) in a sector in the selected period, their individual volatility indexes for the selected period and dividing this by the total number of companies involved (that were in the top4 by market shares). Note that there can be less than four companies in the top four in sectors with less than four companies in total. Formula: $$\mathrm{VOLAT}_s^t = \frac{1}{\sum_{i \in S} \delta_i^t} \sum_{i \in S} \delta_i^t \frac{\left\| m_i^t - m_i^{t-1} \right\|}{\left[\frac{m_i^t + m_i^{t-1}}{2} \right]}$$ where m_i^t is the share of company i in the sector turnover in period t and δ_i^t is a dummy variable taking value one for company i if this company belongs to the top 4 in sector s in year t. **Note: Companies' missing market shares (in periods of inactivity) are not taken into consideration in the formula.* Refer to row 14 – Remarks for further information on the use of this indicator.
3. Result tables in sectoral database		ID_VOLAT_IDX_Yr_NACE&nace
4. Source data used		TU_SEL_AGGREGATES_&YEAR (2000 - 2009), containing data on domestic Turnover (DT) = an estimation of the turnover in Belgium, calculated as the difference between Selected Turnover and Total Exports (based on data from the NBB). TU_BR_ACTIVE_&YEAR (2000-2009) from the Sectoral DB, containing active companies within the selected period.
5. Availability		2001-2009
6. 1 Variable1	**Name**	CD_NACE&nace
	Label	x-digit sector of activity
	Formula	First x digits of the NACE sector
	Comments	The number of digits of the NACE code can be defined via a parameter &nace.
6. 2 Variable2	**Name**	ID_MAX_VOLAT_IDX_CO_&year
	Label	Maximum company volatility index
	Formula	See above
	Comments	

6. 3 Variable3	**Name**	ID_MIN_VOLAT_IDX_CO_&year
	Label	Minimum company volatility index
	Formula	See above
	Comments	
6. 4 Variable4	**Name**	ID_STDEV_VOLAT_IDX_&year
	Label	Standard deviation of company volatility index
	Formula	See above
	Comments	
7. Methodology		Please refer to the final report
8. Literature		Please refer to the final report
9. Last exercise		May 3 2011
10. Responsible		Daniel Neicu
11. Coverage		All NACE 2-3-4 digit sectors over 2001-2009
12. Reliability		See remarks under section 14
13. Annexe(s)		
14. Remarks(s)		• This indicator is to be included in the **composite indicator** calculation; it is different from the fixed period volatility in the sens that it uses a two-year moving computational period, so that the Volatility Index in year *t* is based on data from years *t* and *t-1*. • Companies with NACE codes composed of less than 4 digits are recoded as having missing NACE codes.

<table>
<tr><th colspan="3">SECTOR / MARKET INDICATOR FORM</th></tr>
<tr><td colspan="2">1. Name</td><td>Volatility of market shares</td></tr>
<tr><td colspan="2">2. Description</td><td>A company’s individual volatility index for a selected period (range of years) is the sum of the difference in market shares of that company between two consecutive years within that period, divided by the average market share of the company over the selected period.

An individual company’s market share in a given year is its domestic turnover for that year, divided by the total domestic turnover for that year of all the companies in that sector.

The sectoral indicator VI for a given sector in a selected period is computed by summing, for those companies that have been at least once in the top4 (by market shares) in a sector in the selected period, their individual volatility indexes for the selected period and dividing this by the total number of companies involved (that were ever in the top4 by market shares).

Formula:

$$VOLAT_s^t = \frac{1}{\sum_{i \in S} \delta_i^t} \sum_{i \in S} \delta_i^t \frac{\left| m_i^t - m_i^{t-1} \right|}{\left[\frac{m_i^t + m_i^{t-1}}{2} \right]}$$

where m_i^t is the share of company i in the sector turnover in period t and δ_i^t is a dummy variable taking value one for company i if this company belongs to the top 4 in sector s in year t.

*Note: Companies’ missing market shares (in periods of inactivity) are not taken into consideration in the formula.

Refer to row 14 – Remarks for further information on the use of this indicator.</td></tr>
<tr><td colspan="2">3. Result tables in sectoral database</td><td>ID_VOLAT_IDX_&firstyear_&lastyear_NACE&nace,

ID_TRANSITION_MATRIX_NACE_&nace (optional summary)</td></tr>
<tr><td colspan="2">4. Source data used</td><td>TU_SEL_TRNOV_&YEAR (2000 - 2009), containing data on domestic Turnover (DT) = an estimation of the turnover in Belgium, calculated as the difference between Selected Turnover and Total Exports (based on data from the NBB).

TU_BR_ACTIVE_&YEAR (2000-2009) from the Sectoral DB, containing active companies within the selected period.</td></tr>
<tr><td colspan="2">5. Availability</td><td>2000-2009 (one indicator per sector for the entire period)</td></tr>
<tr><td rowspan="4">6. 1 Variable1</td><td>Name</td><td>CD_NACE_&nace</td></tr>
<tr><td>Label</td><td>x-digit sector of activity</td></tr>
<tr><td>Formula</td><td>First x digits of the NACE sector</td></tr>
<tr><td>Comments</td><td>The number of digits of the NACE code can be defined via a parameter &nace.

As the indicator is dynamic, companies are assigned to the same sector during the entire selected period in order to avoid misleading data on differences in market shares only resulting from changes in NACE codes. The NACE code is assigned by determining the most frequently assigned NACE code for each company during the selected period.</td></tr>
<tr><td>6. 2</td><td>Name</td><td>NR_COMP</td></tr>
</table>

Variable2	**Label**	Number of companies
	Formula	Count of firms appearing in the top4 in a sector within the selected period
	Comments	
6. 3 Variable3	**Name**	VOLAT_IDX_SECT
	Label	Sectoral volatility index
	Formula	See above
	Comments	
6. 4 Variable4	**Name**	MAX_VOLAT_IDX_CO
	Label	Maximum company volatility index
	Formula	See above
	Comments	
6. 5 Variable5	**Name**	MIN_VOLAT_IDX_CO
	Label	Minimum company volatility index
	Formula	See above
	Comments	
6. 6 Variable6	**Name**	STD_VOLAT_IDX
	Label	Standard deviation of company volatility index
	Formula	See above
	Comments	
7. Methodology		Please refer to the final report
8. Literature		Please refer to the final report
9. Last exercise		January 21 2011
10. Responsible		Daniel Neicu
11. Coverage		All NACE 2-3-4 digit sectors over 2000-2009
12. Reliability		See comments under section 14
13. Annexe(s)		
14. Remarks(s)		• This indicator is to be included in the **quick scan** methodology; it is different from the moving periods volatility in the sens that it uses the entire available period (2000-2009) as computational basis, so that the Volatility Index in year *t* is based on data from years *t-n* to *t*. • Our definition of volatility does not directly capture the fact that companies change

	their sector of activity. Indeed, as the volatility of market shares is a dynamic indicator capturing changes over time, we choose for each company its most frequent atributed NACE code over the selected period. • Companies with NACE codes composed of less than 4 digits are recoded as having missing NACE codes.

SECTOR / MARKET INDICATOR FORM		
1. Name		Price-cost Margin (PCM)
2. Description		Profitability measures the difference between the revenues obtained from output and the expense associated with consumption of inputs. Price-cost margin is the difference between price (p) and marginal cost (mc) as a fraction of price ([p-mc]/p). It is usually taken as an indicator of market power because the larger the margin, the larger the difference between price and marginal cost, that is, the larger the distance between the price and the competitive price. The price-cost margin depends on the elasticity of demand and it is also called the Lerner index of market power. Formula: $L_S = \sum_{i \in s} w_i \frac{p_i - MC_i}{p_i}$ for sector s with $w_i = \frac{q_i}{\sum_{k \in s} q_k} = \frac{q_i}{q_S}$
3. Result tables in sectoral database		ID_PCM_NACE&NACE (NACE2, NACE3 and NACE4)
4. Source data used		TU_NBB_&YEAR (2000 - 2009), containing data on: a) Raw materials (code 60/61, 60, 61) = raw materials, consumables, services and other goods b) Labour costs (code 62) = remuneration, social security costs and pensions c) Turnover (code 70)
5. Availability		2000-2009
6. 1 Variable1	**Name**	MS_W_PCM_&YEAR
	Label	
	Formula	W_PCM (weighted average price-cost margin in the sector) 1) Calculate each firm's variable cost=raw materials + social security 2) Calculate each firm's profit= turnover-variable costs 3) W_PCM=sum(each firm's profit in the sector)/sum(each firm's turnover in the sector)
	Comments	
6. 2 Variable2	**Name**	MS_NO_OF_FIRMS_&YEAR
	Label	
	Formula	
	Comments	No of firms= counting the number of firms in corresponding sector based on firms which have raw materials, social security and turnover data.
7. Methodology		See the Final Report for details
8. Literature		Refer to the Final Report
9. Last exercise		June 20, 2011
10. Responsible		Cherry Cheung
11. Coverage		All NACE 2-3-4 digit sectors over 2000-2009
12. Reliability		
13. Annexe(s)		

14. Remarks(s)	1) Only those companies are included that have a NACE code of at least 4 digits (>=4). All other firms (including those with missing NACE code) are dropped. Further analysis on the NACE can be made, e.g. which sector has the most missing NACE… 2) Alternative Calculation of PCM based on other literatures can be done in the future.

SECTOR / MARKET INDICATOR FORM		
1. Name		Labor Productivity
2. Description		Labor productivity LP_s in sector s at time t is calculated as the sum of the value-added VA_j (Euros/hour) of each firm j in the sector at time t over the total number of hours worked H_s in the sector at time t, including both employees and independents: $$LP_{st} = \frac{\sum_{j \in s_t} VA_{jt}}{H_{st}}$$ In order to allow for increased comparability across heterogeneous sectors, growth in labor productivity $\Delta LP_{st} = (LP_{st} - LP_{s,t-1})/LP_{s,t-1}$ is preferred as a measure over absolute levels. Besides the labor productivity in nominal terms, the indicator is also calculated in real terms by using price deflators for NACE 2 digit from 2001 to 2009
3. Result tables in sectoral database		ID_LP_NACE&NACE (NACE2, NACE3 and NACE4)
4. Source data used		TU_NBB_&YEAR (2000 - 2009), containing data on: a) Value added (code 9800) b) Number of hours actually worked: total (full-time and part-time) (code 1013) TU_RSZ_EMPLOYEES_&YEAR (2000-2009), containing data on: a) Number of paid days for full-time workers b) Number of paid hours for part time workers
5. Availability		2000-2009
6. 1 Variable1	**Name**	MS_W_LP_&YEAR
	Label	
	Formula	Sum of value added of each firms in the sector /total number of worked hours in the sector including both employees and independents
	Comments	
6. 1 Variable3	**Name**	MS_W_ LP_CH_&YEAR
	Label	
	Formula	MS_CH_LP (Changes of the Labor Productivity)= (MS_W_LP in year i – MS_W_LP in year i-1)/MS_W_LP in year i
	Comments	
6. 2 Variable4	**Name**	MS_W_LP_RVA
	Label	
	Formula	Sum of value added of each firms in the sector in real term /total number of worked hours in the sector including both employees and independents
	Comments	
7. Methodology		See the Final Report for details
8. Literature		Refer to the Final Report
9. Last exercise		June 20, 2011
10. Responsible		Cherry Cheung
11. Coverage		All NACE 2-3-4 digit sectors over 2000-2009
12. Reliability		
13. Annexe(s)		
14. Remarks(s)		

A DECISION TREE
AS A QUICK SCAN
FOR EFFECTIVE MARKET FUNCTIONING*

CHERRY CHEUNG
HU Brussels

KRISTIEN COUCKE
HU Brussels and Catholic University of Leuven

DANIEL NEICU
HU Brussels and Catholic University of Leuven

Abstract: This paper presents a logical decision tree structure to screen industries for possible malfunctioning using a strategic set of indicators reflecting potential, international and internal competition. Based on this conditional combination of a limited set of key indicators, we classify industries into different groups with a low or high probability that market malfunctioning is present.

*The authors would like to thank Leo Sleuwaegen for his valuable comments on the ideas embodied in this paper.

I. Introduction

The aim of this paper is to generate an analytical framework which can be used for screening manufacturing industries to detect possible competition problems like for instance collusive behaviour. In previous studies (see for instance Office of Fair Trading, 2004; European Commission, 2007) an extensive list of relevant indicators related to innovation, productivity and profitability are calculated for all sectors. However, this extensive approach limits the ability to combine information from the different indicators. Instead of using an extensive list of relevant indicators, we propose to use a logical decision tree structure based on a limited set of important indicators to select possible problematic markets. Our strategic set of indicators will focus on the presence of potential, international and internal competition within industries. The contribution of this paper is twofold. The advantage of using this logical decision tree structure is that it sequentially eliminates those industries where the probability of market malfunctioning is low, following well established theory. Secondly, for those sectors that are classified as possible problematic markets, the method suggests if further in-depth investigation is necessary either at the national or international level.

Going back from the oldest literature in industrial organization, including the Structure Conduct Performance Paradigm (Mason, 1939; Bain, 1951), to more recent work, it is remarkable to observe that no solid relationships have been established linking industry characteristics and market competition and dynamics within sectors, as mentioned by Aiginger and Pfaffermayr (2000). Following the results of many empirical studies, competitive dynamics appear to be strongly dependent on the specific industry context (see for instance Gibrat, 1931; Mansfield, 1962; Schmalensee, 1989; Sutton, 1997; Machado and Mata, 2000). Therefore, recent literature focuses on 'single industry studies’ using ‘structural estimation’ techniques where one specific industry is studied providing estimates for the specific model’s parameters (Reiss and Wolak, 2004).

Our paper has a different set up. We do not try to characterize the competitive interaction that takes place, but we try to sort industries following a limited set of important indicators that have been found to be associated with limited competition in the industry. In the antitrust literature a debate is going on about static versus dynamic competitive interactions, depending on the relevant time framework that is taken into account. We construct a logical decision tree structure based on both concepts of competition and we classify industries into different groups with a low or high probability that market malfunctioning is present, either at a static or dynamic level.

The remainder of the paper is structured as follows. In the first part of the paper, we will focus on the analytical framework that will be used for screening. Section II explains the importance of the different concepts of competition within our logical decision tree structure. In section III we construct the analytical framework that will lead to the classification into different groups of industries. The second part of the paper applies this analytical framework using extensive Belgian micro-level data covering all manufacturing industries with NACE code 101–390 for the period 2000-2009. In section IV data issues and indicators will be described while section V presents the data results of our screening tool to detect anticompetitive behaviour in manufacturing industries. Section VI concludes.

II. Different sources of competition

Different sectors will reflect different competitive market conditions. Taking into account the underlying drivers of these competitive market processes, our decision tree structure has to be seen as a quick scan screening tool, eliminating those industries for which there is a high likelihood that sufficient competitive discipline is exercised. Only those industries failing the test, should be considered for further investigation, taking into account a more elaborate set of other indicators reflecting for instance profits of firms within these industries, productivity growth and the importance of innovative activities (which is not the aim of this paper and therefore not studied here). The next paragraphs focus on the theoretical foundations underlying the structure of the proposed decision tree.

Potential competition

One of the most influential theories in the last decades concerning market performance has been the contestable market literature developed by the American economist William Baumol. Following the contestable market theory (Baumol, Panzar and Willig, 1982), entry is the most important dynamic competitive force, ensuring efficient outcomes even if the industry consists of only a small set of firms. If entry is easy because no barriers to enter the industry are present, incumbent firms have to fear the continuous pressure of possible new competition (Geroski et al, 1988). The threat of new firm entry forces incumbent firms to act competitively, leaving no profit potential for the entrant. This "potential competition" drives incumbent firms towards lower profit margins (Bresnahan and Reiss, 1991), more productive efficiency (Baldwin and Gorecki, 1991), more innovative activities (Geroski and Jacquemin, 1985) or product differentiation (Shapiro and Varian, 1999). Failing to do so, they will pushed out of the industry by newcomers.

Using another perspective, according to the view of the Austrian school of economics, competition should basically be seen as a dynamic process. Competition will be strong and sustained when new firms or entrepreneurs engage in competitive behaviour. In his work, Schumpeter (1942) already emphasized the importance of dynamic competition and called this continuous process of new firms entering the market a process of creative destruction, driving incumbent firms towards efficiency, innovation or upgrading of their products. However, the intensity of creative destruction is related to the life cycle of industries (Acs and Audretsch, 1987; Jovanovic and Chung-Yi Tse, 2006). When entry dynamics are changing substantially over the observation period, it reflects an evolution of the industry moving to the next stage of its life cycle. In that sense, potential competition coming from creative destruction is to some extent path dependent.

International competition

If there is no strong indication that substantial entry, measured on a national base, occurred over time, the next stage investigates if actual competition occurs within the market. However, this raises the question of delineating the relevant market. A crucial tool in competition policy is indeed the definition of the relevant market. The relevant market is the set of products and geographical areas on which firms if acting collusively, could exercise substantial market power (Areeda and Turner, 1978; Horowitz, 1981; Stigler and Sherwin, 1985; Fisher, 2002). As a result of the globalization of trade, relevant markets have widened and the international dimension has raised many competition issues affecting more than one jurisdiction (Salinger, 1990; Gal, 2009).

Belgium is a very open economy where national competition is more and more influenced by the presence of increased imports and strategies of multinational firms (De Backer and Sleuwaegen, 2003). As a result of European integration, following the elimination of trade and production barriers, industry-restructuring processes have been taken place on a European or global level yielding again an indication that relevant markets have widened (Aiginger, 2000; Bowen and Sleuwaegen, 2007). If indeed there is evidence of important cross border flows of goods and services, it is unlikely that the relevant market will be national based and that relevant competition can be studied at a national scale (Sleuwaegen and Van Cayseele,1998; Bishop and Walker, 2002; Simons and Williams, 1993; Massey, 2000).

Hence, if the national market is not the relevant arena for competition, and there is evidence that the industry is characterized by substantial international transactions, these industries will be excluded from further analysis at the national level. If the actual degree of competitive interaction is strong and power of incumbent firms on the relevant market is small, competition is deemed to be effective (White, 2000; Werden, 2000). This is not the same as saying that all these industries that we exclude from further analysis at the national level, are competitive. For industries showing weak potential competition at the national level, there is a danger of a lack of effective competition at a wider geographical scope. Such cases should typically be dealt with by EU competition authorities, in which case the analysis should be done at EU level or if relevant, at the global level.

Internal (within industry) competition

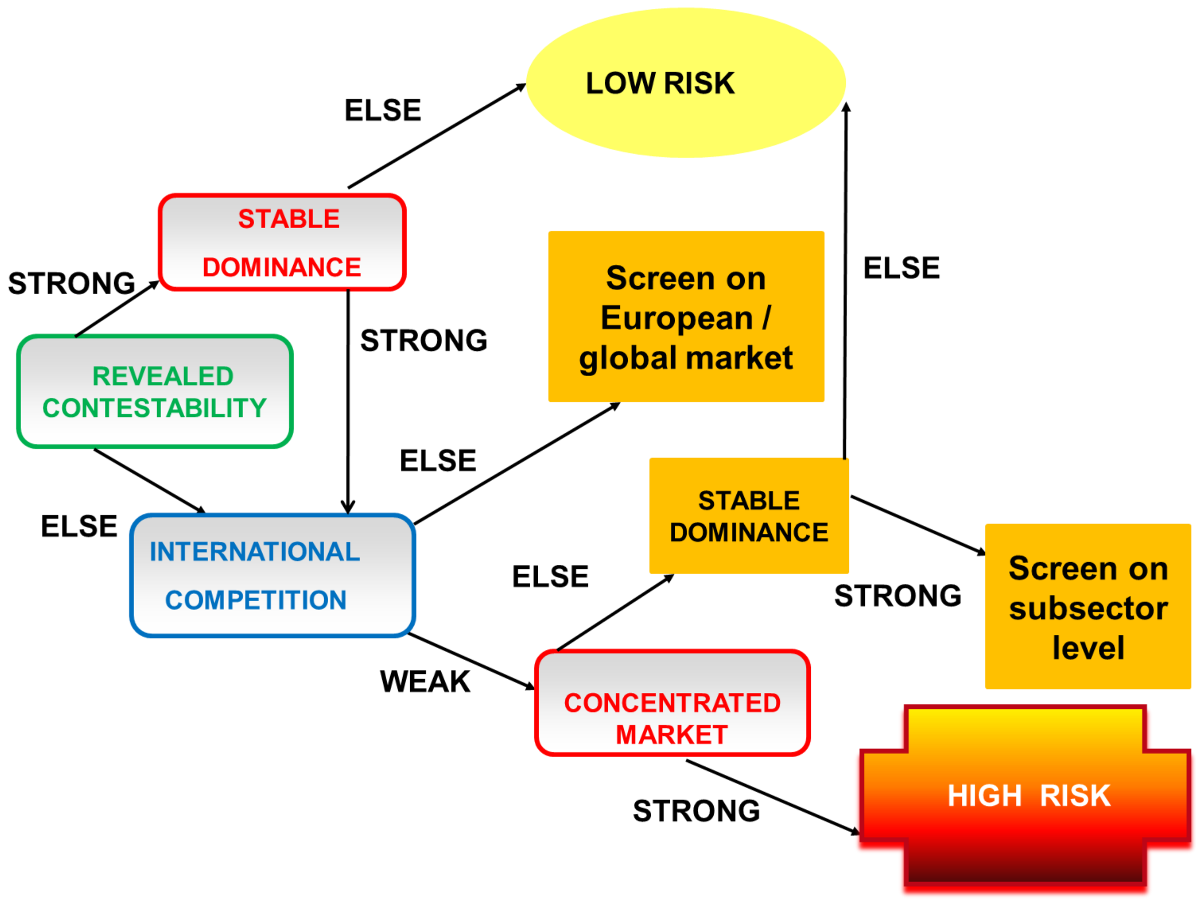
If the relevant market turns out to be national, the next step consists of investigating the competitive conditions prevailing in that market. The focus is typically on market structure. In the traditional industrial organization literature, market structure is generally studied using concentration measures (Mason, 1939; Bain, 1956). However, it is often the case that a traditional concentration ratio does not show substantial change over a time interval (Baldwin, 1995) while considerable industry dynamics still exists, indicated by market share mobility (Davies and Geroski, 1997; Geroski, Machin and Walters, 1997). Therefore, market share mobility and concentration ratios are complementary measures, as these two indicators reveal different aspects of the competitive process within industries (Baldwin and Gorecki, 1994).

As been found in many empirical studies, the dominance of some incumbent firms in an industry characterized by a low volatility of market shares and high concentration ratios, is often associated with collusive behaviour (Shepherd, 1976; Stigler, 1982; Tirole, 1988; Scherer and Ross, 1990; Posner, 2003; Carlton and Perloff, 2005). Considering the Efficient-Structure hypothesis developed by Demsetz (1973) and Peltzman (1977), there is still the possibility that effective competition takes place. However, a direct measurement of all the conditions under which collusive versus competitive behaviour occurs, is rather complicated. In our screening tool, we only point to the high probability that collusive behaviour will occur if dominance goes together with a lack of potential and international competition. Hence, for industries where substantial dominance occurs while market dynamics are absent, we recommend an in depth investigation, based on detailed evidence of the pricing and other market behaviour of incumbent firms (Schmalensee, 1989; Mason, Phillips and Nowell, 1992; Ivaldi, Jullien, Rey, Seabright and Tirole, 2003).

III. Analytical framework

Figure 1 combines the theoretical arguments developed in the previous section in a logical decision tree structure.

Figure 1: Decision tree structure



First node: Revealed contestability

Contestability theory (Baumol, Panzar and Willig, 1983) holds that if the market is open, meaning that there are no or low barriers to enter, new rivals will enter the market when incumbent firms earn excess profits. Such competition includes firms that enter the industry and potential competition by firms that would enter the industry if incumbent firms would set high prices. Since we cannot observe potential competition in our screening tool, we use the actual entry as an indicator of “revealed contestability” and assume that potential competition is highly correlated with this measure. Entry is measured with national firm-level data as the registration of new firms in the industry. A high entry rate typically reflects the absence of important entry barriers (Baldwin and Gorecki, 1991). When an industry reveals strong contestability, the industry will be driven towards a competitive outcome as incumbent firms are unable to exercise market power. Moreover, under continuous competitive pressure of new entrants, incumbent firms are forced to use their production inputs in the most cost efficient way, with a maximal productive efficiency (Helpman, Melitz and Yeaple, 2003; Bernard, Redding and Schott, 2004; De Backer and Sleuwaegen, 2003).

Second node upper part: Stable dominance

Following contestable market theory, in industries with strong revealed contestability, entry will ensure efficient outcomes even if the industry consists of only a small set of firms. However, industries with low structural and regulatory entry barriers do not necessarily reflect competitive markets when incumbent firms can act strategically. More specifically, the strategic behaviour and actions of incumbent firms could distort the competitive process giving entering firms no chance to survive, a possibility that is ruled out by the assumptions of perfectly contestable markets (Martin, 2000). Strategic barriers, generated by the behaviour of incumbent firms for the purpose of pushing new entrants out, are for instance exclusive dealing arrangements, high advertising expenditures, building of overcapacity or drastic price cuts (Aghion and Bolton, 1987; Rasmusen et al., 1991; Segal and Whinston, 2000). In such situations, the revealed contestability as we measure it, is not necessarily reflecting the contestability of the whole industry. If entrants are forced to immediately abandon the market or to operate on a small base, serving only a small niche in the market (Dunne, Roberts and Samuelson, 1988; Hopenhayn 1992), contestability will be very partial and imperfect. Therefore, in industries where revealed contestability does not go together with significant overall market share volatility, competition may still turn out to be imperfect with entry happening only at the fringe, and include possible collusive behaviour among a set of dominant firms in this industry (Cable, 1998). Consequently, industries where strong revealed contestability goes together with stable dominance, reflected by strong market share stability, will not be excluded from further investigation but are referred back to the lower part of the decision tree.

For those industries where strong contestability is revealed and no stable dominance is observed, competition appears dynamic and no structural, regulatory or strategic barriers seem to distort the market functioning. This group of industries is classified as having a low probability that market malfunctioning is present.

Second node lower part: International competition

If there is no revealed strong contestability of the industry, the next question is what happens with competition “within“ the market? In those cases important attention should be given to the definition of the relevant market, i.e. the market where firms actually compete. The data we start from, measures entry on a national base. However, Belgium is an open economy characterized by substantial international competitive pressure in many industries (Coucke and Sleuwaegen, 2008). In industries where imports make up a significant part of the market, the relevant market is deemed to extend Belgian borders.

For such industries revealing weak contestability at the national level, but characterized by a wide relevant market, there is a danger of a lack of effective competition at a wider geographical scope. However, such cases should typically be dealt with by EU competition authorities, in which case the analysis should be done at the European level or if relevant, at the global level.

Third node lower part: Concentrated market

If we do not reveal strong contestability in the industry and the relevant market is national or local, it is important to look at the competitive conditions that prevail in that market. If the market is closed and concentrated, the likelihood that incumbent firms have market power and engage in collusive behavior is high (Landes and Posner, 1981). This type of industry, characterized by a lack of potential, international and internal competition will be classified as having a high risk of market malfunctioning. Those industries are selected for further in-depth study, to investigate if profits in these markets are excessively high and productivity levels and productivity growth limited.

Fourth node lower part: Stable dominance

Finally, there are industries where the relevant market is national or local and apparent concentration measured at the national level is low. For such industries there is still the danger that the measurement is done at the wrong level: too wide industry (NACE3 instead of NACE4) or geography (national instead of regional or local). This is likely to be the case for industries where the concentration ratio score is low but where the volatility of market shares is also very low, reflecting a strong market share stability of firms in the industry. This combination of the two indicators could reflect dominance of some firms in different subsectors of the NACE 3 industry or segments of the national market.

IV. Data and indicators

This section describes the data and defines the indicators related to revealed contestability, international competition, stable dominance and concentrated market.

Source

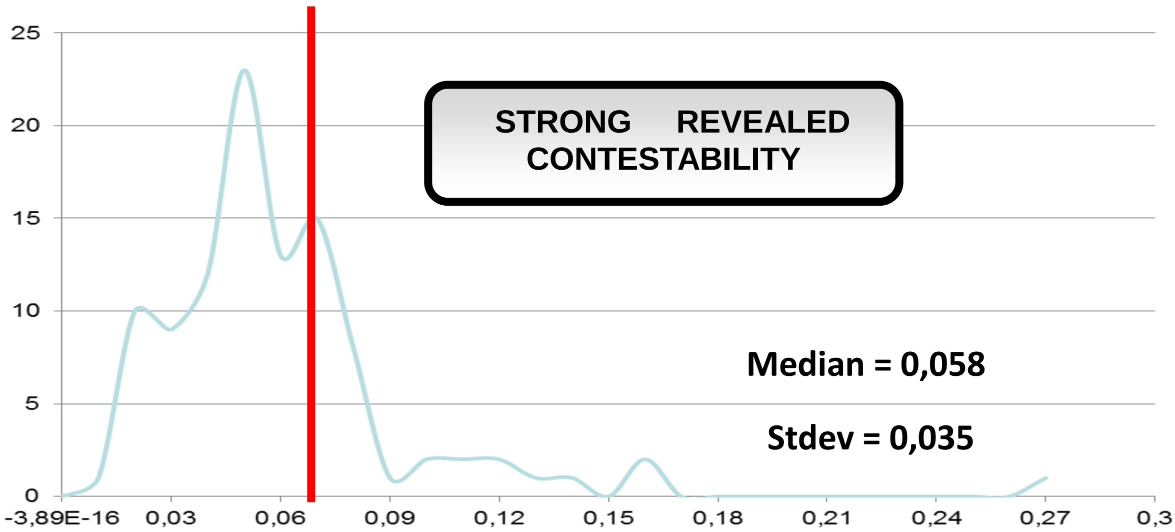
Data are retrieved from different data sources, including VAT declaration, Company Accounts from the National Bank of Belgium, Structural Business Survey and PRODCOM data. All these sources provided by federal institutions (NIS/ADSEI/NBB) contain firm-level data and are available at the FOD Economie. The dataset is aggregated to NACE-3 level[1] covering all the manufacturing industries with NACE code 101–390 in the period 2000-2009. In our analysis, 104 industries and 40.713 manufacturing firms are in the sample.

[1] We recommend to use our analytical framework with the NACE-4 industry level classification but due to data limitations for import penetration based on PRODCOM data, we apply the NACE-3 industry level.

Revealed contestability

Revealed contestability is considered as strong if the average entry rate of the industry over the observation period belongs to the highest tertile of the distribution of average entry rates across all industries (higher than approximately 7% per year).

Figure 2: Number of industries (vertical axis) based on average entry rate 2001-2009 (horizontal axis)



The entry rate is defined as the ratio of the number of firms that enter an industry in a specific year to the number of active firms in that industry in the same year. An entry is defined as a firm that reports a positive turnover for the first time in an industry.[2]

International competition

Exposure to international competition is measured by import penetration which reflects how much of the domestic consumption is supplied through imports (Aiginger and Pfaffermayr, 2000). Import penetration in an industry

[2] Since an entering firm has not reported turnover in the preceding year, we start to calculate the entry rate from the year 2001.

is defined as total imports in the industry divided by total imports plus total turnover minus total exports for the industry (Davis et al, 1996). Imports are obtained from PRODCOM data which classify imports by product category and not by firm-level imports in order to correct for raw materials and intermediate products. International competition is considered weak when the average import penetration of the industry over the observation period belongs to the lowest tertile of the distribution of the average import penetration across all industries, implying that less than approximately 5% of domestic consumption is supplied through imports. Since approximately 15% of all manufacturing industries report no imports based on the PRODCOM data, additional data from the National Bank of Belgium were used, clarifying that the industries that report no imports in PRODCOM have indeed a negligible share of imports in the industry.

Stable dominance

For those industries with a strong revealed contestability, stable dominance reflecting possible strategic anticompetitive behaviour of incumbent firms, is measured by the volatility of market shares. The volatility of market shares (VMS) is in fact an index of relative market share instability (Caves and Porter, 1978; Sakakibara and Porter, 2001; Masatoshi and Yuji, 2006) measured by the average relative changes in domestic market share of the leading firms in an industry over the observation period. We define it as

$$VMS_s = \frac{\sum_{comp^s=1}^{m} \sum_{t=2000}^{2009} \left[\left| MS_{comp^s}^{t+1} - MS_{comp^s}^{t} \right| / AVG(MS_{comp^s}) \right]}{m}$$

with $MS_{comp^s}^{t}$ the market share of a leading firm in year t and active in sector s, defined by its domestic turnover for that year, divided by the total domestic turnover for that year of all the firms in that sector. A firm is selected as a leading firm in an industry when it belongs to the top four largest firms based on domestic market shares, in at least one year of the observation period. The index , $comp^s$ stands for leading firms in sector s. The relative change in domestic market share of a leading firm is measured by the absolute value of the annual domestic market share change, divided by the average domestic market share of that firm during the observation period. This sum of relative changes in domestic market shares is summed up for all leading firms and divided by m, since this number of leading firms m is different between industries. VMS which is directly related to market conduct, can detect possible dominance of one single player or a selected group of players when strong stable dominance is observed in the industry. Stable dominance is strong in an industry when the volatility of market shares over the observation period belongs to the lowest tertile of the distribution of VMS across all industries[3].

Concentrated market

For those industries with a weak revealed contestability and national or local relevant markets, we measure concentration by the Herfindahl Index (HHI) calculated as the sum of squared domestic market shares[4] held by all firms in an industry. An industry is strongly concentrated when the average HHI of the industry over the observation period belongs to the highest tertile of the distribution of the average HHI across all industries, implying an average HHI of more than approximately 1800.

---

[3] The same indicator is used to classify those industries that should be screened on subsector level.

[4] We correct the HHI for exports sales similar to Sleuwaegen and Van Cayseele (1998).

V. Results

High risk industries

Table 1 presents the group of industries that have a high probability of market malfunctioning , lacking apparent potential, international and internal competition. Industries in the table are ranked in decreasing order of HHI but we divide the high risk industries further into two subgroups:

Table 1: High risk industries

NACE	*INDUSTRY DESCRIPTION*
	Subgroup 1
390	Remediation activities, waste management services
351	Electric power generation, transmission and distribution
302	Manufacture of railway locomotives and rolling stock
273	Manufacture of wiring and wiring devices
235	Manufacture of cement, lime and plaster
304	Manufacture of military fighting vehicles
360	Water collection, treatment and supply
233	Manufacture of clay building materials
	Subgroup 2
352	Manufacture of gas; distribution of gaseous fuels through mains
353	Steam and air conditioning supply
254	Manufacture of weapons and ammunition
279	Manufacture of other electrical equipment
370	Sewerage

Applying our analytical framework, all the industries in Subgroup 1 and Subgroup 2 have a high risk on market malfunctioning since these industries are strongly concentrated and revealed contestability is not strong or seems only to happen at the fringe, while international competition is weak. The difference between Subgroup 1 and Subgroup 2 is related to some difference in stable dominance based on the market share volatility threshold. In the industries of Subgroup 1 strong stable dominance is observed[5] meaning that there is market share stability of the leading firms in the industry. In the

[5] Following our definition of strong stable dominance when the volatility of market shares of the industry belongs to the lowest tertile.

industries of Subgroup 2 the volatility of market shares is lower than the average VMS of all industries but higher than the lowest tertile. For all the industries that are selected as high risk industries in Table 1 a high HHI goes together with strong (or almost strong) stable dominance. Cable (1997) indeed showed that market share stability is not independent of the Herfindahl index.

Some of the industries listed in Table 1 are industries where the process of deregulation is just in an early stage and where the foreseen increase in competition is not observed yet during the observation period. Other industries listed in Table 1 are characterized by the fact that the government is usually the largest domestic buyer and where newcomers or foreign companies are in many cases excluded because of national safety or other political reasons (Stigler, 1982b). As such, government procurement can play an important role in the determination of market structure and consequently in the ability of new firms entry (Scherer and Ross, 1990; Geroski, 1991; Kovacic, 1992, Duggan and Scott-Morton, 2006). Finally, some of the industries classified by our analytical framework as high risk industries for market malfunctioning are highly ranked based on the number of antitrust cases in the U.S. and Europe (www.ftc.gov; www.justice.gov; ec.europa.eu/competition/antitrust).

Industries to be screened on European (or global) level

Table 2 presents the group of industries where a national screening tool for market functioning is less relevant since competition takes place at a wider geographical scope. Industries in the table are ranked based on their NACE 3 code.

Table 2: Industries to be screened on European (or global) level

NACE	*INDUSTRY DESCRIPTION*
102	Processing and preserving of fish, crustaceans and molluscs
104	Manufacture of vegetable and animal oils and fats
120	Manufacture of tobacco products
131	Preparation and spinning of textile fibres
151	Tanning and dressing of leather; manufacture of luggage, handbags
152	Manufacture of footwear
191	Manufacture of coke oven products
192	Manufacture of refined petroleum products
202	Manufacture of pesticides and other agrochemical products
203	Manufacture of paints, varnishes and similar coatings, printing ink and mastics
204	Manufacture of soap and detergents, cleaning and polishing preparations
206	Manufacture of man-made fibres
211	Manufacture of basic pharmaceutical products
212	Manufacture of pharmaceutical preparations
221	Manufacture of rubber products
234	Manufacture of other porcelain and ceramic products
237	Cutting, shaping and finishing of stone
239	Manufacture of abrasive products and non-metallic mineral products
241	Manufacture of basic iron and steel and of ferro-alloys
244	Manufacture of basic precious and other non-ferrous metals
259	Manufacture of other fabricated metal products
263	Manufacture of communication equipment
264	Manufacture of consumer electronics
265	Manufacture of instruments and appliances for measuring and testing
266	Manufacture of irradiation, electromedical and electrotherapeutic equipment
267	Manufacture of optical instruments and photographic equipment
268	Manufacture of magnetic and optical media
272	Manufacture of batteries and accumulators
275	Manufacture of domestic appliances
281	Manufacture of general-purpose machinery
291	Manufacture of motor vehicles
301	Building of ships and boats
303	Manufacture of air and spacecraft and related machinery
321	Manufacture of jewellery, bijouterie and related articles

Some of the industries listed in Table 2 are industries where structural entry barriers are characteristic to production conditions (Caves and Porter, 1976; Baumol et al, 1986; Eaton and Lipsey, 1980; Lambson, 1991). In these industries, firms have to bear large fixed sunk costs to enter a sector (Sutton, 1998) and where incumbent firms often want to exploit scale economies supplying the whole European market in order to operate in the most profitable way.

Other industries listed in Table 2 are industries where incumbent firms are forced to rationalize productive operations as a result of increased international competition. In these industries an important share of production activities (mostly labour-intensive production activities) is already relocated abroad and replaced by imports (Coucke, Pennings and Sleuwaegen, 2007), reflecting the situation that these industries have come to the last stage of their life cycle (Gereffi, 1999). Many local firms survive the increased international competition in these industries by increasing their productivity, upgrading their activities or differentiating their products from imported goods (De Backer and Sleuwaegen, 2001).

However, there is still a danger of a lack of effective competition at a wider geographical scope and therefore such cases should be screened by EU competition authorities, in which case the screening should be done at the European level or if relevant, at the global level.

Low risk industries

Table 3 presents the group of low risk industries ranked based on their NACE 3 code. Some of the industries listed in Table 3 are strongly growing industries where entry and innovation are complementary phenomena that are to a large extent driven by the evolution of the industry over the life cycle (Acs and Audretsch, 1987) driven by new product technologies (Gort and Klepper, 1982; Klepper and Grady, 1990).

Table 3: Low risk industries

NACE	*INDUSTRY DESCRIPTION*
101	Processing and preserving of meat and production of meat products
103	Processing and preserving of fruit and vegetables
132	Weaving of textiles
133	Finishing of textiles
139	Manufacture of other textiles
141	Manufacture of wearing apparel, except fur apparel
142	Manufacture of articles of fur
143	Manufacture of knitted and crocheted apparel
161	Sawmilling and planing of wood
162	Manufacture of products of wood, cork, straw and plaiting materials
182	Reproduction of recorded media
201	Manufacture of basic chemicals, fertilisers and nitrogen compounds
205	Manufacture of other chemical products
222	Manufacture of plastics products
231	Manufacture of glass and glass products
232	Manufacture of refractory products
242	Manufacture of tubes, pipes, hollow profiles and related fittings of steel
243	Manufacture of other products of first processing of steel
251	Manufacture of structural metal products
252	Manufacture of tanks, reservoirs and containers of metal
253	Manufacture of steam generators, except central heating hot water boilers
255	Forging, pressing, stamping and roll-forming of metal, powder metallurgy
256	Treatment and coating of metals, machining
257	Manufacture of cutlery, tools and general hardware
261	Manufacture of electronic components and boards
262	Manufacture of computers and peripheral equipment
274	Manufacture of electrical lighting equipment
282	Manufacture of other general-purpose machinery
283	Manufacture of agricultural and forestry machinery
284	Manufacture of metal forming machinery and machine tools
292	Manufacture of bodies for motor vehicles, trailers and semi-trailers
322	Manufacture of musical instruments
323	Manufacture of sports goods
324	Manufacture of games and toys
325	Manufacture of medical and dental instruments and supplies
329	Manufacturing n.e.c.
331	Repair of fabricated metal products, machinery and equipment
332	Installation of industrial machinery and equipment
383	Materials recovery

Other industries listed in Table 3 are industries where incumbent firms face strong international competition (Pennings and Sleuwaegen, 2006) but where local entrepreneurship survived through upgrading of the production activities or by differentiating products from imported goods, supplying a specific niche in the market (Coucke, 2007).

Industries to be screened on subsector level

Table 4: Industries to be screened on subsector level

NACE	***INDUSTRY DESCRIPTION***
	Subgroup 1
105	Manufacture of dairy products
106	Manufacture of grain mill products, starches and starch products
107	Manufacture of bakery and farinaceous products
108	Manufacture of other food products
109	Manufacture of prepared animal feeds
110	Manufacture of beverages
171	Manufacture of pulp, paper and paperboard
172	Manufacture of articles of paper and paperboard
181	Printing and service activities related to printing
236	Manufacture of articles of concrete, cement and plaster
245	Casting of metals
271	Manufacture of electrical motors, generators, and electricity distribution
289	Manufacture of other special-purpose machinery
293	Manufacture of parts and accessories for motor vehicles
309	Manufacture of transport equipment n.e.c.
310	Manufacture of furniture
381	Waste collection
382	Waste treatment and disposal

Table 4 presents the group of industries for which there is a possible danger that the measurement is done at a too wide industry level (NACE 3 instead of NACE 4). These industries have a low HHI but at the same time a strong market share stability in the industry, possibly pointing to dominance and stable leadership in the different subsectors of the NACE 3 aggregation.

VI. Conclusions

In this paper we constructed a structured decision model that can serve as a "quick scan" screening tool for assessing effective market functioning. Our logical decision tree structure is based on four strategic indicators that focus on the presence of potential, international and internal competition within industries. The contribution of this paper is twofold. The advantage of using this logical decision tree structure is that it sequentially eliminates those industries where the probability of market malfunctioning is low, following well established theory. Secondly, for those sectors that are classified as possible problematic markets, the method suggests if further in-depth investigation is necessary either at the national or international level. For possible problematic markets a more extensive list of indicators related to innovation, productivity and profitability should be taken into account, similar to previous studies done at the industry level (Office of Fair Trading, 2004; European Commission, 2007).

The second part of the paper applied the analytical framework using extensive data covering all Belgian manufacturing industries with NACE code 101–390 for the period 2000-2009. In the group of industries that were selected as high risk industries, some industries are characterized by an early stage of market deregulation while other industries belonging to this group are characterized by substantial public procurement. It is analytically reassuring to observe that many of the industries classified as high risk industries are also those industries that have been subject to a significant number of anti-trust investigations in the U.S. and Europe.

VII. References

Acs, Z.J. and D.B. Audretsch (1987). Innovation, market structure and firm size. *Review of Economics and Statistics*, 69, 567-575.

Aghion, P. and P. Bolton (1987). Contracts as a Barrier to Entry. *American Economic Review*, 77, 388-401.

Aiginger, K. (2000). Specialisation of European Manufacturing. *Austrian Economic Quaterly*, 2, 81-92.

Aiginger, K. and M. Pfaffermayr (2000). Profitability differences, market power and competition policy in transition countries – an overview of methodology and econometric models of market power and concentration. *PHARE ACE project NO P-2035.*

Areeda, P.E. and D.F. Turner (1978). Antitrust Law. *Little, Brown and Company, Boston.*

Bain, J.S. (1951). Relation of Profit Rate to Industry Concentration. *Quarterly Journal of Economics,* 65, 293-324.

Bain, J.S. (1956). Barriers to New Competition. Cambridge MA: Harvard University Press.

Baldwin, J.R. (1995). The Dynamics of Industrial Competition: a North American Perspective. *Cambridge: Cambridge University Press*

Baldwin J.R. and P.K. Gorecki (1991). Firm entry and exit in the Canadian manufacturing sector, 1970-198., *The Canadian Journal of Economics*, 24, 300-323.

Baldwin J.R. and P.K. Gorecki (1994). Concentration and Mobility Statistics in Canada's Manufacturing Sector. *The Journal of Industrial Economics*, XLII, 93-103.

Baumol, W., J. Panzar and R. Willig (1982). Contestable Markets and the Theory of Industry Structure, San Diego: Harcourt Brace Jovanovich.

Baumol, W., J. Panzar, and R. Willig (1983). Contestable Markets: an uprising in the theory of industry structure, *American Economic Review*, 73(3), 491-496.

Baumol, W. and R. Willig (1986). Contestability: developments since the book, *Oxford Economic Papers*.

Bernard, A. B., Redding S. and P. Schott (2004). Product Choice and Product Switching. Tuck School of Business Mimeo, *revision of NBER Working Paper No. 9789.*

Bishop, S. and M. Walker (2002). The Economics of EC Competition Law: Concepts, Application and Measurement. *Sweet and Maxwell.*

Bowen, H. and L. Sleuwaegen (2007). European integration: The third step. *International Economics and Economic Policy*, 4, 241-262.

Bresnahan, T. and P. Reiss (1991). Entry and Competition in Concentrated Markets. *Journal of Political Economy,* 99, 977-1009.

Cable, J.R. (1997). Market Share Behavior and Mobility: An Analysis and Time-Series Application. Review of Economics and Statistics, 79, 136-141

Cable, J.R. (1998). Market Share Dynamics and Competition: a Survey of the Empirical Literature. *Aberystwyth Economic Research Papers, School of Management and Business.* The University of Wales, Aberystwyth, No 98-7.

Carlton, D. and J. Perloff (2005). Modern Industrial Organization, 4th Edition. *Pearson, Addison Wesley.*

Caves, R. and M.R. Porter (1976). Barriers to Exit, *in Essays in Industrial Organization in Honor of Joe Bain, R. Masson and P. Qualss, Ballinger,* Cambridge MA, 39-70.

Caves, R., and M. E. Porter (1978). Market Structure, Oligopoly, and Stability of Market Shares. *Journal of Industrial Economics*, 26, 289-313.

Coucke, K. (2007). Globalisation and De-Industrialisation in Belgium: on the Role of Imports and International Sourcing. *Tijdschrift voor Economie en Management*, LII, 13-36.

Coucke, K., Pennings E. and L. Sleuwaegen (2007). Employee layoff under different modes of restructuring: exit, downsizing or relocation. *Industrial and Corporate Change.* Oxford University Press, 16, 161-182.

Coucke, K. and L. Sleuwaegen (2008). Offshoring as a survival strategy: evidence from firms in Belgian manufacturing. Journal of International Business Studies, 39, 1261-1277.

Davies, S.W. and P.A. Geroski (1997). Changes in Concentration, Turbulence, and the Dynamics of Market Shares. *The Review of Economics and Statistics,* 79, 383-391.

Davis, S.J., Haltiwanger, J.C. and S. Schuh (1996). Job Creation and Destruction. *Cambridge: MIT Press.*

De Backer, K. and L. Sleuwaegen (2001). Multinational Enterprises, Market Integration and Trade Structure: What Remains of the Standard Goods Hypothesis, *Weltwirtschaftliches Achiv*,137, 379-403.

De Backer, K. and L. Sleuwaegen (2003). Foreign Ownership and Productivity Dynamics, *Economics Letters*,79, 177-183.

Demsetz, H. (1973). Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, 16, 1-9.

Duggan, M. and F. Scott-Morton (2006). The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing. *The Quarterly Journal of Economics,* 121, 1-30.

Dunne, T., Roberts M.J. and L. Samuelson (1988). Patterns of Firm Entry and Exit in U.S. Manufacturing Industries, *The RAND Journal of Economics*, 19, 495-515.

Eaton, B. and R. Lipsey (1980). Exit Barriers are Entry Barriers: the Durability of Capital as Barrier to Entry, *Bell Journal of Economics*, 11, 721-729.

European Commission. (2007). Implementing the new methodology for product market and sector monitoring: Results of a first sector screening. *Commission staff working document SEC (2007) 1517*.

Fisher, F. (2002). Market Definition: A User's Guide. *Finnish Competition Authority*.

Gal, M. (2009). Antitrust in a Globalized Economy: The Unique Enforcement Challenges Faced by Small and of Developing Jurisdictions. *33 Fordham International L.J.1*

Gereffi, G. (1999). International trade and industrial upgrading in the apparel commodity chain. *Journal of International Economics*, 48, 37-70.

Geroski, P.A. and A. Jacquemin (1985). Industrial change, barriers to mobility, and European industrial policy. *Economic Policy*, 1, 169-204.

Geroski, P., Gilbert, R. and A. Jacquemin (1988). Entry and Strategic Competition, *London: Harwood.*

Geroski, P., Machin, S. and C. Walters (1997). Corporate Growth and Profitability. *Journal of Industrial Economics*, 45, 171-189.

Gibrat, R. (1931), Les Inegalit´es Economiques. Applications: Aux In´egalit´es des Richesses, a la Concentration des Entreprises, Aux Populations des Villes, Aux Statistiques des Familles, etc., d'une Loi Nouvelle: La Loi de l'Effect Proportionnel, Paris: Sirey.

Gort, M. and S. Klepper (1982). Time paths in the diffusion of product innovations. *Economic Journal*, 92, 630-653.

Helpman, D., M. J. Melitz and S. R. Yeaple (2003). Export versus FDI, *NBER Working Paper No. 9439.*

Hopenhayn, H. A. (1992). Entry, Exit and Firm Dynamics in Long Run Equilibrium. *Econometrica*, 60, 1127-1150.

Horowitz, L. (1981). Market Definition in Antitrust Analysis. *Southern Economic Journal*, 48, 1-16.

Ivaldi, M., Jullien, B., Rey, P., Seabright, P. and J. Tirole (2003). Final Report for DG Competition, European Commission. *IDEI, Toulouse.*

Jovanovic, B. and C-Y Tse (2006). Creative Destruction in Industries. *NBER Working Papers 12520, National Bureau of Economic Research, Inc.*

Klepper, S. and E. Grady (1990). The Evolution of New Industries and the Determinants of Market Structure, *Rand Journal of Economy*, 21(1), 27-44.

Kovacic, W.E. (1992). The Influence of Economics on Antitrust Law. *Economic Inquiry*, 30, 294-306.

Lambson, V. (1991). Industry evolution with sunk costs and uncertain market conditions, *International Journal of Industrial 0rganization*, 9, 171-196.

Landes, W.M. and R.A. Posner (1981). Market Power in Antitrust Cases. *Harvard Law Review*, 94, 937-996.

Machado, J. and J. Mata (2000). Box-Cox Quantile Regression and the Distribution of Firm Sizes, *Journal of Applied Econometrics*, 15, 253-274.

Mansfield, E. (1962). Entry, Gibrat's Law, Innovation, and the Growth of Firms, *American Economic Review*, 52, 1023–1051.

Martin, S. (2000). The Theory of Contestable Markets. *Department of Economics, Purdue University.*

Masatoshi, K. and H. Yuji (2006). Market Share Instability and the Dynamics of Competition: a Panel Data Analysis of Japanese Manufacturing Industries. *Review of Industrial Organization*, 28, 165-182.

Mason, E.S. (1939). Price and Production Policies of Large-Scale Enterprise. *American Economic Review*, 29, 61-74.

Mason, C., Philips, O. and C. Nowell (1992). Duopoly Behavior in Asymmetric Markets: An Experimental Evaluation. *The Review of Economics and Statistics*, 74, 662-670.

Massey, P. (2000). Market Definition and Market Power in Competition Analysis – Some Practical Issues. *The Economic and Social Review*, 31, 309-328.

Office of Fair Trading. (2004). Empirical indicators for market investigations. *OFT749a*, see http://www.oft.gov.uk.

Peltzman, S. (1977). The Gains and Losses from Industrial Concentration. *Journal of Law and Economics*, 20, 229-263.

Pennings, E. and L. Sleuwaegen (2006). International Relocation of Production: Where Do Firms Go? *Scottish Journal of Political Economy*, Scottish Economic Society, 53, 430-446.

Posner, R. (2003). Antitrust Law: An Economic Perspective. Chicago: *Chicago University Press.*

Rasmusen, E. B., J. M. Ramseyer and J. S. Wiley (1991). Naked Exclusion, *American Economic Review,* 81, 1137-1145.

Reiss, P. and F. Wolak (2004). Structural econometric modeling: Rationales and examples from industrial organization.

Sakakibara, M. and M.E. Porter (2001). Competing at Home to Win Abroad: Evidence from Japanase Industry. *Review of Economics and Statistics*, 83, 310-322.

Salinger, M. (1990). The Concentration-Margins Relationship Reconsidered, in *Brookings Papers: Microeconomics*, 287-321.

Shapiro, C. and H. Varian (1999). Information rules: a Strategic Guide to the Network Economy. *Boston: Harvard Business School Press.*

Simons, J. and M. Williams (1993). The Renaissance of Market Definition. *Antitrust Bulletin*, 38, 799-857.

Shepherd, W. (1976). Bain's influence on research into industrial organization. In Robert T. Masson and P. David Qualls (eds.), *Essays on Industrial Organization in Honor of Joe S. Bain.* Cambridge, MA: Ballinger, 1-17.

Scherer, F. and D. Ross (1990). Industrial Market Structure and Economic Performance. *3rd edn. Boston: Houghton-Mifflin.*

Schmalensee, R. (1989). Inter-industry studies of structure and performance, in: R. Schmalensee and R. Willig (eds), *Handbook of Industrial Organization*, Volume 2, chapter 16, 951-1010.

Schumpeter, J. A. (1942). Capitalism, Socialism and Democracy, *Cambridge, Harvard University Press.*

Segal, I. R. and M.D. Whinston (2000). Naked Exclusion: Comment, *American Economic Review,* 90, 296-309.

Sleuwaegen, L. and P. Van Cayseele (1998). Competition Policy in Belgium. *In: Martin S. (Eds.), Competition policies in Europe, Contributions to economic analysis*. Amsterdam: North Holland, 185-204.

Stigler, G. (1982). The Economists and the Problem of Monopoly. *American Economic Review*, 72., 1-11.

Stigler, G. (1982b). The Economist as Preacher and Other Essays. *Chicago: University of Chicago Press.*

Stigler, G. and R. Sherwin (1985). The Extent of the Market. *Journal of Law and Economics*, 28, 555-585.

Sutton, J. (1997). Gibrat's Legacy, *Journal of Economic Literature*, 35, 40-59.

Sutton, J. (1998). Technology and Market Structure. Cambridge, MA: MIT Press.

Tirole, J. (1988). The Theory of Industrial Organization. Cambridge, MA: MIT Press.

Werden, G. (2000). Market Delineation under the Merger Guidelines: Monopoly Cases and Alternative Approaches. *Review of Industrial Organization*, 16, 211-218.

White, L. (2000). Present at the Beginning of a New Era for Antitrust: Reflections on 1982-1983. *Review of Industrial Organization*, 16, 131-149.

Entry and Competition in Differentiated Products Markets

Catherine Schaumans and Frank Verboven*

April, 2011

Abstract

We propose a methodology for estimating the competition effects from entry when firms sell differentiated products. We first derive precise conditions under which Bresnahan and Reiss' entry threshold ratios (ETRs) can be used to test for the presence and to measure the magnitude of competition effects. We then show how to augment the traditional entry model with a revenue equation. This revenue equation serves to adjust the ETRs by the extent of market expansion from entry, and leads to unbiased estimates of the competition effects from entry. We apply our approach to seven different local service sectors. We find that entry typically leads to significant market expansion, implying that traditional ETRs may substantially underestimate the competition effects from entry. In most sectors, the second entrant reduces markups by at least 30%, whereas the third or subsequent entrants have smaller or insignificant effects. In one sector, we find that even the second entrant does not reduce markups, consistent with a recent decision by the competition authority.

* *Catherine Schaumans*: Tilburg University. E-mail: catherine.schaumans@uvt.nl. *Frank Verboven*: Catholic University of Leuven and C.E.P.R. (London). E-mail: Frank.Verboven@econ.kuleuven.be. We are grateful to Michelle Goeree, Pasquale Schiraldi, Philipp Schmidt-Dengler, Johan Stennek, Jo Van Biesebroeck and participants at workshops at the FOD Economie, Tilburg University, and the Annual C.E.P.R. conference in Toulouse and for helpful comments and discussions. The project was financed by the Belgian Federal Science Policy Office (Agora) and realized in collaboration with the Federal Public Service Economy (Sector and Market Monitoring Department). We thank the FPS Economy for its contribution in setting-up the dataset and Daniel Neicu for excellent research assistance.

1 Introduction

An important question in industrial organization is how market structure affects the intensity of competition. To address this question a variety of empirical approaches have been developed, each with different strengths and weaknesses depending on the available data.[1] Bresnahan and Reiss (1991) developed an innovative approach applicable to local service sectors: they infer the effects of entry on competition from the relationship between the number of entrants and market size. The intuition of their approach is simple. If market size has to increase disproportionately to support additional firms, entry can be interpreted to intensify the degree of competition. Conversely, if market size increases proportionally with the number of firms, then additional entry is interpreted to leave the degree of competition unaffected. To implement their approach, Bresnahan and Reiss propose the concept of the entry threshold ratio (henceforth ETR). The ETR is the percentage per-firm market size increase that is required to support an additional firm. An estimated ETR greater than 1 indicates that entry leads to stronger competition, whereas an ETR equal to 1 indicates that entry does not intensify competition.

A major strength of Bresnahan and Reiss' methodology is that it can be applied with relatively modest data requirements. One basically needs data on a cross-section of local markets, with information on the number of firms per market, population size and other market demographics as control variables. No information on prices or marginal costs is required. This also makes their approach potentially appealing from a competition policy perspective. It can be used as a first monitoring tool to assess which sectors potentially face competition problems and require more detailed investigation.

A central assumption of Bresnahan and Reiss' methodology is that firms produce homogeneous products: holding prices constant, an additional entrant only leads to business stealing and does not create market expansion. This assumption is potentially problematic since new entrants may be differentiated from existing firms, either because they offer different product attributes or because they are located at a different place. In both cases, additional entry would raise demand (holding prices constant).

In this paper we develop a more general economic model to assess the competition effects from entry. The model allows for the possibility that firms sell differentiated products, i.e. additional entry can create market expansion. We first derive precise conditions under which Bresnahan and Reiss' ETRs can be used as a test for the *presence* of competition effects from entry. We find that this is only possible if products are homogeneous, i.e. additional

[1] For detailed overviews see, for example, Bresnahan (1989), Ackerberg, Benkard, Berry and Pakes (2007) and Reiss and Wolak (2007).

entry only entails business stealing and no market expansion. We then ask when ETRs can be used as a measure for the *magnitude* of competition entry effects. We show that ETRs are generally a biased measure for the percentage markup effect due to entry, except in the special case where products are homogeneous and the price elasticity of market demand is unity. More generally, if products are sufficiently differentiated, ETRs typically tend to underestimate the percentage markup effects from competition.

Our theoretical framework also provides a natural way to extend the Bresnahan and Reiss' approach to obtain an unbiased measure for the magnitude of the markup effects due to entry. We propose to augment the traditional ordered probit entry model with a revenue equation. The entry model specifies the equilibrium number of firms that can be sustained under free entry. The revenue equation specifies per firm revenues as a function of the number of firms and enables one to estimate the total market expansion effects (consisting of both the direct effects from increased product differentiation and any indirect effects through possible price changes). To obtain an unbiased estimate of the markup effects from entry, the traditional ETRs from the entry model should be suitably adjusted by the total market expansion effects estimated from the revenue equation.

To implement our approach, we study a variety of local service sectors, for which revenue data are increasingly becoming available.[2] More specifically, we consider architects, bakeries, butchers, florists, plumbers, real estate agents and restaurants. For each sector, we constructed a cross-section dataset of local markets (towns) in Belgium, with information on market revenues, the number of entrants, market size (population) and market demographics. Estimating the single-equation entry model yields the traditional ETRs, and we estimate these to be close to 1. This would seem to indicate that entry does not lead to intensified competition. In fact, we even estimate some ETRs to be below 1, which would be inconsistent with the hypothesis of increased competition. However, estimation of the revenue equation shows that entry may often lead to important total market expansion, especially for architects, florists and real estate agents. This implies that the traditional ETRs underestimate the competition effects from entry. Accounting for the estimated total market expansion effects leads to stronger competition effects, especially from the second entrant. Third and subsequent entrants have more limited or insignificant competition effects. In one

[2]The increased access to revenue data has recently also been exploited in a variety of other settings. For example, Syverson (2004) uses plant-level revenue data in the ready-mixed concrete industry, to assess how demand factors affect the distribution of productivity. Campbell and Hopenhayn (2005) consider the relationship between market size and the size distribution of establishments. They find that establishments tend to be larger in large markets, consistent with models of large-group competition. Konings, Van Cayseele and Warzynski (2005) and De Loecker and Warzynski (2010) extend Hall's (1988) approach to estimate markups using plant-level data on revenues in combination with variable input expenditures.

sector, bakeries, we find no significant competition effects, not even from the second entrant. Incidentally, this sector has recently been investigated by the local competition authority because of price fixing concerns.

Our paper relates to the growing empirical literature on static entry models. Bresnahan and Reiss (1991) proposed their ordered probit model of free entry to infer competition effects from entry by doctors, dentists, car dealers and plumbers. Asplund and Sandin (1999) and Manuszak (2002) are examples of applications of this model to other sectors. Berry (1992) considered a more general model of entry with heterogeneous firms. Mazzeo (2002), Seim (2006) and Schaumans and Verboven (2008) allow for multiple types of firms or endogenize the choice of type. Other recent work on static entry models has focused on different ways of addressing the multiplicity problem in entry games with firm heterogeneity; see Berry and Reiss (2007) for a recent overview of the literature. In contrast with this recent literature, we maintain the basic entry model that can be applied to market-level data and we focus on the interpretation of ETRs. We show how to augment the entry model with a revenue equation to draw more reliable inferences about the competition effects from entry.

Section 2 presents the theoretical framework, showing under which conditions ETRs can be used as a test for the presence and a measure for the magnitude of competition effects. Section 3 presents the econometric model and Section 4 the empirical analysis. Finally, Section 5 concludes.

2 Theoretical framework

We first describe the model. We then introduce the concept of the ETR, and derive conditions under which ETRs can be used to test for the *presence* of competition effects from entry. Finally, we show how to incorporate revenue data to adjust ETRs to measure the *magnitude* of competition effects from entry in an unbiased way.

2.1 The model

There are N firms, competing in a local market with a population size S. Each firm has the same constant marginal cost $c > 0$ and incurs a fixed cost $f > 0$ (independent of the number of firms).

Demand Firms do not necessarily produce homogeneous products, but in equilibrium they charge the same industry price p. The demand per firm and per capita as a function of this common price p and the number of firms N is $q(p, N)$. This is the traditional

Chamberlinian DD curve (in per capita terms). Similarly, industry demand per capita is $Q(p, N) = q(p, N)N$. Denote the price elasticity of industry demand by $\varepsilon = -Q_p \frac{p}{Q} = -q_p \frac{p}{q}$. We ignore the fact that N can only take integer values here, but we take this into account in the empirical analysis.

We make the following three assumptions about demand.

Assumption 1 $q_p \leq 0$, *or equivalently,* $Q_p = q_p N \leq 0$.

Assumption 2 $q_N \leq 0$.

Assumption 3 $Q_N = q + q_N N \geq 0$.

The first assumption simply says that per-firm or industry demand is weakly decreasing in the common industry price p. The second assumption says that per-firm demand is weakly decreasing in the number of firms N: holding prices constant, additional entry either leads to business stealing (if products are substitutes) or does not affect per-firm demand (if products are independent). Finally, the third assumption says that industry demand is weakly increasing in N: holding prices constant, entry either leads to market expansion because of product differentiation, or leaves industry demand unaffected if products are homogeneous.

These assumptions clearly cover the special case in which products are homogeneous, as in Bresnahan and Reiss (1991). In this case, industry demand per capita can be written as $Q(p, N) = D(p)$, so that $q(p, N) = \frac{D(p)}{N}$. It immediately follows that $q_N = -q/N < 0$ and $Q_N = q + q_N N = 0$. Hence, with homogeneous products entry leads to full business stealing and no market expansion (holding prices constant).

More generally, the assumptions allow for product differentiation with symmetric firms. To illustrate, consider Berry and Waldfogel's (1999) symmetric nested logit model used to study product variety: the first nest includes all firms' products, and the second nest contains the outside good or no-purchase alternative. With identical firms and identical prices, the nested logit per firm and per capita demand function is:

$$q(p, N) = \frac{N^{-\sigma}}{e^{\alpha p} + N^{1-\sigma}},$$

where $\alpha > 0$ is the price parameter and $0 \leq \sigma \leq 1$ is the nesting parameter. It can easily be verified that:

$$\begin{aligned} q_p &= -\alpha(1 - Nq) < 0 \\ q_N &= -\left(\sigma + (1-\sigma)q\right)\frac{q}{N} < 0 \\ Q_N &= (1-\sigma)q(1-q) \leq 0. \end{aligned}$$

If $\sigma = 1$, then $q_N = -q/N$ and $Q_N = 0$, so all firms' products are perceived as homogeneous (relative to the outside good).

Profits and prices Now consider profits and the symmetric equilibrium price in the market. For a common industry price p a firm's profits are

$$\pi = (p - c)\, q(p, N)S - f.$$

Suppose first that all N firms behave as a cartel. In this case, the equilibrium price as a function of N is $p^m(N)$, defined by the first-order condition

$$q(p, N) + (p - c)\, q_p(p, N) = 0.$$

More generally, let the symmetric equilibrium price as a function of the number of firms N be given by $p(N) \leq p^m(N)$. In many oligopoly models, including the Cournot and Bertrand models, this equilibrium price is weakly decreasing in N, $p' \leq 0$. We can then write a firm's equilibrium profits as a function of the number of firms N as:

$$\pi(N) = (p(N) - c)\, q(p(N), N)S - f. \tag{1}$$

In the next two subsections we will decompose profits in two different ways. Define the variable profits per firm and per capita by $v(N) \equiv (p(N) - c)\, q(p(N), N)$, the revenues per firm and per capita by $r(N) \equiv p(N)q(p(N), N)$, and the Lerner index or percentage markup by $\mu(N) \equiv \frac{p(N)-c}{p(N)}$. We can then write

$$\begin{aligned} \pi(N) &= v(N)S - f. &(2) \\ &= \mu(N)r(N)S - f. &(3) \end{aligned}$$

The expression on the first line contains variable profits per firm and per capita, similar to Bresnahan and Reiss (1991). The expression on the second line rewrites variable profits as markups times revenue per firm and per capita. As we will show in the next two subsections, this second expression provides useful additional information to assess the effects of competition on markups, provided that data on revenues are available.

2.2 ETRs to test for the presence of competition effects

Bresnahan and Reiss (1991) introduce the concept of the entry threshold and entry threshold ratio as a test for the presence of competition effects from entry. The entry threshold is the critical market size required to support a given number of firms, and is derived from the zero-profit condition $\pi(N) = 0$. Using (2), this gives

$$S = \frac{f}{v(N)} \equiv S(N).$$

Bresnahan and Reiss argue that entry does not lead to increased competition if the entry threshold increases proportionally with the number of firms. For example, entry would not lead to more competition if a doubling of the market size is required to support twice as many firms. Conversely, entry creates intensified competition if the entry threshold increases disproportionately with the number of firms. For example, competition intensifies if a tripling of the market size would be required to support twice as many firms.

Based on this intuition, Bresnahan and Reiss propose the entry threshold *ratio*, or ETR, as a unit-free measure to test for the presence of competition effects. The ETR is defined as the per-firm entry threshold required to support N firms, relative to the per-firm entry threshold to support $N-1$ firms, i.e.

$$ETR(N) \equiv \frac{S(N)/N}{S(N-1)/(N-1)}. \tag{4}$$

One can then test the null hypothesis, $ETR(N) = 1$, that the N-th entrant does not lead to more competition.

We now assess this interpretation formally, starting from our more general model where products are not necessarily homogeneous, i.e. allowing for market expansion upon entry. Substituting $S(N) \equiv \frac{f}{v(N)}$ in (4), we can write the ETR in a simple form:

$$\begin{aligned} ETR(N) &= \frac{v(N-1)(N-1)}{v(N)N} \\ &\equiv \frac{V(N-1)}{V(N)}. \end{aligned} \tag{5}$$

where $V(N) = v(N)N$ is per capita industry variable profits. The ETR is therefore just the ratio of industry variable profits with N and $N-1$ firms.

It follows immediately from (5) that the $ETR(N) > 1$ if and only if $V'(N) < 0$, i.e. if and only if industry variable profits are strictly decreasing in N. To see under which circumstances this is the case, differentiate $V(N) = v(N)N$ using (1), and rearrange to obtain

$$\begin{aligned} V' &= (q + (p-c)q_p)\, p' N + (p-c)\,(q + q_N N) \\ &= (1 - \mu\varepsilon)\, p' N q + (p-c)\,(q + q_N N)\,. \end{aligned} \tag{6}$$

Suppose first that products are homogeneous, which is the special case considered by Bresnahan and Reiss. In this case, $q + q_N N = 0$ so that the second term in (6) vanishes. Since $1 - \mu\varepsilon \geq 0$, it follows that $V' < 0$ (and hence $ETR(N) > 1$) if and only if $p' < 0$. Similarly, $V' = 0$ if and only if $p' = 0$. We can therefore confirm, and make more precise, Bresnahan and Reiss' justification for using ETRs as a test for the presence of competition effects from entry, when products are homogeneous:

Proposition 1 *Suppose that products are homogenous. $ETR(N) > 1$ if and only if entry leads to a price decrease ($p' < 0$). $ETR(N) = 1$ if and only if entry does not affect the price ($p' = 0$).*

Bresnahan and Reiss also provide examples from oligopoly models to argue that the ETRs are declining in N. Intuitively, entry may be expected to have larger effects on competition if one starts off from few firms with strong market power, as can be confirmed from examples such as the Cournot model. Formally, it follows from (5) that the ETRs are declining if and only if the industry variable profits are convex in N, $V'' > 0$. While this may often be the case, it is not generally true, not even if products are homogeneous. A simple counterexample is a repeated game with price setting firms: profits are monopoly profits for sufficiently low N, and then drop to zero above a critical level for N.[3]

Suppose now that products are differentiated. This means that additional entry implies market expansion (holding prices constant), i.e. $q + q_N N > 0$, so that the second term in (6) becomes positive. It follows immediately that $V' > 0$ (and hence $ETR(N) < 1$) if $p' = 0$. Furthermore, $V' > 0$ is also possible if $p' < 0$, provided products are sufficiently differentiated (since then p approaches p^m or μ approaches $1/\varepsilon$, so that the first term in (6) vanishes and the second term dominates). We can conclude the following about the use of entry thresholds when products are differentiated:

Proposition 2 *Suppose products are differentiated. $ETR(N) < 1$ if entry does not affect the price ($p' = 0$) or even if entry leads to a price decrease ($p' < 0$) provided products are*

[3] In fact, with homogeneous products one can verify that for small N the function V is concave ($V'' < 0$), while for sufficiently large N the function V is convex. In a linear demand Cournot model, the function is convex for $N \geq 2$. So ETRs appear to be increasing for N very small. Yet accounting for the fact that N is an integer, the ETR already drops when moving from 1 to 2 firms.

sufficiently differentiated.

Product differentiation can thus explain occasional findings in applied work of ETRs less than 1. (For example, Bresnahan and Reiss report $ETR(3) = 0.79$ for dentists.) Intuitively, if entry leads to substantial market expansion and does not intensify competition by very much, it is possible that market size increases less than proportionately with the number of firms.

To summarize, Propositions 1 and 2 identify conditions under which the null hypothesis $ETR(N) = 1$ is reasonable as a test for the presence of competition effects. It turns out that this approach is reasonable only if products are homogeneous, but not more generally if products are differentiated.

2.3 ETRs to measure the magnitude of competition effects

Having identified conditions under which ETRs form a reasonable basis to test for the presence of the competition effects from entry, we now ask under which conditions ETRs provide an unbiased measure for the magnitude of the competition effects. Define this magnitude as the percentage drop in the Lerner index, $\mu(N-1)/\mu(N)$.

To address this question, we now start from (3) instead of (2) to rewrite the entry threshold as

$$S(N) = \frac{f}{\mu(N)r(N)}.$$

This can be substituted in the definition of the ETR (4) to rewrite it as:

$$\begin{aligned} ETR(N) &= \frac{\mu(N-1)}{\mu(N)} \frac{r(N-1)(N-1)}{r(N)N} \\ &\equiv \frac{\mu(N-1)}{\mu(N)} \frac{R(N-1)}{R(N)} \end{aligned} \tag{7}$$

where $R(N) = r(N)N$ is the per capita industry revenue function.

It immediately follows that the ETR is an exact measure for the magnitude of the percentage markup drop if and only if industry revenues do not vary with the number of firms, $R(N) = R(N-1)$, i.e. if and only if $R' = 0$ (ignoring that N only takes integer values). Similarly, the ETR underestimates (overestimates) the percentage markup drop if and only if $R' > 0$ ($R' < 0$). To see when this is the case, use $R(N) = p(N)q(p(N), N)N$ to compute

$$\begin{aligned} R' &= (q + pq_p)\, p'N + p\,(q + q_N N) \\ &= (1-\varepsilon)\, p'Nq + p\,(q + q_N N)\,. \end{aligned} \tag{8}$$

As before, suppose first that the products are homogeneous, as in Bresnahan and Reiss. We have that $q + q_N N = 0$, so that the second term in (8) vanishes. For $p' < 0$, we then obtain that $R' < 0$ if $\varepsilon < 1$, $R' = 0$ if $\varepsilon = 1$ and $R' > 0$ if $\varepsilon > 1$. We can conclude the following:

Proposition 3 *Suppose that products are homogeneous. The ETR is a correct measure of the percentage markup drop due to entry, $ETR(N) = \mu(N-1)/\mu(N)$, if and only if $\varepsilon = 1$. It underestimates (overestimates) the percentage markup drop if and only if $\varepsilon > 1$ ($\varepsilon < 1$).*

For example, consider an estimated $ETR = 1.3$, as roughly found for entry by the second and third firm in Manuszak's study of the 19th century U.S. brewery industry. Assuming homogeneous products, this can be interpreted as a markup drop by 30% following the introduction of a second and third competitor, if and only if the price elasticity of market demand is unity.

Proposition 3 shows that it is difficult to draw general conclusions about the direction of bias, since one needs to know the level of the price elasticity of industry demand. But the direction of bias is clear in the special case where industry behaves close to a perfect cartel. In this case, we have that $\varepsilon > 1$ (since marginal cost $c > 0$). Hence, if the industry behaves close to a perfect cartel, the entry threshold would underestimate the magnitude of the markup drop following entry.

Now suppose that products are differentiated, $q + Nq_N > 0$. The second term in (8) is then positive, so that the ETR is more likely to underestimate the markup drop. More precisely, define ε^* as the critical elasticity such that $R' = 0$, i.e.

$$\varepsilon^* \equiv 1 + \frac{q + q_N N}{p' N q / p}$$

For $q + q_N N > 0$ and $p' < 0$, we have that $\varepsilon^* < 1$, so that the ETR would also underestimate the markup drop for an elasticity below 1 but sufficiently close to 1. More precisely, we have:

Proposition 4 *Suppose products are differentiated. The ETR underestimates (overestimates) the percentage markup drop $\mu(N-1)/\mu(N)$ if and only if $\varepsilon > \varepsilon^*$ ($\varepsilon < \varepsilon^*$), where $\varepsilon^* < 1$.*

To summarize, Propositions 3 and 4 imply that the ETR is more likely to underestimate the percentage markup drop from entry if the industry behaves close to a cartel (so that $\varepsilon > 1$) and/or if products are strongly differentiated (substantial market expansion from entry).

To obtain this conclusion we made use of the (per capita) industry revenue function. Provided that revenue data are available, it also suggests a natural way to obtain an unbiased measure of the competition effect from entry. Indeed, using (7) we can write the percentage markup drop as

$$\frac{\mu(N-1)}{\mu(N)} = ETR(N)\frac{R(N)}{R(N-1)}.$$

The markup drop due to entry is thus equal to Bresnahan and Reiss' ETR, multiplied by the percentage industry revenue effects from entry. In the next section, we develop an empirical model that augments the traditional entry model with a revenue function. This leads to the "adjusted ETR" as an unbiased estimate of the competition effects from entry. The approach requires market-level revenue data, in addition to data on the number of entrants and market demographics used in standard entry models.

Remark: absolute margins The above discussion focused on how to obtain an unbiased measure for the magnitude of the competition effect from entry as defined by percentage drop in the Lerner index (or percentage margin), $\mu(N-1)/\mu(N)$. One may also ask this question for the percentage drop in the *absolute* margin, $(p(N-1)-c)\,/\,(p(N)-c)$.[4] One can easily verify that (7) can be rewritten as

$$ETR(N) = \frac{p(N-1)-c}{p(N)-c}\frac{Q(N-1)}{Q(N)}.$$

The bias of the ETR as a competition measure now depends on the reduced form demand function $Q(N)$ instead of the reduced form revenue function $R(N)$. The ETR is an unbiased measure of the percentage drop in absolute margins if and only if $Q' = 0$. Similarly, the ETR underestimates (overestimates) the percentage drop in absolute margins if and only if $Q' > 0$ ($Q' < 0$). We can use $Q(N) = q(p(N), N)N$ to compute

$$Q' = -\varepsilon p' N q/p + (q + q_N N).$$

The counterparts of Proposition 3 and 4 are simple. The ETR is an unbiased estimated of the percentage drop in absolute margins only if products are homogeneous ($q + q_N N = 0$) *and* demand is perfectly inelastic ($\varepsilon = 0$). If either condition is violated, we have $Q' > 0$, so that the ETR will generally underestimate the percentage drop in absolute margins.

This discussion also shows that the appropriate measure of competition depends on data availability. With revenue data (as in most application) it is natural to focus on the percentage drop in the Lerner index $\mu(N)$. With quantity data it is natural to focus on the percentage drop in the absolute margin $p(N) - c$.

[4] We thank Johan Stennek for suggesting us to also look at this measure.

3 Econometric model

We first specify a standard empirical entry model without revenue data in the spirit of Bresnahan and Reiss (1991). We show how to estimate this model and compute ETRs, based on a dataset with the number of firms and market characteristics for a cross-section of local markets. We then show how to extend the standard entry model with a revenue equation, and how to compute adjusted ETRs as an unbiased measure of competition effects from entry.

In both cases the empirical entry model assumes that firm profits are an unobserved, latent variable. But bounds can be inferred based on the assumption that there is free entry, i.e. firms enter if and only if this is profitable.

3.1 Simple entry model

If revenue data are not available, we start from the profit function (2)

$$\pi(N) = v(N)S - f,$$

where $v' < 0$. Both the (per capita) variable profits and the fixed costs component are unobserved. However, bounds can be inferred based on the assumption that there is free entry. Upon observing N firms, we can infer that N firms are profitable, whereas $N+1$ firms are not:

$$v(N+1)S - f < 0 < v(N)S - f,$$

or equivalently

$$\ln \frac{v(N+1)}{f} + \ln S < 0 < \ln \frac{v(N)}{f} + \ln S. \tag{9}$$

Consider the following logarithmic specification for the ratio of variable profits over fixed costs

$$\ln \frac{v(N)}{f} = X\lambda + \theta_N - \omega, \tag{10}$$

where X is a vector of observable market characteristics X, θ_N represents the fixed effect of N firms, and ω is an unobserved error term.[5] Assume that $\theta_{N+1} < \theta_N < \ldots$, i.e. additional firms reduce the variable profits over fixed cost ratio (because of reduced demand and/or reduced markup). We can write the entry conditions as

$$X\lambda + \theta_{N+1} + \ln S < \omega < X\lambda + \theta_N + \ln S.$$

[5] To avoid possible confusion, in the empirical specification we use the subscript N to denote the fixed effect for the N-th firm (as in θ_N). This differs from the previous section where we used the subscript N for the partial derivative with respect to N (as in q_N).

Estimation To estimate the model by maximum likelihood, assume ω is normally distributed $\mathcal{N}(0, \sigma)$. The probability of observing N firms is

$$P(N) = \Phi\left(\frac{X\lambda + \ln S + \theta_N}{\sigma}\right) - \Phi\left(\frac{X\lambda + \ln S + \theta_{N+1}}{\sigma}\right). \tag{11}$$

This is a standard ordered probit model, where the θ_N are the "cut-points" or entry effects. Note that the variance is identified because of the assumption that variable profits increase proportionally with market size S.[6] See Berry and Reiss (2008) for a more general discussion on identification in entry models.

Constructing ETRs Based on the estimated parameters one can compute the entry threshold, i.e. the critical market size to support N firms. Using (9) and (10), evaluated at $\omega = 0$, the entry threshold to support N firms is

$$S(N) = \exp\left(-X\lambda - \theta_N\right). \tag{12}$$

The ETR is the ratio of the per-firm market size to support N versus $N-1$ firms. Using (4), this is

$$ETR(N) = \exp\left(\theta_{N-1} - \theta_N\right)\frac{N-1}{N}. \tag{13}$$

So in our logarithmic specification the ETRs only depend on the differences in the consecutive "cut-points" of the ordered probit model; they do not depend on the market characteristics X.

As shown in the previous section, the ETRs are no good measure of the competitive effects from entry if products are differentiated. Furthermore, even if products are homogenous, ETRs can only be used to test the null hypothesis of no competition effects, but not to measure the magnitude of the competition effects. These considerations motivate augmenting the entry model to include revenue data in the analysis. We turn to this next.

3.2 Simultaneous entry and revenue model

If we observe revenues per firm and per capita $r = r(N)$, we can disentangle the variable profits per capita into a percentage markup and a revenue component, $v(N) = \mu(N)r(N)$. We can then start from the profit function (3):

$$\pi(N) = \mu(N)r(N)S - f,$$

[6] Our specification differs from Bresnahan and Reiss (1991) and more closely resembles Genesove (2000). In contrast with Bresnahan and Reiss, our specification only identifies the ratio of variable profits over fixed costs and not the levels. However, we also identify the variance of the error term.

Upon observing N firms, we can now infer that

$$\mu(N+1)r(N+1)S - f < 0 < \mu(N)r(N)S - f,$$

or equivalently

$$\ln \frac{\mu(N+1)}{f} + \ln r(N+1) + \ln S < 0 < \ln \frac{\mu(N)}{f} + \ln r(N) + \ln S. \tag{14}$$

This again gives rise to the ordered probit model. But since we observe per-firm revenues $r = r(N)$, we can separately specify an equation for revenues and markups (rather than only for variable profits).

We specify revenues per capita to depend on observed market characteristics X, the number of firms N and an unobserved market-specific error term ξ. We consider both a constant elasticity and a fixed effects specification:

$$\ln r = \ln r(N) = X\beta + \alpha \ln N + \xi \tag{15}$$

$$\ln r = \ln r(N) = X\beta + \alpha_N + \xi \tag{16}$$

where X are observed market demographics ξ is an unobserved error term affecting revenues, α is the (constant) elasticity of per-firm revenues r with respect to N, and α_N are fixed entry effects.

To interpret the effect of N on r, one should bear in mind that $r(N) \equiv p(N)q(p(N), N)$. Hence, the elasticity α or the fixed effects α_N capture both the direct effect through increased product differentiation and the indirect effect through a possible price change. More formally, using (8) we can write the elasticity of r with respect to N as:

$$r' \frac{N}{r} = (1 - \varepsilon)\, p' \frac{N}{p} + q_N \frac{N}{q}.$$

The second term $q_N(N/q)$ is the direct effect through increased product differentiation. By assumptions 2 and 3, $q_N(N/q) \in (-1, 0)$: if $q_N(N/q) = -1$, products are homogeneous and there is only business stealing. If $q_N(N/q) = 0$, products are independent and there is only market expansion. The first term is the indirect effect through a possible price change. If the first term is small (because of a modest price effect $p'(N/p)$ and ε relatively close to 1), then we can interpret our estimate of $r'(N/r)$ as the extent of business stealing versus market expansion. For example, in the constant elasticity specification, an estimate of α close to –1 would indicate that entry mainly involves business stealing (homogeneous products), and α close to 0 would indicate that entry mainly involves market expansion (independent products). It will be convenient to follow this interpretation when discussing the empirical results. However, we stress that this interpretation only holds approximately, since α also captures indirect revenue effects through price changes.

Next, we specify the ratio of markups over fixed costs as a function of observed market characteristics X, the number of firms and an unobserved market-specific error term η:

$$\ln \frac{\mu(N)}{f} = X\gamma + \delta_N - \eta. \tag{17}$$

where $\delta_N > \delta_{N+1} > \ldots$, i.e. markups are decreasing in the number of firms.

Substituting the revenue specification (15) or (16) and the markup specification (17) in (14), we can write the entry conditions as

$$X\lambda + \ln S + \theta_{N+1} < \omega < X\lambda + \ln S + \theta_N,$$

where we define

$$\begin{aligned} \lambda &\equiv \beta + \gamma \\ \omega &\equiv \eta - \xi, \\ \theta_N &\equiv \alpha \ln N + \delta_N \quad \text{(constant elasticity revenue specification)} \\ &\equiv \alpha_N + \delta_N \quad \text{(fixed effects revenue specification)} \end{aligned}$$

This gives rise to the following simultaneous model for revenues and the number of firms:

for $N = 0$:

$$r \quad \text{unobserved}$$
$$X\lambda + \ln S + \theta_1 < \omega$$

for $N > 0$:

$$\ln r = X\beta + \alpha_N + \xi$$
$$X\lambda + \ln S + \theta_{N+1} < \omega < X\lambda + \ln S + \theta_N.$$

Estimation This is a simultaneous ordered probit and demand model. It has a similar structure as in Ferrari, Verboven and Degryse (2010), although they derive it from a rather different setting with coordinated entry. The model has the following endogeneity problem. We want to estimate the causal effect of N on r, but N is likely to be correlated with the demand error ξ. Econometrically, the error terms ξ and $\omega \equiv \eta - \xi$ are correlated because they contain the common component ξ. Intuitively, firms are more likely to enter in markets where they expect demand to be high, leading to spurious correlation between the number of firms and total revenues per capita $N \cdot r$, or a bias towards too much market expansion and too little business stealing. Since we will use the estimated market expansion effects to obtain a proper estimate of the competition effects, it is crucial that we do not overestimate market expansion. Fortunately, population size S serves as a natural exclusion restriction to identify the causal effect of N on r. It does not directly affect per capita revenues, yet it is correlated with N, since firms are more likely to enter and cover their fixed costs in large markets. In different contexts, Berry and Waldfogel (1999) and Ferrari, Verboven and Degryse (2010) have used similar identification strategies.

To estimate the model by maximum likelihood, suppose that ξ and η are normally distributed, so that $\omega \equiv \eta - \xi$ is also normally distributed. We then obtain the following likelihood contributions. For markets with $N = 0$ we have

$$P(0) = 1 - \Phi\left(\frac{X\lambda + \ln S + \theta_1}{\sigma_\omega}\right),$$

and for markets with $N > 0$ we have

$$\begin{aligned} f(\ln r)P(N|\ln r) \quad = \quad & \frac{1}{\sigma_\xi}\phi\left(\frac{\xi}{\sigma_\xi}\right) \times \\ & \left(\Phi\left(\frac{X\lambda + \ln S + \theta_N - \left(\sigma_{\omega\xi}/\sigma_\xi^2\right)\xi}{\sqrt{\sigma_\omega^2 - \sigma_{\omega\xi}^2/\sigma_\xi^2}}\right) - \Phi\left(\frac{X\lambda + \ln S + \theta_{N+1} - \left(\sigma_{\omega\xi}/\sigma_\xi^2\right)\xi}{\sqrt{\sigma_\omega^2 - \sigma_{\omega\xi}^2/\sigma_\xi^2}}\right)\right), \end{aligned} \tag{18}$$

where $\xi = \ln r - X\beta - \alpha_N$.

Constructing ETRs and percentage markup drops When the entry model is augmented with revenue data, we can still compute the ETR as before. It is given by

$$ETR(N) = \exp\left(\theta_{N-1} - \theta_N\right)\frac{N-1}{N}.$$

Furthermore, it is now also possible to directly compute the percentage markup drop following entry. Using (17), we can write this percentage markup drop as

$$\frac{\mu(N-1)}{\mu(N)} = \exp\left(\delta_{N-1} - \delta_N\right).$$

To express this in terms of the estimated parameters for the fixed effects revenue specification, we can substitute the definition $\theta_N \equiv \alpha_N + \delta_N$ to obtain:

$$\begin{aligned} \frac{\mu(N-1)}{\mu(N)} \quad &= \quad \exp\left(\theta_{N-1} - \theta_N\right)\exp\left(-\left(\alpha_{N-1} - \alpha_N\right)\right) \\ &= \quad ETR(N)\frac{N}{N-1}\exp\left(-\left(\alpha_{N-1} - \alpha_N\right)\right), \end{aligned} \tag{19}$$

where the second equality follows from the definition of the ETR. Similarly, for the constant elasticity revenue equation, we can substitute the definition $\theta_N \equiv \alpha \ln N + \delta_N$ to obtain

$$\begin{aligned} \frac{\mu(N-1)}{\mu(N)} &= \exp\left(\theta_{N-1} - \theta_N\right) \left(\frac{N-1}{N}\right)^{-\alpha} \\ &= ETR(N) \left(\frac{N}{N-1}\right)^{1+\alpha}. \end{aligned} \tag{20}$$

Consistent with the discussion in Section 2, this shows for both specifications how the ETRs should be adjusted by the estimated revenue parameters to obtain an unbiased estimate for the markup drop after entry. The simple ETRs can only be used as an unbiased measure in the special case where

$$\exp\left(-\left(\alpha_{N-1} - \alpha_N\right)\right) = \frac{N-1}{N},$$

in the flexible specification, and $\alpha = -1$ in the restricted specification. Intuitively, in both cases this requires that entry only leads to business stealing and not to any market expansion.

4 Empirical analysis

We organize the discussion of the empirical analysis as follows. We first present the dataset for the various local service sectors. Next, we discuss the results from estimating the entry model and the revenue model separately. This leads to the construction of traditional Bresnahan and Reiss entry threshold ratios. They do not yet take into account the existence of market expansion from entry, and can be used as a benchmark for our subsequent results. Finally, we present the results for the simultaneous model of entry and demand, leading to estimates of competition effects or "adjusted entry threshold ratios" that take into account market expansion effects.

4.1 Dataset

We analyze seven different local service sectors: architects, bakeries, butchers, florists, plumbers, real estate agents and restaurants. For each sector, we have constructed a cross-sectional data set of more than 800 local markets (towns) in Belgium in 2007. The main variables are firm revenues per capita r, the number of firms N, population size S and other market demographics X.[7]

[7] Firm revenues and the number of firms come from V.A.T. and Business register data from the sectoral database, set up by the Federal Public Service Economy (Sector and Market Monitoring Department). Population size and other market demographics are census data from the FPS Economy (Statistics Belgium).

Selection of sectors Based on our research proposal, the Belgian Federal Ministry of Economic Affairs made available a list of local service sectors at the 4-digit or 5-digit NACE code for empirical analysis. From this list we first eliminated sectors where the relevant market is clearly not local, such as TV-production houses. Furthermore, to avoid possible complications stemming from multi-market competition, we restricted attention to sectors where the average number of establishments per firm is less than 3. Sectors with many chains, such as travel agencies and clothes stores, were therefore also eliminated from the analysis. This resulted in a list of seven local service sectors: architects, bakeries, butchers, florists, plumbers, real estate agents and restaurants. For all these sectors the median number of establishments per company is 1, the 75-percentile is no larger than 2 and the 90-percentile is no larger than 5.

Geographic market definition For each sector, we define the geographic market at the level of the ZIP-code. This roughly corresponds to the definition of a town in Belgium, and it is more narrow than the administrative municipality, which on average consists of about 5 towns. The market definition appears reasonable for the considered sectors, as they relate to frequently purchased goods or to services where local information is important. The extent of the geographic market may of course vary somewhat across sectors. Nevertheless, for simplicity and consistency we decided to use the same market definition for all sectors. To avoid problems with overlapping markets, we only retain the non-urban areas, i.e. towns with a population density below 800 inhabitants per km^2 and a market size lower than 15,000 inhabitants.

Construction of the variables and summary statistics The number of firms N is the number of companies in the market, as constructed from the business registry database. Revenues per firm and per capita r are computed at the company level from the V.A.T. sectoral database. Ideally, we would want to use data at the establishment level but this information is incomplete. As discussed above, we therefore focus on sectors with a low number of establishments per firm (no chains). Furthermore, we restrict attention to companies with at most two establishments in the country.[8]

The data on the number of firms N and revenues r are specific to each of the seven different sectors. In addition to these endogenous variables, we also observe the common variables population size S and a vector of other market demographics X. This vector consists of the market surface, personal income/capita, the demographic composition of the population (%

[8] The results of our analysis are robust when we use alternative selection criteria, e.g. retain companies with at most five establishments.

women, % foreigners, % unemployed and % in various age categories), and a regional dummy variable for Flanders. The vector X enters both the revenue and entry equation. In contrast, population size S only enters the entry equation and therefore serves as an exclusion restriction for the revenue equation to identify the causal effect of N on r.

Table 1 gives a complete list of the variables and their definitions, and presents basic summary statistics for the common variables S and X, as observed for the cross-section of 835 non-urban markets. Table 2 provides more detailed summary statistics for the sector-specific variables, revenues per firm and per capita r and the number of firms N. The top panel shows the number of markets with 0, 1, 2, 3, 4, 5 or more firms. Most sectors have broad market coverage with a common presence of at least one firm per market. This is most notable for restaurants, since there are only 93 markets without a restaurant. The middle and bottom panels of Table 2 show the means and standard deviations for the number of firms N and revenues r across markets.

4.2 Preliminary evidence

We now discuss the results from estimating the entry model and the revenue model separately. This leads to traditional Bresnahan and Reiss entry threshold ratios. It also provides a first indication on the extent of market expansion (as opposed to business stealing) following entry, yet without accounting for endogeneity of N for now.

Entry model Table 3 shows the empirical results per sector from estimating the ordered probit entry model. Consistent with other work, population size $\ln S$ is the most important determinant of firm entry, with a positive and highly significant parameter for all sectors.[9] Several variables of the age structure also tend to have a positive and significant effect across sectors, in particular the %young and %old, relative to the reference group of young adults with age between 25–40 years. The effect of several other variables differs across sectors, both in sign and magnitudes. For example, markets with a high income per capita tend to have more architects, florists and real estate agents, but fewer bakeries. Generally speaking, it is not straightforward to interpret these parameters, as the variables may capture several effects (variable profits, fixed costs) and may be collinear with other variables (e.g. income and unemployment). While the control variables are not of direct interest, it is still important to control for them to allow for different sources of variation across markets.

The ordered probit model also includes the entry effects or "cut-points" θ_N. We transform

[9] Based on (11), the parameter of $\ln S$ can be interpreted as $1/\sigma$, and the parameters of the other demographics as λ/σ.

these parameters to construct the entry thresholds (for a representative market with average characteristics) and the per firm entry threshold ratios (which are independent of the other characteristics). This is based on the expressions (12) and (13) derived earlier.

Table 4 shows the computed entry thresholds and entry threshold ratios. To illustrate, first consider butchers (third column). The entry threshold, i.e. the minimum population size to support one butcher in a town, is 1,166. It increases to 2,736 to support a second butcher and to 4,905 to support a third butcher. The pattern is slightly disproportional, i.e. the minimum population size to support a given number of firms increases disproportionately with the number of firms. This is reflected in the ETRs. For example, $ETR(2) = 1.17$, which means that the minimum population size per firm should increase by an extra 17% to support a second firm. Under the homogeneous goods assumption of the Bresnahan and Reiss model, this can be interpreted as an indication that entry intensifies competition between butchers.

Now consider all sectors. Table 4 shows that the ETRs for the third, fourth or fifth entrant are significantly greater than 1 in about half of the cases, and insignificantly different from 1 in the remaining half. In the traditional Bresnahan and Reiss' framework, this would indicate mixed evidence on the competitive effects of entry from the third entrant onwards. Table 4 also shows that the ETR for the second entrant is only significantly greater than 1 for one sector, butchers; it does not differ significantly from 1 for four sectors; and it is even significantly less than 1 for the remaining two sectors, architects and real estate agents. The latter finding contradicts the competition interpretation of ETRs, as it would suggest that competition becomes weaker when a second firm enters the market. As we will show below, an alternative interpretation is the presence of significant market expansion when a second firm enters the market.

Revenue model Table 5 shows the empirical results per sector from simple OLS regressions of the restricted revenue specification (15), i.e. regressions of $\ln r$ on $\ln N$ and X. Since the model is estimated with OLS, we do not yet account for the endogeneity of N so we should be cautious at this point in drawing causal inferences on market expansion versus business stealing from entry. First, consider the control variables X. In contrast with the entry equation, the parameters are significant for most variables and usually have the same sign across the various sectors. Per capita revenues tend to be larger in markets with a low surface area, a low personal income, a low fraction of unemployed, and a high fraction of kids/young or old (relative to the base young adult group).

Now consider the parameter on $\ln N$. The parameter is negative and significant for five out of seven sectors, and insignificantly different from zero for the remaining two sectors (florists and real estate agents). For the five sectors where the parameter is negative, it is relatively small, varying between –0.15 and –0.39. Overall, this preliminary evidence would suggest that additional entry implies some business stealing but more important market expansion. This would in turn indicate that the ETRs are not a good measure of competition, as this is only the case when entry only leads to business stealing (coefficient for $\ln N$ of –1). However, as already mentioned, we have not yet accounted for the endogeneity of N. Firms tend to locate in markets where they expect demand to be high, leading to a spurious correlation between the number of firms and total market demand and an overestimate of the extent of market expansion. Our full model accounts for this, by estimating the revenue model simultaneously with the entry model, using market size as an exclusion restriction to identify the market expansion effect.

4.3 Results from the full model

We now discuss the main empirical results, from estimating the entry and revenue model simultaneously. We first look at the case of butchers in detail, to give a comparison of the different specifications and methods. We then give a broader overview of all sectors, focusing on the estimated competition effects or adjusted ETRs, which take into account the market expansion effects from entry.

Comparison of different specifications and methods: butchers As discussed in section 3, we consider two specifications for the revenue equation. In the constant elasticity specification (15), the number of entrants appears logarithmically, so $\alpha_N = \alpha \ln(N)$. In the fixed effects specification (16), we estimate the effect of entry α_N on revenues for each market configuration. For both specifications, we compare the results from simultaneous estimation of the demand and entry model with those from estimating the models separately. We focus the comparison on the revenue equation, since the results for the entry equation are very similar across specifications and methods (and given in Table 3 for the single equation estimation).

Table 6 shows the results. The estimated effects of the control variables X are very similar across different specifications, so we do not discuss them further. Our main interest is in the effects of entry on revenues. First consider the constant elasticity specification. When the revenue equation is estimated separately using OLS, we estimate $\alpha = -0.24$ (as already reported in Table 5). In sharp contrast, when the revenue equation is estimated simultaneously with the entry equation, we estimate $\alpha = -0.72$. Hence, accounting for the endogeneity of N implies a considerably higher estimate of business stealing. The market expansion elasticity, $1 + \alpha$, correspondingly drops from 0.76 to 0.28. Intuitively, OLS gives a spurious finding of market expansion, since it does not take into account that entrants tend to locate in markets where the unobserved demand error is high.[10] Nevertheless, the simultaneous model still implies there is some market expansion: an increase in N by 10% tends to raise market revenues by 2.8%. The bottom part of Table 6 shows how α translates into percentage revenue effects $R(N)/R(N+1)$. We see a declining pattern, where the effect on total revenue per capita is 21% for the second entrant, 12% for the third entrant, 8% for the fourth entrant and 6% for the fifth entrant. This smooth pattern is evidently driven by the restricted functional form of the logarithmic specification.

Now consider the unrestricted fixed effects specification. We do not report the different α_N, but immediately discuss the implied percentage revenue effects $R(N)/R(N+1)$. As before, we find large market expansion effects from single equation estimation (e.g. 85% market expansion for the second entrant) and much lower effects when we account for the endogeneity of N (26% for the second entrant). Furthermore, the flexible specification no longer gives a smooth pattern for the entry effects. Only the second butcher leads to significant market expansion. For additional entrants, the extent of market expansion becomes insignificant.

In sum, this discussion shows that both the specification and the method are important to correctly estimate the extent of market expansion. First, it is necessary to account for the endogeneity of entry since otherwise the extent of market expansion will be overestimated. Second, it may be important to consider the possibility of a flexible specification for the entry effects, though this comes at the cost of reduced precision. These conclusions do not just hold for butchers but also for the other sectors we have studied. They will therefore be highly relevant when estimating the competition effects based on the adjusted ETRs.

Competition effects from entry: all sectors Table 7 shows the competition effects from additional entry, as estimated from the simultaneous entry and revenue model. As is clear from (19) and (20), the competition effects can be interpreted as adjusted ETRs: they adjust the traditional ETRs for the extent of market expansion induced by entry. Only if market expansion is small, the competition effects will be close to the traditional ETR's.

The top panel of Table 7 shows the results for the constant elasticity revenue specification. The first row shows the estimated business stealing effects α from the revenue equation. For six out of seven sectors, the estimates are much closer to -1 than in the earlier OLS estimates

[10] More formally, the simultaneous model differs from the single equation model because it accounts for the correlation between the demand and profit error. Table 5 shows that $\sigma_{\omega\xi} = -0.43$, which is negative as expected because the structural error in the entry equation contains the structural error in the demand equation.

of Table 5. This means that the necessary adjustments of the ETRs are much smaller as earlier suggested. Nevertheless, the market expansion elasticity $1+\alpha$ is still important, varying from 0.08 for bakeries to 0.72 for florists.[11]

Based on (20), we can use the α's and the ETRs (very similar to those in Table 4) to compute the markup effects or "adjusted ETRs". For most sectors and market configurations we find significant competition effects from entry. The adjusted ETRs are typically significantly greater than 1, also for entry by the second firm, and they are never significantly below 1. For example, entry by a second restaurant reduces markups by 17% ($\mu(1)/\mu(2) = 1.17$). This contrasts with our earlier estimated simple ETRs, which were often significantly less than 1 for the second entrant (e.g. $ETR(2) = 0.87$ for restaurants). The reason is, of course, that we now adjust for the extent of market expansion. Bakeries are the only sector without significant competition effects from entry in the constant elasticity specification. We already found the traditional ETRs to be close to 1 in this sector. Moreover, it turns out that entry by bakeries largely entails business stealing ($\alpha = -0.92$), so that the adjusted ETRs remain close to and not significantly different from 1.

The bottom panel of Table 7 shows whether these conclusions are confirmed using the more flexible fixed effects revenue specification. The estimated competition effects of the second entrant are broadly similar. In five out of seven sectors, the second entrant has a significant effect on competition. The two exceptions are bakeries (as before) and real estate agents where $\mu(1)/\mu(2)$ does not differ significantly from 1. However, the conclusions regarding competition from the third, fourth or fifth entrant are different from the restricted specification. With the exception of restaurants, we no longer estimate significant competition effects from the third entrant onwards. Note, however, that the standard errors of the estimated $\mu(N-1)/\mu(N)$ have become larger (because of the increased flexibility), so that the competition tests have less power.

Combining the results from the restricted constant elasticity specification (with more precise estimates) and the more flexible fixed effects specification (with larger standard errors), we conclude that in most sectors the second entrant appears to reduce markups by at least 30%, whereas further entrants may not necessarily promote competition further. Bakeries and real estate agents are exceptions to this conclusion. For real estate agents, the fixed effects specification does not estimate significant competition effects from the second entrant, though the standard errors are rather large here.[12] For bakeries, the lack of

[11] Only for real estate agents α is not significant. This suggests considerable market expansion, perhaps capturing that market definition is broader than the town level for this sector.

[12] A lack of competition effects from entry in the real estate sector is consistent with the common practice of more or less uniform percentage commissions. This has also been documented elsewhere, for example

competition effects appears more strongly: both the constant elasticity and the fixed effects specification indicate that the second entrant does not promote competition. Incidentally, this is consistent with a recent decision by the Belgian Council of Competition. In January 2008, the Council convicted the Association of Bakeries for continuing its price fixing policies after prices for bread had been liberalized in 2006.

5 Conclusions

We have proposed a methodology for estimating the competition effects from entry in differentiated products markets, and illustrated how to implement it using datasets for seven different local service sectors. We started from Bresnahan and Reiss' ETRs, and provided conditions under which they can be used as a test for the presence and a measure for the magnitude of competition effects from entry. We subsequently showed how to augment the traditional entry model with a revenue equation. This revenue equation serves to adjust the traditional ETRs by the extent of market expansion due to entry, leading to an unbiased estimate of the competition effects from entry.

Our empirical results show that traditional ETRs are close to one, suggesting limited competition effects, and in some cases even significantly below 1, suggesting entry would reduce competition. Furthermore, we find that entry leads to significant market expansion, which implies that the traditional ETRs underestimate the effects of entry on competition. Accounting for the estimated market expansion, we no longer find adjusted ETRs that are significantly below 1. In most sectors, the second entrant reduces markups by at least 30%, whereas the third or higher entrants have smaller or insignificant effects. In at least one sector, bakeries, we have found that even the second entrant does not create competition, which is consistent with a recent decision by the competition authority.

Our empirical analysis stressed the importance of several specific issues that should be taken into account. First, it is important to account for the endogeneity of the number of entrants in estimating market expansion effects from entry. Failure to do so would result in an overestimate of market expansion effects, and hence an overestimate of the competition effects (adjusted ETRs), as opposed to an underestimate from the traditional ETRs. In our setting, population size arises as a natural instrument, and we found the bias from ignoring the endogeneity issue can be substantial.

Second, it is potentially important to consider a flexible revenue specification to estimate the market expansion effects. Our restricted constant elasticity specification (with $\ln N$) imposes market expansion effects to be declining in N, whereas our more flexible fixed effects

Hsieh and Moretti (2003), who draw implications for the efficiency of entry.

specification allows the effects to vary per consecutive entrant. The flexible specification suggested that the main market expansion effects (and hence required adjustment to the ETRs) come from the second entrant, and less so from the additional entrants. However, this specification also entails less precise parameter estimates. Future research would be desirable to shed further light on this. For example, one may collect more data, or use alternative specifications with more structure from a specific model of product differentiation.

Due to the relative simplicity of our methodology, it was possible to consider quite a number of different local service sectors. Nevertheless, more work on different sectors and different countries would be useful to further evaluate the benefits and limitations of our approach. We hope the increased availability of revenue data at the detailed company level will stimulate such research.

6 References

Ackerberg, D., L. Benkard, S. Berry, A. Pakes (2007), "Econometric Tools for Analyzing Market Outcomes," in Handbook of Econometrics, J.J. Heckman (ed.).

Asplund, M. and R. Sandin (1999), "The Number of Firms and Production Capacity in Relation to Market Size," *The Journal of Industrial Economics*, 47 (1), pp. 69-85.

Berry, S. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), pp. 889-917.

Berry, S. and P. Reiss (2007), "Empirical Models of Entry and Market Structure," Chapter 3 in M. Armstrong and R. Porter, Handbook of Industrial Organization Vol. 3, Elsevier.

Berry, S., Waldfogel, J. (1999), "Free entry and social inefficiency in radio broadcasting," *RAND Journal of Economics*, 30(3), pp. 397-420.

Bresnahan, T. (1989), "Empirical Studies of Industries with Market Power," Handbook of Industrial Organization Vol. 2,

Bresnahan, T. and P. Reiss (1990), "Entry in monopoly markets," *Review of Economic Studies*, 57(4), pp. 531-553.

Bresnahan, T. and P. Reiss (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), pp. 977-1009.

Campbell, J. and H. Hopenhayn (2005), "Market Size Matters," *Journal of Industrial Economics*, 53(1), 1-25.

De Loecker, J. and F. Warzynski (2010), "Markups and Firm-Level Export Status," working paper, Princeton University.

Ferrari, S., F. Verboven and H. Degryse (2010), "Adoption and Usage of New Technologies: Evidence from ATMs", *American Economic Review*, 100(3), 1046-1079.

Genesove, D. (2000), "Why are there so few (and fewer and fewer) two-newspaper towns," working paper Hebrew University of Jerusalem.

Hall, R. (1988), "The Relationship between Price and Marginal Cost in U.S. Industry, Journal of Political Economy, 96 (5), 921-947.

Hsieh, C. and E. Moretti (2003), "Can Free Entry be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry", *Journal of Political Economy*, 111(5), 1076-1122.

Konings, J., P. Van Cayseele and F. Warzynski (2005), "The Effect of Privatization and Competitive Pressure on Firms' Price-Cost Margins: Micro Evidence from Emerging Economies, *Review of Economics and Statistics*, 87(1), 124-134.

Mankiw, G. and M. Whinston (1986), "Free Entry and Social Inefficiency," *RAND Journal of Economics*, 17(1), 48-58.

Manuszak, M. (2002), "Endogenous market structure and competition in the 19th century American brewing industry", *International Journal of Industrial Organization*, 20(5), pp. 673-692.

Mazzeo, M. (2002), "Product Choice and Oligopoly Market Structure", *RAND Journal of Economics*, 33(2), pp. 221-242.

Reiss, P. and F. Wolak (2007), "Structural Econometric Modeling: Rationales and Examples from Industrial Organization," Handbook of Econometrics Vol. 6A, J.J. Heckman (ed.), pp. 4277-4415.

Seim, K. (2006), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37(3), 619-640.

Schaumans, C. and F. Verboven (2008), "Entry and Regulation: Evidence from Health Care Professions," RAND Journal of Economics, 39, 949-972.

Syverson, C. (2004), "Market Structure and Productivity: a Concrete Example," *Journal of Political Economy*, 112(6), 1181-1222.

Table 1: Definition of variables

Name	Definition	Mean	St. Dev.
N	Number of firms with at least one establishment	See Table 2	
r	Revenues per firm and per capita (in €)	See Table 2	
S	Population size or number of inhabitants (in 1,000)	4.53	3.89
Surface	logarithm of surface area (in km^2)	2.71	2.76
GDP	GDP per capita (in 1,000 Euro)	11.15	2.03
%women	Percentage of women	.506	.013
%foreigners	Percentage of foreigners	.043	.057
%unemployed	Percentage unemployed	.057	.028
%kid	Percentage under age of 10 years	.121	.018
%young	Percentage between age of 10 and 25 years	.187	.019
%adult	Percentage between age of 40 and 65 years	.323	.027
%old	Percentage over age of 65	.163	.028
Flanders	Dummy variable equal to 1 for market in Flanders	.398	.490

Notes: The number of observations (markets) is 835. The number of firms N and revenues per firm r are constructed from V.A.T. and Business register data from the sectoral database, set up by the Federal Public Service Economy (Sector and market Monitoring Department). The demographics are census data from the FPS Economy (Statistics Belgium), except for %unemployed which comes from Ecodata.

Table 2: Summary statistics for number of firms and firm revenues

Sector	Archit.	Baker.	Butch.	Florists	Plumb.	Real Est.	Restaur.
NACE code	7111	1071	4722	47761	4322	6831	5610
	Number of markets with						
$N = 0$	144	242	236	260	139	278	93
$N = 1$	83	148	169	147	112	106	74
$N = 2$	76	126	122	130	94	95	65
$N = 3$	79	94	97	85	68	57	57
$N = 4$	68	63	71	62	68	56	37
$N = 5$	39	41	39	44	43	26	37
$N > 5$	337	111	93	94	303	168	472
	Number of firms (sample of all markets)						
mean	6.2	2.5	2.4	2.3	5.1	3.4	11.1
st.dev	7.0	2.8	2.7	2.6	5.1	5.6	12.3
	Revenues per firm and per capita (sample of markets with $N > 0$)						
mean	27.79	65.56	82.09	51.96	108.26	31.68	64.18
st.dev	51.98	76.70	117.8	106.14	231.3	63.32	132.5

Notes: The number of observations (markets) is 835.

Table 3: Ordered probit entry model

	Archit.	Baker.	Butch.	Florists	Plumb.	Real Est.	Restaur.
	Ordered probit entry model (sample of all markets)						
$\ln S$	1.40*	1.62*	1.21*	1.29*	1.34*	1.35*	1.48*
Surface	0.12	-0.04	0.10	0.06	0.15*	-0.09	0.24*
GDP	2.63*	-0.73*	-0.48	0.81*	0.59	2.11*	-0.28
%women	9.27*	-8.58	-0.16	-2.16	-3.57	-0.40	3.63
%foreigners	-0.91	-2.08*	-2.53*	0.18	-1.59*	0.40	-0.04
%unemployed	-4.18*	-2.85	-2.45	-2.36	-2.85	-6.34*	4.95*
%kid	7.41	0.02	-6.69	-7.07	2.44	12.99*	1.29
%young	11.49*	6.99*	7.99*	0.01	1.55	13.20*	9.05*
%adult	2.69	-3.13	-3.75	-7.93*	-0.27	7.55*	9.50*
%old	4.79	10.57*	7.70*	-1.87	-0.10	13.06*	7.08*
Flanders	-0.49*	0.01	0.28*	0.04	-0.05	-0.28	0.59*
θ_N	yes	yes	yes	yes	yes	yes	yes
R^2	0.25	0.29	0.26	.27	0.24	0.25	0.25

Notes: The parameter estimates are based on maximum likelihood estimation of the ordered probit model (11), where the parameters are all multiplied by the standard deviation σ. Hence, the parameter of $\ln S$ can be interpreted as $1/\sigma$, and the parameters of the other demographics as λ/σ. A "*" indicates that the parameter differs significantly from 0 at the 5% level.

Table 4: Entry thresholds and entry threshold ratios

	Archit.	Baker.	Butch.	Florists	Plumb.	Real Est.	Restaur.
	Entry thresholds						
$ET(1)$	692	1387	1166	1405	650	1699	445
$ET(2)$	1137	2610	2736	2873	1251	2818	773
$ET(3)$	1706	4326	4905	5198	2041	4458	1132
$ET(4)$	2527	6446	8027	7864	2845	5896	1572
$ET(5)$	3542	8656	12360	11171	3979	7852	1924
	Entry threshold ratios						
$ETR(2)$	0.82*	0.94	1.17*	1.02	0.96	0.83*	0.87
$ETR(3)$	1.00	1.11*	1.20*	1.21*	1.09	1.06	0.98
$ETR(4)$	1.11*	1.12*	1.23*	1.14*	1.05	0.99	1.04
$ETR(5)$	1.12*	1.07	1.23*	1.14*	1.12*	1.07	0.98

Notes: The entry thresholds (ET) are based on the cut-points θ_N and the other parameter estimates of Table 3, using expression (12) evaluated at the sample means of the variables. The entry threshold ratios (ETR) are based on the cut-points θ_N, using expression (13). All ETs are significant with standard errors varying around 150. For the ETRs, a "*" indicates that the ETR differs significantly from 1.

Table 5: Preliminary regressions for the revenue equation

	Archit.	Baker.	Butch.	Florists	Plumb.	Real Est.	Restaur.
	OLS revenue model (sample of markets with $N > 0$)						
Constant	3.82	11.89*	18.05*	19.57*	16.34*	5.20	11.20*
ln N	-0.15*	-0.39*	-0.24*	-0.02	-0.15*	0.10	-0.25*
Surface	-0.57*	-0.36	-0.53*	-0.43*	-0.50*	-0.52*	-0.45*
GDP	-0.24	-0.69*	-0.86*	-0.75	-1.23*	0.05	-0.81*
%women	-3.10	-9.97*	-15.23*	-15.6*	-11.09*	-11.16	-10.28*
%foreigners	-1.81*	-0.76	-1.50*	-1.89	-1.09	-1.20	-1.48*
%unemployed	-8.74*	-5.95*	-9.66*	-7.70*	-5.61*	-4.19	-5.09*
%kid	13.71*	6.48	7.10	5.53	11.48*	17.80*	10.24*
%young	7.78*	11.63*	6.34*	2.78	13.62*	1.33	11.61*
%adult	1.68	2.95	1.23	-4.03	3.91	2.75	6.81*
%old	10.72*	8.95*	11.42*	3.02	9.76*	6.90	10.45*
Flanders	-0.51*	-0.28*	-0.53*	-0.16	-0.12	-0.19	-0.24*
R^2	.33	.33	.37	.13	.27	.09	.40

Notes: The parameter estimates are based on OLS estimation of the restricted revenue specification (15). A "*" indicates that the parameter differs significantly from 0 at the 5% level.

Table 6: Detailed estimation results for the revenue equation: illustration with butchers

	Constant elasticity model				Fixed effects model			
	Single equation		Simultaneous		Single equation		Simultaneous	
Constant	18.05*	(2.94)	9.76	(3.40)	–		–	
$\ln N$ (α)	-0.24*	(0.06)	-0.72*	(0.09)	(fixed effects)		(fixed effects)	
Surface	-0.53*	(0.05)	-0.18	(0.07)	-0.54*	(0.05)	-0.02	(0.08)
GDP	-0.86*	(0.28)	-0.30	(0.36)	-0.89*	(0.28)	-0.12	(0.40)
%women	-15.23*	(3.83)	-6.78	(3.83)	-15.35*	(3.85)	-3.15	(4.24)
%foreigners	-1.50*	(0.71)	-1.17	(0.88)	-1.56*	(0.72)	-1.15	(0.97)
%unemployed	-9.66*	(1.87)	-7.81*	(2.19)	-9.63*	(1.88)	-7.12*	(2.42)
%kid	7.10	(3.68)	-0.16	(4.11)	7.41*	(3.70)	-3.09	(4.49)
%young	6.34*	(2.67)	5.47	(2.83)	6.51*	(2.69)	5.01	(3.12)
%adult	1.23	(2.47)	-1.72	(3.14)	1.30	(2.48)	-2.99	(3.40)
%old	11.42*	(2.22)	9.53*	(2.41)	11.38*	(2.23)	8.48*	(2.61)
Flanders	-0.53*	(0.11)	-0.14	(0.14)	-0.53*	(0.12)	0.06	(0.16)
$\sigma_{\omega\xi}$	0	(–)	-0.43*	(0.06)	0	(–)	-0.60	(0.08)
$R(2)/R(1)$	1.78*	(0.10)	1.21*	(0.07)	1.85*	(0.20)	1.26*	(0.13)
$R(3)/R(2)$	1.40*	(0.05)	1.12*	(0.04)	1.38*	(0.18)	1.05	(0.13)
$R(4)/R(3)$	1.27*	(0.03)	1.08*	(0.03)	1.29	(0.19)	1.00	(0.14)
$R(5)/R(4)$	1.20*	(0.02)	1.06*	(0.02)	1.04	(0.24)	0.82	(0.17)

Notes: Both the single equation and the simultaneous equation models are estimated by maximum likelihood of the full model (18). The single equation models are the special case in which we set $\sigma^2_{\omega\xi}= 0$, reducing to the earlier ordered probit entry equation and OLS revenue equation. In the restricted constant elasticity model, N enters the revenue equation through $\ln N$, in the flexible fixed effects model it enters through a set of fixed effects α_N. Parameter estimates and standard errors (in parentheses) are only shown for the revenue equation. For the entry equation, they are very similar to the single equation ordered probit results of Table 3. A "*" indicates that the parameter differs significantly from 0 at the 5% level.

Table 7: Markup effects or adjusted entry threshold ratios

	Archit.	Bakeries	Butchers	Florists	Plumbers	Real Est.	Restaur.
	constant elasticity model						
α	-0.48*	-0.92*	-0.72*	-0.28*	-0.53*	0.07	-0.53*
	(0.05)	(0.09)	(0.09)	(0.11)	(0.08)	(0.09)	(0.05)
$\mu(1)/\mu(2)$	1.20*	1.02	1.42*	1.57*	1.35*	1.70*	1.17*
	(0.07)	(0.06)	(0.11)	(0.16)	(0.10)	(0.14)	(0.07)
$\mu(2)/\mu(3)$	1.24*	1.17*	1.33*	1.58*	1.32*	1.58*	1.22*
	(0.06)	(0.05)	(0.07)	(0.10)	(0.07)	(0.09)	(0.06)
$\mu(3)/\mu(4)$	1.26*	1.14*	1.28*	1.37*	1.19*	1.33*	1.21*
	(0.05)	(0.04)	(0.06)	(0.07)	(0.05)	(0.06)	(0.05)
$\mu(4)/\mu(5)$	1.22*	1.07	1.24*	1.31*	1.23*	1.34*	1.08*
	(0.04)	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)	(0.04)
	fixed effects model						
$\mu(1)/\mu(2)$	2.01*	1.19	1.53*	1.73*	1.82*	1.31	1.35*
	(0.19)	(0.11)	(0.16)	(0.26)	(0.25)	(0.25)	(0.13)
$\mu(2)/\mu(3)$	0.99	1.21	1.25	1.40	1.25	0.98	1.40*
	(0.11)	(0.11)	(0.15)	(0.22)	(0.18)	(0.22)	(0.16)
$\mu(3)/\mu(4)$	1.14	1.13	1.21	1.24	1.08	1.55	1.15
	(0.14)	(0.12)	(0.17)	(0.23)	(0.17)	(0.38)	(0.19)
$\mu(4)/\mu(5)$	1.09	0.98	1.01	1.02	1.63*	1.75	0.92
	(0.17)	(0.13)	(0.21)	(0.22)	(0.29)	(0.55)	(0.17)

Notes: The markup effects $\mu(N-1)/\mu(N)$ are computed from (20) for the restricted constant elasticity revenue equation, and from (19) for the more flexible fixed effects revenue specification. For the constant elasticity specification, Table 7 also shows the business stealing effect α, used to adjust the ETR. A "*" indicates that the markup effect differs significantly from 1.

Persistence of Profits in Belgium

Preliminary Version, Please Do Not Distribute

Cherry Cheung

HUBrussels

Stijn Vanormelingen

IESE Business School

HUBrussels

2011

Abstract

Determining the intensity of competition is a key interest in the field of industrial organization. Static measures such as price-cost margins or concentration ratios may inadequately reflect the intensity of competition in a number of cases. A solution is to look at the competitive dynamics and examine the degree of profits persistency. The general idea is that in an efficient market economy, supra-normal profits should quickly disappear as they attract new entrants or imitators. The increase in competitors erodes profits earned by the initially successful incumbent. However, when firms operate in a less competitive environment, profits may be persistent and do not fall back to their competitive level. In order to analyze the persistence of profits in Belgium, we use data on around 200,000 firms between 1999 and 2008, retrieved from their income statements. We apply time series analysis to the data and the results are used to rank the different sectors according to their measured persistency of profits. Several robustness checks are performed and the profits persistency is related to several factors that have an influence on competition intensity.

1 Introduction

Determining the strength of competition in a market is of direct interest to both academics as well as policy makers. They are often interested in evaluating the impact of various policy decisions or variations in the economic environment on competition. Several papers relate a change in the economic environment with a change in competition across different sectors/industries using production data. For example, in the international trade literature, many studies have been devoted to testing of the imports as market-disciplining device (Levinsohn 1993, Harrison 1994). Other studies look at the relation between competition and innovation (Aghion et al., 2005), the link between competition and productivity (Nickel 1996, Syverson 2004), etc. Often, the price cost margin at the industry (market) level is used to measure competition, either directly computed from accounting data or estimated using the Hall (1988) methodology or a variant thereof. Another strand of literature investigates one particular industry in detail and structurally estimate demand and supply in order to infer price cost margins and these price cost margins can be related to the policy change of interest. Notable examples include Porter (1983) and Genesove and Mullin (1998) for homogenous goods markets and Berry et al. (1995) and Nevo (2001) for markets of differentiated products. Other popular measures used in the literature to measure competition are concentration ratio's such as the Herfindahl-Hirschmann Index or Ck ratio's.

All these competition indicators generally focus on a snapshot of a sector taking the implicit assumption that the indicator reaches its long-run equilibrium value in every period. However, there is no guarantee for this to be the case. First, a high price-cost margin at some specific moment in time could just represent a temporary phenomena reflecting a disequilibrium state of the market. Second, these measures do not pick up underlying dynamics in the market. For example, in Schumpeter's creative destruction model, successful firms are able to realize substantial profits in a single period, but they lose their dominant position once a competitor takes over the market with a new innovation. Computing concentration ratio's or price-cost margins for these sectors will erroneously point to a lack of competition in these markets as they ignore the dynamics in the market. To correct for this problem, Mueller (1977, 1986) introduced the so-called persistence of profits concept which explicitly examines the dynamics of market processes applying time-series analysis and uses the results to draw inferences about the nature of competition in the market. The general idea is that firms with an abnormal level of profits in one period are not expected to maintain their high level of profitability in subsequent periods if they are operating in a competitive environment. This will lead to a low measured persistency of profits, for example due to the profits are competed away by imitation or entry of firms attracted by high profits. On the other hand, firms operating in a less competitive environment are more likely to maintain their high profits and profits are expected to be more persistent. This idea has been used in a number of papers and they showed deviations of profit rates from the norm to be substantially persistent. Mueller (1977, 1986) examines 472 firm with 24 years of return on assets data and finds there is persistence of supernormal profits for some firms. The idea has subsequently been used by the Geroski and Jacquemin (1988) for European firms among others. McGahan and Porter (1998) investigate the differential persistence industry, corporate and business segment shocks to profitability and find that industry shocks persist longer. More recently, Glen et al. (2001, 2003) have applied the framework to developing countries and concluded that the intensity of competition is higher compared to advanced countries. Yurtoglu (2004) analyzes the persistence of firm-level profitability on 172 largest manufacturing firms in Turkey from 1985 to 1998 and concludes that firms with the highest initial profit rate and long-run projected profit rate have the highest degree of persistence, which is consistent with the prediction that firms with the higher profit rate should have greater incentive to block entry.

In this study paper, we estimate for the first time the persistence of profits for Belgian firms active in all sectors of the economy. To this end we make use of a unique large panel dataset. Most other studies relied on large publicly listed companies to estimate the persistence of profits[1]. The richness of our dataset allows us to investigate different dimensions of the persistence of profits. First, we are able to make a distinction between large and small firms. Second, we can exploit variation in the persistence of profits across sectors, not only to rank them in terms of competition intensity, but also to explain the heterogeneity in terms of profit persistency using sector characteristics.

Our main results can be summarized as follows. First, we find profits to be persistent although persistency is lower compared to previous studies in other countries. Second, we find that small firms have a substantially lower persistence of profits compared to large firms. This finding can partly explain the difference in profit persistency compared to other studies. Third, the highest persistency is found in sectors such as Mining and Quarrying, Manufacture of Gas, Steam and Air Conditioning Supply which are known to have high entry barriers. Third, profit persistency is negatively correlated with entry and exit rates of firms while it is positively correlated with concentration although this is mainly due to differences between services and manufacturing sectors. The rest of the paper is organized as follows. Section 2 introduces the empirical model applied to measure profits persistency. The dataset is described in Section 3 and the results are presented in Section 4. The final section concludes.

---

[1] For example Glen et al. (2001, 2003) uses a data set consisting of 100 largest listed manufacturing corporations in seven developing countries. Yurtoglu (2004) uses the 172 largest firms listed continuously from 1985 to 1998. Geroski and Jacquemin (1988) use a sample of 134 large European firms, including 51 from the United Kingdom, 28 from West Germany, and 55 from France and Goddard and Wilson (1999) use a sample set of 335 large survival companies from 1972 to 1991.

2 Methodology

The literature on persistence of profits owes a great deal to the work by Mueller (1986) who modeled profitability of a firm as a first order autoregressive process:

$$\pi_{it} = \alpha_i + \lambda_i \pi_{it-1} + \varepsilon_{it} \qquad (1)$$

where π_{it} represents the standardized profitability rate of firm *i* in year *t*. The firm specific parameters to be estimated are α_i and λ_i. ε_{it} represent firm/year specific i.i.d. shocks to profitability. Short-run persistence of profits is picked up by the parameter λ_i and measures how fast profitability returns back to its long term equilibrium after a shock. The estimation equation used to measure persistence of profits is best regarded as a reduced form of a more sophisticated structural model. This model includes not only entry and exit of firms but also the threat of entry, which is obviously mostly impossible to observe. The advantage of the persistence of profits framework is that it does not require any unobservable variables to map competitive dynamics (Geroski 1990, Glen et al. 2003). The drawback is that the framework does not allow us to take a stand on the sources of profit persistency.

In general, one distinguishes three different possibilities for short-run persistency. First, when $\lambda_i = 0$ profitability follows a white noise process. Any abnormal profit earned in period t-1 is immediately eroded away. This can be due to either actual entry or by just the mere threat of entry and one states that firms are operating in a competitive environment. Second, when $0 < \lambda_i < 1$, current and future profitability are positively related and there exists some persistence of profits. The higher λ_i, the higher the persistence of profits and the lower the competitive forces. Ultimately, profitability converges to its long-run equilibrium value given, $\pi_{i,LR} = \frac{\alpha_i}{1-\lambda_i}$. Third, when $\lambda_i = 1$, abnormal profits earned in one period are not threatened at all by (possible) competitors. The profitability process has a unit root and profitability follows a random walk. Note that this is also not very theoretically appealing as this would mean that profitability would ultimately reach an arbitrary high or low value (Geroski and Jacquemin, 1988).

Values of λ larger than 1 would imply profitability rates of firms to blow up over time. Obviously this finding goes against common sense as well as a finding of λ smaller than -1. The same holds for values of λ between 0 and -1, which means profitability would be stationary, but implies profitability to oscillate around its long term average. However, while these values could be dismissed on theoretical grounds we do not impose any of these restrictions in our estimation procedure.

In the absence of (long-run) entry barriers, long-run profitability should be the same for all firms and there is no long-run persistence of profits as measured by $\alpha_i / (1-\lambda_i)$. When there exists long-run persistence of profits, long-run profitability will be positive for some firms and negative for others. A measure for competitive forces in a sector would be the variance of long-term profitability where a large variance points to underlying variables hampering competition. However, most of the literature has focused on the short-run persistency, probably because the easy interpretation of the parameter. We will follow this tradition and devote most of our attention our estimates of the autoregressive coefficient.

In general, equation (1) is estimated at the firm level instead of constructing a panel, assuming (some of) the parameters to be constant across firms and using standard panel data techniques. The only exception is Waring (1996) who estimates equation (1) for a large panel of US firms assuming the short run persistency to be the same for all firms in one sector. If the underlying parameters are indeed constant across firms, this approach is more efficient compared to estimating (1) for each firm separately. However, we would have to assume there is no firm specific long-term persistency in order to retrieve unbiased parameter estimates using OLS, i.e. we have to assume there are no firm fixed effects. Otherwise our estimates for the autoregression parameter will be upward biased as lagged profitability is obviously positively correlated with the firm fixed effect. Moving to a within estimator will not solve the problem as this will introduce a downward bias in the coefficient. We could correct for this by applying dynamic panel data methods (Arellano and Bond, 1991 and Blundell and Bond, 2000) and estimate equation (1) by GMM but we rather choose to follow the standard in the literature and estimate the equation for each individual firm separately.

Estimating a autoregressive model by ordinary least squares will result in consistent estimates for λ when T goes to infinity but will be downward biased in small samples. The bias is inversely proportional to the number of time periods and as we observe each firm only for a limited period of time, this small-sample bias could be important. Patterson (2000) suggests a procedure to correct the point estimates. However, most other persistence of profits studies did compute the small sample bias correction and to improve comparability, we also report the uncorrected estimates[2]. After estimating the equation at the firm level, we aggregate the short-run persistency parameter for different groups of firms. First, we compute average persistency for narrowly defined sectors. Second, we investigate heterogeneity in persistency across different firm sizes. The idea is that large firms are better

---

[2] Note that the bias is equal to $-(1+3\lambda)/T$ with T the number of periods in a first-order autoregressive model. If there is an equal amount of observations for all firms, the small sample bias will not alter the ranking of the firms in terms of competition intensity as correcting for the bias is a monotonic transformation of the parameter estimate, namely $\tilde{\lambda} = \hat{\lambda}T / (T-3) + 1/(T-3)$ with $\hat{\lambda}$ the estimated parameter and $\tilde{\lambda}$ the bias adjusted parameter.

able to protect their supranormal profits from competition compared to small firms[3]. Finally we explain variations in persistency by relating the parameter to different firm level as well as industry level indicators. Obvious candidates for these indicators are entry and exit rates as well as advertising spending, capital intensity, etc., which should pick barriers to entry/exit.

The framework has been used by several researchers and as mentioned before, most of them have reported a generally high value of this statistic in the range 0.4-0.5. Examples include Geroski and Jacquemin (1988), Mueller (1990) and Goddard and Wilson (1999). Glen et al. (2003) have found a slightly lower value for developing countries, namely around 0.2-0.3.4

3 Data

In order to estimate persistency of profits we use firm-level data on total assets and profits before tax are retrieved from the FOD database. The database collects company accounts data of all firms active in Belgium, except for one-man businesses and is constructed using data from the National Bank of Belgium. The result is an unbalanced panel of firms for the period 1999-2008 active in all sectors of the economy. In general, the literature defines the profit rate as the ratio of profit before taxes over either total assets or total sales. However since the smallest firms in Belgium do not have to report sales data, we use profit before taxes over total assets[5].

[3] Shepherd (1972)has shown profit rates increase systematically with size within an industry.
[4] Detail is included in the Appendix Table A1.
[5] Total assets (code 50/58) includes all fixed assets (code 20/58) and current assets (code 29/58). Profits before tax (code 9903) includes operating incomes and charges, taken into account of depreciation, financial and extraordinary operation.

Since we do not observe economic profits, we have to use accounting profits instead. As is well known, the use of accounting profitability measures can diverge from economic profitability. For example, differences in accounting profits across sectors can be caused by different accounting conventions. However, these biases are more likely to be relevant for differences in profitability levels than for differences in the persistence of profits. Only changes in accounting practices over time that differ across industries could be problematic for a comparison of profits persistency across sectors. Moreover Kay and Mayer (1986) found persistently high accounting rates of return indicates persistently high economic rates of return. In addition, as we robustness check we also run the analysis using operating profits over total assets as our profitability measure. Since operating profits do not include depreciation, amortizations and, etc, the measure is less prone to accounting practices. We normalize the profitability ratios by subtracting the yearly average profitability ratio in the Belgian economy[6].

We perform some cleaning on the dataset. First, we restrict the analysis to firms with 5 or more consecutive observations. Second we drop the top and bottom 5 % of profit rates in order to avoid problems with outliers. In the end we are left with an unbalanced panel data set for more than 200,000 companies in Belgium operating from 1999 to 2008.

Tables

Table 1 provides some summary statistics of the profitability rate of firms active in Belgium. For the balanced panel, the average profitability rate is 3.9% and we observe a firm for on average 8.15 years. Not surprisingly, moving to the balanced panel increases the profitability rate which rises to 4.3%.[7] We divide the firms into three size categories based

---

[6] We also experimented with normalizing the profitability ratio with the sector/year average, but this did not change our results.

[7] This highlights one inherent problem with the profit persistency literature, namely that we are obliged to focus on the subset of firms that have survived for a number of periods. However note that

on turnover. Small firms realize a turnover of less than 2 million euros, medium firms realize a turnover between 2 million and 10 million euros and large firms have a turnover of over 10 million euros. In line with expectations and consistent with many empirical and theoretical papers, larger firms have higher profit margins (Sheperd, 1972). Finally we also compute profitability as the ratio of operating profits over total assets which is on average 5.2%.

4 Results

In this section we provide a discussion of our main results. We estimate equation (1) using our large unbalanced panel dataset of over 200,000 Belgian firms. The results for the short-run persistency parameter are reported in Table 2. The average short term persistency parameter equals 0.056, which is low especially in comparison to other studies. However, the standard deviation of the short term persistency is fairly high and equal to .39 pointing to substantial variation across firms. Moreover, it is well known that estimating an autoregressive model using ordinary least squares results in a small-sample bias which could be important since the average time period for our sample is only slightly higher than 8 years. Fortunately we know the size of the bias and can ex post correct our estimates for it. When we apply the procedure described by Patterson (2000), we obtain an unbiased estimate for the average short-run persistence parameter and we find the average λ to be equal to 0.172. This estimate points to a certain extent of short-run persistency, but still substantially lower compared to other studies (cf. Table A 1)[8]. We turn back to this issue on

in our analysis we only constrain the firm to exist for at least 5 periods while other studies focused on large firms being in business for over 15 years.

[8] Moreover, note that most of the papers mentioned do not control for the small sample bias and are as such lower bounds to the true underlying parameter. However, the bias will be lower compared to the present study as the number of observations per firm is higher.

the following pages. Due to the low number of observations per firm, the firm-level λ is often not significantly different from zero at the 10% level.

When we aggregate the short-run persistency parameter using the weighted average with sales as weight, we find a substantially higher persistence of profit which already indicates that large firms are better in insulating their profits from competition, an issue we will treat in more detail in the next section. Moving to the balanced panel, we find the average persistency to be equal to 0.123 (bias corrected: 0.230), higher than for the unbalanced panel, which is in line with our priors as firms that can protect their profits from competitive forces are more likely to survive and consequently more likely to be observed over the whole sample period. The percentage of firms with a short-run profit persistency significantly higher than zero is also higher as the number of observations per firm went up and as such the accuracy of the estimates increased. Finally we run the firm level regressions using operating profits over total assets as our profitability measure. Now, the average short run persistency is slightly higher compared to baseline profitability ratio (profit before taxes over total assets).

In a second step, we look at hetergeneity across different firms in terms of profits persistency. As can be seen from Table 3 large firms are better in protecting their competitive advantage in terms of efficieny or market power from competitive forces. The bias corrected esitmate for short-term persistency of large firms equals .289 compared to .157 for small firms where the categories are defined using the operating revenue of the firms. Note that this can explain part of the result that we find profit persistency to be lower in Belgium compared to previous studies as they used mainly large, even stock-quoted, firms. The rest of the difference is likely to be explained by the different time periods of the empirical analyses.

Third, we turn to sector heterogeneity in profit persistency. When we compute the average of the autoregression parameter for each different NACE 3 digit sector, we can see there exist substantial heterogeneity across firms as displayed in Figure 1.[9] These differences in profit persistency can be used to draw inferences about the strength of competition in a sector. First, we rank the NACE 2 digit sectors in terms of profit persistency. The results are displayed in Table 4. Not surprisingly, the Electricity and Gas sector ranks the highest in terms of profit persistency. Also other sectors which are thought off to have high entry barriers such as the Manufacture of Chemicals and Chemical Products and Manufacture of Rubber and Plastic Products have high levels of persistency. Among the sectors with the lowest persistency are the Forestry and Logging sector as well as the Sewerage and Travel Agency sector. Except for the Sewerage sector[10], these are sectors with low sunk costs and/or simple production technologies.

The ranking of the NACE 3 digit sectors is displayed in Table 5 and Table 6. Sectors with high persistence of profits include Mining and Quarrying, Manufacture of Gas, Water Transport and Steam and Air Conditioning Supply. Again the appearance of these sectors as having high persistence of profits is not surprising and builds up some confidence in the indicator. Turning to the sectors with low persistence of profits, the results are more surprising as sectors such as the Manufacture of Coke Oven Products appear in the list. However, these are typically smaller sectors and the average persistency could be less prone to measurement errors and alike. This is certainly an issue we should take up in future versions of the paper.

---

[9] The average measures of profit persistency displayed here and in the next paragraphs are not corrected for the small sample bias. This is not an issue as the correction of the small sample bias involves a monotone transformation of the parameter for a fixed T. Since we are now only interested in the ranking of the sectors in terms of profit persistency and the observations per firm do not substantially differ across sectors, the ranking of sectors is not altered by the small sample correction.

[10] Note that the Sewerage sector also contains publicly owned companies.

As mentioned before there are various ways to compute persistence of profits. Ideally, the inferences drawn about the competition intensity in a sector are not dependent on the metric/methodology used. In Table 7 we display rank correlations of the aggregate persistence of profits at the sector level between different metrics/methodologies. First we check whether the choice to take a weighted or unweighted average matters for the ranking of the sector. It appears from the first column, that the correlation between the unweighted and weighted average is positive albeit small, especially for the higher the level of aggregation. Second, we check whether moving from the unbalanced to the balanced panel changes results. We find the correlation between the two options to be fairly high around 0.6. Finally, we check whether the choice of the profitability definition drives results and we find this not to be the case (correlations of about 0.8).

An important question is which sector characteristics drive the differences in persistence of profits. The most obvious candidates are clearly entry barriers such as economies of scale and sunk entry costs such as R&D or advertising. Waring (1996) finds both economies of scale as well as R&D intensity to be positively correlated with profit persistency. Instead of looking at possible entry and exit barriers one can also look at the result of the presence (or absence) of these barriers, namely one can look at the churn rate. Other factors that can impact the persistence of profits include the concentration in the sector, the complexity of the production process, the unionization of the sector, capital intensity of the sector, …

We relate profit persistency at the NACE 3 digit level with the churn rate, Herfindahl-Hirschman Index and the capital intensity as measured by the ratio of capital stock over sales in the sector. The results are reported in Table 8. The churn rate is as expected negatively correlated with the persistence of profits at the sector level and thus holds true for the whole sample as well as for services and manufacturing sectors separately. So a higher churn implies lower persistence of profits. Although the HHI index is positively correlated with concentration at for the whole sample, this correlation disappears when looking at the services and manufacturing separately. Note that this is not really surprising as the HHI is not well defined for manufacturing sectors as the measure does not take into account imports which are substantial in a small open economy as Belgium. Finally, capital intensity is negatively correlated with persistency, if anything. This is at first sight a surprising result as capital intensity is expected to pick up returns to scale. However, Waring (1996) has found a similar result and attributes this to capital utilization. Firms rarely produce up t o full capacity and if a competitor earns high profits, they can easily adjust their production level by increasing their capital utilization, thereby eroding the competitors' profits.

5 Conclusion

Determining the intensity of competition is a key interest in the field of industrial organization. Static measures such as price-cost margins or concentration ratios may inadequately reflect the intensity of competition in a number of cases. A solution is to look at the competitive dynamics and examine the degree of profits persistency. The general idea is that in an efficient market economy, supra-normal profits should quickly disappear as they attract new entrants or imitators. The increase in competitors erodes profits earned by the initially successful incumbent. However, when firms operate in a less competitive environment, profits may be persistent and do not fall back to their competitive level. In order to analyze the persistence of profits in Belgium, we use data on around 200,000 firms between 1999 and 2008, retrieved from their income statements Contrary to previous persistence of profits studies we include also small firms into the analysis. We find a certain amount of persistence of profits in the Belgian economy, albeit lower compared to other countries. Furthermore, we show how the inclusion of small firms in the analysis can have important consequences as they have substantially lower persistence of profits compared to large firms.

The richness of the dataset furthermore allows us to examine the persistence of profits along various dimensions. We find Sector heterogeneity to be substantial. The highest persistency is found in sectors such as Mining and Quarrying, Manufacture of Gas, Steam and Air Conditioning Supply which are known to have high entry barriers. Furthermore we relate the persistence of profits with other competition indicators such as the churn rate, concentration and capital intensity. The strongest correlation is the one with the churn rate. Obviously in future versions of the paper we will relate the persistence of profits with other variables that for example should pick up the complexity of the production process. Moreover, this will be done in a multivariate regression framework.

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *The Quarterly Journal of Economics*, 120(2), 701-728.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4), 841-890.

Fisher, F. and McGowan, J. (1983). On the misuse of accounting rates of return to infer monopoly profits. *American Economic Review*, 73, 82-97.

Genesove, D., & Mullin, W. P. (1998). Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914. *The RAND Journal of Economics*, 29(2), 355-377.

Geroski, P. A., & Jacquemin, A. (1988). The Persistence of Profits: A European Comparison. *The Economic Journal*, 98(391), 375-389.

Geroski, P.A.(1990). Modelling persistent profitability. In: Mueller, D.C. (Ed.), The Dynamics of Company Profits: An International Comparison. Cambridge University Press, Cambridge, 15–34.

Glen, J., Lee, K., & Singh, A. (2001). Persistence of profitability and competition in emerging markets. *Economics Letters*, 72(2), 247-253.

Glen, J., Lee, K., & Singh, A. (2003). Corporate profitability and the dynamics of competition in emerging markets: a time series analysis. *The Economic Journal*, 113(491), F465-F484.

Goddard, J. A., & Wilson, J. O. S. (1999). The persistence of profit: a new empirical interpretation. *International Journal of Industrial Organization*, 17(5), 663-687.

Hall R.E. (1988).The Relation between Price and Marginal Cost in U.S. Industry.*Journal of Political Economy* ,96(5), 921-947.

Harrison, A. E. (1994). Productivity, imperfect competition and trade reform : Theory and evidence. *Journal of International Economics*, 36(1-2), 53-73.

Kay, J. and Mayer, C. (1986). On the application of accounting rates of return. *The Economic Journal*, 96, 199-207.

Levinsohn, J. (1993). Testing the imports-as-market-discipline hypothesis. *Journal of International Economics*, 35(1-2), 1-22.

McGahan, A. M., & Porter, M. E. (1999). The Persistence of Shocks to Profitability. *The Review of Economics and Statistics*, 81(1), 143-153.

Mueller, D. C. (1977). The Persistence of Profits Above the Norm. *Economica*, 44(176), 369-380.

Mueller, D. C.(1986). Profits in the Long Run. Cambridge University Press, Cambridge, UK.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, 69(2), 307-342.

Nickell, S. J. (1996). Competition and Corporate Performance. *The Journal of Political Economy*, 104(4), 724-746.

Patterson, K. D. (2000). Bias reduction in autoregressive models. *Economics Letters*, 68(2), 135-141.

Porter, R. H. (1983). A Study of Cartel Stability: The Joint Executive Committee, 1880-1886. *The Bell Journal of Economics*, 14(2), 301-314.

Shepherd, William G. (1972). The Elements of Market Structure. *The Review of Economics and Statistics*, 54, 25.

Syverson, C. (2004). Product Substitutability and Productivity Dispersion. *Review of Economics and Statistics*, 86(2), 534-550.

Yurtoglu, B.B.(2004). Persistence of Firm-level Profitability in Turkey. *Applied Economics*, 36, 615-625.

Tables

Table 1 Summary Statistics

		Average Profitability	Nr. Firms	Obs. Per Firm
Profit Before Tax/Total Assets	Full Sample	0.039	205034	8.15
	Balanced	0.043	89560	10
Size Categories	Small Firms	0.036	101397	8.13
	Medium Firms	0.047	13359	8.68
	Large Firms	0.052	6907	8.82
Operating Profits/ Total Assets	Full Sample	0.050	205376	8.17

Table 2 Results Short Term Persistency

	Average λ	Standard Deviation λ	% Significantly >0
Full Sample			
Unweighted	0.056	0.39	0.178
Weighted	0.171		-
Balanced Panel	0.123	0.36	0.215
Operating Profits/Total Assets	0.074	0.39	0.193

Table 3 Persistence of Profits over Different Size Categories

Category	Av. λ	λ Bias Corr	Nr. Obs.	Criterium
Small	0.040	0.157	97126	OR < 2 million
Medium	0.130	0.243	12689	2 mill. < OR < 10 mill.
Large	0.174	0.289	6556	10 mill. < OR

Profitability measure is profits before tax over total assets. Unbalanced panel.

Table 4 Persistence of Profits per NACE 2 Digit Sector

NACE2	NACE Description	Persistency
High Persistency		
35	Electricity, gas, steam and air conditioning supply	0.1813
22	Manufacture of rubber and plastic products	0.1696
17	Manufacture of paper and paper products	0.1588
20	Manufacture of chemicals and chemical products	0.1573
65	Insurance, reinsurance and pension funding, except compulsory social security	0.1525
36	Water collection, treatment and supply	0.1508
08	Other mining and quarrying	0.1478
14	Manufacture of wearing apparel	0.1424
60	Programming and broadcasting activities	0.1391
75	Veterinary activities	0.1217
Low Persistency		
30	Manufacture of other transport equipment	0.0293
53	Postal and courier activities	0.0268
56	Food and beverage service activities	0.0211
41	Construction of buildings	0.0179
43	Specialised construction activities	0.0156
81	Services to buildings and landscape activities	0.0086
01	Crop and animal production, hunting and related service activities	0.0070
79	Travel agency, tour operator and other reservation service and related activities	0.0062
02	Forestry and logging	-0.0220
37	Sewerage	-0.0234

Unweighted average autoregressive parameter per NACE 2 digit sector. Sectors with lowest and highest profit persistency are reported. Profit before taxes over total assets as profitability measure, unbalanced panel of firms.

Table 5 High Persistency NACE 3 digit Sectors

NACE3	NACE Description	Persistency
089	Mining and quarrying n.e.c.	0.3826
352	Manufacture of gas; distribution of gaseous fuels through mains	0.3185
501	Sea and coastal passenger water transport	0.2913
353	Steam and air conditioning supply	0.2826
152	Manufacture of footwear	0.2602
102	Processing and preserving of fish, crustaceans and molluscs	0.2344
302	Manufacture of railway locomotives and rolling stock	0.2309
104	Manufacture of vegetable and animal oils and fats	0.2202
143	Manufacture of knitted and crocheted apparel	0.2186
822	Activities of call centres	0.2122
601	Radio broadcasting	0.2114
261	Manufacture of electronic components and boards	0.1835
651	Insurance	0.1834
201	Manufacture of basic chemicals, fertilisers and nitrogen compounds, plastics and synthetic rubber in primary forms	0.1826
274	Manufacture of electric lighting equipment	0.1820
782	Temporary employment agency activities	0.1706
222	Manufacture of plastics products	0.1706
236	Manufacture of articles of concrete, cement and plaster	0.1689
643	Trusts, funds and similar financial entities	0.1649
171	Manufacture of pulp, paper and paperboard	0.1646

Unweighted average autoregressive parameter per NACE 3 digit sector. Sectors with lowest and highest profit persistency are reported. Profit before taxes over total assets as profitability measure, unbalanced panel of firms.

Table 6 Low Persistency NACE 3 Digit Sectors

NACE3	NACE Description	Persistency
663	Fund management activities	0.0065
681	Buying and selling of own real estate	0.0057
242	Manufacture of tubes, pipes, hollow profiles and related fittings, of steel	0.0042
493	Other passenger land transport	0.0039
582	Software publishing	0.0032
431	Demolition and site preparation	0.0019
813	Landscape service activities	-0.0026
268	Manufacture of magnetic and optical media	-0.0027
291	Manufacture of motor vehicles	-0.0045
439	Other specialised construction activities	-0.0104
370	Sewerage	-0.0234
266	Manufacture of irradiation, electromedical and electrotherapeutic equipment	-0.0238
279	Manufacture of other electrical equipment	-0.0302
799	Other reservation service and related activities	-0.0539
301	Building of ships and boats	-0.0649
303	Manufacture of air and spacecraft and related machinery	-0.0658
783	Other human resources provision	-0.0741
272	Manufacture of batteries and accumulators	-0.1000
191	Manufacture of coke oven products	-0.2106
652	Reinsurance	-0.2183

Unweighted average autoregressive parameter per NACE 3 digit sector. Sectors with lowest and highest profit persistency are reported. Profit before taxes over total assets as profitability measure, unbalanced panel of firms.

Table 7 Correlation between Different Approaches

	Spearman's rho Between Different Approaches		
	Unweighted & Weighted	Unbalanced & Balanced Samples	Before Tax Profits & Operating Profits
NACE2	0.0133	0.6643	0.8015
NACE3	0.2191	0.5404	0.7843
NACE4	0.3208	0.5526	0.7211

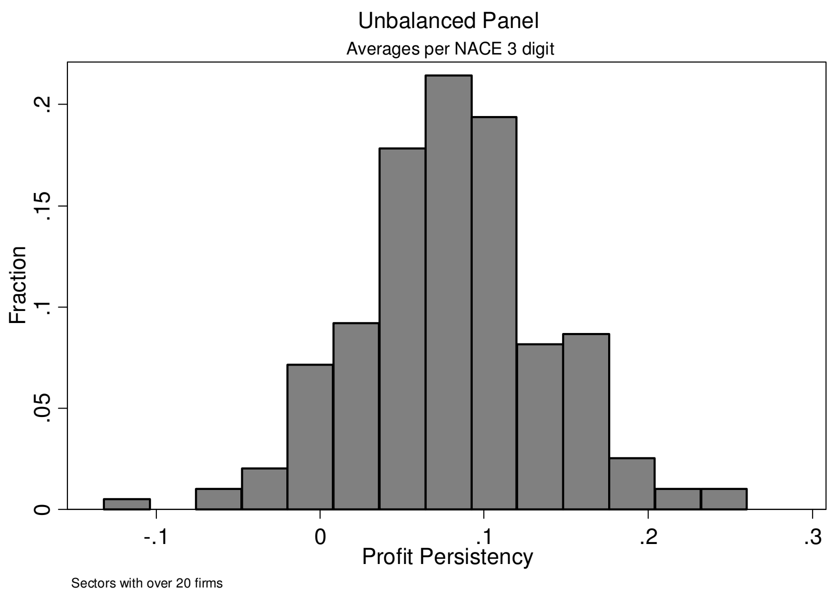
Table 8 Correlation Persistence with Other Indicators

	Churn Rate	Concentration	Capital Intensity
All	-0.283	0.193	-0.091
Manufacturing	-0.229	-0.081	0.086
Services	-0.219	0.002	-0.315

Spearman rank correlation between indicators and persistence of profits at the NACE 3 digit sector level.

Figures

Figure 1 Profit Persistency per NACE 3 digit sector



Appendices

Table A 1 Overview Studies Profit Persistency. [a]

Author	Country	Sample Period	Obs./firm	No. firms	Sample Mean (λ)
Geroski & Jacquemin (1988)	UK	1947-1977	29	51	0.488
	France	1965-1982	18	55	0.412
	Germany	1961-1981	21	28	0.410
Schwalbach et al. (1989)[b]	Germany	1961-1982	22	299	0.485
Mueller (1990)	US	1950-1972	23	551	0.183
Cubbin and Geroski (1990)	UK	1948-1977	30	243	0.482
Khemani & Shapiro (1990)	Canada	1964-1982	19	129	0.425
Odagiri & Yamawaki (1990)	Japan	1964-1982	19	376	0.465
Schohl (1990)	Germany	1961-1981	21	283	0.509
Waring (1996)	US	1970-1989	20	12, 986	0.540
Glen et al. (2001)	Brazil	1985-1995	11	56	0.013
	India	1982-1992	11	40	0.221
	Jordan	1980-1994	15	17	0.348
	Korea	1980-1994	15	82	0.323
	Malaysia	1983-1994	12	62	0.349
	Mexico	1984-1994	11	39	0.222
	Zimbabwe	1980-1994	15	40	0.421
Yurtoglu (2004)	Turkey	1985-1998	14	172	0.380

Source: Glen et al. (2001), for all except Glen et al (2001) and Yurtoglu (2004)

[a] All references are from Glen et al (2001), except Glen et al (2001) and Yurtoglue (2004).

[b] Based on nominal profit on capital, before tax.