Industry similarities, comparative advantage and structural transformation: the case of Portugal, 2005.

Miguel Lebre de Freitas* (UA, GEE, NIPE)

Susana Salvado (UNL, GEE)

Work in Progress

October 2007

Abstract: this paper provides an empirical assessment of the "sophistication" of the Portuguese pattern of trade in 2005 and its possible implications for the process of structural transformation. Using trade data at the product (NC4) level for 96 countries in 2005, we first compute a measure of sophistication for each product (PRODY), as a weighted average of the per capita incomes of countries that export it. Following Hausmann and Klinger (2006), we then use the structure of international trade in 2005 to assess the similarity between each pair of products, in terms of the capabilities they use. The method consists in estimating the extent to which a country having comparative advantage in one product increments the probability of the same country having comparative advantage in another product. Contrary to Hausmann and Klinger, our measure of "proximity" is subject to a statistical scrutiny. Implementing a probit model with robustness checks, we show that a large number of branches between products are not significant. For each product, we then explore some measures assessing how distant it is from products with higher income content. We then investigate the extent to which upscale products in which a country didn't develop RCA are surrounded by products in which the country already has developed RCA. These measures are then aggregated at the country level to assess the opportunities of each country in the process of structural transformation.

Key-Words: International trade, specialization patterns, Structural Transformation, Industry Heterogeneity, The Portuguese Economy.

JEL: C14, F14

---

* Corresponding author. Campus Universitário, 3810-193 Aveiro, Portugal. Tel: 351-234-370200. Fax: 351-234-370215. E-mail: afreitas@egi.ua.pt. The authors aknowledge…

1. Introduction

The current debate on the Portuguese economy is dominated by the view that the country' specialization pattern – arguably dominated by low-skilled labour intensive products – impairs the convergence process towards higher per capita income levels. As new trading partners competing in similar segments emerge, there is an increasing believe that the future performance of the Portuguese economy depends critically on its ability to shift its specialization pattern towards more sophisticated goods. In the case of Portugal, an extensive literature already exists focusing on the role of institutions, especially those in the labour market, as a major constraint to the process of structural transformation. Less attention has been given, however, to *industry heterogeneity* as a main barrier to industry reallocation.

The idea that a country' economic performance is to a large extent determined by the type and characteristics of the sectors wherein the country specializes has been stressed by many authors (for example, Kaldor, 1966, Thirlwall, 1979, Grossman and Helpman, 1991, Hausmann et al., 2005, Hausmann and Rodrick, 2006). Empirically, however, such proposition is difficult to test. Dalum et al. (1999) shows that the characteristics of the specialization pattern are important to explain growth differentials, but the results are sensitive to alternative classifications of sectors into different technological categories that the author considers. Feenstra and Rose (2000) find a strong relation between what they call "advanced export structure" and high productivity levels and fast growth rates. More recently, Hausmann et al. (2005) proposed a quantitative measure of the income content of each product, by calculating the weighted average of the GDP per capita of the countries that export it. This measure allows products to be ranked according to its "productivity content" or "sophistication" level. Using this index, the authors then construct a quantitative measure of the overall specialization pattern of each country, called EXPY. The authors found that this measure is highly correlated with capita incomes and that it is also a good predictor of subsequent growth. The conclusion is that "poor countries export poor country goods and rich countries export "rich country goods".

This paper draws on these findings and addresses the extent to which the Portuguese economy will able to move towards a more "sophisticated" pattern of trade.

Such a process will be labelled *Structural Transformation*. The analysis follows Hausmann and Klinger (2006). The authors argue that, because industries differ in terms of the production capabilities they need, the ability of a country to move towards a pattern of specialization with higher value added depends on the existence of the required capabilities. Because, in turn, capabilities develop and are accumulated in each economy as a bequest of previously existing activities, the implication is that the capacity of economies to move to new activities will depend on the existence of potential activities that can use the capabilities bequested. This means that that process of structural transformation is path dependent.

The authors proposed a new measure of similarity between products, given by the probability of having comparative advantage in one of them, conditioned on the probability of having comparative advantage in the other. They then relate the probability of a country developing comparative advantage in a product to the density of the country's current production relative to that good. Aggregating the data to the national level, the authors investigate whether the speed of structural transformation is affected by the current pattern of specialization and the structure of the product space.

In this paper, we develop an alternative method to measure the proximity between pairs of goods. Instead of computing non-parametric conditional probabilities for all pairs of goods, we run a probit model to estimate the increment in the probability of a country having RCA in one product, given that it has RCA in other product. The advantage of this methodology is that is subjects the proximity measures between each pair of goods to a statistical scrutiny. As robustness check, we evaluate the extent to which the t-ratio capturing the cross effect between each pair of products remains significant with the presence of a third product, and we let this third product to vary along all products in the product space. Estimating the matrix and performing the robustness involves running more than 2 billion regressions. We show that the number of non-significant connections is quite large.

Using the information on the number of significant segments linking each product, the fraction of these that represent upscale movements, the estimated proximities between products and the existing pattern of revealed comparative advantages, we then develop alternative indicators to access the extent to which the Portugal' current specialization patterns helps or impairs the process of structural transformation.

2. The Hasumann-Klinger forest

Hausmann and Klinger (2006) argue that the process of structural transformation is highly path dependent. The idea is that the production of each particular good requires a set of "specific" inputs, "specificity" meaning that these inputs would be much less productive if deployed in some other activity. This includes technical knowledge, physical assets, intermediate inputs, labour skills, access to markets, public infrastructure and specific regulatory requirements. Because industries differ in terms of the production capabilities they need (i.e the human, physical and institutional capabilities to produce cotton trousers are closer to those needed to produce cotton shirts than those needed to produce computer monitors), the ability of a country to move towards a pattern of specialization with higher value added depends on the existence of the required capabilities.

A problem, however, is that capabilities are not uniformly distributed across economies. The neo-classical theory of comparative advantages goes along with this idea, stressing that a move towards rich country goods may be impaired by lack of appropriate endowments. For example, because poorer countries have a higher relative endowment of, say, unskilled labour, they are more likely to export goods that use unskilled labour intensively. A limitation of the neo-classical model is that it takes factor endowments as given. The recent developments in growth theory have however emphasized that, with the exceptions of natural resources, endowments shall be seen as endogenous to the development process (Lucas, 1990, Hall and Jones, 1999, Murphy et al., 1989). At the microeconomic level, Jovanovic and Nyarko (1996) explore the idea of industry *heterogeneity*. In particular, they build a model of learning by doing and technological upgrading at the individual level, whereby experience provides agents with information that improves their productivity in the given technology (vertical shift). Because gains in that direction are limited, agents may also "jump" to new technologies (horizontal shift). The degree of similarity of the new technology to the old determines how transferable the accumulated knowledge is.

These developments point to the case that, in general, capabilities shall not be seen as exogenous in respect to the development process. On the contrary, resource endowments develop and are accumulated in each economy as a bequest of previously existing activities. The implication is that the capacity of economies to move to new activities will depend on the existence of potential activities that can use the capabilities bequest by previous activities. This means that that process of structural transformation is largely dependent on which industries the country already learned to operate.

To illustrate this problem, Hausmann and Klinger (2006) used the metaphor of a forest, where each tree represents a product and the forest represents the product space. In that forest, each tree is placed at some distance from the other trees, the distance capturing the degree to which the production capacities of one product can be used in other product. Moving to trees at larger distances involves the need for productive capabilities that have not been previously accumulated.

Because some industries use skills that are common to a large number of industries, some parts of the forest are more dense than others. In this metaphor, firms are monkeys that live on trees and the process of structural transformation involves the monkeys jumping around from tree to tree. Because some trees generate more income than others, each monkey would like to move to high productivity trees ("rich-country goods"). However, because smaller jumps are less costly than larger jumps, the ability of the country to engage in *upscale jumps* depends on having a path to nearby goods that are increasingly of higher value. Hausmann and Rodrick (2006) and Rodrick (2006) argue that industrial policy has a role in removing coordination failures and helping the tribe to move toward areas in the forest with higher growth potential.

3. Measuring proximity

To build the forest map, Hausmann and Klinger (2006) used data on international trade. Their argument is that the similarity of capabilities (or proximity between trees) is related to the likelihood that countries have revealed comparative advantage in both goods. To develop a measure, they used product level data on exports and estimated, for each pair of products, the conditional probability that countries having revealed comparative advantage on the first product also have on the other. These conditional probabilities are then used to develop a measure of "revealed proximity" between each two products, that maps the product space. These distances intend to capture the degree of similarity of the requisite capabilities.

Because the conditional probability of a country having comparative advantage in a good i given that it has comparative advantage in good j, P(i/j), is different from the conditional probability of a country having comparative advantage in a good j given that it has comparative advantage in good i, P(j/i), the simple application of conditional probabilities would lead to an asymmetric matrix. Arguing that these conditional probabilities may tend to extreme values in cases where only few countries have comparative advantage in one of the goods, Hausmann and Klinger (2006) imposed symmetry in the matrix of proximities, by measuring "proximity" between each two goods i and j as the minimum of the two above mentioned conditional probabilities.

In this paper, we adopt an alternative method to estimate the "proximities" between pairs of goods. In particular, for each one of the 1245 products in our sample, we implement a probit model, assessing the extent to which the probability of a country having RCA in one product increases by the fact of having RCA in another product. Since we use NC-4 classification encompassing 1245 products, we have a total of 1244x1244 cells to estimate in the matrix of all possible relations between pairs of products. To each pair of goods, we then estimate the increment in probability - the marginal effect – which captures technologic *proximity* between these two products (see Appendix 1 for details).

Table 1 exemplifies the implementation of such an exercise for a particular product, corresponding to one line of our 1245X1245 matrix. The product *i* at hand is "6302 – bed linen, table linen, toilet linen and kitchen linen". The first column in the table (1) displays the estimated marginal effects, β_{ij}, which correspond to the estimated increase in the probability of having RCA in each product j given that the country has RCA in product 6302. The advantage of our method is that it provides a significance test for the estimated proximity. Column (2) of Table 1 displays the *t-test*s corresponding to the proximities estimated, consisting on the ratio between the estimated coefficient and the standard error, which in this case has a normal distribution (actually, in Table 1, the j-products are displayed by decreasing order of *t*). In the bottom line of Table 1, we see that the number of significant segments according to this criteria, is only 625, thus roughly half of the potential significant connections.

This evidence challenges Hausmann and Klinger (2006). Because these authors considered all possible connections leading to each good, they are likely to be overestimating the available options in the process of structural transformation.

Not yet implemented:

A well known problem of estimating with a large number of explanatory variables is that the significance with a particular combinations of variables in a equation does not necessary guarantees significance with other combination of variables in the equation. In other words, a significant *t-test* does not preclude the possibility of the corresponding coefficient be spurious, capturing the influence of an omitted variable. Following Sala-i-Martin (1996), we check the robustness of all regressions involving two products, by testing the resilience of its t-test to the presence of a third product in which the country has comparative advantage. We run all possible third variables, which means that we run 1243 equations with 3 products, for each one of the 1244*1244 cells to be estimated in the matrix of proximities. We assume that the relation between the two variables is robust if the t-statistics of the marginal effect remains significant in at least 95% of the 1243 regressions performed with three variables. That being the case, we turn to the original estimation with two variables and we accept the estimated measure of proximity (see Appendix 2, for details). The results of this robustness test are displayed in column (3) of Table 1. We mark with 1 the cases that pass the robustness check and with 0 the cases in which the robustness test failed. In the bottom line of the table, we see that the number of significant "proximities" after the robustness check has decreased to z (w% of the potential significant connections).

….

Another novelty of our estimation is that we do not impose symmetry in the matrix of "proximities". The rationale is that the increment in the probability of having RCA in product *i* because of RCA in product j does not need to be the same as in the increment in the probability of having RCA in product j due to RCA in i. As an example, think on automobiles and carpets: a country having comparative advantage in producing automobiles may explore a synergy, developing the activity of producing carpets for these automobiles. However, a country producing carpets does not endow it with RCA in producing automobiles. Because we are interested in the connections departing from each good in which a country already has RCA and the connections leading to each (upscale) good in which the country has not RCA, both directions are important. The disadvantage of not assuming a symmetric matrix is that we will not be able to map the product space in a two dimension plan, as done by Hausmann and Klinger (2006).

4. Measuring the productivity content of traded goods and the "sophistication" of specialization patterns: PRODY and EXPY

In order to map the opportunities of a country in terms of increasing the level of sophistication of its exports, one needs a measure of sophistication at the product level. Hausmann et al. (2005) proposed the PRODY index. The index intends to capture the degree of "sophistication" of sectors, by relating it with the level of development of countries that exporting it. Formally, the index is defined, for each good, as the weighted average of per capita incomes of countries exporting that good, where the weights are related to the country's Revealed Comparative Advantage in that good (a formal explanation in Appendix 3). Sectors with high values of PRODY are, by construction, those where high income countries play a major role with respect to the other trading partners. The implied assumption is that the presence of higher wages is stronger where comparative advantage are determined by factors other than labour cost, such as know how, technology, etc. Hausmann and al., (2005) also computed an index, called EXPY, intending to capture the productivity content associated to a country' export vector. This is an weighted average of the PRODYs of the products exported by that country. The authors found a strong relationship between EXPY and GDP per capita.

Figure 1 displays our estimations of EXPYs, using trade data for the year of 2005 at the product (NC4) level for 96 countries, and IMF figures for Per Capita GDP in 2005 (PPPs). The figure confirms the high correlation between this indicator and per capita incomes. Portugal, is located slightly above the regression line, meaning that its income performance is slightly better than that implied by its specialization pattern. This path is common to the developed countries in general and is likely to reflect better policies and institutions in these countries.

Figure 2 displays the relationship between our estimated PRODYs and the Balassa index of RCA in the product space, for the case of Portugal[1]. The figure roughly suggests a negative relationship between the two, meaning that products in which the country specializes more are of lower income content.

---

[1] Both variables are in Logs. Null coefficients of RCA became missing values.

Because the figure is silent in respect to sector sizes, we display in Figure 3 a similar relationship, but only for products weighting at least 0,2% in the Portugal' export basket (bubble areas measure shares on exports). The figure still reveals a negative relationship and with a steeper slope, suggesting that the negative relationship between RCA and value is more pronounced in these large sectors.

Figure 4 repeats the exercise, but only for products in which Portugal has positive RCA. As in the other two figures, the slope of the regression line is negative.

5. How valuable is your knowledge?

We next turn to the question of to which extent the goods in which the country developed RCA may help or impair the process of structural transformation. To analyze this question, we compute, for each product, a number of indicators.

Columns (1), (2) and (3) of Table 2 display, for each product *i*, respectively, the value of the Balassa RCA index in Portugal, the PRODY corresponding to this good and the percentage of products in the sample with superior PRODY.

The three products displayed in Table 2 are products in which Portugal exhibits a significant RCA. Their PRODY levels are quite low in terms of the other products. For example, the product 6109 "T-Shirts" is in percentile 83% meaning that only 17% of the 1245 products have lower income content.

Column (4) of Table 2 displays the number of *significant*[2] connections involving each good (the same information is displayed in the bottom line of Table 1, column (2) for the case of 6302 "Bed linen"). Note that these may be computed departing from a good in which the country has RCA (outward move) or arriving to a good in which the country still didn't develop RCA (inward move). At this stage, we are concerned with the former concept. In the following section we will address inward movements. For the case of 6302 bed linen, we see that the number of (non-robust) significant segments is 625. In the case of 6403 "Footwear", the number of significant connections is 922.

---

[2] At this stage, by significant connections we refer those with 97,5% significance in the first round estimations. Robustness checks are still to be performed.

Not all significant connections, however, are conducive to products with superior income content (PRODY). Column (5) of Table 2 reports the proportion of significant segments leading to products with lower PRODY. In the case of 6302 "bed linen", this proportion is 21%. This compares to 11% in the case of T-shirts and 17% in the case of Footwear.

A different question relates to the size of the potential upscale moves. Columns (6), (7) and (8) of Table 2 display the number of segments departing from each product *i* to products *j* with PRODY at least 0%, 10% and 50% higher than i. For example, in the case of 6302 "bed linen", only 492 out of the 625 significant connections are conducive to products with at least the same PRODY level; 448 segments are conducive to products with PRODYs at least 10% larger; and 184 segments are conducive to segments with PRODYs at least 50% higher.

Figure 5 reports similar calculations for each 5% increase in the required upscale move. For each product, the figure displays a downward step-function, m_{ik}, returning the number of significant segments leading to products with at least k% higher value, with k increasing 5% at the time. This figure allows comparisons between goods without the need to mix information on marginal probabilities and PRODYs. For each good, the function m_{ik} gives an idea on how "dense" and how "rich" is the part of the forest where the good is located. Comparing the 3 products at hand, we see that 6403 "Footwear" has more significant connections, leads to more positive increments in value and has in general more connections to goods located in the richer part of the forest[3].

Figure 6 displays, for the case i= "t-shirts", the relationship between "distances" to other goods j (defined as the inverse of the "proximity" measure) and the increment in value associated to each move, $\lambda_{ij} = \ln(PRODY_j) - \ln(PRODY_i)$. Notably, these good looks conducive to a large spectrum of values, at low distances (between 1 and 2).

---

[3] Note however, that the horizontal axes measures percentage changes, so that the origin does not correspond to equal PRODYs. A different view would be obtained if the horizontal axis consisted in the PRODY levels themselves.

The last column of Table 2 also displays a parametric index measuring the value of a product neighbourhood. It basically consists on a weighted sum of the increments in value associated to each segment, where the weights are the marginal probabilities, β_{ij}:

$$v_i = \sum_{j:\beta_{ij},\lambda_{ij}>0} \beta_{ij}\lambda_{ij} \quad \text{, with.}$$

This index sums all possible upscale movements, weighted by the "proximity" measure. The index will be higher, the larger the number of significant branches leading to products of higher income content, the higher the income levels in related products and the higher the proximity of these products. As shown in Table 2, this parametric value is larger for the product 6109 "T-Shirts", than for 6403 "footwear" and 6302 "Bed Linen". This measure, however, suffers from a basic aggregation problem, not distinguishing whether a large value is due to the presence of short distances, large increments in value or a large number of segments.

6. The open forest

A different question is the extent to which a country having comparative advantage in a number of products may help its production to move towards a new product in which still didn't develop RCA. This perspective is "inward", in the sense that we are concerned with all significant branches leading to a specific product[4].

To assess how close a given product is to the other products in the product space, Hausmann and Klinger propose a measure (path) that consists in summing up the estimated proximities in the row for that product in the matrix. Our corresponding measure will be denoted "in-path":

$$inpath_j = \sum_i \beta_{ij}$$

[4] In the Hausmann-Klinger framework, the two problems are similar, because they imposed a symmetric matrix.

Table 3 displays the 15 largest and the 15 lowest in-path values obtained with our estimated matrix[5]. As expected, the densest part of the forest (top panel) is dominated by manufactured products, meaning that they involve skills and assets that are much closer to those required by other goods and hence they facilitate the process of structural transformation. In contrast, the sparsest products (bottom panel) tend to be un-processed agricultural goods and natural resources. The reason is that these productions require very specific endowments and do not easily prepare a country to enter other goods.

If proximities are important to the process of structural transformation, then the probability of developing RCA in a particular product in the future should be affected by the ease with which the current capabilities can be adapted to the new product.

To capture this, Hausmann and Klinger (2006) propose a measure of density for each product, given by:

$$density_{j,c} = \frac{\sum_j \beta_{ij} x_{cj}}{\sum_j x_{cj}}, \ \forall \beta_{ij} > 0, \text{ where } x_{cj} = \begin{cases} 1 & if \quad RCA_{cj} > 1 \\ 0 & if \quad otherwise \end{cases}$$

We can interpret this measure as the extent to which goods in which a country already has RCA *surround* the particular new product under consideration. This variable varies between 0 and 1, with higher values indicating that the country developed capabilities in many nearby products and therefore is more likely to have RCA also in this good in the future.

Having this measure of "proximity to the new product" in mind, we can check how close and how valuable are the "unoccupied trees" that a country has. This question is explored by graphing for every country c the difference between Ln(PRODY) and Ln(EXPYc) against the inverse of the density measure for each product. Hausmann and Klinger (2006)' estimations suggest that one should observe rapid structural transformation in countries with many *upscale* goods exhibiting high density.

---

[5] Because robustness checks are not yet concluded, at this stage we are assuming that all segments are relevant. Thus, the following discussion is for illustration, only. As soon as we have run all robustness checks, we will implement the corrected "in-path" measure.

In Figure 6, such an exercise is implemented for Portugal, Spain, Netherlands, Costa Rica, Saudi Arabia and Kazakhstan.

Comparing the different figures, we first observe that Saudi-Arabia and Kazakhstan, both highly specialized in oil production, have larger distances to new goods than the remaining countries (note that in these two cases, the x-axes are scaled between 0 and 70, to allow for larger distances). Other characteristic of the open forest for these two countries is that most goods are located above the country EXPY level. This is an expected results, as the product "2709 – Petroleum, oils, crude" is a *poor country good* (75% of the 1245 products have higher income content).

Costa Rica is also an interesting case. It is a quite specialized country, with 6 products accounting for more than 40% of its. These include "Electronic integrated circuits and microassemblies", "Parts and accessories for use with machines of heading", Instruments and appliances used in medical, surgical, dental or veterinary" and some tropical products, namely bananas. This highly specialized pattern does not provide the country with capabilities enough to have an easy process of industrial transformation. The first upscale product is at a distance of 3 and most products are at distances superior to 5.

Comparing now Portugal, Spain and the Netherlands, we observe that Spain' empty forest is closer than that in Portugal and the Netherlands. Starting at an inverse density of 2, Spain has more upscale unoccupied trees in the neighborhood of products in which it has already developed comparative advantage than the other two countries. In Netherlands, upscale unoccupied trees are concentrated in a density region between 2.5 and 3.00. The empty forest for Portugal is much sparser than that of the other two countries. That is, some empty trees are moderately close to products in which Portugal has already developed comparative advantages, but most of them are relatively far away. Only few (upscale) goods lie in (inverse) densities between 2 and 3.

Figure 7 analyses the open forest for Portugal in greater detail. In particular it focus only in the upscale segment of figure 6 and it splits them in two cases: products which Portugal already exports (upper panel) but in which still didn't developed RCA; products that the country does not export at all (lower panel). The motivation for this exercise is that it should be easier for a country to develop comparative advantages in products in which it already accumulated some experience. Focusing on the upper panel, only, we see that "printed books", "cocoa beans" and "air or vacuum pumps" are all goods with higher value than the country' current EXPY and at the same time are quite surrounded by goods in which the country already developed comparative advantage. By contrast, "Floor coverings of plastics in rolls", "Cermets and articles thereof, including waste and scrap" and "Sheet piling of iron or steel" have larger income content but are less related to the county current RCA.

To continue: compute measures of open forest, at the country level

Table 1: Illustrating the forest: proximity of 6302 – "Bed Linen…. " to other products

NC - 4	6302 - Bed linen, table linen, toilet linen and kitchen linen. (Prody = 14,49)	(1) β_{ij}	(2) t	(3) Rob.	(4) β_{ij} - Rob.	(5) Prody $_j$	(6) λ_{ij}
8517	**Electrical apparatus for line telephony or line telegraphy**	0,73	4,71			25,65	0,57
4910	**Calendars of any kind, printed, including calendar blocks.**	0,70	4,29			21,27	0,38
3305	**Preparations for use on the hair**	0,61	4,15			20,35	0,34
6307	**Other made up articles, including dress patterns.**	0,67	4,06			14,08	-0,03
9018	**Instruments and appliances used in medical, surgical, dental or veterinary ...**	0,67	4,06			25,94	0,58
9615	**Combs, hair-slides and the like; hairpins, curling pins, curling grips**	0,75	4,00			22,07	0,42
7215	**Other bars and rods of iron or non-alloy steel.**	0,74	3,92			15,24	0,05
3909	**Amino-resins, phenolic resins and polyurethanes, in primary forms.**	0,74	3,92			24,21	0,51
8214	**Other articles of cutlery (for example, hair clippers, butchers' or kitchen ...**	0,74	3,92			20,46	0,35
2005	**Other vegetables prepared or preserved**	0,59	3,86			14,85	0,02
3920	**Other plates, sheets, film, foil and strip, of plastics**	0,59	3,86			27,86	0,65
4821	**Paper or paperboard lables of all kinds, whether or not printed.**	0,59	3,86			20,32	0,34
8441	**Other machinery for making up paper pulp, paper or paperboard**	0,59	3,86			28,23	0,67
8504	**Electrical transformers, static converters (for example, rectifiers) and in ...**	0,59	3,86			22,19	0,43
8536	**Electrical apparatus for switching or protecting electrical circuits, or fo ...**	0,59	3,86			22,20	0,43
4015	**Articles of apparel and clothing accessories**	0,65	3,85			13,08	-0,10
4817	**Envelopes, letter cards, plain postcards and correspondence cards, of paper ...**	0,65	3,85			16,78	0,15
3215	**Printing ink, writing or drawing ink**	0,65	3,85			29,94	0,73
6110	**Jerseys, pullovers, cardigans, waist-coats and similar articles, knitted or ...**	0,65	3,85			12,46	-0,15
6117	**Other made up clothing accessories, knitted or crocheted**	0,65	3,85			21,39	0,39
3002	**Human blood; animal blood prepared for therapeutic uses**	0,73	3,84			32,63	0,81
9505	**Festive, carnival or other entertainment articles**	0,73	3,84			20,53	0,35
4823	**Other paper, paperboard, cellulose wadding and webs of cellulose fibres**	0,51	3,81			20,83	0,36
8443	**Printing machinery used for printing by means of the printing type, blocks**	0,51	3,81			28,90	0,69
6112	**Track suits, ski suits and swimwear, knitted or crocheted.**	0,72	3,76			11,92	-0,19
6115	**Panty hose, tights, stockings, socks and other hosiery**	0,72	3,76			14,65	0,01
6804	**Millstones, grindstones, grinding wheels and the like**	0,72	3,76			32,67	0,81
7608	**Aluminium tubes and pipes.**	0,72	3,76			21,68	0,40
8211	**Knives with cutting blades, serrated or not (including pruning knives)**	0,72	3,76			25,28	0,56
8306	**Bells, gongs and the like, non-electric, of base metal; statuettes and othe ...**	0,72	3,76			22,73	0,45
8302	**Base metal mountings, fittings and similar articles suitable for furniture**	0,64	3,75			19,99	0,32
8525	**Transmission apparatus for radio-telephony, radio-broadcasting**	0,64	3,75			25,36	0,56
6211	**Track suits, ski suits and swimwear; other garments.**	0,64	3,75			10,60	-0,31
2309	**Preparations of a kind used in animal feeding.**	0,57	3,73			20,11	0,33
5208	**Woven fabrics of cotton, containing 85 % or more by weight of cotton**	0,57	3,73			12,49	-0,15
6301	**Blankets and travelling rugs.**	0,57	3,73			11,95	-0,19
602	**Other live plants (including their roots), cuttings and slips**	0,70	3,69			10,47	-0,32
5515	**Other woven fabrics of synthetic staple fibres.**	0,70	3,69			15,59	0,07
7418	**Table, kitchen or other household articles and parts thereof, of copper**	0,70	3,69			16,90	0,15
6201	**Men's or boys' overcoats, car-coats, capes, cloaks, anoraks**	0,63	3,65			11,90	-0,20
6206	**Women's or girls' blouses, shirts and shirt-blouses.**	0,63	3,65			8,99	-0,48
9503	**Other toys; reduced-size (scale) models and similar recreational models**	0,63	3,65			21,63	0,40
3210	**Other paints and varnishes**	0,63	3,65			17,70	0,20
5903	**Textile fabrics impregnated, coated, covered or laminated with plastics**	0,63	3,65			24,05	0,51
…	**…**	…	…			…	…
…	**…**	…	…			…	…
…	**…**	…	…			…	…
…	**…**	…	…			…	…
…	**…**	…	…			…	…
703	**Onions, shallots, garlic, leeks and other alliaceous vegetables**	-0,07	-0,42			4,40	-1,19
1203	**Copra.**	-0,04	-0,45			5,35	-1,00
2524	**Asbestos.**	-0,08	-0,66			11,58	-0,22
4105	**Tanned or crust skins of sheep or lambs, without wool on**	-0,13	-0,74			5,14	-1,04
2403	**Other manufactured tobacco and tobacco substitutes**	-0,13	-0,82			31,94	0,79
2709	**Petroleum oils, crude**	-0,16	-0,95			14,33	-0,01
2612	**Uranium or thorium ores and concentrates.**	-0,10	-1,57			1,26	-2,44
			N. signif.				V_i
	Summary		**625**			**-**	**104,11**

Table 2 – Summary statistics

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		RCA	Prody	Prody Rank (%)	ni	% Inferior connections	m_k			V_i
							0%	10%	50%	
6109	**T-shirts, singlets and other vests, knitted or crocheted.**	8,19	10,90	83,8	602	11%	533	516	385	175,60
6302	**Bed linen, table linen, toilet linen and kitchen linen.**	16,83	14,49	73,3	625	21%	492	448	184	104,11
6403	**Footwear with outer soles of rubber, plastics, leather or composition leath …**	8,60	12,99	78,5	922	17%	766	733	429	158,28

Table 3 – The fifteen goods in the densest part of the forest and the fifteen goods in the least dense part of the forest

		In Path
3812	Prepared rubber accelerators	543,90
1518	Animal or vegetable fats and oils	537,12
2916	Unsaturated acyclic monocarboxylic acids	531,32
5901	Textile fabrics coated with gum or amylaceous substances	530,47
8457	Machining centres, unit construction machines (single station)	529,85
8715	Baby carriages and parts thereof.	528,70
4006	Other forms (for example, rods, tubes and profile shapes) and articles	526,44
2909	Ethers, ether-alcohols, ether-phenols, ether-alcohol-phenols	523,20
7414	Cloth (including endless bands), grill and netting, of copper wire	520,47
3703	Photographic paper, paperboard and textiles, sensitised, unexposed.	519,81
9001	Optical fibres and optical fibre bundles	518,70
5909	Textile hosepiping and similar textile tubing	518,49
9107	Time switches with clock or watch movement or with synchronous motor.	518,00
2821	Iron oxides and hydroxides	517,90
5206	Cotton yarn (other than sewing thread), containing less than 85 % by weight ...	517,74
…	…	…
7309	Reservoirs, tanks, vats and similar containers for any material	29,27
8431	Parts suitable for use principally with the machinery of headings 84.25 to ...	27,65
7323	Table, kitchen or other household articles and parts thereof, of iron or st ...	27,54
5203	Cotton, carded or combed.	25,97
2609	Tin ores and concentrates.	19,95
7308	Structures (excluding prefabricated buildings of heading 94.06)	15,00
8413	Pumps for liquids, whether or not fitted with a measuring device; liquid el ...	11,84
1201	Soya beans, whether or not broken.	-0,94
1203	Copra.	-6,21
5201	Cotton, not carded or combed.	-10,95
1202	Ground-nuts, not roasted or otherwise cooked	-11,28
2612	Uranium or thorium ores and concentrates.	-14,30
714	Manioc, arrowroot, sweet potatoes and similar roots	-20,98
2709	Petroleum oils, crude	-27,58
1801	Cocoa beans, whole or broken, raw or roasted.	-42,27

Figure 1: EXPY and Per capita incomes

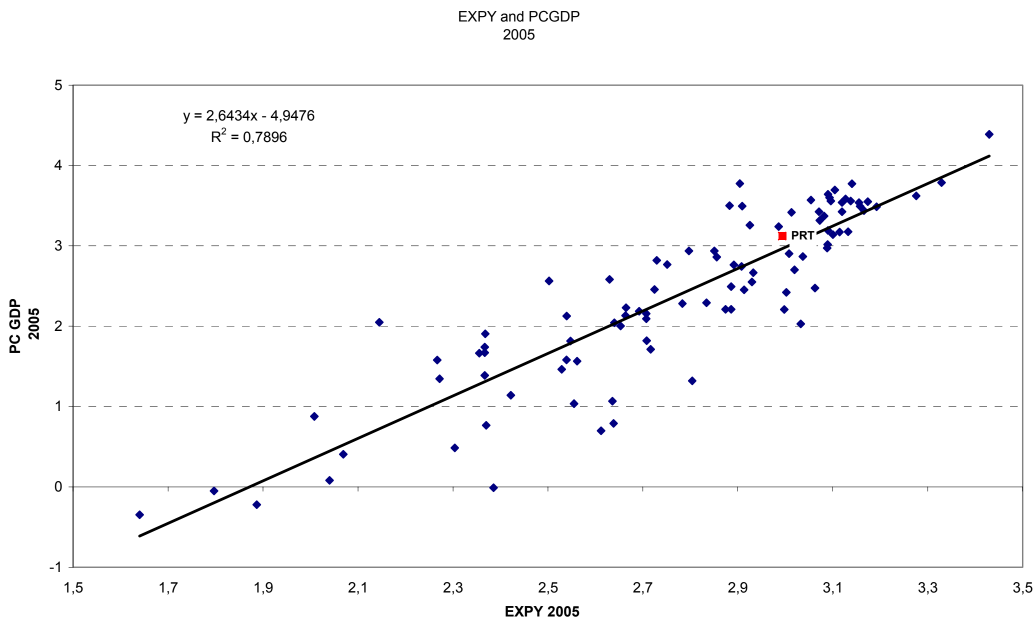


Figure 2: RCA and PRODY, Portugal 2005

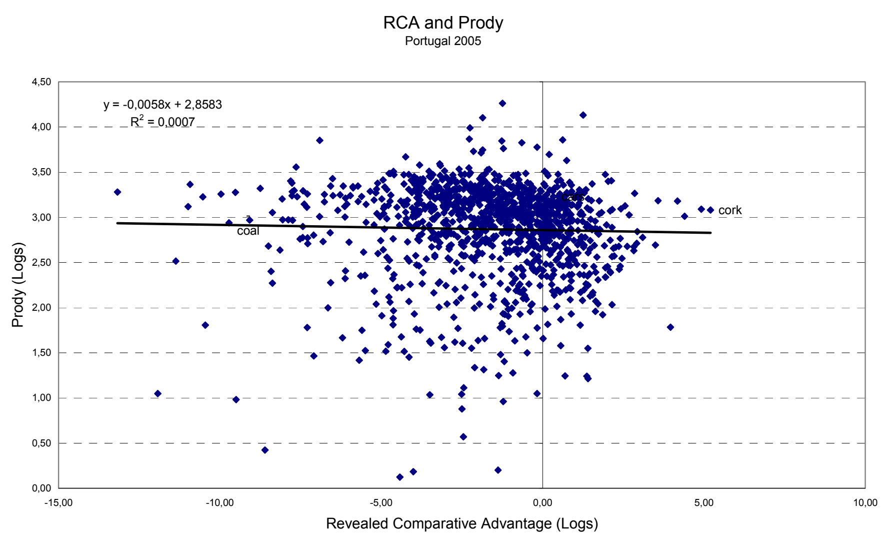


Figure 3: RCA and PRODY, Portugal 2005 (products weighting more than 0,2%)

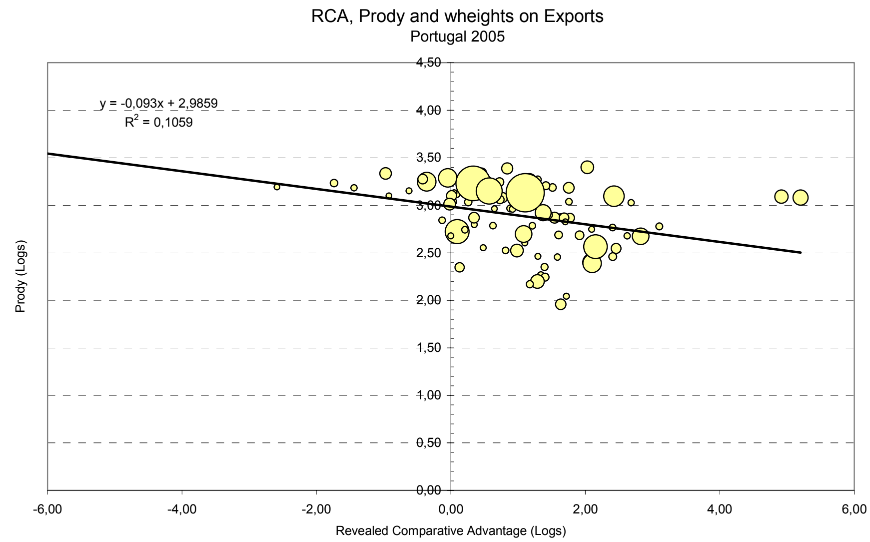


Figure 4: RCA and PRODY, Portugal 2005 (products with positive RCA, only)

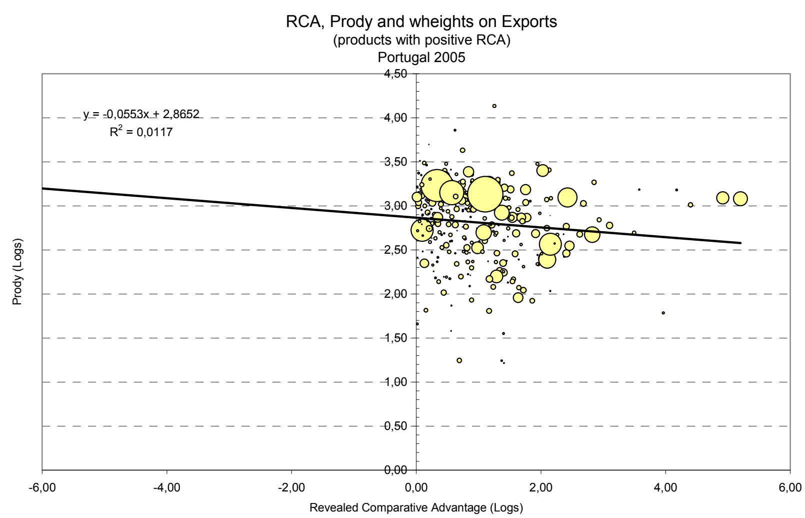


Figure 5 – The number of significant segments leading to k% increase in value

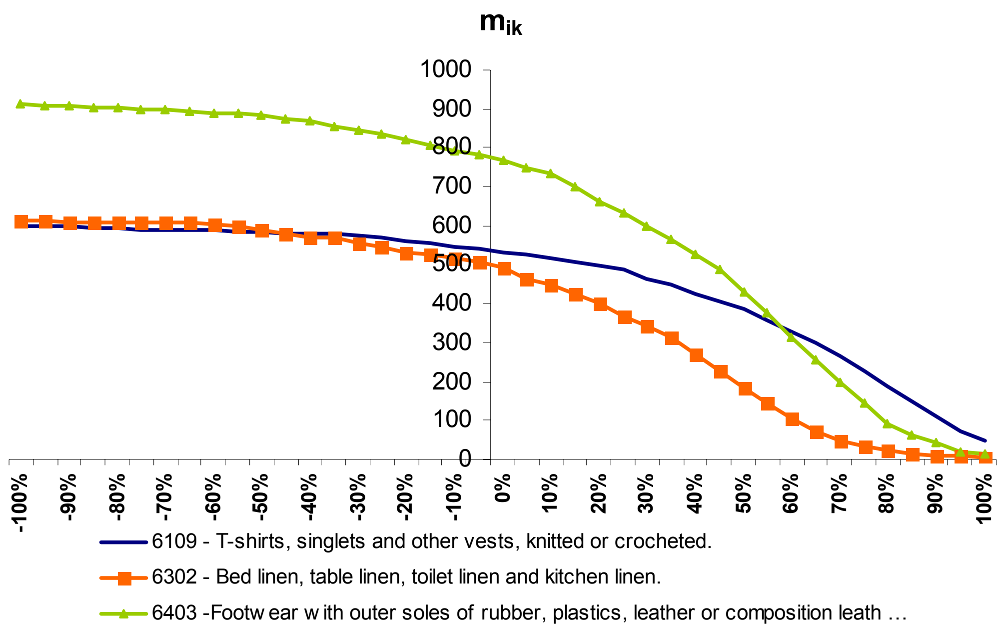


Figure 6 – Distances versus values (T-shirts)

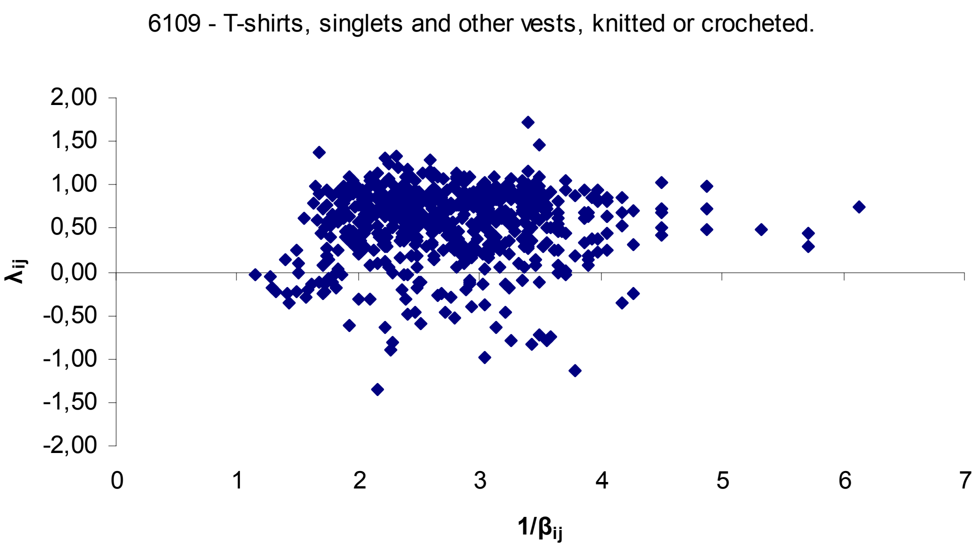
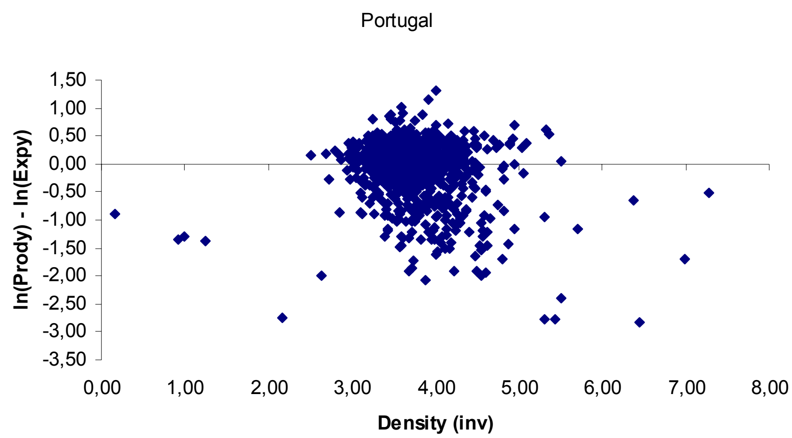
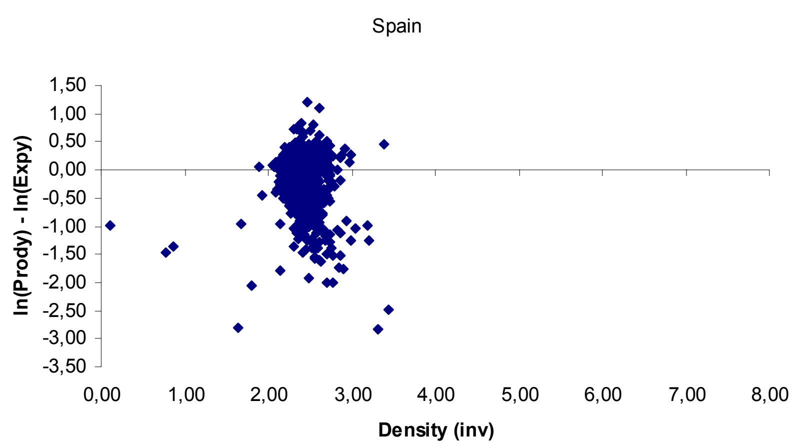
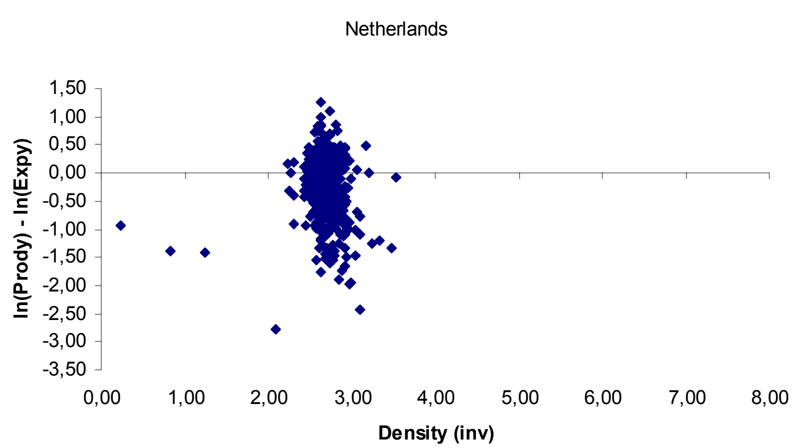


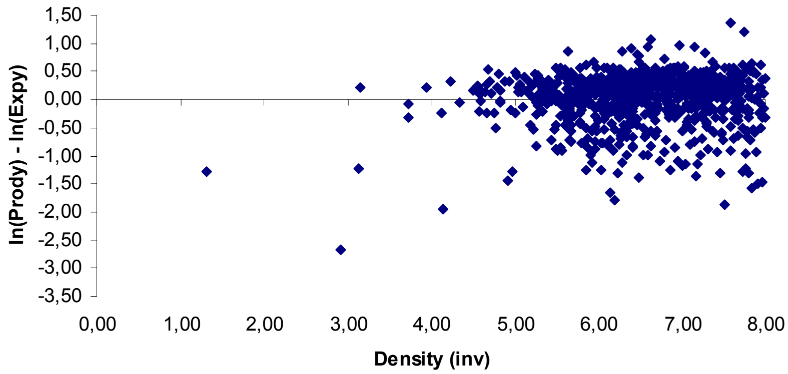
Figure 7 – Visual representation of the “open forest” by country, 2005



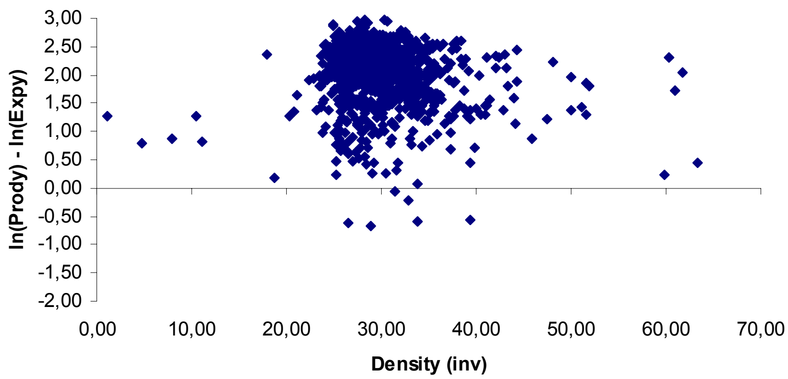




Costa Rica



Saudi Arabia



Kazakhstan

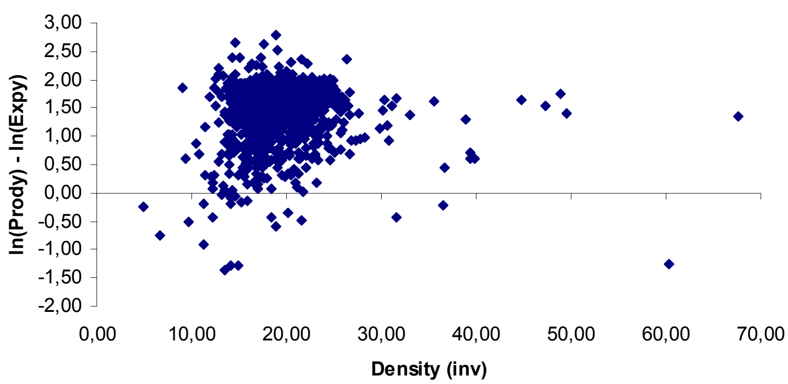
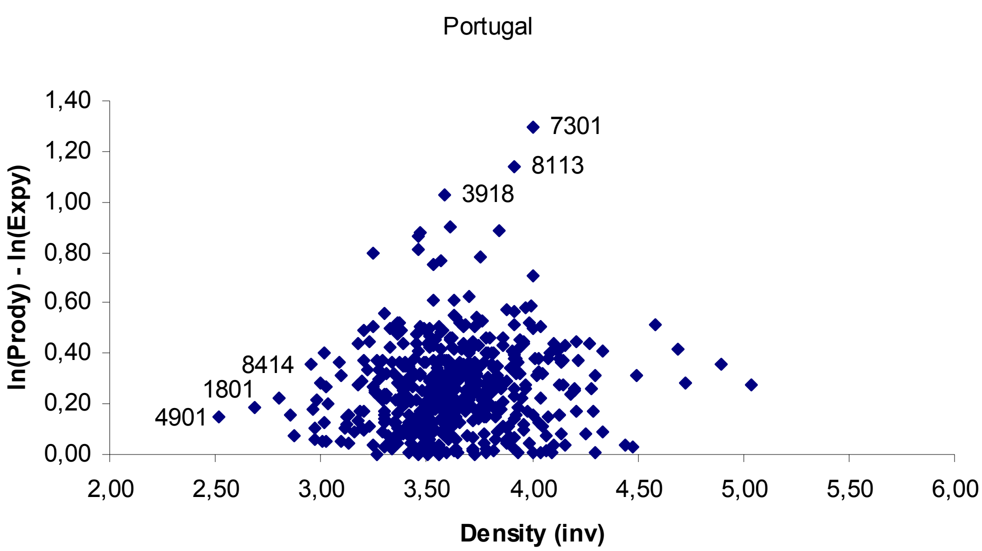
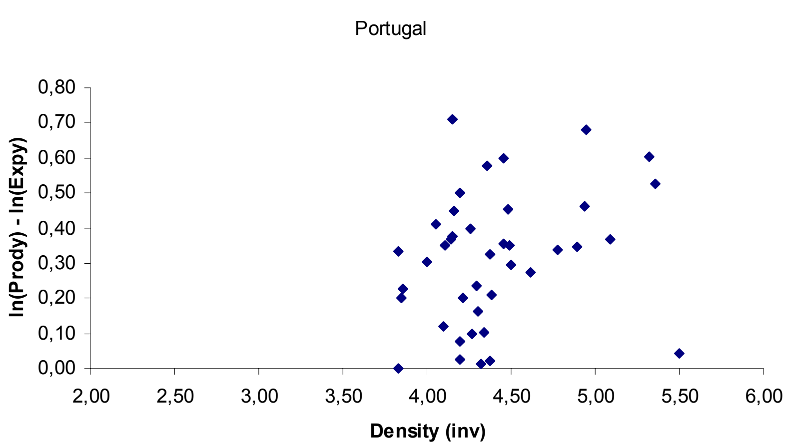


Figure 8: The open forest for Portugal: exported products and non-exported products



1801 Cocoa beans, whole or broken, raw or roasted.
3918 Floor coverings of plastics in rolls
4901 Printed books, brochures, leaflets and similar printed matter
7301 Sheet piling of iron or steel
8113 Cermets and articles thereof, including waste and scrap.
8414 Air or vacuum pumps, air or other gas compressors and fans



References

Dalum, B. K. and B. Verspagen, (1999), "Does specialization matter for growth?", Industrial and Corporate Change, vol. 8 (2), 267-288.

Feenstra, R. C. and A. K. Rose, (2000), "Putting things in order: patterns of trade dynamics and growth", The review of Economics and Statistics, vol. 82(3), 369-382.

Grossman, G. M. and E. Helpman, (1991), "Innovation and growth in the global economy", MIT Press, Cambridge.

Hall, R., and C. Jones, 1999. "Why do some countries produce so much more output per worker than others?", The Quarterly Journal of Economics 114 (1), 83-116

Hausmann, R. J. Hwang and D. Rodrik, (2005), "What you export matters", NBER working paper 11905.

Hausmann, R. and B. Klinger, (2006), "Structural transformation and patterns of comparative advantage in the product space", Center for International Development, Harvard University.

Hausmann, R. and D. Rodrik, (2006), "Doomed to choose: industrial policy as predicament", draft.

Jovanovic, B., Nyarko, Y. (1996). “Learning by doing and the choice of technology”. Econometrica 64 (6), 1299-1310.

Levine, R. and D. Renelt (1992), “A sensitivity analysis of cross-country growth regressions”, American Economic Review 82(4), 942-63.

Kaldor, N, (1966), "Causes of the Slow Rate of Growth of the United Kingdom", Cambridge University Press.

Lucas, R., 1990, “Why doesn’t capital flow from rich countries to poor countries?”, *American Economic Review* 80, 92-96.

Murphy, K., Shleifer, A, Vishny, R., 1989. Industrialization and the big push, Journal o Politica Economy 97, 1003-1026.

Sala-i-Martin, X. (1996), “I just run four million regressions”, NBER Working Paper No 6252.

Thirwall, A. P., (1979), "The Balance of Payments constraint as an explanation of international growth differences", Banca Nazionale del Lavoro Quarterly Review, vol. 3, 245-253.

Appendix 1 – Estimated proximities

Using UN – COMTRADE 2005 database, we assess whether having revealed comparative advantage (RCA) in a product affects the probability of having RCA in another product. In other words, we wonder whether the probability of having RCA in product *Y* is conditional on having RCA in product *X*. Let y_i be a dummy variable equal to 1 if country *i* has RCA in product *Y* and 0 otherwise and, similarly, let x_i be a dummy variable equal to 1 if country *i* has RCA in product *X* and 0 otherwise. Therefore, we are interested in estimating a model like:

$$P(y=1|x)=G(\beta_0+\beta_1 x) \text{ being } G(.) \text{ a distribution function.}$$

In such a model if $\beta_1=0$, the probability of having RCA in product *Y* is not dependent on having RCA in product *X* and, thus, we will consider those products to be unrelated.

Actually, if $\beta_1=0$, an estimate of $P(y=1|x)=G(\beta_0)$ will be identical to the percentage of countries having RCA in product *Y*.

Since we use NC-4 classification, encompassing 1245 products, we can have a total of 1245x1245 possible relations between the RCA of any pair of products. Therefore, it is possible to construct a 1244x1244 non-symmetrical matrix where in each cell we can record the estimated marginal effect on $P(y=1|x)$ given that there is RCA in product *X*. Those marginal effects are given by $G(\hat{\beta}_0+\hat{\beta}_1)-G(\hat{\beta}_0)$ whenever $\beta_1 \neq 0$.

We assume *G*(.) to be the standard normal cumulative distribution function:

$$G(z)=\Phi(z)\equiv\int_{-\infty}^{z}\phi(v)dv \text{ with } \phi(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

Thus, we estimate PROBIT models for each possible pair of products using the 96 available observations (i.e. the number of countries in the database) on each pair of dummy variables (x,y) and report the marginal effect on $P(y=1|x)$ of having RCA in product *X* whenever $\beta_1 \neq 0$.

We interpret high values of the marginal effect based on the PROBIT estimation as a greater technologic *proximity* between the two products.

Appendix 2 – Robustness check

For robustness check, given a pair (x, y), we shall say that $\beta_1 \neq 0$ if we reject the null $H_0 : \beta_1 = 0$ in 90% of the cases where is included an additional dummy variable recording existence of RCA in some other product that is not *X* or Y. Let that third product be *Z* so that we estimate:

$P(y = 1|x) = G(\beta_0 + \beta_1 x + \beta_2 z)$ with *G* equal to the normal CDF

and check the rejection of $H_0 : \beta_1 = 0$ in 90% of the 1243 possible products *Z*.

After such a robustness check, we can safely say that having RCA in product *X* will affect the probability of having RCA in product *Y* and that is an effect of the technological relation between those two products and not due to an indirect effect through a third product (*Z*) absent from the relation.

If, however, the model does not pass the test described above, we conclude that the probability of having RCA in product *Y* is not conditional on having RCA in product *X* and hence $P(y = 1|x) = G(\beta_0)$, which will simply be identical to the percentage of countries having RCA in product *Y* as already mentioned above.

Appendix 3: Definitions of PRODY and EXPY

For each product I, the Prody index is computed as:

$$PRODY_i = \sum_{c=1}^{C} \sigma_{ci} Y_c \text{ , where } \sigma_{ci} = \frac{RCA_{ci}}{\sum_c RCA_{ci}} \text{, } RCA_{ci} = \frac{X_{ic}/X_c}{X_i/X} \text{.}$$

Y_C is real GDP per capita in the c-th country (c=1,…….,C) exporting good *i*, and the weights σ_{ci} normalize the Balassa index of Revealed Comparative Advantage (RCA) of the c-country with respect to all the countries exporting in the same sector.

The income content of a country exports, EXPY, is computed, for each country, as:

$$EXPY_c = \sum_i s_i PRODY_i$$

Where $s_i = \frac{X_{ic}}{X_c}$ is the share of product i in te exports of country c.