UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

BE CAREFUL! A SHORT NOTE ON A POSSIBLE BIAS IN (TRADE) STRUCTURAL CHANGE ANALYSIS

MASSIMO TAMBERI

QUADERNO DI RICERCA n. 341

May 2010

Collana curata da:
Massimo Tamberi

BE CAREFUL! A SHORT NOTE ON A POSSIBLE BIAS IN (TRADE) STRUCTURAL CHANGE ANALYSIS[1]

Massimo Tamberi[2]

ABSTRACT

I was accustomed to think that the world, in this of "modern economic growth", is becoming less specialized: the invention of new goods is, only partially, a "Schumpeterian" process, in the sense that new goods sometimes replace old ones, but it can also happens that these new goods simply are added to the old ones. Also in modern theoretical literature emerges (at least) the idea that producers use an increasing variety of intermediate goods and that consumers are likely better with more variety of goods in their hands. The process of change in the produced/consumed goods is one of the aspects of the broader spectrum of economic structural change that accompanies economic growth, "structural change" being one of the basic stylized facts of growth according to the Nobel Lecture of Simon Kuznets.

In this paper, I will suggest that sector disaggregated data, necessary to study structural change, contain a bias that "hides" this process and causes a drift. By using trade data (because of their higher sector disaggregation richness) I'll first show that there is a tendency for a steady increase in sector concentration. Next, I will argue that this is due to the impossibility to properly register product innovation and finally, through a very rough model and an empirical example of two countries, I will also suggest that this reflects differently in developing and developed countries.

JEL: F43, O14, O40

Keywords: International Trade, Structural Change, Economic Growth

[1] I would like to thank Aleksandra Parteka for useful comments

[2] *Università Politecnica delle Marche*
Dipartimento di Economia
P.le Martelli, 8
60121 Ancona (Italy)
e-mail: m.tamberi@univpm.it

A PUZZLING EMPIRICAL EVIDENCE: IS THERE A DRIFT IN SECTORAL TRADE (AND PRODUCTION) DATA?

Some recent models of economic growth are based on the idea that economic growth depends on an increase in the variety of goods (for example Grossman and Helpman, 1991)[3]. In other literature traditions, as in NEG models, utility functions try to catch the "love for variety" of consumers (Fujita, Krugman, Venables, 2000): both the supply and demand sides are involved.

In short, authors of relevant theoretical contributions convey the idea that economic growth is somehow connected to an increasing variety of goods in the market. This connection to economic growth makes structural change an intrinsic dynamic phenomenon. These themes are not completely new. As mentioned in the abstract, Simon Kuznets has discussed in length about them many years ago (see his contributions of 1972 and 1973 for a synthetic approach).

From an empirical point of view structural change can be referred to a very broad and disaggregated framework: here, I limit my attention only to sector and product change composition. In this respect I will use a detailed sector definition, while most past works mainly focused on the shift of production and employment only among macro-sectors (i.e. agriculture, industry, services)[4].

We could observe that usually, from an empirical point of view, we do not have data on products, but mainly on sectors. Nevertheless, trade data could be highly product-specific: I will use them at different sector disaggregation levels to show some possible limits in structural change analysis. It is implicit in the analysis that the highlighted problems could also be referred to production data.

The main point of this note is that structural change analysis clashes against a intrinsic limit of sector statistics.

To have a first indication, it would be useful to throw a glance at figure 1, where an index of concentration of the sector structure of world export is represented (I used a normalized Herfindhal index[5]). I excluded petroleum (code 3310) from data (COMTRADE, 4 digit[6]) whose price waves in the seventies are too much disturbing (I will come back to this point in appendix notes). One could

---

[3] Instead, in a Schumpeterian world there is a perfect process of creative destruction (Aghion P, Howitt P., 1992): new goods completely replace old ones. As a consequence, the "number" of goods in the market is constant but we observe a growth in their good quality.

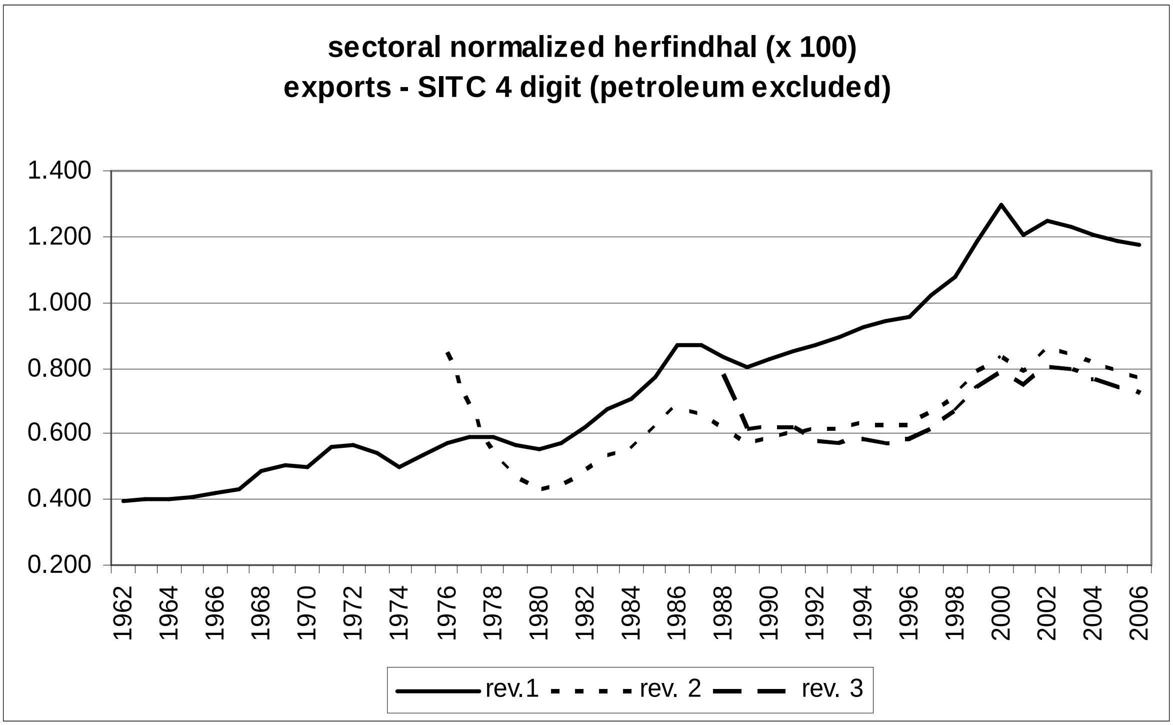
[4] The main exception being studies on sector composition of the industrial sector.

[5] Normalized Herfindahl is $NH[0,1]=(nH-1)/(n-1)$ (where H is the classic Herfindahl index). It is known that, since Herfindhal index H is related to the coefficient of variation CV, according to $H=(CV^2+1)/n$, its normalized version is also $NH=CVN^2$ (CVN being the normalized coefficient of variation). I use the normalized version of the index because I will compare datasets with different number of sectors.

[6] I will use a 4 digit disaggregation scheme throughout the paper: a 4 digit level, on the one hand, is sufficient to evidence clear results, and, on the other hand, more easily tractable. Nevertheless, I will occasionally provide comparisons with different disaggregation schemes (as in table 2).

go farther and work without commodities or primary products, but by excluding petroleum, it is sufficient to get a “smooth” figure without strong perturbations[7].

Fig. 1



Looking at the rev. 1 line, the result is very clear: world trade seems to have become increasingly concentrated. We get the same indication whatever the revision of trade data we use: with rev. 2 and rev. 3 there is a decrease in the first years, nevertheless, also in their case, the “long-run” growth seems confirmed.

A possible interpretation of the previous evidence could be that people have access to an ever more decreasing variety of goods at world level, at least if we measure them with (detailed) sector shares (obviously, the production structure could move in a different direction). Paradoxically, this should have happened in a period of great expansion of world integration[8]. After the above-mentioned considerations on theoretical insights, and also in the light of our personal lives[9], this interpretation does not sound too serious.

In fig. 1, it is possible to appreciate the smoothness of the curve: the outlined phenomenon becomes evident early (very soon for rev. 1, after a few years for rev.2 and 3) and steadily proceeds in time.

[7] The general conclusion would remain the same even if we included petroleum, but the figure would be less “attractive”, in the sense that short term perturbations would be strongly visible.

[8] World trade is supposed to increase welfare because it raises the variety of goods available to people. For a well-known empirical work, see Broda and Weinstein (2006).

[9] Many people would remember when the first refrigerators appeared in the market, and the first TV sets. There were not cellular phones, computers, credit cards, on-line shops, satellites over our heads, etc. etc.: many of these goods did not substitute previous (non-existing) similar goods.

A possible more detailed view of what happened is available in table 2, where a transition matrix of sector shares of exports (beginning-final year) is showed.

Table 1

Sector export transition matrix, initial, final and ergodicdistribution

SITC rev. 1 - 1962-2006

	first quartile	second quartile	third quartile	fourth quartile
first quartile	0.953	0.036	0.007	0.004
second quartile	0.735	0.147	0.088	0.029
third quartile	0.367	0.469	0.143	0.020
fourth quartile	0.318	0.318	0.227	0.136
Initial distribution*	0.721 (449)	0.164 (102)	0.079 (49)	0.035 (22)
Final distribution*	0.849 (529)	0.098 (61)	0.037 (23)	0.014 (9)

* Number of sectors in parenthesis

Clearly, we are in front of a polarizing phenomenon, with a lot sectors going under very small dimensions and a few growing bigger. In short:

- there is a high persistence of only small sectors (value in the first cell)
- a part from the previous case, values on the diagonal are not the highest, highlighting strong dynamics; values generally decrease from left to right (one exception)
- a comparison between the initial (1962) and final (2006) distributions shows that there is a falling share of sectors in the upper tail of the distribution and a growing share in the lower tail.

I would like to stress the point that sectors whose size is bigger than 1/100 of world export move from 12 to 21, while sectors with a size lower than 1/10000 move form 67 to 189.

As we will see later, some of the "small sectors" in 2006 do not report any export values, i.e., they are "non active sectors" (but which had positive export values in 1962).

Finally, on this point, note that four of the first six biggest sectors in 2006 are hi-tech sectors (7149 - office machines n.e.s (not elsewhere specified); 7143 - statistical machines-cards or tapes; 7249 - telecommunications equipment n.e.s.; 7293 - thermionic valves and tubes, transitors, etc.): only one of them was in the top ten rankings in 1962. Moreover, two of them are "n.e.s." sectors, a possible indication that new and not properly classifiable goods had been introduced in those residual categories.

I also propose another way to look at this structural and continuous change in the sector structure of world trade statistics. In the following lines, I have presented simple estimations of a power law which can be formulated in the following way:

$$q_i \approx \beta r_i^{-\alpha}$$

where q is the 'size' of the occurrence/event and r its rank; β and α are unknown parameters.
A power law like this one is known to represent a large set of very different phenomena, and it is sometimes seen as an example of self-organizing processes. It is able to approximate data used in physical and social sciences (and recently to describe traffic to internet pages). In economics, it can be found in economic geography (size of cities, with the name of Zipf law), in income distribution (where it is better known as the Pareto law), in industrial economics (firm size distribution) . I am not interested in the doubtful meaning of this law, if there is one, but in the possible changes of its empirical parameters.
The linearization of this power law gives us a simple functional form to be tested:

$$\ln(q_i) \approx \ln(\beta) - \alpha \ln(r_i)$$

In the context of this paper, q_i are sector sizes of exports (as reported in COMTRADE) while r_i are their rank positions. I run simple cross-countries OLS estimations[10] of this equation every 4 years (to save time for needless calculations), over the entire period under analysis (i.e.: 1962, 1966, …, 2006), and I did it for all three available statistical revisions of trade data. I am interested in the evolution, if there is one, of parameter α, which is an index of the steepness of the function. Coefficients α, and estimations in general, are highly significant (t and F test) in all regressions.
You can find two different presentations of the outcome: in fig. 2, you can see the time trend of the coefficients, while in fig. 3, I have presented the same data but “shifting to the left” all curves in order to appreciate changes after a given number of years, after the initial year of each statistical revision.
In fig. 2, estimated coefficients have a very clear tendency to steadily decrease, independently of the revision adopted. There is a drift in the data to be explained.
In some way, results replicate what we saw in fig. 1: downward curves in fig. 2 mean that relative differences among sector sizes tend to increase in time. In order to have a quantitative indication of this increment, consider that:

$$ln(q_i/q_{i-1}) = \alpha\, ln(r_{i-1}/r_i).$$

[10] For a discussion on the estimation problems of the power law, limited to city distribution, see Henderson and Thisse (2004), pp. 2346-2350.

With coefficient α near the unity, as in the early years of the analysis, the first sector in the ranking has a size that is approximately twice the second. With a coefficient near two, as in the final years (rev.1), this ratio becomes four.

A second relevant point is that new revisions of the dataset increase the level of the coefficients: they seem to correct the bias.

Fig.2

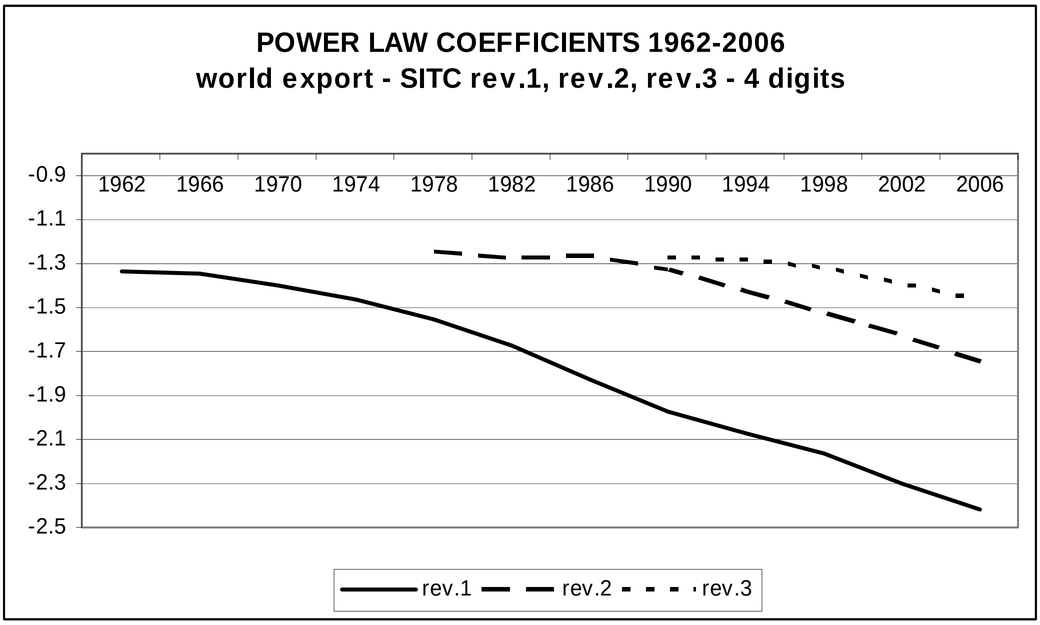
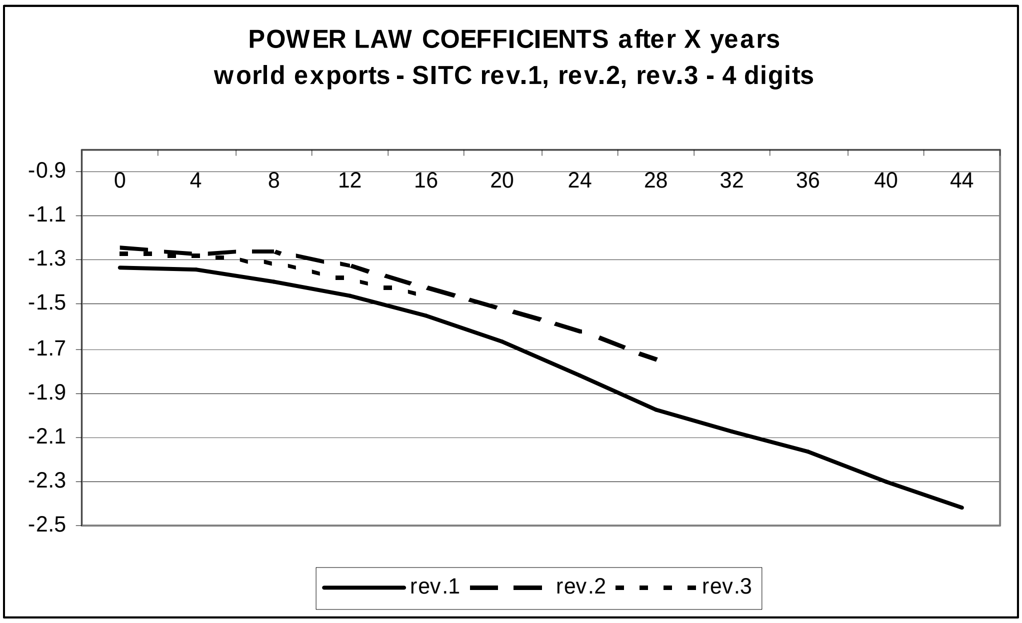


Fig. 3



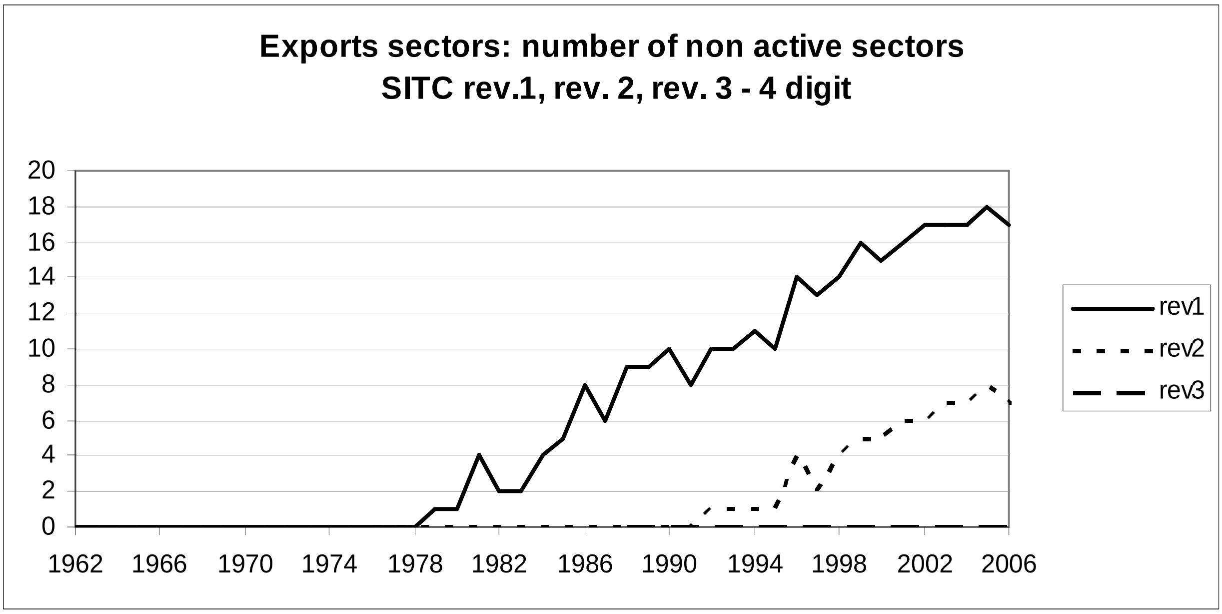
Finally, fig. 3 points out that when a new revision of data becomes active, coefficients are sufficiently, even if not precisely, near unity (according to which, as said, the second sector has a

"size" of about ½ of the first, and so on) and these changes seem to accelerate after 8-10 years from the beginning.
I would especially like to draw your attention to the first conclusion. As in the case of the concentration index, we see that there is an evident time trend in the transformation of the sector structure of trade data, and this evolution, surprisingly, consists in a rise of sector concentration.

WHAT DRIVES CONCENTRATION

In my opinion, the evidenced phenomenon has nothing to do with economics but with statistics: the positive (negative) trend shown in figure 1 (2) mainly depends on a statistical bias due to our inability to register product innovation. Moreover, in the following sections, I suggest that this problem has impact on developing economies different from developed economies.
To better highlight this issue, I will follow Cadot et al. (2007) and somehow also Amurgo-Pacheco and Pierola (2008). In figure 4, I have shown the total number of sectors for which there are no reported exports in COMTRADE for the period 1962-2006 and for the whole world economy. Throughout this paper I will call sectors with no export value "non active sectors": they are a complement to the "active lines" of Cadot et al. (cit.).

Fig. 4



The value, or rather, the evolution of these statistics is clearly not random: even if the values show a very low level of non active sectors, it is evident that there is a growing trend, i.e., there is a slow but steady growth in the number of sectors that "disappear"[11]. We could again conclude that the process of development and globalization, at least according to trade data, (strictly) consists in a decreasing number of goods available to people.

First of all, note that this process is more evident when a long time span is used, and it becomes visible only after a certain number of years have passed from the first year of a specified system of classification: with SITC rev. 1, starting in 1962, we find that the phenomenon appears for the first time 18 years later, i.e., in 1980. i.e. after that year the number of non active sectors, as seen, starts growing. If we use SITC rev. 2, starting in 1976, we find that the same phenomenon appears for the first time 16 years later, i.e., in 1992. Finally, if we use SITC rev. 3, starting in 1988, we do not observe the phenomenon at all (after 18 years). Also note that the number of non active sectors is always lower the more recent is the classification system we have used (rev.1>rev.2>rev.3)[12]. An analogous conclusion also holds for the concentration index and for the coefficients of the estimated power law.

Another useful observation is that the share of non active sectors depends on data aggregation. This share does not exist for classification less disaggregated than 4 digits, while it is higher for more disaggregated classification. This point is summarized in table 2. You can see that the appearance and time trend of the share of non active sectors depends on the scheme of disaggregation. In particular, at five digits, it is a significant phenomenon since it accounts for about 13% of the total number sectors in 2006. Here, I would like to underline that we find the same kind of positive trend in the growth of non active sectors at higher disaggregation schemes, but at the 5 digit level it is sensibly stronger than at the 4-digit level, while in the 3 digit case it is not visible.

Table 2

Percentage of non active sectors in trade statistics (SITC rev.1)

various sector aggregation

	1962	1985	2006
3 digit	0	0	0
4 digit	0	0.8	2.7
5 digit	0	3.1	13.4

Obviously, this possibly have an impact on indexes of sector concentration and, indeed, relative differences in concentration indexes between the first and the final year are more accentuated when we use finer disaggregation (evidence not showed here).

[11] Look also at table 2 and relative comments for a deepening.

[12] The total number of sectors characterising SITC revisions is growing in time: at 4 digits, rev. 1 has 623 sectors, rev.2 has 783, and rev.3 has 1031. I do not know if this sectors-number increase somehow reflects an implicit or explicit statisticians' and economists' perception that the economy begins to progressively have an increasing variety of goods.

As figures 4 and table 2 clearly show, there is a steady trend in the number of non active sectors in export data at world level. This means that zeros are not casually distributed in the matrix but that they are correlated to time, and this could constitute a relevant problem in the analysis of structural transformation, and, in fact, it does, as I will show.

My proposal, here, is to interpret the positive trend of the non active sectors as the "destructive" side of the Schumpeterian creative-destruction process that characterizes the innovation activity even if, as I will argue later, this evidence about the rising number of non active sectors is only the tip of the iceberg.

The problem, based on my interpretation, is that we cannot appreciate the creative side of innovations: we can see dead (i.e. non active) sectors, but we cannot see new sectors! Why we cannot see new sectors is because we have to work within a framework, that is, with a number of sectors, determined at the beginning of the period of analysis, on the basis of what the economy was in that period. Clearly, SITC rev. 1 reflects the "relevant" sectors at the beginning of the sixties. If, in the following years, new products were invented, they do not had their own "cell" to be collocated, but they had to be added to some of the already defined sectors. Perhaps, there are not big problems if the new goods completely replace old ones also in our statistical definition: in this case the process of creative destruction is also replicated in our statistics (I leave aside questions connected to the monetary evaluation of goods of different quality). But this is not necessarily always true. Let us suppose that a tool used by farmers, say a scythe, is replaced in the economy by a kind of machinery, say a threshing-machine. This latter will probably be registered with a digit different from the first. Let us say that the threshing-machine "goes" into mechanics, while the scythe disappears from "metal products". This, as a consequence, could lead to a growing concentration. In particular, we could have a possible extreme case with a sector with no export (metal product) and one with an increased share (mechanics)

This problem is even more evident when new goods are radically new and do not substitute old ones: "where" to classify them[13]?

Finally, as previous considerations should have already clarified, the process is also at work when (if) we do not observe the appearance of "zeros" in our sector matrix: substitution between old and new goods is not instantaneous and it requires time for several reasons (and also because in our statistical datasets we have more sectors than products). Steady trends showed by the concentration index and power law highlight this point.

[13] A common procedure is to add them to "residual" categories in statistical tables.

So far, I discussed the evidence of a rising sector/product concentration at world level and for the whole economy. Obviously, if we worked with a subset of sectors or with specific countries we may have different dynamics.

CREATIVE-DESTRUCTION, LEADERS AND FOLLOWERS: A SIMPLE SKETCH AND MEASURE OF SECTOR ECONOMIC STRUCTURE

In order to better highlight the question, I have outlined a very rough model based on Lucas (2000) and Krugman's (1979) models. These two models are close in spirit with a leader and follower economy. The first grows on the basis of some kind of AK model (or, more generally, a model with constant per capita growth), while there is a growth-premium for the follower economy, because of international technology transfers. There is no explicit theory in my (and their) model since I (and they) assume theoretical works of past years and authors.

I follow Lucas' analytical representation, but I borrow Krugman's idea of economic growth as an increasing variety of goods. In other words, I hypothesize that growth of income and increase in the variety of goods are two sides of the same coin. At the economic and technological "frontier" there is a constant rate of growth $\dot{n} = \alpha n$, where n is the number of varieties produced in the economy.

Given the limited scope of this section, I do not need either a process of specialization (as in Krugman) or a "law" governing the succession of take-offs of different follower countries (as in Lucas', as I have only two countries).

At the beginning of our story (*t=0*) the L (leader) and F (follower) economies have the same living standard. In a traditional world, economies are characterized by the same low variety of (traditional) goods.

The two economies differ only because they start off in different periods. This has an impact on their rate of growth. The L economy's take-off is in *t=1*; and this economy starts growing according to the general law:

$$\dot{n}_L = \alpha n_L$$

F is any economy which takes off in t>1, and its rate of growth takes advantages of a catching-up process, according to:

$$\dot{n}_F = \alpha n_F \left(\frac{n_L}{n_F} \right)^{\beta}$$

where β is a (positive) parameter determining the intensity of the "backwardness advantage" originated by the relative gap n_L/n_F between the two economies.

I suppose that n at $t=0$ is $n_0=25$ for all countries, think of those 25 goods as "subsistence goods" [14]. Suppose that the follower take-off is in $t=180$. Expressed in "real" time, our story could start in 1800 as in Lucas', and 180 match 1980, a date around which we observed take-offs of a large portion of Asia. Moreover, in 1980, non active sectors appeared for the first time in world trade statistics when we use SITC rev.1, as seen in figure 4.

Table 3 is a synopsis of this section, and I will follow its lines to explain my point.

Consider the same parameter values used by Lucas ($\alpha=0.02; \beta=0.025$). If we observe our two economies in $t=210$, we will have $n_L^{210} = 1599$ and $n_F^{210} = 317$, i.e., economy L produces about 5 times more (goods) than the follower, while n_L/n_F was approximately 35 before the F take-off (in $t=179$)[15] (lines a, b and c of table 3). Nothing new under the sun.

In the final year, the developed L country, generally speaking, is a country which has accumulated a lot of knowledge, technically and organizationally speaking, and which produces a lot of things and is a very differentiated economy. F, instead, is still a concentrated economy. It is not developed because, by definition, it is characterized by the production/consumption of a relatively low variety of goods. The sector/product matrix of this economy is full of zeros, and the rise in the variety of goods in F is a way to consider the "imitation" process of a follower economy facing leader economies.

Suppose then that the process of creating goods (i.e., the process of modern economic growth) is partially Schumpeterian: $\alpha = (\gamma-\delta)$, with both γ and δ constant and positive, and $\gamma>\delta$. Now, let us suppose that $\gamma=0.03$ and $\delta=0.01$. This means that 3% of new goods are added every year, but that one third of this process of creation is a simple substitution of old disappearing goods[16]. Details of this process of creative-destruction are showed in lines d-e-f of the table: 1599 goods in the final year are the difference between the 2362 goods "invented" from t=1 to t=210 and the 787 "died" in the same period (plus the 25 of t=0).

[14] 25 has no specific meaning. It is there just to have a starting point, but I chose this value in order to have results that were not too far from the empirical ones which can be seen in figures 4 and 5; cfr. also note 12.

[15] The 2% rate of growth, that is known as the secular rate of growth of per capita income, turns out to be good also in order to represent the "real" structural evolution of the economy, at least if we accept the idea that each sector classification is correct when it is created. Utilizing a constant yearly rate of growth of 2%, I chose 25 sectors in 1800 in order to get a number that is very close to the 623 sectors of SITC rev.1 at 4 digits in 1962. In 1978, my imaginary economy had 832 sectors, quite close to the 783 of the rev. 2, and finally, it reached 1035 sectors in 1988, when rev. 3 accounted for 1031. As mentioned in the text, the number of sectors of the modelled economy reaches a total amount of 1599 in 2100.

[16] There is no particular meaning for the chosen values, except for the fact that their difference must be equal to 0.02, the hypothesized value for α. If I had chosen higher values for γ and δ, my results would have been in the same direction even if at a different scale.

Table 3
SYNOPTIC TABLE - Evolution of L and F economies and their statistical description

		description	T=0 (1800)	T=179 (1979)	T=210 (2010)
	economy ….				
a	n_L (varieties present in L)	$n_L = n_0[1+0.02]^t$	25	866	1599
b	n_F (varieties present in F)	$n_F(t) = [1+0.02]n_F(t-1)[n_L / n_F]^{\beta}$	25	25	317
c	Relative gap n_F/n_L (%)	(b/a)	1	3%	20%
d	Cn_L (n created from t=1 to T)	$\sum_{t=1}^{T} 0.03 n_L(t-1)$	0	1261	2362
e	Dn_L (n destroyed from t=1 to T)	$\sum_{t=1}^{T} 0.01 n_L(t-1)$	0	420	787
f	"New" existing variety	c-d (N.B.: a = d-e+n_0)	25	841	1574
	… and its description				
g	Registered varieties		-	866	866
h	L (and world) non active sectors	$\sum_{t=180}^{T} 0.1(Dn_L)$	-	0	37 (=10 % of 787-420)
i	L (and world) active sectors	g-h	-	866	829
j	L non active sectors (% of 866)	h/g	-	0%	4.2%
k	F non active sectors (% of 866)	(g-b)/g	-	97%	63%

Now, let us suppose that economists in *t=179* decide to represent the sector characterization of our two economies, and that statisticians are able to correctly state that $n = 866$ in the world economy in that year. For the following 30 years, the statistical representation of the economy will follow this convention, i.e., the economic structure will be represented by those $\overline{n} = 866$ sectors initially defined (line g of the table).

Also suppose that statisticians do not always recognize the creative-destruction process of the economy, also because new and old goods, even when they satisfy the same use, could be radically different in several fundamental aspects (as already underlined). As a consequence, it can sometimes happen that new goods are classified in sectors that are different from those of the old goods they are replacing. This kind of classification problems[19] depends on the fact that defining sectors is highly subjective outside the world of perfect competition, where only homogeneous goods exist[20].

Whenever this happens in our statistical representation, we will have a sector with no production (the old goods) and another "old" sector with a raised production (as the new goods are added to an already existing sector).

Imagine that in our sketched economy this problem has to do with a few cases, say about 10%. Starting from t=180, 10% of the disappearing goods are not replaced by new goods in the sectoral matrix of the economy. They become non active sectors.

In table 3, the effect can be seen and differentiated between the F and L economies.

Since the L economy produces all the existing goods, in the final year it will have 788 sectors with a positive production plus 37 non active sectors (i.e., with no registered production: see lines h and j). If we suppose that each variety is produced in the same amount and with the same price, as in many NEG models, this will mean that in *t=179* there was a situation of perfect equi-distribution since each sector share was $1/\overline{n}$. As a consequence, whatever the index of concentration we use, we will obtain a raised concentration in *t=210*: the L economy appears more specialized in the final year[21].

---

[19] A real life example could be the Railway locomotives-steam-and tender (4 digit code 7311 in sitc rev.1) which was not traded (registered) for the first time in 1984 and which definitely disappeared from international trade after 1988, while other kinds of railway locomotives, a very clear substitute, were (already) classified in different sectors.

[20] A sector can be defined and classified on the basis of many different characteristics: the type of final good (say for final consumption or as intermediate), the type of used material (natural or artificial), the technological process, the social or perhaps sociological characteristics ("traditional" or "modern"), etc..

[21] The most unfavourable hypothesis for my thesis would be if the 1599 goods were distributed equally among active sectors, i.e., if in the final year, we will have 767 "sectors" accounting for 2 goods and 25 for 3 (plus the 79 "non active sectors"). If we supposed that (innovation is not evenly spread and that) new goods are mainly registered in a reduced fraction of the initially defined 866 sectors, the final distribution would also be more unequal (concentrated).

But what happens if we observe the F economy? Here, there is a completely different picture. Before the take-off (*t=179*), the F economy is represented by the same 866 sectors, but only 25 of them have a positive production (3% of the total, see line k of the table). The process of imitation tells us that in this economy the number of variety starts growing rapidly after take off due to the growth premium, and in *t=210,* this economy will have 317 goods, as already observed.

If we accept the idea that, generally speaking, patterns of growth exist, we can think that those 292 (317 minus the initial 25) goods already existed in the L economy in previous years. They are "new" goods in the F economy but old in the L economy. Going to a little extreme, and only to simplify, we can suppose that these 292 goods were already present among the 866 goods of t=179, and because of that, in our statistical tables, they have their appropriate "cells". A a consequence, the F economy in *t=210* is statistically represented by 317 active sectors, about 37% of the total (line k). Each of these sectors accounts for a share of 1/317, while 549 non active sectors still remain.

The conclusion for the F economy is that we observe a clear "de-concentration" and "de-specialization" process since the initial condition of the F economy[22] was of only 25 active sectors, each one with a share of 1/25, and 841 non active sectors.

The picture for the follower economy would not substantially change even in the F economy were added sectors which, in a second moment, disappeared from the world (and F!) economy: in this case the active sectors of F would be 317-something, with "something" being between 0 and 37 (that is the number of non active sector in the final year).

AN EMPIRICAL CONFIRMATION THROUGH AN EXAMPLE OF TWO COUNTRIES

I suggested that the pattern of world exports structure representation, with its increase in the number of non active sectors, does reflect in the developed economies matrix, but not in that of the developing countries. This can empirically be confirmed by the simple empirical evidence in table 4 where I compare world export structure of France and Bangladesh, two countries at the extreme ranges of the scale of development[24]: developed countries export structure (France) turns out to be quite similar to that of the world's export structure, while the developing countries (Bangladesh) have a different pattern. Data are again SITC rev1, 4 digit, excluding petroleum.

[22] Remember the initial and simplifying hypothesis that each good is produced in the same amount.
[24] In this final section, I will use those two countries as practical examples for developed and developing countries.

Table 4

World, France and Bangladesh export sector structure compared- 2006

	World-France	World-Bangladesh
Correlation index (on exp. values)	0.80	0.15
Similarity index[25] (on exp. shares)	0.72	0.10

This evidence could in part depend on the high share of developed countries' export in world demand, but if this is true for developed countries taken all together, it could be less compelling in the case of an isolated country such as France.

Under these circumstances, it is obvious that we have to expect that world export trends, like the ones shown in the first sections, influence the economic structure of the developed countries more intensely than that of the developing economies.

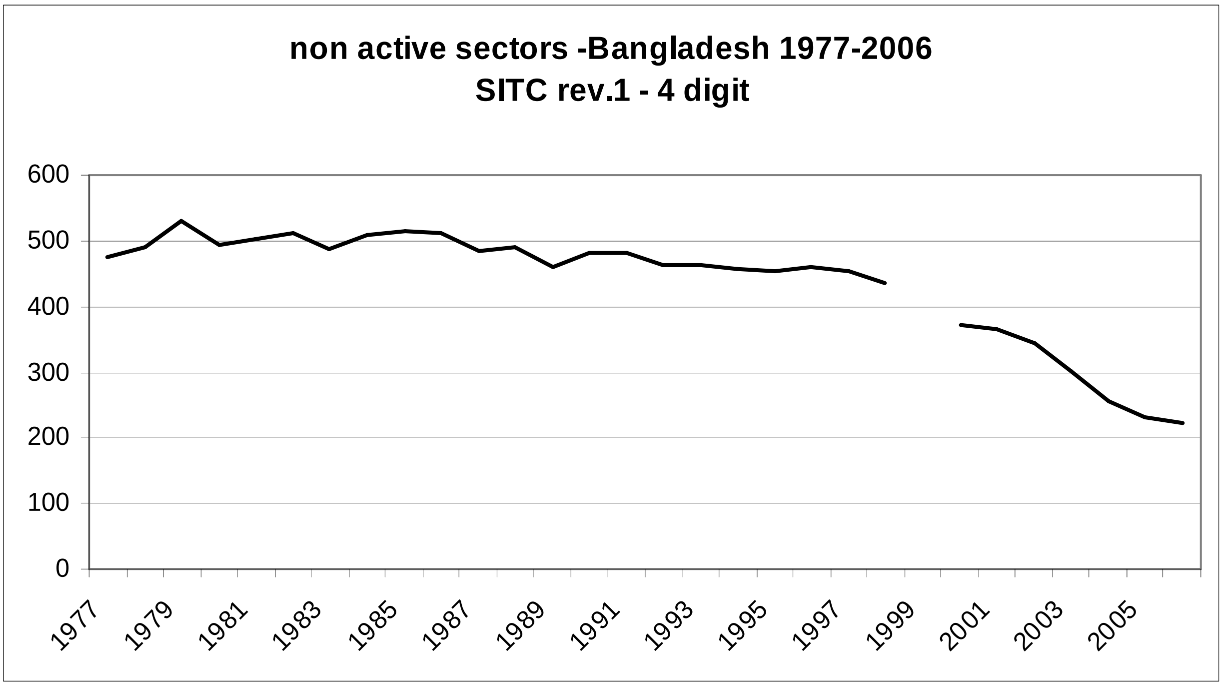
This empirical similarity depends on the fact that, as in the model above, a rich country produces (exports, in this empirical case) in almost all the existing sectors.

To do that, following the rough empirical indications of figure 4, I want to show my very simple statistics again, i.e., the number of non active sectors, both for Bangladesh and for France. It could be useful to remember that SITC rev.1, at 4 digit, has 623 sectors.

In the following figures 5 and 6, you can observe the result of my simple analysis. Figure 5 displays the evidence from Bangladesh where the number of non active sectors steadily declines, indicating a progressive de-specialization of that economy. Bangladesh had about 150 active sectors in 1977 (the first available year for this country), i.e., about 25% of the total. Note that this means that it did not export in about 470 sectors. After 30 years of development, Bangladesh had about 400 active sectors in 2006 (i.e. 220 non active sectors), around two thirds of the total.

[25] The similarity index is built according to: $SI = 1 - \left(\sum_s \left| \frac{x_{sc}}{x_c} - \frac{x_{sw}}{x_w} \right| \right) / 2$ (s: sector, c: country, w: world), and it ranges from 0 (non similarity) to 1 (perfect similarity)

Fig.5

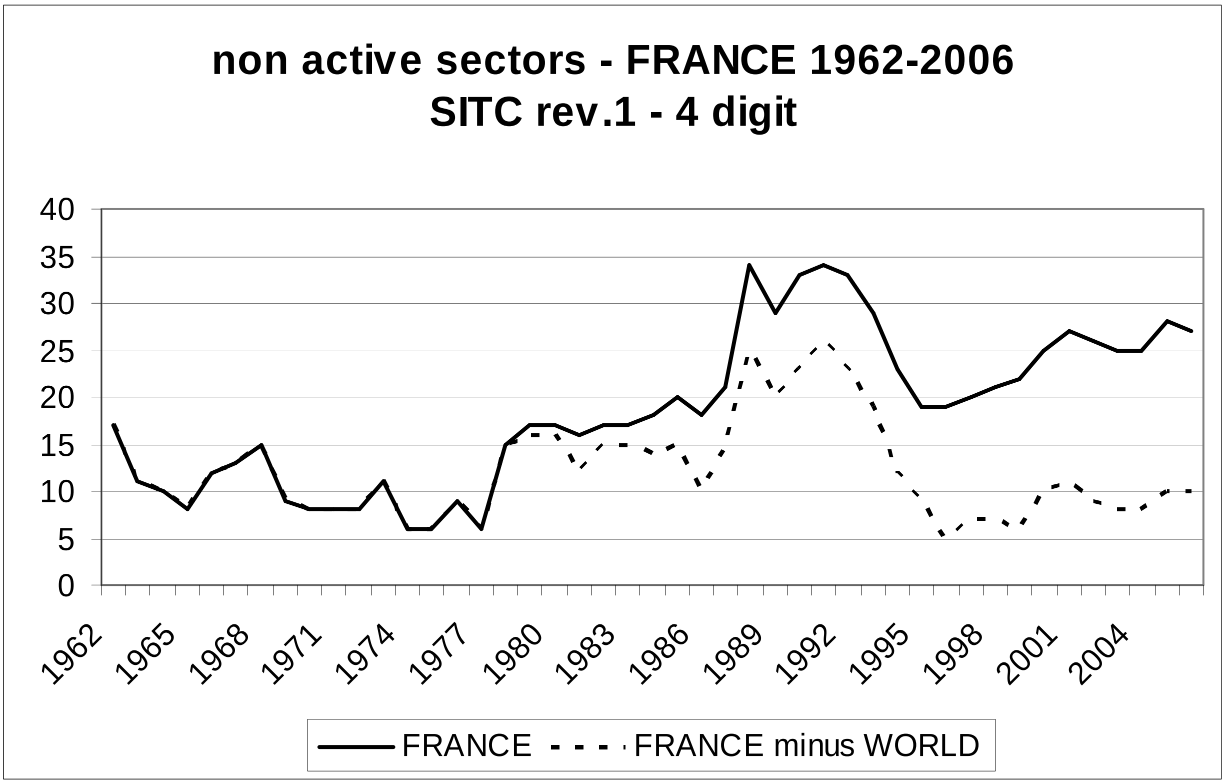


A completely different picture emerges for the advanced French economy (fig. 6), but, as we will see in a moment, we could reach a wrong conclusion if we did not consider the phenomenon showed in figure 4.

In fig. 6 two lines are showed:

- The continuous line is analogous to the previous line for Bangladesh, i.e., it is the total amount of non active sectors
- A second, dashed line is added in the French case. It represents the difference in the number of active sectors between France and those in the world. In practice, the dashed line represents the non active French sectors when we previously eliminate from the overall set of sectors those which are non active at world level.

Fig. 6



These two lines overlap initially , but they start diverging from 1980, the first year in which a sector disappeared in world statistics. We should observe that:

- First, there is a wave in the eighties, but note that it could be not even perceived using the y-axis scale of Bangladesh.
- Second, and more important in the present paper, looking at the continuous line, is that the total number of non active sectors seems to display a rising trend, a possible indication of decreasing export variety for this country. Nevertheless, we can easily appreciate that this is due to the rising number of "disappearing" sectors at world level (the difference between the continuous and the dashed lines). In fact, if we do not consider them, as in the case of the dashed line, no trend can be seen in French data. This suggests that there is no French "concentration" process, at least if measured through this rough method. If a concentration trend is found, it depends on world dynamics.
- Third, France has between 5 and 27 non active sectors (dashed line), less than 5% of the total, or, from a complementary perspective, it has a positive value of exports for more than

95% of the sectors (at 4 digit-SITC rev.1)[26]. This is not completely obvious, and it is not a consequence of the aggregation scheme (as discussed in the following).

An example could clarify this point. In 2006, France had 27 non active sectors (continuous line in fig. 6). From fig.4, we know that 17 of them were non active at world level. As a consequence, the exclusively "French" non active sectors were 10 (dashed line).

The simple empirical evidence of figures 5 and 6 is in line with the conclusion I obtained with the model presented above, both for what concerns the general problem and also for the similarity (dissimilarity) among export structures of the developed (developing) countries and the world average.

Finally, I would like to stress that these results also hold with a 5 digit disaggregation scheme and indeed, we could say that they are even clearer. In fact, even if I do not show here detailed data, I can remark that, at 5 digits, and considering only sectors that are active at world level, France still exports in more than 95% of the world sectors both in the first and in the final year, exactly as in the 4 digit case. On the other hand, Bangladesh moves from around 10% in 1962 to 50% in 2006 of the world's active sectors, the same kind of trend as at four digits (i.e., a clear decrease in specialization), but with lower values.

CONSEQUENCES AND CONCLUSIONS: A TALE OF TWO STORIES OF ECONOMIC PROCESSES AND STATISTICAL ARTEFACTS

As we have seen in the first part of the analysis, the highlighted bias in the sectoral representation of an economy becomes serious as long as the same statistical revision is used. When a new statistical revision becomes available, the bias can be considered sufficiently light for some years[27]. Nevertheless, in many cases, researchers need to work with a long term series, especially when structural questions and time evolution are the central targets of analysis. For example, any analysis of sectoral structural change generally requires a long term series of data.

In this case, the statistical bias I revealed might substantially influence the results.

A practical example can highlight this point. I refer to a recent strand of studies where some authors (Imbs and Wacziarg, 2003; Cadot, Carrère and Strauss-Kahn, 2007; Koren and Tenreyro, 2007), investigating the evolution of what is called the "extensive margin" in international trade, arrive to a conclusion that developed countries progressively specialize their production in a decreasing

[26] Even this is a trivial observation. Note that if France exports in so many sectors, it does not mean that it has a relative comparative advantage in all of them, but simply that it is able to export a very high variety of goods.

[27] According to the evidence presented in this paper, less than ten years.

“number” of goods (they found an increasing sector concentration). Given that the opposite is true for developing countries, they concluded that there is a U-shaped relationship between an index of overall specialization (usually an index of sector concentration of production and/or trade) and the level of development (per capita income or labour productivity). Cadot, Carrère and Strauss-Kahn (cit.), together with usual concentration indices, also added other “statistics” in order to better highlight this problem, specifically the number of “active lines”, that is, sectors with a positive value of export (taking ideas from Klinger and Lederman, 2004)). They find an inverse relationship between the number of active sectors and the concentration index; moreover both their indexes of specialization show that a “typical” country, along its development path, first de-specializes (increases the number of active sectors) and then specializes (decreases the number of active sectors)

I completely agree with the usual interpretation for the declining part of the curve: poor countries, in their process of acquisition of knowledge, human capital and technology, succeed in progressively diversifying their production structure. This has also been deeply analyzed in recent studies (Hausmann and Rodrik, 2003).

Nevertheless, while I do not refuse previous explanations proposed, I do think that the increasing part of the curve, which refers to developed economies, is at least a partial consequence of the statistical bias that I have discussed throughout the paper, even if I can not say to what extent precisely.

Without correcting for sectors disappearing at world level, we find an increase in “active sectors” at low level of income (Bangladesh) and a decrease in them at high income level (France): an inverse U curve. Nevertheless, if we “clean” the data, i.e., eliminate sectors that become non active at world level, we do not find any decrease in active sectors in the French case.

As a consequence, the inverse U-shaped curve is a tale of two different stories: the decreasing part effectively reflects a genuine economic process, i.e., the de-specialization of a rapidly growing follower economy, while the increasing part is due to a statistical artefact, deriving from the fact that we are not able to properly register product innovation. Obviously, my explanation is not necessarily alternative, but I think that it still poses a problem. Moreover, my evidence above is only indicative, and not conclusive. You can read it as a parable, or perhaps better, this evidence is like the tip of the iceberg.

Scholars in some way are aware of the problem. In fact, datasets are periodically revised. Nevertheless, some of literature’s conclusions about empirical structural change in an economy could still be biased.

In order to convince you that I am not wrong, think about the “classical” agriculture-industry-services pattern (probably an extreme example) in the present context: the economy starts concentrated (in agriculture) and finishes concentrated again (in tertiary). Somewhere in between, it is “differentiated”. But should we conclude that at the end of the period we have the same variety of goods than at the beginning? The same consideration, even in a completely different scale, can be done for more disaggregated data.

In reality, both agriculture and services (and industry) are much more differentiated macro-sectors today than several decades ago. As an example, while “traditional” services where concentrated in small-scale retail-commerce, some personal services (household servants, etc) and a few more, “modern” services include very complex and differentiated financial services, with many products and actors, transports of numerous kind, several levels of commerce, not to speak of “amusement” services, from “disneyland-type” centers, to gymn centers, “discos” etc.

But, in the three-sector picture, this growing variety cannot emerge[28]. Moreover, a growing variety can never emerge in any sector representation in which the number of sectors is fixed for the whole period of analysis.

The evidence proposed in this short paper does not exclude that other explanations can be valid at the same time. Moreover, imposing an equal size of production and equal prices for all goods, I limit my attention to a quantitative interpretation. It is obvious that it would be possible to enrich this picture by considering other dimensions of the problem. Nevertheless, I think that the bias I tried to describe so far should be considered.

As I said above, I do not propose a solution to the problem, but I could make a small suggestion. If we go back to the evidence in figures 1 and 6 and if the positive trend of world trade concentration is a statistical artefact, as I have suggested throughout this paper, a possible solution could be to normalize single country data to the world average, thus eliminating the positive trend we observed (an example could be to use relative concentration indexes).

[28] The American economy was “concentrated” in agriculture at the beginning of the XIX century when 70% of GDP came from agriculture. Today, it is concentrated in services, since tertiary value added represents 80% of GDP. Would you say that today USA economy offers the same (or a lower) variety of goods to its citizens?

BIBLIOGRAPHICAL REFERENCES

Aghion P, Howitt P. (1992), *A model of growth through creative destruction*, Econometrica, 60 (2), pp. 232-251

Amurgo-Pacheco A, Piérola M.D. (2008), *Patterns of export diversification in developing countries: intensive and extensive margins*, Policy Research Working Paper n.4473, World Bank

Broda C., Weinstein D.E. (2006) *Globalization and the gains from variety*, Quarterly Journal of Economics, vol. 121, May, pp. 541-585

Cadot O., C. Carrère and V. Strauss-Kahn (2007). *Export diversification: what's behind the hump?*, Centre for Economic Policy Research Discussion Paper No. 6590.

Fujita M., Krugman P., Venables A.J. (2001), *The spatial economy: cities, regions, and international trade*, MIT Press

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

Hausmann, R., Rodrik D. (2003), *Economic development as self-discovery*, Journal of Development Economics, 72, pp. 603-33.

Henderson J. V., Thisse J. (2004), *Handbook of regional and urban economics*, vol. 4 Cities and Geography, Elsevier

Imbs J., Wacziarg R. (2003), *Stages of diversification*, American Economic Review, 93 (1), pp. 63-86

Klinger B., Lederman D. (2004), *Discovery and development, an empirical exploration of new products*, World Bank Policy Research Working Paper No. 3450

Koren M., S. Tenreyro (2007), *Volatility and development*, Quarterly Journal of Economics, 122(1), pp. 243-287.

Krugman p. (1979), *A model of innovation, technology transfer, and world distribution of income*, Journal of Political Economy, 87 (2), pp. 253-266

Kuznets S. (1972), *Innovations and adjustment in economic growth*, Swedish Journal of Economics, 74 (4): 431-451.

Kuznets S. (1973), *Modern economic growth: findings and reflections*, American Economic Review, LXIII (3): 247-258.

Lucas R. (2000*), Some macroeconimcs for the 21st century*, Journal of Economic Perspectives, 14(1), pp. 159-168