A LATERITE AND ECONAFRIQUE POLICY PAPER

www.laterite-africa.com

Export similarity, growth and counter-factual analysis

By Sachin Gathani and Dimitri Stoelinga

8/31/2011

In this paper we develop new insights on economic growth and propose a new tool for counter-factual analysis, building on a measure of the export similarity of countries. We show that the export similarity of pairs of countries is a good proxy for their similarity in capabilities and can therefore be used to identify accurate comparators for a country of interest. We show that the match between countries and their closest comparators can be quite remarkable; a synthetic combination of a country's most similar exporters often perfectly matches economic growth in the reference country over a long period of time (in many cases more than 40 years). We build on this idea to propose an innovative technique with which policy makers can conduct counter-factual analysis. We call this approach Proximity Controls and illustrate, using three case studies, how it can be used to measure the impact of major events (e.g. a crisis, a major policy change, etc). We use various Proximity Control strategies to measure the long term impact of the 1997 East Asian financial crisis on growth in Indonesia, the impact of Ivory Coast's decade long political crisis, and the impact of Kenya's dual consecutive domestic crisis, the December 2007 election and the 2008-2009 drought..

**We are extremely grateful to Miguel Almunia, Jonathan Argent, Ricardo Hausmann, Cesar Hidalgo, and Rupert Simons for their insights and advice. All errors in this paper are our own.*

Introduction

The objective of this policy paper is to introduce a new way of thinking about economic comparisons and counter-factual analysis, building on a measure of the export similarity of countries. We hope to convince the reader of the value of targeted, data-driven, cross-country economic comparisons and of the benefits of analyzing the global economy as a network of countries with points of similarity, rather than a group of individual countries with a set of different macro-economic performance indicators. Analyzing the global economy from a network perspective also enables us to develop new types of metrics, which we show lead to interesting insights about economic development, and to introduce new visual tools that researches can use to better represent and understand economic activity.

This piece of work is specifically targeted at economic policy makers. What we hope they will get out of this is: (i) a number of data-driven strategies with which they can identify optimal comparator countries for a country of interest (be it at the sector level or comparisons across time); (ii) an innovative technique to carry out aggregate counter-factual analysis at the sector or country levels, which we call Proximity controls and which is largely inspired by the synthetic controls methodology of Abadie et Gardeazabal (2003); and (iii) new insights about economic growth, in particular the fact that: countries with similar export structures tend to grow at similar rates, countries that deviate from these shared growth rates tend to converge back towards them, and the connectedness of countries in the exports proximity network is very informative about the dynamics of two key drivers of economic growth: diversification and specialization.

These techniques are all derived from a measure of the export similarity between countries, which we show is predictive of how similar countries are in terms of a whole range of other indicators, including GDP per capita, growth, imports, educational attainment, institutional performance, export sophistication, etc. This paper builds on previous and current work by Hausmann, Hidalgo et al - who introduced the concept of the product space (2006) and have just recently also introduced a new metric of the export similarity between countries (2011, forthcoming) - as well as the synthetic controls methodology introduced by Abadie et Gardeazabal (2003) and further developed by Abadie, Diamond et Hainmueller (2008).

The three main lessons that this paper takes away from the exports research of Hausmann, Hidalgo et al are that: (i) to produce a certain product with a comparative advantage a country needs to have the right capabilities mix (including non tradable-capabilities such as property rights, regulation, infrastructure, specific labor skills); (ii) it is possible to estimate how similar the capabilities required to produce a pair of products *(i,j)* are, by measuring the likelihood that countries that export product *i* with a comparative advantage also export product *j* with a comparative advantage; and (iii) it is possible to transpose this measure of similarity between products, to measure the export similarity between countries. Where Bahar, Hausmann et Hidalgo (2011) use a continuous Revealed Comparative Advantage (RCA)[1] vector to measure the export similarity between countries, we use a measure based on a cut-off of the RCA vector,

[1] See Annex 1 for a definition of Revealed Comparative Advantage

distinguishing between products for which a country has a Revealed Comparative Advantage (RCA>1) and products for which a country does not have a comparative advantage (RCA<1).

We use this measure of the similarity between countries – which we call *Proximity* - to identify the most appropriate comparators for a certain country of interest. Our main assumption in this paper - and we present evidence to support this claim – is that if countries have a revealed comparative advantage in similar products, then they have a more similar capabilities mix than countries that have a revealed comparative advantage in very different products. In particular, we show that countries that have the most similar export structure also tend to have the most similar performance (both in terms of levels and trends) on a whole range of social and economic indicators. On this basis, we argue that *Proximity* is a good proxy for the similarity in capabilities between countries and, by extension, also a good way to identify comparators.

We then show – based on different strategies - that it is possible to construct a testable control region for a country's performance on a certain variable of interest using a linear combination of its closest comparators. We call this method *Proximity controls*. It draws on lessons from the synthetic controls methodology, developed by Abadie et Gardeazabal (2003) and in particular the techniques used by Abadie, Diamond et Hainmueller (2008) to test the validity of the synthetic controls they construct. We illustrate how this approach works by estimating the economic impact of three major events: (i) the impact of the 1997 East Asian financial crisis on GDP per capita growth in Indonesia; (ii) the impact of Ivory Coast's decade long political crisis on its GDP per capita (focusing on the 1999-2009 period); and (iii) the impact of Kenya's dual domestic crisis, the December 2007 election and the 2008-2009 drought, on its GDP per capita.

This paper proceeds as follows: we briefly describe the data utilized, before introducing the concept of Export Proximity and the properties of the Export Proximity Space; next, we explore the relationship between the growth rates of countries that are close to each other in the export proximity, providing a number of insights on economic development and introducing ways to identify comparator country's for a country of interest; we then propose three different strategies policy makers can use to develop Proximity Controls for reference countries and estimate the impact of a major event. We apply and test the relevance of these tools using three examples: Indonesia (1997), Ivory Coast (1999-2009), and Kenya (2007-2009). We close with a discussion on the policy implications and limitations of the Export Proximity Space.

Data

The data used to calculate the export similarity patterns as well as all export related data is from the BACI database, which is a world trade database developed by CEPII at a high level of product disaggregation. BACI is developed using a procedure that reconciles the declarations of the exporter and the importer, based on original data provided by the United Nations Statistical Division (COMTRADE database)[2]. BACI provides bilateral values and quantities of exports at the HS 6-digit product disaggregation, for more than 200 countries. However, we limit this study to countries with a population greater than 3 million because as the economics of small

[2] See http://www.cepii.fr/anglaisgraph/bdd/baci.htm

economies, in particular island economies, often do not apply to larger countries. Our sample is thereby reduced to 124 countries. The Export Proximity measures we derive in the entire study are based on 1995 and 2005 data.

All other economic indicators (GDP per capita data, GDP growth, etc) have been taken from the World Development Indicators database, unless otherwise indicated. These cover the period 1980-2009. Any data including monetary values, is expressed in terms of constant 2000 USD.

The export proximity space and how it relates to capabilities.

In this section we introduce the Export Proximity Space – which we will show has some properties that can be used to deepen our understanding of how the global exports industry works. The export proximity space - inspired by Hausmann et al's product space - is a network that links countries to each other based on how similar their exports are. Countries that have similar exports will be close to each other in the export space; countries that have very different exports packages will be further away. The logic behind the export proximity space is exactly the same as the logic behind the product space, except that instead of linking products to products, it links countries to countries. In the product space, products that require similar capabilities to be produced are close to each other, while products that require a different set of capabilities are further away. For example, it is very likely that laptops and 3G mobile phones would be closer to each other in the product space than laptops and bath-tubs, for the simple reason that they require more similar technologies and skills to be produced than bath-tubs. In the same way, we show that countries that are close to each other in the export proximity space have a comparative advantage in similar products because they have a similar capabilities-mix and vice-versa. Therefore one would expect similar countries such as Holland and Germany for example, which have much more similar capabilities (e.g. institutions, human capital, infrastructure, access to technology, etc) than let's say Holland and Chad, to be closer to each other in the export proximity space than Holland and Chad.

To measure how similar the exports of a pair of countries are we introduce a measure called *Proximity*. *Proximity* is based on the concept of revealed comparative advantage (see Annex 1) and is calculated in a very similar way to the distance between products in the product space (see Hausmann et Klinger, 2006). We first specify how Proximity is calculated, before illustrating some properties of the export proximity measure which indicate that Proximity is likely to be a good proxy for the similarity of capabilities between pairs of countries.

Measure of export proximity

Our measure of the export similarity of two pairs of countries is inspired by the work of Hausmann et Klinger on the proximity between products. We define the export similarity – or *proximity* - of two countries as the number of common products in which both countries have a revealed comparative advantage, weighted by the total number of products in which the most diverse of the two countries has a revealed comparative advantage (the most diverse country being the one with the highest number of products with a revealed comparative advantage). The

reason we base the measure on the concept of revealed comparative advantage, is to ensure that we capture only information on important products for a country's exports.

Formally, this measure of proximity between two countries *a* and *b* at time *t* can be written as:

$$Proximity_{a,b,t} = \frac{\sum_{i=1}^{P}\sum_{j=1}^{P} X_{a,i,t}X_{b,i,t}}{max\left(\sum_{i=1}^{P} X_{a,i,t}, \sum_{i=1}^{P} X_{b,i,t}\right)}$$

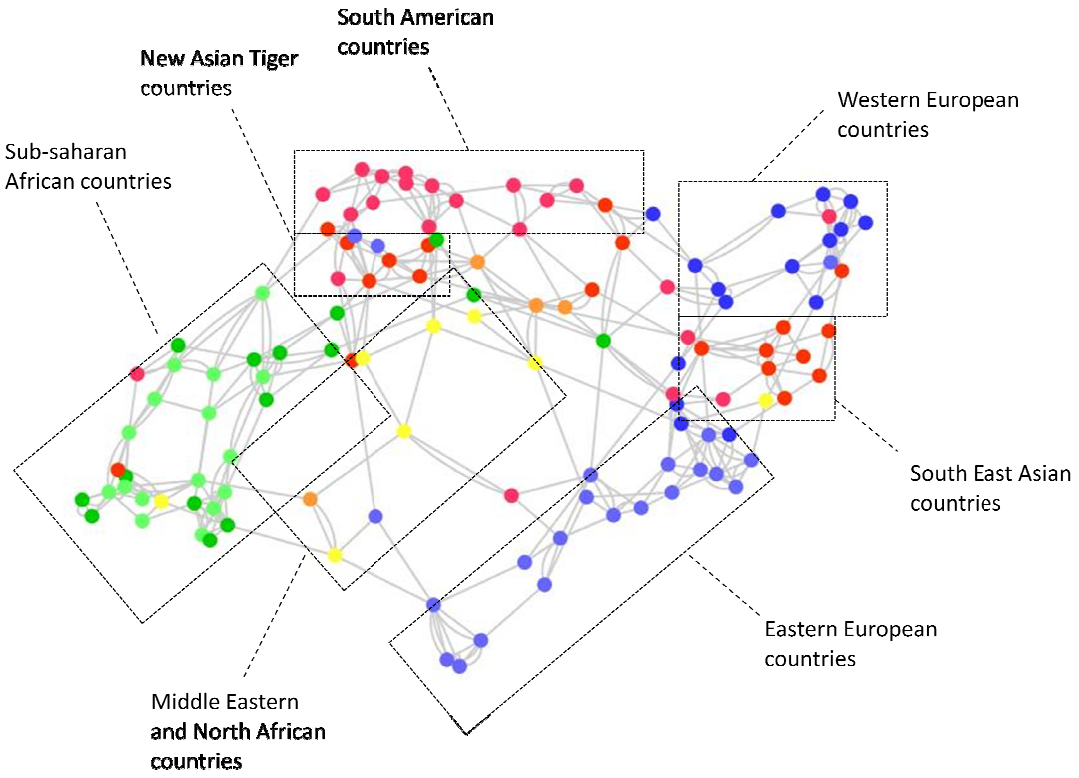
$$where\ X_{a,i,t} = \begin{cases} 1\ if\ RCA_{a,i,t} > 1 \\ 0\ otherwise \end{cases}$$

The reason the denominator is the maximum of the total number of products in which either country has a revealed comparative advantage is to ensure that this measure of proximity is symmetric (i.e. $Proximity_{a,b} = Proximity_{b,a}$) and to minimize the proximity of countries with different levels of diversification. Had the minimum been used in the denominator rather than the maximum (which would also have ensured symmetry), then the similarity between a relatively less diversified exporter and a more diversified exporter would have been overstated. It is important to note that this is just one of many possible measures of export similarity – Bahar, Hausmann et Hidalgo (2011) for example use a Pearson correlation based on the continuous RCA vector to calculate similarity, and one could think of other ways to structure a measure of similarity. Each approach has advantages and disadvantages. By using an RCA cutoff at 1 and including in the denominator the maximum diversity of a pair of countries, we are making a clear choice of: (i) focusing only on the significant exports of a country; and (ii) minimizing the proximity of two countries with very different diversity levels. This serves the purposes of this analysis very well, but might not always be the most optimal approach.

The graph below illustrates what the global export proximity space based on this measure looks like. The nodes represent countries, while the edges between them represent the export proximity distance between countries. The further away countries are from each other in the network, the more different their areas of revealed comparative advantage; the closer, the more similar their areas of comparative advantage. The network representation below only depicts the three closest neighbors of countries in the export proximity space and does not differentiate between incoming and outgoing connections. We highlight groupings of different countries to give the reader a sense of what the export proximity space looks like. At a first glimpse, it seems to make sense.

Figure 1: Network representation of the export proximity space[3]

[3] Note that the distance between countries in this representation is inversely proportional to their Proximity score. The higher the Proximity score the closer countries are in the export proximity space.



This measure of export proximity can be expanded to include triplets, quadruplets, quintuplets, etc, of countries, rather than simply pairs (the same should apply to Hausmann et al's measure of export proximity). The export proximity of a group of three countries for example would be the intersection of their product spaces (i.e. the total number of products in which all three countries have a revealed comparative advantage in) divided by the export diversity of the country with the highest diversity. By construction the proximity of a pair of countries *(a,b)* will be bigger or equal to the proximity of a triplet of countries *(a,b,c)*; the proximity of a triplet of countries (a,b,c) bigger or equal to the proximity of a quadruplet (a,b,c,d), and so forth. We call this the degree of Proximity: the degree is 2 when we measure the Proximity of pairs of countries, 3 with triplets, 4 with quadruplets, etc. The maximum degree is N-1; N being the total number of countries in the sample. While one loses information in terms of Proximity as the degree increases, one gains a lot of information in terms of the number of possible combination of countries for which we have information. Assuming there are 100 countries in the sample, and given that Proximity is symmetric, there would be 2,475 different combinations of pairs of countries, 53,900 different possible combinations of triplets of countries, 980,000 possible combinations of quadruplets, and more than 15 million combinations of quintuplets.

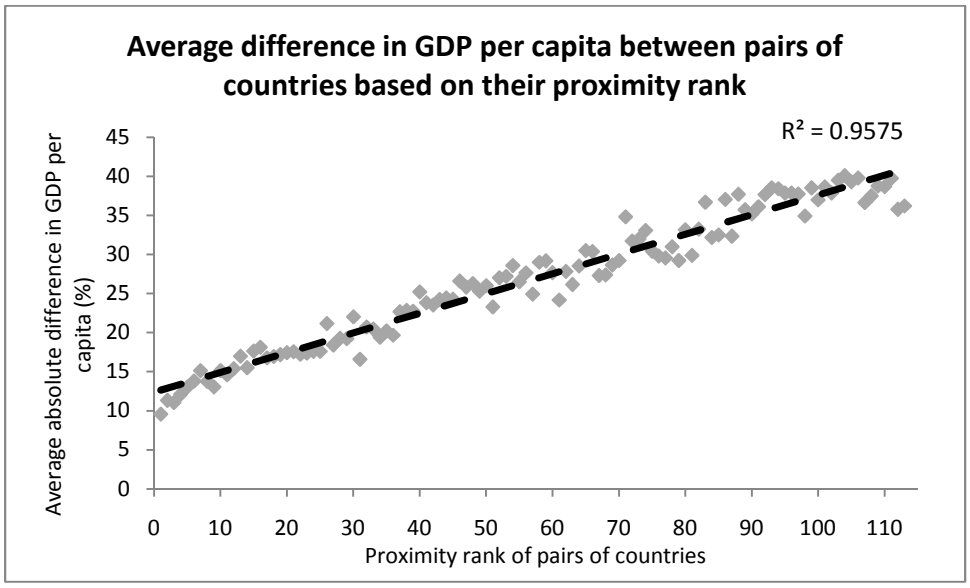
Properties of the export proximity space

In this paper we use the export proximity measure as a proxy of how similar countries are in terms of their capabilities. The reason a proxy is needed, is because some capabilities – that can play an important role in determining whether a country has a competitive edge or not in the production of a certain product – are not directly observable or measurable. Examples include business regulations, the efficiency of institutions, specific skills required to produce a certain

good, the adequacy of the infrastructure mix for the production of a certain product, etc. While it is impossible to prove that export proximity is a good proxy for the similarity in capabilities between countries, we can point to a number of properties of the export proximity space which strongly suggest this is the case. Also, in Annex 2, we show that export proximity provides much stronger signals and correlations than alternative variables, in particular GDP per capita and years of schooling (a proxy for human capital).

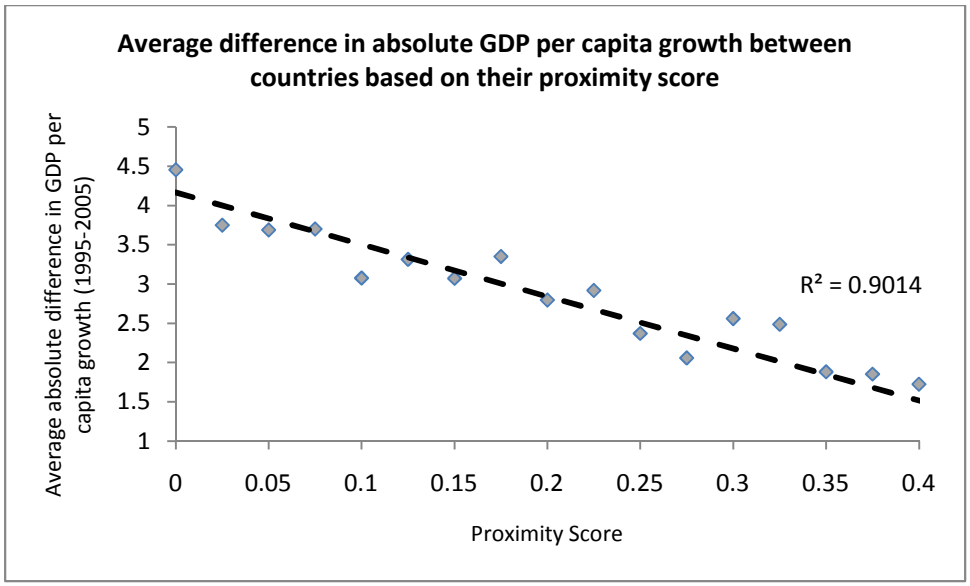
Property 1: On average, the closer countries are to each other in the export proximity space, the more similar their GDP per capita levels

The graph below depicts the average absolute difference in GDP per capita between countries, based on their export proximity rank with another country. The closest country to a reference country in the export proximity space is ranked one, the second closest ranked 2, and so forth. As can be seen in the graph below, the average difference in GDP per capita between countries and their closest comparator is about 9%. The further away countries are from each other in the export proximity space, the greater on average the difference between their GDP per capita levels. This relationship is linear and strongly statistically significant.



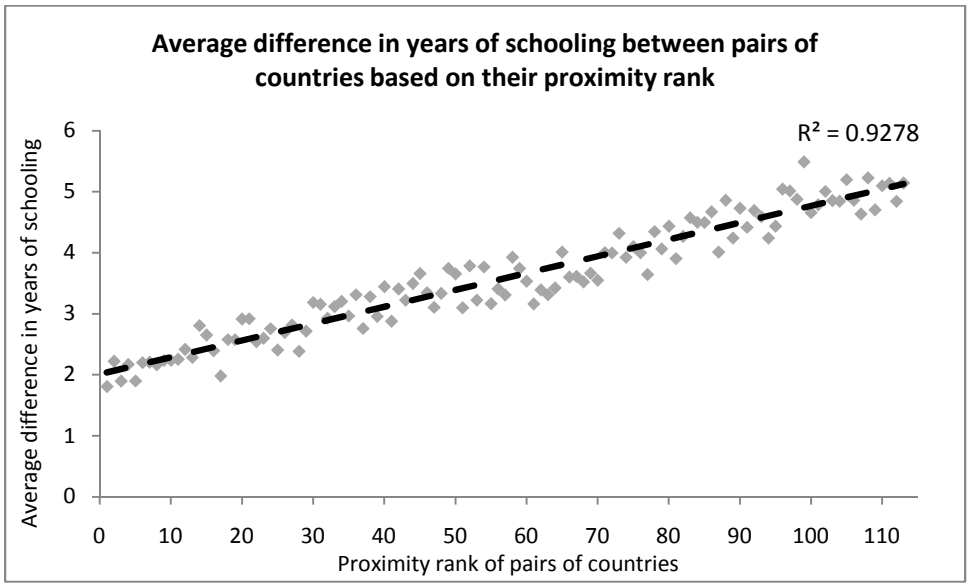
Property 2: On average, the closer countries are to each other in the export proximity space, the more similar their economic growth rates

As can be seen in the graph below, the greater the export proximity between countries, the smaller on average the absolute difference between their GDP per capita growth rates (we take the average annual GDP per capita growth rate during 1995-2005). The relationship is strongly statistically significant.



Property 3: On average, the closer countries are to each other in the export proximity space, the more similar their levels of human capital

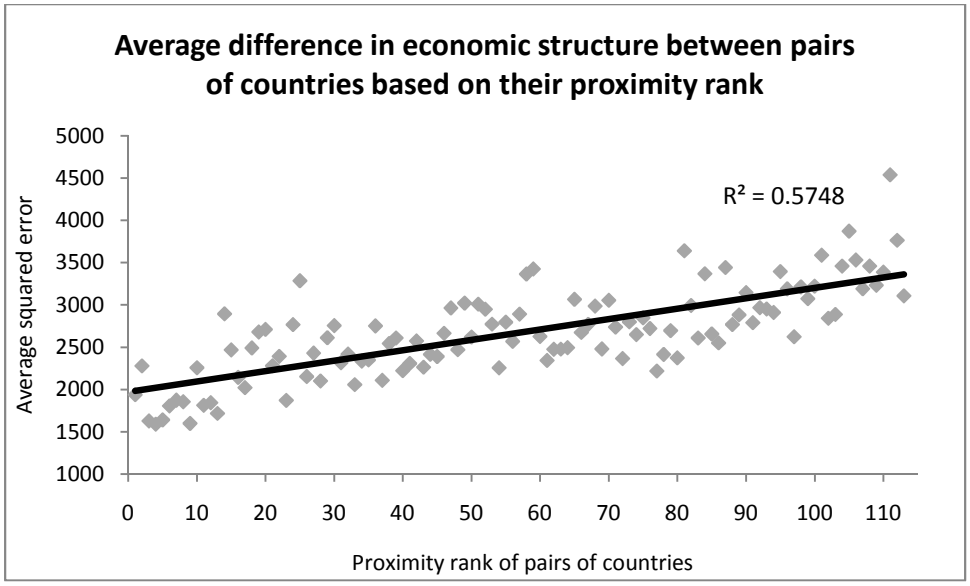
The same finding for GDP per capita and growth also applies to human capital. We find that countries with similar exports have more similar levels of average schooling per capita than countries with very different exports. This relationship holds when controlling for differences in GDP per capita.



Property 4: On average, the closer countries are to each other in the export proximity space, the more similar their economic structure

To determine whether countries have a similar economic structure or not we calculate the sum of the squared difference between these countries on a number of key economic indicators. We call this measure the squared error and apply it to all pairs of countries in our sample based on differences between countries on the following structural economic indicators: gross fixed capital formation (%GDP), domestic savings (%GDP), exports (%GDP), imports (%GDP), agriculture (%GDP), industry (%GDP) and services (%GDP). The smaller the squared error, the more similar a pair of countries are on average on the selected indicators.

As can be seen in the graph below there is a strong relationship between proximity and squared error. On average, countries that are closer to each other in the export proximity space fit each other's structural economic indicators better than countries that are further apart. This relationship holds when controlling for differences in GDP per capita.

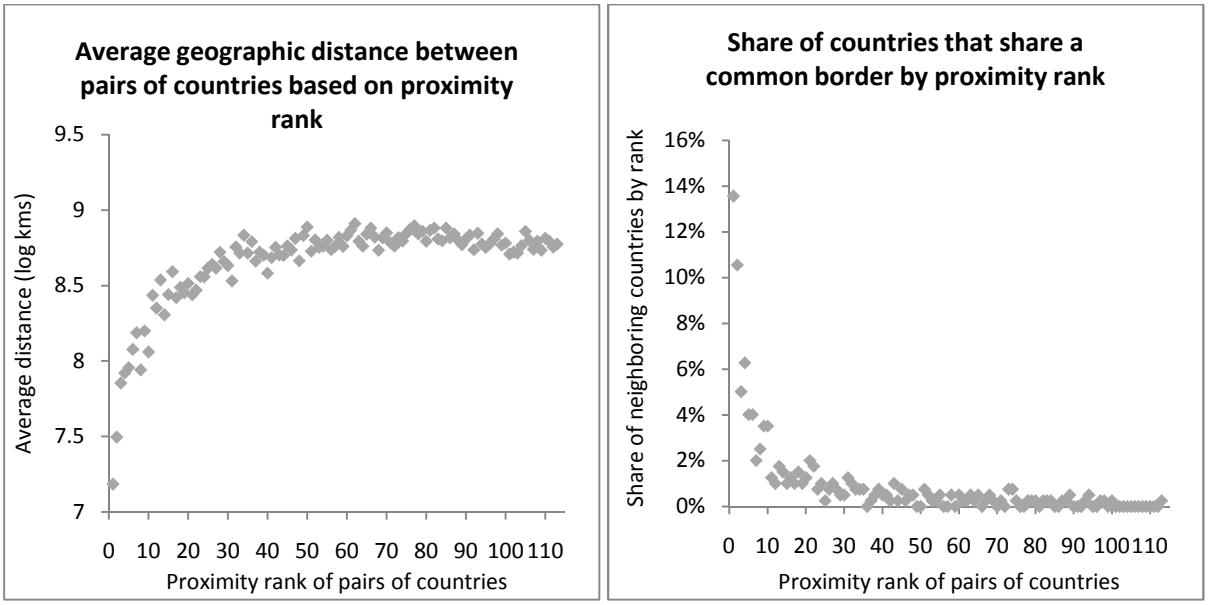


Property 5: On average, countries that closer to each other in the export proximity space, are also geographically closer to each other

Let us make a not-so-controversial assumption that capabilities are more easily transferable between countries that are closer to each other geographically than countries that are further apart. This seems to be a reasonable assumption to make and there are many examples that would justify why. The most obvious example is that of regional economic partnerships, such as the EAC or the EU, which promote the free movement of people, goods and capital between neighboring countries, thereby facilitating the transfer of capabilities. Infrastructure is also a good example in the sense that countries can make use of existing infrastructure in neighboring countries such as ports, airports, and financial infrastructure. Add to that the fact that countries that are close to each other are also more likely to share similar natural resources (e.g. oil, minerals, etc) and a similar climate (meaning they can growth similar crops, etc), and our prior quickly becomes that countries that are close to each other geographically should also have similar capabilities. If capabilities were really intricately linked to the comparative advantage

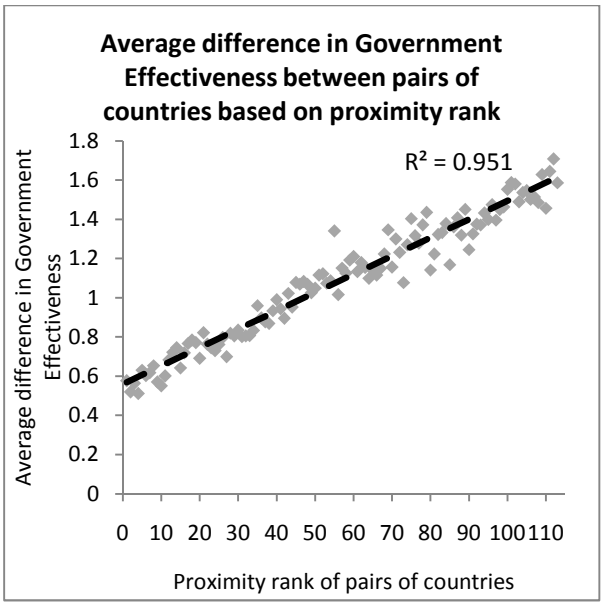
patterns of countries – as the authors strongly believe - then we would expect countries that are close to each other geographically to also have similar patterns of comparative advantage. Do we observe this in the data?

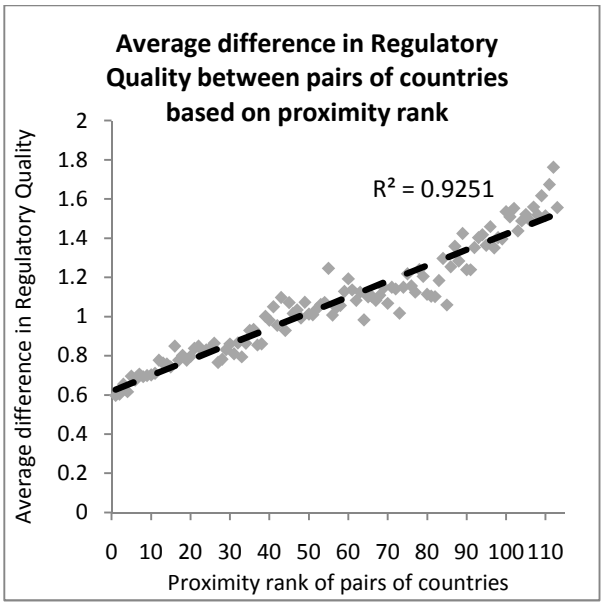
The answer is yes. As can be seen in the graphs below, as countries move further apart from each other in the export proximity space, the geographic distance between them also increases – exponentially initially. Moreover, pairs of countries that share a common border are also much more likely to be closer to each other in the export proximity space. In other words, neighboring countries tend to have similar comparative advantage patterns, thereby confirming our initial intuition. This is also one of the key findings of Bahar, Hausmann and Hidalgo's work (2011) on the producer space.



Property 6: On average, the closer countries are to each other in the export proximity space, the more similar their institutional quality

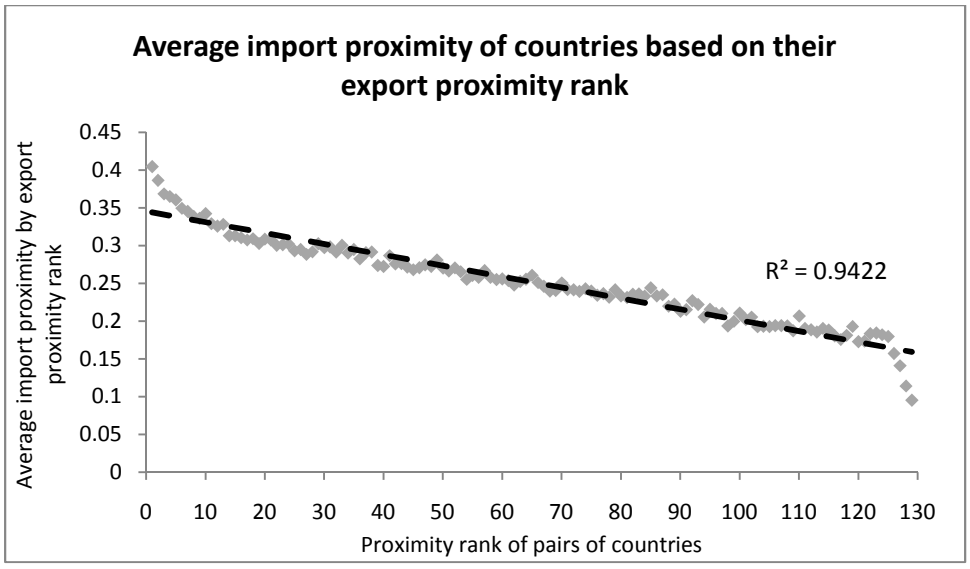
To measure how similar the institutional performance of pairs of countries are we use the Kaufmann Governance Indicators on government effectiveness and regulatory quality and measure the absolute differences in scores between countries. Again we find that countries that are closer to each other in the export proximity space tend to have more similar institutional quality (effectiveness and regulatory quality) than countries that are further apart. The relationships are strongly statistically significant and hold when controlling for differences in GDP per capita.





Property 7: On average, the closer countries are to each other in the export proximity space, the more similar their imports

To test whether countries that export similar products also import similar products, we create a measure of Import Proximity which mirrors the methodology we used in the export space. Based on this measure, pairs of countries with a higher import proximity have a more similar import package than pairs of countries with lower import proximity levels. As can be seen in the graph below we find a very strong correlation between how close countries are to each other in the export proximity space and how similar their import package is. On average, countries that export similar products are also more likely to import similar products, which means that they require similar economic inputs. While similarity in the export proximity space would indicate that countries have similar capabilities (they have the right capabilities mix to produce a certain product with a comparative advantage), similarity in the import space would indicate the corollary: that countries lack similar capabilities, and hence need to import them from abroad.

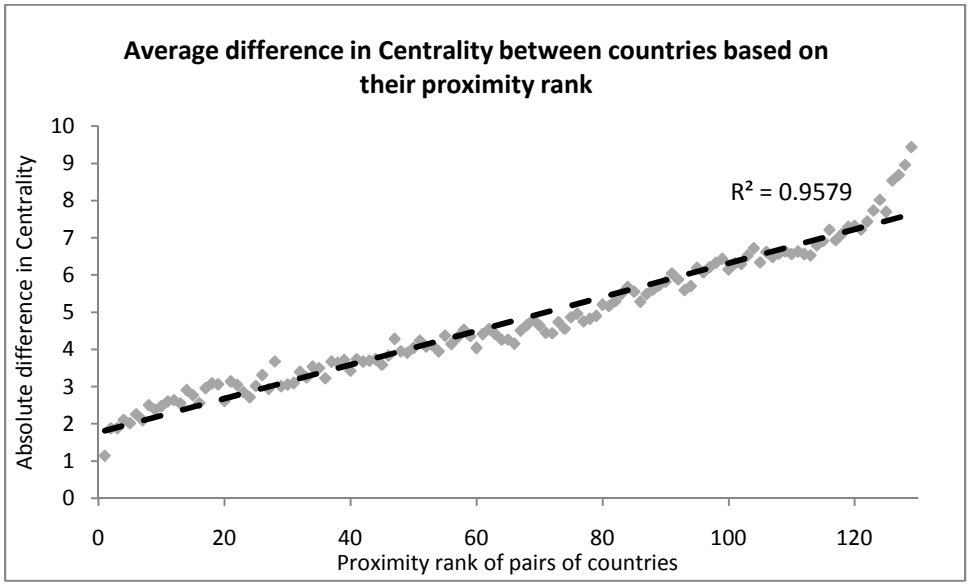


Property 8: On average, countries that are close to each other in the export proximity space are at a similar stage in their economic development process

To show this we introduce a new simple measure of the connectedness of a country in the export proximity space, which we call Centrality. We define the Centrality of a country as the sum of its proximities with all other countries in the export proximity space:

$$Centrality_i = \sum_{j=1}^{N} Proximity_{i,j}$$

Given that Centrality is the sum of proximities, countries with high levels of Centrality are more strongly connected to other countries in the export proximity network. They therefore compete on global product markets with more countries than countries that have lower levels of Centrality. A country can have a high level of Centrality if it is highly diversified - and therefore competes on multiple product markets – and/or if its products are ubiquitous (i.e. also exported by many other countries). Not surprisingly, countries that are close to each other in the export proximity space have very similar levels of Centrality (see graph below). We show in the next few paragraphs that this similarity suggests that countries that are close to each other in the export proximity space are at a similar stage in their economic development process.



As can be seen in the graph below, we find a very clear U-shaped relationship between economic development and Centrality[4]. Centrality first increases along with GDP per capita, before dropping again after a certain threshold. The relationship is non-monotonous and points towards the existence of three distinct phases of economic development:

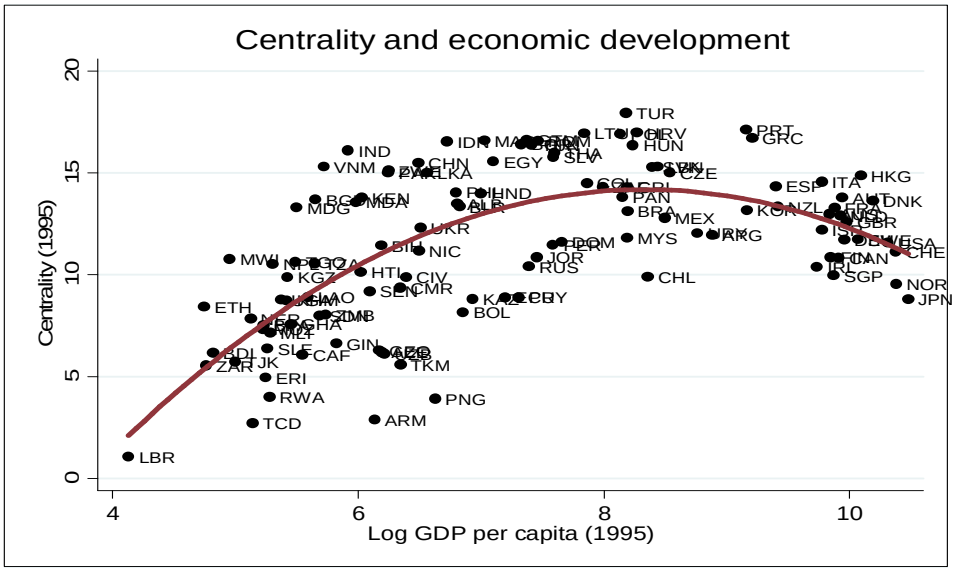
(i) Phase 1: countries start-off with very low levels of Centrality and find themselves at the periphery of global export markets;

[4] We have excluded from the sample countries in which mineral and oil exports account for more than 50% of the export package.

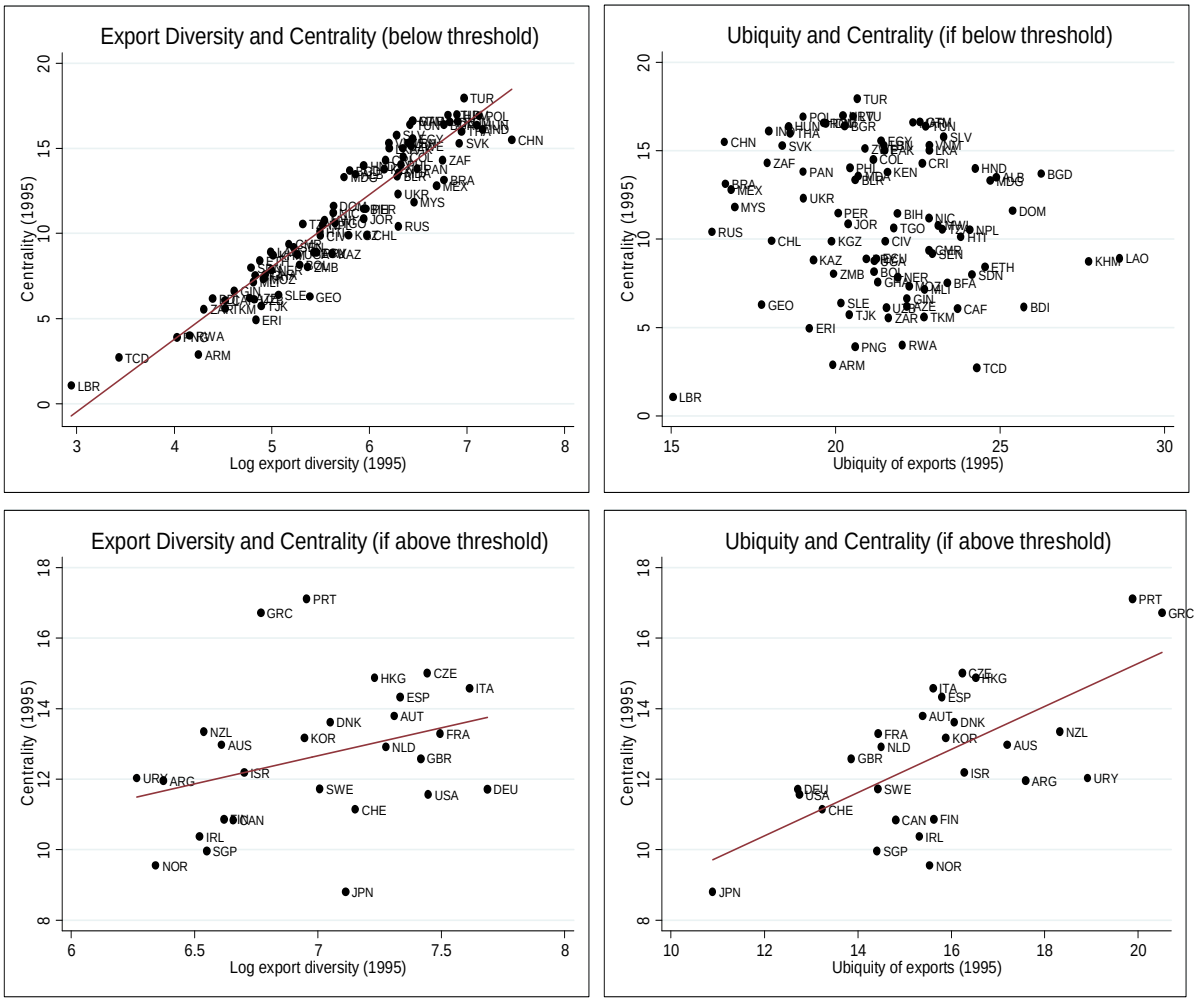
(ii) Phase 2: as they develop, countries diversify into new product markets, leading to increased competition and higher levels of Centrality;

(iii) Phase 3: after a certain development threshold, estimated at about USD4150 (=exp(8.32)) in constant 2000 US dollars (which is where countries like Chile and Croatia were in 1995), countries start to diversify into more specialized and sophisticated products where competition is limited to a few highly developed economies – this in turn leads to a drop in Centrality.

The relationship between Centrality and GDP per capita is best captured by a quadratic curve, which is highly statistically significant (with R^2=0.45).



The curve above is the result of two key forces of economic development: diversification and specialization. Before a certain development threshold is achieved, economic diversification is the force driving a country's position in the export proximity space; after that threshold is achieved, specialization takes over as the driving force. This can clearly be seen in the four graphs below, which respectively depict the relationship between export diversification and Centrality, and export ubiquity and Centrality, below and above the GDP per capita threshold. As in Hausmann et Hidalgo (2009), we define export diversification as the sum of products a country has a revealed comparative advantage in, and export ubiquity as the average number of countries which export the same products as the country of interest. The smaller the ubiquity of a product, the higher its level of sophistication.



These results are in line with findings by McMillan and Rodrik (2011) who identify a stylized relationship between the productivity ratio of agriculture over non-agriculture and the average labor productivity of the economy. As in the case of Centrality, the relationship between these two variables is non-monotonous: countries with low levels of economic development have a high agriculture/non-agriculture productivity ratio, reflecting the fact that they have few industries with low levels of productivity (i.e. they are non-diversified); as countries develop and enter new industries, the gap between the productivity of the still traditional agriculture sector and the non-agriculture sector widens (this corresponds to the diversification phase discussed above); finally, after a certain threshold of economic development, labor increasingly starts moving away from the traditional low-productivity sectors towards more modern sectors, thereby harmonizing productivity in the economy and pushing the agriculture/non-agriculture productivity ratio higher again (this corresponds to the specialization phase discussed above).

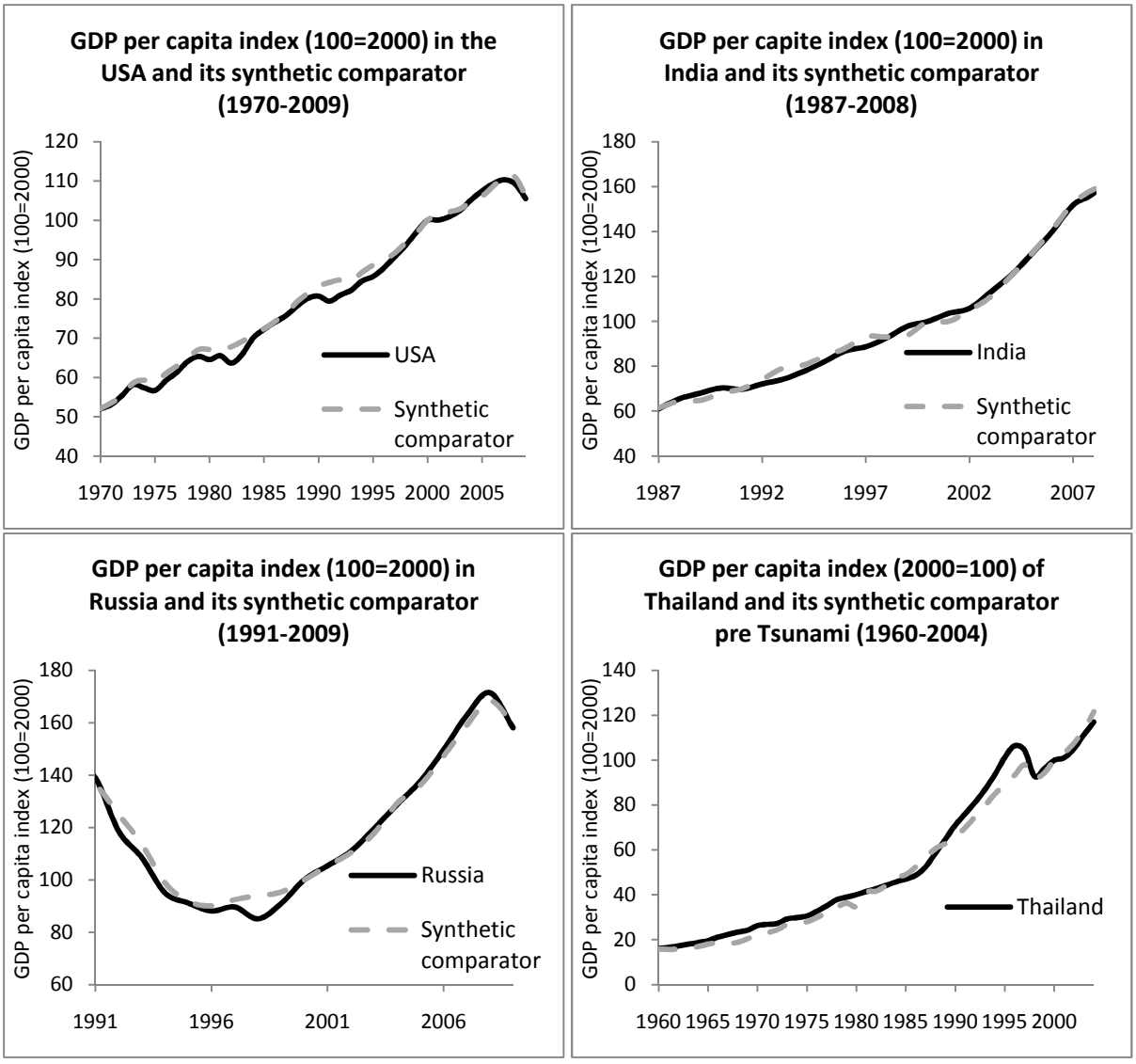
The properties above show that the closer countries are to each other in the export proximity space, the more similar their levels of GDP per capita and economic growth, the more aligned their economic structure, human capital and institutional indicators, the closer they are geographically, the more similar their import structure, and the more likely it is that they are at a similar stage in their economic development process. These characteristics convincingly make

the case that pairs of countries that have a high proximity score tend to have more similar economic capabilities than countries with low levels of proximity.

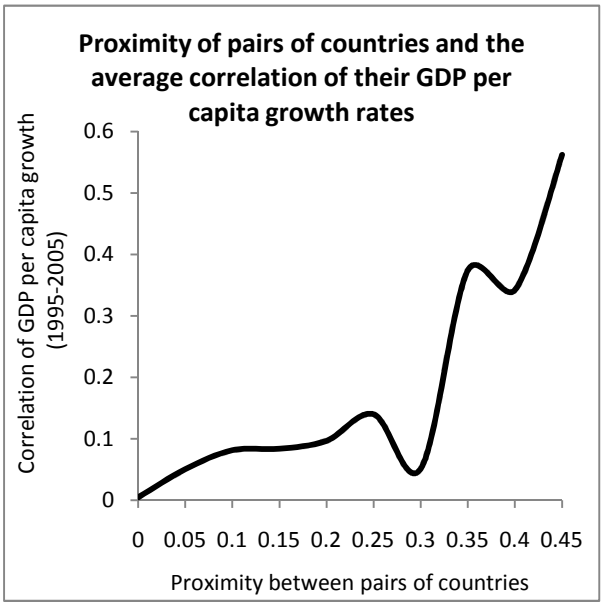
How similar is similar? Export proximity, growth and comparator countries

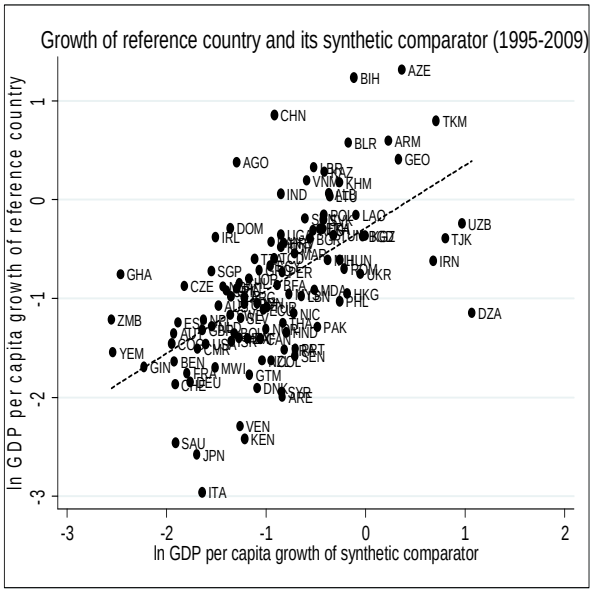
Let us further test this assumption that countries that have similar exports are also good comparators for each other, by comparing their GDP growth rates over time. To do that, for each country of reference we create a synthetic comparator, constructed by simply averaging the GDP per capita index (100 in 1995) of the three countries that are closest to it in the 1995 export proximity space.

We find some remarkable results. The examples below (and in Annex 2), highlight just how similar – in the long term – the growth rate of countries can be to that of their synthetic comparators.



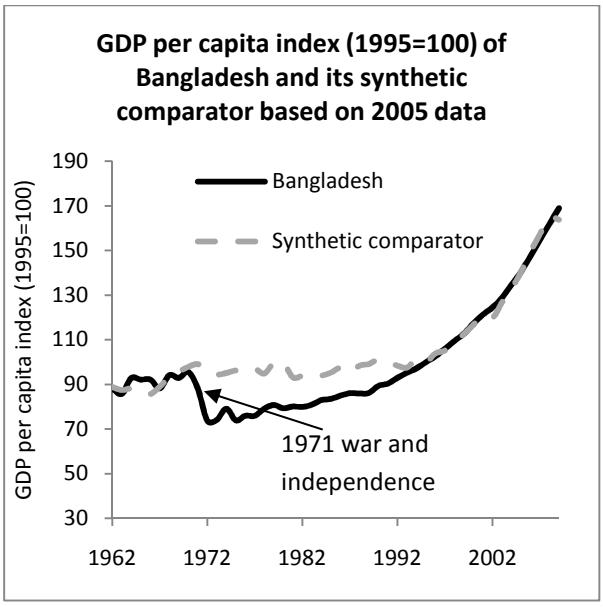
USA's average GDP per capita growth rate between 1970-2009 was 1.855% per year, compared to 1.856% for its synthetic comparator; India's GDP per capita growth rate between 1987-2008 was 4.64%, compared to 4.73% for its synthetic comparator; Russia's average GDP per capita growth rate between 1991-2009 was 0.7%, compared to 0.747% for its synthetic comparator; Thailand's average GDP per capita growth rate between 1960-2004 (pre-tsunami) was 4.67%, compared to 4.76% for its synthetic comparator. These are just a few examples out of many, but they underline one very important point: countries with similar exports can have almost identical growth rates and growth patterns in the long run. Of course this is not always the case, in particular for oil exporters, countries that have experienced domestic shocks (positive or negative), or small economies with highly volatile growth rates (where comparatively small events in the economy can lead to large swings in economic growth). But on average, as the two graphs below reveal, countries with similar exports have similar growth patterns. The closer countries are to each other in the export proximity space, the more correlated on average their growth rates.

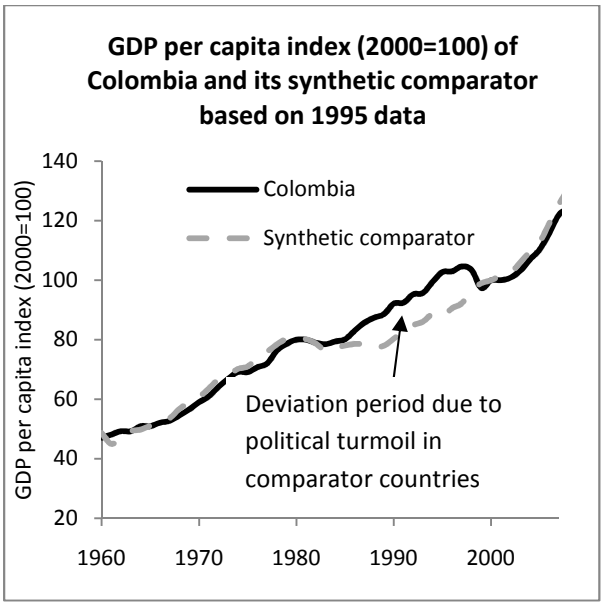


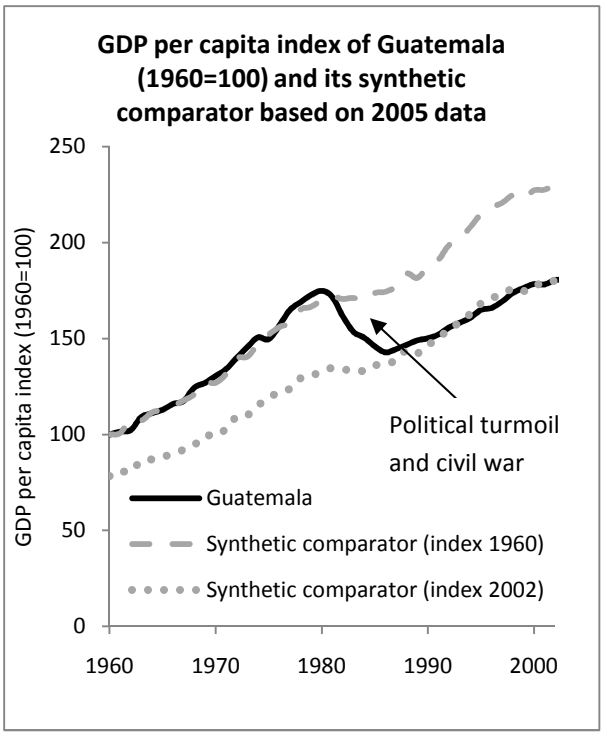


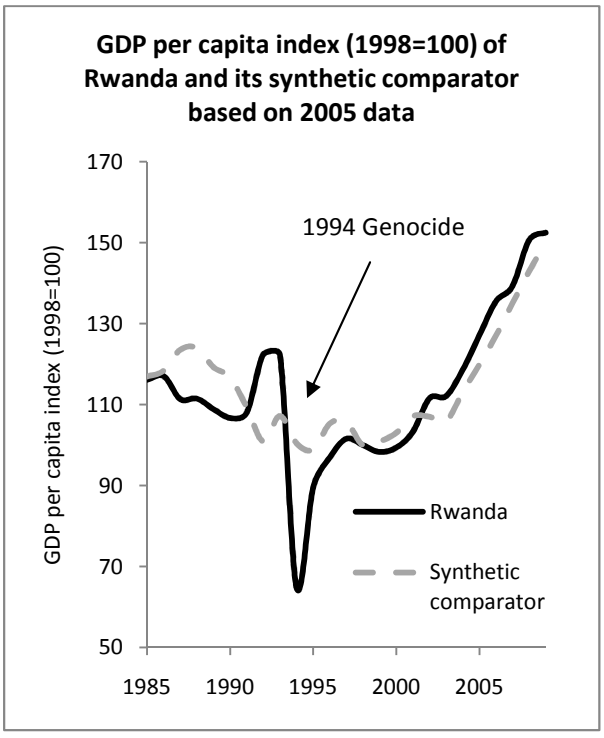
We also find anecdotal evidence suggesting that countries that deviate from their shared growth path tend to converge back towards it in the long run - *by shared growth path we refer to the periods of time when the reference country and the synthetic comparator grow at a similar rate.* Some notable examples, which highlight this fact, are Bangladesh, Colombia, Guatemala and Rwanda; other examples are provided in Annex 3. We use different approaches to construct the synthetic comparators for these countries: for Bangladesh, we construct a synthetic comparator by averaging its 4 closest comparators in the 1995 export proximity space; for Colombia, we average its 5 closest comparators and replace one of its closest comparators, El Salvador, which was conflict ridden, by the next closest comparator; for Guatemala, we average its 3 closest comparators in the 2005 proximity space, but eliminate ex-USSR countries from the sample; lastly, for Rwanda, we create a synthetic comparator by averaging the GDP per capita indexes of its third and fourth closest comparators, namely Ethiopia and Zambia, as the first two (Burundi and Niger), have undergone high levels of political turmoil over the past two decades.

As can be seen in the graphs below, after positive or negative shocks, these countries eventually converge back towards their shared growth path.









The reason we observe such a close relationship between the growth rates of countries that have similar exports is beyond the scope of this study, but there are a number of possible explanations worth exploring in future research: (i) countries that have similar exports have similar growth rates because they compete in the same global product markets and hence are affected in the same way by changes or shocks in those markets; (ii) countries that have similar exports have similar endowment structures (capital, labor, technology, etc), similar balanced

growth paths (this is a result of growth theory), and therefore grow at similar rates in the long run; and (iii) growth is a continuous function of a country's capabilities vector or a proxy thereof (where the capabilities vector is a vector that captures all the capabilities present in a country).

From the perspective of a policy maker, the fact that countries with similar exports tend to have highly correlated growth rates, leads to three important insights: (i) countries that are close to each other in the export proximity space are the most appropriate comparators for policy makers interested in benchmarking a country's economic performance; (ii) policy makers can draw lessons from deviations in the growth patterns of a reference country and its closest comparators (e.g. deviations which could be due to a certain policy interventions); and, (iii) it makes sense to analyze a country's economy (whatever the variable of interest, be it over time, or at the sector level) in comparison to a group of countries with similar characteristics/capabilities rather than independently.

On this last point, in network theory there are a variety of algorithms and approaches that enable researchers to identify communities of nodes (or in the case of the export proximity space, groups of countries). We define a community of nodes as groups of nodes that are densely connected amongst themselves, while being sparsely connected to the rest of the network. In the context of the export proximity network, communities would consist of groups of countries that have very similar exports to each other, but a different export structure to countries outside the group. Based on what we have seen above, we would expect such communities of countries to have very similar economic characteristics.

To illustrate this we run one such algorithm on the 1995 export proximity space[5] and identify a number of relatively cohesive communities of countries:

Communities	**Some common characteristics of these economies (1995)**
{Tadjikistan, Azerbaijan, Uzbekistan, Turkmenistan}	• Neighboring countries in post-USSR transition, with negative growth rates • Size of agriculture, industry and service sectors balanced (about 1/3 each) • Trade deficits • Demographic transition in process, with fertility rates around 3.5
{Mali, Sudan, Ethiopia, Burkina Faso}	• Similar growth rates (around 3.3% GDP per capita growth in 1995) • A large agriculture sector (45% of GDP), combined with a small industrial sector (15% of GDP) • Non-trade intensive countries, with a large trade deficit (9%) • Low domestic savings (10% of GDP) • Very high fertility rates (6.4 children per woman)
{Philippines, Vietnam, Sri Lanka, Tunisia}	• Large services sector (50% of GDP) and comparatively large industrial sectors (29%) • Trade-intensive countries (trade over GDP >80%) running trade deficits (7.8%) • Healthy investment and savings rates (24% and 17% respectively) • Demographic transition in process (with fertility rates around 2.9)
{Central African Republic, Burundi, ex-Zaire, Rwanda}	• Conflict or post-conflict countries with highly volatile growth rates • Largely rural economies (agriculture over GDP>45%) • Negative savings rates (except Zaire, due to mineral exports) • Very high fertility rates (6.2 children per woman)

[5] The tool used is MCODE, a plugin for Cytoscape which enables quick identification of clusters in a network. We use this approach to illustrate a point; the exact methodology is not central to the discussion at hand.

{Myanmar, Cambodia, Laos}	• Largely rural economies (55% of GDP), with very small industrial sectors • Large trade deficits (10% of GDP) • Low investment (13% of GDP) and savings rates
{Russia, Belarus, Ukraine}	• Neighboring countries in post-USSR transition with negative growth rates • Highly industrialized economies (>35% of GDP) • Trade intensive countries (>90% for Belarus and Ukraine) • High savings rates (>20% of GDP) • Very low fertility rates (1.3 per woman), leading to negative population growth
{Thailand, India, China}	• High growth countries (GDP per capita growth around 7.9%) • Very high investment and savings rates (>25%)

The examples above are the result of one such clustering algorithm; communities can vary in size and intensity depending on the how the data is structured, which algorithm is used, and what criteria or specifications the researcher or policy maker uses. The important conclusion from the discussion above is that the export proximity space provides policy makers with various data-driven approaches to identifying comparator countries with similar economic characteristics/capabilities, thereby making economic comparisons more meaningful. The export proximity space provides the extra flexibility of making such cross-country comparisons possible at the sector level and over time: it is possible for example to identify the Asian country which in 1975 had the most similar agro-processing sector to Rwanda today. This has many useful applications for policy makers interested in cross-country economic comparisons, growth diagnostics, industrial policy development, exports analytics, etc.

Proximity controls for counter-factual analysis

In this section we exploit the properties of the export proximity measure to introduce a data-driven method with which policy makers and researchers can infer the impact of a major event or policy on a region and variable of interest. The methodology we put forward is inspired by the synthetic controls methodology introduced by Abadie et al, and builds on a common idea, which is that it is possible to construct a control region of a certain region of interest using a linear combination of other "control" regions.

In the synthetic controls methodology the counterfactual is constructed using the linear combination of control regions that minimizes the difference between the synthetic region and the region of interest on a certain number of aggregate variables. For example, in their paper on the impact of terrorism on economic growth in the Basque region, Abadie et Gardeazabal construct a synthetic Basque region using the linear combination of control regions (in this case other Spanish regions) that minimizes the difference between the synthetic Basque region and the actual Basque region on the following indicators: Real GDP per capita, the investment ratio, population density, sector shares as a percentage of GDP, and human capital indicators (illiteracy rate and primary and secondary education enrollment rates). The authors show that the synthetic Basque region not only does a good job in fitting the values of the Basque region on these economic determinants before the beginning of terrorist activity (this is by construction), but also perfectly matches economic growth in the Basque country for a period of

20 years before the beginning of terrorist activity. While these results and ensuing placebo checks indicate that the constructed synthetic Basque region is a valid control, the methodology is nevertheless based on the assumption that we know which determinants are the most appropriate to match two distinct regions.

Export proximity introduces an alternative way of developing a valid control region using a linear combination of other regions. Rather than selecting which determinants are important - and based on that constructing a synthetic control region that best fits the treatment region on these determinants – we propose using just one measure: how close countries are to each other in the export proximity space. As we have shown in the previous sections, the countries that are closest to each other in the export proximity space have a similar performance on a broad range of indicators. On average, one could say that they are quite similar, and hence we argue they can be used in various ways to construct control regions.

We propose three different methods to construct Proximity Controls for a region of interest. Each method is applied to a case study, respectively: a) the impact of the East Asian Financial crisis on GDP per capita growth in Indonesia; b) the impact of the 1999 coup d'Etat and ensuing political crisis in Ivory Coast on GDP per capita growth; and c) the impact of Kenya's dual domestic crises (election violence in December 2007-2008 and the 2008-2009 drought) on GDP per capita growth in the country[6]. We check the validity of the resulting control regions with three separate tests, which we will detail below.

Method 1 – Proximity Controls by replacement and the impact of the East Asian financial crisis on growth in Indonesia

To measure the impact of the East Asian Financial Crisis on growth in Indonesia (one of the countries that was the hardest hit), we construct a Proximity control of Indonesia using a simple replacement strategy based on the export proximity measure. We then test whether the Proximity Control is a valid control, by (i) checking if this fictive region fits Indonesia on a number of indicators before the beginning of the crises, (ii) checking if the results are very sensitive to changes in the composition of the Proximity Control, and (iii) by running a falsification test.

The East Asian Financial Crisis in Indonesia

The East Asian financial crisis began in July 1997 and its contagion effect raised fears of a global economic meltdown. The crisis began with the devaluation of the Thai baht after it was hit by severe international speculative attacks. The baht devalued swiftly and lost half its value, which led the government to float the currency. As asset prices crashed and debt defaults increased, the resulting panic spread to other crisis countries as lenders withdrew significant credit causing a credit crunch and bankruptcies.

[6] There is no particular reason why one case study was selected over another. The only criteria we had was to find some interesting case studies to apply to the African context, given that Laterite/Econafrique is a Rwanda-based company with an African focus.

Indonesia, South Korea and Thailand were the countries most affected by the crisis. In Thailand, the stock market dropped 75%, per capita income dropped form $8,800 to $8,300 between 1997 and 2005 and politically led to the ouster of its Prime Minister. South Korea faced numerous corporate bankruptcies, led to a weakening of the South Korean won and its national debt-to-GDP ratio more than doubled. In Indonesia, the rupiah also came under speculative attacks, leading to a strong recession and the eventual departure of President Suharto.

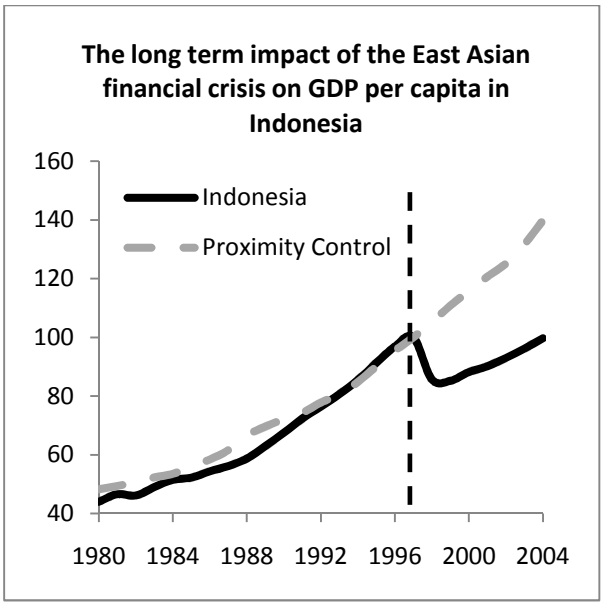
By 1999, there were signs that most of the countries had begun to recover from the crisis.

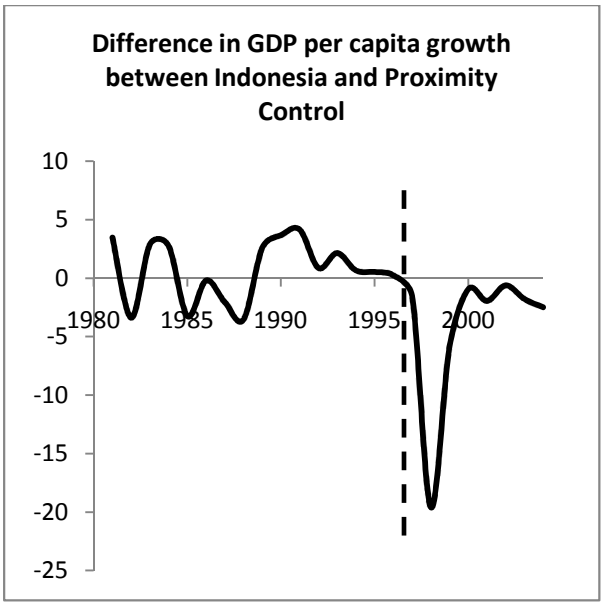
Constructing a Proximity Control for Indonesia

To construct a Proximity Control for Indonesia we use a simple replacement strategy. We replace Indonesia by the average indexed GDP per capita of its three closest comparators in the 1995 export proximity space, eliminating from the sample countries in the immediate vicinity of Indonesia that were also severely affected by the crisis, such as Thailand, Korea, the Philippines, Laos, Vietnam, Singapore, Hong Kong, Taiwan, and Malaysia. As can be seen in the table below, the three closest comparators to Indonesia in 1995 - that were comparatively less affected by the crisis - are: Portugal, China and India.

Country	Closest comparators	Proximity
Indonesia	Portugal	0.360143
	China	0.325648
	India	0.322997

The graph below, which focuses on the 1980-2004 period, shows that a simple average of Portugal, China and India's indexed GDP per capita (index=100 in 1997) is quite an accurate predictor of Indonesia's GDP per capita growth during the seventeen years preceding the Asian Financial Crisis (1980-1997), indicating that the constructed Proximity Control could indeed be a good control region for Indonesia. The graph also reveals that the East Asian financial crisis seems to have impacted Indonesia in two ways: (i) it shaved off an approximate 23.1% of Indonesia's potential GDP per capita in the immediate aftermath of the crisis (1997-1999); and (ii) Indonesia settled on a slower growth path thereafter (between 1999-2004). While Indonesia grew slightly faster than the Proximity Control during the five years preceding the crisis (5.55% per capita growth during 1992-1997 vs 5.07%), it grew one and a half percentage points slower in the four years after the crisis had settled (3.19% between 2000-2004, vs 4.76%). By 2004 Indonesia's GDP per capita was only 71.3% of what it could have been (i.e this is equivalent to a GDP per capita loss of 28.7% or about US$103bn when extrapolated).





Testing the validity of the Proximity Control

We conduct three tests to check the validity of the control: a fitness test, the "bow-tie test" and a falsification test.

Fitness test

First we check whether the Proximity Control closely fits the treatment region (Indonesia) on a number of key economic indicators before the beginning of the financial crisis. The more similar the treatment and control regions are on structural economic indicators such as investment, savings, trade and the sectoral distribution of the economy, the better the control. A perfect control region would match the treatment region exactly on each and every single indicator.

Indicator (averages between 1992-1997)	Indonesia	Proximity Control
Gross fixed capital formation (% of GDP)	27.7	26.9
Gross domestic savings (% of GDP)	31.7	27.1
Household final consumption expenditure (% of GDP)	60.3	58.5
Agriculture (% of GDP)	17.3	17.8
Industry (% of GDP)	41.6	33.8
Manufacturing (% of GDP)	*24.0*	*22.9*
Services (% of GDP)	41.1	48.4
Exports of goods and services (% of GDP)	26.9	18.9
Imports of goods and services (% of GDP)	26.1	21.1
External balance on goods and services (% of GDP)	0.8	-2.3

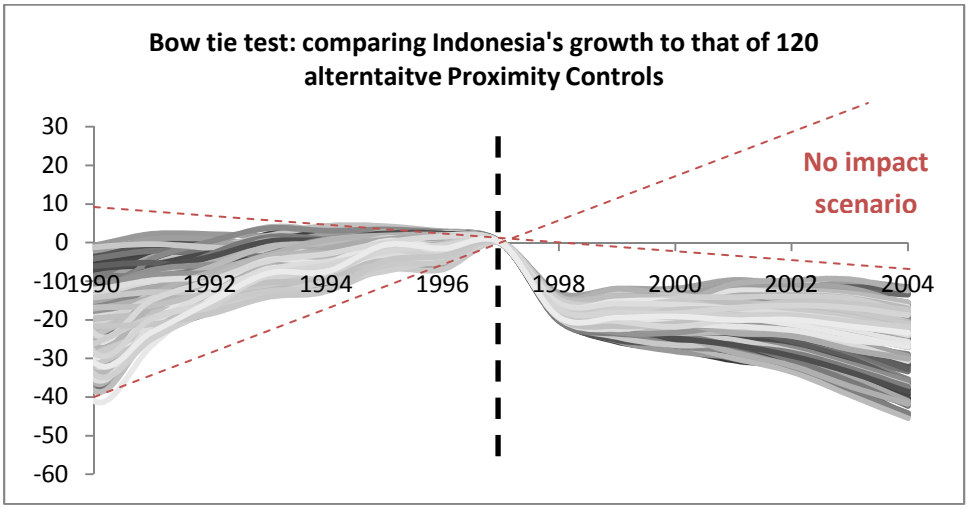
As we can see in the table above, there are some strong similarities between Indonesia and the Proximity Control region. The similarity between the two regions is based on high investment and savings rates, similar household consumption rates, as well as a similar share of agriculture

and manufacturing over GDP. Indonesia is more industry and trade intensive than the Proximity Control, but given the similar manufacturing rates the difference in industry and exports could be attributable to Indonesia's petroleum sector. The decision of whether the control is similar enough to the treatment region is left to the researcher or policy maker. While calculations show that there are other combinations of three countries that would be better fits to Indonesia on the selected indicators, the authors believe that this Proximity Control is good enough, given that key indicators for growth such as investment, savings and the consumption rates are almost identical.

Bow-tie test

We call the second test the bow tie test. The "bow-tie" test consists in changing the composition of the Proximity Control using other countries close to Indonesia in the 1995 export proximity space. Rather than selecting only the three closest countries to Indonesia, we try all the different combinations of triplets out of Indonesia's ten closest countries (excluding the countries that were severely affected by the crisis) and compare their growth performance to that of Indonesia. If the Proximity Control we have constructed is valid, then changes in the composition of the Proximity Control should yield similar results. If this is not the case, then the observed impact could be due to the idiosyncratic growth performance of a single country in the Proximity Control region, which would mean the control is not valid.

The ten countries closest to Indonesia in the export proximity space that were not severely affected by the East Asia crisis are: Portugal, China, India, Turkey, Sri Lanka, Romania, Pakistan, Morocco, Croatia, and Panama; and there are 120 different possible combinations of the latter. Each line in the graph below represents the difference between the GDP per capita index of Indonesia and one of the 120 newly constructed Proximity Controls. According to the graph, Indonesia had a lower GDP per capita index than most Proximity Controls in the period leading to the crisis, which means that it grew quicker during that period as it started from a lower base to end-up at the same point in 1997. However, had there been no impact in 1997, Indonesia should have continued on its steady growth path (7% on average in GDP per capita terms between 1990-1997) and surpassed the GDP per capita indexes of its Proximity Controls. Instead its GDP per capita index plummeted compared to all the alternative controls after 1997.



If all alternative controls were valid estimates of Indonesia's potential GDP per capita had the East Asia financial crisis not happened, our estimates of the impact of the crisis would range between a GDP per capita loss of about 23% and 42% by 2004 (with an average of 31.5%) – this is taking into account the fact that Indonesia was growing faster than almost all alternative Proximity Controls before the crisis. This is in line with our estimate of 28.7% and suggests that the results we obtain are not unique to the Proximity Control composed of the {China, India, Portugal} region.

The reason we call this the bow-tie test is because the resulting graph resembles a bow-tie; the further we move away from the anchor year (1997), the greater the difference on average between GDP per capita indexes. If the difference between the treated and control regions are similar before and after an event, then we get a symmetrical bow-tie. If there has been an impact, and the difference between the treated and control regions is significant after the event, then we get a distorted bow-tie. In the case of Indonesia we clearly observe a distorted bow-tie.

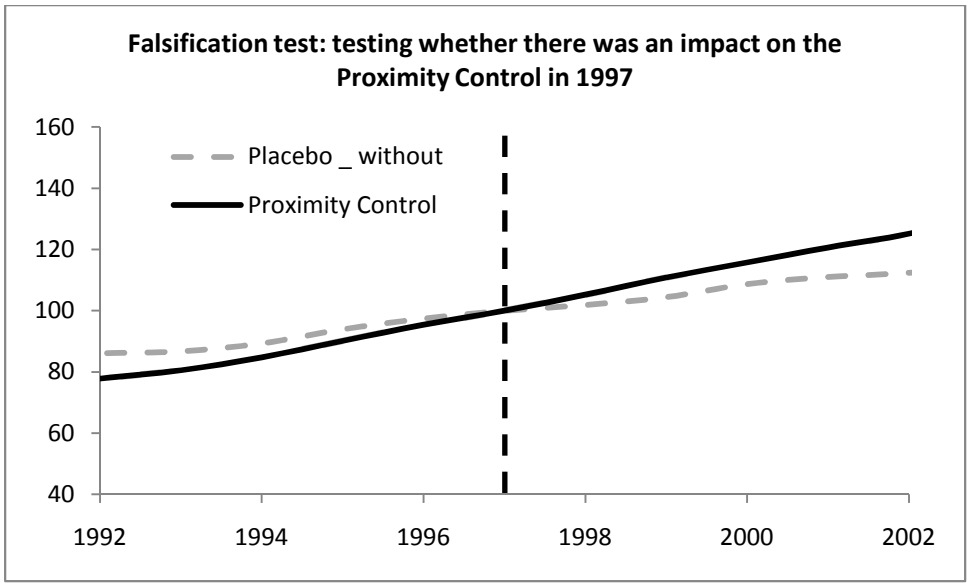
Falsification test

By definition a control region can only be a valid control if it did not experience the treatment itself (the East Asian financial crisis). Given how inter-connected the global economy is, no country escaped the effects of the East Asian financial crisis, so it is impossible to find a valid control (based on a combination of other countries) for what would have happened in Indonesia had the Asian financial crisis not occurred at all. What we can attempt to measure though, is the difference between the impact of the crisis on Indonesia, which was at the center of the storm, and the impact of the crisis on other countries which were at the periphery.

To test the validity of our Proximity Control, we create a control of the Proximity Control – which we call a Placebo region – using the same methodology. If the Proximity Control were an accurate control region for Indonesia, before and after the financial crisis, then we would expect to observe no difference between the Proximity Control and the Placebo region during the period under consideration, indicating that the Proximity Control was not disproportionately affected by the East Asian financial crisis.

To construct the Placebo region we create an aggregate {China, India, Portugal} region and measure its export proximity to other countries in the 1995 export proximity space. As in the case of Indonesia and its Proximity Control, the Placebo region will consist of the three closest countries to the {China, India, Portugal} region that were not severely affected by the East Asian financial crisis (Indonesia, Thailand, Korea, the Philippines, Laos, Vietnam, Singapore, Hong Kong, Taiwan, and Malaysia). We also take Turkey out of the sample, which was affected by its own financial crisis between 1999-2001. As can be seen in the table below the three closest countries to the {China, India, Portugal} region were the Czech Republic, Italy and Poland.

Region	Closest comparators	Proximity
{China, India, Portugal}	Czech Republic	0.39497
	Italy	0.375717
	Poland	0.299635



While the resulting Placebo region is not a good predictor of growth in the Proximity Control – on average the Proximity Control region grew faster between 1992-2002 – the graph above suggest that the East Asian financial crisis did not disproportionately affect either regions. Both regions grew at approximately the same rate before and after the 1997 crisis and the growth differential between them was approximately the same (2.43% per capita per year in the five years before the crisis; 2.38% in the five years after).

Average growth (% per capita)	Placebo region	Proximity Control	Difference
1992-1997	2.65%	5.07%	2.43%
1998-2004	2.53%	4.91%	2.38%

The Proximity Control region therefore appears to be a valid control for Indonesia: (i) the Proximity Control region fits Indonesia quite closely on a number of key economic variables before the East Asian financial crisis; (ii) changes in the composition of the Proximity Control region yield similar results, which means the observed impact is not due to the idiosyncratic growth patterns of a single country; and (iii) the Proximity Control was not disproportionately affected by the crisis.

Method 2 – Proximity Controls using weights and the impact of the 1999-2009 political crisis on growth in Ivory Coast

To measure the impact of Ivory Coast's political crisis on growth in the country between 1999-2009), we construct a Proximity control of Ivory Coast using a simple weighting strategy based on the export proximity measure. As in the case of Indonesia, we test whether the Proximity Control is a valid control, by (i) checking if the latter fits Ivory Coast on a number of indicators

before the beginning of the crises, (ii) checking if the results are very sensitive to changes in the composition of the Proximity Control, and (iii) by running a falsification test.

Ivory Coast's 1999-2009 political crisis

The event that arguably triggered Ivory Coast's recent political and economic downfall was the 1999 coup d'Etat, which removed former president Henri Bedie from power. The political situation was already tense, in particular because of the north-south polarization of the country and the repeated attempts by the establishment to block Northerners from voting in national elections due to their lack of "Ivoireness". But it was this event that really marked the beginning of Ivory Coast's downfall.

Since the coup d'Etat political unrest has been a constant. After the 1999 coup d'Etat elections were organized and the now deposed President Laurant Gbagbo took power. Disenfranchised and discriminated against, the North resorted to an armed rebellion in 2002 by forming *Force Nouvelle* (FN) an armed group of rebels from a diverse variety of backgrounds with a common goal to end the marginalization and reintegrate the North in national politics. The FN launched as a troop mutiny in 2002 and then escalated into a full-scale rebellion in which thousands were killed. The short civil war led to a tenuous peace plan in 2005 that permitted Outtara to run for the highest office but had split the country between a rebel-held North and a government-held South.

These sentiments were the background for the long-postponed Presidential election, which took place in late November 2010 and led to the most recent political crisis in the country. In this mini-case study we focus on the economic toll of Ivory Coast's political unrest before the most recent election crisis, namely the 1999-2009 period.

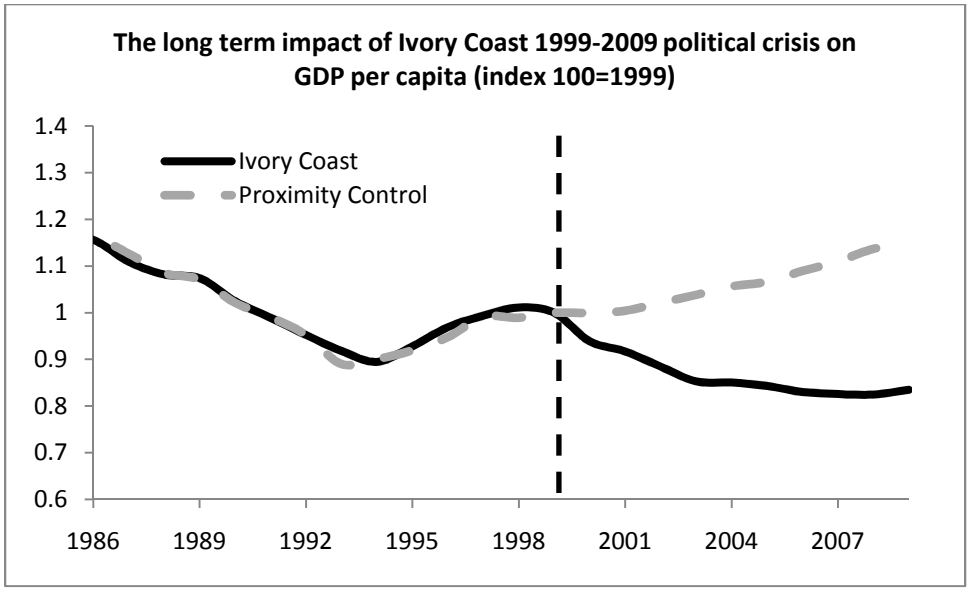
Constructing a Proximity Control for Ivory Coast

We use a very simple weighting strategy to construct a Proximity Control for Ivory Coast. In the previous example we constructed the Proximity Control by taking a simple average of the three closest countries to Indonesia in the 1995 export proximity space that were not severely affected by the crisis. Here we weight each country by its proximity score. The greater the proximity score of a country, the greater its contribution to the Proximity Control. Again we consider only the three closest countries to the treated region - Ivory Coast – in the 1995 export proximity space. These are: Cameroon, Ghana and Togo.

Country	Proximity to Ivory Coast	Contribution to Proximity Control
Cameroon	0.266393	37.39%
Ghana	0.241803	33.94%
Togo	0.204225	28.67%

As the graph below reveals, the {Cameroon (37.39%), Ghana (33.94%), Togo (28.67%)} region almost perfectly matches Ivory Coast's growth patterns between 1985-1999 (a 14 year period). However, Ivory Coast and the Proximity Control start to diverge significantly after 1999, which

confirms that 1999 was indeed a turning point for the Ivorian economy. Based on this Proximity Control Ivory Coast lost about 26.8% of its potential GDP per capita between 1999-2009 (or about $6.2bn in current US dollars when extrapolated).



Testing the validity of the Proximity Control

Again, we conduct three tests to check the validity of the control: a fitness test, the "bow-tie test" and a falsification test.

Fitness test

As we did in the case of Indonesia, we first check whether the Proximity Control closely fits the treatment region (Ivory Coast) on a number of key economic indicators before the 1999 coup d'Etat. As a point of comparison we also create a Synthetic Ivory Coast using the synthetic controls methodology developed by Abadie (2003). The synthetic control is constructed: (i) using a sample of 93 countries; (ii) using as predictor variables: gross fixed capital formation, gross domestic savings; agriculture, industry and services over GDP; and exports and imports over GDP; and (iii) using as reference period the 1990-1998 period. The resulting Synthetic Control region is composed of Sudan (29.8%), Kenya (25.5%), Senegal (22.7%), Swaziland (14.4%), Panama (3%), Bulgaria (2%), Cameroon (1.8%), Ethiopia (0.4%) and Fiji (0.1%).

Indicator (averages between 1992-1997)	Ivory Coast	Proximity Control	Synthetic Control
Gross fixed capital formation (% of GDP)	12.0	16.1	15.9
Gross domestic savings (% of GDP)	19.2	11.0	9.8
Household final consumption expenditure (% of GDP)	68.6	77.2	77.8
Agriculture (% of GDP)	26.7	34.6	27.7

Industry (% of GDP)	21.7	26.3	21.7
Manufacturing (% of GDP)	*17.3*	*13.6*	15.1
Services (% of GDP)	51.7	39.1	50.6
Exports of goods and services (% of GDP)	37.7	24.9	28.4
Imports of goods and services (% of GDP)	30.5	30.4	34.8
External balance on goods and services (% of GDP)	7.1	-5.5	-6.4

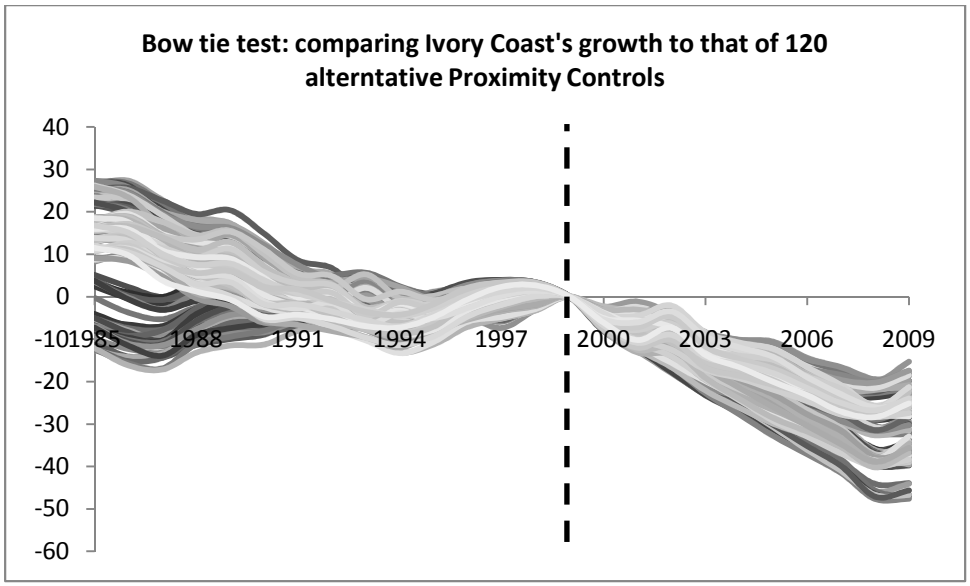
As can be seen in the table above, the fitness test for Ivory Coast fails. The Proximity Control and Ivory Coast differ on many characteristics, in particular: Ivory Coast is more export oriented (largely due to the cocoa and coffee industries) and during 1992-1997 ran a big trade surplus compared to the Proximity Control region's trade deficit; its economy is more services based with a smaller agricultural sector than in the Proximity Control; and savings during the period under consideration were higher than investments, while in the Proximity Control the opposite was true.

However, when we compare the Proximity Control to the Synthetic Control region, we find that both regions are quite similar on many of those same characteristics: they have similar levels of investment, domestic savings and household consumption, and both regions run a trade deficit. The synthetic control region by construction minimizes the difference between the treated region and the control regions on the selected economic predictors before the occurrence of an event. This means that the resulting synthetic control region is the linear combination of countries that best fits the treated region on the variables under consideration.

This leads to two insights: (i) the Proximity Control is not very different from the linear combination that best matches Ivory Coast on the selected variables; and (ii) there does not seem to be a linear combination of countries that would perfectly match Ivory Coast on the selected variables. Hence either there is no good control for Ivory Coast (based on a linear combination of other countries) or these are not the right variables to judge whether a control region is a good enough fit for Ivory Coast or not.

Bow-tie test

While the fitness test does not succeed, the bow-tie test does. As in the case of Indonesia, we create a 120 different alternative Proximity Controls by changing the composition of the countries in the Proximity Control to include different combinations of Ivory Coast's 10 closest countries in the 1995 export proximity space. Ivory Coast's 10 closest comparators are: Cameroon, Ghana, Togo, Ecuador, Paraguay, Malawi, Benin, Senegal, Kenya and Madagascar. The only difference this time around is that all countries are weighted by their respective proximity to Ivory Coast. Each line in the graph below represents the difference in the indexed GDP per capita (index=100 in 1999) of Ivory Coast and each of the alternative Proximity Controls.



The graph reveals that while Ivory Coast cannot be distinguished from the alternative Proximity Controls during the 1985-1999 period, it is a clear and undisputable outlier after 1999. Based on these alternative controls, estimates of Ivory Coast's lost GDP per capita between 1999 and 2009 range from -15.3% to -47.5%, with an average of about -30.2%. This is not very far off the -26.8% estimate resulting from the preferred Proximity Control (the standard deviation of these estimates is 7.7%).

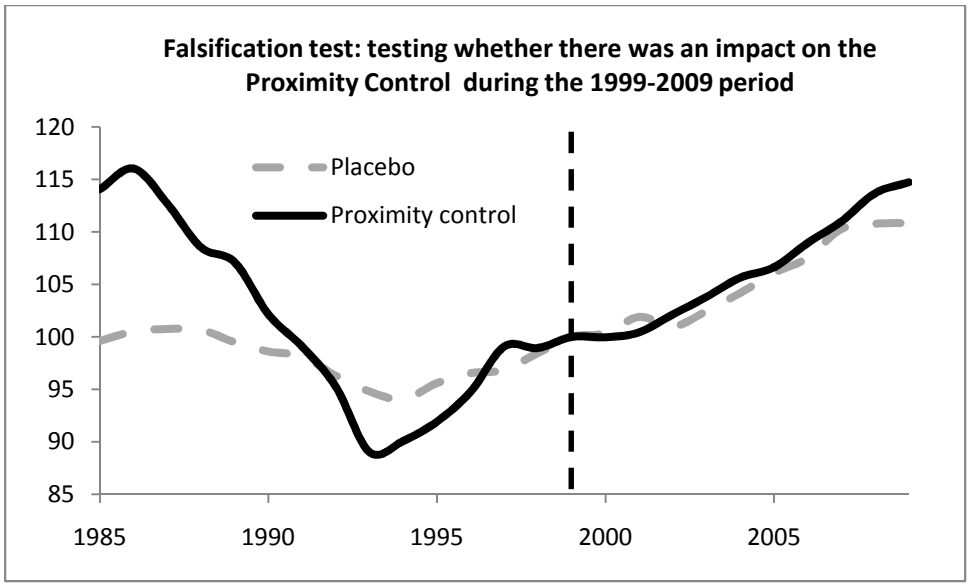
Falsification test

To test whether the Proximity Control itself was not affected by any major events during the period under consideration we construct a control of the aggregate {Cameroon (37.39%), Ghana (33.94%), Togo (28.67%)} region, which we call the Placebo region. To construct the Placebo Region we calculate the proximity of the aggregated Proximity Control region to that of other countries. The closest three countries in the export proximity space to the {Cameroon (37.39%), Ghana (33.94%), Togo (28.67%)} region are Benin, Kenya and Senegal[7].

Country	Proximity	Contribution to Placebo region
Benin	0.196875	34.6%
Kenya	0.19403	34.1%
Senegal	0.178125	31.3%

As can be seen in the graph below, the resulting Placebo region closely matches growth in the Proximity Control during the 1997-2008 period, suggesting that the Proximity Control region was not subject to any major economic events during the between 1997-2009.

[7] Note that we have taken Malawi out of the sample as Malawi's growth during the period under consideration was very volatile ranging between -10% in 1994 to 16% in 1995 and -5% in 2001.



In balance the Proximity Control appears to be a valid control for Ivory Coast as (i) it predicts growth in Ivory Coast 14 consecutive years prior to the crisis; (ii) changes in the composition of the Proximity Control region yield similar results; and (iii) the Proximity Control was not disproportionately affected by the crisis itself. However the question of whether this is a valid control for Ivory Coast is less clear cut than in the case of Indonesia, where there was a much better fit between the treatment and control region on the selected economic variables

Method 3 – Proximity Controls using an elimination strategy and the impact of Kenya's dual domestic crisis (the 2007 election crisis, and the 2008-2009 drought) on growth

To measure the impact of Kenya's December 2007 election crisis and the ensuing 2008-2009 drought, we construct a Proximity control of Kenya using an elimination strategy. As in the case of Indonesia and Ivory Coast, we test whether the Proximity Control is a valid control, by (i) checking if the latter fits Kenya on a number of indicators before the beginning of the dual domestic crises, (ii) checking if the results are very sensitive to changes in the composition of the Proximity Control, and (iii) by running a falsification test.

Kenya's December 2007 election and 2008-2009 drought

Kenya experienced two major domestic crises in 2008-2009: the major outbreak of violence that followed the December 2007 presidential elections and the 2008-2009 drought. The December 2007-2008 Kenyan crisis refers to the political, economic, and humanitarian crisis that erupted in Kenya after incumbent President Mwai Kibai was declared the winner of the presidential election held on December 27, 2007 that pitted him against the populist opposition candidate Raila Odinga. With widespread reports of ballot irregularities from polling stations, the now discredited Electoral Commission of Kenya declared Kibaki the winner for another five-year term and was hastily sworn in on December 30, 2007. The nation erupted. Western Kenya and the Nairobi slums bore the brunt of the violent clashes. Planned nationwide protests were frequently

disrupted in the ensuing days while several attempts at international mediation were thwarted. Talks resumed and were finally concluded on February 28th through an agreement that came to be known as the National Accord and Reconciliation Act, which ushered in a power-sharing agreement. It was reported that by the end of the violence, between 800-1500 people had died in the clashes and between 180,000 – 250,000 people displaced

The 5-week violence in Kenya had serious economic ramifications. There is no clear estimate of the economic cost of the Kenyan election crisis, but rough estimates note a 2% decline in GDP growth in 2008 compared with 7.1% growth in 2007. However, this decline in GDP growth however also reflects other factors: (i) a major drought in 2008-2009 which heavily impacted livestock and agricultural production; (ii) the global food crisis in 2008; and (iii) the global financial crisis which started in the summer of 2008.

So how can one distinguish between the impact of Kenya's domestic crises and the global crises which all had an impact on its economy? While it is difficult at the aggregate level to disentangle the effects of the election violence from Kenya's drought (which were both major domestic crises that affected the economy overall and in particular the country's rural areas), we will show that it is possible using the Proximity controls methodology to isolate the impact of these dual domestic crises from the effects of external crises such as the global financial crises and the global food crises.

Constructing a Proximity Control for Kenya

In the Indonesia and Ivory Coast examples we selected the three countries that were closest to the treated region in the export proximity space and constructed a Proximity control by simply averaging out or weighting the GDP per capita indexes of the latter. Here, in addition to the condition of export similarity, we add one condition which is that the resulting Proximity control region should also be a significant exporter of at least 70% of Kenya's significant export products. What matters here is not the number (it could have been 50%, 80% or even 100%, depending on the objectives of the researcher), but the fact that one could argue that a control region is not really representative of a reference country if it does not account for at least a minimum share of that country's exports.

One way of constructing such a Proximity control for Kenya would be to simply go down the list of countries with the highest levels of Proximity to Kenya and add them to the Proximity control region until the 70% threshold is achieved. But this would be inefficient. Assume for example that the two closest countries to Kenya in the 2005 export proximity space, namely Guatemala and Syria, had exactly the same exports. Then they would both be very close to Kenya in the export proximity space, yet including Syria on top of Guatemala in the Proximity control region would not increase the latter's coverage of Kenya's exports.

To avoid this, we propose an eliminatory approach which consists of the following iterative steps:

(i) Identify the closest country to Kenya in the export proximity space, include it in the Proximity control region and measure what share of Kenya's significant export products are now accounted for;

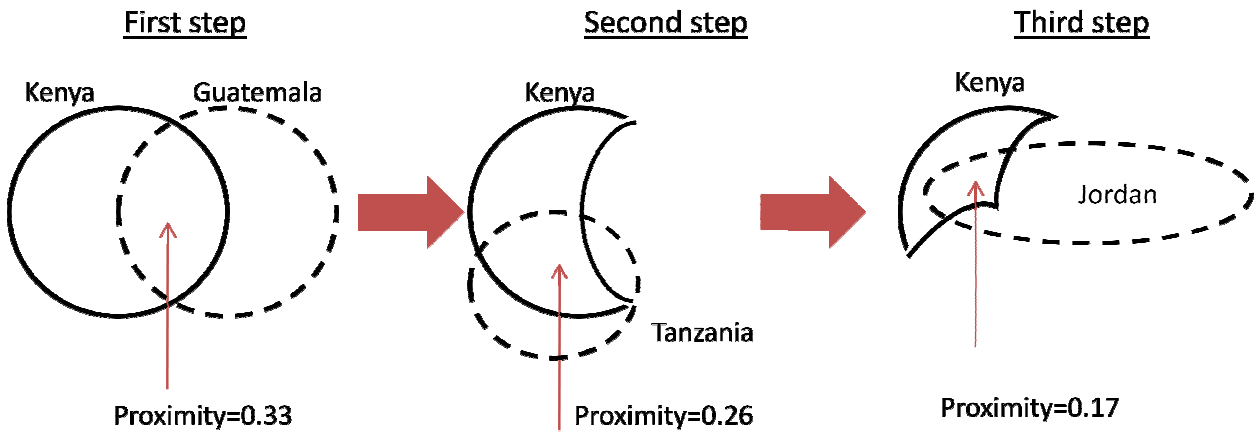
(ii) If less than 70%, eliminate from the sample all the products that this country has in common with Kenya and recalculate export proximity levels based on this narrowed down sample;

(iii) Again, identify the country that is closest to Kenya in the narrowed down export proximity space, include it in the Proximity control region and measure what share of Kenya's significant export products are now accounted for;

(iv) If less than 70%, eliminate from the sample all the products that this country has in common with Kenya and recalculate export proximity levels;

(v) Repeat these steps until 70% of Kenya's export base has been covered.

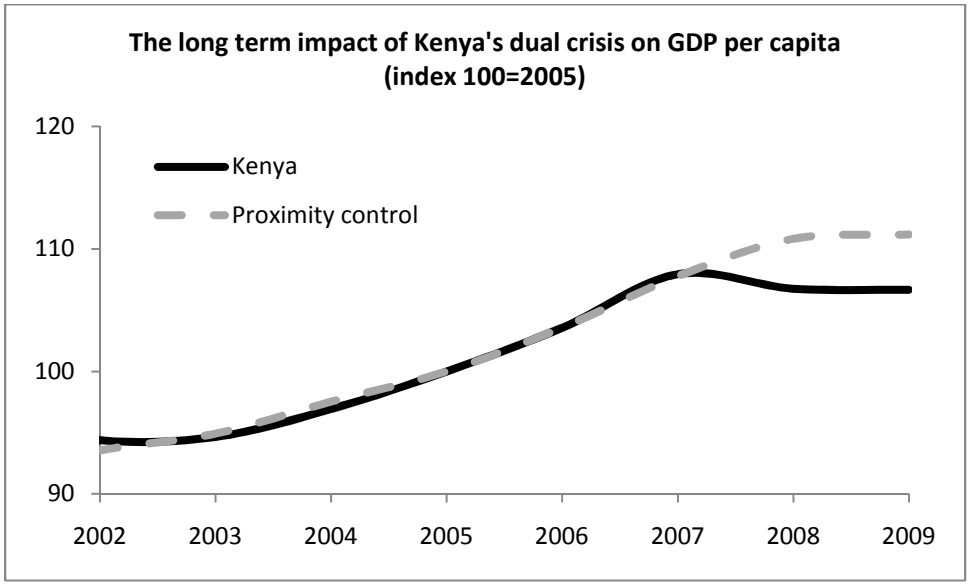
We can explain this more effectively with an illustration:



As can be seen in the figure above, to construct a Proximity Control for Kenya we first start by identifying Kenya's closest comparator in the 2005 export proximity space, namely Guatemala. We then eliminate all the products that Guatemela and Kenya have in common from the sample. We then recalculate Kenya's export proximity levels with other countries and find that the closest comparator to Kenya in this narrowed down space is Tanzania. If we now eliminate all the products Tanzania has in common with Kenya, we find that Jordan is next closest country to Kenya, etc. After repeating this process five times, we get the following Proximity Control for Kenya, which accounts for 72.6% of its exports:

Country	Adjusted Proximity to Kenya	Share of Kenya's exports accounted for	Contribution to Proximity Control
Guatemala	0.331081	33.1%	45.6%
Tanzania	0.260606	17.4%	24.0%
Jordan	0.172481	12.0%	16.6%
Nepal	0.144404	6.6%	9.1%
Uganda	0.109649	3.4%	4.7%
	Totals	**72.6%**	**100%**

The resulting Proximity Control region is an excellent predictor of growth in Kenya in the 5 years leading to the crisis (2002-2007). As in the case of Indonesia and Ivory Coast, we see that Kenya and its proximity control diverge after the beginning of the election crisis and the ensuing drought (2008-2009). Based on this Proximity Control, we estimate that by 2009 Kenya had lost 4.06% of its potential GDP per capita (a loss of approximately 1.25bn in current US dollars when extrapolated).



Testing the validity of the Proximity Control

In line with our approach in the Indonesia and Ivory Coast cases, we conduct three tests to check the validity of the control: a fitness test, the "bow-tie test" and a falsification test.

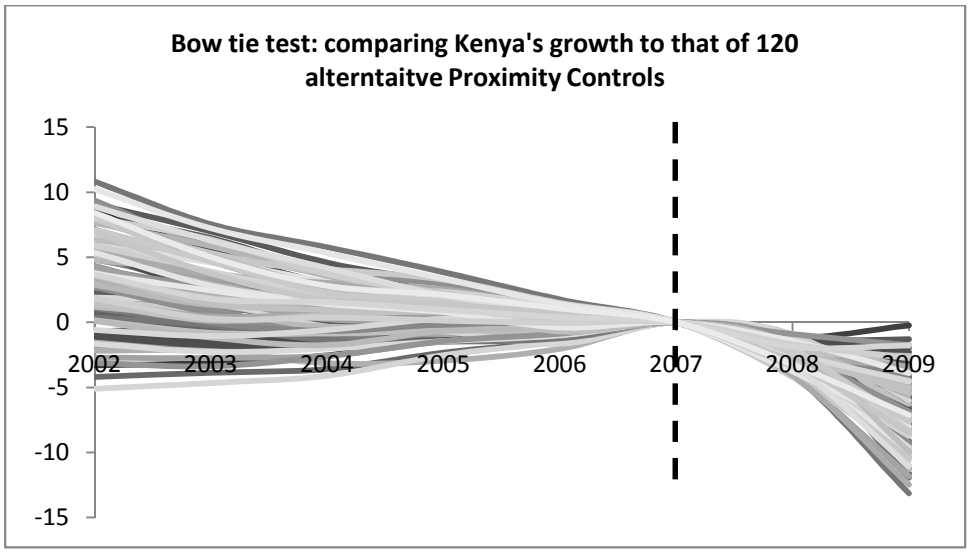
Fitness test

As can be seen in the table below the similarity between Kenya and its Proximity Control is based on: similar investment and household consumption rates, a low savings rate, a large service sector and a significantly negative trade balance. The main difference between Kenya and the Proximity is the relative size of the agriculture and industry sectors; Kenya is more agriculture intensive, whereas the Proximity Control is more industry intensive. The country that is driving the wedge between Kenya and the Proximity Control region in Jordan, which is an outlier: Jordan has a negative savings rate, an almost non-existent agricultural sector and a trade deficit of about 30%. If Jordan is eliminated from the sample some of the indicators realign. Despite these differences however Kenya and the Proximity Control seem to have quite a similar economic reality, with modest investment rates, low savings rates, a large service sector and a large trade deficit.

Indicator (averages between 2002-2007)	Kenya	Proximity Control
Gross fixed capital formation (% of GDP)	17.8	20.9
Gross domestic savings (% of GDP)	9.4	4.8
Household final consumption expenditure (% of GDP)	72.9	82.2
Agriculture (% of GDP)	26.7	16.9
Industry (% of GDP)	17.6	26.4
Manufacturing (% of GDP)	11.0	*16.1*
Services (% of GDP)	55.7	56.7
Exports of goods and services (% of GDP)	26.1	28.9
Imports of goods and services (% of GDP)	33.7	47.0
External balance on goods and services (% of GDP)	-7.7	-18.0

Bow-tie test

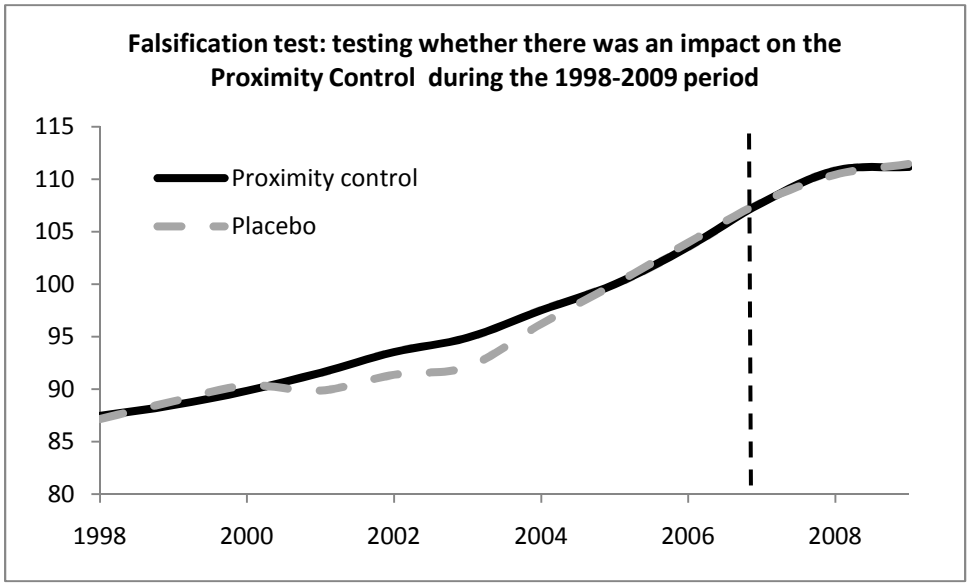
The-bow tie test shows that all alternative Proximity Controls lead to the same conclusion[8]. While it is difficult to distinguish Kenya from the rest of the Proximity Controls before the beginning of the December 2007 crisis, Kenya performs worse than every single alternative Proximity Control in the 2007-2009 period, bar one. When taking prior growth differentials into account estimates of impact range from a GDP loss of 0% by 2009 to a maximum of 16.1%, the average being 7.2%. This is not very far off our estimate of 4.06% (which is within 1 standard deviation).



[8] For computational purposes, we construct the alternative Proximity Controls using a simple replacement strategy.

Falsification test

To test whether the Proximity Control also experienced an impact in 2007, we construct a Placebo region of the aggregate Proximity Control region. We use a simple replacement strategy to construct the Placebo region. We replace the aggregate {Guatemala(45%), Tanzania (24%), Jordan(16.6%), Nepal (9.1%), Uganda(4.65%)} by its three closest comparators in the export proximity space, namely: Syria, Jordan and Pakistan. As we can see in the graph below, the Proximity Control region does not seem to have been disproportionately impacted by any economic event between 1998-2009.



The Proximity Control region therefore appears to be a valid control for Kenya: (i) in broad terms the Proximity Control has similar economic characteristics to Kenya; (ii) changes in the composition of the Proximity Control region yield similar results, which means the observed impact is not due to the idiosyncratic growth patterns of a single country; and (iii) the Proximity Control was not itself affected by Kenya's 2007 election crisis and the ensuing drought.

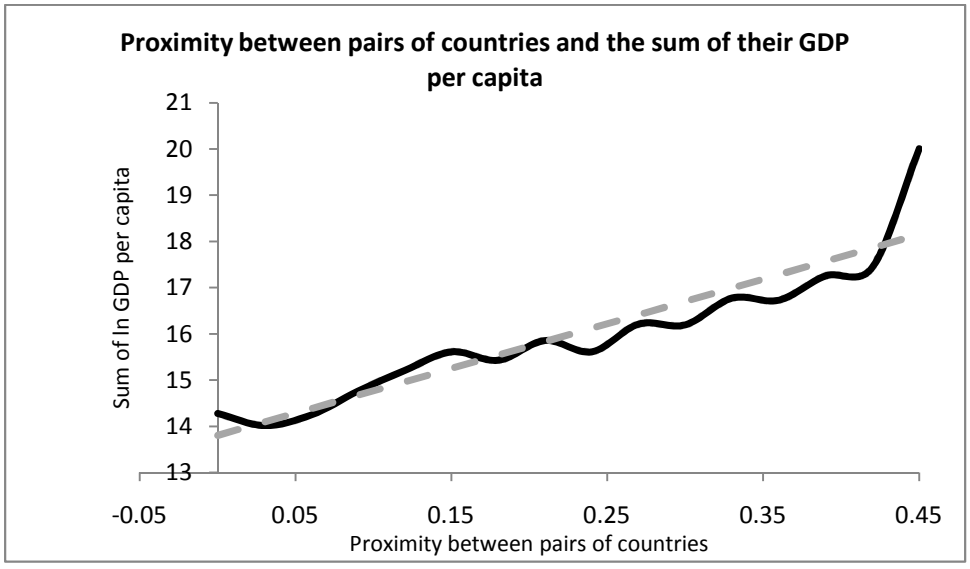
Closing remarks

What the examples above highlight is that a country's position in the export proximity space matters for its economic growth and that we can use export proximity to identify appropriate comparator countries for a reference country. This can be done at the country level, at the sector level and over time.

Identifying comparator countries at the country or sector levels is useful for policy-makers as it enables them to: (i) benchmark their country's performance against countries with similar capabilities and export structures; (ii) identify direct competitors, as comparator countries compete in similar product markets; and (iii) learn from the successes or failures of economic policies in countries with similar capabilities, rather than countries with a very different set of economic conditions and capabilities.

Another way to use the export proximity measure is to find comparators over time. This would enable a policy maker to answer a question such as: which country in 1960 had the most similar export structure to a certain reference country today? In particular, it would enable policy makers to identify what Lin and Monga (2010) call "compass" economies. Lin and Monga argue that countries that have experienced high rates of growth throughout the 20th century, have done so by emulating the economic success of wealthier countries with similar endowments (their *compass* economy). The export proximity space provides one way of going about that.

Of course there are also limitations to the techniques we have introduced. The first limitation is that this method works better for wealthier countries. Data shows somewhat counter-intuitively that the wealthier a pair of countries - and the more complex their economies - the more similar their export structures. As the graph below shows, proximity between countries increases as the sum of their log GDP per capita increases.



The second limitation –referring to the Proximity Controls methodology in particular - is that there is no unique and correct strategy to create linear combinations of comparator countries. In the three case studies above we proposed three different strategies to create a Proximity

Control, and all yielded similar outcomes. And there are many more ways of constructing successful controls. We have not included this in the final paper, but methodologies that work equally well, include: using the proximity between triplets and quadruplets of countries rather than pairs, generating random weighted combinations of a country's 5 closest comparators and selecting the combination that best matches certain criteria (e.g. that best matches the growth rate of the reference country, or that best matches the reference country on selected indicators of interest, or that maximizes the share of a country's export package that is accounted for), generating groups of countries in the export proximity space using various clustering algorithms, etc. How good the resulting Proximity Control is will only partly depend on the approach. What matters is how well the Proximity Control resists to robustness checks such as the three we used in each of the case studies (the fitness test, the bow tie test and the placebo test).

The third limitation of this study is the measure of export proximity itself. The measure we propose is not continuous, we do not provide a valid justification as to why we should chose as a threshold RCA>1 (as opposed to RCA>0.5 for example, or RCA>2), and by including in the denominator the maximum diversity of a pair of countries, we are inevitably omitting from the results information about the export diversity of one of the two countries, which is not an accurate representation of reality. There are many different ways to measure the export similarity between pairs of countries, or triplets of countries (etc), and each will come with a number of advantages and disadvantages. There is no perfect measure however and the researcher will always have to make a choice between the amount of information he/she collects about the similarity between two countries (e.g. by lowering the RCA threshold or using the continuous RCA vector as in Bahar, Hausmann, et Hidalgo, you capture more information about the export similarity between two countries) and the relevance of that information. The export similarity measure proposed here is therefore not an ideal measure, but one possible measure, which serves the purposes of this exercise.

The ideas and approaches discussed in this paper are just one of many possible policy applications of the export proximity space. In particular there are two major areas which are not covered: (i) the behavior of the export proximity space as a whole, rather than from the reference point of a specific country of interest; and (ii) the dynamics of the export proximity space over time. Both will lead to interesting insights about economic development and should be further researched.

References

Abadie, Gardeazabal , "The Economic Costs of Conflict: a Case-Control Study for the Basque Country", American Economic Review (2003)

Abadie, Diamond, Hainmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program", 2009

Bahar, Hausmann, Hidalgo, *"The Structure and Dynamics of the Network of Export Similarity",* Forthcoming (2011)

Hausmann, Hidalgo , "*Country Diversity, Product Ubiquity and Economic Divergence",* Forthcoming (2010)

Hausmann, Klinger, " *Structural Transformation and Patterns of Comparative Advantage in the Product Space*" CID Working Paper No. 128

Hausmann, Pritchett, Rodrik , "*Growth Accelerations"*, Journal of Economic Growth, 10, 303-329, 2005

Hidalgo, Hausmann , "*A Network View of Economic Development"*, Developing Alternatives (2008) 12(1): 5-10

Hidalgo, Hausmann , "*The Building Blocks of Economic Complexity"*, Proc. Natl. Acad. Sci. (2009) 106(26):10570-10575

Hidalgo, Klinger, Barabasi, Hausmann*, "The Product Space Conditions the Development of Nations"*, Science (2007) 317: 482-487

Ju, Lin, Wang, *"Endowment Structures, Industrial Dynamics and Economic Growth*", No 5055, Policy Research Working Paper, Sept 2009

Lin, *" New structural economics: a framework for rethinking development",* No 5197, Policy Research Working Paper Series from The World Bank

Lin, Monga, *"Growth identification and facilitation: the role of the state in the dynamics of structural change",* No 5313, Policy Research Working Paper Series from The World Bank

Lin, Monga, *"The growth report and new structural economics "* , No 5336, Policy Research Working Paper Series from The World Bank

McMillan, Rodrik, *"Globalization, structural change and productivity growth"* , Version February 2011 (forthcoming)

Technical Annexes

Annex 1 – Revealed Comparative Advantage

Revealed Comparative Advantage (RCA) indices are measures aimed at determining whether a country has a comparative advantage in a certain product or not and what the extent of this comparative advantage or disadvantage is.

The measure of Revealed Comparative Advantage (RCA) we use is Balassa's index, developed in 1965. Balassa's RCA index defines country *i*'s comparative advantage in product *j* as:

$$RCA_{a,i} = \frac{E_{a,i}}{E_{w,i}} / \frac{E_a}{E_w} \text{ <=> } RCA_{a,i} = \frac{E_{a,i}}{E_a} / \frac{E_{w,i}}{E_w}$$

where $RCA_{a,i}$ is the revealed comparative advantage of country *a* in product *i*, $E_{a,i}$ is total exports of country *a* in product *i*, $E_{w,i}$ is total global export of product *i*, E_a is total exports of country *a*, and E_w total global exports. Basically what this formula measures is a country's share of world exports of a specific product divided by its share of total world exports. A country is said to have a revealed comparative advantage in a certain product when it's RCA in that product is greater than 1, i.e. when the country's share of world exports of that product is greater than the country's share of global exports. This is the definition of revealed comparative advantage we use from this point forward.

Annex 2 – Comparing Export Proximity to alternative variables

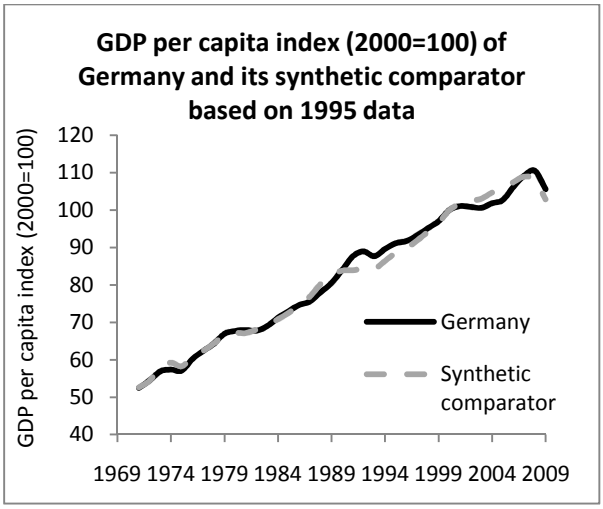
In the table below we compare export proximity to other similarity measures constructed using proximity in terms of GDP per capita and years of schooling. We find that on average, export proximity is much more predictive of differences between countries on other key economic variables than similarity measures based on individual variables such as GDP per capita or years of schooling. The reason export proximity provides much stronger signals is because it captures a lot more information about the similarity between countries.

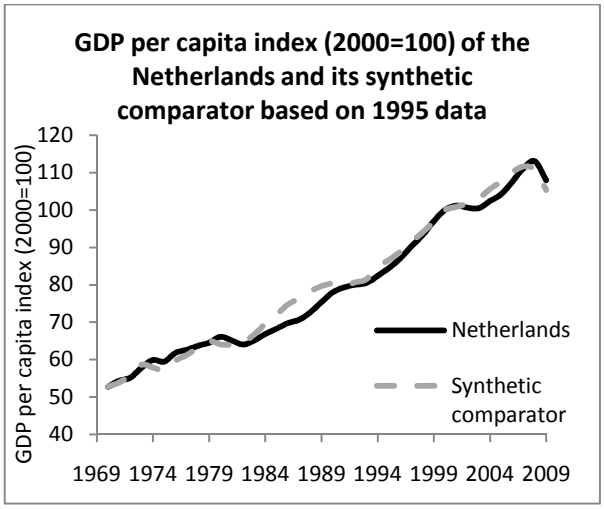
Variable of interest (1995 data)	Regression with similarity ranking based on years of schooling	Regression with similarity ranking based on GDP per capita	Regression with similarity ranking based on export proximity
Average difference in governance effectiveness between pairs of countries by rank	• Observations: 120 • T-statistic: 7.9 • R^2=31.2	• Observations: 120 • T-statistic: 15.75 • R^2=74.75	• Observations: 120 • T-statistic: 52.8 • R^2=95.33
Average difference in regulatory quality between pairs of countries by rank	• Observations: 120 • T-statistic: 13.26 • R^2=58.1	• Observations: 120 • T-statistic: 17.58 • R^2=74.44	• Observations: 120 • T-statistic: 39.06 • R^2=92.65

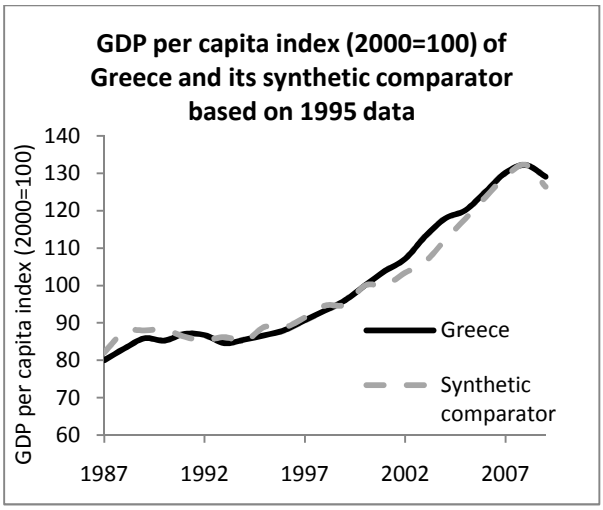
Average difference in years of schooling between pairs of countries by rank	Not applicable	• Observations: 120 • T-statistic: 14.5 • R^2=69.46	• Observations: 120 • T-statistic: 41.51 • R^2=93.31
Average difference in years of economic structure (squared error) between pairs of countries by rank	• Observations: 120 • T-statistic: 3.14 • R^2=3.02	• Observations: 120 • T-statistic: 2.92 • R^2=13.83	• Observations: 120 • T-statistic: 12.03 • R^2=59.35
Average difference in GDP per capita between pairs of countries by rank	• Observations: 120 • T-statistic: 7.11 • R^2=34.79	Not applicable	• Observations: 120 • T-statistic: 42.62 • R^2=95.16
Average difference in GDP per capita growth (1995-2005) between pairs of countries by rank	• Observations: 120 • T-statistic: -1.63 • R^2=1.31	• Observations: 120 • T-statistic: 6.3 • R^2=31.81	• Observations: 120 • T-statistic: 6.72 • R^2=28

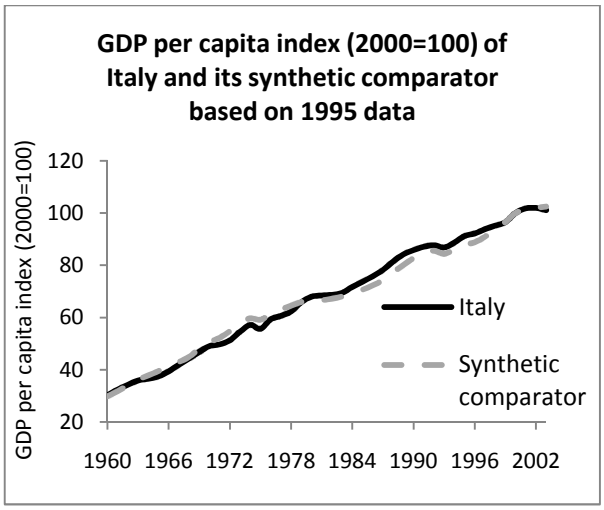
Annex 3 – Additional examples highlighting the fact that countries that are close to each other in the export proximity space tend to have similar growth patterns

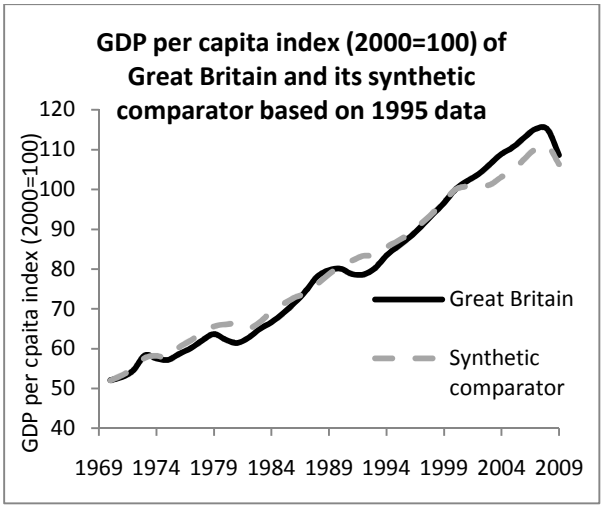
In this Annex we provide additional examples of countries which have shown similar growth patterns of time to their Proximity controls over a long period of time. We use Proximity data from 1995 or 2005 to construct the Proximity controls (the date is specified).

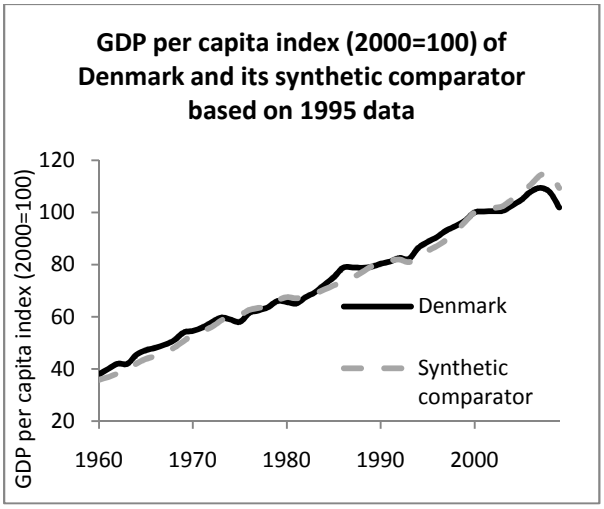


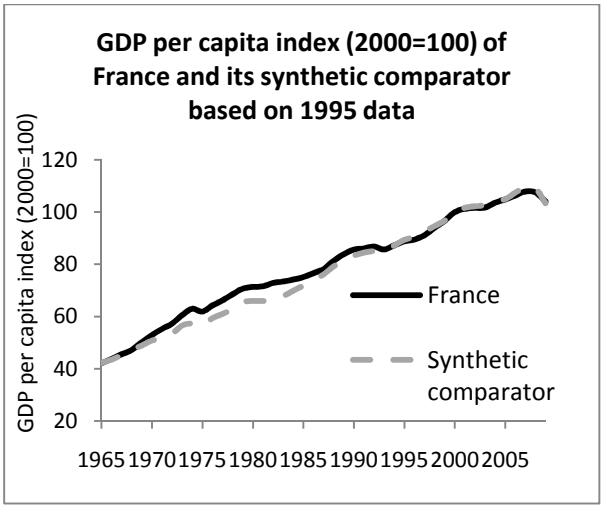


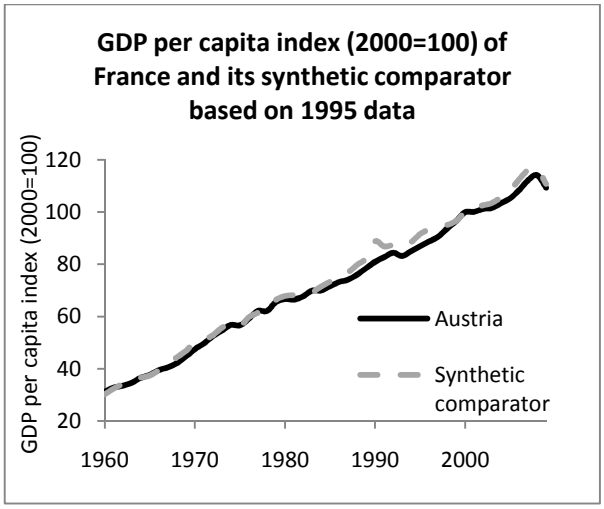

GDP per capita index (2000=100) of Greece and its synthetic comparator based on 1995 data
GDP per capita index (2000=100)
140
130
120
110
100
90
80
70
60
1987
1992
1997
2002
2007
Greece
Synthetic comparator

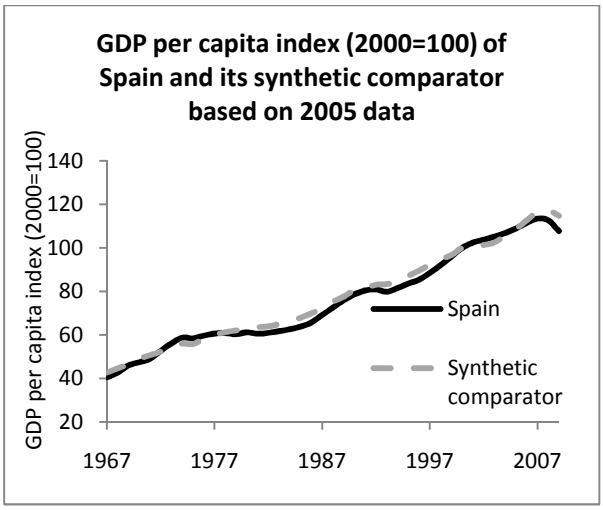

GDP per capita index (2000=100) of Italy and its synthetic comparator based on 1995 data
GDP per capita index (2000=100)
120
100
80
60
40
20
1960
1966
1972
1978
1984
1990
1996
2002
Italy
Synthetic comparator

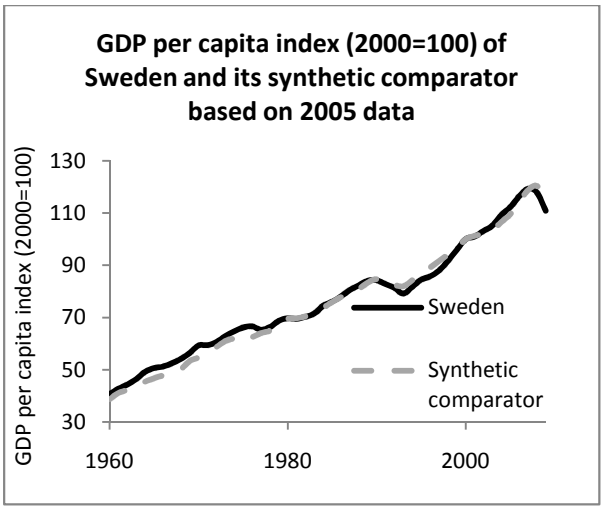

GDP per capita index (2000=100) of Great Britain and its synthetic comparator based on 1995 data
GDP per cpaita index (2000=100)
120
110
100
90
80
70
60
50
40
1969
1974
1979
1984
1989
1994
1999
2004
2009
Great Britain
Synthetic comparator

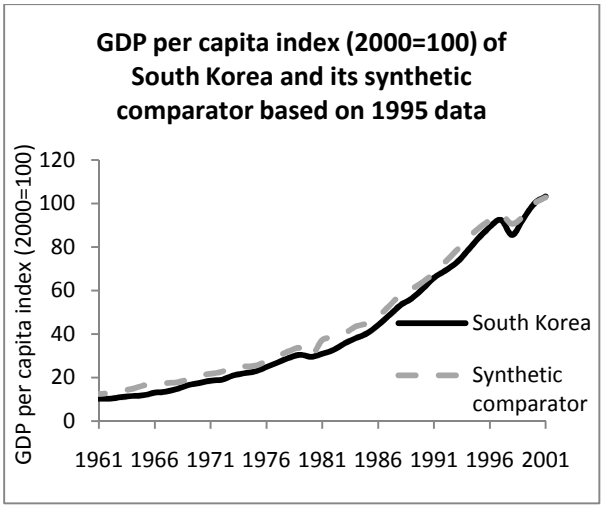

GDP per capita index (2000=100) of Denmark and its synthetic comparator based on 1995 data
GDP per capita index (2000=100)
120
100
80
60
40
20
1960
1970
1980
1990
2000
Denmark
Synthetic comparator

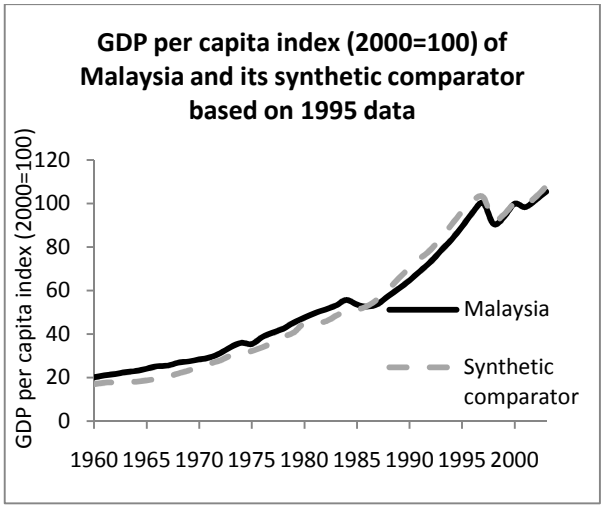

GDP per capita index (2000=100) of France and its synthetic comparator based on 1995 data
GDP per capita index (2000=100)
120
100
80
60
40
20
1965
1970
1975
1980
1985
1990
1995
2000
2005
France
Synthetic comparator

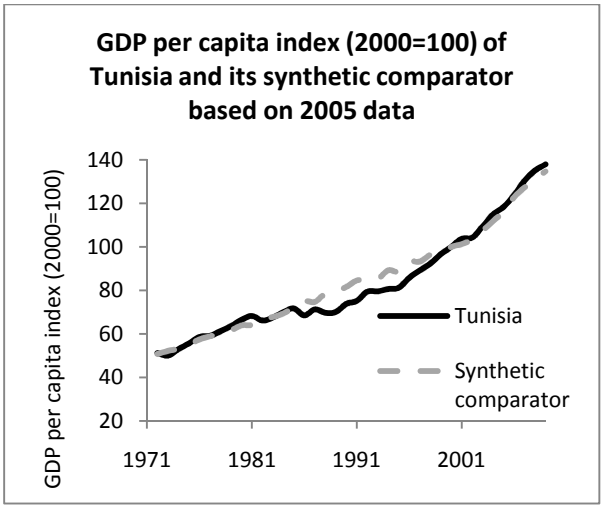

GDP per capita index (2000=100) of France and its synthetic comparator based on 1995 data
GDP per capita index (2000=100)
120
100
80
60
40
20
1960
1970
1980
1990
2000
Austria
Synthetic comparator

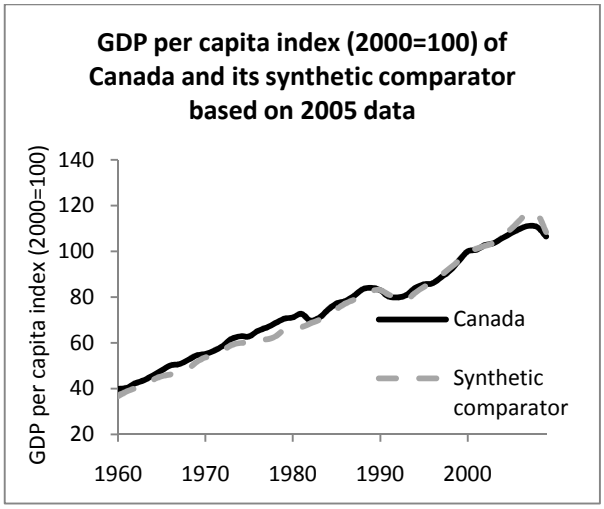










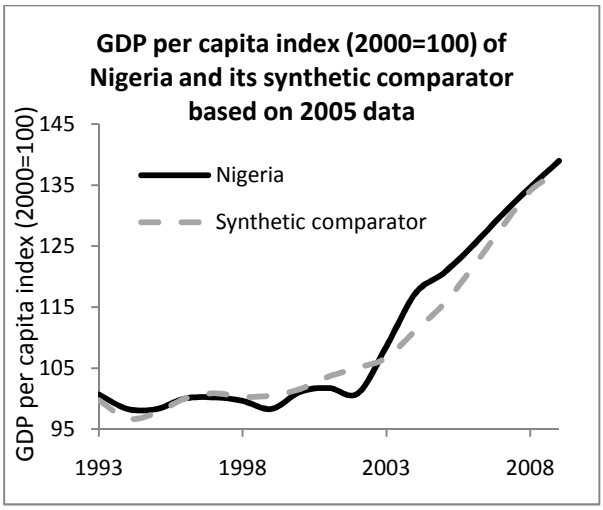


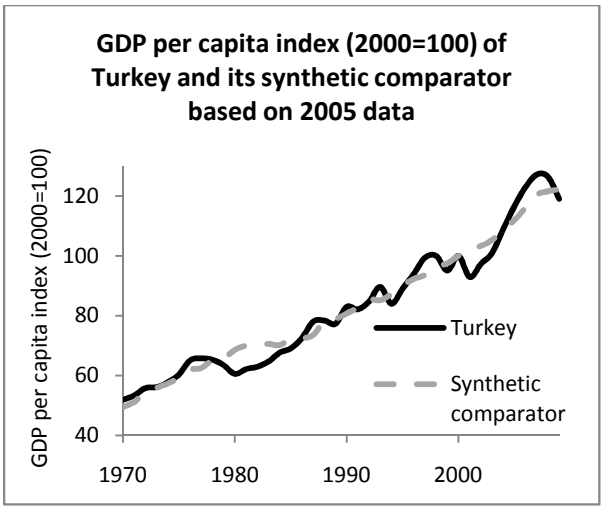
The synthetic comparators above have been constructed using a simple average of the GDP per capita index of the three closest countries in the export proximity space to the country of interest. The only

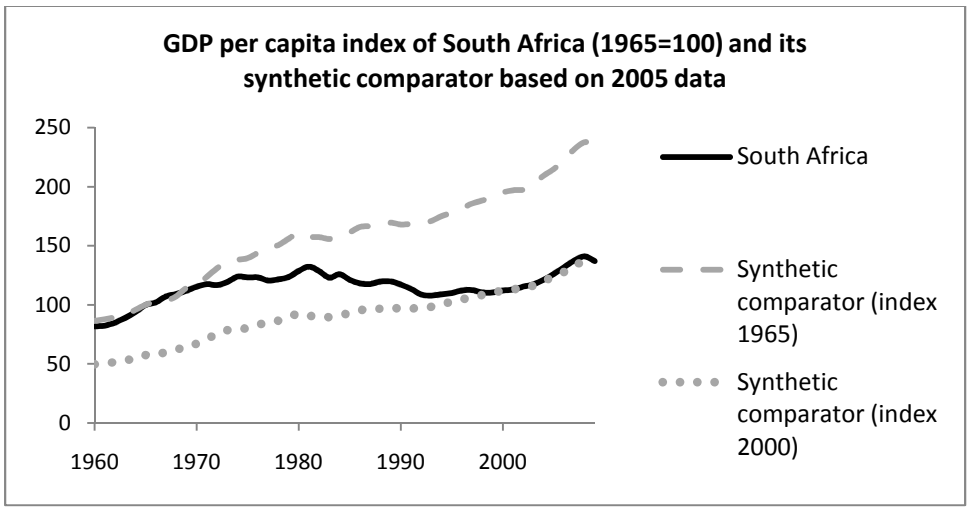
exceptions are Canada, for which ex-USSR countries were eliminated from the sample; and Sweden and Spain, for which their 7 closest comparators were used.

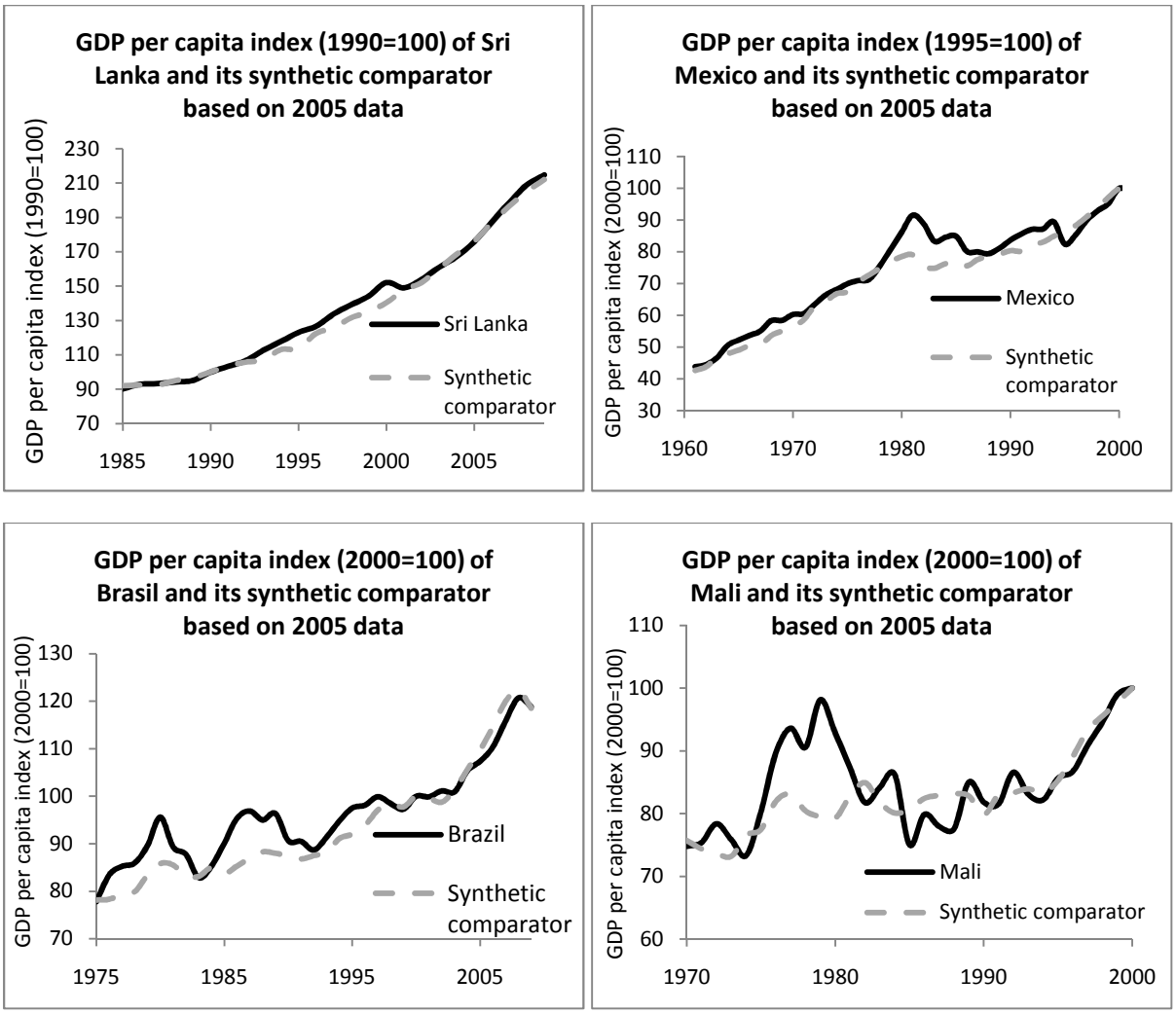
Annex 4 – Additional examples highlighting the fact that countries that deviate from their shared growth path tend to converge back towards it

In this Annex we provide additional examples of countries which - after a positive or negative shocks - converge back to similar growth rates with their synthetic comparators. We use Proximity data from 1995 or 2005 to construct the Proximity controls (the date is specified).









We have constructed these synthetic comparators as follows: Nigeria – top 3 2005 data; Turkey – top 6 2005 data; South Africa – top 7 2005 data, eliminating ex-USSR countries; Sri Lanka – top 3 2005 data, eliminating ex-USSR countries; Mexico – top 3, 2005 data, eliminating ex-USSR countries; Brazil – top 5, eliminating ex-USSR countries; Mali – top 3, eliminating Niger which has experienced high levels of political turmoil.