Sachin Gathani and Dimitri Stoelinga*

# Export Similarity Networks and Proximity Control Methods for Comparative Case Studies

**Abstract:** In the paper we explore just how similar the growth trajectory of countries with similar exports is, and exploit this similarity to conduct counter-factual analysis. We find that 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. We call this method Proximity Controls and apply it to the case of Indonesia's 1997 financial and political crisis. We also highlight applications to the cases of political instability in Ivory Coast, election violence in Kenya and Greece's debt crisis.

***Corresponding author: Dimitri Stoelinga,** Laterite Ltd, Kigali, Rwanda
e-mail: dstoelinga@laterite-africa.com (www.laterite-africa.com)
**Sachin Gathani:** Laterite Ltd, Kigali, Rwanda

## 1 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 as 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 and visual tools, which we show lead to interesting insights about growth and economic development.

This piece of work is specifically targeted at economic policy makers. What we hope they will gain from it is: (i) a number of data-driven strategies with which they can identify optimal comparator countries for a country of interest; (ii) an innovative technique to carry out aggregate counter-factual analysis at the sector or country level, which we call proximity controls and which is largely inspired by the synthetic controls methodology of Abadie and 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 and that countries that deviate from these shared growth rates tend to converge back towards them.

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, and institutional performance. This paper builds on previous and current work by Hausmann et al. [Hausmann and Klinger (2006), Hausmann et al. (2007), Hausmann and Hidalgo (2008), Hausmann and Hidlago (2009), and Bahar et al. (2012)] – who introduced the concept of the product space (2006) and have recently proposed a new metric of the export similarity between countries (2012) – as well as the synthetic controls methodology introduced by Abadie and Gardeazabal (2003) and further developed by Abadie et al. (2010).

The three main lessons that this paper takes away from the exports research of Hausmann et al. are that: (i) to produce a certain product with a comparative advantage a country needs to have the right capabilities[1] 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, in order to measure the export similarity between countries. Where Bahar et al. (2012) use a continuous revealed comparative advantage (RCA)[2] vector to measure the export similarity between countries, we use a discrete measure based on a cut-off of the RCA vector, 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. 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 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

**1** In this paper we define capabilities as all the inputs, infrastructure (soft and hard), processes, technology and skills required to produce a certain product with a comparative advantage. Amongst others, this includes endowments (minerals, geography, etc.) and non-tradable capabilities such as property rights, regulations, infrastructure, labor skills, etc.

**2** See Annex 1 for a definition of revealed comparative advantage.

closest comparators. We call this method *Proximity Controls*. It draws on lessons from the synthetic controls methodology, developed by Abadie and Gardeazabal (2003) and in particular the techniques used by Abadie et al. (2010) to test the validity of the synthetic controls they construct. We illustrate how this approach works by estimating the economic impact of Indonesia's financial and political crisis, triggered by the East Asian financial crisis in 1997. We also highlight alternative case studies in Annex 2, such as the impact of Ivory Coast's decade long political crisis on its GDP per capita (focusing on the 1999–2009 period); the impact of Kenya's election violence on GDP per capita (2007); and the impact of the current financial crisis on Greece's economy (2007–2011).

This paper proceeds as follows: we briefly describe the data utilized, before introducing *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 space, providing a number of insights on economic development and introducing ways to identify comparator countries for a country of interest; we then propose a strategy that 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 the case of Indonesia's financial and political crisis. We close with a discussion on the policy implications and limitations of the export proximity space.

## 2 Data

The data used to calculate the export similarity patterns 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).[3] 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 >3 million as the economics of small economies often do not apply to larger countries. Our sample is thereby reduced to 130 countries. The export proximity measures we derive throughout the study are based on 1995, 2005 and 2010 data; the base year is specified in each case.

All other economic indicators (GDP per capita data, GDP growth, etc.) have been taken from the World Development Indicators database, except where indicated otherwise. Any data referring to monetary values is expressed in

**3** Gaulier and Zignago (2010).

terms of constant 2000 USD. Education data on years of schooling has been taken from the Barro and Lee (2010) dataset.[4]

# 3 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, countries that are close to each other in the export proximity space export products that require a similar capabilities-mix.

We first show how proximity is calculated and why this measure was selected over alternatives, before illustrating some properties which indicate that our measure of proximity is likely to be a good proxy for the similarity of capabilities between pairs of countries.[5]

## 3.1 Measure of Export Similarity and Alternatives Considered

Our purpose in selecting a measure of export similarity is to identify the most appropriate comparators for a country of interest. So we need to identify a metric

4 Barro et al., April 2010.

5 Please note that this measure of proximity can be expanded to include triplets, quadruplets, quintuplets, etc., of countries, rather than simply pairs. A measure of proximity based on *n*-tuplets, would measure the similarity in the exports of *n* countries, resulting in exponentially increasing combinations of countries.

that provides the best possible signal of how similar the economies of a pair of countries are. To do that we compare how well each of the potential export similarity indices introduced below predict the similarity in GDP per capita between countries with similar exports and their long-term growth rates.

In this paper we use a discrete measure of export similarity between a pair of countries inspired by the measure of distance between products introduced by Hausmann and Klinger (2006) (see Annex 1 for an explanation of revealed comparative advantage). We define this measure as the number of common products in which a pair of countries has a revealed comparative advantage (i.e., RCA>1), 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). Formally, this measure of export similarity between two countries *a* and *b* at time *t* can be written as:

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

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

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 similarity is symmetric (i.e., $\text{Discreet}_{Sim_{a,b}} = \text{Discreet}_{Sim_{b,a}}$) 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. By selecting proximity over alternative measures 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. It is important to note that we could have used other cut-off rates rather than RCA>1; the results obtained using a cut-off rate of 0.5, which is sometimes used in the literature (see Bahar et al. 2012), does not defer significantly.

An alternative approach to measuring export similarity is the export similarity index introduced by Bahar et al. (2012), which is calculated using the Pearson correlation between the continuous RCA vectors of pairs of countries. Its continuous nature means it captures information both on whether countries have similar exports or not and on the respective intensity of these exports. This gives it a theoretical advantage over discrete methods, including the method proposed above,

which only captures information on the similarity of products exported, but not on their respective intensity. Formally, this measure of export similarity between two countries *a* and *b* at time *t* can be written as:

$$\text{Pearson}_{a,b,t} = \frac{\sum_{i=1}^{P}\left(X_{a,i,t}-\bar{X}_{a,t}\right)\left(X_{b,i,t}-\bar{X}_{b,t}\right)}{\sqrt{\sum_{i=1}^{P}\left(X_{a,i,t}-\bar{X}_{a,t}\right)^2 \sum_{i=1}^{P}\left(X_{b,i,t}-\bar{X}_{b,t}\right)^2}} \quad (2)$$

where $X_{a,i,t} = log_{10}(\text{RCA}_{a,i,t}+0.1)$ and $\bar{X}_{a,t}$ is the average of all $X_{a,i,t}$ over all products for country *a* at time *t*.

The measure is negative for pairs of countries that export different sets of products and positive for countries that have a similar exports-mix. In addition, this measure distinguishes between products that are exported by one country only and products that are exported by neither.

Another continuous measure of export similarity to consider is the Finger and Kreinin (1979) export similarity index. This simple measure of export similarity entails working out the product share of total exports for each country and, thereafter, for each pair of countries, summing the minimum of the two countries' shares for a given product across all products. If one of the two countries does not export a given product then the index records a zero for that product; if both countries export a product, then the measure captures the minimum of the product share of the two countries. This measure is therefore reflective of whether countries have similar exports. It also weights larger export products more than smaller ones. This measure of similarity between country *a* and *b* at time *t*, which we call *F&K*, is defined as:

$$F\&K_{a,b,t} = \sum_{i=1}^{N}[\min(S_{a,i,t}, S_{b,i,t})] \quad (3)$$

where $S_{a,i,t}$ *is the* share of product *i* over country's *a*'s total exports at time *t*.

To identify which metric is the most likely to yield good comparators, we test how well these measures predict the similarity in the GDP per capita levels and growth rates of reference countries versus their most similar exporters. We compare countries on two measures: GDP per capita using 1995 and 2010 data and compounded annual GDP per capita growth over the 1995–2010 period (see Table 1). We measure the goodness-of-fit by calculating, for each country, the average GDP per capita and GDP per capita growth rate of its three most similar exporters (selected using each of the different measures of export similarity); we then look at the $R^2$ of the resulting linear regression between the reference countries and their comparators.

No single measure minimizes the difference between countries and their most similar exporters across all time periods, but both proximity and the Pearson

**Table 1** How Well do Each of these Measures Predict the Similarity Between a Country and its Most Similar Exporters?

| Selected Measure of Export Similarity | $R^2$ on GDP per capita 1995* | $R^2$ on GDP per capita 2010** | $R^2$ on Compounded Annual GDP per capita Growth Rate 1995–2010* |
|---|---|---|---|
| Proximity (RCA>1) | 0.61 (*t*-statistic=10.6) | 0.69 (*t*-statistic=13.6) | 0.27 (*t*-statistic=6.1) |
| Pearson's Correlation | 0.59 (*t*-statistic=10.1) | 0.77 (*t*-statistic=16.0) | 0.21 (*t*-statistic=5.4) |
| F&K | 0.49 (*t*-statistic=8.4) | 0.45 (*t*-statistic=10.0) | 0.06 (*t*-statistic=2.5) |

*Similarity measures using 1995 exports data; **similarity measured using 2010 exports data.

correlation based measures perform better than the Finger and Kreinin export similarity index (Table 1). The proximity and Pearson correlation indexes yield very similar results; for some countries the proximity measures perform better, for others – in particular when it involves countries with very concentrated exports where export intensity matters a lot, such as oil exporters – the Pearson correlation measure performs better. In this paper we choose to use the proximity measure moving forward, as it does a better job in matching growth over the 1995–2010 period and fits the 1995 GDP per capita data better, which we use as a base year for the case of Indonesia.

Figure 1 illustrates what the global export proximity space – based on the selected measure – looks like. The nodes represent countries, while the edges between them represent the link between a reference country and its most similar exporters. 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 in Figure 1 only depicts the three closest neighbors of countries in the export proximity space; it is therefore a directed network with arrows going from the reference country to its three closest exporters. We highlight geographic groupings of countries to give the reader a sense of what the export proximity space looks like. At a first glimpse, it seems to make sense – countries in the same continent seem to have more similar exports than countries in other continents.

## 3.2 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,

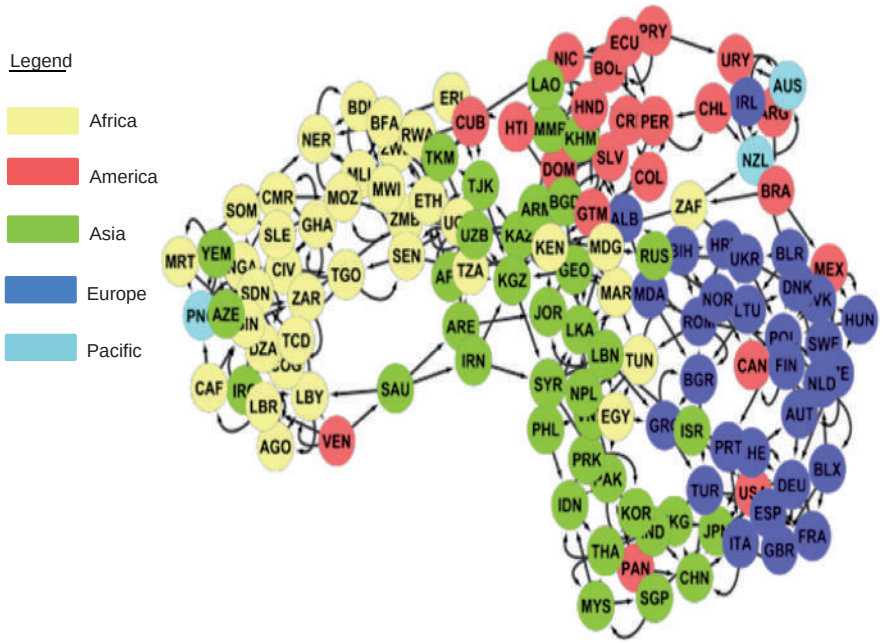



**Figure 1** Network Representation of the Export Similarity Space (Based on 2010 Export Data).[6]

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. We also 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). It is important to note that we focus only on the 110 closest exporters (out of a 130) for each country, as the *Proximity* measure fails to provide an adequate signal for countries with highly concentrated exports (in particular oil exporters) such as Iraq, Angola, Chad, Libya, Venezuela, Congo Brazzaville, Central African Republic, Liberia, Somalia, Algeria, Azerbaijan, Sudan, Papa New Guinea, Democratic Republic of Congo, Turkmenistan, Mauritania, Saudi Arabia and Yemen. Because of the limited diversity of their export base, these countries tend to have low export similarity levels with other countries. Moreover, especially in the case of oil exporters, there is a mismatch between their socio-economic indicators (which for countries like Saudi Arabia matches that of devel-

6 This network representation was designed using Cytoscape 2.8.3. For more information see http://www.cytoscape.org/.

oped nations) and the structure of their non-oil economy. Continuous measures of export similarity, such as the Pearson correlation introduced by Bahar et al. (2012), perform better for these countries.

**Property 1:** ***On average, the closer countries are to each other in the export proximity space, the more similar their GDP per capita levels.*** Figure 2 depicts the average absolute difference in log GDP per capita between countries, based on their export proximity rank with other countries. The closest country to a reference country in the export proximity space is ranked one, the second closest ranked 2, and so forth. 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. A linear regression fits this association between the average absolute difference in log GDP per capita between pairs of countries and their Proximity rank well and is strongly statistically significant ($t$-statistic=42.8; $R^2$=0.95).

**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 Figure 3 below, the greater the export proximity between countries, the smaller on average the absolute difference between their GDP per capita growth rates (we compare countries based on their compounded annual GDP per capita growth rate during 1995–2005). This is a powerful association that we will elaborate on further in the ensuing sections.

**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 (see Figure 4). We find that countries with similar exports have more similar levels of average

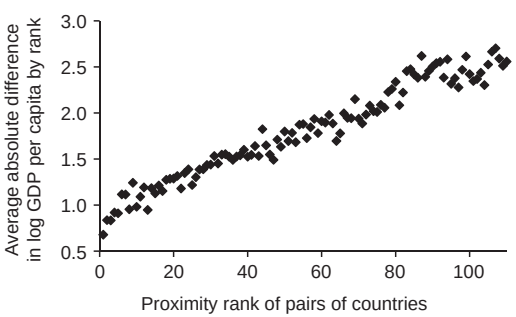



**Figure 2** Average Difference in GDP per capita Between Pairs of Countries based on Proximity Rank (Closest 110 Exporters, 2010 Data).

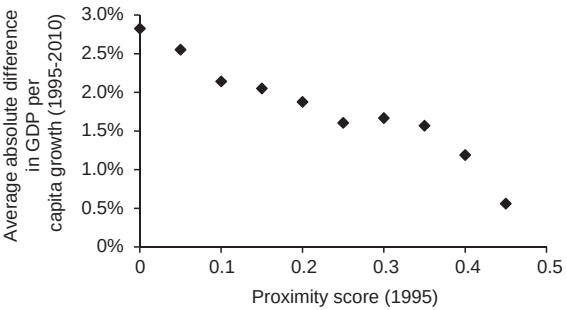



**Figure 3** Average Difference in Absolute GDP per capita Growth Between Countries on Proximity Score (1995–2010).

years of schooling 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 macroeconomic structure.*** To determine whether countries have a similar economic structure or not we calculate the correlation between all pairs of countries on a number of key macro-economic indicators. The selected indicators include: gross fixed capital formation (%GDP), gross domestic savings (%GDP), exports (%GDP), imports (%GDP), agriculture (%GDP), industry (%GDP), and services (%GDP). As can be seen in Figure 5 there is a strong association between the correlation levels of pairs of countries across the selected indicators and their proximity rank. 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.

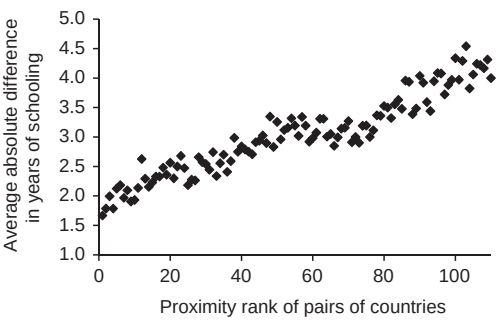



**Figure 4** Average Difference in Years of Schooling Between Pairs of Countries based on Proximity Rank (Closest 110 Exporters, 2010 Data).

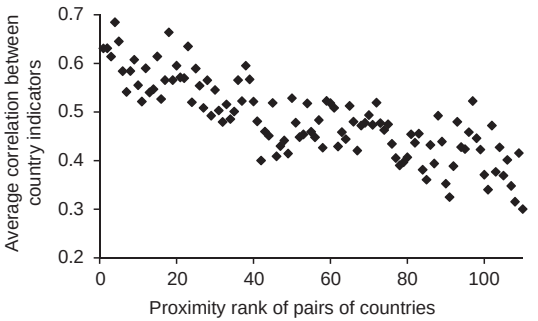



**Figure 5** Average Difference in Economic Structure Between Pairs of Countries based on Proximity Rank (2010 Data).

**Property 5:** ***On average, countries that are closer to each other in the export proximity space are also geographically closer to each other.*** As can be seen in the Figures 6 and 7, 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. This is also one of the key findings of Bahar et al.'s work (2012) on the producer space. The most likely explanation as to why this might be the case is that capabilities are more easily transferable between countries that are closer to each other geographically than countries that are further apart, be it because of regional partnerships (e.g., the European Union, the East African Community, etc.), similar geography and climatic conditions, shared natural resources, direct transportation links, and the continuous movement of people, capital and goods between neighboring countries.

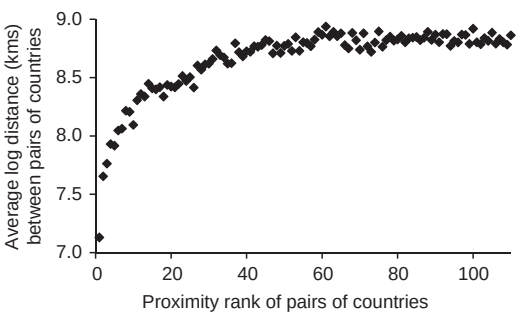



**Figure 6** Average Geographic Distance Between Pairs of Countries based on Proximity Rank (110 Closest Exporters, 2010 Data).

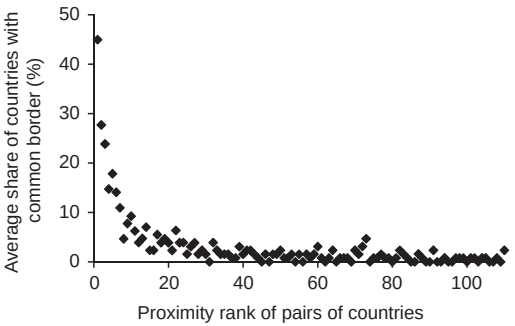



**Figure 7** Share of Countries that Share a Common Border by Proximity Rank (110 Closest Exporters, 2010 Data).

**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 World Governance Indicators (WGI) on Government Effectiveness and measure the absolute differences in scores between pairs of countries. Again we find that countries that are closer to each other in the export proximity space tend to have a more similar institutional performance (here we use Government Effectiveness as a proxy) than countries that are further apart. The relationships hold when controlling for differences in GDP per capita (Figure 8).

**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 Figure 9 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. 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 products that require them to be produced.

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

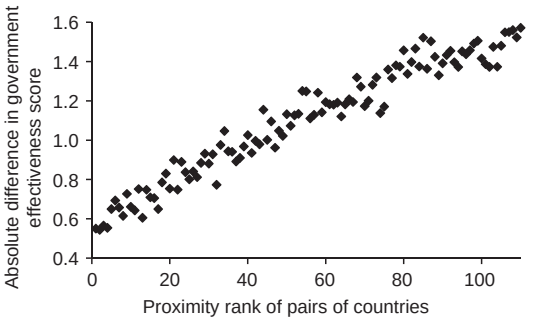



**Figure 8** Average Difference in Government Effectiveness Between Pairs of Countries based on Proximity Rank (110 Closest Exporters, 2010 Data).

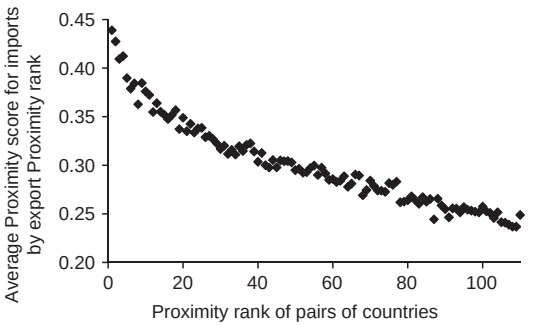



**Figure 9** Average Import Proximity of Countries based on their Export Proximity Rank (110 Closest Exporters, 2010 Data).

economic growth, the more aligned their macro-economic structure, human capital and institutional indicators, the closer they are geographically, and the more similar their import structure is. 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.

Moreover, 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 socio-economic variables. We illustrate this in Table 2 using the example of GDP per capita and years of schooling, which are aggregate estimates of economic development and human capital. The most likely explanation as to why export proximity provides much stronger signals is because it captures a lot more information about the similarity between countries and the complexity of their economies. This is the main difference in the

**Table 2** Comparing Proximity to Pair-wise Similarity Measures Calculated using GDP per capita and Years of Schooling (Based on 2010 Data Except for Growth Comparison which are based on 1995–2010 Data).

| Dependent Variable | Explanatory Variable: Similarity Ranking based on Years of Schooling | Explanatory Variable: Similarity Ranking Based on GDP per capita | Explanatory Variable: Similarity Ranking based on Export Proximity |
|---|---|---|---|
| Average difference in governance effectiveness between pairs of countries by rank | – Observations: 110<br>– *t*-statistic: 14.9<br>– $R^2$=69.9 | – Observations: 110<br>– *t*-statistic: 18.2<br>– $R^2$=83.8 | – Observations: 110<br>– *t*-statistic: 43.5<br>– $R^2$=94.6 |
| Average difference in years of schooling between pairs of countries by rank | Not applicable | – Observations: 110<br>– *t*-statistic: 14.3<br>– $R^2$=70.4 | – Observations: 110<br>– *t*-statistic: 32.3<br>– $R^2$=90.2 |
| Average difference in macroeconomic structure (squared error) between pairs of countries by rank | – Observations: 110<br>– *t*-statistic: –6.4<br>– $R^2$=36.6 | – Observations: 110<br>– *t*-statistic: –6.9<br>– $R^2$=29.6 | – Observations: 110<br>– *t*-statistic: –14.6<br>– $R^2$=67.8 |
| Average difference in log GDP per capita between pairs of countries by rank | – Observations: 110<br>– *t*-statistic: 19.7<br>– $R^2$=80.9 | Not applicable | – Observations: 110<br>– *t*-statistic: 42.8<br>– $R^2$=95.4 |
| Average difference in GDP per capita growth (1995–2005) between pairs of countries by rank (top 20 countries by rank) | – Observations: 20<br>– *t*-statistic: 0.4<br>– $R^2$=0.00061 | – Observations: 20<br>– *t*-statistic: –0.05<br>– $R^2$=0.0002 | – Observations: 20<br>– *t*-statistic: 7.26<br>– $R^2$=78.33 |

literature between the approaches proposed by Bahar et al. (2012), which focus on complex networks, and those that focus on factors and economic aggregates such as Lin and Monga (2010).

## 4 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 (see Annex 3 for full list). We use 1995 as the base year as this is the first year for which BACI data is available; selecting 1995 also enables us to compare growth rates forwards and backwards in time. We find some remarkable results. The examples in Figures 10–13, which represent a diverse mix of countries, highlight just how similar – in the long-term – the growth rate of countries can be to that of their synthetic comparators.

The USA's compounded annual GDP per capita growth rate between 1970 and 2009 was 1.81% per year, compared to 1.88% for its synthetic comparator, constructed using the average of Great Britain, Germany and France (see Figure 10). India's compounded annual GDP per capita growth rate between 1990 and 2010 was 4.71%, compared to 4.85% for its synthetic comparator, constructed using China, Hong Kong and Turkey (Figure 11). South Korea's compounded GDP per capita growth rate between 1960 and 2000 was 5.3%, compared to 5.9% for its synthetic comparator, made out of the combined indexed GDP per capita of China,

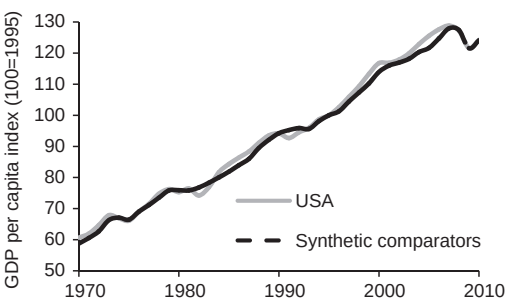



**Figure 10** GDP per capita Index (100=1995) in the USA and its Synthetic Comparator (1970–2010).

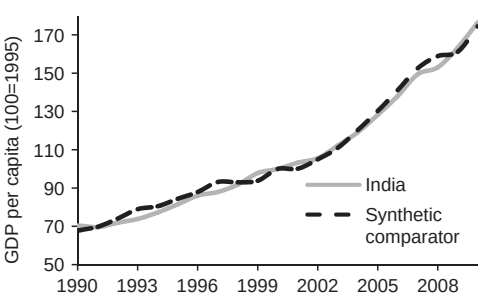



**Figure 11** GDP per capita Index (100=1995) in India and its Synthetic Comparator (1990–2010).

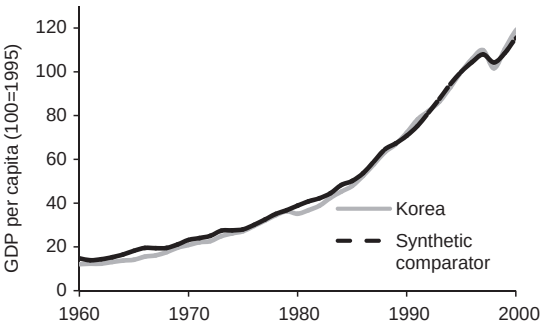



**Figure 12** GDP per capita Index (100=1995) in South Korea and its Synthetic Comparator (1960–2000).

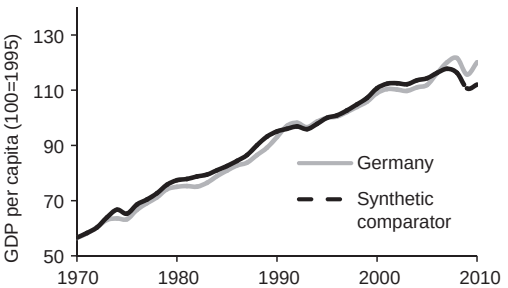



**Figure 13** GDP per capita Index (100=1995) in Germany and its Synthetic Comparator (1970–2010).

Thailand and Hong Kong (Figure 12). Germany's compounded annual GDP per capita growth rate during 1970–2010 was 1.91% compared to 1.84% for its synthetic comparator (Figure 13). Germany's synthetic comparator was constructed using Germany's two closest exporters in the 1995 export proximity space: Italy and France.

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 (we do not imply any relationship of causality). Of course this is not always the case, in particular for oil exporters, countries that have experienced domestic shocks (positive or negative), small economies with highly volatile growth rates (where comparatively small events in the economy can lead to large swings in economic growth), or outliers such as China on the positive side and countries like the Democratic Republic of Congo, Haiti and Eritrea on the negative side. But on average, countries with similar exports

have similar growth patterns (see Figures 14 and 15). The association between the compounded annual GDP per capita growth rates of reference countries and comparators selected using 1995 as the base year are positive and statistically significant during the 15 years preceding and following 1995, i.e., during the 1980–1995 period as well as the 1995–2010 period ($t$-statistic >5 in both cases).

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*

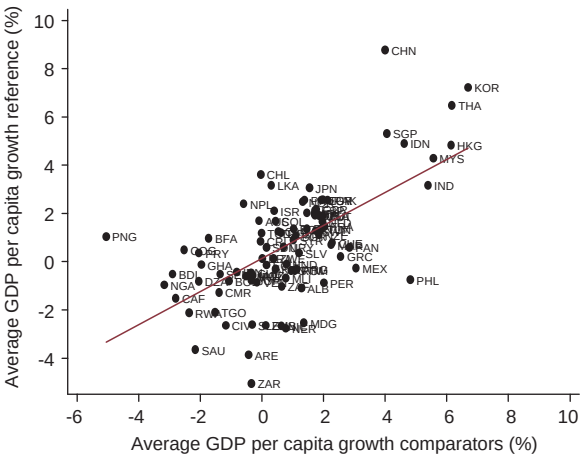



**Figure 14** GDP per capita Growth: Reference vs. Comparators (1980–1995).

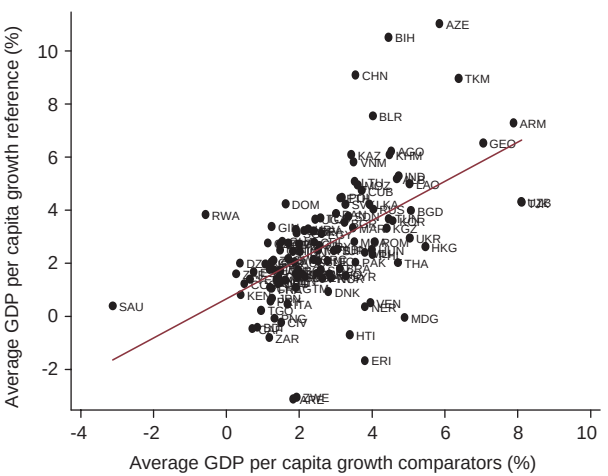



**Figure 15** GDP per capita Growth: Reference vs. Comparators (1995–2010).

*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 and Guatemala. We use different approaches to construct the synthetic comparators for these countries: for Bangladesh (see Figure 16), we construct a synthetic comparator by averaging its four closest comparators in the 1995 export proximity space; for Guatemala (see Figure 17), we average its three closest comparators in the 2005 proximity space, but eliminate ex-USSR countries from the sample. As can be seen in the figures above, 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

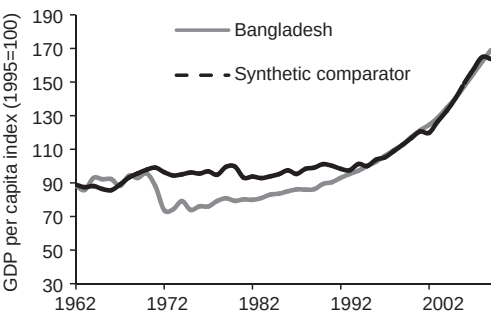



**Figure 16** GDP per capita Index (1995=100) of Bangladesh and its Synthetic Comparator based on 1995 Data.

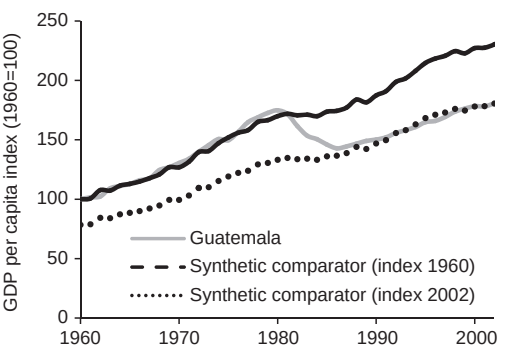



**Figure 17** GDP per capita Index of Guatemala (1960=100) and its Synthetic Comparator based on 2005 Data.

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. In addition to selecting a country's most similar comparators, researchers can identify potential comparators by using clustering algorithms to identify nodes of countries in the export proximity space, corresponding to groups of countries with very similar exports to each other and therefore, in all likelihood, also similar economic characteristics. The export proximity space also 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.

## 5 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 and Gardeazabal (2003); Abadie et al. (2010), 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 and Gardeazabal (2003) 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 illustrate how the proximity control method works using the case of Indonesia's financial and political crisis which was triggered in 1997. We check the validity of the resulting control regions with two complementary tests, which we will detail below. The strategy we propose to construct a proximity control for Indonesia is just one of many possible strategies. In Annex 2, we have included other case studies as well, where we use different methods to arrive at a valid proximity control. In the case of Ivory Coast, we weighted countries in the proximity control by their proximity score; in the case of Kenya we use an elimination strategy to measure the impact of Kenya's dual domestic crises (election violence in December 2007–2008 and the 2008–2009 drought)[7]; in the case of Greece we

7 See Sachin Gathani and Dimitri Stoelinga (2011).

simply take the average of Greece's seven closest comparators that were not as severely affected by the Euro crisis.[8]

## 5.1 Proximity Controls and the Impact of the Indonesian Financial and Political Crisis on GDP per Capita in Indonesia

To measure the impact of Indonesia's financial and political crisis (Indonesia was one of the countries that was the hardest hit by the Asian financial crisis), we construct a proximity control of Indonesia using a simple strategy based on the export proximity measure. We then test whether the proximity control is a valid control, by (i) checking if this synthetic region fits Indonesia on a number of indicators before the beginning of the crisis, (ii) checking if the results are very sensitive to changes in the composition of the proximity control, and (iii) by running a falsification test.

### 5.1.1 The Indonesian Financial and Political Crisis

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 countries, encouraging lenders to withdraw significant credit and causing a credit crunch and bankruptcies on a massive scale.

Indonesia, South Korea and Thailand were the countries most affected by the crisis. In Indonesia, the rupiah was also subject to severe speculative attacks, leading to a strong recession. This crisis came on top of a political legitimacy crisis which had been brewing since mid-1996, following the July 27 riots at the headquarters of one of the opposition parties (PDI) which sparked the beginning of a popular movement to challenge the Suharto regime.[9] Suharto's ill health, the legitimacy crisis and the collapse of the economy made the eventual departure of President Suharto inevitable. By 1999, there were signs that economically most of the countries had begun to recover economically from the East Asian financial

**8** 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 highlight how this methodology works.
**9** See Stefan Eklöf (2004).

crisis. In Indonesia political uncertainty continued through to the first popular presidential election in 2004.

### 5.1.2 Constructing a Proximity Control for Indonesia

To construct a proximity control for Indonesia we start by selecting a base year for the analysis. Our objective is to match Indonesia to a synthetic combination of similar exporters that fit Indonesia's growth path prior to the 1997 financial and political crisis. We select 1995 as the base year, as this comes before the start of the political turmoil, which began mid-1996, and before the onset of the East Asian financial crisis. We then eliminate all countries from the 1995 export proximity space that were directly and severely affected by the East Asian financial crisis, including Thailand, Korea, the Philippines, Hong Kong, Laos, and Malaysia. This leaves us with a pool of countries that were comparatively less affected, from which we select Indonesia's three closest exporters: Portugal, China and India (see Table 3).

We define Indonesia's proximity control as the linear combination of these three countries that best matches indexed GDP per capita (100=1995) in Indonesia during the 1980–1995 period. We find the most optimal linear combination by generating 5000 random combinations of these three countries and selecting the one that minimizes the difference between Indonesia's growth path and that of the proximity control. The resulting contribution of these comparator countries to Indonesia's proximity control is: China (48.35%), Portugal (27.03%) and India (24.62%).

This proximity control has very similar macro-economic characteristics to Indonesia (see Table 4). The similarity between the two regions is based on high investment and savings rates, a similar share of agriculture and manufacturing over GDP, and an almost identical trade balance and urbanization rate. Indonesia is more industry and trade intensive than the proximity control, but the difference in industry and exports is most probably attributable to Indonesia's petroleum sector.

In addition to matching Indonesia on key indicators, the selected proximity control almost perfectly matches growth in Indonesia during the 1980–1996

**Table 3** Contribution of Comparators to Proximity Control.

| Country | Closest Comparators | Proximity | Contribution to Proximity Control |
|---|---|---|---|
| Indonesia | Portugal | 0.360 | 27.03% |
| | China | 0.325 | 48.35% |
| | India | 0.322 | 24.62% |

**Table 4** Comparing Macro-economic Variables in Indonesia and Proximity Control (1995 Data).

| Indicator (1995 data) | Indonesia | Proximity Control |
|---|---|---|
| Gross fixed capital formation (% of GDP) | 28.4 | 28.7 |
| Gross domestic savings (% of GDP) | 30.6 | 31.8 |
| Agriculture (% of GDP) | 17.1 | 17.6 |
| Industry (% of GDP) | 41.8 | 37.2 |
| Manufacturing (% GDP) | 24.1 | 25.5 |
| Services (% of GDP) | 41.1 | 45.2 |
| Exports of goods and services (% of GDP) | 26.3 | 19.7 |
| Imports of goods and services (% of GDP) | 27.6 | 21.1 |
| External balance on goods and services (% of GDP) | −1.3 | −1.3 |
| Gross national expenditure | 101.3 | 101.1 |
| Urban (% population) | 35.5 | 35.3 |

period, i.e., before the financial crisis (see Figure 18). During this period, the compounded annual GDP per capita growth rate of the proximity control was 5.1%, compared to 5% for Indonesia. Figure 18 also reveals that while growth in Indonesia and its proximity control were almost identical during the 1980–1996 period, they started diverging in 1997. The East Asian financial crisis seems to have impacted Indonesia in two ways: (i) it shaved off an approximate 21% off Indonesia's potential GDP per capita in the immediate aftermath of the crisis (1996–1999); and (ii) Indonesia settled on a slower growth path thereafter (between 1999 and 2004). While Indonesia grew just 0.3% points slower than the proximity control during the 4 years preceding the crisis (6.1% per capita growth during 1992–1996 vs. 6.4%), it grew 2.6% points slower in the 4 years after the crisis had settled (3.2% between 2000 and 2004, vs. 5.8%). This could in part be attributed to the ongoing political uncertainty leading to the 2004 elections. By

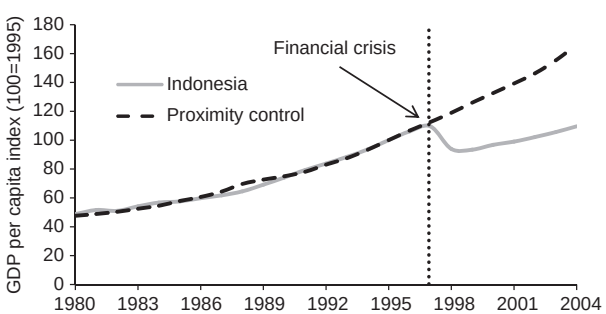



**Figure 18** The long-term impact of the Indonesian financial and political crisis on GDP per capita.

2004 Indonesia's GDP per capita was only 66% of what it could have been. This is equivalent to a GDP per capita loss of 33.5% or a cumulative opportunity cost of about US$528bn (in current US$) when extrapolated between 1997 and 2004.

### 5.1.3 Testing the Validity of the Proximity Control

We conduct two tests to check the validity of the control: a random permutations test and a falsification test on the proximity control itself.

#### 5.1.3.1 Random Permutations Test

First, we propose an inference technique – which we call the random permutations test – that enables us to test the sensitivity of our results to changes in the composition of the proximity control. Random permutation testing consists in randomly changing the composition of the proximity control using other countries close to Indonesia in the 1995 export proximity space and random weights. Rather than selecting only the three closest countries to Indonesia, we randomly select different combinations of triplets out of Indonesia's 10 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 (in terms of countries and weights) 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.

This inference exercise enables us to estimate the distribution of differences between growth in the country of interest and randomly generated proximity controls, consisting of countries that are "similar enough" to the country of interest. One way of picturing this, is to imagine that each of these proximity controls is one possible path the country of interest could have taken had the event or intervention not taken place. If the impact of an event/intervention on the country of interest is large, then the estimated difference should be large regardless of the proximity control.

The 10 countries closest to Indonesia in the export proximity space that were not severely affected by the East Asia crisis are: Portugal, China, India, Vietnam, Sri Lanka, Romania, Pakistan, Morocco, Croatia, and Tunisia.[10] We create 5000

**10** We exclude Turkey, which experienced a major financial crisis itself during the 1999–2001 period.

random combinations of triplets of these countries, with random weights assigned to each. We compare their growth performance during the 1990–2004 period to that of Indonesia.

These random permutations enable us to conclude that if the universe of possible alternative growth paths for Indonesia created by these permutations is representative of what would have happened had the financial and political crisis not occurred – something we argue is likely because of how similar the growth of countries that are close to each other in the export proximity space is – we can state with 95% confidence that by 2004 the crisis had cost Indonesia between 3.6% and 37.2% of GDP. However, not all 5000 proximity controls are valid controls for Indonesia, as they are not accurate predictors of how well Indonesia performed before the financial and political crisis. This explains the large spread in these initial estimates.

To overcome this problem, we measure how well these alternative linear combinations of comparator countries match growth in Indonesia before the East Asian financial crisis (during the 1990–1996 period)[11] and select the 5% of linear combinations (or 250 combinations) that do the best job. The reasoning is that if these linear combinations of similar exporters were best at estimating growth in Indonesia before the crisis, then they are also more likely to be the most accurate reference point for what would have happened in Indonesia had the crisis not occurred. As can be seen in the Figure 19, this approach reduces the intervals of our estimates. While we cannot reject the null hypothesis that the difference in the indexed GDP per capita of Indonesia and that of the 250 linear combinations

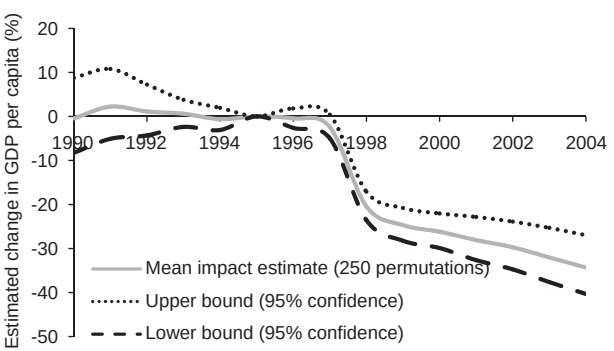



**Figure 19** Estimated Impact of Indonesian Financial and Political Crisis on GDP per capita (Based on 250 Best Permutations).

**11** We measure how well proximity controls match growth in Indonesia before the East Asian financial crisis using the squared sum of the difference between the indexed GDP per capita of Indonesia and each individual proximity control between 1990–1997.

of comparators was zero between 1990 and 1996, Indonesia and these 250 alternative proximity controls start to diverge in 1997. In other words, the mean seems to be an accurate predictor of growth in Indonesia during the 1990–1996 period.

These estimates are based on very different linear combinations of countries than the initial proximity control (see Table 5). While the contribution of China to the mean of the 250 permutations is not very different from its contribution to the proximity control (i.e., close to 50%), the shares of India and Portugal have been significantly reduced and replaced by countries such as Sri Lanka, Pakistan, Tunisia and Morocco. Romania, Croatia and Vietnam contribute very little to the 250 selected permutations (see Table 6). This increased diversity of the impact estimates gives us confidence that the latter are robust to changes in the composition of countries and their respective contribution to the proximity controls.

According to this narrowed down set of 250 permutations, we estimate that the East Asian financial crisis had cost Indonesia up to 34.4% of GDP by 2004 (with a 95% confidence interval ranging from 27% to 40.1%). This is very much in line with our initial estimated impact of 33.5% of GDP by 2004.

#### 5.1.3.2 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

**Table 5** Results Based on Random Permutations of Top 250 Alternative Proximity Controls.

| Indicator | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 |
|---|---|---|---|---|---|---|---|
| Estimated cumulative GDP loss based on Proximity control | −20.8% | −25.6% | −26.7% | −28.5% | −29.7% | −31.5% | −33.5% |
| Minimum estimate permutations | 0.2% | 0.7% | 2.0% | −2.6% | −2.4% | −2.0% | −3.6% |
| Maximum estimate permutations | −25.3% | −30.3% | −32.9% | −36.2% | −39.3% | −42.5% | −45.5% |
| 95% confidence interval (upper bound) | −6.99% | −7.76% | −6.10% | −6.23% | −4.37% | −3.01% | −3.6% |
| 95% confidence interval (lower bound) | −24.09% | −27.72% | −29.26% | −31.03% | −32.60% | −35.01% | −37.16% |

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 exact same methodology we used to construct the proximity control. 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 its Placebo region during the period under consideration, indicating that the proximity control did not itself experience any positive or negative shocks.

To construct the placebo region we create an aggregate [China (48.35%), Portugal (27.03%), India (24.62%)] region. We define its export proximity to a given country in the 1995 export proximity space as the weighted average of the pair-wise proximity scores of China, Portugal and India with that country. As in the case of Indonesia and its proximity control, the placebo region will consist of the three closest countries to the [China (48.35%), Portugal (27.03%), India (24.62%)] region that were not severely affected by the East Asian financial crisis (Indonesia, Thailand, Korea, the Philippines, Laos, Singapore, Hong Kong, and Malaysia). We also take Turkey out of the sample, which was affected by its own financial crisis between 1999 and 2001. As can be seen in the Table 7 the three

**Table 6** Contribution of Comparator Countries to Mean of 250 Best Permutations.

| Country | Contribution to Mean | Country | Contribution to Mean |
|---|---|---|---|
| China | 49.9% | Tunisia | 5.7% |
| Sri Lanka | 12.1% | Morocco | 2.5% |
| India | 11.5% | Romania | 0.6% |
| Portugal | 9.7% | Croatia | 0.4% |
| Pakistan | 7.3% | Vietnam | 0.2% |

**Table 7** Contribution of Comparators to Placebo Region.

| Country | Closest Comparators | Proximity | Contribution to Placebo Region |
|---|---|---|---|
| Indonesia | Italy | 0.591 | 98.3% |
| | Spain | 0.562 | 2.6% |
| | Germany | 0.525 | 0.2% |

closest countries to this aggregate region in the 1995 export proximity space were Italy, Spain and Germany.

While the resulting Placebo region is not a good predictor of growth in the proximity control during the 1990–1997 period – on average the proximity control region grew much faster – the results signal that the East Asian financial crisis did not disproportionately affect either regions (see Table 8). Both regions grew at approximately the same rate before and after the 1997 crisis and the growth differential between them was approximately the same (4.83% per capita per year in the 7 years before the crisis; 4.73% in the 7 years after).

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.

## 6 Closing Remarks

What the example above highlights 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

**Table 8** GDP per capita growth in the Placebo Region and the Proximity Control.

| Average GDP per capita growth | Placebo region | Proximity control | Difference |
|---|---|---|---|
| 1990–1997 | 1.31% | 6.15% | 4.83% |
| 1997–2004 | 1.20% | 5.93% | 4.73% |

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 Figure 20 shows, proximity between countries increases as the sum of their log GDP per capita increases. This does not mean that the proximity control method does not yield satisfactory results for less developed countries: the examples of Ivory Coast and Kenya in Annex 2 are a testament to that fact. What it means is that the proximity measure captures more information about the similarities and dissimilarities of developed countries.

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 case study above we propose a specific strategy to create a proximity control. Yet 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 five 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

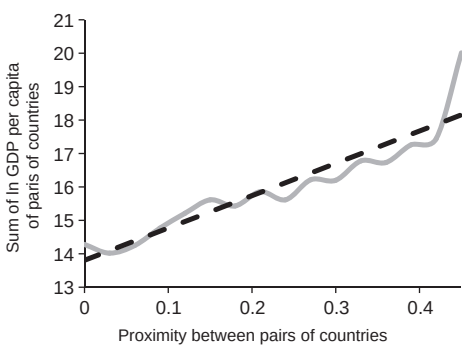



**Figure 20** Proximity Between Pairs of Countries and the Sum of their GDP per capita.

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 controls resist to robustness checks such as the two we used in the Indonesia case study (the random permutations tests and the falsification test), but we could think of many more potential checks.

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 , e.g., 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 et al., 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.

**Acknowledgements:** We are extremely grateful to Miguel Almunia, Jonathan Argent, Ricardo Hausmann, Cesar Hidalgo, Michele Savini, and Rupert Simons for their insights and advice.

# 7 Technical Annexes

## 7.1 Annex 1 – Revealed Comparative Advantage

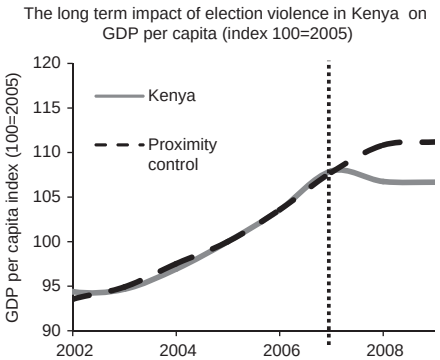
The measure of revealed comparative advantage (RCA) we use is Balassa's index, developed in 1965 (Balassa 1986). Balassa's RCA index defines country *i*'s comparative advantage in product *j* as:

$$\mathrm{RCA}_{a,i}=\frac{\frac{E_{a,i}}{E_{w,i}}}{\frac{E_a}{E_w}} \quad <=> \quad \mathrm{RCA}_{a,i}=\frac{\frac{E_{a,i}}{E_a}}{\frac{E_{w,i}}{E_w}}$$

where $\mathrm{RCA}_{a,i}$ 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 its RCA in that product is >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 in this paper.

## 7.2 Annex 2 – Other Interesting Examples of Aggregate Impact

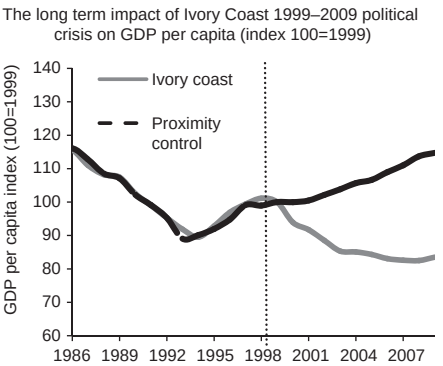
Kenya – Impact of election violence on GDP per capita (2007)




| Country | Contribution to proximity control (%) |
|---|---|
| Guatemala | 45.6 |
| Tanzania | 24.0 |
| Jordan | 16.6 |
| Nepal | 9.1 |
| Uganda | 4.7 |

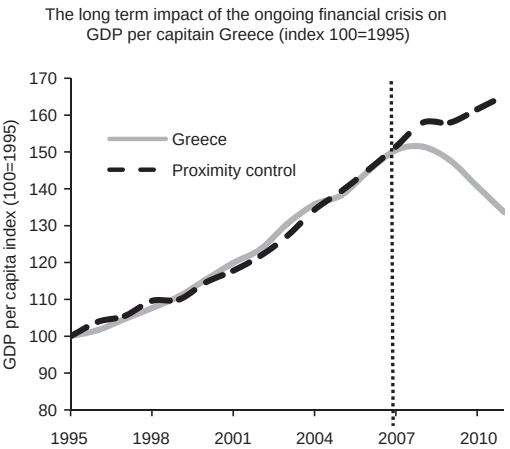
(**Annex 2 Figure** Continued)

Ivory Coast – Impact of political crisis on GDP per capita in Ivory Coast (1999–2009)




| Country | Contribution to proximity control |
|---|---|
| Cameroon | 37.39% |
| Ghana | 33.94% |
| Togo | 28.67% |

Greece – Impact of financial crisis on GDP per capita in Greece (2007–2011)




| Country | Contribution to proximity control (%) |
|---|---|
| Turkey | 14.3 |
| Croatia | 14.3 |
| Bulgaria | 14.3 |
| Hungary | 14.3 |
| Morocco | 14.3 |
| Tunisia | 14.3 |
| Lebanon | 14.3 |

## 7.3 Annex 3

Comparing the Long-term Compounded Annual GDP per capita Growth Rates of Countries and their Three Closest Exporters in the 1995 Export Proximity Space (1995–2010).

| Reference Country | Most Similar Exporter | Second Most Similar Exporter | Third Most Similar Exporter | Compounded GDP per capita Growth Reference (%) | Compounded GDP per capita Growth Comparators (%) |
|---|---|---|---|---|---|
| Angola | Congo | Liberia | Yemen | 6.2 | 0.5 |
| Albania | Bangladesh | Vietnam | Dominican republic | 5.2 | 5.1 |
| UAE | Bangladesh | Madagascar | Honduras | –3.1 | 5.1 |
| Argentina | Uruguay | Australia | New Zealand | 2.7 | 2.6 |
| Armenia | Iran | Azerbaijan | Cuba | 7.3 | 2.0 |
| Australia | New Zealand | South Africa | Argentina | 2.0 | 1.5 |
| Austria | Czech republic | Germany | Italy | 1.7 | 1.1 |
| Azerbaijan | Uzbekistan | Turkmenistan | Tajikistan | 11.0 | 8.1 |
| Burundi | Congo, Dem. Rep. | Rwanda | Central African Republic | –0.4 | 1.2 |
| Burkina Faso | Niger | Mali | Ethiopia | 3.2 | 3.8 |
| Bangladesh | Sri Lanka | Vietnam | Albania | 4.0 | 3.9 |
| Bulgaria | Romania | Turkey | Poland | 3.5 | 4.1 |
| Bosnia Herzegovina | Albania | Bangladesh | Sri Lanka | 10.5 | 4.7 |
| Belarus | Ukraine | Lithuania | Russian federation | 7.5 | 5.0 |
| Bolivia | Paraguay | Peru | Nicaragua | 1.8 | 1.2 |
| Brazil | South Africa | Bulgaria | Slovak Republic | 1.8 | 0.3 |
| Central African Republic | Guinea | Congo, Dem. Rep. | Burundi | –0.5 | 1.2 |
| Canada | Finland | Sweden | South Africa | 1.6 | 1.8 |
| Switzerland, Liechtenstein | Germany | Austria | Japan | 1.1 | 1.4 |

(**Annex 3** Continued)

| Reference Country | Most Similar Exporter | Second Most Similar Exporter | Third Most Similar Exporter | Compounded GDP per capita Growth Reference (%) | Compounded GDP per capita Growth Comparators (%) |
|---|---|---|---|---|---|
| Chile | Peru | New Zealand | Argentina | 2.8 | 2.1 |
| China | Hong Kong | India | Czech republic | 9.1 | 5.5 |
| Ivory Coast | Cameroon | Ghana | Togo | –0.2 | 2.0 |
| Cameroon | Ivory Coast | Ghana | Tanzania | 1.5 | 1.5 |
| Congo | Congo, Dem. Rep. | Nigeria | Central African Republic | 1.2 | 1.2 |
| Colombia | Guatemala | El Salvador | Tunisia | 1.5 | 1.9 |
| Costa Rica | Guatemala | El Salvador | Honduras | 2.5 | 1.9 |
| Cuba | Mozambique | Ghana | Guinea | 4.7 | 3.6 |
| Czech republic | Italy | Austria | Germany | 2.7 | 1.7 |
| Germany | Italy | France | Czech republic | 1.2 | 1.7 |
| Denmark | Netherlands | Sweden | Poland | 0.9 | 1.2 |
| Dominican republic | Haiti | Bangladesh | Honduras | 4.2 | 3.4 |
| Algeria | Libya | Iran | Saudi Arabia | 2.0 | 1.2 |
| Ecuador | Ivory Coast | Chile | Senegal | 1.7 | 1.5 |
| Egypt | Pakistan | Morocco | Turkey | 3.1 | 3.6 |
| Eritrea | Ethiopia | Burkina Faso | Sudan | –1.7 | 3.1 |
| Spain | Italy | France | Belgium-Luxembourg | 1.7 | 1.7 |
| Ethiopia | Sudan | Mali | Burkina Faso | 4.4 | 3.3 |
| Finland | Sweden | Austria | Canada | 2.4 | 1.7 |
| France | Germany | Italy | United kingdom | 1.0 | 1.4 |
| United kingdom | USA | Germany | France | 2.0 | 1.4 |
| Georgia | Azerbaijan | Russian federation | Kazakhstan | 6.5 | 5.9 |

(**Annex 3** Continued)

| Reference Country | Most Similar Exporter | Second Most Similar Exporter | Third Most Similar Exporter | Compounded GDP per capita Growth Reference (%) | Compounded GDP per capita Growth Comparators (%) |
|---|---|---|---|---|---|
| Ghana | Guinea | Ivory Coast | Cameroon | 2.8 | 1.2 |
| Guinea | Ghana | Central African Republic | Mauritania | 3.4 | 1.5 |
| Greece | Turkey | Portugal | Croatia | 2.1 | 3.0 |
| Guatemala | El Salvador | Costa Rica | Colombia | 1.1 | 2.6 |
| Hong Kong | China | Thailand | India | 2.6 | 3.5 |
| Honduras | Costa Rica | Dominican republic | El Salvador | 1.6 | 1.5 |
| Croatia | Hungary | Portugal | Romania | 3.3 | 4.0 |
| Haiti | Dominican republic | Bangladesh | Nepal | –0.7 | 1.6 |
| Hungary | Poland | Croatia | Slovak Republic | 2.5 | 3.2 |
| Indonesia | Thailand | Hong Kong | Portugal | 2.4 | 4.7 |
| India | China | Hong Kong | Turkey | 5.3 | 3.5 |
| Israel | Greece | Switzerland, Liechtenstein | Hungary | 1.6 | 2.4 |
| Italy | Germany | Czech republic | France | 0.5 | 1.4 |
| Jordan | Lebanon | Kenya | Guatemala | 2.7 | 3.1 |
| Japan | Switzerland, Liechtenstein | Germany | USA | 0.7 | 1.2 |
| Kazakhstan | Kyrgyzstan | Russian federation | Ukraine | 6.1 | 4.4 |
| Kenya | Zimbabwe | Costa Rica | El Salvador | 0.8 | 1.9 |
| Kyrgyzstan | Kazakhstan | Uzbekistan | Moldova, republic of | 3.3 | 3.4 |
| Cambodia | Lao | Myanmar | Bangladesh | 6.1 | 5.0 |
| Korea, republic of | Hong Kong | Thailand | China | 3.6 | 5.5 |
| Lao | Cambodia | Myanmar | Bangladesh | 5.0 | 4.5 |
| Lebanon | Morocco | Tunisia | Greece | 2.5 | 3.5 |

(**Annex 3** Continued)

| Reference Country | Most Similar Exporter | Second Most Similar Exporter | Third Most Similar Exporter | Compounded GDP per capita Growth Reference (%) | Compounded GDP per capita Growth Comparators (%) |
|---|---|---|---|---|---|
| Liberia | Angola | Congo | Congo, Dem. Rep. | 11.0 | 4.5 |
| Sri Lanka | Vietnam | Philippines | Tunisia | 4.2 | 3.5 |
| Lithuania | Poland | Croatia | Romania | 5.1 | 3.2 |
| Morocco | Tunisia | Sri Lanka | Lebanon | 3.3 | 4.5 |
| Moldova, republic of | Lithuania | Ukraine | Lebanon | 2.8 | 3.5 |
| Madagascar | Bangladesh | Myanmar | Vietnam | –0.1 | 5.1 |
| Mexico | Thailand | Brazil | Turkey | 1.6 | 4.7 |
| Mali | Burkina Faso | Ethiopia | Sudan | 2.4 | 2.4 |
| Mozambique | Tanzania | Ghana | Cuba | 4.9 | 1.9 |
| Mauritania | Somalia | Senegal | Yemen | 1.4 | 2.2 |
| Malawi | Tanzania | Bangladesh | Togo | 1.6 | 1.9 |
| Malaysia | Singapore | Thailand | Indonesia | 2.5 | 1.9 |
| Niger | Burkina Faso | Ethiopia | Sudan | 0.3 | 2.4 |
| Nigeria | Congo, Dem. Rep. | Cameroon | Ethiopia | 2.8 | 1.2 |
| Nicaragua | Honduras | Costa Rica | Dominican republic | 2.1 | 2.8 |
| Netherlands | Belgium-Luxembourg | Denmark | USA | 1.8 | 1.5 |
| Norway | Finland | Sweden | Russian federation | 1.4 | 1.8 |
| Nepal | Bangladesh | Haiti | Madagascar | 2.0 | 5.1 |
| New Zealand | Australia | Denmark | Ireland | 1.3 | 1.9 |
| Pakistan | Egypt | Morocco | Sri Lanka | 2.0 | 2.6 |
| Panama | Thailand | Morocco | Tunisia | 3.9 | 4.7 |
| Peru | Chile | Pakistan | Honduras | 3.2 | 2.4 |

(**Annex 3** Continued)

| Reference Country | Most Similar Exporter | Second Most Similar Exporter | Third Most Similar Exporter | Compounded GDP per capita Growth Reference (%) | Compounded GDP per capita Growth Comparators (%) |
|---|---|---|---|---|---|
| Philippines | Sri Lanka | Vietnam | Thailand | 2.3 | 3.9 |
| Papa New Guinea | Yemen | Congo, Dem. Rep. | Guinea | –0.1 | 0.6 |
| Poland | Romania | Hungary | Slovak Republic | 4.5 | 4.1 |
| Portugal | Turkey | Croatia | Greece | 1.4 | 3.0 |
| Paraguay | Nicaragua | Ivory Coast | Bolivia | 0.6 | 2.8 |
| Romania | Poland | Bulgaria | Slovak Republic | 2.8 | 3.2 |
| Russian federation | Ukraine | South Africa | Belarus | 4.0 | 5.0 |
| Rwanda | Congo, Dem. Rep. | Burundi | Central African Republic | 3.8 | 1.2 |
| Saudi Arabia | UAE | Libya | Iran | 0.4 | 1.8 |
| Sudan | Ethiopia | Mali | Burkina Faso | 3.8 | 3.1 |
| Senegal | Tanzania | Mauritania | Togo | 1.4 | 1.9 |
| Singapore | Malaysia | Hong Kong | Japan | 3.1 | 2.5 |
| Sierra Leone | Ghana | Guinea | Senegal | 2.1 | 1.5 |
| El Salvador | Guatemala | Costa Rica | Sri Lanka | 1.7 | 1.9 |
| Slovak Republic | Romania | Poland | Hungary | 4.2 | 4.1 |
| Sweden | Finland | Austria | Denmark | 2.2 | 1.8 |
| Syria | Bangladesh | Lebanon | Morocco | 1.5 | 5.1 |
| Chad | Somalia | Mauritania | Sudan | 3.7 | 2.2 |
| Togo | Ivory Coast | Malawi | Cameroon | 0.2 | 1.5 |
| Thailand | Indonesia | Hong Kong | China | 2.0 | 2.0 |
| Tajikistan | Uzbekistan | Azerbaijan | Turkmenistan | 4.3 | 8.1 |
| Turkmenistan | Azerbaijan | Uzbekistan | Sudan | 9.0 | 5.9 |

(**Annex 3** Continued)

| Reference Country | Most Similar Exporter | Second Most Similar Exporter | Third Most Similar Exporter | Compounded GDP per capita Growth Reference (%) | Compounded GDP per capita Growth Comparators (%) |
|---|---|---|---|---|---|
| Tunisia | Morocco | Sri Lanka | Vietnam | 3.7 | 3.5 |
| Turkey | Portugal | Greece | India | 2.5 | 2.6 |
| Tanzania | Uganda | Zambia | Madagascar | 3.3 | 2.4 |
| Uganda | Tanzania | Malawi | Myanmar | 3.7 | 1.9 |
| Ukraine | Russian federation | Belarus | Bulgaria | 2.9 | 4.0 |
| Uruguay | Argentina | New Zealand | Tunisia | 2.6 | 2.0 |
| USA | United kingdom | Germany | France | 1.5 | 1.2 |
| Uzbekistan | Azerbaijan | Turkmenistan | Tajikistan | 4.3 | 5.9 |
| Venezuela | Kazakhstan | Russian federation | Ecuador | 0.5 | 3.4 |
| Vietnam | Sri Lanka | Philippines | Bangladesh | 5.8 | 3.9 |
| Yemen | Somalia | Mauritania | Papa New Guinea | 1.4 | 2.2 |
| South Africa | Australia | Zimbabwe | Brazil | 1.6 | 1.9 |
| Congo, Dem. Rep. | Nigeria | Congo | Burundi | –0.8 | 1.7 |
| Zambia | Tanzania | Zimbabwe | Uganda | 2.1 | 1.9 |
| Zimbabwe | South Africa | Guatemala | Egypt | –3.1 | 0.3 |

# References

Abadie, A. and J. Gardeazabal (2003) "The Economic Costs of Conflict: A Case Study of the Basque Country," American Economic Review, 93(1):113–132.

Abadie, A., A. Diamond and J. Hainmueller (2010) "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," Journal of the American Statistical Association, 105:493–505.

Bahar, D., R. Hausmann and C. Hidalgo (2012) *International Knowledge Diffusion and the Comparative Advantage of Nations*. CID Working Paper, 235.

Balassa, B. (1986) "Comparative Advantage in Manufactured Goods – A Reappraisal," The Review of Economics and Statistics, 68:315–319.

Barro, R. J. and J.-W. Lee (2010). "A New Data Set of Educational Attainment in the World, 1950–2010," NBER Working Papers 15902, National Bureau of Economic Research, Inc.

Finger, M. and M. Kreinin (1979) "A Measure of 'Export Similarity' and Its Possible Uses," Economic Journal, 89:905–912.

Gaulier, G. and S. Zignago (2010) BACI: *International Trade Database at the Product-level The 1994–2007 Version*. CEPII Working Paper 2010–2023.

Hausmann, R. and B. Klinger. (2006) *Structural Transformation and Patterns of Comparative Advantage in the Product Space*. Working Paper Series rwp06-041, Harvard University, John F. Kennedy School of Government.

Hausmann, R. and C. Hidalgo (2008) "A Network View of Economic Development", Developing Alternatives 12(1):5–10.

Hausmann, R. and C. Hidalgo (2009) "The Building Blocks of Economic Complexity," Proc. Natl. Acad. Sci. USA 106(26):10570–10575.

Hausmann, R. and C. Hidalgo (2011) "The Network Structure of Economic Output," Journal of Economic Growth, Springer 16(4):309–342.

Hausmann, R., C. Hidalgo, B. Klinger and A.-L. Barabasi (2007) "The Product Space Conditions the Development of Nations," Science 317:482–487.

Lin, J. and C. Monga (2010) *The Growth Report and New Structural Economics*. No. 5336, Policy Research Working Paper, World Bank.