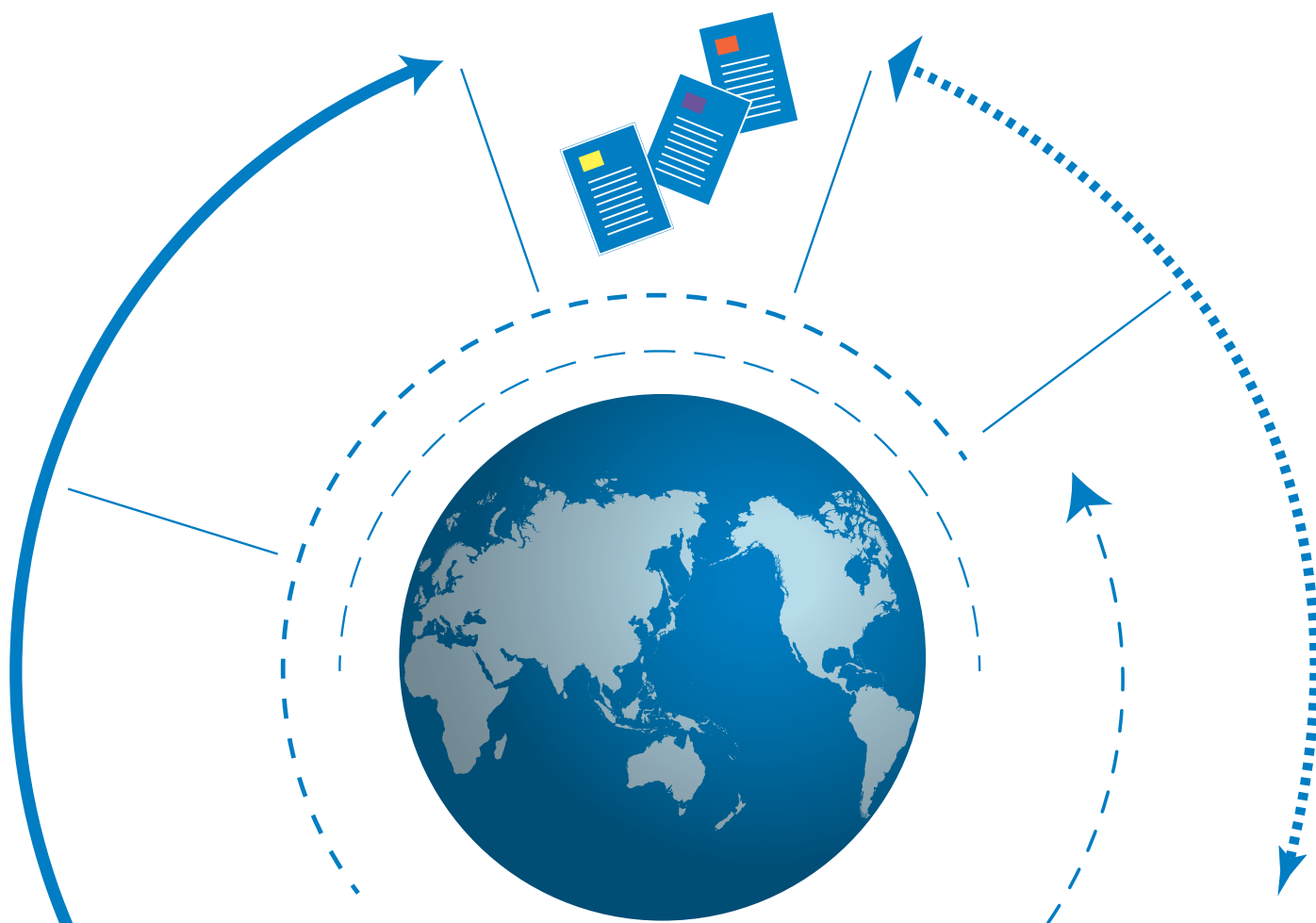




Productive capabilities: An empirical investigation of their determinants

Christian Daude, Arne Nagengast and Jose Ramon Perea



DEVELOPMENT CENTRE WORKING PAPERS

This series of working papers is intended to disseminate the Development Centre's research findings rapidly among specialists in the field concerned. These papers are generally available in the original English or French, with a summary in the other language.

Comments on this paper would be welcome and should be sent to the OECD Development Centre, 2 rue André Pascal, 75775 PARIS CEDEX 16, France; or to dev.contact@oecd.org. Documents may be downloaded from: <http://www.oecd.org/dev/wp> or obtained via e-mail (dev.contact@oecd.org).



THE OPINIONS EXPRESSED AND ARGUMENTS EMPLOYED IN THIS DOCUMENT ARE THE SOLE RESPONSIBILITY OF THE AUTHORS AND DO NOT NECESSARILY REFLECT THOSE OF THE OECD OR OF THE GOVERNMENTS OF ITS MEMBER COUNTRIES

©OECD (2013)

Applications for permission to reproduce or translate all or part of this document should be sent to rights@oecd.org

CENTRE DE DÉVELOPPEMENT DOCUMENTS DE TRAVAIL

Cette série de documents de travail a pour but de diffuser rapidement auprès des spécialistes dans les domaines concernés les résultats des travaux de recherche du Centre de développement. Ces documents ne sont disponibles que dans leur langue originale, anglais ou français ; un résumé du document est rédigé dans l'autre langue.

Tout commentaire relatif à ce document peut être adressé au Centre de développement de l'OCDE, 2 rue André Pascal, 75775 PARIS CEDEX 16, France; ou à dev.contact@oecd.org. Les documents peuvent être téléchargés à partir de: <http://www.oecd.org/dev/wp> ou obtenus via le mél (dev.contact@oecd.org).



LES IDÉES EXPRIMÉES ET LES ARGUMENTS AVANCÉS DANS CE DOCUMENT SONT CEUX DES AUTEURS ET NE REFLÈTENT PAS NÉCESSAIREMENT CEUX DE L'OCDE OU DES GOUVERNEMENTS DE SES PAYS MEMBRES

©OCDE (2013)

Les demandes d'autorisation de reproduction ou de traduction de tout ou partie de ce document devront être envoyées à rights@oecd.org

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	4
PREFACE.....	5
RÉSUMÉ	6
ABSTRACT.....	6
I. INTRODUCTION	7
II. METHODS AND EMPIRICAL STRATEGY	10
III. RESULTS.....	18
IV. CONCLUSION	25
ANNEX.....	27
REFERENCES	32
OTHER TITLES IN THE SERIES/ AUTRES TITRES DANS LA SÉRIE	35

ACKNOWLEDGEMENTS

The authors would like to thank Jake Schurmeier for his capable research assistance and René Orozco who helped editing the paper. All remaining errors are the authors' responsibility.

PREFACE

Economic development requires an eventual change in productive specialisation. This process, known as structural transformation, refers to the reallocation of economic activity across broad economic sectors, and specifically away from agriculture. Recent research has associated structural transformation, and ultimately development, to a change in the type of goods a country produces and exports. This change entails a gradual move towards goods that embody greater physical and/or human capital. Thus, this line of work has also been able to create a variable that embodies the “economic complexity” of a country, otherwise a proxy for the productive capabilities of an economy.

This paper focuses on unveiling the determinants of this variable. In line with the breadth of the related literature, there is a large set of factors that can potentially affect the degree of economic complexity of a country. This, in turn, heightens the uncertainty in selecting the adequate empirical model that explains the behaviour of our dependent variable most accurately. To address this challenge, the paper adopts an econometric methodology that considers all possible combinations of specifications, across a broad set of potential regressors.

The analysis singles out a few variables that are significantly associated with economic complexity. This is the case, for instance, of commodity terms of trade, energy availability, government consumption, capital per worker, arable land or the accrual of capital flows. An identification we deem relevant, not only because of the exhaustive empirical investigation to which these variables have been subjected to; but also because some of them are susceptible to policy intervention.

This work contributes to the OECD Development Centre’s interest to disentangle a more accurate account of which factors, and therefore which policies, facilitate and sustain development. In this way, I hope it will provide valuable guidance on an issue that remains of critical importance for many countries.

Mario Pezzini

Director

OECD Development Centre

December 2013

RÉSUMÉ

Les contributions récentes à la littérature sur la croissance économique ont fait valoir que la structure d'une économie, mesurée par ses capacités de production, est un facteur déterminant pour les différences de développement inter-pays. Les capacités productives sont considérées être hautement prédictives de la croissance économique future, mais leurs déterminants au niveau des pays sont restés inconnus. Dans cet article, nous explorons empiriquement les déterminants à l'aide d'un cadre d'étalement du modèle qui peut gérer un très grand nombre de variables explicatives, sans la nécessité d'une sélection de modèle. Afin d'estimer notre spécification de panel dynamique, nous proposons un nouveau calcul de la procédure de moyenne bayésienne d'estimation classique basé sur l'estimateur simple et efficace corrigé par un estimateur à effets fixes. Notre référence et analyse de robustesse considèrent un grand nombre de variables, périodes d'échantillonnage et priors de modèle. Nous constatons que le stock existant de capacités (mesuré par la variable dépendante retardée), les termes de l'échange sur les matières premières, la disponibilité de l'énergie, la consommation publique, le capital par travailleur, les terres arables et les flux de capitaux montrent une association importante et robuste sur les capacités.

ABSTRACT

Recent contributions to the growth literature have argued that the structure of an economy, as measured by its productive capabilities, is a key determinant for inter-country differences in development. Productive capabilities have been shown to be highly predictive of future economic growth, yet their country-level determinants have remained unknown. In this paper, we empirically explore their determinants using a model averaging framework that can handle a very large number of explanatory variables without the need for model selection. In order to estimate our dynamic panel specification, we propose a novel Bayesian Averaging of Classical Estimates procedure based on the simple and efficient bias-corrected LSDV estimator. Our baseline and robustness analysis consider a large number of variables, sample periods and model priors. We find that the existing stock of capabilities (as measured by the lagged dependent variable), commodity terms of trade, energy availability, government consumption, capital per worker, arable land and capital inflows show a strong and robust association with capabilities.

JEL Classification: C11; O11; O14; O33; F43

Keywords: capabilities; economic complexity; exports; growth; Bayesian model averaging.

I. INTRODUCTION

Development economists have often argued that the type and range of goods that a country produces and exports is at the core of structural transformation, and ultimately economic development. In essence, economic development is often defined as the transition of production factors from traditional activities, largely in agriculture towards value creation embodying greater physical and human capital (Kuznets, 1955 ; Kaldor ,1963). A major part of the literature has focused on the special role of the abundance of natural resources in developing economies. Primary commodities have been argued to lead to deteriorating terms of trade for natural resource countries (Prebisch, 1950), greater volatility in export revenues (Easterly and Kraay, 2000), Dutch disease (Corden, 1982), and poor institutional quality (Sachs and Warner, 1995; Ross, 2001; Collier and Hoeffler, 2005). More recent research has shifted the focus from differences in sector shares to differences in product characteristics. Accordingly, the process of structural transformation involves a change in the type of goods that are produced, i.e. shifting production from simple to complex or sophisticated goods (Lall, 2000; Hausmann and Klinger, 2006; Hausmann et al., 2007). In this literature, what a country produces is indicative of its developmental stage, with advanced economies generally producing and exporting high-technology and high-skill products.

In a recent contribution to this literature, Hidalgo and Hausmann (2009) propose a network theoretic measure of the productive capabilities of an economy based purely on trade data. They do so using the “method of reflections”, which combines the level of export diversification of a country and the average ubiquity of the goods it exports using a bipartite network algorithm. The rationale behind this approach is that countries that export a large range of products (i.e. are highly diversified) are likely to have more capabilities. Similarly, products that are exported by relatively few countries (i.e. products with low ubiquity) seem to require many of hard-to-find capabilities. Hidalgo and Hausmann (2009) show that after controlling for initial levels of development the capabilities measure is highly predictive of future economic growth.¹ This work has attracted considerable interest in the literature and international development institutions (Jankowska et al., 2012; Lederman & Maloney, 2012; UNESCAP, 2011 and many others).

If capabilities play such an important role for economic growth, it is crucial to understand the process of capability formation. Hausmann and Hidalgo (2011) build a theoretical model that

¹ See also Ourens (2013) for a use of alternative data sources and several robustness checks of this finding.

features strong path dependence in capability accumulation, which they define as the *quiescence trap*. Since many products require the combination of several capabilities, countries with few capabilities will have a lower probability of finding alternative uses for any additional capability than countries having a large pool of capabilities. In this way, countries with low capabilities will have lower incentives to accumulate additional capabilities. According to the authors, this trap helps to explain the persistent differential in income levels between countries, since initial differences in capability endowments would be amplified overtime.

The way the relationship between a country's export profile and development is portrayed in the literature on capabilities also suggests a very specific set of policy recommendations, in particular with regards to industrial policies. According to Hausmann et al. (2011) countries should follow a targeted diversification strategy that favours expansion into more sophisticated activities and aim to increase the number of exported products according to two criteria. First, selecting those which are most sophisticated and hence promise the greatest increase in the country's capabilities. Second, choosing them according to their proximity, i.e. the match with existing capabilities.

All things considered, the literature on capabilities arrives at an interesting crossroad, which motivates our study. On the one hand, the analysis draws bold conclusions about the link between the theoretical construct of capabilities and economic development, some of which depart substantially from conventional trade theory. Yet, this contrasts with the ambiguity that still surrounds the definition of capabilities: Hidalgo and Hausmann (2009) describe capabilities as those requirements that are necessary for producing a certain set of goods and are essentially non-tradeable. Generally, these would include a wide range of factors such as aspects of governance (e.g. rule of law, property rights), the level of infrastructure, and the skill endowment of the labour force. The vice and virtue of the computation of the capabilities variable is that the method of reflections only requires international trade data as input. In doing so, it effectively circumvents the question of which country-level determinants account for the level of capabilities.

The lack of clarity on the determinants of capabilities also has ramifications for policy recommendations. If capabilities can be improved by exporting more sophisticated goods, the main policy advice that results is a targeted support of industries based on the sophistication of their products, and their proximity to the current set of capabilities. Even though it is reasonable to assume that the products a country exports without market intervention are a measure of their productive capabilities, to start exporting a particular good at any cost because of the above considerations might not necessarily improve the capabilities of that country. This will be particularly true if structural factors such as the market structure, the factor endowments and infrastructure of a country are not favourable for producing and exporting that particular good. However, if the country-level determinants of capabilities were known an alternative course of action would be the improvement of a particular set of structural variables (e.g. education, infrastructure, business environment, etc.) that can be seen as conducive to the accumulation of

capabilities, which would then facilitate the inception and growth of more sophisticated industries.

Our study aims to shed light on the determinants of capabilities through an empirical analysis that considers a broad set of potential variables. These encompass institutional, macroeconomic, skill-related, and other “non-tradeable” features of an economy, as suggested in Hidalgo and Hausmann (2009). Specifically, the set of variables we consider can be grouped into different categories related to macroeconomic conditions, financial development, economic structure, external sector, demographics, education and skills, infrastructure and importance of the primary sector. In all, this comprehensive approach allows us to combine variables where the room for policy influence is more evident, as would be case for a variable such as the fiscal deficit; to those that are relatively unaffected by policy, either for being structural in nature (e.g. demography), or for being partially dependent on external factors (e.g. terms of trade). Consequently, the extent to which the policymaking process can influence the capabilities endowments of a country will largely depend on which type of determinant turns out to be more important.

The consideration of a large set of potential explanatory variables requires a robust estimation procedure, for which Bayesian Model Averaging (BMA) is the method of choice. As we will see below, BMA is well-suited to account for the uncertainty associated with model selection, a feature that gains importance in empirical specifications that consider many explanatory variables. In our setting, it is desirable to use a fixed effects specification to account for unobserved heterogeneity across countries. Due to the persistence of the capabilities measure over time, this variable needs to be included in lags in the estimation equations. However, the lagged dependent variable in the presence of country-fixed effects leads to the well-known Nickel bias of the least squares dummy variable (LSDV) estimator (Nerlove, 1971; Nickell, 1981). Here, we develop a new BMA procedure based on the simple and efficient bias corrected LSDV estimator Kiviet (1995, 1999) to estimate the dynamic panel specification.

The remainder of the paper proceeds as follows. Section II details the construction of the measure on capabilities, as defined in Hidalgo and Hausmann (2009); it describes the BMA methodology and introduces of the bias-corrected LSDV estimator into the BMA framework. Section III discusses the results and their robustness, and section IV concludes.

II. METHODS AND EMPIRICAL STRATEGY

In this section, we present in detail how the capabilities variable used as dependent variable in our analysis is constructed. Second, we discuss in detail the Bayesian model averaging approach used in this study, as well as our preferred estimator, the priors and computational implementation. Finally, we introduce the definitions and sources of the explanatory variables included in the empirical analysis.

Capabilities measure

The capabilities measure proposed by Hidalgo and Hausmann (2009) is based on an algorithm that combines information of the diversification of a country and the ubiquity of the products it exports. Whether a country is effectively exporting a good is assessed by its revealed comparative advantage RCA_{cp} following Balassa (1965) as the ratio of the export share x_{cp} of product p in country c to the export share of product p in world trade

$$RCA_{cp} = \frac{x_{cp}/\sum_p x_{cp}}{\sum_c x_{cp}/\sum_c \sum_p x_{cp}}$$

A country is assumed to be competitive in exporting a product p if its revealed comparative advantage (RCA_{cp}) is greater than 1, i.e. if the share of product p in country c 's export basket is greater than the share of the same good globally. Therefore, we can define the following indicator function for country c having a comparative advantage in product p as:

$$|RCA|_{cp} = \begin{cases} 0 & \text{if } RCA_{cp} < 1 \\ 1 & \text{if } RCA_{cp} \geq 1 \end{cases}$$

The observed level of diversification $K_{c,0}$ of a country c is then defined as the number of products in which the country has a revealed comparative advantage. Conversely, the ubiquity $K_{p,0}$ of a product p is defined as the number of countries exporting that product with a revealed comparative advantage:

$$K_{c,0} = \sum_p |RCA|_{cp}$$

$$K_{p,0} = \sum_c |RCA|_{cp}$$

Next, these two indicators are combined using a bipartite network representation of countries and products, in which countries and products are connected if a country has a revealed comparative advantage greater than one in that particular product category

$$K_{c,N} = \frac{1}{K_{c,0}} \sum_p |RCA|_{cp} K_{p,N-1}$$

$$K_{p,N} = \frac{1}{K_{p,0}} \sum_c |RCA|_{cp} K_{c,N-1}$$

where $K_{c,N}$ is the N th country moment and $K_{p,N}$ is the N th moment on the product side. Iterating the equations for $K_{c,N}$ and $K_{p,N}$ gradually extracts more information about product sophistication, on the product side, and capabilities, on the country side, and this procedure was iterated until convergence. Then, the capabilities measure was computed based on a sample of 103 countries for which data for the entire time period from 1976-2010 is available.

The actual value of the measure is sensitive to the overall connectivity in the network, which changes over time, and hence comparisons across years are only meaningful if the measure is standardised across countries for each year:

$$capab_c = \frac{K_{c,\infty} - E(K_{c,\infty})}{\sqrt{Var(K_{c,\infty})}},$$

where time subscripts have been omitted to simplify notation. A value of zero in this measure corresponds to a country having the same capabilities as the world average; a value of one corresponds to a country that is one standard deviation above world average. When observing changes over time in this measure, one can determine whether a country has improved its position relative to other countries, while of course it is likely that on average all countries have improved their capabilities over time.

The standardisation of the capabilities measure has consequences for our statistical analysis. Since the dependent variable is by definition a *relative* measure, the actual levels of the explanatory variables are not meaningful and only changes in their values relative to other countries are of importance. Hence, all the independent variables are standardised accordingly for a given cross-section of the panel. This is equivalent to running a regression with standardised coefficients. Therefore, the coefficients are to be interpreted as how many standard deviations the dependent variable changes for a standard deviation increase in the explanatory variables.

Bayesian model averaging

As discussed in the introduction, a large number of variables could in principle determine the productive capabilities of a country. If the number of explanatory variables (K) is substantially smaller than the number of observations (N), a simple regression could be run to identify the most important determinants. However, if K is close to N , statistical estimates become too imprecise to yield robust and meaningful results. A popular approach in the growth literature to overcome this problem has been to use Bayesian model averaging (Fernández et al., 2001; Sala-i-Martin et al., 2004; Ciccone and Jarocinski, 2010; Moral-Benito, 2012). Bayesian model averaging focuses on the fact that there is substantial model uncertainty, i.e. that it is not clear which model out of the many possible specifications given a set of K explanatory variable has generated the data. Model uncertainty is accounted for by estimating all possible models and by evaluating the probability of each model given the data. Subsequently, the (posterior) model probability can be used to calculate a proxy for the importance of a variable (the posterior inclusion probability) and for calculating weighted averages of the mean and variance of the coefficients.

Formally, there are $j = 1, \dots, 2^K$ different models formed by all combinations of the K explanatory variables. For a given model M_j with a subset of independent variables x^j consider the panel regression of the form

$$y_{it} = x_{it}^j \beta^j + \varepsilon_{it}^j$$

where y_{it} is the capabilities measure of country i in time period t , β^j are the marginal effects of the explanatory variables included in the model and ε_{it}^j is an error term. In this article, we use Bayesian model averaging of classical estimates following Raftery (1995) and Sala-i-Martin et al. (2004), for which standard econometric methods can be used to estimate this equation. A key statistic in the model averaging framework is the posterior model probability $P(M_j|y)$ of model M_j . We follow Raftery (1995) in using the Schwarz asymptotic approximation to the integrated likelihood which in a panel context is given by

$$P(M_j|y) = \frac{P(M_j)(NT)^{-k^j/2} SSE_j^{-(NT)/2}}{\sum_i^{2^K} P(M_i)(NT)^{-k^i/2} SSE_i^{-(NT)/2}}$$

where $P(M_j)$ is the prior model probability, NT is the number of observations (N is the cross-section and T the time dimension), k^j is the number of parameters included in the model and SSE_j is the sum of squared residuals of model regression j . Note that the posterior model probability comprises the degrees of freedom adjustment $(NT)^{-k^j/2}$, as information criteria do, independent of the prior. An important summary statistic is the posterior inclusion probability (PIP), which is a proxy for the importance of a variable in explaining changes in the dependent variable. The PIP of explanatory variable k is given by the sum of the posterior probabilities of all models in which the variables was included:

$$PIP_k = P(\theta_k \neq 0|y) = \sum_{\theta_k \neq 0} P(M_j|y)$$

In a Bayesian framework the posterior density is simply the sum of the posterior densities weighted by their posterior model probabilities. Accordingly, the posterior mean of the coefficients is given by:²

$$E(\theta|y) = \sum_j^{2^K} P(M_j|y) \hat{\theta}^j$$

where $\hat{\theta}^j$ is the estimate of θ using the estimator discussed below for the set of explanatory variables included in model M_j . Following Leamer (1978) the posterior variance of θ is given by

$$\begin{aligned} Var(\theta|y) &= \sum_j^{2^K} P(M_j|y) Var(\theta|y, M_j) \\ &\quad + \sum_j^{2^K} P(M_j|y) (\hat{\theta}^j - E(\theta|y))^2 \end{aligned}$$

which accounts for the variance of the estimated parameters in individual models as well as the variance of the estimates across models.

The estimator

Here we consider a dynamic panel specification with persistence of the dependent variable and country-specific effects³ to account for unobserved heterogeneity across countries, given by:

$$y_{it} = \alpha y_{it-1} + x_{it}\beta + f_i + u_{it}$$

Under the assumption that $E(u_i|y_i, x_i, f_i) = 0$, the least square dummy variable (LSDV) estimator defined as:

$$\hat{\theta}_{LSDV} = (X'AX)^{-1}X' Ay$$

is unbiased, where X includes both the lagged dependent variable and other explanatory variables and $A = I_N \otimes (I_T - \frac{1}{T} \iota_T \iota_T')$ is the symmetric and idempotent matrix subtracting individual-specific means. However, it is well known that the LSDV estimator is biased when $E(u_i|y_i, x_i, f_i) \neq 0$ and T is small (Nerlove, 1971; Nickell, 1981). Kiviet (1995, 1999) derived an analytical expression of the small sample bias of the LSDV estimator up to order $O(N^{-2}T^{-2})$ and proposed a bias-corrected estimator based on his results:

² Summing over the models where $\theta_k \neq 0$ instead of over all models yields the posterior mean of the coefficients conditional upon inclusion.

³ We also implicitly account for time-specific effects due to the transformation of the data resulting from the standardisation of the variables.

$$\hat{\theta}_{LSDVC} = \hat{\theta}_{LSDV} - E(\hat{\theta}_{LSDV} - \theta)$$

The bias corrected LSDV estimator (LSDVC) generally has been shown to be a more efficient estimator than alternative IV and GMM-estimators that also address the bias arising from the lagged dependent variable (Anderson and Hsiao, 1981; Arellano and Bond, 1991). For example, Judson and Owen (1999) use Monte Carlo simulations to show that for panels of all sizes the bias corrected estimator consistently has the lowest root mean squared error in comparison to OLS, Anderson-Hsiao and GMM estimators. Moral-Benito (2012) proposes a maximum likelihood estimator whose first moment conditions correspond to those of the GMM problem to deal with the endogeneity of the lagged dependent variable in a Bayesian model averaging context. However, the small-sample properties of this estimator have so far not been investigated. Hence, in this study we combine the estimates of the LSDVC estimator using the model averaging framework outline above. In the robustness analysis, we show that the model averaging results of the LSDV and the LSDVC estimator differ appreciably from each other and that the bias of the LSDV estimator for our dataset is not negligible.

The prior

We use the relatively standard binomial distribution over the model size K as a prior distribution (Sala-i-Martin et al., 2004):

$$K \sim \text{Bin}(K, \varphi)$$

$$E(K) = K\varphi = K \frac{\bar{k}}{K} = \bar{k}$$

where $\varphi = \bar{k}/K$ is the prior inclusion probability of each variable and \bar{k} is the prior expected model size. In this framework, the prior model probability of model M_j with k^j different explanatory variables is

$$P(M_j) = (\varphi)^{k^j} (1 - \varphi)^{K - k^j}.$$

This specification only requires the choice of the prior expected model size, which we set to $\bar{k} = 10$. This is slightly larger than in previous studies on the determinants of growth in cross-section regressions with $\bar{k} = 7$ (Sala-i-Martin et al., 2004) and in the panel setting with $\bar{k} = 5$ (Moral-Benito, 2012). However, as there is currently little prior knowledge on the determinants of productive capabilities and the measure has potentially many determinants we chose a slightly larger expected model size as a reflection of this uncertainty.

We investigate the robustness of our results to our prior assumptions in two ways. First, we examine the effect of the prior expected model size of the binomial prior on our results. In the alternative specification \bar{k} is set to 5 and 15. Second, we use a random rather than a fixed prior specification. Using a fixed φ prior has been criticised by Ley and Steel (2009) on the grounds that the results are more sensitive to the choice of the prior expected model size parameter \bar{k} than

alternative hierarchical prior specifications which allow φ to be random. To address this issues, Ley and Steel (2009) propose the use of a beta-binomial prior over the model size K

$$\begin{aligned} K &\sim \text{Bin}(K, \varphi) \\ \varphi &\sim \text{Beta}(a, b), \end{aligned}$$

where $a, b > 0$ are hyperparameters and φ is now a probabilistic variable that is drawn from a beta distribution. The prior expected model size in this framework is

$$E(K) = \frac{a}{a+b} K.$$

Ley and Steel (2009) propose a choice of $a = 1$ and $b = (K - \bar{k})/\bar{k}$, which allows setting the prior expected model size as in the case of the binomial prior and hence makes the results directly comparable. In our robustness analysis, we set $\bar{k} = \{5, 10, 15\}$ to compare the results with those of the fixed binomial prior specification.

Computational implementation

The number of models that has to be evaluated grows exponentially with the number of covariates included in the regression. A complete evaluation of all models with K covariates would require running 2^K regressions. For $K = 30$, this is equivalent to $1.07 \cdot 10^9$ model evaluations, which is computationally infeasible. As a solution, algorithms have been developed which carry out Bayesian model averaging without the need for evaluating every possible model (Koop, 2003). Here, we employ the commonly used Markov Chain Monte Carlo Model Composition (MC³) algorithm originally proposed by Madigan, York and Allard (1995). The MC³ constructs a Markov Chain in model space whose stationary distribution converges to the posterior model probability distribution by sampling from regions in model space where the posterior model probability is high. Specifically, the MC³ simulates a chain of models $M^{(s)}$ for $s = 1, \dots, S$ samples where $M^{(s)}$ is sampled from the set of all possible models $\{M_1, \dots, M_{2^K}\}$. The birth-and-death sampler is used to implement the MC³, in which the candidate model $M^{(s)}$ is identical to the current model $M^{(s-1)}$ except for one variable. In each iteration, one of the K covariates is drawn at random. If the explanatory variable already forms part of the current model $M^{(s-1)}$, then the variable is deleted in the candidate model $M^{(s)}$. Similarly, if the drawn explanatory variable is not included in the current model, then it is added in the candidate model. The posterior model probability for the candidate model is computed and the proposal is accepted with probability:

$$\min\left\{\frac{P(M^{(s)}|y)}{P(M^{(s-1)}|y)}, 1\right\}.$$

If the candidate model is not accepted, then the Markov Chain remains at the current model ($M^{(s)} = M^{(s-1)}$). The algorithm was run until convergence, which was verified by the correlation

coefficient between the relative frequencies of model visits and the analytical posterior model probabilities for the 10 000 best models (Fernández, Ley and Steel, 2001).

Data

Our data sample is conditioned by the need to incorporate a comprehensive set of variables, representative of all the dimensions that could potentially affect capabilities.⁴ While a more complete definition of the variables and their sources is included in the appendix, here we simply list the variables used in the baseline specification, along the following thematic classification.

- *Macroeconomic conditions*: GDP volatility, GDP growth, current account balance, government consumption over GDP, inflation, total reserves over GDP, currency overvaluation.
- *Financial Development*: money and quasi-money as percentage of GDP, domestic credit over GDP.
- *Physical capital*: capital per worker, gross capital formation over GDP.
- *External sector*: Trade openness, capital account openness, foreign direct investment over GDP, capital flows over GDP, terms of trade.
- *Demographics*: population, population growth, age dependency ratio, life expectancy.
- *Education/Skills*: average years of schooling, percentage of population with secondary education, percentage of population with tertiary education.
- *Infrastructure*: number of telephone lines per 100 people, energy use, energy production,
- *Importance of primary sector*: agricultural land as percentage of total land, arable land as percentage of total land.

The implementation of BMA requires having a balanced panel, a condition that significantly reduces the number of countries for which the previous variables are available. This condition leaves our baseline specification with 43 countries, listed below according to latest 2012 World Bank's income group classification.⁵

⁴ The rationale for the different groups of variables is discussed in the introduction.

⁵ See <http://data.worldbank.org/news/new-country-classifications> for details.

Lower middle income	Upper middle income	High income
Bolivia, Cote d'Ivoire, Cameroon, Egypt, Ghana, Guatemala, Honduras, Sri Lanka, Morocco, Pakistan, Philippines, Paraguay, Senegal, El Salvador.	Colombia, Costa Rica, Dominican Republic, Ecuador, Jamaica, Jordan, Mexico, Malaysia, Peru, Thailand, Turkey.	Australia, Canada, Switzerland, Denmark, Finland, Ireland, Israel, Italy, Japan, Republic of Korea, Netherlands, Norway, New Zealand, Sweden, Trinidad and Tobago, United States, Uruguay

Note: Classification based on Gross National Income per capita of 2012 by the World Bank with the following brackets: lower middle income (<USD 1 036 - USD 4 085), upper middle income (USD 4 086 - USD 12 615), high income (USD 12 616 or more).

With these countries and variables, we create 5-year-average periods between the years 1976 and 2010, therefore leading to seven time periods. As a robustness check, we estimate two alternative samples that in both cases include the same number of countries and a shorter time span (i.e. only the most recent four periods, between 1991 and 2010).

III. RESULTS

Next, we present the results of estimating the alternative specifications and applying the Bayesian model averaging method discussed in the previous section to analyse the determinants of productive capabilities. We first discuss our baseline results. Then we perform a series of robustness tests considering alternative estimation methods, samples, additional variables and different assumption regarding the priors. Finally, we extend the analysis towards interpreting our results separately in the two dimensions (diversification and ubiquity) that combine in the capability measure.

Baseline Results

In Table A.1 (see Annex), we present the estimates for our baseline model that are based on the Bayesian Averaging of Classical Estimates (BACE) LSDVC estimator considering the long panel (from 1976 to 2010). A series of interesting results emerge. On the one hand, in terms of importance of the potential variables considered, we find three variables that stand out with a PIP of 1: the lagged dependent variable, commodity terms of trade and real exchange rate deviations from trend. The substantial autocorrelation in capabilities, with an average coefficient of around 0.51, shows that it is actually useful to think about capabilities as a stock that moves slowly and has to be accumulated over time. On the other hand, variables that *a priori* would be considered to play an important role in the accumulation of capabilities such as human capital, infrastructure, trade and financial openness do not seem to have a robust effect, in the sense that they have low PIPs.

Commodity terms of trade have a negative impact on capabilities. Therefore, improvements in the ratio of export commodity prices versus import commodity prices are associated with a decline in the capabilities index. This effect could in principle be due to different reasons that operate mainly through a combination of impeding diversification and product sophistication. Commodity exporting countries might specialise in exporting less sophisticated products, for example, due to Dutch disease effects that reduce the competitiveness of other tradeable goods. Furthermore, as commodity sectors tend to be very intensive in physical capital, it could be that they create very little employment and spill-over effects within their sector. When the relative prize of commodity exports increases, this feature might become more severe. Finally, upstream and downstream spill-overs might also be limited in natural resource activities. For example, in the mining sector it could potentially involve moving into logistics, transport or machinery,

activities that require very different skills, inputs and industrial capabilities that might not be close in technological terms to capabilities accumulated in the mining production stage.

The third variable that has a PIP of 1 and also a negative, albeit quantitatively very small, impact on capabilities is the undervaluation of the real exchange rate. This is an interesting finding, as the literature generally argues that an undervalued real exchange rate fosters economic growth through self-discovery or learning-by-doing externalities that are associated with higher capabilities (Rodrik, 2008). Nevertheless, it is consistent with other studies that do not find a significant trade effect of currency undervaluation, such that currency undervaluation might foster growth by other channels than through raising export capabilities: e.g. facilitating the Lewis transition from the traditional rural areas to the modern sector – or keeping labour cheap and allowing therefore for more savings in the corporate sector that can be reinvested in the context of low financial development (see Glüzman et al, 2012). A potential channel of transmission from real appreciation towards capabilities might be that an overvalued exchange rate often reduces the relative price of imported capital goods.

Additionally, there is a group of five variables that also have a relatively high PIP. These variables are energy use intensity, government consumption, the capital stock per worker, capital inflows and arable land.

Energy use intensity has a positive effect on capabilities, which is expected from a variable that is a proxy of energy availability. Government expenditures also have a positive relationship with capabilities, likely reflecting the provision of public services needed for enabling the accumulation of capabilities by the private sector. This is an interesting result, as in the economic growth literature, government consumption is usually found to have a negative correlation with economic growth (e.g. Sala-i-Martin, 2004). However, both results are not contradictory. For example, it might be that the distortions caused by the way these expenditures are financed and the crowding-out effect of an increase in government expenditure on private investment in the absence of perfect access to international capital markets reduce economic growth, despite the fact that they might also create public goods needed for the accumulation of capabilities. The capital stock per worker has a positive effect on capabilities, such that relatively capital-abundant economies tend to have higher capabilities. This might be due to the fact that sophisticated technologies are often embodied in physical capital.

The other two variables, capital flows and arable land have a negative impact on capabilities. The channel of transmission for capital inflows has to differ from the real exchange rate, as the results discussed above imply that appreciations linked to inflow episodes would tend to have a positive impact on capabilities. In our view, the most reasonable alternative might be through macroeconomic instability and financial crises often associated with capital inflow episodes (Kaminski and Reinhart, 1999). This would also be consistent with the evidence that the type of crises associated with capital inflow reversals (sudden stops, banking and debt crises) are frequently followed by a sustained decline in total factor productivity (Blyde et al, 2010). Arable

land also has a negative effect on capabilities, which probably reflects the fact that countries with relative abundance of land tend to specialise in primary goods with weak capabilities.

To provide some intuition for the magnitude of the coefficients, in particular for the policy-relevant variables, let us consider some examples for the most recent period, 2005-10, of our dataset. In our sample, Japan had the highest value of capabilities with 2.02, the Comoros the lowest value with -1.96 and Colombia was at the median with -0.13. According to our results, raising the level of government spending from the one of the Dominican Republic to the one in Sweden (average of 7.41% GDP vs. 26.4% GDP for the 2005-10 period) would be associated with an increase in the capabilities indicator by 0.15. Similarly, fostering investment to increase the stock of capital per worker from the level of Ghana to the one of the US (average of 10 373 USD at constant 2005 prices vs. 220 928) would raise the level of capabilities by 0.19.

Robustness analysis

Next, we explore a series of robustness checks of our main results. First, we run the model averaging procedure without correcting for the endogeneity bias deriving from the lagged dependent variable. Second, we consider the possibility that the impact of certain variables on capabilities has changed over time by restricting our analysis to the period from 1991 onwards. We also analyse the robustness of our results to different priors regarding the model size by varying the prior expected size, as well as using a random instead of a fixed prior specification.

LSDV estimation

Using the LSDV estimator without correcting for the bias deriving from the lagged dependent variable in the presence of country-fixed effects leads quantitatively to slightly different results from our baseline model (see Table A.2 in the annex). First, the coefficient of the lagged dependent variable has a downward bias using the LSDV estimator, which is expected since the coefficient is positive (Nickell, 1981). Furthermore, the PIP of the real exchange rate deviations from trend, energy use and government consumption is substantially reduced while the one for capital per worker is inflated. The altered PIPs in combination with the bias deriving from the lagged dependent variable also affect the associated average coefficients of the variables. The differences in the estimation results underline the importance of using the correct estimator when considering our dynamic panel dataset.

Shorter sample from 1991 onwards

Since our baseline specification spans a time-period of more than 35 years in which the world economy has undergone substantial structural changes, it is important to consider whether the determinants of capabilities might have concomitantly changed over time. This alternative

specification confirms some of the findings of our main results but also provides some important qualifications. The lagged dependent variable and the commodity terms of trade continue to show a PIP close to one and the sign of the average coefficient remains unchanged. Capital per worker appears even more closely linked to high levels of capabilities and is not only important in terms of the PIP, but now also has the highest average effect on capabilities. In contrast, the undervaluation of the real exchange rate has a substantially reduced PIP and the average estimate also reverses its sign, which turns positive in the short panel in accordance with Rodrik (2008). Similarly, energy use, government consumption and to a lesser degree capital inflows turn out to be much less closely linked to capabilities in the last 20 years than considering the whole time span.

Some additional variables show a stronger association with export capabilities during this period. For example, some factors that can be influenced via policy such as tertiary education and foreign direct investment seem to play an increasingly important role in the last two decades. This might reflect two trends of globalization recently highlighted by analysts and policy makers. First, the reduction in transportation and communication costs have increased the fragmentation of the production process in global value chains, with FDI becoming an important vehicle in this field, such that vertical integration and trade in tasks has become more relevant (WEF, 2012; OECD, 2013). Second, technological change has been skilled-biased, such that the complementarities between human capital, innovation and upgrading have strengthened.⁶ Both of these trends should in principle facilitate the accumulation of export capabilities in countries successful in inserting themselves in global value chains through FDI and accumulating human capital.

Interestingly, agricultural land is also positively linked with capabilities. However, structural demographic factors associated to low levels of development, such as population growth and the dependency ratio that put a drag on the accumulation of capabilities have also gained in importance.

Robustness to different prior specifications

In a Bayesian framework, the prior probability distribution and its parameters need to be chosen by the researcher. In this section, we investigate the robustness of our results to the prior expected model size \bar{k} and the particular probability distribution of the prior. First, we consider the effect of changing the prior expected model size of the binomial prior on our results. In the alternative specifications \bar{k} is set to 5 and 15 instead of 10, as in the previous section (Figure A.1, Panel a and b in the Annex). Second, we use the beta-binomial prior specification proposed by Ley and Steel (2009), which has been argued to be less sensitive to the choice of the prior expected model size parameter \bar{k} due to its probabilistic formulation. Again we consider a prior expected model size of 5, 10 and 15 in this alternative specification.

⁶ See Acemoglu (1998, 2002) and Autor et al (1998) for some theory and evidence on this issue.

Decreasing the prior expected model size to $\bar{k} = 5$ favours models with fewer variables, which leads to half the probability mass being concentrated on models with five explanatory variables. Increasing the prior expected model size to $\bar{k} = 15$ leads to putting more weight on models with more variables and shifts the posterior model probability distribution to the right as well as making it more spread out. Varying the prior expected model size of the binomial prior does not qualitatively change our results with regards to the PIPs (Figure A.1 – Panel b in Annex). The lagged dependent variable, commodity terms of trade and exchange rate undervaluation are always included in all models independent of the prior expected model size. Similarly, there is a set of variables including energy use, government consumption, capital per worker, capital flows and arable land that have a PIP between one and 0.2 for all values of \bar{k} . The other explanatory variables are not important in any specification and always have a PIP between 0.2 and zero. In general, the PIP of every variable for $\bar{k} = 5$ is lower than in our preferred specification. The PIP of most variables is reduced because the lagged dependent variables, commodity terms of trade and exchange rate undervaluation are always included in all regressions, which leads to all other variables being excluded in small models. This reduces the PIP of these variables since small models receive the largest weighting. For $\bar{k} = 15$ the PIP of every variable is generally higher than in our preferred specification and the converse reasoning applies in this case.

In contrast to the binomial prior with the same value of \bar{k} the beta-binomial prior is more spread out across different model sizes and has more probability mass on small models. The model likelihood, corrected for the combinatorial effect of overweighting medium-sized models, is also more concentrated on small models with virtually no probability mass on models with more than 10 variables. As a consequence the use of the beta-binomial prior leads to smaller posterior model sizes than the binomial prior (Figure A.1 – Panel c in Annex). The lower posterior model size has repercussions on the PIP of the explanatory variables, which is also generally lower for most variables (Figure A.1 – Panel d in Annex). As described above, this is due to the fact that certain variables are (almost) always included in the regressions and therefore leave less room for other variables in small models. Despite these small quantitative differences, the general results remain qualitatively unchanged. The lagged dependent variable and commodity terms of trade are still the most important variables. Exchange rate undervaluation remains important, but is now no longer always included for $\bar{k} = 5$ and $\bar{k} = 10$. Energy consumption, government consumption, capital per worker, capital flows and arable land still have a PIP which is markedly larger than zero. All other variables have a PIP very close to zero as before. The only exception is GDP growth volatility which turns out to be important for small model priors.

Effects through the building blocks of capabilities: diversification and ubiquity

In the preceding econometric analysis, we have treated the capabilities variable as an abstract measure for the export potential of a country. Additional insights into the relation between the capabilities measure and our set of explanatory variables can be gained by considering its relation with the international export data that it is based on. The capabilities variable is computed by combining the diversification of a country – the 0th country moment – with the

ubiquity of the goods – the 0th product moment – it exports using a bi-partite network representation of world trade. Intuitively, diversification and the average ubiquity of goods have the following relation to capabilities. A highly diversified country requires many different skills to produce the large number of goods it exports and therefore should have a high level of capabilities. Similarly, a country whose goods are on average not very ubiquitous has skills that are scarce and therefore also should have a high level of capabilities relative to other countries. The 1st country moment is the average ubiquity of the goods a country exports. The 2nd country moment is the average diversification of the countries that export the products a country exports and so forth. That means that initially even country moments are more related to the diversification of a country and odd country moments are more related to the ubiquity of goods. Note that the two sources of information become increasingly intertwined and, when standardised, eventually converge to the same values in absolute terms, which is the measure of capabilities proposed by Hidalgo and Hausmann (2009).

Additional regressions using lower country moments can therefore determine whether the explanatory variables that turned out to be strongly associated with high capabilities are more related to a) the diversification of a country, b) the average ubiquity of the products it exports or c) a non-trivial combination of the two sources of information. To this end we re-ran our model averaging algorithm for the long sample with the same variables as before, but instead of using the capabilities measure as the dependent variable, we included the diversification (the 0th country moment) and the average ubiquity of the goods a country exports (the 1st country moment) in the regressions.

Figure A.2 in the Annex shows the PIP of all explanatory variables arranged by their importance in our baseline estimation as a function of country moments. As one might expect, there is a lot of persistence in the export structure of countries and hence the lagged dependent variable is always included in all regressions. The same applies – albeit at lower PIP values – for capital per worker. The results also indicate when the capabilities measure really adds some value and extracts higher-order information about capabilities that is neither exclusively contained in a country's diversification nor average product ubiquity. Here we can differentiate between two cases. First, there could be a strong relation of an explanatory variable with capabilities that does not fully show up with either export diversification or average product ubiquity. This is the true for the variables that turned out to be key determinants of capabilities in the previous section, which show only a medium (commodity terms of trade, currency undervaluation, government consumption) and a very low association (energy use capital flows, arable land) with product ubiquity and almost no relation with diversification. This means that price conditions and government consumption exert their influence on capabilities particularly through their impact on product upgrading, but that additional information extracted by the amalgamation of product and country information accounts for the major part of the effect. Second, an explanatory variable could be irrelevant for capabilities, while considering diversification and average product ubiquity alone might lead us to conclude that there is an association. This is the case for the dependency ratio and population size, which seem linked to average product ubiquity but not capabilities, and for tertiary schooling, GDP growth volatility, foreign exchange reserves,

agricultural land, capital account openness, phone lines per capita and life expectancy, which matter for diversification but not for capabilities.

IV. CONCLUSION

As recent contributions to the literature on international trade and economic growth have argued, the structure of exports – in particular the sophistication of exports baskets – is a good predictor of subsequent economic growth (Hausmann et al, 2007; Hausmann and Hidalgo, 2011). Therefore, it is important to understand what structural conditions of the economy as well as variables amenable to policy action are associated with export capabilities. Our paper's contribution is to analyse empirically what variables from a large set of potential drivers are significantly and robustly associated with export capabilities. In order to do so, we have used a Bayesian model averaging approach that can handle a very large number of explanatory variables without the need for model selection.

Our results show that commodity terms of trade are negatively associated with export capabilities, such that an increase in the relative price of commodity exports to manufacturing reduces export capabilities. This result is robust to different sample periods, estimation methods and priors regarding the prior expected model size. Furthermore, our disaggregated analysis on the moments associated with diversification and ubiquity show that the negative effect on export capabilities of commodity terms of trade changes takes place mainly through the product upgrading channel related to ubiquity rather than product diversification. We also find a positive and significant effect of the lagged dependent variable, which emphasises the importance of considering capabilities as a stock that evolves rather slowly over time. In the case of real exchange rate undervaluations, while the evidence generally indicated that it is important for economic growth, the sign of its effect is less clear. Considering a sample starting in the mid-1970s yields a negative effect of undervaluations on capabilities. This is consistent with Schröder (2013) who actually finds a negative impact of real exchange rate deviation from their equilibrium values, but does not support the results of Rodrik (2008) of a positive link to economic growth, nor his conjecture that undervaluated exchange rate reduce costs associated with self-discovery externalities or coordination failures. However, when considering a shorter panel from the 1990s onwards, the PIP of this variable is significantly reduced and the sign of the average estimated coefficient is reverted, becoming actually consistent with Rodrik's paper. Therefore, our paper is not conclusive regarding the use of exchange rate undervaluation as a strategy to upgrade and/or diversify the export basket of a country. More research is needed to sort out this issue.

Additional variables that have a relatively important and robust effect on capabilities include energy availability, government consumption, capital per worker, arable land and capital

inflows. The first three variables are positively related to export capabilities, probably reflecting the availability of important inputs and public goods, as well as embodied technological change in capital that allows for more diversification and product upgrading. The other two have a negative impact on capabilities, showing that countries relatively abundant in land might specialise too much in not very sophisticated goods, as well as the likely negative effects of capital inflows on financial stability. However, these results are somewhat weaker, when considering the last two decades only.

Finally, focusing on the most recent period, a couple of policy relevant variables appear to have a more important role in determining export capabilities. In particular, FDI attraction and tertiary education have a positive and robust impact on export capabilities from 1991 onwards. A possible interpretation is that these policy drivers have become more relevant in an era of globalization that has allowed a greater fragmentation of the production process across national borders and where technological change has been skilled-biased. If this process is to intensify, policies that facilitate the accumulation of the relevant human capital and facilitate the integration into global value chains might be successful in strengthening also export capabilities.

Looking forward, many questions remained unanswered. In particular, the need to have a relatively long and balanced panel has not allowed us to explore the importance of additional variables, such as different economic and political institutions as well as regulation such as production market regulations for entry, exit and competition in markets. Exploring these aspects would be a necessary step to map results into more concrete areas of policy recommendations, beyond promoting investment and human capital accumulation. In the meantime, some countries might learn from other countries that have successfully upgraded and/or diversified their economies despite having a large commodity export sector.

ANNEX

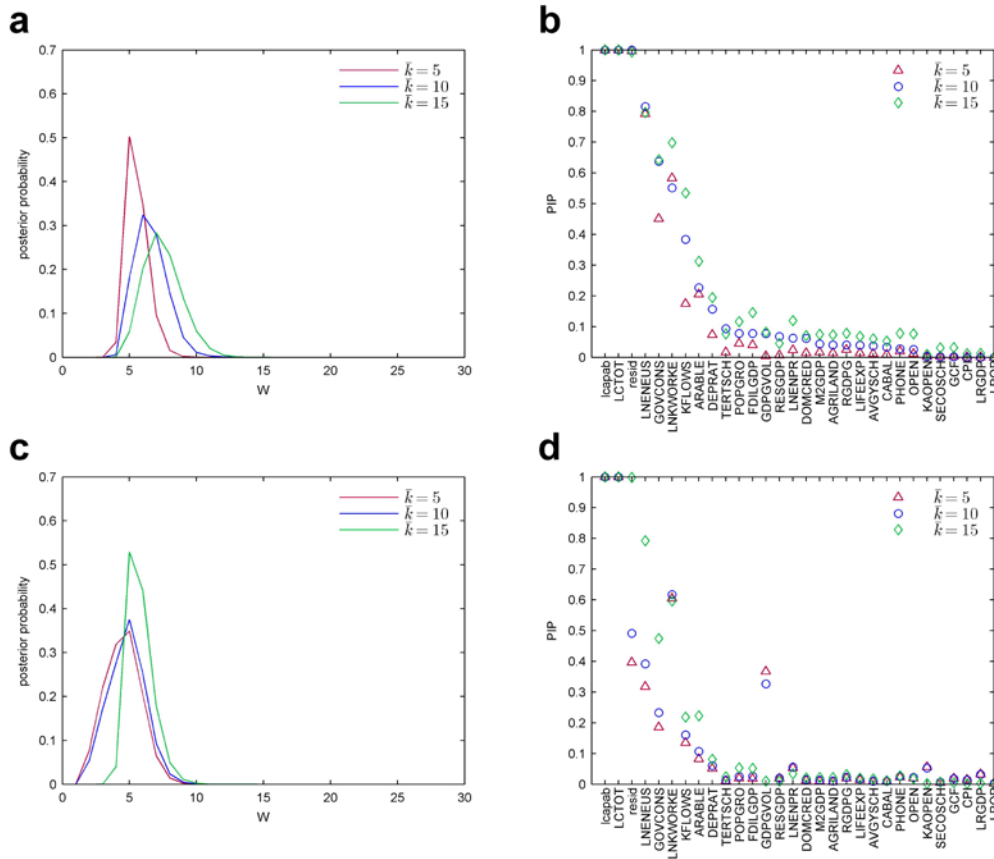
Table A.1. BACE results for the LSDVC estimator, long panel (5-yr averages)

	<i>PIP</i>	$E(\theta y)$	$E(\theta y, I)$	$Var(\theta y)$	$p(\theta > 0 y, I)$
Capabilities(t-1)	1.000	0.505	0.505	0.113	1.000
Commodity terms of trade	1.000	-0.051	-0.051	0.030	0.000
RER undervaluation	1.000	-0.009	-0.009	0.043	0.040
Energy use	0.816	0.027	0.033	0.203	0.708
Government consumption	0.638	0.042	0.066	0.093	1.000
Capital per worker	0.552	0.074	0.134	0.221	1.000
Capital flows	0.384	-0.012	-0.032	0.044	0.000
Arable land	0.226	-0.018	-0.081	0.226	0.000
Dependency ratio	0.156	-0.004	-0.025	0.078	0.017
Tertiary education	0.093	0.001	0.006	0.075	0.886
Population growth	0.078	0.001	0.008	0.057	0.828
Foreign direct investment	0.078	0.003	0.036	0.054	1.000
GDP growth volatility	0.077	-0.002	-0.021	0.033	0.000
Foreign exchange reserves	0.068	0.001	0.012	0.033	1.000
Energy production	0.062	-0.001	-0.021	0.112	0.000
Domestic credit	0.062	0.000	0.004	0.058	0.747
M2	0.043	0.001	0.018	0.068	1.000
Agricultural land	0.040	0.002	0.056	0.154	1.000
Real GDP growth	0.040	0.000	-0.010	0.023	0.000
Life expectancy	0.040	0.000	0.004	0.082	0.639
Average years of schooling	0.037	0.000	-0.006	0.119	0.041
Current account balance	0.033	0.000	0.006	0.028	0.719
Telephone lines per capita	0.028	0.001	0.039	0.126	1.000
Trade openness	0.026	0.002	0.064	0.073	1.000
Capital account openness	0.003	0.000	-0.047	0.045	0.000
Secondary education	0.002	0.000	-0.025	0.050	0.000
Gross capital formation	0.002	0.000	-0.012	0.026	0.000
Inflation	0	0	0	-	-
Real GDP	0	0	0	-	-
Population	0	0	0	-	-

Table A.2. PIP and $E(\theta|y)$ for the LSDV estimator and for the short panel specification

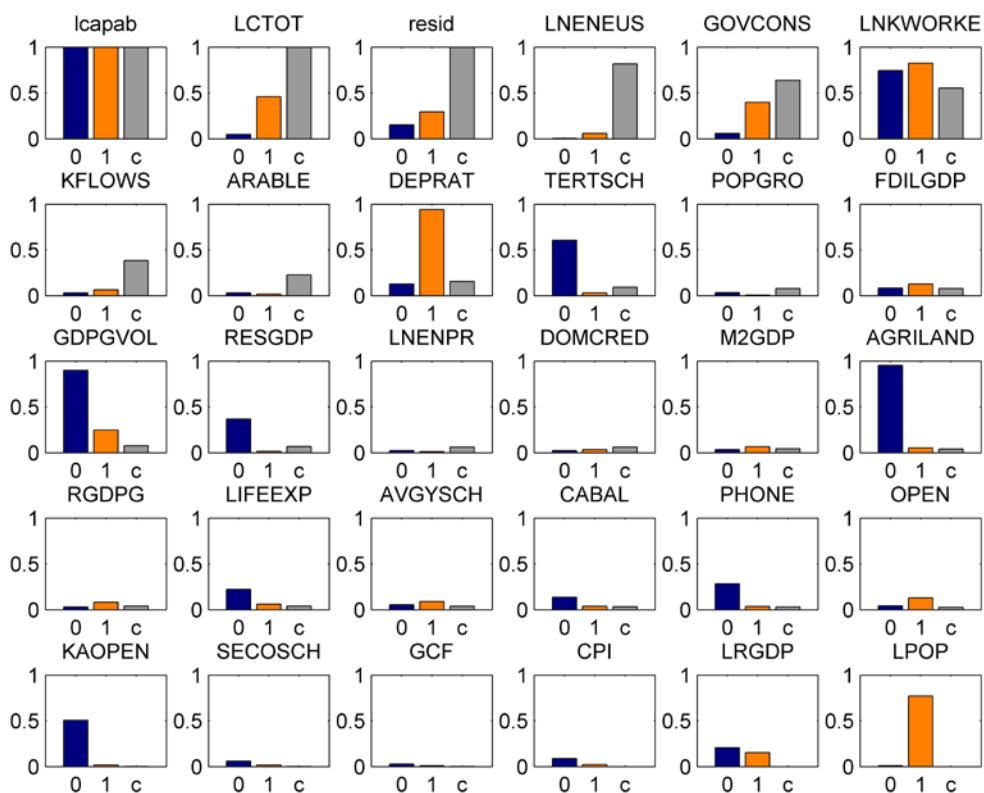
	PIP			$E(\theta y)$		
	LSDVC	LSDV	1991 –	LSDVC	LSDV	1991 –
Capabilities(t-1)	1.000	1.000	0.999	0.505	0.479	0.495
Commodity terms of trade	1.000	1.000	0.942	-0.051	-0.057	-0.029
RER undervaluation	1.000	0.031	0.317	-0.009	0.000	0.006
Energy use	0.816	0.054	0.008	0.027	0.003	0.000
Government consumption	0.638	0.206	0.044	0.042	0.010	0.001
Capital per worker	0.552	0.880	0.998	0.074	0.149	0.559
Capital flows	0.384	0.321	0.162	-0.012	-0.009	-0.002
Arable land	0.226	0.031	0.266	-0.018	0.000	-0.076
Dependency ratio	0.156	0.048	0.332	-0.004	-0.001	-0.034
Tertiary education	0.093	0.027	0.295	0.001	0.000	0.033
Population growth	0.078	0.029	0.840	0.001	0.000	-0.051
Foreign direct investment	0.078	0.056	0.728	0.003	0.001	0.023
GDP growth volatility	0.077	0.139	0.124	-0.002	-0.003	-0.002
Foreign exchange reserves	0.068	0.030	0.037	0.001	0.000	0.000
Energy production	0.062	0.047	0.031	-0.001	-0.002	0.001
Domestic credit	0.062	0.029	0.260	0.000	0.000	-0.012
M2	0.043	0.029	0.063	0.001	0.000	-0.001
Agricultural land	0.040	0.049	0.397	0.002	0.003	0.083
Real GDP growth	0.040	0.046	0.048	0.000	0.000	0.000
Life expectancy	0.040	0.029	0.024	0.000	0.000	0.001
Average years of schooling	0.037	0.030	0.088	0.000	0.000	0.009
Current account balance	0.033	0.034	0.012	0.000	0.000	0.000
Telephone lines per capita	0.028	0.048	0.095	0.001	0.002	0.007
Trade openness	0.026	0.092	0.025	0.002	0.003	0.000
Capital account openness	0.003	0.064	0.016	0.000	-0.001	0.000
Secondary education	0.002	0.031	0.015	0.000	0.000	0.001
Gross capital formation	0.002	0.042	0.085	0.000	0.000	-0.003
Inflation	0	0.030	0.051	0	0.000	0.001
Real GDP	0	0.050	0.072	0	-0.001	-0.018
Population	0	0.032	0.006	0	-0.001	0.003

Figure A.1



(a) Posterior model probability distribution for different values of the prior expected model size of the binomial prior. (b) Posterior inclusion probability of all variables for different values of the prior expected model size of the binomial prior. (c) Posterior model probability distribution for different values of the prior expected model size of the beta-binomial prior. (d) Posterior inclusion probability of all variables for different values of the prior expected model size of the beta-binomial prior.

Figure A.2



Posterior inclusion probability of all explanatory variables arranged by their importance in our baseline estimation (grey bars) as a function of country moments. 0: 0th country moment – diversification (blue bars). 1: 1st country moment – average ubiquity of the goods a country exports (orange bars). c: converge even country moment - capability variable (grey bars).

Category	Variable Name	Abbreviation	Description	Source	Summary Statistics (mean ; std dev)
Lagged Dependent Variable	Capabilities	Icapab	Author's calculations	Hidalgo and Hausmann (2009)	(0.3; 0.97)
Macroeconomic conditions	GDP Growth Rate	RGDPG	Annual percentage growth in GDP is calculated using Heston et al. (2012) PPP converted GDP per capita (chain series), at 2005 constant prices.	Heston et al. (2012)	(0.04; 0.03)
	GDP growth Volatility	GDPGVOL	standard deviation of the annual percentage growth in GDP over the observation period. Annual percentage growth in GDP is calculated using PPP converted GDP per capita (chain series), at 2005 constant prices.		(0.03; 0.02)
	GDP per capita	LRGDP	log of PPP converted GDP per capita (chain series), at 2005 constant prices.		(8.39 ; 1.3)
	Current Account Balance	CABAL	Current account balance is measured as the sum of net exports of goods, services, net income, and net current transfers as a percentage of GDP.		(-1.91; 4.74)
	Government Consumption	GOVCONS	General government final consumption expenditure as a percentage of GDP		(15.18; 5.82)
	Inflation (Consumer Price Index)	CPI	Annual percentage change in the cost to the average consumer of a basket of goods using the Laspeyres formula.		(29.76; 205.28)
	International Reserves	RESGDP	Total Reserves minus gold as a percentage of GDP in current US\$		(0.1; 0.09)
	Currency Undervaluation	resid	Author's calculations using Heston et al. (2012)		(-0.07; 0.42)
	Money Supply as a Percentage of GDP	M2GDP	Money and quasi money as a percentage of GDP		(57.1; 38.95)
	Domestic Credit to the Private Sector	DOMICRED	Domestic credit to the private sector as a percentage of GDP		(54.21; 45.34)
Physical capital	Capital Formation	GCF	Gross capital formation as a percentage of GDP	World Development Indicators	(21.96; 5.22)
	Capital Per Worker	LNKWORKE	Natural log of Capital per worker following Daude (2012)	Heston et al. (2012)	(10.75; 1.11)
	Capital Flows	KFLOWS	Private capital flows as percentage of GDP	World Development Indicators	(1.15; 3.15)
	Trade Openness	OPEN	Trade as a percentage of GDP	World Development Indicators	(66.84; 31.48)
External Sector	Capital Account Openness	KAOPEN	Index measuring a country's degree of capital account openness	Chinn & Ito (2008)	(0.45; 1.5)
	Foreign Direct Investment Liabilities	FDLGD	Foreign direct investment liabilities as a percentage of GDP.	Lane and Milesi-Ferretti (2007); Heston et al. (2012)	(0.2; 0.19)
	Commodity Terms of Trade	LCTOT	Natural log of commodity terms of trade. Commodity terms of trade is measured as the ratio of commodity export prices to commodity import prices, with each price weighted by the share of the relevant commodity in the country's GDP or total trade.	Tyrell and Spatafora (2009)	(4.61; 0.04)
	Population	LPOP	Natural log of population		(9.53; 1.24)
Demographics	Population Growth	POPGR0	Annual percentage growth in population	World Development Indicators	(1.62; 0.99)
	Dependency Ratio	DEPRAT	Age dependency ratio as a percentage of working-age population		(67.62; 17.38)
	Life Expectancy	LIFEEXP	Life expectancy at birth		(69.34; 8.29)
Education/Skills	Average Years of Schooling	AVGYSCH	Average total years of school		(7.34; 2.98)
	Secondary Schooling Level of Attainment	SECOSCH	Percentage of population aged 25 and over with secondary school completion as their highest level of attainment	Barro and Lee (2010)	(18.49; 12.2)
	Tertiary Schooling Level of Attainment	TERTSCH	Percentage of population aged 25 and over with tertiary school completion as their highest level of attainment		(8.55; 6.87)
Infrastructure	Phone Lines	PHONE	Number of telephone lines per 100 people		(20.19; 21.15)
	Energy Use	LNENEUS	Natural log of energy use measured in kilogram of oil equivalent per capita	World Development Indicators	(7.12; 1.05)
Importance of primary sector	Energy Production	LNENPR	Natural log of energy production measured in kiloton of oil equivalent		(9.21; 2.07)
	Agricultural Land	AGRLAND	Agricultural land as a percentage of total land area	World Development Indicators	(38.74; 20.72)
	Arable Land	ARABLE	Arable land as a percentage of total land area		(13.65; 10.91)

REFERENCES

- ACEMOGLU, D. (2002), "Directed Technical Change", *Review of Economic Studies*, 69(4), pp. 781-809.
- ACEMOGLU, D. (1998), "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality", *Quarterly Journal of Economics*, 113(4), pp. 1055-89.
- ANDERSON, T. W. and C. HSIAO (1981), "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, 76, pp. 589-606.
- ARELLANO, M. and S. BOND (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, pp. 277-297.
- AUTOR, D.H., L.F. KATZ and A.B. KRUEGER (1998), "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, 1998, 113(4), pp. 1169-1213.
- BALASSA, B. (1965), "Trade Liberalisation and 'Revealed' Comparative Advantage", *The Manchester School of Economic and Social Studies* 33(2), pp. 99-123.
- BARRO, ROBERT and JONG-WHA LEE (2010), "A New Data Set of Educational Attainment in the World, 1950-2010", *National Bureau of Economic Research Working Paper series*, no. 15902.
- BLYDE, J.S., C. DAUDE and E. FERNÁNDEZ-ARIAS (2010), "Output Collapses and Productivity Destruction", *Review of World Economics* 146(2), pp. 359-387.
- CHINN, MENZIE D. and HIRO ITO (2008), "A New Measure of Financial Openness", *Journal of Comparative Policy Analysis*, Volume 10, Issue 3, pp. 309-322, September.
- CICCONE, A. and M. JAROCINSKI (2010), "Determinants of Economic Growth: Will Data Tell?", *American Economic Journal: Macroeconomics*, 2, pp. 223-247.
- COLLIER, P. and A. HOEFFLER (2005), "Democracy and resource rents", *Working Paper*, Department of Economics, University of Oxford.
- CORDEN, W.M. and J.P. NEARY (1982), "Booming Sector and De-industrialization in a Small Open Economy", *Economic Journal*, 92(368), pp. 825-848.
- DAUDE, C. (2012), "Development Accounting: Lessons for Latin America", *OECD Development Centre Working Papers*, No. 313, OECD Publishing. doi: 10.1787/5k97f6ws6llp-en
- EASTERLY, W. and A. KRAAY (2000), "Small States, Small Problems? Income, Growth, and Volatility in Small States", *World Development*, 28(11), pp. 2013-2027.
- FERNÁNDEZ, C., E. LEY and M. STEEL (2001), "Model Uncertainty in Cross-Country Growth Regressions", *Journal of Applied Econometrics*, 16, pp. 563-576.
- GLÜZMAN, P.A., E. LEVY-YEYATI and F. STURZENEGGER (2012), "Exchange rate undervaluation and economic growth: Díaz Alejandro (1965) revisited," *Economics Letters*, 117, pp. 666-672.
- HAUSSMANN, R. and C. HIDALGO (2011), "The Network Structure of Economic Output", *Journal of Economic Growth*, 16, pp. 309-342.

- HAUSSMANN, R. and B. KLINGER (2006), "Structural Transformation and Patterns of Comparative Advantage in the Product Space", *CID Working Paper*, No. 128, Harvard University.
- HAUSMANN, R., J. HWANG and D. RODRIK (2007), "What You Export Matters", *Journal of Economic Growth*, 12, pp. 1-25.
- HAUSSMANN, R., C. HIDALGO, S. BUSTOS, M. COSCIA, S. CHUNG, J. JIMENEZ, A. SIMOES and M. YILDRIM (2011), "The Atlas of Economic Complexity: Mapping Paths to Prosperity" Center for International Development, Harvard University
- HESTON, A., R. SUMMERS and B. ATEN (2012), "Penn World Table Version 7.1", Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- HIDALGO, CA. and R. HAUSMANN, "The Building Blocks of Economic Complexity", *Proceedings of the National Academy of Sciences*, 106(26) (2009), pp. 10570-10575.
- JANKOWSKA, A., A. NAGENGAST and J.R. PEREA (2012), "The Product Space and the Middle-Income Trap: Comparing Asian and Latin American Experiences," *OECD Development Centre Working Paper* 311.
- JUDSON, R.A. and A.L. OWEN (1999), "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists", *Economic Letters*, 65(1), 9-15.
- KALDOR, N. (1963), "Capital Accumulation and Economic Growth." In: Lutz, F. A. and D. C. Hague (eds.), *Proceedings of a Conference Held by the International Economics Association*, Macmillan, London.
- KAMINSKI, G.L. and C.M. REINHART (1999), "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems", *American Economic Review*, 89(3), pp. 473-500.
- KIVIET, J.F. (1999), "Expectation of Expansions for Estimators in a Dynamic Panel Data Model; some Results for Weakly Exogenous Regressors", in: C. Hsiao, K. Lahiri, L.-F. Lee and M.H. Pesaran (Eds.), *Analysis of Panel Data and Limited Dependent Variables*, Cambridge University Press, Cambridge.
- KIVIET, J.F. (1995), "On Bias, Inconsistency, and Efficiency of various Estimators in Dynamic Panel Data Models" *Journal of Econometrics*, 68, pp. 53-78.
- KOOP, D. (2003), "Bayesian Econometrics", John Wiley & Sons, New York, NY.
- KUZNETS, S. (1955), "Economic growth and income inequality", *American Economic Review*, 45 (1), pp. 1-28.
- LALL, S. (2000). "The technological structure and performance of developing country manufactured exports, 1995-1998". *Oxford Development Studies*, 28(3), pp. 337-369.
- LEAMER, E. (1978), "Specification Searches", John Wiley & Sons, New York, NY.
- LEDERMAN, D. and MALONEY, W. (2012), "Does What You Export Matter? In Search of Empirical Guidance for Industrial Policies", *Latin American Development Series*, World Bank, Washington, DC.
- LEY, E. and M.F. STEEL (2009), "On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression", *Journal of Applied Econometrics*, 24 (2009), pp. 651-674.
- MADIGAN, D., J. YORK and D. ALLARD (1995), "Bayesian Graphical Models for Discrete Data", *International Statistical Review*, 63 (1995), pp. 215-232.
- MORAL-BENITO, E. (2012), "Determinants of Economic Growth: A Bayesian Panel Data Approach", *The Review of Economics and Statistics*, 94(2) (2012), pp. 566-579.
- NERLOVE, M. (1971), "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross-Sections", *Econometrica: Journal of the Econometric Society*, pp. 359-382.
- NICKELL, S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica: Journal of the Econometric Society*, pp. 1417-1426.
- OECD (2013), *Interconnected Economies: Benefiting from Global Value Chains*, OECD Publishing, Paris, doi: 10.1787/9789264189560-en

- OURENS, G. (2013), "Can the Method of Reflections Help Predict Future Growth", Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).
- PREBISCH, R. (1950), "The economic development of Latin America and its principal problems", *Economic Bulletin for Latin America*, 7, pp. 1-12.
- RAFTERY, A.E. (1995), "Bayesian Model Selection in Social Research", *Sociological Methodology*, 25, 111-163.
- RODRIG, D. (2008), "The Real Exchange Rate and Economic Growth: Theory and Evidence", *Brookings Papers on Economic Activity*, Fall, pp. 365-412.
- ROSS, M.L. (2001), "Does oil hinder democracy?" *World Politics* 53: pp. 325-361.
- SACHS, J. and A. WARNER (1995), "Natural resource abundance and economic growth". *NBER Working Paper* No. 5398, December.
- SALA-I-MARTIN, X., G. DOPPELHOFER and R.I. MILLER (2004), "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", *American Economic Review*, 9, pp. 813-835.
- SCHRÖDER, M. (2013), "Should countries undervalue their currencies?" *Journal of Development Economics*, 105, pp. 140-151.
- SPATAFORA, N. and I. TYTELL (2009), "Commodity Terms of Trade: The History of Booms and Busts", *IMF Working Papers*, 09/205, International Monetary Fund, Washington, DC
- UNESCAP (2011), "Economic and social survey of Asia and the Pacific 2011", United Nations Publication.
- WEF (2012), "The Shifting Geography of Global Value Chains: Implications for Developing Countries and Trade Policy", World Economic Forum.

OTHER TITLES IN THE SERIES/ AUTRES TITRES DANS LA SÉRIE

The former series known as “Technical Papers” and “Webdocs” merged in November 2003 into “Development Centre Working Papers”. In the new series, former Webdocs 1-17 follow former Technical Papers 1-212 as Working Papers 213-229.

All these documents may be downloaded from:

www.oecd-ilibrary.org/development/oecd-development-centre-working-papers_18151949

or obtained via e-mail (dev.contact@oecd.org).

Recent working papers:

Working Paper No. 320, *Capital Flows in Asia-Pacific: Controls, Bonanzas and Sudden Stops*, by Margit Molnar, Yusuke Tateno and Amornrut Supornsinchai, September 2013.

Working Paper No. 319, *The rationale for higher education investment in Ibero-America*, by José Joaquín Brunner, August 2013.

Working Paper No. 318, *How redistributive is fiscal policy in Latin America: The case of Chile and Mexico*, by Barbara Castelletti, July 2013.

Working Paper No. 317, *Opening the Black Box of Contract Renegotiations: An Analysis of Road Concessions in Chile, Colombia and Peru*, by Eduardo Bitran, Sebastián Nieto-Parra and Juan Sebastián Robledo, April 2013.

Working Paper No. 316, *The Politics of Transport Infrastructure Policies in Colombia?*, by Sebastián Nieto-Parra, Mauricio Olivera and Anamaría Tibocho, April 2013.

Working Paper No. 315, *What Drives Tax Morale?*, by Christian Daude, Hamlet Gutiérrez and Ángel Melguizo, November 2012.

Working Paper No. 314, *On the Relevance of Relative Poverty for Developing Countries?*, by Christopher Garroway and Juan R. de Laiglesia, September 2012.

Working Paper No. 313, *Development accounting: lessons for Latin America*, by Christian Daude, July 2012.

Working Paper No. 312, *South-South migration in West Africa: Addressing the challenge of immigrant integration*, by Jason Gagnon and David Khoudour-Castéras, April 2012.

Working Paper No. 311, *The Product Space and the Middle-Income Trap: Comparing Asian and Latin American Experiences*, by Anna Jankowska, Arne Nagengast and José Ramón Perea, April 2012.