

**How costly is rent-seeking to diversification :
an empirical approach**

Felipe STAROSTA DE WALDEMAR

2010.08



How costly is rent-seeking to diversification: an empirical approach *

Felipe STAROSTA DE WALDEMAR[†]

December 2009

Abstract

The empirical U-shaped pattern between product diversification and economic development has been widely examined, but here we analyze the determinants of diversification. We find that a high level of rent-seeking activities has a large impact on the diversification of nations: in countries where rent-seeking is a widespread practice, the number of products being exported will be smaller and its value more concentrated in certain goods. Our analysis embraces a large sample of more than 130 countries between 1995 and 2007, using a highly disaggregated export database comprising more than 5000 products. To establish this relationship we use a Generalized Method of Moments estimation, controlling for endogeneity originated from reverse causality. These empirical predictions contribute to the idea that resources allocated to harm diversification are an important binding constraint for developing countries.

J.E.L.: O11, F1, O31

Key-words: Product diversification, international trade, economic development

* **Acknowledgments:** I am grateful to all members of the Development and Globalization research group for their useful discussions. I would also like to thank Jean-Claude Berthélemy, Sandra Poncet, Patrick Sevestre, Akiko Suwa-Eisenmann and Jean Imbs for their helpful comments, as well as the participants of the Doctoral Days at University of Paris 1 in November 2009. All errors remain mine.

[†]PhD student - Centre d'Economie de la Sorbonne - UMR 8174 - Université Paris 1 Panthéon Sorbonne. felipe.starosta-de-waldemar@univ-paris1.fr

“From South Korea to Singapore, history shows that countries thrive when they invest in their people and in their infrastructure. . . when they promote multiple export industries, develop a skilled work force and create space for small and medium-sized businesses that create jobs.” Barack Obama’s speech on 07/11/2009, in Ghana.

1 Introduction

The development agenda is still characterized by the idea that the presence of many exporting sectors is very important to reach economic prosperity¹. In that sense, in the past few years, a growing literature showing a pattern between diversification and economic development has emerged². However, the determinants of diversification were not deeply explored. Parteka and Tamberi (2008) searched for possible determinants of export diversification, such as geography, size and institutions. But these authors did not solve for endogeneity rising from reverse causality from specialization to institutions, and so proper statistical inference was not properly done. This article tries to fill this gap and looks on the link between the level of rent-seeking in a country and its pattern of diversification. We use a highly disaggregated export database covering thirteen years of international trade and more than 130 countries, between 1995 and 2007. So when we focus on export diversification and look at rent-seeking activities as a possible determinant, we turn to possible explanations found in the development economics literature³.

In this direction, we analyze product diversification, meaning a wider range of goods being produced and exported by a country through time. As we look into the literature, we find many mechanisms of how the amount of rent-seeking activities in a country impacts its productive structure, making it less diversified. Public rent-seeking affects innovation, and specially inside-the-frontier innovations, which is defined as new products for the technological frontier of a given country. Rent-seeking can also influence the process of diversification in a self-discovery frame-

¹Justin Lin (Lin, 2009), chief economist of the World Bank, recently underlined the relevance of the productive structure and an active industrial policy to promote economic development

²Imbs and Wacziarg (2003), Klinger and Lederman (2006) and Cadot, Carrere, and Strauss-Kahn (2007) documented this relationship, which has a U-shaped format: at low levels of income per capita, countries have concentrated productive structures, but as they move along the development path, their production (exports) become much more diversified. That process goes up to a level of development, in which the forces of specialization become more powerful and the now advanced countries specialize in some sectors, resulting in a more concentrated productive structure.

³Even though we acknowledge the importance of the international trade literature (as the results found in Amurgo-Pacheco and Pierola (2008)) for this matter, the use of traditional gravity-like variables is not in the scope of this paper.

work, as in Hausmann and Rodrik (2003), where uncertainty costs on the individual product level diminishes the level of investment done by entrepreneurs. This last article shares the same views as in Hausmann, Hwang, and Rodrik (2007) and Hausmann and Klinger (2006), where the diversification of a country's productive structure towards high earning goods can boost exports and economic growth.

More recently, Cuberes and Jerzmanowski (2009), even though focusing on growth reversals, showed empirically and theoretically how democracy, through a higher level of barriers to entry, impacts diversification. Their model and findings are related to the work of Acemoglu (2008), where in "oligarchic" societies there are bigger barriers to entry to new entrepreneurs, as a result of the concentration of political power with major producers.

Regarding rent-seeking, Murphy, Shleifer, and Vishny (1993) defines it as a redistributive activity that consumes resources. The rent-seeking activities can take the form of corruption, protection costs or attempts to obtain political power and restraint political rights from opponents (Svensson, 2000), amongst other possibilities. Following Murphy, Shleifer, and Vishny (1993), we focus on the detrimental effect of public rent-seeking on innovative activities. By the fact that innovations are risky, an entrepreneur is discouraged to innovate if the returns of its successful project are expropriated and the burden of a failure falls on his shoulders (uncertainty costs à la Hausmann and Rodrik (2003)). Rent-seeking also impacts innovation if the licenses needed to implement new activities are expensive and the entrepreneurs are opponents of the established producers, who will lobby in order to prevent the creation of new goods and services. Credit constraints faced by innovators and the long-term maturation of new goods are the other factors that makes rent-seeking so costly to diversification. Combining all these factors together, the probability of the introduction of new goods is severely diminished. By consequence, economic diversification is hampered.

The quest for clear and solid answers on the diversification process goes back a long way in the economic literature. Concerns about diversification and specialization and its impact on economic growth and development have been an important matter subject since Smith and Ricardo. We can also find distinguished remarks in the works of development pioneers such as Rosenstein-Rodan and

Lewis, and in policies derived from the Prebisch-Singer hypothesis for Latin American economic development. More recently, Lederman and Maloney (2008) pointed that export concentration hampers economic growth. In a similar way, Hesse (2008), in a dynamic panel, shows that export concentration is detrimental to economic growth. Al-Marhubi (2000) and Agosin (2007) found the same result in a cross-sectional regression. Hesse (2008) does also a good review on the empirical works on the subject.

We find that rent-seeking has a negative impact on product diversification. At first, we explore the theoretical link on how rent-seeking affects diversification. Empirically, our results show that the level of rent-seeking has a positive impact on an index of specialization (the opposite of diversification). Most importantly, we control for the problem of endogeneity using Generalized Method of Moments (GMM), in order to capture the relationship going from rent seeking to diversification. To give support to our estimations, we also use external instruments in the GMM framework. This is proven to be robust to different specifications, using a different index for diversification, changing the proxy for rent-seeking and controlling for the institutional environment and geographical conditions.

The goal of this paper is to estimate the impact of rent-seeking on diversification in a panel data setting, which it comes up to be quite important. To reduce the Theil index of one point, our estimations show a need for a 1.39 reduction in the corruption index. Even if this seems a big evolution, it is much smaller than the jump in the income level needed for a similar change in the productive structure.

This paper is organized as follows: in section 2 we present our theoretical framework, based on how rent-seeking impacts diversification; Section 3 introduces our database on international trade and the measures of diversification and rent seeking. Section 4 contains our empirical methodology, with our results using GMM, including robustness checks; Section 5 wraps up our story with concluding remarks.

2 Rent-seeking and diversification

Several arguments (Hesse, 2008; Berthelemy, 2005) have been used to support the vital role of diversification for economic development and growth.⁴ Differently, we underline the mechanisms in which rent-seeking impacts diversification and then we test it empirically. Before doing so, we summarize the finding on previous works on diversification.

To our knowledge, only a few papers have looked into the determinants of diversification. In a recent paper, Malik and Temple (2009) briefly study the determinants of export concentration. Through Bayesian methods, they show that market access, climate variability, geography and institutional quality have influences on export concentration. This results will be also of important guidelines in our empirical approach. As we explore the micro foundations of product diversification, the framework was consolidated by Hausmann and Rodrik (2003), where economic development is originated from the discovery of what a country is a good producer. The authors show that in the existence of uncertainty costs to entrepreneurs, diversification is restrained by imitation and other failures generated by a *laissez-faire* equilibrium, as entrepreneurs are not able to appropriate the full amount of their investments. If we link product diversification to economic growth, we find the conclusions made by Hausmann, Hwang, and Rodrik (2007), where the components of the export basket are good predictors of economic growth. To these authors, structural transformation, i.e., the change in developing countries to a export basket composed by high yield products, have an important impact on growth.

Following Hausman and Rodrik's story, Klinger and Lederman (2006) found that a good part of the explanation for the lack of diversification is related to the presence of market failures. Market failures disfavor the positive incentives entrepreneurs are faced with when they decide to invest in new activities that can be possibly "discovered" by the country. These discoveries are not necessarily innovations to the global technological frontiers, such as patents, but in its most part they are new products to the countries' export basket.

⁴In this line, we emphasize the role of a reallocation of resources in higher level activities (therefore boosting productivity) in learning-by-doing in the manufacturing sector (Matsuyama, 1992), reducing volatility (Ramey and Ramey, 1995; Imbs, 2007; Koren and Tenreyro, 2007; Stanley and Bunnag, 2001), and diminishing risks to economic agents through the financial system (Acemoglu and Zilibotti, 1997).

While analyzing why developing countries have larger growth reversals, Cuberes and Jerzmanowski (2009) found that less democratic countries, with larger barriers to entry for new firms, have lower sectoral diversification. They developed a model where only one sector pays off at each time so less democratic countries, with higher barriers to entry, have a more concentrated productive structure in the sectoral level. This concentration leads to a higher probability of growth reversals. The authors empirical tests, explaining diversification by democracy, income and its squared term show that our empirical methodology is proven to be correct. But our theoretical framework is different, as we look at the product level and we explore how rent-seeking impacts diversification directly, but the contribution of Cuberes and Jerzmanowski (2009) is very important in understanding the link between diversification and the global environment of a country.

Switching to our framework, Murphy, Shleifer, and Vishny (1993) developed the idea that rent-seeking has an impact on growth, especially in innovative activities. More importantly, we can relate these to the empirical works of Mauro (1995) and Svensson (2000)⁵. It is well established in this literature that rent-seeking activities have a negative impact on growth, and this leads us to think that rent-seeking has an effect on diversification. A larger presence of rent-seeking harms the diversification of production and exports, meaning a smaller variety of goods being produced. We follow Murphy, Shleifer, and Vishny (1993) and we enumerate a series of reasons why the introduction of new goods are significantly affected by rent-seeking. These effects are the main channels why rent-seeking impacts diversification (new products or innovations), so we use it as the core of our framework.

The first one is political, as the entrepreneurs that introduce innovations are not part of the government and do not have established lobbies. They are more exposed to heavy bribes and expropriation, especially when the interest of the establishment and the innovators are opposed. This is one of the main reasons that the productive structure will be hard to diversify, as there are limited options of investments and the established economic and political elite will do the necessary efforts to block this process. This mechanism is the same one as developed by Acemoglu (2008). Secondly, Murphy, Shleifer, and Vishny (1993) observe that the lack of credit is a binding

⁵Svensson (2000) focus on the relationship between aid and rent-seeking, showing that in countries that have larger social competing groups, there exists an association between aid and corruption.

constraint, as these entrepreneurs do not have the possibility of assuming additional costs due to bribery and other forms of illegal payments. Rent-seeking imposes another barrier to the launch of new products as this extra fixed cost has to be overcome, in the same way as Cuberes and Jerzmanowski (2009) introduce a fixed cost of entry for a new firm generated by barriers to entry.

Thirdly, innovations are long term projects, as it takes more time and steps to create the new product and to reap its profits, so the possibility of the extraction of rent-seeking is greater in this kind of process. Hence, a higher level of rent-seeking will impact more the introduction of new products. The last reason why new procedures are more susceptible to rent-seeking deserves a careful treatment, as we analyze the importance of risk in the diversification in two different ways. In one way, the process of introducing innovations has a higher level of embedded risk, as entrepreneurs are not sure if the project will succeed. If it is the case, there is still the possibility of *ex post* rent-seeking, meaning that a good part of the specific returns of this activity will be captured from the entrepreneur. If the innovation is not successful, the entrepreneur will have to face all the costs of the failure. Therefore, with a high level of rent-seeking, entrepreneurs are less prone to invest. If entrepreneurs do not know if an important part of their profits will be reaped by any kind of rent-seeking practice undermines new investments. This could be related to the Klinger and Lederman (2006) idea where new discoveries are not realized because entrepreneurs cannot appropriate the total of their benefits due to market failures.

Differently, Acemoglu and Zilibotti (1997) studied the importance of risk in economic development, where the latter is accompanied by market expansion and different diversification possibilities. According to the authors, diversification allows the reallocation of resources towards more productive uses and diminishes growth variability. In their model, agents dislike risks and they try to diversify the economies correlated projects in order to spread out this risk. To Acemoglu and Zilibotti (1997), “the more sectors that are open, the higher the proportion of their savings that agents are willing to put in risky investments”. In this framework, agents will not be encouraged to innovate if the country’s productive structure is too concentrated, because of the risky nature of the innovation process. The economy would then be stagnated on few activities of low level of productivity, where the level of idiosyncratic risks is not high but its resources are easily appropriated by a political powerful elite. The process of diversification is then affected.

When it comes to introducing innovation and the role of entrepreneurial activity, Baumol (1990) analyzed the interaction between the systems of payoffs of a society and the allocation of its entrepreneurial activity. Entrepreneurial activity can either have incentives to innovate or to focus on unproductive activities. If the payoff favours rent-seeking activities, it is hard to believe that entrepreneurs will innovate, and this effect will reinforce itself, because of the increasing returns in rent-seeking activities (Murphy, Shleifer, and Vishny, 1993). In their work on rent-seeking and resource booms, Baland and Francois (2000) also integrated the idea that rent-seekers prey on entrepreneurs, diminishing the introduction of new products and services.

All the above literature points out the evidence of a noxious effect of rent-seeking in the process of introducing new goods, through the channels proposed by Murphy, Shleifer, and Vishny (1993). The impact of rent-seeking on diversification seems clear and we test it empirically in section 4, after presenting the measures of the main variables and the data sources in the next section.

3 Data and measures

3.1 Measures of diversification

It is important to clarify that our framework is based on product diversification⁶ - and not sectoral diversification. Even though they are closely related subjects, the level of aggregation in the datasets, measuring goods or sectors, has a huge impact on analysis. This is confirmed by the work of Benedictis, Gallegati, and Tambari (2009) and Parteka and Tambari (2008). The approach on the link between diversification and development based on a higher level of aggregation was analysed by Kalemli-Ozcan, Sorensen, and Yosha (2003), Koren and Tenreyro (2007) and Cuberes and Jerzmanowski (2009). Differently, we explore the relationship between rent-seeking and diversification on a product basis.

We define product diversification as the widening of the range of goods in a country's export basket and a more equal distribution of the values between the products in a country exports.

⁶We recognize that geographical diversification has an important impact on income (Jansen, Lennon, and Piemartini, 2009) and output (Malik and Temple, 2009) through the reduction of volatility, but in this paper we do not explore this dimension of diversification

Hence, diversification has two important dimensions: the first one is when a country produces and exports goods that it didn't do it before, as is the case of innovations and discoveries (Klinger and Lederman, 2006). Next, diversification can originate from a more even distribution of total value of exports within a fixed number of goods.

Both dimensions are of important matter. The case of discoveries or innovations fits correctly our theoretical hypothesis and empirical testing, as we discussed how rent-seeking can influence the introduction of new goods, as in Murphy, Shleifer, and Vishny (1993). Nonetheless, we think that these same channels influence a more even distribution export (or production) basket. We make the assumption that a low value of exports in a high number of goods combined with big export values that are concentrated in a few products can be originated by the same problems to the introduction of new goods: large scale investments are not made due to the inexistence of the correct system of incentives and a high level of rent-seeking. We conclude that a more even distribution within goods means that most of these failures have been surpassed.

In order to measure export diversification, we alternate between two widely used indexes. The Theil and the Herfindahl specialization indexes are absolute measures as they compare product distribution with a uniform distribution (Benedictis, Gallegati, and Tamberi, 2009). To measure export diversification, we follow Cadot, Carrere, and Strauss-Kahn (2007) and we construct the Theil index as:

$$T_{i,t} = \frac{1}{n} \sum_{g=1}^n \frac{X_{g,i,t}}{B_{i,t}} \ln \left(\frac{X_{g,i,t}}{B_{i,t}} \right) \quad \text{where} \quad B_{i,t} = \frac{\sum_{g=1}^n X_{g,i,t}}{n_{i,t}} \quad (1)$$

Where $X_{g,i,t}$ is the export value of each individual good in each country and year and n is the number of products exported for each country-year pair⁷. Differently, a normalized Herfindahl

⁷We follow Cadot, Carrere, and Strauss-Kahn (2007) and we normalize the number of products for the 5017 products existing in the Harmonized System as in the Baci database. This is done in order to compare the index between all countries in our sample.

index is built as:

$$H_{i,t} = \frac{\sum_g (Q_{g,i,t})^2 - \frac{1}{n}}{1 - \frac{1}{n}} \quad \text{where} \quad Q_{g,i,t} = \frac{X_{g,i,t}}{\sum X_{g,i,t}} \quad (2)$$

Where $Q_{g,i,t}$ is the export value in a individual product divided by total exports.

3.2 Data source

To measure rent-seeking, we use the index Control of Corruption, compiled by the Governance Indicators in the World Bank (Kaufmann, Kraay, and Mastruzzi, 2009). The variable “measures the extent to which public power is exercised for private gain, including petty and grand forms of corruption, as well as “capture” of the state by elites and private interests”, which is one definition of rent-seeking. Even though rent-seeking could take other forms, this index is a very good approximate, as in Svensson (2000) and Mohtadi and Roe (2003)⁸. The use of this kind index is proven to be a standard procedure in the economic literature.

To use this variable in a simple way, we followed Croix and Delavallade (2009) and we transformed this variable to have values between 0 and 5, in which a lower value (close to 0) means that the country has a low level of corruption, and a high value (near 5) means that the country has a high level of corruption. The data shows that low-income countries are the more corrupt ones, and they also have a highly concentrated export basket.

Rent-seeking can be also captured by an alternative variable, the Political Rights from the Freedom House Index. In this case, we follow Croix and Delavallade (2009) who explain that “few political rights for the population indicate a strong concentration of power in the hands of a few. And those who hold the power are presumably rent-seekers”. This definition is quite straight and follows the idea of Acemoglu (2008), and it also proves to be a good approximation of rent-seeking. This variable measures the lack of political right and ranges between 1 and 7, where a low value means broad political rights for the population and high levels represents the lack of political rights for its citizens.

⁸Although they use the Corruption Index from ICRG, the definitions of both variables are quite identical

To calculate both indexes on trade data, we used the BACI database (Gaulier and Zignago, 2008), compiled by the CEPII (Centre d'Etudes Prospectives et d'Informations Internationales) using the data reported by the United Nations COMTRADE. The BACI database was built to provide international trade analysis at the product-level. Covering 13 years (1995-2007) and almost 200 countries, it is disaggregated using the Harmonized System (HS) at the 6-digit level, distinguishing more than 5000 products. This high level of desegregation is the thinnest one available for an international comparison in a panel setting, and merging with the data on rent-seeking and GDP, we are able to keep more than 130 countries of all levels of economic development⁹. The list of countries used in our estimations and descriptive statistics for each group can be found in the appendix.

We see that countries have different productive structures, depending on their development level. Low income countries (following the World Bank classification) have highly concentrated export baskets¹⁰. They also are the ones with a high level of rent-seeking. In the opposite way, high income countries have diversified export baskets and a low level of rent-seeking. Descriptive statistics show that these variables are correlated, and our theoretical mechanisms and empirical treatment will reflect the correct impact of rent-seeking on diversification.

So what are some of the factors that determine this difference in the productive structure? As it described by the literature and seen in our results, a higher level of economic development is associated with greater diversification. Therefore, we look at those factors that impact economic growth and are the ones that could mostly affect the process of diversification in a direct way.

⁹The HS 6-digit level of classification is an important evolution from the usual 4-digit Standard International Trade Classification (SITC) covering 1200 products. As a matter of fact, some studies done on diversification uses the Feenstra, Lipsey, Deng, Ma, and Mo (2005) database, which is built on the revision 2 of the SITC classification, in a 4 digit format, covering around 1000 products. One of the important contributions of our work is indeed the fact that using a more disaggregated database we will provide a more robust study on diversification. The HS classification can provide us a better picture of the country's productive structure in the product level. Largely used when studying sector shares in the economy, the UNIDO (United Nations Industrial Development Organization) database has only 28 sectors so it doesn't allow the same kind of study on the product level.

¹⁰Benin, for example, has a Theil index that on average stays around 6 and in 1995 a single export product line - cotton - accounted for 70% of total exports value. Differently, a high income country as France has a Theil index of around 2, and no product exceeds more than 4% of the total value of exports. Also, while Benin has an average of 830 export lines, France has a greater number of products, reaching more than 5000 product export lines.

4 Econometric methodology and results

4.1 Methodology: baseline equation and preliminary results

As it can be seen at Figure 1, for a sample of 131 countries in 2007, a scatter plot of our data show a negative relation between product concentration and level of development, respectively measured by the Theil index and GDP per capita. Also in the Figure 1 below, a simple pooled OLS regression (equation (3), following Imbs and Wacziarg (2003)) of export diversification on per capita revenue reveals the same pattern found in previous studies: low-income countries have highly concentrated structures, but it diversifies as long their income per capita becomes higher. After a level of revenue that characterizes high-income countries, specialization restarts and countries concentrate their exports.

$$DIV_{i,t} = \alpha + \gamma_t + \beta_1 GDP_{i,t} + \beta_2 GDP_{i,t}^2 + \epsilon_{i,t} \quad (3)$$

Where DIV is the index of diversification (a high value means *less* diversification and a low value means *more* diversification), the right-hand-side variables are the GDP per capita and this same term in its quadratic form. ϵ is the error term and α is the constant. We add time dummies (γ) to control for possible year effects in trade patterns through time¹¹.

We continue to follow the empirical framework of the articles studying the relationship between diversification and development. Our point of departure is that the parametric evidence found in Imbs and Wacziarg (2003) and Cadot, Carrere, and Strauss-Kahn (2007) (and followed by Cuberes and Jerzmanowski (2009)) capture this relation in a quadratic form, and it has significant similarities to the non-parametric estimations found in these same articles¹². Our model of departure, including rent-seeking, is then:

$$DIV_{i,t} = \alpha + \gamma_t + \beta_1 RT_{i,t} + \beta_2 GDP_{i,t} + \beta_3 GDP_{i,t}^2 + \beta_4 POP_{i,t} + \epsilon_{i,t} \quad (4)$$

¹¹The results of the estimation of equation (3) are not shown, but can be obtained with the author.

¹²We did a test proposed by Lind and Mehlum (2007) to verify the existence of the U-shaped relationship, which we find it to be the correctly present.

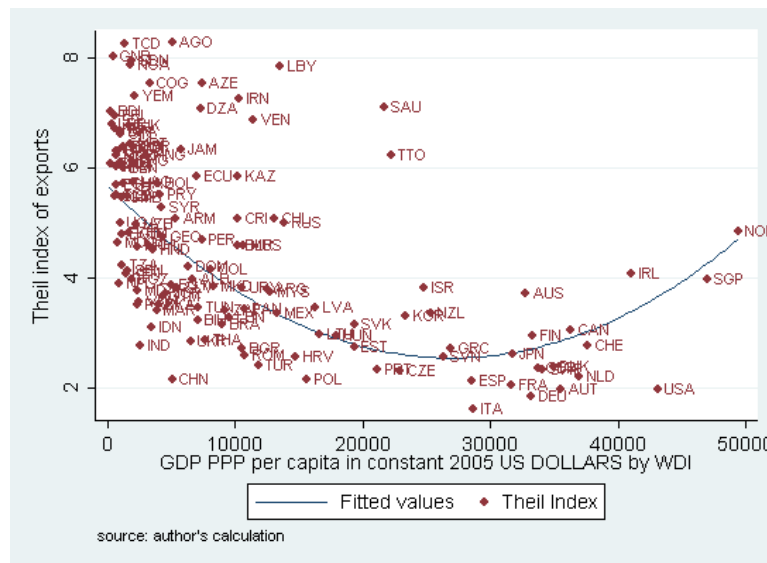


Figure 1: Diversification and development

Where RT is the proxy of rent-seeking. We also add the population variable (POP) to control for the size of the economy, as it is suggested in Berthelemy (2005), where a larger number of producers and a bigger internal market would lead to a higher level of diversification. But at this stage we will not include other possible determinants of diversification, like geography or institutional environment, with the objective of not entering in the same procedure (a horse race) on the cross country determinants of growth, which found 145 possible determinants but with small robustness.

Our empirical methodology, as will see in the next section, is to use System-GMM. With this methodology we can make empirical corrections and specially for endogeneity. Reverse causality is the main reason we do not present our results in Pooled OLS. But preliminary results using Pooled OLS show a stable and robust pattern: the coefficient of the rent-seeking variable is positive and strongly significant and it is true for all four pairs of combinations between the index of diversification and the proxy for rent-seeking. So a higher level of corruption is correlated with a higher level of specialization measured by the Theil or the Herfindahl index. These preliminary results reflects the correlation between rent-seeking and diversification, meaning that countries with higher levels of rent-seeking are more specialized countries¹³. In turn, if we use Fixed Effects

¹³For the coefficients of GDP and GDP^2 , they have the expected signs and they are often significant. The

estimations, these results are the same in its most part¹⁴.

4.2 Generalized Method of Moments

Results using Pooled OLS and Fixed Effects show a positive correlation between rent-seeking and specialization. Countries that have higher levels of rent-seeking activities have also more concentrated productive structures, through the reasons we explained in section 2. But both methods used to estimate this relationship do not account for a major problem in econometrics that we find in our model: endogeneity. Endogeneity arises when one or more independent variables are correlated with the error term. In this sub-section, we try to solve for the endogeneity issue that arises both from reverse causality (simultaneity bias) and omitted variable bias¹⁵ in order to estimate the impact of rent-seeking on diversification correctly.

In our model, the simultaneity bias is originated from reverse causality. When we try to estimate the impact of rent-seeking on diversification, it is important to distinguish in which sense the causal direction works, and we need special methods to correct for potential problems. In our case, there could be a reinforcement process, going from diversification to the level of rent-seeking. More specialized productive structures facilitate the capture of resources by a powerful elite, creating more rent-seeking. This is the case for the capture of natural resources and the degradation of the institutional quality (Couttenier, 2008; Isham, Woolcock, Pritchett, and Busby, 2005; Ades and Tella, 1999; Anthonsen, Lofgren, and Nilsson, 2009). Where the revenues of a natural resource are quite important, rent-seeking activities are also more present, like in the case of oil exporting countries or other nations exporting valuable primary commodities. This could exist in the form of capture of the revenues from a few big companies owned by a elite interested in maintaining the productive and social structure. It could also happen as a lower level of checks and balances originated by revenues not coming from taxation on productive activities and also

coefficient of the population variable is negative and significant, meaning that a larger population facilitates diversification.

¹⁴The presence of individual-specific effects was verified through a Breusch-Pagan test. In our case, we allow the individual specific effect to be correlated with the regressors, as a Hausman test suggested when it was made on our regression.

¹⁵In the Pooled OLS, the estimations could not be consistent due to reverse causality and omitted variables problems. When using Fixed-Effects, it is especially the reverse causality issue that generates inconsistent estimates, because this model of panel data could be used to obtain consistent estimators under certain assumptions.

by the possibility of choosing investments (for the infra-structure to extract the natural resource, for exemple) where rents are easily captured. So, in Pooled OLS and Fixed Effects, the coefficient estimated to the rent-seeking variable is inconsistent and biased, as this term could be correlated with the error term. In this scenario, the coefficient could be upward biased, as we expect that more specialized countries will have a larger amount of rent-seeking activities. The treatment for this problem is fundamental to obtain sound and robust results.

Omitted variables bias often leads to inconsistency of OLS estimates. We think that the omitted variable issue is also quite possible in our model, because we could find possible determinants of diversification that are correlated with rent-seeking (such as institutions or geography), but whose omission of the regression leads to inconsistent estimations, leading to the overestimation of the effect of rent-seeking on diversification.

To solve for the endogeneity problem, we use the Generalized Method of Moments (GMM). System GMM solve for the omitted variables problem by controlling for possible unobserved heterogeneity (time variant and invariant) between rent-seeking and diversification. Most importantly, to solve for the reverse causality issue, we need eligible instruments. In this direction, this method uses lagged regressors as instruments and also estimates by lagging regressors. One of the reasons we use system GMM is that we do not have many suitable external instruments, and so estimations using Two-Stage Least Squares are not a appropriate method of estimation, due to the impossibility of testing overidentifying restrictions. Even though there are common instruments for the corruption index or other institutional variables found in the literature, such as legal origin or religion, they are constant through time, not allowing a fixed effects estimation in a panel data setting.

System GMM has been widely used in the growth literature, as an improvement from the Difference GMM¹⁶. Arellano and Bover (1995) and Blundell and Bond (1998) built a system of equations where we estimate the equation in levels (equation 5) using lagged values of the difference

¹⁶Arellano and Bover (1995) and Blundell and Bond (1998) show that in the difference GMM the lagged levels of the explanatory variables (if they are persistent over time) are weak instruments for the regression in differences, creating a small-sample bias and raising the asymptotic variance of the estimator (Calderon and Serven, 2004).

of regressors as instruments and we use lagged values of the level of regressors as instruments for the regressors in the equation in first differences (equation 6).

$$DIV_{i,t} = \eta_i + \gamma_t + \beta_1 RT_{i,t} + \beta_2 GDP_{i,t} + \beta_3 GDP_{i,t}^2 + \beta_4 POP_{i,t} + \epsilon_{i,t} \quad (5)$$

$$DIV_{i,t} - DIV_{i,t-1} = \beta_1 (RT_{i,t} - RT_{i,t-1}) + \beta_2' (X_{i,t} - X_{i,t-1}) + \epsilon_{i,t} - \epsilon_{i,t-1} \quad (6)$$

Where X contains GDP, GDP² and POP in order to simplify our equation (6). This set of equations can solve for the problem of endogeneity of the rent-seeking variable and to obtain unbiased and consistent estimates of the parameters we assume the following moment conditions:

$$E[RT_{i,t-s} \cdot (\epsilon_{i,t} - \epsilon_{i,t-1})] = 0 \text{ for } s \geq 2; t = 3, \dots, T \quad (7)$$

$$E[X_{i,t-s} \cdot (\epsilon_{i,t} - \epsilon_{i,t-1})] = 0 \text{ for } s \geq 2; t = 3, \dots, T \quad (8)$$

$$E[(RT_{i,t-1} - RT_{i,t-2}) \cdot (\eta_i + \epsilon_{i,t})] = 0 \quad (9)$$

$$E[(X_{i,t-1} - X_{i,t-2}) \cdot (\eta_i + \epsilon_{i,t})] = 0 \quad (10)$$

To be consistent, the system GMM estimators depend on the validity of the moment conditions (7) to (10). We test this using the Sargan/Hansen test and the Arellano-Bond test. The Hansen test (as we use a robust estimation) checks the null hypothesis that the instruments are exogenous. It is a test of the over-identifying restrictions of the model, and the failure to reject the test validates the assumptions of the model, so a higher p-value of Hansen statistic is a positive result. Differently, the Arellano-Bond test looks at the null hypothesis of no autocorrelation in the error term $\epsilon_{i,t}$. As the Hansen test, the failure to reject the null gives support to the model. The test for the first-order serial correlation usually rejects the null hypothesis, but this should not be a problem, as it is expected. But the presence of second-order correlation would endanger the

validity of some lags of the instruments, so we should fail to reject the null.

We use the standard treatment for the lags of the endogenous variable (rent-seeking) as in Roodman (2006)¹⁷. That means we specify 2 lags and longer for the equation in difference and 1 lag for the equation in level. Also, we will use this treatment to the other explanatory variables. We will see that treating the other explanatory variables as predetermined, by using 1 lag and earlier for the differenced equation and lag 0 for the level equation, won't change the results.

Moreover, we follow Roodman (2006) and we use a two-step system GMM estimator to correct for heteroscedasticity and serial correlation. Although the two-step robust estimator is more efficient than the one-step estimator¹⁸, it could result in a significant downward bias of the standard deviation. In order to correct for this, we use the Windmeijer (2005) finite-sample correction to the standard errors, which makes two-step robust estimates more efficient rather than the one-step procedure.

Roodman (2009) also highlights the special care we should give to the instrument count. He criticizes recently published papers showing that their system GMM estimations are flawed due to a high number of instruments. Roodman (2009) establishes as rule of thumb that the number of instruments should be smaller than the number of groups in the panel. We stick with his recommendations and we report the number of instruments used in our regressions, as we “collapse” the blocks in the instruments matrix¹⁹. In accordance with the author, we will comment the results from difference-in-Hansen tests for all subgroups of the instruments used in our estimations, although we will not report them in order to save space.

Furthermore, our estimations have two specificities that slightly differ from the usual application of system GMM found in the literature. At first, we do not use the lagged dependant variable as a explicative variable. We think that the diversification process is a static process, as past realizations of diversification will not influence future changes in diversification. Therefore, diversification is a static notion of a productive structure that modifies itself with time.

¹⁷Roodman (2006) elaborates a complete pedagogical paper on using GMM with Stata, which we follow thoroughly.

¹⁸Cameron and Trivedi (2005) explain the theory on the optimality of the two-step GMM estimator.

¹⁹As Roodman (2009) explains, “this allows us to combine instruments through addition in smaller sets”, and to reduce the number of instruments.

Secondly, we use external instruments in order to validate the relationship between rent seeking and diversification. Although it is not the most common procedure, the system GMM allows the inclusion of external instruments (Roodman, 2006). Our objective with the use of external instruments is to eliminate any further concerns with the endogeneity of the rent-seeking variables. We use the instruments in two ways: we first add them as additional instruments in our estimations and we also remove the rent-seeking variable and replace it only by the external instruments.

As an instrument, we choose the percentage of women holding seats in parliament²⁰. We follow Dollar, Fisman, and Gatti (2001) as they found in a cross country study that higher the representation of women in a government chamber, the lower the level of corruption. As we find in our data, corruption and the percentage of seats held by women in national parliaments are negative correlated. Furthermore, we believe that the participation of women in the national chamber has no other effect on diversification other than the one through corruption.

4.3 Main results

In Table (1), using system GMM, we see that through all specifications the coefficient for the rent-seeking variable is positive and significant. We conclude that there is a positive impact of rent-seeking on specialization. A higher level of rent-seeking, measured by the corruption variable, results into a more concentrated productive structure. This event occurs as a consequence of the constraints faced by entrepreneurs on the introduction of new activities as explained in Murphy, Shleifer, and Vishny (1993).

In regard to each specification, column (1) is our baseline system GMM specification with the characteristics explained in section 4.2. In column (2), we changed the lags of the control variables (GDP, GDP² and POP) and we treat them as predetermined. All the results of the first column remain the same. Column (3) was estimated slightly different, as we dropped all lags of the corrupt variable to provide an extra proof of the exogeneity of the instruments, as we explained in the end of section 4.2. Column (4) follows column (3) but we also change the lags as in column (2), and we find the same results as in the previous tables. In column (5), we change the diversification index

²⁰This data is compiled by the Inter-Parliamentary Union.

from Theil to Herfindahl and we estimate in the same way as column (1). All our main results remain unchanged, even though the control variables are no longer significant albeit having the usual sign.

The coefficient of the corrupt variable is 0.716 in column (1) and of a similar magnitude in columns (2) to (4). A reduction of one point on the corruption index means 0.716 less in the Theil index, that is, a more diversified export basket. To compare with the effect of GDP and GDP² on diversification, we calculate the change needed in these independent variables to a one point change on the Theil index. To decrease the Theil index by one point it is needed a 1.39 fall in the corruption index. This change in the corruption index is a figure that was almost reached by a few countries during our 12 year period of analysis. Differently, for the same one point reduction in the Theil index, it is needed at least a 160% increase in the GDP level of the country (or 8300 US dollars of increase in the GDP per capita), which is a unrealistic growth scenario for the same time period. Furthermore, the increase needed to the GDP variable is much higher for low-income countries, due to the non-linearity of the GDP term.

In respect to the control variables, they are significative and they have the expected sign in the first four columns. GDP and GDP² show the U-shaped relationship between diversification and development, while the negative sign on the population variable reflects the size effect of the population towards a more diversified productive structure. The Hansen test in all specifications confirms the validity of the set of instruments. To be valid, the results from the GMM estimations need to provide exogenous instruments and so it is the case in Table (1), where we can't reject the null of the validity of the set of instruments. We do not report the Difference-in-Hansen tests of exogeneity of instrument subsets for all subgroups of instruments in each specification, but we find that we can't reject the null of the validity of those instruments subsets for all specifications in Table (1). In addition, the Arellano-Bond test (AR-1 and AR-2) show that the model is correctly specified, as we reject the null hypothesis of no autocorrelation of first order in the disturbance term ϵ , but we can't reject autocorrelation of second order. This result is the same for all specifications in Table (1).

All this empirical evidence assert the significant impact of rent-seeking on diversification as

Table 1: Generalized Method of Moments

VARIABLES	(1) Theil	(2) Theil	(3) Theil	(4) Theil	(5) Herfindahl
corrupt	0.716** (0.323)	0.707** (0.330)	0.582** (0.293)	0.607* (0.317)	0.065** (0.031)
gdp	-1.211e-04*** (4.545e-05)	-1.171e-04*** (4.277e-05)	-1.276e-04*** (4.574e-05)	-1.240e-04*** (4.375e-05)	-2.553e-06 (4.213e-06)
gdpsq	3.115e-09*** (9.304e-10)	2.974e-09*** (9.189e-10)	2.964e-09*** (9.760e-10)	2.904e-09*** (9.625e-10)	1.250e-10 (7.917e-11)
population	-2.521e-09*** (4.921e-10)	-2.543e-09*** (5.032e-10)	-2.746e-09*** (4.273e-10)	-2.745e-09*** (4.704e-10)	-6.973e-11 (4.565e-11)
Constant	2.931*** (1.032)	2.943*** (1.041)	3.426*** (0.929)	3.350*** (0.981)	-0.075 (0.101)
Observations	987	987	987	987	987
Number of i	131	131	131	131	131
Number of instruments	56	59	45	48	56
Hansen Test	0.189	0.264	0.345	0.382	0.470
AR(1)	0.0190	0.0168	0.0157	0.0150	0.0291
AR(2)	0.130	0.143	0.177	0.174	0.819

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Estimated coefficients of time dummies are not reported here. Robust standard errors are in parentheses. Hansen test and the Arellano-Bond (AR(1) and AR(2)) tests are P values.

we expected with our theoretical framework. This effect affects nations that are diversifying their productive structure, and the explanations for the second half of the U-shaped curve can be found in Imbs and Wacziarg (2003). In order to see if this is not specific to the indexes used, in Table (2) we change the proxy for rent-seeking and the index used to measure diversification. We see that, as in Table (1), all specifications exhibit a positive and significant impact of the lack of political rights on specialization. Fewer political rights (*more* in our index) means a more concentrated export basket, and we could translate this effect similarly to the economic interpretation above.

Again we find that in all specifications the Hansen test can't reject the null of the validity of our instruments. Also, the difference-in-Hansen tests confirms the exogeneity of the different instruments subsets. The Arellano-Bond tests give more support to the validity of the model, as in every single column of Table (2) we reject the null of no first order autocorrelation and we do not reject the null of second-order autocorrelation.

The specifications in Table (2) follow the same pattern as in Table (1). Column (1) is estimated using the specifications described in section 4.2. In column (2) we treat GDP, GDP² and POP as

predetermined variables. In column (3) we do not use the lagged levels and lagged differences of the political rights variables as instruments, following our external instrument strategy. Column (4) combines both specificities of columns (2) and (3). Finally, column (5) has the same specification of column (1) but using the Herfindahl index at the place of the Theil index.

Table 2: Generalized Method of Moments

VARIABLES	(1) TT	(2) TT	(3) TT	(4) TT	(5) H
prights	0.358*** (0.114)	0.356*** (0.108)	0.490*** (0.161)	0.487*** (0.169)	0.021* (0.012)
gdp	-1.027e-04*** (3.944e-05)	-9.420e-05** (3.754e-05)	-6.715e-05 (4.204e-05)	-5.986e-05 (4.277e-05)	-5.261e-06* (3.090e-06)
gdpsq	1.757e-09* (9.291e-10)	1.528e-09* (9.006e-10)	1.095e-09 (8.776e-10)	8.810e-10 (8.553e-10)	9.593e-11 (6.963e-11)
population	-2.449e-09*** (5.248e-10)	-2.400e-09*** (5.173e-10)	-2.120e-09*** (7.068e-10)	-2.007e-09*** (7.528e-10)	-8.291e-11* (5.039e-11)
Constant	3.659*** (0.533)	3.623*** (0.506)	2.967*** (0.762)	2.956*** (0.791)	0.061 (0.049)
Observations	1336	1336	1336	1336	1336
Number of i	131	131	131	131	131
Number of instruments	60	63	48	51	60
Hansen Test	0.397	0.563	0.700	0.824	0.608
AR(1)	0.0265	0.0252	0.0310	0.0303	0.0226
AR(2)	0.814	0.804	0.890	0.882	0.792

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Estimated coefficients of time dummies are not reported here. Robust standard errors are in parentheses. Hansen test and the Arellano-Bond (AR(1) and AR(2)) tests are P values.

Thus, Table (2) gives further evidence of the impact of rent-seeking on diversification. This effect is captured independently of the variable used to measure rent-seeking and the index of diversification. We also changed the lags of control variables (GDP, GDP² and POP), treating them as predetermined or endogenous, and the results are robust. In the next subsection we will take into account other empirical findings found in the contemporary literature on diversification and development in order to discuss how could our results be sensitive (or not) to the introduction of other determinants of export diversification.

4.4 Robustness checks

Following Parteka and Tamberi (2008), Benedictis, Gallegati, and Tamberi (2009) suggest three groups of variables that are the possible determinants of export diversification: size, geography and institutions. We already control for the size of the population and the economy (POP and GDP). Concerning geography, we normally control for their effect as they are constant through time for each country (fixed effect). Nevertheless, we include geography variables in our baseline system GMM regression that comes from Column (1) of Table (1). The variables used are the distance from major markets or the distance from the nearest navigable river or coastline, from the Gallup, Sachs, and Mellinger (1999) dataset, and we also use a dummy for countries that are landlocked. In the Table (3) in the appendix we see that the coefficient of the corrupt variable remains significant and positive when we add these three different measures of geography, and all usual tests for the system GMM estimation validates the assumptions of our model. The size of the coefficient of the rent-seeking variable is still high and comparable to the estimates found when we do not add the geography variables.

With reference to the impact of institutions on diversification, the literature mentioned above made no clear theory of how could this be originated other than Cuberes and Jerzmanowski (2009). When using the corruption and political rights variables, we are trying to capture the effect of rent-seeking on product diversification, and we explain how this phenomenon works in our review of literature. We use two broad institutional indexes to test if our rent-seeking variables are still significative when controlling for these effect: ICRG, which is a composite index of institutions; and Polity2, a political institutional index. The columns (3) and (4) on Table (3) show that the introduction of these variables do not affect our results.

We have found a negative and significant linear impact of rent-seeking on diversification, which is robust to different specifications. We do not believe that there is a non-linearity in this effect, as Dutt and Traca (2009) found in the relationship between corruption and bilateral trade flows. They were analyzing trade flows and it's own dynamic, whereas we study the diversification of the productive structure. But to test this, we look at the relationship between our corruption variable and our levels of diversification using a non-parametric technique (Kernel-weighted local

polynomial smoothing). We see at Figure (2) on the appendix that non-linearity doesn't seem to exist in our context.

5 Concluding Remarks

The U-shaped pattern between diversification and development has been quite explored, but the search for the determinants of product diversification is still an open debate. In this article, we find that rent-seeking has an impact on diversification. In countries where the level of rent-seeking is high, the introduction of new goods will be smaller. New activities made by unestablished entrepreneurs are more prone to public rent-seeking, as they are long term and riskier projects, and these entrepreneurs are often financially constrained and usually opposed to an established lobby. Hence, in the presence of a high level of rent-seeking activities, the diversification process is damaged. This is specially true for low-income countries, where the effect of rent-seeking on diversification is greater than the wealth effect (measured by the GDP and GDP²).

In order to estimate the impact of rent-seeking on product diversification we use a highly disaggregated (on the product level) export database, between 1995 and 2007. We control for the endogeneity problem originated by reverse causality using system GMM. Our results are robust to the change of index for diversification and the proxy for rent-seeking, and also for the introduction of geographical and institutional variables. Consequently, we go beyond the description of the relationship between diversification and development and we give theoretical reasons and empirical estimations to why rent-seeking impacts diversification.

Following the consensus on the importance of diversification as a driver of economic growth, our findings contribute to understand how should diversification be managed. Arguments for the importance of the productive structure are a contemporary topic, as we find in a recent presentation by the World Bank chief economist (Lin, 2009). Knowing that an active industrial policy is crucial to economic development (Rodrik, 2004), these policies need to be implemented with care, as they lead to opportunities of corruption and rent-seeking (Dutt, 2009). Therefore, a positive industrial policy with the objective of diversifying the productive structure needs to incorporate the struggle against rent-seeking in its agenda.

Appendices

A Robustness check

Table 3: Robustness check

VARIABLES	(1) Theil	(2) Theil	(3) Theil	(4) Theil	(5) Theil
corrupt	0.592** (0.260)	0.837*** (0.308)	0.714** (0.317)	0.898*** (0.274)	0.671** (0.290)
gdp	-1.115e-04*** (3.769e-05)	-3.862e-05 (4.265e-05)	-1.151e-04** (4.479e-05)	-1.089e-04** (4.743e-05)	-1.180e-04*** (4.512e-05)
gdpsq	2.772e-09*** (9.119e-10)	1.774e-09** (8.782e-10)	3.031e-09*** (9.261e-10)	2.851e-09*** (9.697e-10)	2.968e-09*** (9.004e-10)
population	-2.469e-09*** (4.510e-10)	-2.413e-09*** (5.005e-10)	-2.376e-09*** (4.746e-10)	-2.522e-09*** (4.676e-10)	-2.490e-09*** (5.020e-10)
dister	0.001*** (0.000)				
airdist		2.423e-04*** (5.684e-05)			
landlocked			0.370 (0.270)		
icrg				0.793 (1.157)	
polity2					0.001 (0.027)
Constant	2.904*** (0.799)	1.197 (1.027)	2.812*** (1.004)	1.932* (1.168)	3.046*** (0.953)
Observations	971	987	987	852	974
Number of i	129	131	131	114	130
Nb of instruments	57	57	57	68	68
Hansen Test	0.281	0.163	0.222	0.409	0.542
AR(1)	0.0121	0.0122	0.0183	0.100	0.0182
AR(2)	0.207	0.158	0.133	0.142	0.234

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Estimated coefficients of time dummies are not reported here. Robust standard errors are in parentheses. Hansen test and the Arellano-Bond (AR(1) and AR(2)) tests are P values.

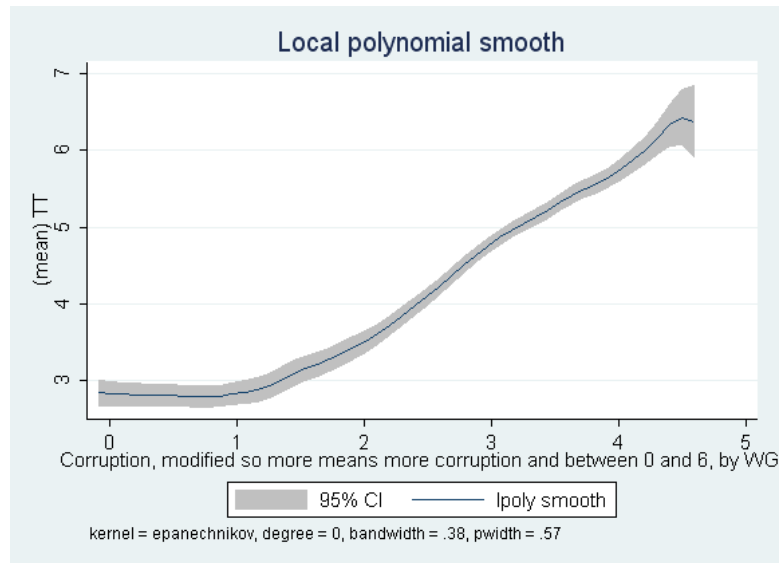


Figure 2: Diversification and rent-seeking: robustness check

B Country List and grouped statistics

LOW INCOME COUNTRIES (LIC): Bangladesh, Burkina Faso, Burundi, Cambodia, Central African Republic, Congo, Dem. Rep., Eritrea, Ethiopia, Guinea-Bissau, Liberia, Madagascar, Malawi, Mali, Mozambique, Niger, Rwanda, Sierra Leone, Togo, Uganda.

LOWER MIDDLE INCOME (LMIC): Angola, Armenia, Azerbaijan, Benin, Bosnia and Herzegovina, Cameroon, Chad, China, Congo, Rep., Cote d'Ivoire, Gambia, The, Georgia, Ghana, Guinea, Haiti, Honduras, India, Indonesia, Kenya, Kyrgyz Republic, Lao PDR, Mauritania, Moldova, Mongolia, Morocco, Nepal, Nicaragua, Nigeria, Pakistan, Papua New Guinea, Philippines, Senegal, Sri Lanka, Sudan, Tajikistan, Tanzania, Uzbekistan, Vietnam, Yemen Rep., Zambia.

UPPER MIDDLE INCOME(UMIC): Albania, Algeria, Argentina, Belarus, Bolivia, Brazil, Bulgaria, Chile, Colombia, Costa Rica, Croatia, Dominican Republic, Ecuador, Egypt, Arab Rep., El Salvador, Estonia, Guatemala, Iran, Islamic Rep., Jamaica, Jordan, Kazakhstan, Latvia, Lebanon, Lithuania, Macedonia, FYR, Malaysia, Mauritius, Mexico, Panama, Paraguay, Peru, Poland, Romania, Russian Federation, South Africa, Syrian Arab Republic, Thailand, Tunisia, Turkey, Ukraine, Uruguay, Venezuela.

HIGH INCOME COUNTRIES(HIC): Australia, Austria, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Korea, Rep., Libya, Netherlands, New Zealand, Norway, Portugal, Saudi Arabia, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Trinidad and Tobago, United Kingdom, United States.

Table 4: Summary statistics for LIC

Variable	Mean	Std. Dev.	Min.	Max.	N
gdp	716.591	258.328	259.294	1702.191	137
corrupt	3.224	0.406	1.731	3.936	137
TT	5.807	0.932	3.918	8.037	137

Table 5: Summary statistics for LMIC

Variable	Mean	Std. Dev.	Min.	Max.	N
gdp	2154.416	1128.145	833.426	7413.789	297
corrupt	3.267	0.351	2.371	4.234	297
TT	5.132	1.497	1.94	8.292	297

Table 6: Summary statistics for UMIC

Variable	Mean	Std. Dev.	Min.	Max.	N
gdp	8088.230	3074.158	3429.97	19327.148	322
corrupt	2.725	0.572	1.015	3.801	322
TT	4.041	1.264	1.999	7.42	322

Table 7: Summary statistics for HIC

Variable	Mean	Std. Dev.	Min.	Max.	N
gdp	28173.2	8086.485	11885.384	49358.906	231
corrupt	1.093	0.809	-0.088	3.374	231
TT	2.951	1.21	1.53	7.878	231

C Descriptive statistics of main variables

Table 8: Summary statistics for the whole sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Theil	4.399	1.655	1.517	8.292	1703
Herfindahl	0.135	0.188	0.002	0.926	1703
gdp	9181.416	10649.952	136.829	49358.906	1699
corrupt	2.603	1.015	-0.088	4.589	1172
population	45120725.915	145019661.735	1122457	1318309760	1703
prights	3.474	2.111	1	7	1703
women_par	13.685	9.404	0	49	1336
airdist	3927.511	2370.118	140	9590	1703
distr	354.925	469.74	7.952	2385.58	1677
icrg	0.554	0.209	0.111	1	1440
polity2	3.999	6.103	-10	10	1678

D Data Source

We use two other independent variables in our regressions, following the common framework in the literature. We use the Gross Domestic Product (GDP) per capita, taken from the World Development Indicators. It is measured in PPP values in constant 2005 US dollars, so we can compare our results to one of the main articles in this area (Cadot, Carrere, and Strauss-Kahn, 2007). We also use the size of population taken from the World Development Indicators.

Variable	Description	Source
gdp	Gross Domestic Product PPP per capita in constant 2005 US DOLLARS	World Development Indicators
population	Country Population	World Development Indicators
corrupt	Control of Corruption	World Governance Indicators
prights	Lack of political rights	Freedom House
women_par	Percentage of the number seats women hold in Parliament	Inter-Parliamentary Union survey
airdist	km to closest major port (Rotterdam, New York or Tokyo)	Gallup et al. (1999)
distr	mean distance to coast or river	Gallup et al. (1999)
icrg	ICRG Indicator of Quality of Government	ICRG
polity2	Revised combined Polity Score	PolityIV database

References

- ACEMOGLU, D. (2008): "Oligarchic Versus Democratic Societies," *Journal of the European Economic Association*, 6(1), 1–44. 3, 6, 10
- ACEMOGLU, D., AND F. ZILIBOTTI (1997): "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth," *Journal of Political Economy*, 105(4), 709–51. 5, 7
- ADES, A., AND R. D. TELLA (1999): "Rents, Competition, and Corruption," *American Economic Review*, 89(4), 982–993. 14
- AGOSIN, M. R. (2007): "Export Diversification and Growth in emerging economies," Series Documentos de Trabajos 233, Universidad de Chile. 4
- AL-MARHUBI, F. (2000): "Export Diversification and Growth: An Empirical Investigation," *Applied Economics Letters*, 7(9), 559–62. 4
- AMURGO-PACHECO, A., AND M. D. PIEROLA (2008): "Patterns of export diversification in developing countries : intensive and extensive margins," Policy Research Working Paper Series 4473, The World Bank. 2
- ANTHONSEN, M., A. LOFGREN, AND K. NILSSON (2009): "Rent Dependency: Bad for quality of Government," QoG Working Paper 23, The Quality of Government Institute. 14
- ARELLANO, M., AND O. BOVER (1995): "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68(1), 29–51. 15
- BALAND, J.-M., AND P. FRANCOIS (2000): "Rent-seeking and resource booms," *Journal of Development Economics*, 61(2), 527–542. 8
- BAUMOL, W. J. (1990): "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy*, 98(5), 893–921. 8
- BENEDICTIS, L., M. GALLEGATI, AND M. TAMBERI (2009): "Overall trade specialization and economic development: countries diversify," *Review of World Economics*, 145(1), 37–55. 8, 9, 22
- BERTHELEMY, J.-C. (2005): "Commerce international et diversification économique," *Revue d'économie Politique*, 115(5), 591–611. 5, 13
- BLUNDELL, R., AND S. BOND (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115–143. 15
- CADOT, O., C. CARRERE, AND V. STRAUSS-KAHN (2007): "Export Diversification: What's behind the Hump?," CEPR Discussion Papers 6590, C.E.P.R. Discussion Papers. 2, 9, 12, 27
- CALDERON, C., AND L. SERVEN (2004): "The effects of infrastructure development on growth and income distribution," Policy Research Working Paper Series 3400, The World Bank. 15
- CAMERON, C. A., AND P. K. TRIVEDI (2005): *Microeconometrics : Methods and Applications*. Cambridge University Press. 17
- COUTTENIER, M. (2008): "Relationship between natural resources and institutions," Documents de travail du Centre d'Economie de la Sorbonne 060, Université Panthéon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne. 14

- CROIX, D., AND C. DELAVALLADE (2009): "Growth, public investment and corruption with failing institutions," *Economics of Governance*, 10(3), 187–219. 10
- CUBERES, D., AND M. JERZMANOWSKI (2009): "Democracy, Diversification and Growth Reversals," *Economic Journal*, 119(540), 1270–1302. 3, 6, 7, 8, 12, 22
- DOLLAR, D., R. FISMAN, AND R. GATTI (2001): "Are women really the "fairer" sex? Corruption and women in government," *Journal of Economic Behavior & Organization*, 46(4), 423–429. 18
- DUTT, P. (2009): "Trade protection and bureaucratic corruption: an empirical investigation," *Canadian Journal of Economics*, 42(1), 155–183. 23
- DUTT, P., AND D. TRACA (2009): "Corruption and Bilateral Trade Flows: Extortion or Evasion," *Review of Economics and Statistics*, (forthcoming). 22
- FEENSTRA, R. C., R. E. LIPSEY, H. DENG, A. C. MA, AND H. MO (2005): "World Trade Flows: 1962-2000," NBER Working Papers 11040, National Bureau of Economic Research, Inc. 11
- GALLUP, J. L., J. D. SACHS, AND A. MELLINGER (1999): "Geography and Economic Development," CID Working Papers 1, Center for International Development at Harvard University. 22
- GAULIER, G., AND S. ZIGNAGO (2008): "BACI: A World Database of International Trade Analysis at the Product-level," Cepii working paper, CEPII Working Paper. 11
- HAUSMANN, R., J. HWANG, AND D. RODRIK (2007): "What you export matters," *Journal of Economic Growth*, 12(1), 1–25. 3, 5
- HAUSMANN, R., AND B. KLINGER (2006): "Structural Transformation and Patterns of Comparative Advantage in the Product Space," Working paper series, Harvard University, John F. Kennedy School of Government. 3
- HAUSMANN, R., AND D. RODRIK (2003): "Economic development as self-discovery," *Journal of Development Economics*, 72(2), 603–633. 3, 5
- HESSE, H. (2008): "Export diversification and Economic Growth," Working Paper 21, Commission on Growth and Development. 4, 5
- IMBS, J. (2007): "Growth and volatility," *Journal of Monetary Economics*, 54(7), 1848–1862. 5
- IMBS, J., AND R. WACZIARG (2003): "Stages of Diversification," *American Economic Review*, 93(1), 63–86. 2, 12, 20
- ISHAM, J., M. WOOLCOCK, L. PRITCHETT, AND G. BUSBY (2005): "The Varieties of Resource Experience: Natural Resource Export Structures and the Political Economy of Economic Growth," *World Bank Economic Review*, 19(2), 141–174. 14
- JANSEN, M., C. LENNON, AND R. PIERMARTINI (2009): "Exposure to External Country Specific Shocks and Income Volatility," CEPR Discussion Papers 7123, C.E.P.R. Discussion Papers. 8
- KALEMLI-OZCAN, S., B. E. SORENSEN, AND O. YOSHA (2003): "Risk Sharing and Industrial Specialization: Regional and International Evidence," *American Economic Review*, 93(3), 903–918. 8
- KAUFMANN, D., A. KRAAY, AND M. MASTRUZZI (2009): "Governance matters VIII : aggregate and individual governance indicators 1996-2008," Policy Research Working Paper Series 4978, The World Bank. 10

- KLINGER, B., AND D. LEDERMAN (2006): "Diversification, innovation, and imitation inside the Global Technological Frontier," Policy Research Working Paper Series 3872, The World Bank. 2, 5, 7, 9
- KOREN, M., AND S. TENREYRO (2007): "Volatility and Development," *The Quarterly Journal of Economics*, 122(1), 243–287. 5, 8
- LEDERMAN, D., AND W. F. MALONEY (2008): "In search of the missing resource curse," Policy Research Working Paper Series 4766, The World Bank. 4
- LIN, J. Y. (2009): "New Structural Economics - A Framework for Rethinking Development," mimeo. 2, 23
- LIND, J. T., AND H. MEHLUM (2007): "With or Without U? - The appropriate test for a U shaped relationship," MPRA Paper 4823, University Library of Munich, Germany. 12
- MALIK, A., AND J. R. W. TEMPLE (2009): "The geography of output volatility?," *Journal of Development Economics*, 90(2), 163–178. 5, 8
- MATSUYAMA, K. (1992): "Agricultural productivity, comparative advantage, and economic growth," *Journal of Economic Theory*, 58(2), 317–334. 5
- MAURO, P. (1995): "Corruption and Growth," *The Quarterly Journal of Economics*, 110(3), 681–712. 6
- MOHTADI, H., AND T. L. ROE (2003): "Democracy, rent seeking, public spending and growth," *Journal of Public Economics*, 87(3-4), 445–466. 10
- MURPHY, K. M., A. SHLEIFER, AND R. W. VISHNY (1993): "Why Is Rent-Seeking So Costly to Growth?," *American Economic Review*, 83(2), 409–14. 3, 6, 8, 9, 18
- PARTEKA, A., AND M. TAMBERI (2008): "Determinants of Export Diversification: an Empirical Investigation," Working Papers 327, Universita' Politecnica delle Marche (I), Dipartimento di Economia. 2, 8, 22
- RAMEY, G., AND V. A. RAMEY (1995): "Cross-Country Evidence on the Link between Volatility and Growth," *American Economic Review*, 85(5), 1138–51. 5
- RODRIK, D. (2004): "Industrial Policy for the Twenty-First Century," Working paper series, Harvard University, John F. Kennedy School of Government. 23
- ROODMAN, D. (2006): "How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata," Working Papers 103, Center for Global Development. 17, 18
- (2009): "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158. 17
- STANLEY, D. L., AND S. BUNNAG (2001): "A New Look at the Benefits of Diversification: Lessons from Central America," *Applied Economics*, 33(11), 1369–83. 5
- SVENSSON, J. (2000): "Foreign aid and rent-seeking," *Journal of International Economics*, 51(2), 437–461. 3, 6, 10
- WINDMEIJER, F. (2005): "A finite sample correction for the variance of linear efficient two-step GMM estimators," *Journal of Econometrics*, 126(1), 25–51. 17