ERUDITE
ÉQUIPE DE RECHERCHE SUR L'UTILISATION
DES DONNÉES INDIVIDUELLES EN LIEN
AVEC LA THÉORIE ÉCONOMIQUE

Sous la co-tutelle de :
UPEC • UNIVERSITÉ PARIS-EST CRETEIL
UPEM • UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE

DOCUMENT DE TRAVAIL
N° 03-2015

Export diversification and the legacy of the Soviet Union

Patrick Domingues, Felipe Starosta de Waldemar

UNIVERSITÉ
PARIS-EST

Export diversification and the legacy of the Soviet Union

Patrick Domingues[†] **and Felipe Starosta de Waldemar**[‡]

March 2015

Abstract:

This paper contributes to the analysis of countries' export diversification dynamics using a quasi-natural experiment provided by the fall of the USSR. Using cross-sectional estimations, this paper offers evidence on the exogenous relationship between the Soviet legacy and export diversification using a sample of former members of the Soviet Union and European countries. We measure this legacy using a proxy variable based on a country's energy production infrastructure, as this dimension reflected Moscow's influence because it was used as a foreign policy tool. The energy production technology gap (EPTG) variable is constructed as the Euclidean distance between a country's share of each of the six main electricity production sources and the corresponding ratio for Russia. We find that the greater the EPTG, the higher the degree of export diversification *ceteris paribus*. An inefficient legacy, unfavorable to diversification, which arose from the Soviet period, constituted a blockage in the export diversification process after the fall of the USSR.

KEYWORDS: export diversification, Soviet Union, institutions, energy

JEL CLASSIFICATION: O11, P28, F14

[†] University Paris Est Créteil (UPEC), ERUDITE: patrick.domingues@u-pec.fr
[‡] University Paris-Sud, RITM: felipe.starosta-de-waldemar@u-psud.fr

"Communism is Soviet power plus the electrification of the whole country." Vladimir Lenin

Introduction:

Several analyses have been conducted to identify the determinants of export diversification. However, few studies have succeeded in empirically addressing the issue because many potential determinants, such as institutional quality, may suffer from endogeneity problems. In this paper, we propose to address this endogeneity issue using a quasi-natural experiment provided by the fall of the USSR to investigate the economic effect of institutional legacy on export diversification.

As highlighted by numerous contributions, the institutions of former USSR members and some central and eastern European countries were influenced by Moscow during the Soviet period (*e.g.,* Ericson, 1991), inducing a legacy effect that could affect their pattern of diversification during the post-Soviet period. We investigate this effect, using the similarity of the country's energy production structure with that of Russia after the fall of the USSR as a proxy for the country's exposure to Russian influence during the Soviet period. We construct this proxy variable as the energy production technology gap (EPTG) between each country and Russia. This EPTG is a distance calculated in a six-dimensional Euclidean space derived from six main sources of electricity production. The existence of similar energy production structures between two countries is either the consequence of the availability of the same natural resources or the result of the influence of one country on the other. Russia has influenced the energy dependency of its neighbors (Balmaceda, 2008), as in many former Soviet and eastern European countries, the design of energy production infrastructure was influenced following centralized decisions from Moscow. As a consequence, the energy production structure does not reflect the natural resource endowment of each country.

There is substantial evidence of the influence of this centralized decision making regarding national energy infrastructure in various World Energy Outlook reports (WEO: 1998, 1999) compiled by the International Energy Agency. First, there is significant heterogeneity between other countries in the region and Russia. For example, the 1998 WEO notes: "The Transition Economies have significant variations in their electricity output mix. Electricity generation in Poland and Kazakhstan is almost entirely based on coal; natural gas is the most significant source of power for Russia, Moldova, Uzbekistan and Belarus". Second, the pricing in the energy sector did not follow an efficient and rational framework. A substantial share of the sector was subject to non-payment or barter between companies, which imposed huge costs on the economy. These chains of payments led to financial slumps and also had an effect on production and investment. In summary, the underpricing of energy resulted in major distortions and considerable economic costs. Third, and more important, the basic logic of the energy sector during the Soviet era was not efficiency. As noted by the 1999 WEO, "Energy has traditionally played a central role in the Russian economy. During the Soviet era, the energy sector was developed to provide resources for heavy industry and national defense, as well as to earn foreign exchange to finance imports." This evidence indicates that the energy production infrastructure reflected the institutional quality and decision-making system of the former Soviet Union.

First, Ericson (1991) noted just after the collapse of the USSR, the economic system in the Soviet economy is characterized by central planning, put in place administratively through direct commands and detailed instructions. Second, once the former USSR had collapsed into different republics (the Commonwealth of Independent States; CIS henceforth), a common energy system remained a major feature of the union treaty signed by these republics (Gleason, 1992). Third, the sovereign energy policy perspective (Cherp and Jeweel, 2011) highlights that energy security focuses on the configuration on power, interest and alliances. Cherp and Jewell (2011) offer evidence on this subject using the example of the negotiations between Russia and Ukraine over the extension of the lease of military bases in Crimea.

More important, Balmaceda (2008) offers a detailed review of the centralized energy issue, as the central topic of her book concerns how the former USSR countries managed and coped with their energy dependency *vis-à-vis* Russia. Cheap and abundant energy supplies were one of the factors that held the Soviet Union together. The author also notes that the 2006 energy crisis between Russia and Ukraine generated alarm among EU countries that Russia's dominance of the energy markets in the former USSR was excessive. Furthermore, Russia maintains considerable influence over trade in gas and oil in the former Soviet republics that extends beyond "the infrastructural dependencies inherited from the Soviet Union." In summary, this dependency offers such considerable rents to domestic elites that it ultimately has negative implications for a state's economic and political development. Therefore, it is not only the lack of natural resources or energy dependency that has a negative effect on the economy but also the rents that arise for local elites that result from this dependency, which are defined by Balmaceda (2008) as the "rents of energy dependency".

In this paper, we show that the legacy of the Soviet Union represents an impediment to export diversification in a sample of former members of the Soviet Union and European countries. The mechanism for this effect is that the quality of the economic and political institutions established by the centralized superpower did not favor the elements that encourage diversification, such as the emergence of entrepreneurs, higher human capital, privatization or high-quality infrastructure (Omgba, 2014). The logic in the Soviet Union was reflected in the country's preferred system of energy production infrastructure. Hence, the closer to Moscow a nation was at the beginning of its economic transition, the fewer incentives it had to encourage diversification.

Export diversification is a key issue for a country's development, a subject that was revived by the seminal contribution of Imbs and Wacziarg (2003). According to their argument, countries diversify as they begin to develop until they reach a turning point at a high income level and begin to re-concentrate their productive structure. However, most studies examine the effects of diversification on other economic variables. We can highlight its effect on volatility reduction (Ramey and Ramey, 1995; Stanley and Bunnag, 2001; Imbs, 2007, Koren and Tenreyro, 2007) and diminished risks to economic agents through the financial system (Acemoglu and Zilibotti, 1997). Furthermore, Cadot, Carrère, and Strauss-Kahn (2013) note that the effect of diversification on economic growth is also related to the literature on the natural resource curse, in which a more concentrated productive structure has negative effects on GDP growth.

This U-shaped relationship between diversification and development remained unexplained, but as we find in the literature review of Cadot, Carrère, and Strauss-Kahn (2013) on the determinants and drivers of diversification, different explanations have been proposed to fill this gap. Various studies analyzing diversification suggest myriad empirical determinants without shedding light on the underlying mechanisms that explain their effects. Concerning this specific literature that explores the determinants of product export diversification, Cadot, Carrère, and Strauss-Kahn (2013) find that among several determinants, such as infrastructure and human capital, institutional quality is a major positive determinant of export diversification. In contrast, Parteka and Tamberi (2013) contend that country size and distance to major markets are important determinants of product export diversification.

However, as previously mentioned, one of the main difficulties in capturing this effect of institutional quality on export diversification is the exogeneity assumption regarding the institutional variable. An interesting strategy to address this issue is provided by Omgba (2014). The author focuses on former colonies that are also oil exporting countries for the period 1995-2011. He notes that the export diversification patterns of these countries also have an institutional and political background. His main explanatory variable is the length between the beginning of oil production and the date of a country's political independence. The author argues that the greater the difference between these two dates, the more likely it is that a country has a diversified export basket relative to another oil exporting nation. The reason is that if oil production began before or near to independence, the local political elites established an institutional environment favorable only to the exploitation of natural resources. No further programs for developing infrastructure or human resources were established. Conversely, if oil production began many years after independence, a non-resource sector (*e.g.*, agricultural products) would have been sufficiently developed. This leads to a higher level of export diversification.

In our paper, we use the energy production sources, captured by the EPTG variable, to address this issue. As presented above, the decision regarding how to produce electrical power was largely made in Moscow and disconnected from the real natural resources in each country; hence this decision did not obey economic and political incentives. Therefore, export diversification did not influence the sources of energy production, a variable that captures the Soviet Union's legacy. As a consequence, we can exclude a problem of reverse causality.

Apart from the literature on the determinants of export diversification, this paper contributes to the literature on the legacy that Soviet institutions have on economic variables after the collapse of the USSR. Here, we present a selection of papers that are closely connected to our study. Berkowitz et al. (2012) focus on the link between finance and growth, exploiting the exogenous variation in banking across Russian regions generated by the creation of specialized banks in the final years of the Soviet Union. First, the authors find that the locations of these specialized banks were exogenous to economic factors. Second, they show that the increase in financial activities did not create economic growth – although increased lending did induce a surge in employment. In contrast, Ben Yishay and Grosjean (2014) focus on initial endowments and economic reforms for a sample of 27 former communist countries in Europe and Central Asia. They consider initial endowments in two respects: institutional (the political history of the region) and economic (the distribution of the natural resources sector at the beginning of transition). They find that deeply rooted determinants of institutional quality and the concentration of natural resources at the beginning of the transition period explains part of the variation in the success of economic reforms. Countries that had political and economic elites that controlled less of the centralized rents were unable to prevent the liberalization of economic gains in the economy. Other papers also focus on the legacy of the former USSR. Libman and Obydenkova (2013) examine the persistence of corruption and whether Communist rule contributed to this major problem. They focus on an intra-country comparison by analyzing sub-national regions of Russia, and their focus is the effect of members of the Communist Party. Libman and Obydenkova (2013) find that legacy of the former Soviet Union is a highly powerful factor in explaining the variation in corruption across different regions in Russia: where there were higher levels of membership in the Communist Party in 1970, there were higher levels of corruption in 2010. Shurchkov (2012) also explores intra-country variation in Russia. She focuses on the influence of political newcomers on entrepreneurial activity in 71 different regions over the 1994-2006 period. Shurchkov (2012) shows that regions with governors supported by Putin had fewer small and medium-sized enterprises than did regions controlled by old elites. Her argument is that these new elites supported by Putin attempted to dominate all of a region's economic rents, and therefore, entrepreneurial activity stalled at low levels.

Regarding these two literatures, the contribution of this paper is twofold. First, it contributes to the literature on export diversification. As presented above, few studies have convincingly identified the determinants of diversification, due to endogeneity problems. Here, we demonstrate the effect of the Soviet legacy using exogenous variation in the implementation of centralized power from Moscow in many former USSR republics and European countries. To measure this, we use the energy production technology gap (EPTG) in terms of the sources of inputs for electricity generation. The second contribution concerns the effect of the Soviet Union's institutional effect on an economic variable at the country level. On the one hand, export diversification is an important contributor to the development of nations. However, no previous article has investigated its pattern while devoting particular attention to this group of countries. On the other hand, different studies, at different levels of analysis, have highlighted the influence of the institutional legacy of former Soviet Union countries on various economic variables. We find that the greater the EPTG, the higher the degree of export diversification *ceteris paribus*. An inefficient legacy, unfavorable to diversification, which arose from the Soviet period, served to impede the export diversification process after the fall of the USSR.

The remainder of this paper is structured as follows. In Section 2, we present our hypothesis, the variables and the data. The empirical strategy and main results using OLS estimations are then presented in Section 3, in addition to an IV strategy following Lewbel's (2012) method. Section 4 concludes.

Section 2: Data and hypothesis

2.1 Data

This paper argues that the level of export diversification in a country tends to have a negative relationship with its energy production technology gap (EPTG) with Russia. The EPTG is treated as a proxy for the institutional environment experienced by former members of the USSR and, to a lesser extent, by countries in Central and Eastern Europe (CEE).

To test this proposition, our dependent variable is a country's level of export diversification. We follow Cadot, Carrère and Strauss-Kahn (2011) and construct the Theil diversification index as:

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

where $X_{j,i,t}$ is the export value of each good "j" in each country "i" and year "t", and *n* is the number of products exported for each country-year pair. A higher level of the Theil index indicates that a country is *more* concentrated – the opposite of diversified. Therefore, in our estimations, a variable taking a negative sign means that it has a positive effect on diversification. To construct this variable, we use export data from the BACI dataset. It covers 5017 products for all countries in the world, with product-level data at the 6-digit level of disaggregation[1], from 1992 to 2010. Note that the level of disaggregation is important for economic analysis, as explained by De Benedictis et al. (2009), and we focus on the most disaggregated level possible for a cross-country study.

As robustness checks, we perform a series of tests that modify our dependent variable. Our primary objective here is to show that our explanatory variable has strong predictive power for various measures of diversification – or changes to the composition of nations' export baskets. In our first test, we use a Theil diversification index using import data to assess whether our observed effect is not just a spurious correlation. We expect the effect of the EPTG to not be significant on the Theil variable for imports, as in this case, there is no reason that our EPTG variable should have an effect. Second, we use the average value of the export diversification index for the full period to avoid capturing a year-specific effect (such as a price fluctuation or high demand due to an exogenous shock).

Third, we consider instead the Herfindahl index, measured following Cadot, Carrère and Strauss-Kahn (2011):

$$H_{i,t} = \frac{\sum_j (s_j)^2 - 1/n}{1 - 1/n}$$

where $s_j = x_j / \sum_{j=1}^{n} x_j$ is the share of exports of product j (with export value x_j) in total exports. The Herfindahl index is positively correlated with the Theil index, and it also measures the diversification of the export basket.

Next, our objective is to construct different variables that capture similar characteristics of the composition of a nation's export basket, but in alternative ways. Hence, we also test the effect of the EPTG on other dependent variables. In doing so, we switch to a variable proposed by Brenton and Newfarmer (2008), namely, international export market penetration:

$$IEMP_{i,t} = \frac{\sum_{j \in I_{ji}} \sum_m Y_{jim}}{\sum_{j \in I_{ji}} \sum_m Z_{jm}}$$

The IEMP variable has the ability to provide us with the bilateral dimension of export diversification. It measures the ratio of the actual number of bilateral trade flows to the potential bilateral trade flows. I_{ij} represents the set of products "j" for a country "i" where positive value of exports exist, Y_{ijm}= 1 for $X_{ijm} > 0$, and Y_{ijm}= 0 otherwise. Z_{im}= 1 for $M_{im} > 0$, else Z_{im}= 0, where X_{ijm} is the value of exports of product "j" from exporter "i" to importer "m", and M_{ik} is the value of imports of product *j* by importer *m*.

The last measure of the diversification and structural transformation of a nation's export basket that we use as a dependent variable is density. We use the Product Space

[1] This dataset is constructed based on COMTRADE data using an original procedure that reconciles the declarations of exporters and importers (Gaulier and Zignago, 2010).

framework that was developed by Hidalgo et al. (2007) and applied at the country level by Kali et al. (2013). The density variable measures how connected a country's export basket is and it is constructed as follows:

$$Density_{i,t} = \sum_{j=1}^{z_i} \left(\frac{e_j}{\sum_{l \in R_i} e_l} \omega_j^i \right)$$

$$where \quad \omega_j^i = \frac{\sum_{l \in R_i, j \neq k} \varphi_{jk}}{\sum_{j \neq k} \varphi_{jk}}$$

$$and \quad \varphi_{jk} = \min[\Pr(j|k), \Pr(k|j)]$$

where "l" indexes all the products in country's specialization basket "Rx"[2], "j" and "k" are all products that exist in the product space, z_i is the number of products for which a country has revealed comparative advantage and e_j is the export value of product "j" for country i. φ_{jk} (called proximity) is calculated using the minimum of the conditional probabilities of exporting a product "j" with comparative advantage (following the definition of Balassa, 1964) knowing that one also exports the product "k" with comparative advantage.

In Table 1 below, we report the pairwise correlation between these measures of the structure of countries' export baskets. While having different definitions, the evidence below reveals a strong correlation, indicating similar patterns for all countries in our sample.

Table 1 : Pairwise correlation

Dependent Variable	Theil	Herfindahl	Density	IEMP
Theil	1			
Herfindahl	0,8643	1		
Density	-0,8881	-0,5939	1	
IEMP	-0,5774	-0,2913	0,7993	1

Furthermore because our goal is to identify the effect of the EPTG with Russia on export diversification, we require a variable to measure this gap. As we wish to use this EPTG as an explanatory variable, special attention should be paid to the potential reverse causality issue. As described in the introduction, institutions and other determinants of diversification suffer from a problem of reverse causality. This problem emerges because diversification could also have an effect on these explanatory variables. On the one hand, efficient institutions are an incentive for entrepreneurs to develop new activities and diversify their exports because an efficient institutional framework creates very few barriers to doing so. However, on the other hand, one could assume that in a country with a more concentrated export basket, the rents of exports will also be concentrated in a few selected sectors, thereby increasing the opportunity for rent seeking and corruption.

As a consequence, to measure the true effect of the institutional legacy on diversification, we need an exogenous variation in the institution environment that avoids all problems of reverse causality. So we constructed this measure as a country-level variable calculated for the initial year of the period of study during all that will remain time invariant.

[2] For further information on the Product Space, see Hidalgo et al. (2007).

We define the variable EPTG$(i, Russia)$ as the distance between each country i and Russia in a six-dimensional Euclidean space derived from six main sources of electricity production, following this formula:

$$EPTG(i, Russia) = \sqrt{\sum_{x=1}^{n}(x_{i,t0} - x_{Russia,t0})^2}$$

Where $\sum_{x=1}^{n}(x_{i,t0} - x_{Russia,t0})^2$ is the sum of the squared difference between each country i and $Russia$ in terms of share of electricity produced by each of the x sources of electricity at time t_0. The n sources of electricity refer to the inputs used to generate electricity. For this study, we consider six main sources of electricity production: oil, nuclear, gas, coal, hydroelectric sources and renewable sources excluding hydropower (*i.e.*, geothermal, solar, tidal, wind, biomass, and biofuels). Data on electricity production come from the World Bank.

2.2 Hypothesis

As mentioned above, in this paper we investigate the economic effect of institutional legacy on export diversification using a quasi-natural experiment provided by the fall of the USSR. Table 2, Column 1 presents the average value of the Theil diversification index for three different groups of countries from 1989 to 2010. This column indicates that former members of the USSR display a higher value of the Theil index, meaning that they are less diversified than CEE countries, which are in turn less diversified than other European countries.

In other words, during the post-USSR period, former members of the Soviet Union seem to have followed a path of diversification that differed from that of the rest of Europe. Our hypothesis is that this statistical result could be explained by an institutional legacy of the USSR's institutional framework experienced by these former soviet countries and, to a lesser extent, by CEE countries (because these countries were also influenced by Moscow during the Soviet period). To capture this effect, we use the country's EPTG with Russia after the fall of the USSR as a proxy for its institutional proximity with Russia during the Soviet period (this gap is calculated following the formula described in section 2.1).

Table 2 : Average Theil and average energy production technology gap (EPTG) decompose by group

Groups	(1) Average Theil	(2) Average EPTG
Former members of USSR	4,727	56,045
CEE	2,815	61,557
Other European countries	2,661	62,123

Source: author's calculations

Table 2, Colum 2 reports the energy production proximity to Russia for three different groups of countries in 1992 following the fall of the USSR (CIS). This column indicates that former members of the USSR (except Russia) are closer to Russia than CEE countries or other European countries. This means that the electricity production in former Soviet republics is much more similar to that in Russia than the rest of Europe in terms of the shares of the different inputs used to generate electricity. We can also see that, albeit to a lesser extent, the electricity production in CEE countries is also more similar to the Russian one than other European countries.

Furthermore, Figure 1 highlights that, in 1992, countries close to Russia exhibit a lower degree of export diversification (reflected by a higher value of the Theil index) and emphasizes that this relationship between the EPTG with Russia and the Theil index follows a quadratic form.

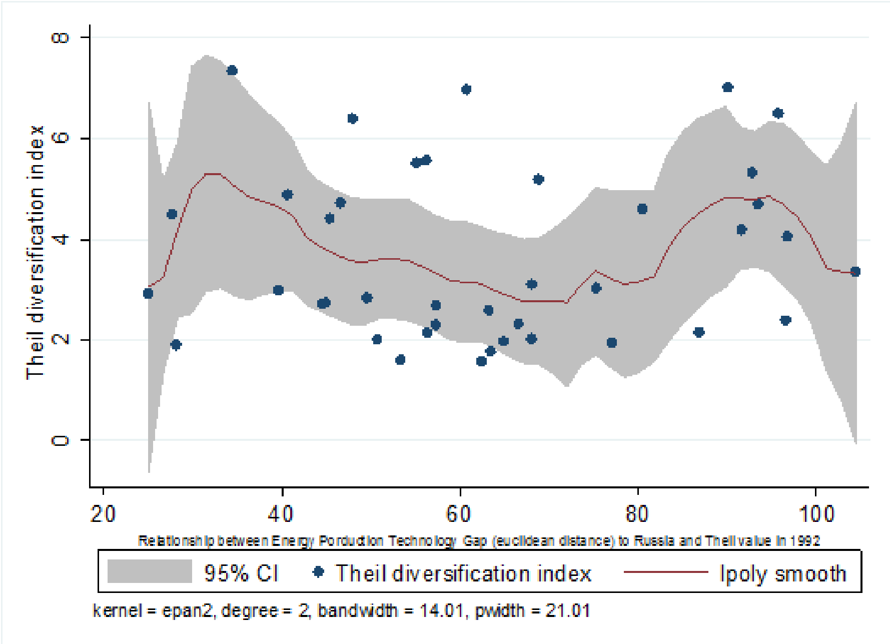

Figure 1: Relationship between the energy production technology gap (Euclidean distance) with Russia and Theil index value in 1992

Following these preliminary results, the estimation equation can be written as follows:

$$Diversification\ index_i = \beta_0 + \beta_1\ EPTG_i + \beta_2\ EPTG_{i}\_sq + \theta_k X'_i + \varepsilon_i \quad (1)$$

where Diversification index_i represents the alternative indicators of the export basket structure presented above, *EPTG* and *EPTG_sq* are the variables capturing proximity to Russia and its squared term, respectively (because the relationship between our left-hand variable and the EPTG appears to be quadratic), X'_i is a set of control variables presented in Table 3[3] and ε_i is the disturbance term.

Before turning to our results, one important point has to be discussed. Because the period of interest for this study is that after the fall of the USSR, the EPTG variable should be constructed for 1989. However, we cannot select this date as our initial year of study. First, some European countries became fully independent after this date, for example the split of Czechoslovakia or the independence of some of the former members of the CIS. Second, as we employ numerous export diversification indicators (presented above) derived from the

[3] Table B1 in the appendix presents the sample analyzed in the study.

UNCOMTRADE dataset, we face the problem that international trade data only began to be directly reported for certain countries in the mid-1990s. For example, as mentioned by Van der Marel and Dreyer (2014), data from UNCOMTRADE are only recorded for Russia after 1996. Therefore, we have to consider 1996 as the starting date of our period of study, and hence we constructed our EPTG variable for this year, and we restrict the period of study from 1996 to 2010. Nevertheless, this could violate the exogeneity assumption described above. To avoid this issue, we must ensure that the different shares of inputs used for electricity production in Russia and the other countries remained stable from 1992 to 1996.

Table 3 : Descriptive statistics (1996)

Variable	Obs	Mean	Std. Dev.	Min	Max
Energy production technology gap (EPTG)	43	61,037	21,254	25,852	107,031
Theil	43	3,133	1,364	1,528	7,561
Herfindahl	43	0,057	0,121	0,002	0,637
Density	43	0,243	0,117	0,017	0,528
International Export Market Penetration	43	0,012	0,016	0,001	0,059
RGDP_pc	43	14074,49	11324,15	1216,569	48861,54
Population	43	16,021	20,594	0,270	82,200
Natural Resources	42	3,789	8,957	0	46,564
School Enrollment, primary	35	27,968	12,096	7,031	53,547
School Enrollment, secondary	35	53,274	13,981	18,841	81,104
Investment	43	0,198	0,082	0,057	0,435
Rule of Law	43	0,378	1,121	-1,693	1,907
Oil	43	99,493	211,102	0	1233,585
Coal	43	1965,508	6560,482	0	34359,05
Natural Gas	43	112,3518	365,993	0	1724,184
Distance to Moscow	43	1934,20	753,108	675,42	3906,67
Landlocked	43	0,3488	0,482	0	1
Air distance to major ports	40	1782,55	1366,478	140	5300
Distance to coast or river (km)	42	277,685	541,863	11,039	2291,68
Theil without energy products	43	2,949	1,253	1,5136	6,9301
Switzerland EPTG	43	70,713	27,044	0	120,429

Figure 2 in the appendix depicts the evolution of the different shares of inputs used for electricity production in Russia from 1992 to 1996. As we can see, this figure indicates that this evolution is quite stable during this period, and hence the use of data from 1996 to construct our EPTG variable could be reasonably considered to remain exogenous in 1996.

Figures 3, 4, and 5 depict the evolution of the different shares of inputs used for electricity production for the three different groups of countries considered in this study, from 1992 to 1996. All of these figures display the same pattern: the evolution is stable during this period. We consider that the use of data from 1996 to construct our EPTG variable does not invalidate the exogeneity assumption.

Finally, Figure 6 provides another test regarding the evolution of the EPTG between Russia and the three different groups of countries considered in this study from 1992 to 1996. The EPTGs remain stable during this period, which confirms the results derived from the previous tables.

Section 3: Empirical strategy and results

3.1 OLS results

We analyze the relationship between the Soviet Union’s legacy and export diversification, and hence we begin by using the export concentration index in 1996 as a dependent variable. Because our variable of interest is a time-invariant variable, we first perform an OLS analysis without country fixed effects. Following Acemoglu et al. (2003) ‘‘*our interest is in the historically determined component of institutions (that is more clearly exogenous), hence not in the variations in institutions from year-to-year. As a result, this regression does not (cannot) control for a full set of country dummies.*’’

Table 4 presents the initial results of several OLS estimations. Column 1 reports the results derived from a traditional benchmark model based on explanatory variables commonly used in empirical studies of the determinants of diversification (Cadot, Carrère and Strauss-Kahn, 2013). Although numerous explanatory variables differ across authors and studies, income per capita, its quadratic term and population are three variables that are included in many studies, and hence these three variables constitute our benchmark model (Omgba, 2014). We note that the results presented in this column, which are consistent with those found in the literature, indicate a negative quadratic relationship between real GDP and diversification and a negative influence of population size.

Column 2 presents the results of the regression between our variable of interest, *EPTG*, its quadratic term, *EPTG_sq*, and the dependent variable, the *Theil* diversification index. Furthermore, because natural resources can influence the choice of energy production infrastructure established in a country, we add the variable *Natural resources*, which captures the total natural resource rents as percentage of GDP, to control for the natural resource endowment of each country included in our sample. Finally, because we cannot add country fixed effects because our variable of interest is a time-invariant country-level characteristic, we add a dummy variable that takes value 1 if a country is a former Soviet republic and 0 otherwise to account for unobservable variables specific to former members of the USSR. The results indicate a significant and negative correlation between *EPTG* and export diversification and a positive relationship between *EPTG_sq* and *Theil*. This two coefficients are statistically and significantly different from zero at the 1 percent level, meaning that the greater the EPTG, the higher the degree of export diversification *ceteris paribus*. In other words, we find that in 1996, when a country’s energy production infrastructure is more similar to that of Russia, that country is less diversified. This effect persists until a turning point, where the more different a country is from Russia, the less concentrated its exports will be.

To check the robustness of this first result, Column 3 and Column 4 present the results for different specifications that successively use different dummy variables for different geographical regions to add greater flexibility to our previous specification to account for unobservable characteristics related to these regions (Table B1 in the Appendix presents these different dummies). The results presented in these two columns are relevant to the results from Column 2, and the relationship between the EPTG with Russia and diversification is negative and follows a quadratic pattern. Furthermore, Column 5 provides a robustness check for another specification for which we use alternative measures of the natural resource endowment. The results do not contradict the relationship highlighted previously; we still find a negative effect of the EPTG on the degree of export diversification. Moreover, this last specification better controls for unobservable variables because the R-squared is higher for this specification. As a consequence, we will use this specification in the remainder of the study to perform further robustness checks.

One important robustness check for our preferred specification concerns the potential spurious correlation that may drive our previous results. The use of a variable that captures import diversification as a dependent variable is a simple approach to address this issue because if there were a spurious correlation issue, this variable would be significant. There is no theoretical justification suggesting that the EPTG would have an effect on import diversification. Column 6 provides the results using this alternative dependent variable; it clearly indicates that our variables of interest are no longer significant, meaning that our previous results are not driven by a spurious correlation.

The strategy used to obtain the results in Table A1 in the appendix follows the same logic as that used for Table 4, except that we now use the average of the export concentration index over the period 1996-2010 as the dependent variable (instead of the value of the Theil index for 1996), and we evaluate all the covariates in the initial year of 1996. This strategy is used to verify that the previous results provided in Table 4 are not associated with particular characteristics of the year 1996, as well to emphasize that the negative impact of the Soviet legacy induced a lasting effect on export diversification. The two main results highlighted in this table are the following. First, we find a significant and negative relationship between *EPTG* and export diversification and a positive relationship between *EPTG_sq* and the average value of the Theil index. Thus, the relationship between our variables of interest and our new dependent variable exhibits the same pattern as that reported in Table 4. This means that when the energy production infrastructure of a particular country in 1996 is similar to that of Russia, that country is less diversified over the period of study after controlling for the natural resource endowment and unobservable regional characteristics until a turning point. Second, as the results from Column 6 indicate that our variables of interest are no longer significant, we conclude that our previous results are not driven by a spurious correlation.

In Table 5, we perform a series of robustness checks to control for other possible determinants of export diversification, based on the articles presented above. Again, we focus on variables that could capture the unobserved heterogeneity that we do not entirely control for because we use OLS estimations and to avoid an omitted variables issue.

In the first column, we introduce the “Rule of Law” variable, taken from the World Governance Indicators (WGI: Kaufmann, Kraay and Mastruzzi, 2010) to account for the institutional environment in 1996. This variable could have an effect on export diversification, as higher quality would favor the introduction of new export products. For instance, entrepreneurs and business man are more incited to invest in new activities if they know that the risk of expropriation is lower. However, although the sign is as expected – a higher quality institutional environment reduces export concentration – the coefficient is not significant. Moreover, our EPTG variable capturing the Soviet legacy remains negative and significant, and is now of larger magnitude. That implies that this EPTG variable does not capture the overall institutional quality in 1996 but the institutional proximity to Moscow before the fall of the Soviet Union.

Table 4 : Initial results: OLS using data for 1996

VARIABLES	(1) Theil	(2) Theil	(3) Theil	(4) Theil	(5) Theil	(6) Imports Concentration
EPTG		-0.080***	-0.087***	-0.090***	-0.093***	-0.069
		(0.025)	(0.029)	(0.028)	(0.026)	(0.041)
EPTG_sq		0.001***	0.001***	0.001***	0.001***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc	-0.000***	-0.000*	-0.000***	-0.000***	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc_sq	0.000***	0.000*	0.000***	0.000***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population	-0.016**	-0.013***	-0.015***	-0.015***	0.002	0.006
	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.010)
URSS		0.637*	-1.178*			
		(0.366)	(0.591)			
URSS_a				-1.117**	-0.895*	1.421*
				(0.471)	(0.520)	(0.832)
URSS_b				-0.272	0.763	1.260
				(0.601)	(0.759)	(0.878)
URSS_c				0.016	0.832	1.190
				(0.519)	(0.545)	(0.954)
CEE			-1.763***	-1.420***	-1.281***	0.744
			(0.438)	(0.355)	(0.361)	(0.657)
Ex-Yugoslavia			-1.359***	-0.992**	-0.946**	0.029
			(0.495)	(0.390)	(0.363)	(0.625)
East Mediterranean			-0.904*	-0.692	-1.376***	0.383
			(0.467)	(0.434)	(0.496)	(0.686)
Natural Resources		0.048***	0.049***	0.036***		
		(0.012)	(0.012)	(0.010)		
Oil					-0.004***	-0.002*
					(0.001)	(0.001)
Coal					-0.000***	-0.000
					(0.000)	(0.000)
Natural gas					0.002***	0.001
					(0.001)	(0.000)
Constant	4.789***	5.369***	7.744***	7.214***	6.541***	4.265***
	(0.370)	(0.886)	(1.159)	(0.913)	(0.832)	(1.277)
Observations	43	42	42	42	43	43
R-squared	0.447	0.740	0.828	0.873	0.887	0.560

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The Energy production technology gap (EPTG), our variable of interest, is calculated as the Euclidean distance.

Table 5 : OLS Results, Robustness checks, alternative mechanisms

VARIABLES	(1) Theil	(2) Theil	(3) Theil	(4) Theil	(5) Theil	(6) Herfindahl
EPTG	-0.100***	-0.098***	-0.099***	-0.095***	-0.104***	-0.004*
	(0.028)	(0.031)	(0.030)	(0.027)	(0.032)	(0.002)
EPTG_sq	0.001***	0.001***	0.001***	0.001***	0.001***	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc	0.000	-0.000	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc_sq	-0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population	-0.000	-0.006	-0.005	0.002	-0.006	0.000
	(0.004)	(0.007)	(0.006)	(0.004)	(0.006)	(0.000)
URSS_a	-0.923*	-1.040	-1.217*	-0.824	-1.008*	-0.023
	(0.497)	(0.633)	(0.695)	(0.582)	(0.541)	(0.040)
URSS_b	0.587	0.594	0.276	0.812	0.975	0.129**
	(0.750)	(0.652)	(1.087)	(0.770)	(0.640)	(0.051)
URSS_c	0.580	0.354	0.021	0.911	0.230	0.068
	(0.618)	(0.749)	(1.185)	(0.639)	(0.535)	(0.072)
CEE	-1.370***	-1.318***	-1.574*	-1.227***	-1.326***	-0.017
	(0.367)	(0.334)	(0.761)	(0.419)	(0.340)	(0.028)
Ex-Yugoslavia	-1.036***	-0.723**	-0.960	-0.927**	-0.936**	-0.013
	(0.354)	(0.291)	(0.741)	(0.386)	(0.357)	(0.027)
East Mediterranean	-1.487***	-1.141*	-1.056	-1.349**	-1.284**	-0.051
	(0.508)	(0.637)	(0.698)	(0.531)	(0.601)	(0.039)
Oil	-0.004***	-0.002	-0.002	-0.004***	-0.003	-0.000
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.000)
Coal	-0.000***	-0.000	-0.000	-0.000***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Natural gas	0.002***	0.001	0.001	0.002***	0.002	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Rule of law	-0.303				-0.483**	-0.024
	(0.264)				(0.221)	(0.028)
School enrollment, primary		0.004	0.016		0.003	-0.000
		(0.011)	(0.030)		(0.010)	(0.001)
School enrollment, secondary			0.015			
			(0.036)			
Investment				0.789	-4.346	-0.186
				(1.622)	(2.971)	(0.171)
Constant	6.565***	7.011***	6.146**	6.478***	6.702***	0.115
	(0.872)	(0.736)	(2.272)	(0.891)	(0.940)	(0.080)
Observations	43	35	35	43	35	35
R-squared	0.892	0.876	0.878	0.888	0.904	0.808

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The Energy production technology gap (EPTG), our variable of interest, is calculated as the Euclidean distance

In Column 2, we introduce educational attainment variables taken from Barro and Lee (2010).[4] The idea here is that a better-educated and better-trained population could introduce new technologies (to the country, not to the world) and thus launch new activities and new products. We first introduce the percentage of primary schooling attained by the population and observe that this variable is not significant – and takes the opposite sign relative to our expectation. We then also include the percentage of secondary schooling attained by the population and find a similar result.

In Column 4, we introduce an investment variable, the share of gross capital formation in total output, taken from the Penn World Tables 8.0 (Feenstra et al., 2013). One would expect that a higher level of investment is key to export diversification, as it allows for the creation of new business and ventures that introduce new export products or activities. Note that the coefficient associated with this variable is not significant and has a positive sign, again the opposite of what one would expect.

In Table 5, Column 5, we introduce all of these variables simultaneously, and our results are similar to those of the previous estimate, although rule of law becomes significant at the 5 percent level.

The last column of Table 5 reports the results of a similar estimation to that in Column 5, but here we use the Herfindahl export diversification index as the dependent variable. This index of diversification is often used as an alternative measure of export diversification, and it is highly positively correlated with the Theil index but is not constructed in an identical manner. Unsurprisingly, the results in Column 6 display a similar effect to those in Column 5: the EPTG variables have a quadratic effect on the Herfindahl index.

Table A2 in the appendix presents the results for the specifications that are identical to those used in Table 6, except that we now use the average of the export concentration index over the period 1996-2010 as the dependent variable (instead of the value of the Theil index for 1996). Regarding the covariates, we again use their initial value for the year 1996. We use this dependent variable to confirm that the results in Table 5 are not driven by particular characteristics of the year 1996. All of these results highlight that the negative impact of the Soviet legacy induced a lasting effect on export diversification over the period 1996-2010.

In Table 6[5], we check the robustness of our results to other potential threats, using the specification from Column 5 in Table 5 as a benchmark. In the first column, we remove Germany from the sample, as this country was divided into two countries prior to 1990 and was thus subject to influence from both Western countries and Soviet power. When we drop this country, our results remain unchanged. In Columns 2 to 4, we control for a series of geographical variables, as one could argue that our EPTG variable captures different geographical characteristics and that could affect export diversification. In Column 2, we introduce the distance to Moscow from each country's capital. In Column 3, we introduce a dummy variable for countries that are landlocked. In Column 4, we simultaneously introduce two geographical variables, the distance to major markets and the proportion of a country's

---

[4] Note that the Barro and Lee's (2010) dataset provides observations for every 5 years (1990, 1995, 2000, etc.). As a consequence, we use the value for 1995 in our estimation.

[5] Our results are robust to the use of the average Theil value as the dependent variable. These results are available upon request.

territory within 100 km of the coast (or a navigable river), obtained from Gallup et al. (1999). None of these geographical variables are significant. More important, the main results remain unchanged. In Column 5 we reconstruct the Theil variable but we do not account for exports of energy products, such as oil, coal and natural gas. The usual results for the EPTG variable remain stable. Finally, in Column 6, we conduct a falsification test as we calculate an EPTG variable replacing the Russian energy structure with its Swiss counterpart. The motivation is that this variable should have no effect on export diversification, and the choice of Switzerland is due to the historically neutral geopolitical position of this country. Indeed, as expected, there is no effect of the Swiss EPTG on export diversification.

In Table 7, we use our benchmark specification but use different dependent variables, either their values in 1996 or their average values, as we did for the Theil index. We alternate between the two variables presented in Section 2.1. Globally, the interest here is to show that our hypothesis holds for similar measures of the export basket structure. This implies that what we capture is not simply a spurious correlation between our EPTG and Theil variables. Note that all dependent variables in Table 7 have the opposite signs of the Theil index: this means that the higher the value of each variable, the more “positive” is the export structure of this country. Therefore, we expect that the coefficients for the EPTG measure should be the inverse of those in the previous tables – a positive value for the linear variable and a negative value for the quadratic measure (an inverted U-shaped relationship).

In Column 1, we use the international export market penetration, defined as above. The interest here is that diversification also involves bilateral trade flows. This variable captures each country’s share of all (actual plus potential) bilateral trade flows. We find that the EPTG, as a proxy for the Soviet legacy, has the same effect as in the previous tables. As further away from the Soviet Union, the greater the international export market penetration of a country is. This effect persists until a turning point, after which, the further a country is from the former Soviet Union, the weaker the country’s international export market penetration is. This pattern is identical to what we find in Column 2, where instead of the value in 1996, we use the average value of this variable for the period 1996–2010, to avoid potential measurement issues due to price fluctuations in the export basket.

Finally, in Columns 3 and 4 (using the average value for the same period as above), we estimate the relationship using the density variable. This variable captures how dense the export structure of a country is. A higher level of density means that the country produces and exports many goods that have the same capabilities as other products involved in international trade. Therefore, the country could diversify and modify its export basket. Our results in Columns 3 and 4 indicate that density has a significant effect on this variable, following the same u-shaped pattern as above.

Table 6 : OLS Results, Robustness checks, other potential threats

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Theil	Theil	Theil	Theil	Theil without energy exports	Theil
EPTG	-0.104***	-0.099***	-0.105**	-0.076**	-0.105***	
	(0.032)	(0.032)	(0.042)	(0.032)	(0.031)	
EPTG_sq	0.001***	0.001***	0.001**	0.001**	0.001***	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Swiss EPTG						0.026
						(0.018)
Swiss EPTG_sq						-0.000
						(0.000)
RGDP_pc	0.000	0.000	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc_sq	-0.000	-0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population	-0.006	-0.007	-0.006	-0.009	-0.005	-0.023**
	(0.011)	(0.007)	(0.007)	(0.006)	(0.006)	(0.009)
URSS_a	-1.009*	-0.621	-1.023	-1.230**	-1.358**	-1.660**
	(0.534)	(0.649)	(0.791)	(0.516)	(0.523)	(0.670)
URSS_b	0.973	1.122	0.942	0.217	1.106	-0.322
	(0.918)	(0.703)	(1.167)	(0.801)	(0.669)	(0.800)
URSS_c	0.229	0.185	0.197	-0.524	0.145	-1.007
	(0.731)	(0.598)	(1.127)	(1.532)	(0.578)	(0.815)
CEE	-1.327***	-1.118**	-1.345*	-1.450**	-1.407***	-1.788***
	(0.386)	(0.399)	(0.707)	(0.586)	(0.347)	(0.543)
Ex-Yugoslavia	-0.936**	-0.764**	-0.939**	-0.936*	-1.160***	-1.418**
	(0.366)	(0.332)	(0.390)	(0.493)	(0.395)	(0.586)
East Mediterranean	-1.284*	-0.988*	-1.295	-1.032**	-1.396**	0.025
	(0.684)	(0.508)	(0.807)	(0.370)	(0.648)	(0.811)
Oil	-0.003	-0.003*	-0.003	-0.002	-0.003	-0.001
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Coal	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Natural gas	0.002	0.001	0.002	0.001	0.001	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Rule of law	-0.483**	-0.436*	-0.482*	-0.347	-0.619***	-0.257
	(0.227)	(0.229)	(0.229)	(0.266)	(0.194)	(0.303)
School enrollment, primary	0.003	0.001	0.003	-0.004	-0.003	-0.018
	(0.010)	(0.010)	(0.010)	(0.007)	(0.009)	(0.011)
Investment	-4.345	-5.416	-4.398	-2.715	-4.073	-3.058
	(3.072)	(3.672)	(3.123)	(2.411)	(2.714)	(4.316)
Distance to Moscow		0.000				
		(0.000)				
Landlocked			0.015			
			(0.463)			
Distance to major ports				-0.000		
				(0.000)		
Distance to coast/river				0.000		
				(0.001)		
Constant	6.704***	6.208***	6.742***	7.172***	6.411***	5.263***
	(1.128)	(0.976)	(1.630)	(1.402)	(0.799)	(1.201)
Observations	34	35	35	33	35	35
R-squared	0.900	0.913	0.904	0.941	0.911	0.802

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The Energy production technology gap (EPTG), our variable of interest, is calculated as the euclidean distance.

Table 7 : OLS results using alternative left hand side variables				
VARIABLES	(1) International Export Market Penetration	(2) Average International Export Market Penetration	(3) Density	(4) Average Density
EPTG	0.001***	0.001***	0.011***	0.013***
	(0.000)	(0.000)	(0.003)	(0.003)
EPTG_sq	-0.000**	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_cp	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_cp_sq	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Population	0.000***	0.000***	0.002***	0.002***
	(0.000)	(0.000)	(0.001)	(0.001)
URSS_a	0.011**	0.006	0.117*	0.008
	(0.004)	(0.004)	(0.056)	(0.063)
URSS_b	0.021**	0.013**	0.049	-0.066
	(0.008)	(0.006)	(0.077)	(0.089)
URSS_c	0.018**	0.010	0.071	-0.094
	(0.007)	(0.006)	(0.075)	(0.089)
CEE	0.013***	0.009**	0.174***	0.067
	(0.004)	(0.004)	(0.054)	(0.063)
Ex-Yugoslavia	0.008**	0.005	0.119***	0.035
	(0.003)	(0.003)	(0.035)	(0.045)
East mediterranean	0.006	0.006	0.113*	0.088
	(0.004)	(0.004)	(0.059)	(0.059)
Oil	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Coal	-0.000**	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Natural gas	-0.000***	-0.000**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Rule of law	0.002	0.002	0.034	0.040
	(0.002)	(0.002)	(0.027)	(0.035)
School enrollment, primary	0.000	-0.000	-0.001	-0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Investment	0.015	0.019	0.430	0.513
	(0.019)	(0.017)	(0.349)	(0.393)
Constant	-0.039***	-0.030***	-0.296**	-0.194
	(0.010)	(0.007)	(0.120)	(0.146)
Observations	35	35	35	35
R-squared	0.979	0.974	0.912	0.892

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The Energy production technology gap (EPTG), our variable of interest, is calculated as the euclidean distance

3.2 Instrumentation strategy

The OLS results above reveal a stable pattern: a quadratic relationship between the EPTG and diversification. This finding is robust to the inclusion of different control variables, different dependent variables and spurious correlation. Note that we estimate our effect in a cross-sectional framework, and thus we cannot include country fixed effects. However, our results are also robust to the inclusion of different types of dummy variables that capture region-specific factors. Furthermore, as explained above, we exclude the problem of reverse causality, as the EPTG measure is exogenous to the countries' export baskets. Nevertheless, one could argue that the estimated coefficient for the EPTG variable captures a common, unobserved factor that determines both the EPTG and the diversification variables: the results would then be biased due to omitted variables. The most common solution to this issue would be to employ a standard two-stage least squares estimator to eliminate this problem. However, finding external instruments for our framework is quite hard. Hence, to estimate our equation, we appeal to a recently developed instrumentation technique for cases in which appropriate instruments are not available.

We apply the two-stage estimator proposed by Lewbel (2012) that uses heteroscedasticity for identification. This method permits the identification of structural parameters in regressions with endogenous explanatory variables in the absence of standard identifying information such as external instruments. Following Lewbel (2012), identification is then achieved by obtaining explanatory variables that are uncorrelated with the product of heteroskedastic errors, which he shows is a common feature of models in which error correlations are due to a common, unobserved factor. Lewbel (2012) develops in detail the moment conditions of a new method to identify parameters in models with an unobserved common factor. The identification method through heteroscedasticity is similar to that in works by Rigobon (2003) and Klein and Vella (2010), but he extends their methods[6]. Most important, Lewbel's (2012) method does not rely on exclusion restrictions as in a standard IV estimation, which is generally the most difficult requirement to satisfy for an instrument. Moreover, it allows for an IV estimation without external instruments. Therefore, the IVs are generated within the model, exploiting the exogenous control variables already included in the OLS estimates.

To demonstrate the intuition underlying the approach, we follow the notation of the estimated equation in the previous section (i.e., equation 1). We form a group of exogenous variables Z, where Z ∈ X. Note that X are the control variables in the estimated equation 1, and in our case Z = X, as we do not have external instruments. First, *EPTG* and *EPTG_sq* are regressed (separately) on the Z vector, which contains our X' variables in equation (1), followed by the retrieval of the residuals $\hat{\xi}$. These residuals are used to construct our standard instrumental variables in the second stage as $(Z_i - \bar{Z}_i) * \hat{\xi}_i$. The main assumption on which this IV strategy relies is the fact that the $\hat{\xi}_i$ are heteroskedastic. If this hypothesis is true (the Breusch-Pagan test is rejected), we can then estimate the second stage as a regular IV estimation.

The identification through heteroscedasticity proposed by Lewbel (2012) has been used in recent papers[7]. Emran and Hou (2013) apply it to household consumption in China, as they are interested in the effect of access to markets. They apply estimations methods in

[6] Lewbel (2012) discusses the difference between the estimator developed in that paper and previous methods.
[7] Lewbel (2012) illustrates the method using the example of the estimation of Engel curves.

addition to Lewbel's (2012). Most important, the results estimated from the latter's methodology are quite similar to those of a traditional IV estimation. Chowdury et al. (2014) focus on whether sectoral (production) diversification has an effect on the choice of exchange regime. They also employ OLS, standard IV and the identification through heteroscedasticity approach. We also highlight that their results using this methodology support the results they obtain using the two more common estimation methods.

Table 8 presents the results from the second stage of this IV estimation. Following Lewbel (2012), we use the Breusch-Pagan test for heteroskedasticity. Our results indicate that the null of a homoscedastic error is rejected for the two variables (*i.e.*, EPTG and its squared term) with a *p*-value of less than 0.01. In the second stage, we use the set of instruments described above to estimate the effect of EPTG and its squared term. Furthermore, each regression is accompanied by the p-value of the Hansen test for overidentification. These tests indicate that the instruments based on heteroskedasticity satisfy the exogeneity condition: the null of exogeneity of the set of Lewbel IVs cannot be rejected with a *p*-value equal to at least 0.23. The results in all columns of Table 8 show that proximity to the Soviet Union legacy has an effect on export diversification, irrespective of the left-hand side variable selected. Thus, the further from the Soviet Union, the higher the country's export diversification until a turning point, after which, the further a country is from the former Soviet Union, the lower is its diversification. The results using the two-stage Lewbel estimator confirm our main finding.

Table 8 : Lewbel Two-stage Estimator

VARIABLES	(1) Theil	(2) Average Theil	(3) International Export Market Penetration	(4) Average International Export Market Penetration	(5) Density	(6) Average Density
EPTG	-0.092***	-0.078**	0.001***	0.001***	0.010***	0.010***
	(0.032)	(0.038)	(0.000)	(0.000)	(0.003)	(0.003)
EPTG_sq	0.001***	0.001**	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	5.253***	5.135***	-0.002	-0.000	-0.004	-0.004
	(1.094)	(1.131)	(0.004)	(0.003)	(0.101)	(0.103)
Observations	35	35	35	35	35	35
R-squared	0.683	0.670	0.961	0.959	0.756	0.789
Hansen (p-value)	0.278	0.237	0.436	0.424	0.496	0.230

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The independent variables in the above models are: variables of interest (EPTG and its squared term) and the usual controls variables present in Table 5. The test of heteroskedasticity in the first stage of Lewbel two-stage estimator (2012) are as follows:
1st-Stage Regression (Dependent Variable) Breusch-Pagan/Cook-Weisberg test for heteroskedasticity
EPTG Chi square(1) = 8,82 *p*-value = 0.003
EPTG_sq Chi square(1) = 7,93 *p*-value = 0.005

Section 4: Conclusion

This paper contributes to the literature on the empirical determinants of export diversification by highlighting the effect of the Soviet legacy. By exploiting a quasi-natural experiment provided by the collapse of the USSR, we find that countries that were closer to the Moscow regime had a significantly less diversified export basket until a turning point, *ceteris paribus*.

To capture the effect of Moscow's influence, we measure the energy production technology gap (EPTG) with the Soviet Union. This strategy is based on considerable evidence indicating that energy production infrastructure served as a foreign policy tool of the Soviet regime to secure greater influence in the region. By using this approach, we avoid the endogeneity problem due to reverse causality in the empirical estimations that is common in the export diversification literature, as diversification could also potentially determine many of its explanatory variables. There is no reason that the export diversification of a country would influence the choice of energy sources made by another country.

Our main results indicate that the implantation of economic and political institutions similar to those advocated by the Communist regime served as an impediment to the product diversification of countries, as it hindered the development of the environment necessary for entrepreneurs to introduce new activities and new products. These results are robust to alternative specifications using different control variables. Moreover, we use a two-stage estimator proposed by Lewbel (2012), which relies on a heteroskedasticity based-identification strategy, which confirms that our main result is robust

References

Acemoglu, Daron, Johnson, Simon, Robinson, James, Thaicharoen, Yunyong, 2003. Institutional causes, macroeconomic symptoms: volatility, crises and growth. Journal of Monetary Economics, 50, 49-123.

Acemoglu, Daron, Zilibotti, Fabrizio, 1997. Was Prometheus Unbound by Chance? Risk, Diversification, and Growth. Journal of Political Economy, 105(4), 709–51.

Balassa, Bela, 1964. The Purchasing Power Parity Doctrine - A Reappraisal. Journal of Political Economy, 72(6), 584–596.

Balmaceda, Margarita M., 2008. Energy Dependency, Politics and Corruption in the Former Soviet Union: Russia's Power, Oligarchs' Profits, and Ukraine's Missing Energy Policy, 1995-2006. London: Routledge.

Barro, Robert, Lee, J.W., 2010. A New Data Set of Educational Attainment in the World, 1950-2010. National Bureau of Economic Research Working Papers 15902, National Bureau of Economic Research, Inc.

Ben Yishay, Ariel, Grosjean, Pauline, 2014. Initial Endowments and Economic Reform in 27 Post-Socialist Countries. Journal of Comparative Economics 42 (4), 892 – 906.

Berkowitz, Daniel, Hoekstra, Mark, Schoors, Koen, 2012. Does finance cause growth? Evidence from the origins of banking in Russia. National Bureau of Economic Research Working Papers 18139, National Bureau of Economic Research, Inc.

Brenton, Paul, Newfarmer, Richard, 2008. Watching more than the discovery channel. In: Newfarmer, Richard, Shaw, William, Walkenhorst, Peter. (Eds.), Breaking into new Markets. The World Bank: Washington D.C.

Cadot, Olivier, Carrère, Celine, Strauss-Kahn, Vanessa, 2011. Export Diversification: What’s behind the Hump? The Review of Economics and Statistics, 93(2), 590–605.

Cadot, Olivier, Carrère, Celine, Strauss-Kahn, Vanessa, 2013. Trade Diversification, income, and growth: What do we know? Journal of Economic Surveys, Volume 27, Issue 4, pages 790–812,

Cherp, Aleh, Jewell, Jessica, 2011. The three perspectives on energy security: intellectual history, disciplinary roots and the potential for integration. Current Opinion in Environmental Sustainability, 3:202–212

Chowdhury, Mohammad T. H., Battacharya, Prasad S., Mallick, Debdulal, Ulubasoglu, Mehmet A., 2014. An empirical inquiry into the role of sectoral diversification in exchange rate regime choice. European Economic Review, 67, 210-227.

De Benedictis, Luca, Gallegati, Marco, Tamberi, Massimo, 2009. Overall trade specialization and economic development: countries diversify. Review of World Economics, 145(1), 37–55.

Emran, M. Shahe, Hou, Zhaoyang, 2013. Access to markets and rural poverty: evidence from household consumption in China. Review of Economics and Statistics, 95(2), 682 – 697.

Ericson, Richard E., 1991. The Classical Soviet-Type Economy: Nature of the System and Implications for Reform. The Journal of Economic Perspectives, Vol. 5, No. 4, pp. 11-27

Feenstra, Robert C., Inklaar, Robert, Timmer, Marcel P., 2013. The Next Generation of the Penn World Table available for download at www.gddc.net/pwt

Gaulier, Guillaume, Zignago, Soledad, 2010. BACI: International Trade Database at the Product-Level. The 1994-2007 Version. Working Papers 2010-23, CEPII research center.

Gallup, John L., Sachs, Jeffrey D., Mellinger, Andrew D., 1999. Geography and Economic Development. International Regional Science Review 22, 2: 179-232.

Gleason, Gregory, 1992. The Federal Formula and the Collapse of the USSR. Publius: The Journal of Federalism Vol. 22, No. 3 pp. 141-16

Hidalgo, Cesar, Klinger, Bailey, Barabasi, Albert L., Hausmann, Ricardo, 2007. The Product Space Conditions the Development of Nations. Science, 317(5837), 482–487.

Hummels, David, Klenow, Peter J., 2005. The Variety and Quality of a Nation’s Exports. American Economic Review, 95(3), 704–723

Imbs, Jean. 2007. Growth and Volatility. Journal of Monetary Economics, 54(7), 1848-1862.

Imbs, Jean, Wacziarg, Romain. 2003. Stages of Diversification. American Economic Review, 93(1), 63–86.

International Energy Agency (IAE), 1998. World Energy Outlook. IAE Publications: Paris

International Energy Agency (IAE), 1999. World Energy Outlook IAE Publications: Paris

Kali, Raja, Reyes, Javier, McGee, Joshua, Shirrell, Stuart, 2013. Growth Networks. Journal of Development Economics 101: 216–27.

Kaufmann, Daniel, Kraay, Aart, Mastruzzi, Massimo, 2010. The worldwide governance indicators: methodology and analytical issues. Policy Research Working Paper Series 5430, The World Bank.

Klein, Roger, Vella, Francis, 2010. Estimating a class of triangular simultaneous equations models without exclusion restrictions. Journal of Econometrics 154, 154 – 164.

Koren, Miklos, Tenreyro, Silvana, 2007. Volatility and Development. The Quarterly Journal of Economics, 122(1), 243–287.

Lewbel, Arthur, 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. Journal of Business & Economics Statistics 30:1, 67-80.

Libman, Alexander, Obydenkova, Anastassia, 2013. Communism or communists? Soviet legacies and corruption in transition economies. Economics Letters 119, 101-103.

Omgba, Luc D., 2014. Institutional foundations of export diversification patterns in oil-producing countries, Journal of Comparative Economics 42 (4), 1052 – 1064.

Parteka, Aleksandra, Tamberi, Massimo, 2013. What Determines Export Diversification in the Development Process? Empirical Assessment. The World Economy, Volume 36, Issue 6, pages 807-826.

Ramey, Garey, Ramey, Valerie A., 1995. Cross-Country Evidence on the Link Between Volatility and Growth. American Economic Review, 85(5), 1138–51.

Rigobon, Roberto, 2003. Identification through Heteroskedascity. The Review of Economics and Statistics, 85 (4), 777 - 792.

Shurchkov, Olga, 2012. New elites and their influence on entrepreneurial activity in Russia. Journal of Comparative Economics 40, 240-255.

Stanley, Denise L., Bunnag, Sirima, 2001. A New Look at the Benefits of Diversification: Lessons from Central America. Applied Economics, 33(11), 1369–83.

Van der Marel, Erick, Dreyer, Iana, 2014. Beyond Dutch disease: when deteriorating rule of law affects Russian trade with and investment from advanced economies. Economics of Transition 2014 1–24 DOI: 10.1111/ecot.12039

Appendix A

Table A1 : Initial results: OLS using average value for the dependent variable (1996-2010)						
VARIABLES	(1) Average Theil	(2) Average Theil	(3) Average Theil	(4) Average Theil	(5) Average Theil	(6) Imports Concentration (Average)
EPTG		-0.049*	-0.055*	-0.059**	-0.062*	-0.024
		(0.027)	(0.032)	(0.028)	(0.033)	(0.032)
EPTG_sq		0.000**	0.001*	0.001**	0.001**	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc	-0.000***	-0.000*	-0.000***	-0.000***	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc_sq	0.000***	0.000**	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population	-0.016***	-0.015***	-0.017***	-0.016***	-0.005	-0.002
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.006)
URSS_t_t		0.708*	-0.825			
		(0.377)	(0.651)			
URSS_a				-0.746	-0.503	1.271
				(0.440)	(0.511)	(0.797)
URSS_b				0.359	1.313	0.495
				(0.484)	(0.896)	(0.773)
URSS_c				0.761	1.565**	0.601
				(0.463)	(0.585)	(0.761)
CEE			-1.451***	-0.999**	-0.847*	0.170
			(0.504)	(0.391)	(0.437)	(0.510)
Ex-Yugoslavia			-1.194**	-0.711*	-0.607	-0.075
			(0.515)	(0.363)	(0.413)	(0.619)
East Mediterranean			-0.831	-0.550	-0.919	0.854
			(0.512)	(0.451)	(0.553)	(0.768)
Natural Resources		0.055***	0.055***	0.039***		
		(0.010)	(0.009)	(0.008)		
Oil					-0.002***	-0.000
					(0.001)	(0.001)
Coal					-0.000	-0.000
					(0.000)	(0.000)
Natural gas					0.001***	-0.000
					(0.000)	(0.000)
Constant	5.117***	4.772***	6.775***	6.075***	5.696***	3.528***
	(0.364)	(0.886)	(1.236)	(0.933)	(1.066)	(1.051)
Observations	43	42	42	42	43	43
R-squared	0.510	0.761	0.814	0.887	0.869	0.594

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The Energy production technology gap (EPTG), our variable of interest, is calculated as the euclidean distance

Table A2 : OLS Results : Robustness using average value for the dependent variable (1996-2010)

VARIABLES	(1) Average Theil	(2) Average Theil	(3) Average Theil	(4) Average Theil	(5) Average Theil	(6) Average Herfindahl
EPTG	-0.072**	-0.104**	-0.104**	-0.065*	-0.110***	-0.004**
	(0.030)	(0.038)	(0.036)	(0.032)	(0.030)	(0.001)
EPTG_sq	0.001**	0.001**	0.001**	0.001**	0.001***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc	-0.000	-0.000	-0.000	-0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
RGDP_pc_sq	0.000	0.000	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population	-0.008*	-0.005	-0.005	-0.006	-0.005	0.000
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.000)
URSS_a	-0.540	-0.235	-0.224	-0.398	-0.218	-0.001
	(0.482)	(0.601)	(0.725)	(0.593)	(0.509)	(0.022)
URSS_b	1.079	1.149	1.170	1.386	1.569**	0.088**
	(0.901)	(0.758)	(1.081)	(0.926)	(0.736)	(0.031)
URSS_c	1.232*	2.181**	2.203*	1.684**	2.036***	0.178***
	(0.601)	(0.795)	(1.149)	(0.663)	(0.636)	(0.044)
CEE	-0.965**	-0.578	-0.562	-0.768	-0.594	-0.011
	(0.461)	(0.402)	(0.719)	(0.503)	(0.411)	(0.021)
Ex-Yugoslavia	-0.727*	-0.287	-0.271	-0.579	-0.527	-0.014
	(0.412)	(0.344)	(0.704)	(0.433)	(0.338)	(0.019)
East Mediterranean	-1.067*	-0.971	-0.976	-0.879	-1.129*	-0.042
	(0.539)	(0.679)	(0.737)	(0.607)	(0.600)	(0.024)
Oil	-0.002***	-0.002	-0.002	-0.002*	-0.004*	-0.000*
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.000)
Coal	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Natural gas	0.001***	0.001	0.001	0.001*	0.002	0.000*
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Rule of law	-0.401				-0.528**	-0.029
	(0.255)				(0.219)	(0.020)
School enrollment, primary		0.007	0.007		0.007	0.000
		(0.011)	(0.028)		(0.010)	(0.000)
School enrollment, secondary			-0.001			
			(0.032)			
Investment				1.176	-5.161	-0.245
				(1.665)	(3.363)	(0.141)
Constant	5.727***	6.279***	6.334**	5.603***	5.975***	0.087
	(0.989)	(1.152)	(2.307)	(1.140)	(1.124)	(0.051)
Observations	43	35	35	43	35	35
R-squared	0.877	0.881	0.881	0.871	0.914	0.940

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The Energy production technology gap (EPTG), our variable of interest, is calculated as the euclidean distance

APPENDIX B

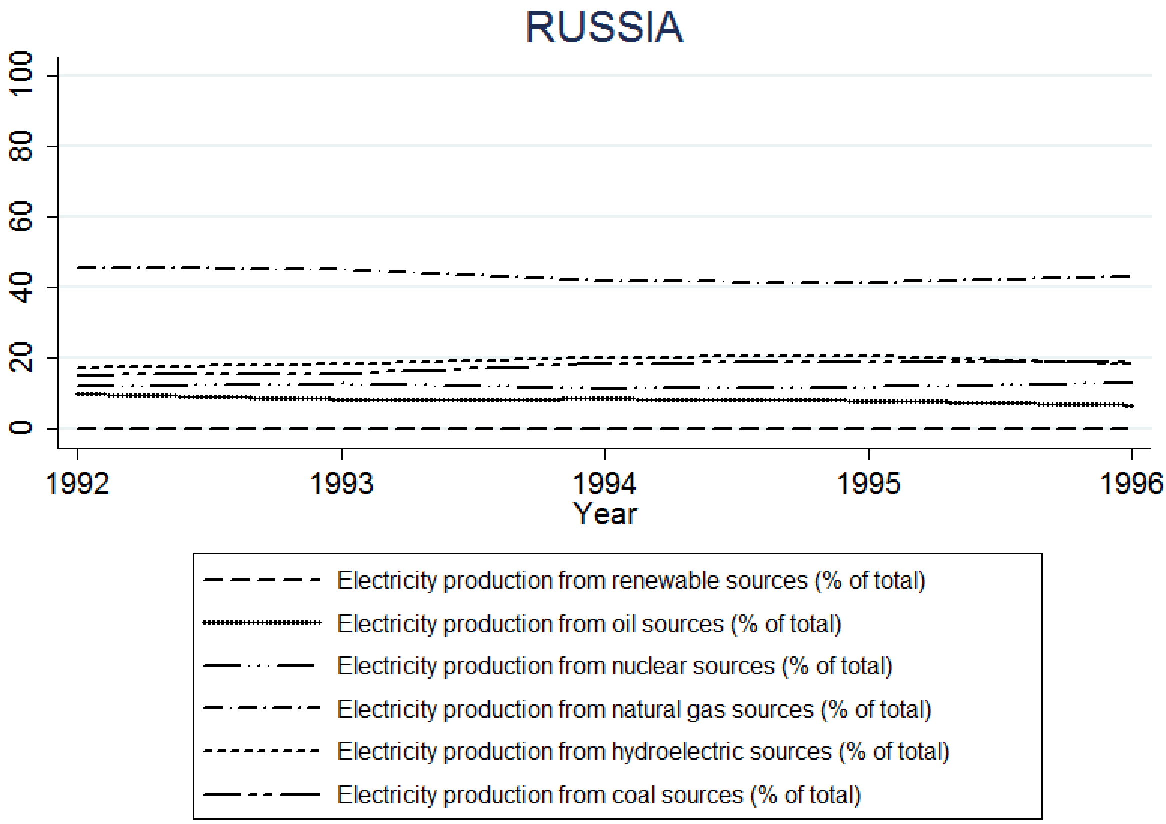

Figure 2, Evolution of the different shares of inputs used for the electricity production in Russia (1992 - 1996).

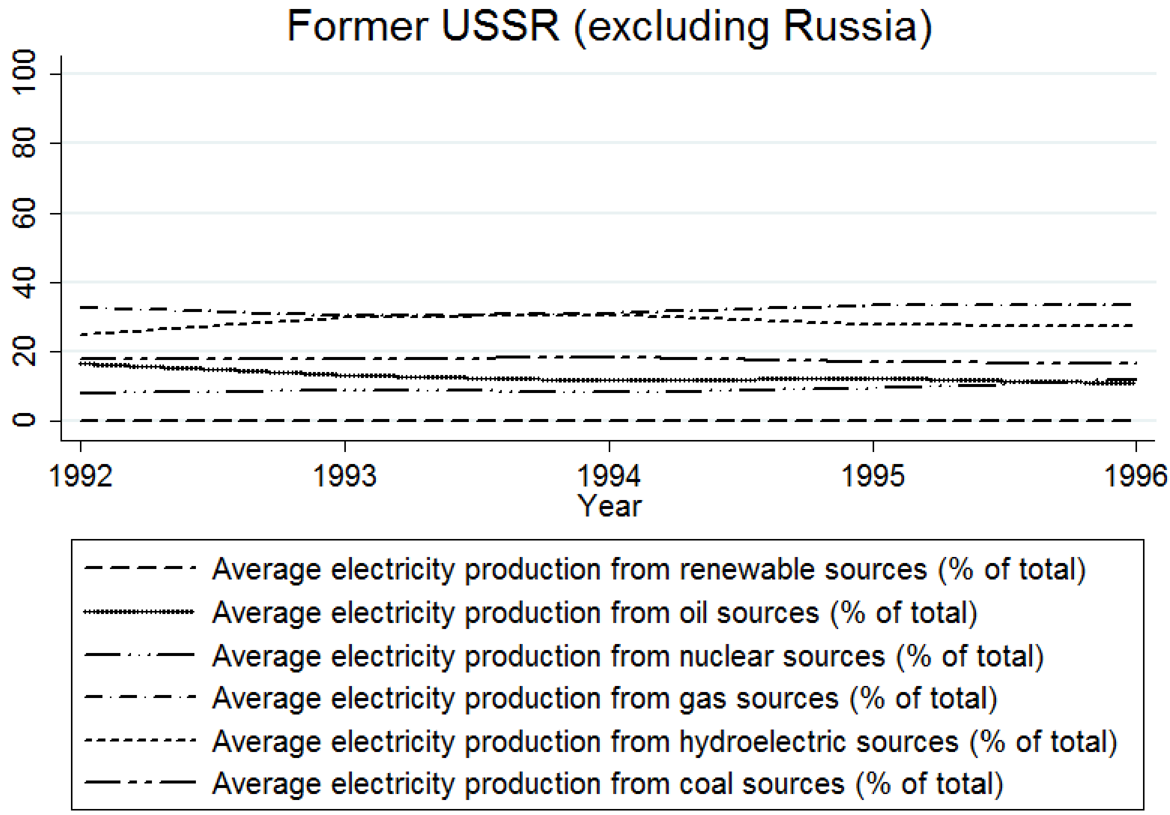

Figure 3, Evolution of the different shares of inputs used for the electricity production of former URSSS (1992 - 1996).

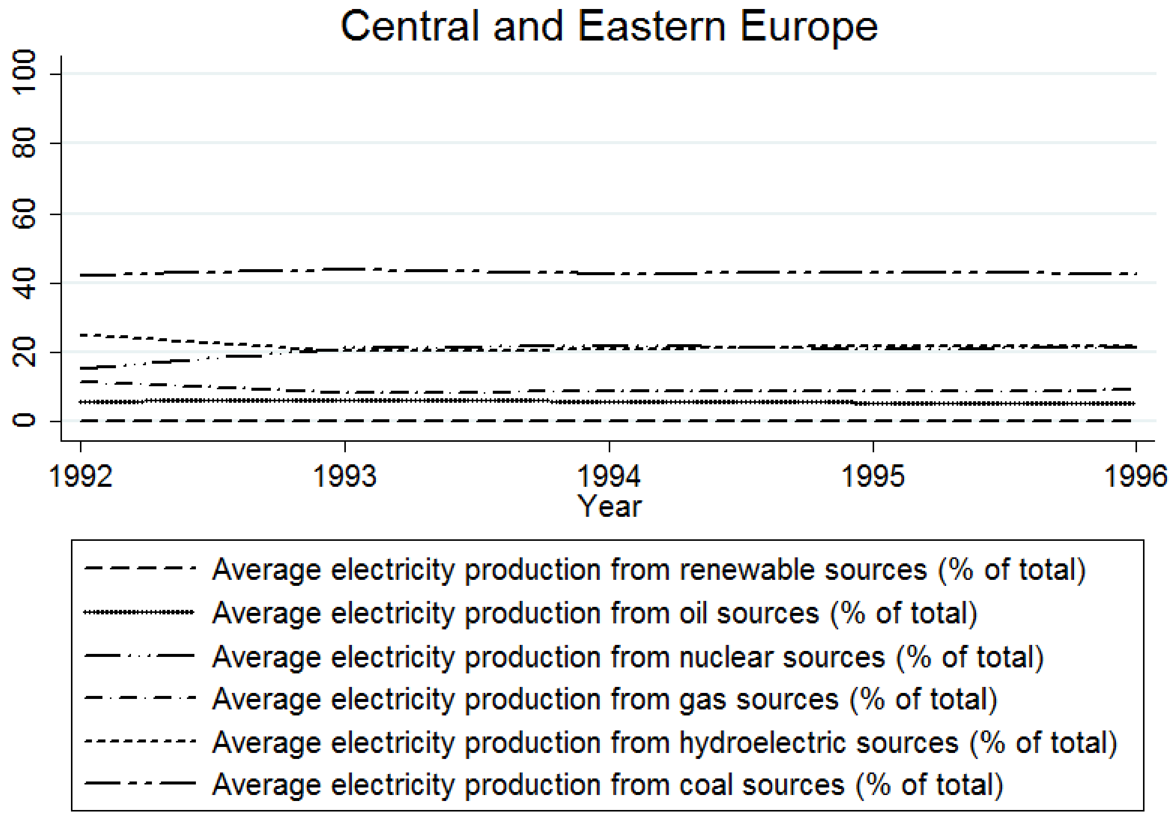

Figure 4, Evolution of the different shares of inputs used for the electricity production of Central and Eastern Europe (1992 - 1996).

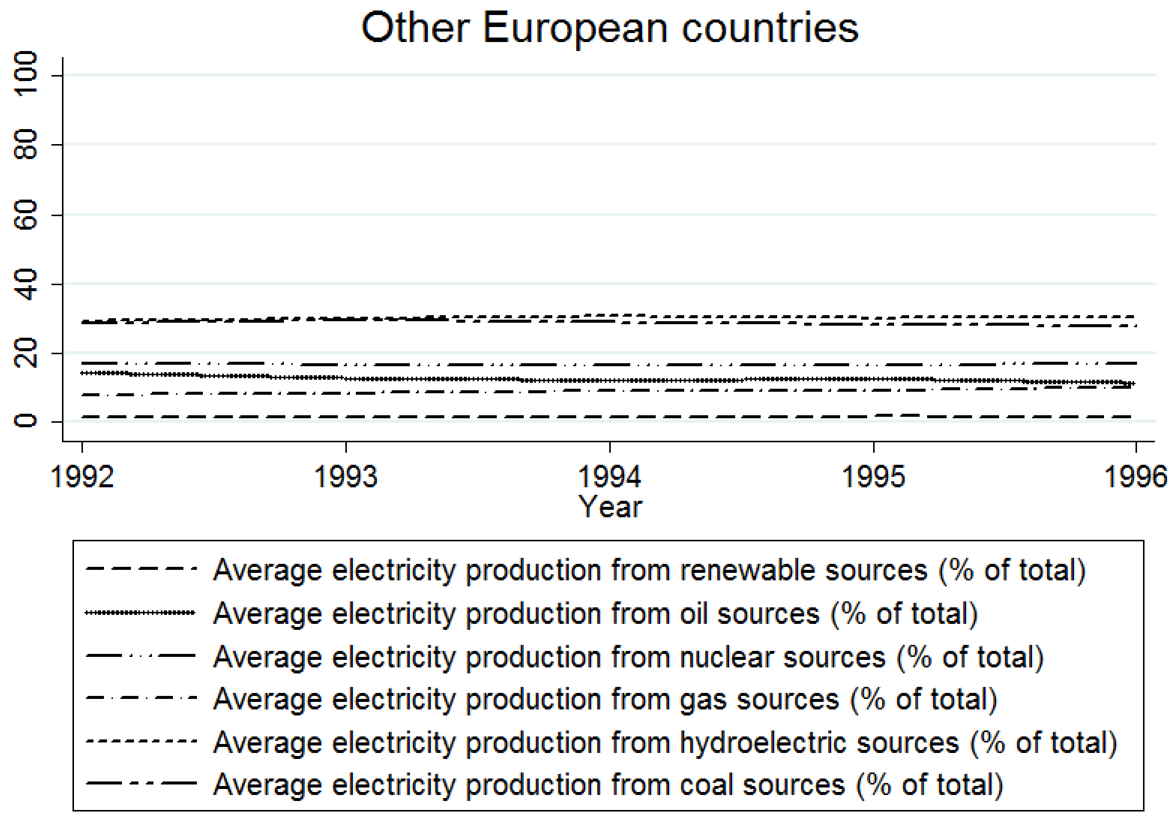

Figure 5, Evolution of the different shares of inputs used for the electricity production of other European country (1992 - 1996).

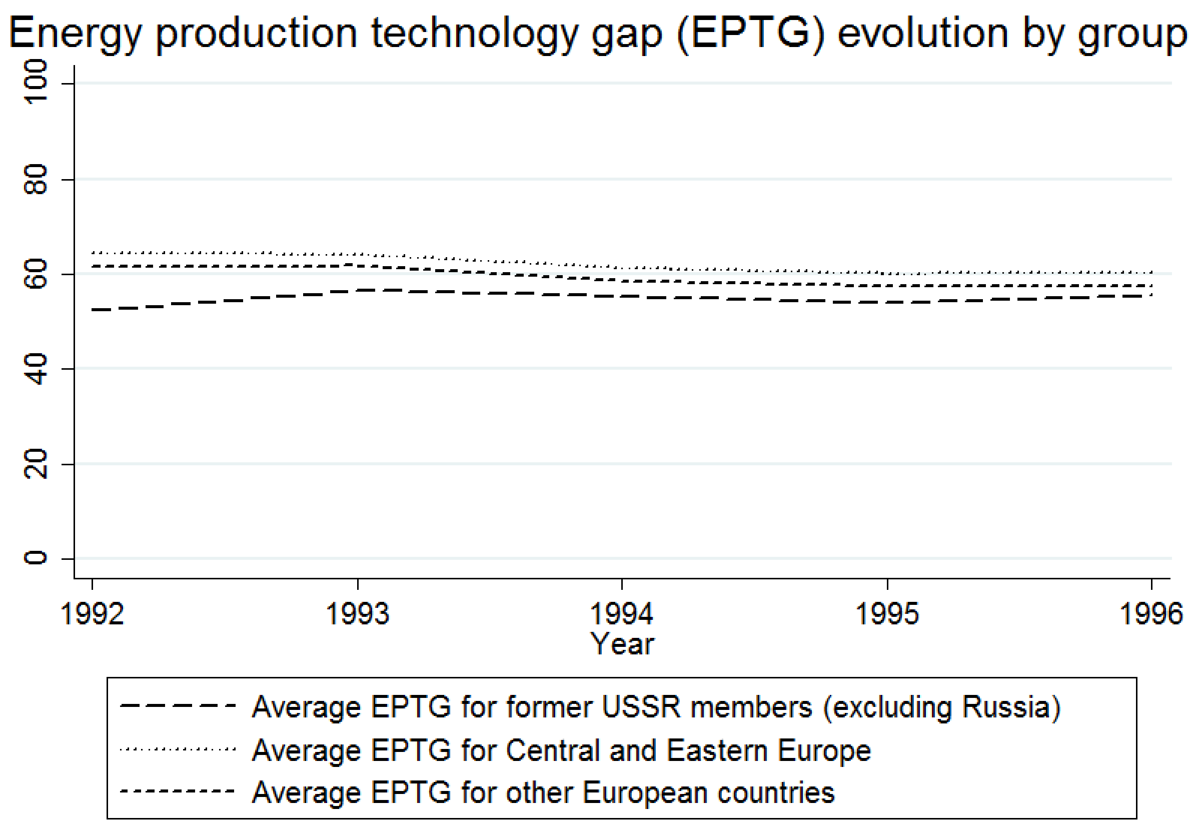

Figure 6, evolution of the energy production technology gap (EPTG) between Russia and the three different groups of countries (1992 – 1996)

Table B1 : Sample								
COUNTRIES IN THE SAMPLE	URSS	URSS_a	URSS_b	URSS_c	CEE	Ex-Yougoslavia	East-Mediterrean	Western Europe
Albania	0	0	0	0	1	0	0	0
Armenia	1	0	1	0	0	0	0	0
Austria	0	0	0	0	0	0	0	1
Azerbaijan	1	0	1	0	0	0	0	0
Belarus	1	1	0	0	0	0	0	0
Bosnia and Herzegovina	0	0	0	0	0	1	0	0
Bulgaria	0	0	0	0	1	0	0	0
Croatia	0	0	0	0	0	1	0	0
Cyprus	0	0	0	0	0	0	1	0
CzechRepublic	0	0	0	0	1	0	0	0
Denmark	0	0	0	0	0	0	0	1
Estonia	1	1	0	0	0	0	0	0
Finland	0	0	0	0	0	0	0	1
France	0	0	0	0	0	0	0	1
Georgia	1	0	1	0	0	0	0	0
Germany	0	0	0	0	0	0	0	1
Greece	0	0	0	0	0	0	0	1
Hungary	0	0	0	0	1	0	0	0
Iceland	0	0	0	0	0	0	0	1
Ireland	0	0	0	0	0	0	0	1
Italy	0	0	0	0	0	0	0	1
Kazakhstan	1	0	0	1	0	0	0	0
KyrgyzRepublic	1	0	0	1	0	0	0	0
Latvia	1	1	0	0	0	0	0	0
Lithuania	1	1	0	0	0	0	0	0
Macedonia. FYR	0	0	0	0	0	1	0	0
Moldova	1	1	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0	1
Norway	0	0	0	0	0	0	0	1
Poland	0	0	0	0	1	0	0	0
Portugal	0	0	0	0	0	0	0	1
Romania	0	0	0	0	1	0	0	0
SlovakRepublic	0	0	0	0	1	0	0	0
Slovenia	0	0	0	0	0	1	0	0
Spain	0	0	0	0	0	0	0	1
Sweden	0	0	0	0	0	0	0	1
Switzerland	0	0	0	0	0	0	0	1
Tajikistan	1	0	0	1	0	0	0	0
Turkey	0	0	0	0	0	0	1	0
Turkmenistan	1	0	0	1	0	0	0	0
Ukraine	1	1	0	0	0	0	0	0
United Kingdom	0	0	0	0	0	0	0	1
Uzbekistan	1	0	0	1	0	0	0	0

Table B2 : Data sources

Variable	Source
Energy production technology gap (EPTG)	WDI (World Bank)
RGDP_pc	PWT 8.0
Population	WDI (World Bank)
Natural Resources	WDI (World Bank)
School Enrollment, primary	Barro & Lee (2010)
School Enrollment, secondary	Barro & Lee (2010)
Investment	PWT 8.0
Rule of Law	WGI (World Bank)
Oil	US energy administration
Coal	US energy administration
Natural Gas	US energy administration
Distance to Moscow	www.distance.to
Air distance to major ports	Gallup et al. (1999)
Distance to coast or river	Gallup et al. (1999)