Economic complexity and jobs: an empirical analysis

Adam, Antonis and Garas, Antonios and Lapatinas, Athanasios

February 2019

Economic complexity and jobs: an empirical analysis

Antonis Adam[1], **Antonios Garas**[2],
and Athanasios Lapatinas[1,3,*]

[1]Department of Economics, University of Ioannina, P.O. Box 1186, 45110 Ioannina, Greece
[2]ETH Zürich, Chair of Systems Design, Weinbergstrasse 56/58, 8092 Zürich, Switzerland
[3]European Commission, Joint Research Centre (DG-JRC),
Via E. Fermi 2749, TP 361, Ispra (VA), I-21027, Italy

Abstract

This paper analyses the impact of economic complexity on the labour market using annual data on OECD countries for the period 1985-2008 and averaged data over the period 1990-2010 for 74 developed and developing countries with a large number of controls. We show that moving to higher levels of economic sophistication leads to less unemployment and more employment, showing that economic complexity does not induce job loss. Our findings remain robust across alternative econometric specifications. Furthermore, we place the spotlight on the link between products' embodied knowledge (sophistication) and labour market outcomes at the micro-level. We build a product-level index that attaches a product to the average level of unemployment (or employment) in the countries that export it. With this index, we illustrate how the development of sophisticated products is associated with changes in the labour market and show that the economic sophistication of an economy captures information about the economy's job creation and destruction.

Keywords: Economic Complexity, Product Sophistication, Unemployment, Employment

*Corresponding author (athanasios.lapatinas@ec.europa.eu). The final revisions of both text and empirical strategy of the article have been conducted when Athanasios Lapatinas took service at the European Commission, Joint Research Centre. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use, which might be made of this publication.

1 Introduction

Progressing to a more complex economy by developing new sophisticated products is a process of creative destruction that directly affects the labour market by creating and destroying jobs. Although it is easy to highlight positive and negative effects in particular cases, it is much more difficult to analyze and measure the overall outcome of the economic complexity advancements throughout the economy. During the last 20 years, new sophisticated products have radically transformed sectors/industries, leading to a destruction of obsolete jobs and, in parallel, to the creation of new ones [32]. On the one hand, product sophistication can displace labour by reducing or eliminating the demand for particular goods and/or services. In addition, it can reduce employment within particular occupations, for example, through the introduction of machines and robots. On the other hand, the development of new sophisticated products can have positive effects, by creating new jobs, as the recent boom in the information and communication technology (ICT) industries manifests. The question then is: what is the net impact of the transformation of the productive structure towards more sophisticated industries on employment and unemployment?

A number of recent contributions have introduced the measure of economic complexity in order to explain structural transformations and economic growth as a process of information development and of learning how to produce and export more complex products [1, 5, 24–27, 33, 37–40, 49, 55]. It is showed that more developed countries produce more diversified and complex products, while less developed countries produce fewer and simpler products. This strand of the literature shows that the development path of a country lies in its capacity to accumulate the 'knowledge' part of the Solow residual. Knowledge is required to produce varied and more sophisticated goods; moreover, embedded in countries' productive structures, it also explains the differences in their economic performance [27, 49, 52]. The relevant question then becomes: how much knowledge is there in an economy?

In recent years, the search for an answer to this question has given rise to the elaborate metric called Economic Complexity Index (ECI), which quantifies the amount of knowledge/know-how materialized in the country's production. To elucidate this, consider producing and exporting an electronic product, like computer hardware. Computer hardware is a product that requires specific ICT and physical capital inputs, specific knowledge and cognitive skills, e.g information technology (IT) skills. This implies that the observation that a country is exporting computer hardware gives us information about the presence of specific economic capabilities in its economy [5]. Using the 'bucket of Legos' metaphor of Hidalgo and Hausmann [39] and assuming that Lego-buckets represent countries and each Lego-piece in the bucket represents an economic capability, we can infer the shape (or other properties) of the Lego-pieces contained in the different buckets by looking only at the Lego-models developed with those pieces. In same way, the network of internationally traded products between countries signals the underlying capabilities of a given country that enables the production patterns, structures and links evident in this network.

The ECI captures the reflection of the economy's capabilities in the goods produced and exported by quantifying the network representation of the relatedness and proximity between products traded internationally [41]. When a country produces a good that is located in the core of the product space, many other related goods can also be produced with the given capabilities. But for the goods lying in the network's periphery this does not hold because they require different production capabilities. The core is formed, for example, by metal products, electronics, machinery and chemicals while the periphery is populated by product-categories like fishing, animal agriculture, cereal agriculture, garments and textiles.[1] The ECI methodology assigns lower values to countries that export products located at the periphery of the product space, i.e. products that require low capabilities, and higher values to countries that export commodities located in the center of the product space [39, 40]. This means that if the bipartite network connecting countries to the products they export is the result of a larger tripartite network connecting countries and products to the capabilities embodied in their production, then becoming a sophisticated (complex) economy is the process of moving from the periphery to the core of the international trade network of products.

In the relevant literature, the ECI has been validated as a highly predictive measure of future economic growth [36, 39]. Moreover, Hartmann *et al.* [35] have recently shown that economic sophistication is also associated with the quality of economic institutions and the average level of income inequality. The authors also point out that structural transformations and industrialization played a major role in the rise of a new middle class by creating new jobs and learning opportunities for workers.

On the other hand, one of the most notable stylised facts of the last decades is deindustrialisation, i.e. the decline in manufacturing employment in the industrialised world. Rodrik [49] documents a significant deindustrialisation trend in recent years that goes beyond the advanced, post-industrial economies. Buera and Kaboski [23], Matsuyama [45], Nickell *et al.* [46], Rowthorn and Ramaswamy [51] argue that deindustrialisation is explained by 'the relative productivity hypothesis': faster growth in manufacturing productivity leads to relative price changes and shifts in the economy's productive structure [19]. Other recent works show that manufacturing employment declines as a result of globalization and the strengthening of the manufacturing sectors of developing economies [11, 12, 47]. Advanced economies seem to have lost considerable employment because of globalization and labour-saving technological progress.

Recent advancements in artificial intelligence (AI), machine learning, robotics and related ICT technologies have revived the debate and concerns about 'technological unemployment'. The

[1]See Figure 1 in Hidalgo *et al.* [40] for the network representation of the product space for 775 SITC-4 product classes exported in the 1998-2000 period.

hypothesis of skill-biased technological change [13, 14, 16, 18, 21, 22, 29, 42, 43] has been resurrected: it being pointed out that the introduction of sophisticated production methods reduced the demand for low-skilled workers and increased the demand for medium- and high- skilled ones. In addition, there is evidence of a labour supply shift from middle-income manufacturing to low-income service occupations, because the latter has a lower probability of being replaced by machines [15, 30, 34]. At the same time, because of technology's rapid growth, the prices of sophisticated (ICT-related) products are falling, which makes the high-skills-related occupations relatively productive. This leads to higher demand for skilled labour and to an increase in occupations involving cognitive tasks [2, 43].

The above studies show a clear relationship between economic sophistication and the labour market. Here, we contribute to this literature and to the strand of works on economic complexity by documenting a robust effect of a country's level of economic complexity on its labour market outcomes. We find that higher sophistication of exported products is associated with lower unemployment. This result is found to be consistent across various specifications, and in both a panel setting, in which we examine a limited sample of OECD countries for the period 1985-2008 and a cross-sectional setting, in which we use a global sample of 74 developed and developing countries taking averages over the period 1990-2010.

Furthermore, we have developed a product-level index that allows the classification of products according to the level of unemployment (or employment) that they are associated with [35]. Using this index, we illustrate how the development of sophisticated products is associated with changes in the labour market. Our results suggest that countries' labour markets are conditioned by their 'product space' and hence by the level of sophistication embodied in the production of exported goods. Therefore, our index might be a promising policy tool that could be used to estimate the changes in labour market outcomes (unemployment and employment rates) we would expect if a country were to modify its product mix by adding or removing a product.

The remainder of the paper is structured as follows. Section 2 discusses the data. Section 3 presents the methods used in the paper. Section 4 describes the econometric analysis for studying the impact of economic complexity on labour market outcomes, discusses the control variables and the instruments of ECI included in the model and presents and discusses the results. Section 5 introduces two indexes of the unemployment and employment rates expected for the producers and exporters of 773 different products in the Standard Industrial Trade Classification at the four-digit level (SITC-4 Rev.2). Using these indexes, we put the spotlight on the links between products' sophistication and the unemployment and employment rates, at the micro-level. We illustrate that the development of more complex products is associated with lower unemployment and higher employment rates. Finally, in Section 6, we draw our conclusions.

2 Data

Data on labour market data are taken from the International Labour Organization (ILO) as reported in the World Bank's World Development Indicators database. Even though our main focus is on the overall unemployment and employment rates, in the econometric analysis we also experiment with the associated variables for subgroups of the population, i.e. young people and, men and women. To ensure comparability of our results in all instances, we use the same data source, i.e. the ILO national estimates.

To study the effect of *ECI* on the various labour market outcomes we use two different datasets. The first one includes OECD countries only, while the second includes a total of 74 countries both developed and developing. The OECD sample provides a more reliable dataset regarding the set of control variables (which is discussed in 4.1 and presented in Table 1) and its availability over a longer period of time. Specifically, given the availability of controls, the OECD sample covers the period 1985-2008 and includes the following countries: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Korea Rep., Latvia, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Rep., Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

In contrast, for non-OECD countries we have a limited set of control variables. In order to maximize the number of countries included in the sample we consider a simple cross-section, taking averages for the period 1990-2010 to the detriment of having temporal variations but for much fewer countries (mainly for the OECD countries as above). In other words, the benefit of having more countries in the sample and available controls is counterbalanced by the absence of time variation. However, we feel that this additional analysis gives valuable insights and further robustification of our results.[2]

We also use freely available international trade data from the Massachusetts Institute of Technology (MIT)'s Observatory of Economic Complexity (`http://atlas.media.mit.edu`). We chose the SITC-4 rev.2 dataset, which provides the longest time series, combining information from a dataset compiled by Feenstra *et al.* [31] for the years 1962-2000 and the UN Comtrade dataset from 2001 to 2008 (`https://comtrade.un.org`), and details about the products exported by every country.

We measure economic complexity using the improved *ECI* (*ECI*+). *ECI*+ measures the diversity and sophistication of a country's export structure corrected by how difficult it is to export each product. It combines information on the diversity of a country, i.e. the number of products it exports, and the ubiquity of its products, i.e. the number of countries that export these

[2]Data definitions and summary statistics for this dataset are given in Table 5

products [39]. *ECI+* is estimated from data connecting countries to the products they export and is freely available at MIT's Observatory of Economic Complexity. The index is calculated by applying the methodology described in Albeaik *et al.* [5] to the international trade data from the MIT's Observatory of Economic Complexity (a brief description of this methodology is discussed in Section 3). Albeaik *et al.* [5] show that *ECI+* outperforms the original *ECI* in its ability to predict economic growth and in the consistency of its estimators across different econometric specifications. *ECI+* captures information about an economy's level of development that is different from what is captured by, for example, gross domestic product (GDP) growth or GDP per capita. *ECI+* incorporates the idea that institutions, knowledge and technology are prerequisites for economic growth but, in contrast to other indexes of growth, *ECI+* is measured with simple linear algebra techniques that determine the knowledge intensity of economies endogenously – from the countries' export data [5]. In a very recent working paper, Albeaik *et al.* [4] show that the definition of *ECI+* is equivalent to the Fitness Complexity metric proposed by Tacchella *et al.* [54].

3 Methods

To calculate the improved measure of economic complexity (*ECI+*) used in this work, we rely on the methodology described in Albeaik *et al.* [5]. In short, let us assume that we have trade information for l number of countries and k products. We can calculate the total exports of a country corrected by how difficult it is to export each product using

$$\mathrm{X_c^1} = \sum_{\mathrm{p}} \frac{\mathrm{X_{cp}}}{\sum_{\mathrm{c}} \frac{\mathrm{X_{cp}}}{\mathrm{X_c^0}}}, \tag{1}$$

where $X_c^0 = \sum_p X_{cp}$ is the total exports of country c and $1/\sum_c \frac{X_{cp}}{X_c^0}$ measures how difficult it is for country c to export product p.

We then use this corrected value of total exports (equation 1) to calculate the second order correction:

$$\mathrm{X_c^2} = \sum_{\mathrm{p}} \frac{\mathrm{X_{cp}}}{\sum_{\mathrm{c}} \frac{\mathrm{X_{cp}}}{\mathrm{X_c^1}}}, \tag{2}$$

where X_c^2 represents again the proportion that a product represents of the average country.

Iterating this to the limit:

$$\mathrm{X_c^N} = \sum_{\mathrm{p}} \frac{\mathrm{X_{cp}}}{\sum_{\mathrm{c}} \frac{\mathrm{X_{cp}}}{\mathrm{X_c^{N-1}}}}, \tag{3}$$

and normalizing X_c at each iteration step by its geometric mean:

$$\mathrm{X_c^N} = \frac{\mathrm{X_c^N}}{(\prod_{\mathrm{c'}} \mathrm{X_{c'}^N})^{\frac{1}{[\mathrm{C}]}}} \tag{4}$$

where $[C]$ is the number of countries in the sample. We estimate $ECI+$ as the total exports of a country corrected by how difficult it is to export each product, minus the average proportion that the country represents in the total exports of a product (which accounts for the size of a country's export economy):

$$ECI_c^+ = \log(X_c^\infty) - \log(\sum_p \frac{X_{cp}}{X_p}). \tag{5}$$

Likewise, but putting the spotlight on products rather than on countries, the improved product complexity index ($PCI+$) is defined as the following iterative map:

$$\mathrm{X_p^N} = \sum_\mathrm{c} \frac{\mathrm{X_{cp}}}{\sum_\mathrm{p} \frac{\mathrm{X_{cp}}}{\mathrm{X_p^{N-1}}}} \tag{6}$$

with the initial condition $\mathrm{X_p^0} = \sum_\mathrm{c} \frac{\mathrm{X_{cp}}}{\mathrm{X_c^0}}$ being the average proportion of product p in country c. Again, normalizing at each step X_p by its geometric mean:

$$\mathrm{X_p^N} = \frac{\mathrm{X_p^N}}{(\prod_\mathrm{p'} \mathrm{X_{p'}^N})^{\frac{1}{[\mathrm{P}]}}} \tag{7}$$

where $[P]$ is the number of products in the sample, we define the product complexity index, corrected by how difficult it is to export each product,

$$PCI_p^+ = \log(X_p) - \log(X_p^\infty) \tag{8}$$

where X_p is total world trade of product p.

To summarize, $ECI+$ and $PCI+$ denote, respectively, the total exports of a country, corrected by how difficult it is to export each product, and the total trade in a product, corrected by how easy it is to export that product [5]. For simplicity of notation, we will hereafter call these measures ECI and PCI respectively.

4 Regression analysis

We study the effect of economic complexity, measured by the ECI, on various labour market outcomes, using the datasets described in Section 2.

According to Hausmann *et al.* [38] higher economic complexity is associated with higher productivity. The ECI ranks traded goods in terms of their implied productivity. This signals an important source of endogeneity in the relationship considered and an obvious problem of reverse causality. Furthermore, given that ECI is an alternative measure of structural transformations, i.e., the reallocation of factors of production from traditional to modern activities, its simultaneity with labour market outcomes cannot be neglected. In order to mitigate the endogeneity of the independent variables, we follow a fixed effects, two-stage least squares/instrumental variables (FE 2SLS/IV) strategy for both datasets (OECD panel sample and cross-section world sample).

We regress the baseline specification described by the following equation:

$$\mathrm{y}_{i,t} = \alpha_0 + \beta_1 ECI_{i,t} + \beta_k controls_{i,t} + \gamma_i + \delta_t + u_i. \tag{9}$$

Here, labour market outcomes for country i in period t are expressed as a function of the ECI, a set of control variables, time δ_t and country γ_i fixed effects, and a stochastic term u_i.[3] The main dependent variable in all regressions is the overall unemployment rate. To examine the robustness of our results and to generalize our findings, we also replicate our analysis for the unemployment and employment rates, among young people and men and women separately.

4.1 OECD panel sample

To correctly specify our regression model we use two broad groups of control variables out of the full set, which is listed in Table 1. The first group includes macroeconomic controls, i.e. *Inflation*, *Imports* and *Output Gap*. The inflation rate controls for the standard Phillips curve relationship (e.g. Wyplosz [58]). The *Output Gap* controls for the business cycle, whereas imports as proportion of GDP (*Imports*) controls for the effect of international trade. The second group of control variables includes those that control for the effect of labour market institutions.[4] More specifically, we use the average tax wedge, denoted *Tax Wedge* (see Daveri and Tabellini [28]), and variables that control for the main characteristics of the wage bargaining system as in Aidt and Tzannatos [3] (namely *Union Density, Coverage, Centralization* and *Coordination*). Finally, the generosity of the unemployment benefits system is captured through the variable *Replacement* (see for example Lichter [44], Scarpetta [53]).[5].

Table 2 presents our main results. In all cases except column (6), we use lagged and differenced values of the main independent variable (ECI) for up to four years as instruments. Changes in the ECI over the previous four years are likely to have a direct impact on the level of ECI in the current year: if a country managed to improve its economic complexity over the past

[3]The cross-section model does not include country and time fixed effects.

[4]For variable definitions, data sources and summary statistics see Table 1.

[5]To examine the robustness of our results we also introduce a series of additional variables that capture the strictness of government regulation in the labour market. Specifically, we employ an index of employment protection legislation (EPL), an index that measures the strictness of regulation in the economy (*Regulation*), the proportion of public expenditure spent on active labour market programmes (as a percentage of government spending) and the variable *Min Wage*, which measures the generosity of the minimum wage scheme.

Table 1: Variable sources, definitions and summary statistics; OECD sample

Variable	**Definition**	**Source**	**Mean**	**Std. Dev.**
ECI	Economic Complexity Index	Observatory of Economic Complexity	1.02	4.15
Inflation	% change in annual Consumer Price Index	International Monetary Fund, International Financial Statistics	10.34	31.09
Imports	Imports of goods and services as % of GDP	OECD National Accounts	32.88	16.96
Output Gap	The difference between actual and potential real (GDP) as a per cent of potential real GDP.	IMF World Economic Outlook	-0	0.023
Union Density	Net Union Membership as a proportion of total number of wage and salary earners in employment	Visser, 2015	35.58	19.27
Centralization	Level at which bargaining takes place, higher values indicate more centralized level of bargaining	Visser [56]	2.71	1.48
Coordination	Coordination of Wage setting, higher values indicate more centralized wage setting institutions	Visser [56]	2.95	1.41
Union Coverage	Number of Workers covered by wage bargaining	Visser [56]	35.58	19.27
Replacement	Net Unemployment Replacement Rate for an Average Single Production Worker (with no children)	OECD Social and Welfare Statistics	0.543	0.019
Tax Wedge	The ratio between the amount of taxes paid by an average single worker (a single person at 100% of average earnings) without children and the corresponding total labour cost for the employer.	OECD Revenue Statistics	-1.06	44.31
EPL	The procedures and costs involved in dismissing individuals or groups of workers and the procedures involved in hiring workers on fixed-term or temporary work agency contracts.	OECD Employment Database	2.18	0.83
Min Wage	Minimum Wage Setting insitutions categorical variable. Higher values indicate higher level of min wage setting institutions	Visser [56]	1.24	0.92
Regulation	Summary measure of a wide array of regulatory provisions in the economy	OECD Public Sector, Taxation and Market Regulation database	6.96	1.08
ALMP	Expenditure on labour market policies (LMP) targeted at groups of persons with difficulties in the labour market, as % of GDP	OECD, Labour Market Policy Statistics	0.01	7.07
Articles	(log) Number of journal articles in scientific and technical journals, in the fields of physics, chemistry, biology, mathematics, clinical medicine, biomedical research, engineering and technology, earth and space sciences. Scientific and technical article counts are from journals classified by the Institute for Scientific Information's Science Citation Index (SCI) and Social Sciences Citation Index (SSCI)	World Bank, World Development Indicators	9.546	1.344

Table 2: Fixed Effects 2SLS, OECD sample

	(1)	(2)	(3)	(4)	(5)	(6)
ECI	-9.371**	-16.699***	-9.375**	-18.468***	-12.461***	-15.023**
	(-2.140)	(-2.926)	(-2.135)	(-3.667)	(-2.862)	(-2.092)
Inflation	-0.341***	-0.366***	-0.341***	-0.320***	-0.142**	-0.217**
	(-6.595)	(-6.399)	(-6.586)	(-4.608)	(-2.403)	(-2.499)
Imports	-0.080**	-0.025	-0.080**	-0.056	-0.056	-0.089**
	(-2.267)	(-0.626)	(-2.233)	(-1.365)	(-1.623)	(-2.235)
Output Gap	-21.645***	-15.952**	-21.647***	-18.145**	-14.505**	-11.360*
	(-3.422)	(-2.387)	(-3.422)	(-2.448)	(-2.344)	(-1.766)
Union Density	0.142***	0.164***	0.142***	0.055	-0.033	0.026
	(3.089)	(2.962)	(3.090)	(0.867)	(-0.661)	(0.323)
Centralisation	0.149	0.524	0.148	0.606**	0.417	0.185
	(0.586)	(1.620)	(0.583)	(2.046)	(1.346)	(0.650)
Coordination	-0.815***	-0.911***	-0.814***	-0.859**	-0.898***	-0.840**
	(-2.887)	(-2.621)	(-2.884)	(-2.392)	(-2.908)	(-2.443)
Union Coverage	0.022	0.030	0.022	0.033*	0.123***	0.006
	(1.103)	(1.492)	(1.103)	(1.702)	(2.986)	(0.535)
Replacement	4.339**	-3.155	4.341**	4.234	-0.202	-7.478*
	(2.135)	(-1.492)	(2.136)	(1.020)	(-0.050)	(-1.898)
Tax Wedge	0.130***	0.066	0.129***	0.100*	0.086*	0.099*
	(2.600)	(1.238)	(2.591)	(1.774)	(1.689)	(1.871)
EPL		2.525***				
		(2.598)				
Min Wage			0.045			
			(0.174)			
Regulation				-0.806*		
				(-1.819)		
ALMP					1.769***	
					(5.649)	
Observations	403	362	403	280	344	252
R-sq	0.483	0.494	0.483	0.485	0.575	0.486
F-test	11.23	9.039	11.55	6.126	9.487	7.231
DWH-test	0.528	0.694	0.526	1.140	0.0103	0.574
Weak-id	21.64	18.03	21.44	17.08	22.67	22.87
LM-Underid	53.63	48.85	53.26	38.71	53.26	15.79
Hansen(p-value)	0.347	0.914	0.346	0.791	0.960	

Note: Dependent variable: unemployment rate. *ECI* is instrumented. To save space, the first stage results are not included in the Table; the results are available upon request. Clustered t- statistics in parentheses. DWH is the Durbin- Wu- Hausman test of endogeneity of the regressors. Rejection of the null suggests that the IV regression is required. LM-Underid gives the Kleibergen-Paap Wald test of weak identification, with the null hypothesis indicating that the model is weakly identified. Weak-id gives the F statistic for weak identification. Hansen test (p-value) gives the p-value of the Hansen test of overidentification. Rejection of the null implies that the overidentifying restrictions cannot be rejected. * p<0.10, ** p<0.05, *** p<0.01

four years, it is likely to achieve a higher level of ECI in the current year. This reasoning is verified by the first-stage results (not reported here), since in all cases, the coefficients of the lagged and differenced ECI have a positive sign and are statistically significant at the 1% level. Regarding excludability, while we do not have a precise theory for why the lagged and differenced ECI should have no direct effect on current labour outcomes, it seems plausible to expect that these variables affect labour market outcomes in the current year only through the current level of ECI. More specifically, and taking into account Angrist and Krueger [7]'s caution against blindly using lags as instruments, we run the baseline model including also the four lagged variables of ECI as independent variables. Our results, which are available upon request, show that the variables considered as instruments do not belong in the main model, hence the use of lagged and differenced ECI might not be problematic. It has been shown in the literature [8, 20, 50] that lagged differences of the independent variable are appropriate instruments, provided that they pass the tests for overidentifying restrictions and the tests for their relevance and weakness [32].

In the first column, we present our baseline specification, whereas columns (2)-(5) report the results from our robustness checks regarding the inclusion of additional variables. The coefficient of ECI in each regression is statistically significant, at least at the 5% level. Most importantly, the estimated effect is negative, suggesting that economic complexity is associated with lower unemployment. Our main regression results also imply a substantial effect: a one standard deviation increase in ECI is associated with a one standard deviation decrease in the unemployment rate, or equivalently a decrease in the unemployment rate of four percentage points.

Regarding the rest of the independent variables, our results can be summarized as follows: The inverse relationship between inflation and unemployment is verified by our model. $Inflation$ is negative and statistically significant at the 1% level in all columns. Similarly, and as indicated by our priors, *Output Gap* is negative and statistically significant in all cases. The estimated coefficient implies that a one standard deviation increase on the output gap, i.e. the difference between actual and potential output, is associated with a 0.5% decline in the unemployment rate. On the other hand, $Imports$ and $TaxWedge$, even though correctly signed, are not always statistically significant. Finally, concerning the institutional variables, it seems that their statistical significance and sign change depending on the specification. In general, however, when a variable is found to be significant, its sign is the one expected by our theoretical priors.

The columns (2)-(5) of Table 2 show that our results regarding ECI are robust independently of the explanatory variables included in the analysis. In column (2), we introduce the variable EPL, which is an index of the overall degree of employment protection in the economy, in column (3) a variable that measures the level of the minimum wage, in column (4) an index of the level of regulation of the product markets in the economy, where higher values indicate lower regulation, and, finally, in column (5), spending on active labour market policies as a proportion of GDP.

Table 3: Fixed Effects 2SLS, Unemployment and Employment in Specific Groups

	(1) Youth Unemployment	(2) Male Unemployment	(3) Female Unemployment	(4) Employment Male	(5) Employment Female	(6) Employment Total
ECI	-33.412***	-16.902***	-14.391***	10.720***	3.237	7.028**
	(-3.558)	(-3.547)	(-3.458)	(3.058)	(0.946)	(2.416)
Inflation	-0.427***	-0.269***	-0.115*	0.163***	-0.006	0.077*
	(-2.966)	(-3.852)	(-1.675)	(3.187)	(-0.095)	(1.674)
Imports	-0.126**	-0.064*	-0.071**	0.009	-0.047**	-0.021
	(-1.994)	(-1.754)	(-2.466)	(0.361)	(-2.170)	(-1.090)
Output Gap	-26.181**	-16.241***	-14.936**	18.866***	13.464***	15.891***
	(-2.172)	(-2.579)	(-2.484)	(4.280)	(2.620)	(3.823)
Union Density	-0.032	0.113***	-0.146***	-0.103***	0.094**	-0.001
	(-0.358)	(2.668)	(-2.923)	(-3.309)	(2.003)	(-0.044)
Centralisation	1.209**	0.432	0.990***	-0.363	-0.714***	-0.536**
	(2.144)	(1.470)	(2.999)	(-1.623)	(-2.660)	(-2.389)
Coordination	-2.267***	-1.209***	-1.661***	1.108***	1.018***	1.074***
	(-3.803)	(-3.821)	(-4.336)	(4.325)	(3.371)	(4.208)
Union Coverage	0.022	0.016	0.024	-0.016	-0.017	-0.017
	(1.035)	(1.595)	(1.602)	(-1.562)	(-1.328)	(-1.533)
Replacement	-23.547***	-5.252**	-5.930**	-0.485	-3.710	-2.038
	(-4.497)	(-2.137)	(-2.194)	(-0.253)	(-1.543)	(-1.107)
Tax Wedge	0.227**	0.112**	0.105**	-0.139***	-0.066**	-0.101***
	(1.979)	(2.140)	(2.047)	(-3.707)	(-2.027)	(-3.479)
Observations	338	338	338	338	338	338
R-sq	0.463	0.592	0.566	0.510	0.714	0.634
F-test	7.920	15.27	11.31	11.55	20.46	17.37
DWH-test	0.561	0.922	0.0299	0.00643	0.0000431	0.00000224
Weak-id	23.99	23.99	23.99	23.99	23.99	23.99
LM-Underid	56.64	56.64	56.64	56.64	56.64	56.64
Hansen(p-value)	0.507	0.479	0.644	0.514	0.697	0.599

Note: Dependent variable: as noted in columns. *ECI* is instrumented. To save space, the first stage results are not included in the Table; the results are available upon request. Clustered *t*- statistics in parentheses. DWH is the Durbin- Wu- Hausman test of endogeneity of the regressors. Rejection of the null suggests that the IV regression is required. LM-Underid gives the Kleibergen-Paap Wald test of weak idendification, with the null hypothesis indicating that the model is weakly identified. Weak-id gives the F statistic for weak identification. Hansen test (p-value) gives the p-value of the Hansen test of overidentification. Rejection of the null implies that the overidentifying restrictions cannot be rejected. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

With the exception of the last variable, the signs of the rest of the variables turn out as expected, with $MinWage$ being the only variable that is statistically insignificant.

To further convince the reader about our finding, in column (6) we adopt an alternative instrument of ECI, namely the (log) number of journal *articles* published in scientific and technical journals in a given year. This index calculates the total number of papers in the fields of physics, biology, chemistry, mathematics, clinical medicine, biomedical research, engineering and technology, and earth and space sciences. Higher values are associated with higher scientific effort and output, which are directly related to the intensity of process and product innovation in the economy. Hence, we naturally expect *articles* to influence economic sophistication as measured by the ECI. Going back to Hidalgo and Hausmann Hidalgo and Hausmann [39]'s metaphor of 'Legos bucket', it is reasonable to assume that new knowledge appearing in scientific articles is materialized in new Lego pieces and, in turn, new Lego models. Regarding the exogeneity of the instrument it is plausible to assume that changes in the number of journal articles do not have a direct impact on labour market institutions and outcomes. The first-stage results (not reported in Table 2) confirm the positive and statistical significant relationship between *articles* and ECI at the 1% level and the negative effect of the latter on unemployment is confirmed in the second-stage results reported in column (6).

In Table 3, we estimate the same model as in column (1) of Table 2; however this time, we use as dependent variables the unemployment rates of specific groups, i.e. the unemployment rate for individuals aged 15-24 (column 1), the male unemployment rate (column 2) and the female unemployment rate (column 3). We also estimate the male employment rate (column 4), the female employment rate (column 5) and the total employment rate (column 6). In all cases, the results of the baseline model are verified. Economic complexity has a negative effect on unemployment for all groups and a positive effect on employment. The only exception to this is column (5); in this case the ECI does not exert a statistically significant effect on the employment of females, even though the variable is again positively signed. Similarly, for the rest of the explanatory variables, the same picture emerges as in the baseline model, with the exception of *Replacement*, which, in columns (1) to (3), is negatively signed and statistically significant in contrast to our theoretical priors.

Supplementary regressions are also presented in Table 4, in which we examine the dynamics of the underlying relationship. The underlying assumption tested is that the effect of all explanatory variables comes with an one-year lag. In column (1), we estimate equation 9 but all variables are introduced with a time lag, while in column (2) ECI is considered without a lag. An additional benefit of the time lag assumption is that it further allows us to control for potential endogeneity between the explanatory variables and ECI. In columns (3) and (4), we also consider the dependent variable with a time lag. In column (3) we estimate a simple ordinary least squares (OLS) panel fixed effects model, whereas in column (4) we employ the Arellano-Bond estimator. In all cases, our main variable of interest remains statistically significant and its magnitude does not change much.

Table 4: Dynamic model; OECD sample

	(1) All Independent Lagged	(2) Lagged ex.ECI	(3) Lagged Unemployment	(4) Arellano Bond
ECI	-9.024*	-12.833**	-7.749**	-3.710**
	(-1.821)	(-2.390)	(-1.972)	(-2.287)
Inflation	-0.305***	-0.320***	-0.168***	-0.065*
	(-4.812)	(-4.944)	(-3.161)	(-1.836)
Imports	-0.130***	-0.101**	-0.072***	0.005
	(-3.357)	(-2.423)	(-2.643)	(0.851)
Output Gap	-14.268**	-13.272*	-12.742***	-12.011***
	(-2.019)	(-1.877)	(-2.742)	(-3.904)
Union Density	0.116**	0.119***	0.043	-0.002
	(2.533)	(2.596)	(1.568)	(-0.380)
Centralisation	0.272	0.309	0.056	-0.252*
	(1.068)	(1.194)	(0.354)	(-1.717)
Coordination	-0.793***	-0.744***	-0.275	0.077
	(-2.808)	(-2.748)	(-1.530)	(0.636)
Union Coverage	0.026	0.029	0.019*	0.003
	(1.193)	(1.406)	(1.717)	(0.820)
Replacement	3.078	3.130	0.798	2.137**
	(1.494)	(1.460)	(0.495)	(2.448)
Tax Wedge	0.119**	0.133**	0.022	0.007
	(2.080)	(2.369)	(0.682)	(0.544)
Unemployment (-1)			0.756***	0.940***
			(11.814)	(20.992)
Observations	407	410	397	402
R-sq	0.400	0.390	0.806	
F-test	5.951	5.894	35.29	41.86
DWH-test			5.280	
LM-Underid	55.22	59.21	25.09	
Sargan(p-value)				0.00
Hansen(p-value)	0.481	0.283	0.647	
AR(1)(p-value)				0.00190
AR(2)(p-value)				0.401

Note: Sargan (p-value) and Hansen(p-value) give the p-value of the Sargan and Hansen test of overidentifying restrictions respectively. AR(1)(p-value) and AR(2)(p-value) give the p-value for the rest of first and second order autocorrelation. See also Table 2.

Table 5: Data sources, definitions and summary statistics; world sample

Variable	Definition	Source	Mean	Std. Deviation
ECI	Economic Complexity Index	Observatory of Economic Complexity	0.003	0.97
Inflation	% change in annual Consumer Price Index	International Monetary Fund, International Financial Statistics	37.03	117.02
Imports	Imports of goods and services as % of GDP	World Bank, World Development Indicators	42.8	21.58
Output Gap	The difference between actual and potential real (GDP) as a per cent of potential real GDP.	IMF World Economic Outlook	-0.001	0.003
Taxes	Share of Tax Revenue to GDP	World Bank, World Development Indicators	15.59	6.88
Articles	(log) Number of journal articles in scientific and technical journals, in the fields of physics, chemistry, biology, mathematics, clinical medicine, biomedical research, engineering and technology, earth and space sciences. Scientific and technical article counts are from journals classified by the Institute for Scientific Information's Science Citation Index (SCI) and Social Sciences Citation Index (SSCI)	World Bank, World Development Indicators	5.94	2.7
Genetic Diversity	The expected heterozygosity (genetic diversity) of a country's contemporary national population, as developed by Ashraf and Galor [10] and Ashraf and Galor [9]. This measure is based on migratory distances from East Africa to the year 1500 locations of the ancestral populations of the country's component ethnic groups in 2000 and on the pairwise migratory distances among these ancestral populations	Ashraf and Galor [10]	72.63	2.75
Secular Values Index	12-item measure of distance from sacred sources of authority	World Values Survey	0.38	0.1

4.2 Cross-section, world sample

To generalize our findings from the previous section to a wider set of countries, we focus our analysis on the global sample of 74 developed and developing countries. To maximize the number of countries used in the regression, we employ only a subset of the control variables used in the econometric analysis of the previous section, and we use averages from 1990 to 2010. We end up with a maximum of 74 observations for our cross-country specification. Data definitions and summary statistics for the world sample are given in Table 5.

Applying a fixed-effects 2SLS/IV regression in a cross-section setting requires a set of external instruments. We experiment with three instruments of ECI and examine the robustness of our results using different subsets of these instruments. Firstly, we employ again the measure of the (log) number of journal articles published in scientific and technical journals in a given year (denoted *Articles*).

The second instrument considered is an index of genetic diversity. Following the comparative development literature (Ashraf and Galor [10]), genetic diversity, predominantly determined during the prehistoric 'out of Africa' migration of humans, explains modern ethnic diversity and economic prosperity. Following the relevant literature, "*higher diversity therefore enhances society's capability to integrate advanced and more efficient production methods, expanding the economy's production possibility frontier and conferring the benefits of improved productivity*" (Ashraf and Galor [10], p. 3). Therefore, the proportion of ethnic diversity explained by prehistoric diversity is expected to be correlated with economic complexity, without having a direct effect on contemporary unemployment and employment rates.

The third instrument used is the *Secular Values Index*, which is a 12-item measure of the distance from 'sacred' sources of authority in each country [57]. It is a continuous scale in the [0,1] range, where 0 (1) denotes the less (more) secular position. Countries that hold high beliefs in 'sacred' sources of authority are expected to be less modernized and less prone to innovation and adoption of sophisticated methods of production.

The top part of Table 6 shows the output of the first-stage regressions for the external instruments.[6] The results indicate that ECI is positively associated with both the amount of research undertaken in an economy (*Articles*) and the *Secular Values* Index. On the other hand, it seems that higher levels of economic complexity are associated with lower genetic diversity. The hypothesis of weak identification is rejected in all instances, since the value of the relevant test (F-statistic of the first-stage estimation: *Weak-id*), is well above 10.

The second-stage regression results verify the negative and statistically significant relationship between ECI and unemployment rate. The estimated effect is relatively smaller than in the OECD sample: a one standard deviation increase in ECI results in an approximately 1.15 percentage points decline in unemployment. With respect to the rest of the (second stage) results, the main conclusions drawn from the OECD sample remain qualitatively intact.

To examine the robustness of our results in columns (2) and (3), we experiment using different subsets of the instruments employed in the main specification (column 1). Once again, the asso-

[6] To save space, the first-stage results for the independent variables are not included in the Table. Results are available upon request.

Table 6: Cross-section regressions

	(1) Total unempl.	(2) Total unempl.	(3) Total unempl.	(4) Total unempl.	(5) Youth unempl.	(6) Male unempl.	(7) Female unempl.
Articles	0.245***	0.262***	0.297***		0.252***	0.233***	0.245***
	(8.425)	(8.484)	(9.714)		(8.477)	(7.428)	(8.425)
Gen. Diversity	-0.078***		-0.096***	-0.026	-0.080***	-0.068**	-0.078***
	(-3.124)		(-3.904)	(-0.786)	(-3.151)	(-2.379)	(-3.124)
Secular Values	2.361***			4.722***	2.377***	2.251***	2.361***
	(4.560)			(6.579)	(4.623)	(4.363)	(4.560)
ECI	-2.467**	-2.046*	-2.849**	-1.873*	-3.616*	-7.519	-4.059***
	(-2.420)	(-1.818)	(-2.564)	(-1.756)	(-1.873)	(-1.570)	(-2.816)
Inflation	-0.007***	-0.007***	-0.008***	-0.007***	-0.009*	-0.050	-0.010***
	(-3.529)	(-3.312)	(-3.530)	(-3.361)	(-1.952)	(-1.448)	(-3.322)
Imports	0.028	0.007	0.029	0.025	0.033	0.198	0.016
	(0.811)	(0.278)	(0.839)	(0.768)	(0.516)	(1.370)	(0.378)
Output Gap	-351.945*	-400.238**	-342.541*	-366.557*	-870.566**	-1261.259	-243.160
	(-1.863)	(-2.254)	(-1.807)	(-1.916)	(-2.287)	(-0.600)	(-0.877)
Taxes	0.192**	0.168**	0.202**	0.176**	0.351**	0.131	0.221**
	(2.196)	(1.985)	(2.288)	(2.005)	(2.071)	(0.361)	(2.157)
Constant	5.761***	6.750***	5.706***	5.847***	13.016***	28.951***	7.719***
	(3.811)	(4.605)	(3.678)	(3.998)	(4.137)	(3.516)	(3.476)
Observations	70	74	70	70	69	59	70
R-sq	0.115	0.109	0.101	0.125	0.0691	0.119	0.0950
F-test	4.905	4.536	4.951	4.927	3.347	1.227	2.503
DWH-test	0.566	0.353	2.540	0.0663	2.905	0.0101	3.389
Weak-id	43.43	71.99	47.23	22.10	42.32	39.70	43.43
LM-Underid	26.76	26.25	25.18	15.61	25.42	19.82	26.76
Hansen(p-values)	0.0155		0.0210	0.00559	0.0339	0.00399	0.0805

Notes: See Table 2. To save space we do not report the first stage results for the second stage independent variables.

ciation of ECI with unemployment is negative and statistically significant. Finally, in columns (5), (6) and (7) we examine the effect of economic complexity on youth, male and female unemployment rates, respectively. Qualitatively, the results are similar to the ones obtained from the OECD panel dataset.

5 Products complexity and the labour market

The ECI methodology provides a useful toolbox that allows us to compute indexes that quantify economic sophistication, for both countries and products. For example, using the same methodology that computes ECI, but placing the spotlight on products rather than on countries, we can calculate the PCI (see Section 3). This index quantifies the sophistication of each product according to the amount of knowledge/know-how involved in its production, reflected by the countries that export the product [37]. In other words, when a product is located in the center of the product space i.e. in the core of the international trade network of products, it ranks higher in the PCI because its production requires more knowledge/know-how. Recently, Hartmann *et al.* [35], using the ECI methodology, introduced a measure that associates products with income inequality and showed how the development of new products is associated with changes in income inequality. However, the labour market effects are key to understanding national income disparities, since income differences are, by definition, based on differences in the labour productivity and/or employment level, among other factors. Here, we introduce a measure that links a product to the average unemployment and employment rates of the countries that export it. In this way, we illustrate how labour markets are affected by the level of products' sophistication and we quantify the influence of countries' level of economic complexity on their labour markets' outcomes.

Following Hartmann *et al.* [35], we define the *Product Unemployment Index* (PUI) (resp. *Product Employment Index*, PEI) as the average unemployment rate (resp. employment rate) faced by the countries that export the focal product, normalized by the importance of this product to the total exports of the countries that export it. More precisely, we decompose the relationship between economic complexity and unemployment and employment rates into individual economic sectors, by creating product-level estimators of these rates that are expected for the countries exporting a given product.

5.1 Product unemployment and employment indexes

Assuming that we have trade data for l countries and k products, we can fill the $(l \times k)$ matrix $\mathbf{M}$ so that its matrix element $M_{cp} = 1$ if country c has Revealed Comparative Advantage for product p and zero otherwise (see Section 3). For our case, the international trade data from MIT's Observatory of Economic Complexity contains information for 33 OECD countries and 773 products from 1985 to 2008, classified in accordance with the SITC at the 4-digit level. A visualization of matrix $\mathbf{M}$ for this dataset, which is used to calculate the ECI and the PCI is shown in Figure 1.

Table 7: PUI by industry: averages across time and across 4-digit categories

SITC4	Industry	PUI	SITC4	Industry	PUI
76	Telecommunications and sound-recording and reproducing apparatus and equipment	5.52	79	Other transport equipment	7.76
88	Photographic apparatus, equipment and supplies and optical goods	5.69	93	Special transactions and commodities not classified according to kind	7.76
97	Gold, non-monetary (excluding gold ores and concentrates)	5.91	12	Tobacco and tobacco manufactures	7.77
87	Professional, scientific and controlling instruments and apparatus	6.46	65	Textile yarn, fabrics, made-up articles, n.e.s., and related products	7.78
34	Gas, natural and manufactured	6.57	11	Beverages	7.84
35	Electric current	6.73	91	Postal packages not classified according to kind	7.87
73	Metalworking machinery	6.76	53	Dyeing, tanning and colouring materials	7.87
77	Electrical machinery, apparatus and appliances	6.77	54	Medicinal and pharmaceutical products	7.89
25	Pulp and waste paper	6.83	52	Inorganic chemicals	7.90
96	Coin (other than gold coin), not being legal tender	7.02	58	Plastics in non-primary forms	7.92
72	Machinery specialized for particular industries	7.03	69	Manufactures of metals, n.e.s.	7.92
43	Animal or vegetable fats and oils	7.07	24	Cork and wood	7.95
74	General industrial machinery and equipment	7.09	64	Paper, paperboard and articles of paper pulp, of paper or of paperboard	7.95
75	Office machines and automatic data-processing machines	7.1	26	Textile fibres and their wastes	7.97
41	Animal oils and fats	7.17	83	Travel goods, handbags and similar containers	8.03
71	Power-generating machinery and equipment	7.22	84	Articles of apparel and clothing accessories	8.05
89	Miscellaneous manufactured articles, n.e.s.	7.35	67	Iron and steel	8.07
68	Non-ferrous metals	7.4	85	Footwear	8.09
59	Chemical materials and products	7.45	66	Non-metallic mineral manufactures	8.17
81	Prefabricated buildings; sanitary, plumbing, heating and lighting fixtures and fittings	7.49	61	Leather, leather manufactures, and dressed furskins	8.21
28	Metalliferous ores and metal scrap	7.52	27	Crude fertilizers, other than those of Division 56, and crude minerals	8.35
29	Crude animal and vegetable materials	7.53	56	Fertilizers (other than those of group 272)	8.36
23	Crude rubber (including synthetic and reclaimed)	7.53	57	Plastics in primary forms	8.49
78	Road vehicles	7.55	82	Furniture and parts thereof; bedding, mattresses, mattress supports, cushions and similar stuffed furnishings	8.53
51	Organic chemicals	7.57	63	Cork and wood manufacture	8.59
21	Hides, skins and furskins, raw	7.59	62	Rubber manufactures, n.e.s.	8.61
33	Petroleum, petroleum products and related materials	7.6	55	Essential oils and resinoids and perfume materials; toilet, polishing and cleansing preparations	8.93
42	Fixed vegetable fats and oils, crude, refined or fractionated	7.61	32	Coal, coke and briquettes	9.17
22	Oil-seeds and oleaginous fruits	7.65			

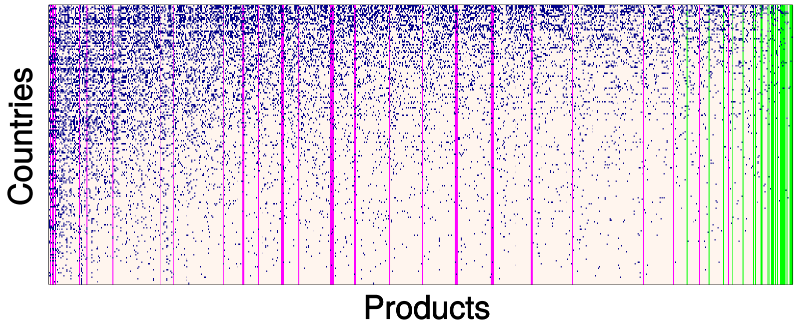

Figure 1: **Matrix representation of the links between a country and the products it exports.** A visualization of this matrix for the year 2010, where a dark point indicates that country c exports a given product p. The matrix is sorted using the NODF algorithm [6], which highlights the existence of countries that are very well diversified and countries that export only a small set of products. Highlight in green is the position of the 40 products with the highest PCI values and in red the 40 products with the lowest PCI values. It is clear that the more diversified countries are those that produce the more complex products.

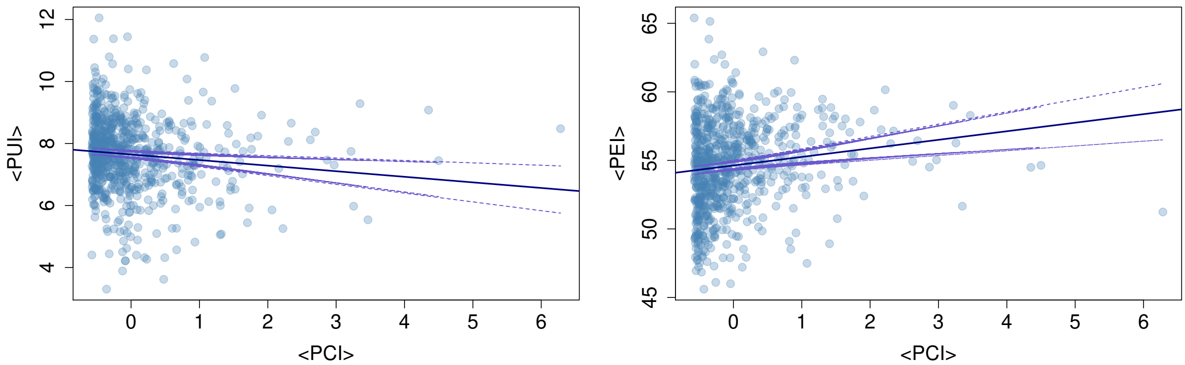

Figure 2: **PUI and PEI against PCI.** The solid lines represent the fit of a linear model and the dashed lines represent a 95% prediction interval based on the fitted linear model.

Every product p generates some value for the country c that exports it. Therefore, for every product p we can calculate the fraction s_{cp}:

$$s_{cp} = \frac{X_{cp}}{\sum_{p'} X_{cp'}}, \tag{10}$$

where X_{cp} is the total export value of product p when exported by country c, while $\sum_{p'} X_{cp'}$ is the value of all exports of country c. If U_c (resp. E_c) is the unemployment (resp. employment) rate of country c, we can calculate the PUI_p and the PEI_p for every product, as:

$$\mathrm{PUI}_p = \frac{1}{N_p} \sum_c M_{cp} s_{cp} U_c, \tag{11}$$

Table 8: List of the five products with the highest and lowest PUI and PEI values during the period 1985-2008

SITC4	Product name	Product section	PUI	PEI
Five products with highest PUI				
571	Polymers of ethylene, in primary forms	Chemicals and related products, n.e.s.	11.9	
8994	Umbrellas, sun umbrellas, walking-sticks, etc	Miscellaneous manufactured articles	11.4	
8933	Floor coverings, wall or ceiling coverings	Miscellaneous manufactured articles	11.3	
6624	Non-refractory ceramic bricks, tiles, pipes	Manufactured goods	11.3	
2450	Fuel wood/wood charcoal	Crude materials	11.2	
Five products with lowest PUI				
7612	Monitors and projectors etc	Machinery & transport equipment	4.5	
8982	Musical instruments	Miscellaneous manufactured articles	4.4	
3330	Petroleum oils	Mineral fuels, lubricants	4.4	
7638	Video-recording	Machinery and transport equipment	4.3	
8852	Watches and clocks	Miscellaneous manufactured articles	3.9	
Five products with highest PEI				
3330	Petroleum oils	Mineral fuels, lubricants		65.5
3414	Gas	Mineral fuels, lubricants		65.2
8851	Watches and clocks	Miscellaneous manufactured articles		63.4
2516	Pulp and waste paper	Crude materials		63.0
2222	Oil-seeds and oleaginous fruits	Crude materials		62.7
Five products with lowest PEI				
6597	Floor coverings, etc.	Manufactured goods		47.2
571	Polymers of ethylene, in primary forms	Chemicals and related products		47.0
1211	Tobacco, unmanufactured	Beverages and tobacco		47.0
6624	Non-refractory ceramic bricks, tiles, pipes	Manufactured goods		46.3
4235	Fixed vegetable fats and oils	Animal and vegetable oils		46.2

Notes: PUI: Product Unemployment Index; PEI: Product Employment Index. Average value for 1985-2008

$$\mathrm{PEI}_p = \frac{1}{N_p} \sum_c M_{cp} s_{cp} E_c, \tag{12}$$

where $N_p = \sum_c M_{cp} s_{cp}$ is a normalization factor.

Utilizing the information we have for the unemployment and employment rates for the OECD countries we are able to calculate the above indexes. For every year in the period 1985-2008 we calculate all product-related indexes, i.e. PCI, PUI and PEI, and we obtain their mean value for each product. Table 7 lists the averages of PUI across the sample and across 4-digit SITC4 categories of the 2-digit SITC4 industries. Industries are sorted in order of increasing PUI. Table 7 reveals that the industry group with the lowest average proportion of the total unemployment rate is *'Telecommunications and sound-recording and reproducing apparatus and equipment'*. Similarly, the more sophisticated industry/product categories appear to have the lowest PUI. At the other end of the spectrum, the *'Coal, coke and briquettes'* industry has the highest PUI. As the reader can easily verify, primary sector industries (with low product-sophistication), appear to be associated with higher rates of unemployment. This is also implied by Table 8 which lists the five products with the highest and lowest PUI and PEI values during the period 1985-2008.

In addition, we test the existence of a bivariate relationship between PCI and PUI and PEI. Thus, we calculate Pearson's correlation coefficient for both pairs, i.e. PUI against PCI and PEI against PCI. If a relation exists, it should allow us to derive expectations of whether or not the products' complexity can be associated with the unemployment and employment rates. In the case of PUI against PCI, the correlation coefficient is $\rho = -0.10$ with p-value = 0.0061, while for the case of PEI against PCI it is $\rho = 0.14$ with p-value = 0.0002. In Figure 2, we present the scatter plots of PUI and PEI against PCI for all 773 products in our dataset together with the fitted linear models. The slopes of the linear fits are the corresponding correlation coefficients.

The statistically significant negative (resp. positive) correlation between PUI (resp. PEI) and PCI indicates that the sophisticated products are associated with countries that bear relative low unemployment rates (resp. high employment rates). This adds to our previous discussion about economic complexity at the country level, as it allows us to understand which sets of products are leading to more employment and less unemployment based on their sophistication.

6 Conclusions

Our analysis illustrates that the labour market performance of a country is highly predicted by the mix of products that a country produces and exports. Both in a panel and in a cross-country setting we have verified that there is a robust negative (resp. positive) relationship between unemployment (resp. employment) and product sophistication. Moreover, the relationship between these two variables is verified by instrumental variables (IV) estimation techniques. Hence, the evidence presented in this paper suggests that a country's level of economic sophistication, determines its labour market outcomes.

In detail, countries that produce more sophisticated products generally have lower unemployment rates and higher employment rates. As higher product sophistication results in higher growth rate, there seems to be a capitalization effect at work [17, 48]: the present value for firms creating new jobs is higher when product sophistication increases, and, according to our estimates, this effect is not symmetrical across industries. We built the PUI and the PEI, which associate exported products with the average level of countries' employment and unemployment rates, respectively. With these indexes we show how the development of sophisticated products is associated with changes in the labour market. This result is important from a policy perspective. Using the proposed indexes, it is possible to design sectoral reallocation policies and smart specialization strategies that promote activities/ sectors that are associated with lower unemployment and higher employment. Adding to the above, our analysis provides additional insights for the *ex post* policy evaluation process. As many tax and subsidy policies are associated with sectoral reallocation, our indexes can provide a quantitative measure of the average unemployment cost (or gain) due to the implemented policy.

In sum, this study examined labour market outcomes at the macroeconomic level, but went beyond the standard institutional and economic factors to explain unemployment. We identified economic complexity as an explanatory variable of the observed differences in labour market outcomes across countries. An interesting way to build further on this would be to identify the exact inclusive institutions and technological capabilities that can have a mitigating effect on 'technological unemployment'.

References

[1] Abdon, A.; Felipe, J. (2011). The product space: What does it say about the opportunities for growth and structural transformation of Sub-Saharan Africa? *Economics Working Paper Archive* **670**, Levy Economics Institute.

[2] Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of economic literature* **40(1)**, 7–72.

[3] Aidt, T. S.; Tzannatos, Z. (2008). Trade unions, collective bargaining and macroeconomic performance: a review. *Industrial Relations Journal* **39(4)**, 258–295.

[4] Albeaik, S.; Kaltenberg, M.; Alsaleh, M.; Hidalgo, C. A. (2017). 729 new measures of economic complexity (Addendum to Improving the Economic Complexity Index). *arXiv preprint arXiv:1708.04107* .

[5] Albeaik, S.; Kaltenberg, M.; Alsaleh, M.; Hidalgo, C. A. (2017). Measuring the Knowledge Intensity of Economies with an Improved Measure of Economic Complexity. *arXiv preprint arXiv:1707.05826* .

[6] Almeida-Neto, M.; Guimaraes, P.; Guimarães, P. R.; Loyola, R. D.; Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117(8)**, 1227–1239.

[7] Angrist, J. D.; Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives* **15(4)**, 69–85.

[8] Arellano, M.; Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics* **68(1)**, 29–51.

[9] Ashraf, Q.; Galor, O. (2013). Genetic diversity and the origins of cultural fragmentation. *The American economic review* **103(3)**, 528–533.

[10] Ashraf, Q.; Galor, O. (2013). The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *The American Economic Review* **103(1)**, 1–46.

[11] Autor, D.; Dorn, D.; Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the US. *American Economic Review* **103(6)**, 2121–68.

[12] Autor, D. H.; Dorn, D.; Hanson, G. H.; Song, J. (2014). Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics* **129(4)**, 1799–1860.

[13] Autor, D. H.; Katz, L. F.; Krueger, A. B. (1998). Computing inequality: have computers changed the labor market? *The Quarterly Journal of Economics* **113(4)**, 1169–1213.

[14] Autor, D. H.; Levy, F.; Murnane, R. J. (2002). Upstairs, downstairs: computers and skills on two floors of a large bank. *ILR Review* **55(3)**, 432–447.

[15] Autor, D. H.; Levy, F.; Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics* **118(4)**, 1279–1333.

[16] Bartel, A.; Ichniowski, C.; Shaw, K. (2007). How does information technology affect productivity? Plant-level comparisons of product innovation, process improvement, and worker skills. *The quarterly journal of Economics* **122(4)**, 1721–1758.

[17] Bean, C.; Pissarides, C. (1993). Unemployment, consumption and growth. *European Economic Review* **37(4)**, 837–854.

[18] Beaudry, P.; Green, D. A.; Sand, B. M. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics* **34(S1)**, S199–S247.

[19] Bernard, A. B.; Smeets, V.; Warzynski, F. (2017). Rethinking deindustrialization. *Economic Policy* **32(89)**, 5–38.

[20] Blundell, R.; Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics* **87(1)**, 115–143.

[21] Bound, J.; Johnson, G. E. (1989). *Changes in the Structure of Wages during the 1980's: An Evaluation of Alternative Explanations. Tech. rep.*, National Bureau of Economic Research.

[22] Bresnahan, T. F.; Brynjolfsson, E.; Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics* **117(1)**, 339–376.

[23] Buera, F. J.; Kaboski, J. P. (2009). Can traditional theories of structural change fit the data? *Journal of the European Economic Association* **7(2-3)**, 469–477.

[24] Bustos, S.; Gomez, C.; Hausmann, R.; Hidalgo, C. A. (2012). The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS one* **7(11)**, e49393.

[25] Caldarelli, G.; Cristelli, M.; Gabrielli, A.; Pietronero, L.; Scala, A.; Tacchella, A. (2012). A network analysis of countries' export flows: firm grounds for the building blocks of the economy. *PloS one* **7(10)**, e47278.

[26] Cristelli, M.; Gabrielli, A.; Tacchella, A.; Caldarelli, G.; Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. *PloS one* **8(8)**, e70726.

[27] Cristelli, M.; Tacchella, A.; Pietronero, L. (2015). The heterogeneous dynamics of economic complexity. *PloS one* **10(2)**, e0117174.

[28] Daveri, F.; Tabellini, G. (2000). Unemployment, growth and taxation in industrial countries. *Economic Policy* **15(30)**, 47–104.

[29] David, H. (2015). Why are there still so many jobs? The history and future of workplace automation. *The Journal of Economic Perspectives* **29(3)**, 3–30.

[30] David, H.; Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *The American Economic Review* **103(5)**, 1553–1597.

[31] Feenstra, R. C.; Lipsey, R. E.; Deng, H.; Ma, A. C.; Mo, H. (2005). *World trade flows: 1962-2000. Tech. rep.*, National Bureau of Economic Research.

[32] Feldmann, H. (2013). Technological unemployment in industrial countries. *Journal of Evolutionary Economics* **23(5)**, 1099–1126.

[33] Felipe, J. (2012). *Inclusive Growth, Full Employment, and Structural Change: Implications and Policies for Developing Asia.* Anthem Press.

[34] Goos, M.; Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *The review of economics and statistics* **89(1)**, 118–133.

[35] Hartmann, D.; Guevara, M. R.; Jara-Figueroa, C.; Aristarán, M.; Hidalgo, C. A. (2017). Linking Economic Complexity, Institutions, and Income Inequality. *World Development* **93**, 75–93.

[36] Hausmann, R.; Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth* **16(4)**, 309–342.

[37] Hausmann, R.; Hidalgo, C. A.; Bustos, S.; Coscia, M.; Simoes, A.; Yildirim, M. A. (2014). *The atlas of economic complexity: Mapping paths to prosperity.* Mit Press.

[38] Hausmann, R.; Hwang, J.; Rodrik, D. (2007). What you export matters. *Journal of economic growth* **12(1)**, 1–25.

[39] Hidalgo, C. A.; Hausmann, R. (2009). The building blocks of economic complexity. *proceedings of the national academy of sciences* **106(26)**, 10570–10575.

[40] Hidalgo, C. A.; Klinger, B.; Barabási, A.-L.; Hausmann, R. (2007). The product space conditions the development of nations. *Science* **317(5837)**, 482–487.

[41] Inoua, S. (2016). A Simple Measure of Economic Complexity. *arXiv preprint arXiv:1601.05012* .

[42] Juhn, C.; Murphy, K. M.; Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of political Economy* **101(3)**, 410–442.

[43] Katz, L. F.; Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics* **107(1)**, 35–78.

[44] Lichter, A. (2016). *Benefit Duration and Job Search Effort: Evidence from a Natural Experiment. IZA Discussion Papers 10264*, Institute for the Study of Labor (IZA). URL `https://EconPapers.repec.org/RePEc:iza:izadps:dp10264`.

[45] Matsuyama, K. (2009). Structural change in an interdependent world: A global view of manufacturing decline. *Journal of the European Economic Association* **7(2-3)**, 478–486.

[46] Nickell, S.; Redding, S.; Swaffield, J. (2008). The uneven pace of deindustrialisation in the OECD. *The World Economy* **31(9)**, 1154–1184.

[47] Pierce, J. R.; Schott, P. K. (2016). The surprisingly swift decline of US manufacturing employment. *The American Economic Review* **106(7)**, 1632–1662.

[48] Pissarides, C. A. (1992). Loss of skill during unemployment and the persistence of employment shocks. *The Quarterly Journal of Economics* **107(4)**, 1371–1391.

[49] Rodrik, D. (2006). What's so special about China's exports? *China & World Economy* **14(5)**, 1–19.

[50] Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics* **71(1)**, 135–158.

[51] Rowthorn, R.; Ramaswamy, R. (1999). Growth, trade, and deindustrialization. *IMF Staff papers* **46(1)**, 18–41.

[52] Saviotti, P. P.; Frenken, K. (2008). Export variety and the economic performance of countries. *Journal of Evolutionary Economics* **18(2)**, 201–218.

[53] Scarpetta, S. (1996). Assessing the role of labour market policies and institutional settings on unemployment: A cross-country study. *OECD Economic studies* **26(1)**, 43–98.

[54] Tacchella, A.; Cristelli, M.; Caldarelli, G.; Gabrielli, A.; Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific reports* **2**, 723.

[55] Tacchella, A.; Cristelli, M.; Caldarelli, G.; Gabrielli, A.; Pietronero, L. (2013). Economic complexity: conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control* **37(8)**, 1683–1691.

[56] Visser, J. (2015). The Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts. ICTWSS Database, version 5.1. Amsterdam institute for Advanced Labour Studies, Amsterdam. *Amsterdam: Amsterdam Institute for Advanced Labour Studies* .

[57] Welzel, C. (2013). *Freedom rising.* Cambridge University Press.

[58] Wyplosz, C. (2001). Do we know how low should inflation be. *Why price stability* , 15–33.