Inequality Accounting for a Large Cross-Section of Countries

(Very preliminary version. Please do not quote without permission)

Davide Fiaschi and Angela Parenti*
University of Pisa

May 15, 2014

Abstract

We propose a novel nonparametric methodology to estimate how cross-country differences in investment rates, stocks of human capital, growth rates of employment and initial levels of productivity explain the change in the distribution of labour productivity for a sample of 84 countries in the period 1960-2008. We find that the initial level of labour productivity and the investment rate have decreased dispersion and polarization (to large extent only the first); the stock of human capital and growth rate of employment have increased dispersion and polarization (only marginally the first); and that unobservable characteristics of countries have had a modest impact on dispersion, but have played a crucial role in the emergence of polarization in 2008.

Keywords: convergence, inequality, distribution dynamics, polarization, nonparametric methods.

JEL: *C21; E62; R11; O52*

*Dipartimento di Economia e Management, Via Ridolfi 10, 56124, Pisa, Italy, e-mail: dfiaschi@ec.unipi.it - aparenti@ec.unipi.it, phone/fax: ++39.050.2216208/++39.050.2216384.

Contents

1 Introduction **3**

2 Methodology **4**

2.1 Modeling Growth 5

2.2 Distribution Effect 5

2.2.1 Counterfactual Stochastic Kernels and Ergodic Distributions . 6

2.2.2 The Distributional Effect of Individual Variables and the Marginal Growth Effect 7

2.3 Test of Residual Growth 9

3 Empirical Analysis **9**

3.1 The Estimate of Growth Model 11

3.1.1 The Estimate of Best Growth Model 13

3.2 Distributional Impact of Explanatory Variables 18

3.2.1 Initial labour Productivity 19

3.2.2 Investment Rate 20

3.2.3 Human Capital 21

3.2.4 Employment Growth 22

3.2.5 Dummies 23

4 Concluding Remarks **25**

A Country List **29**

B Descriptive Statistics **30**

1 Introduction

The causes of inequality in the countries'productivities is a still very debated issue. Many years of empirical research have not been able to select among several growth paradigms proposed in literature (see Durlauf et al. (2005) and Durlauf (2009)), suggesting that there not exists just an unique explanatory variable of the development of countries. Taken as granted the existence of complementary sources of long-run growth, the actual literature provides only a limited information on the individual contributions of the possible *proximate* explanatory variables to the overall inequality, despite this information becomes crucial to formulate economic policy recommendations (especially in presence of public budget constraints).[1] In particular, one the most important issue is the role of factor accumulation (physical and human capital) versus total factor productivity (TFP) as sources of convergence or divergence (see, e.g., the findings of Henderson and Russell (2005) versus Easterly and Levine (2001) and the discussion in (Durlauf et al., 2005, p. 605)).

This paper contributes to this literature applying a new methodology developed in Fiaschi et al. (2012b) to quantify the impact of individual variables on cross-country distribution of labour productivity via a counterfactual approach based on semiparametric growth regressions in the period 1960-2008. In particular, we explore the capacity of the well-known augmented Solow growth model proposed by Gregory Mankiw, David Romer and David Weil (MRW) in 1992 (Mankiw et al. (1992)) to account for the inequality in the distribution of labor productivity. We depart from the current approach adopted in growth regressions i) adopting a very parsimonious growth model where, the omitted variable bias, that generally plague the augmented Solow model, is resolved by including some dummies optimally chosen to maximize the goodness-of-fit of the model; and ii) allowing for a nonlinear impact of explanatory variables on growth rate. In this approach dummies have a role similar to TFP in growth accounting, catching the unexplained part of countries'growth; TFP

[1] Here we refer to Durlauf et al. (2009), who argue that: "current debate in growth economics, which has focused on the role of fundamental factors such as geography and institutions as opposed to proximate factors such as macroeconomic policy. The division between fundamental theories and proximate theories is not well defined, in the abstract, but in practice fundamental theories refer to slower moving factors that create an environment out of which neoclassical growth dynamics emerge. [...] Religion, geography, ethnic fractionalisation and institutions are sometimes distinguished as representing fundamental rather than proximate growth determinants." In this paper we directly consider potential proximate variables as investment rates, growth rate of employment and education, suggested by the augmented Solow model (see Mankiw et al. (1992)).

is indeed a measure of "our ignorance" as clearly discussed in Easterly and Levine (2001).

The augmented Solow model with dummies reveals a remarkable power in reproducing the observed distribution of labor productivity across countries (about 92.5% of total observed variance of labour productivity), with all the explanatory variables (labor productivity in 1960, investment rate, stock of human capital proxied by the share of labor force with secondary education, and growth rate of employment) statistically significant at 5% conventional level. As expected investment rate and the stock of human capital appear positively correlated with the average growth rate of labor productivity, while the growth rate of employment is negatively correlated; all three variables however present non-linear relationships (expecially investment rates and the stock of human capital). The labor productivity in 1960 has an inverted U-shape relationship with the growth rate of labor productivity, casting some doubt on the presence of conditional convergence in the sample, i.e. on the presence of decreasing returns or international technological spillovers (these are the explanations of conditional convergence).[2]

Summarizing the main findings on the source of inequality across countries are: i) initial levels of labour productivity, proxy for diminishing returns and/or technological "catching up" have decreased dispersion and polarization to large extent; i) investment rates have only marginally decreased dispersion and polarization; iii) the accumulation of human capital has only marginally increased dispersion and polarization; iv) the growth rate of population has substantially increased the dispersion and the polarization; finally, v) the dummies has a modest impact on dispersion, but have played a crucial role in the emergence of polarization.

Paper is organized as follows: Section 2 briefly describes the current state of literature and presents the proposed methodology; Section 3 presents the result of empirical analysis, and, finally, Section 4 contains some concluding remarks. Appendix gathers descriptive statistics of variables and other technical stuff.

2 Methodology

In the following we briefly expose the methodology used to measure the distributional impact of each explanatory variables and refer to Fiaschi et al. (2012b) for

[2]These results broadly support the presence of parameter heterogeneity as discussed in Durlauf et al. (2005) and Durlauf et al. (2009).

more technical details. The methodology is based on six steps: i) estimation of a semiparametric growth regression model (Section 2.1); ii) calculation of counterfactual productivity (Section 2.2.1); iv) estimation of counterfactual stochastic kernels (Section 2.2.1); v) estimation of counterfactual ergodic distributions (Section 2.2.1); vi) evaluation of the distributional effects of a variable and estimation of its marginal growth effect (Section 2.2.2); ii) test on the distributional effects of growth residuals (Section 2.3).

2.1 Modeling Productivity Growth

Assume there exist N regions, and define by $y_i(t)$ labor productivity of country i at time t. labor productivity of region i at time $T > 0$, therefore, can be expressed as:

$$y_i(T) = y_i(0)e^{g_i T}, \tag{1}$$

where g_i is the annual rate of growth of productivity in country i, between periods 0 and T.

Assume that g_i is a function of K explanatory variables, collected in vector $\mathbf{X}_i = (X_{i,1}, ..., X_{i,K})$, and of a residual component υ_i accounting for unobservable factors, that is:

$$g_i = \varphi(\mathbf{X}_i, \upsilon_i). \tag{2}$$

Differently from other approaches to growth regressions, we model the growth rate g_i by a semiparametric model, that is:[3]

$$g_i = m(\mathbf{X}_i) + \upsilon_i = \alpha + \sum_{j=1}^{K} \mu_j(X_{i,j}) + \upsilon_i, \tag{3}$$

where α is a constant term, $\mu_j(\cdot)$ are one-dimensional nonparametric functions operating on each of the K elements of $\mathbf{X}_i$, and υ_i is an error term with the properties: $E(\upsilon_i|\mathbf{X}_i) = 0$, $var(\upsilon_i|\mathbf{X}_i) = \sigma^2(\mathbf{X}_i)$ (i. e. the model allows for heteroskedasticity).

2.2 Distributional Effects of Individual Variables

Denote by $\mathbf{X}_{i,\underline{k}}$ the vector of all explanatory variables but $X_{i,k}$ for country i, i. e.:

$$\mathbf{X}_{i,\underline{k}} = (X_{i,1}, ..., X_{i,(k-1)}, X_{i,(k+1)}, ..., X_{i,K})$$

[3]Notation refers to Härdle et al. (2004).

Eq. (3) can be rewritten as:

$$g_i = \alpha + \mu_k(X_{i,k}) + \sum_{j \neq k} \mu_j(X_{i,j}) + \upsilon_i. \tag{4}$$

Substituting Eq. (4) into Eq. (1) leads to the following expression for productivity:

$$\begin{aligned} y_i(T) &= y_i(0)e^{[\alpha+\mu_k(X_{i,k})+\sum_{j\neq k}\mu_j(X_{i,j})+\upsilon_i]T} = \\ &= \underbrace{y_i(0)e^{[\alpha+\sum_{j\neq k}\mu_j(X_{i,j})]T}}_{y_{i,\underline{k}}(T)}\underbrace{e^{\mu_k(X_{i,k})T}}_{e^{g^M_{i,k}T}}\underbrace{e^{\upsilon_i T}}_{e^{g^R_i T}}, \end{aligned} \tag{5}$$

where $y_{i,\underline{k}}(T) = y_i(0)e^{[\alpha+\sum_{j\neq k}\mu_j(X_{i,j})]T}$ is the level of productivity in period T obtained by "factoring out" the effect of $X_{i,k}$; $g^M_{i,k} = \mu_k(X_{i,k})$ is the part of the annual growth rate of y_i explained by $X_{i,k}$, capturing the "marginal" effect of $X_{i,k}$ on g_i and, finally, $g^R_i = \upsilon_i$ is the annual "residual growth", not explained by the variables in $\mathbf{X}_i$. The modelling of growth in Eq. (5) will be the basis for the identification of the distributional effects of the *k-th* variable.

2.2.1 Counterfactual Stochastic Kernels and Ergodic Distributions

We define the *counterfactual productivity* $y^{CF}_{i,k}(T)$, the productivity level that a country would attain at time T if there were no differences within the sample in terms of the *k-th* variable (whose values are collected in the N-dimensional vector $\mathbf{X}_k$). That is, $y^{CF}_{i,k}(T)$ aims at capturing the effect on the productivity distribution of the cross-sectional distribution of the *k-th* variable. To isolate this effect, we will impose to each country the *cross-section average value* of the variable.

Hence, the *counterfactual growth rate* of country i with respect to the k-th variable, $g^{CF}_{i,k}$, is defined as:

$$g^{CF}_{i,k} \equiv \alpha + \sum_{j\neq k} \mu_j(X_{i,j}) + \mu_k(\bar{X}_k) + \upsilon_i, \tag{6}$$

where $\bar{X}_k = N^{-1}\sum_{j=1}^{N} X_{k,j}$, and $\mu_k(\cdot)$ is the smoothed function relative to the *k-th* variable, obtained from the estimation of Eq. (3). The counterfactual productivity of country i in period T, relative to variable k, is therefore defined as:

$$y^{CF}_{i,k}(T) \equiv y_i(0)e^{g^{CF}_{i,k}T} = y_i(0)e^{[\alpha+\sum_{j\neq k}\mu_j(X_{i,j})+\mu_k(\bar{X}_k)+\upsilon_i]T}. \tag{7}$$

Counterfactual productivities are the bases to compute *counterfactual stochastic kernels*. Specifically, the *actual* and *counterfactual* stochastic kernels are respectively defined as $\phi(\mathbf{y}(T)|\mathbf{y}(0))$ and $\phi^{CF}(\mathbf{y}_k^{CF}(T)|\mathbf{y}(0))$, where $\mathbf{y}(0)$, $\mathbf{y}(T)$ and $\mathbf{y}_k^{CF}(T)$ are the vectors collecting individual productivities at times 0 and T.[4]

The actual stochastic kernel $\phi(\cdot)$ maps the distribution of (relative) productivity in period 0 into the distribution of (relative) productivity in period T. The counterfactual stochastic kernel $\phi^{CF}(\cdot)$, instead, maps the distribution of (relative) productivity in period 0, into the distribution of counterfactual relative productivities in period T. Therefore, the counterfactual stochastic kernel highlights, for every initial productivity level, the probability distribution over productivity levels at time T if the cross-country heterogeneity in the variable k is suppressed. This implies that the possible differences with respect to the probability distribution based on the actual stochastic kernel depends on the *k-th* variable, in particular on its distribution across country.

For actual and counterfactual stochastic kernels we estimate the corresponding ergodic distributions, i.e. the *actual* and the *counterfactual ergodic distribution*, following the procedure proposed by Johnson (2005).[5] The ergodic distribution highlights whether the estimated distribution dynamics over the period of interest has completely exhausted its effects or, otherwise, significant distributional changes are expected in the future.

2.2.2 The Distributional Effect of Individual Variables and the Marginal Growth Effect

The distributional effect of a variable is evaluated in terms of two aspects: i) its capacity to make actual and counterfactual stochastic kernels different; ii) its *marginal growth effect* conditioned to initial productivity, which highlights if a variable is a source of convergence or divergence and which part of the distribution of productivity is particularly affected by the variable.

As regards the differences between between actual and counterfactual kernels,

[4]In general, a stochastic kernel is an operator mapping the density of a variable at time t into its density at time $t+\tau$, $\tau>0$, and indicates for each level of the variable in period t its the probability distribution in period $t+\tau$. That is, the relation between the densities and the stochastic kernel is: $f_{t+\tau}(z)=\int_0^\infty g_\tau(z|x)\,f_t(x)\,dx$, where z and x are two levels of the variable, and $g_\tau(z|x)$ is the stochastic kernel. To estimate the stochastic kernel $g_\tau(z|x)=g(z,x)/f(x)$ we estimated the joint density of z and x, $g(z,x)$, and the marginal density of x, $f(x)$. In the estimation of $g(z,x)$ we followed Johnson (2005), who used the *adaptive kernel estimator* discussed by (Silverman, 1986, p. 100), in which the window of the kernel (Gaussian in our case) increases when the density of observations decreases.

[5]Specifically, the ergodic distribution solves $f_\infty(z)=\int_0^\infty g_\tau(z|x)\,f_\infty(x)\,dx$.

consider the value of (log) actual productivity in period T, $y_i(T)$, as a function of the counterfactual productivity, $y_{i,k}^{CF}(T)$:

$$\log\left(y_i(T)\right) = \log\left(y_{i,k}^{CF}(T)\right) + \left[\mu_k(X_{i,k}) - \mu_k(\bar{X}_k)\right]T + \upsilon_i T. \tag{8}$$

Taking the expected value of $\log\left(y_i(T)\right)$ conditional to $y_i(0)$, we get:

$$E\left[\log\left(y_i(T)\right)|y_i(0)\right] = E[\log(y_{i,k}^{CF}(T))|y_i(0)] + E[\mu_k(X_{i,k}) - \mu_k(\bar{X}_k)|y_i(0)]T, \tag{9}$$

which provides a condition for the equality of the expected values of productivity based on actual and counterfactual kernels, i.e. for the absence of distributional effect of the $k-th$ variable:

$$E[\mu_k(X_{i,k})|y_i(0)] = \mu_k(\bar{X}_k). \tag{10}$$

The result in Eq. (10) depends on the fulfilment of the following two conditions:

1. $E[\mu_k(X_{i,k})|y_i(0)] = E[\mu_k(X_{i,k})]$, i. e. $\mu_k(X_{i,k})$ and $y_i(0)$ are independent, that is the impact of the *k-th* variable on productivity in country i is independent from the initial productivity level.

2. $E[\mu_k(X_{i,k})] = \mu_k(E[X_{i,k}]) = \mu_k(\bar{X})$, i. e. $\mu_k(\cdot) = \beta_k X_{i,k}$, that is the *marginal* impact of the *k-th* variable is constant, i.e. the term $X_{i,k}$ has a linear effect on growth.

Condition in Eq. (10) therefore implies:

$$E[\mu_k(X_{i,k})|y_i(0)] = E[\mu_k(X_{i,k})] = \mu_k(E[X_{i,k}]) = \mu_k(\bar{X}_k). \tag{11}$$

The use of a semiparametric specification is therefore necessary to identify possible differences between the actual and counterfactual stochastic kernels, even when the marginal effect of the *k-th* variable is independent of the initial productivity level (i.e. when the condition $E[\mu_k(X_{i,k})|y_i(0)] = E[\mu_k(X_{i,k})]$ is fulfilled).

As regards the *marginal growth effect* of the *k-th* variable, this is defined by $g_{i,k}^M = \mu_k(X_{i,k})$ in Eqq. (3)-(5).

The marginal effect of the *k-th* variable on the distribution dynamics is identified by estimating the marginal growth $\mathbf{g}_k^M$ conditioned on the initial level of productivity, i. e. by estimating $\phi^M(\mathbf{g}_k^M|\mathbf{y}(0))$. If the estimate of the marginal effect does not result statistically different from its unconditional mean, i. e. $\phi^M(\mathbf{g}_k^M|\mathbf{y}(0)) = E[\mathbf{g}_k^M]\ \forall \mathbf{y}(0)$, then the *k-th* variable has no distributional effects. On the contrary, if $\phi^M(\mathbf{g}_k^M|\mathbf{y}(0))$

is statistically different from its unconditional mean and, in particular, it is an increasing (decreasing) function of $\mathbf{y}(0)$, then the *k-th* variable is a source of divergence (convergence).

Since the estimation of the marginal effect in semiparametric models is performed through the backfitting technique, it requires as identification assumption that: $E_{\mathbf{X}_k}[\mu_k(\mathbf{X}_k)] = 0$ (see Härdle et al., 2004, pp. 212-222). Therefore, the unconditional mean of marginal growth will always be equal to zero in the estimation of the semiparametric terms in the growth regression.

2.3 Test of Distributional Effects of Residual Growth

Fiaschi et al. (2012a) discuss as a final step a test for the goodness of fit, i. e. for the presence of possible misspecification in the model for some ranges of initial productivity. In particular, Eq. (5) suggests to consider $\hat{\mathbf{g}}^R$, defined as $\hat{\mathbf{g}}^R \equiv log\left(\frac{\mathbf{y}(T)}{\hat{\mathbf{y}}(T)}\right)$, to test that:

$$E[\hat{\mathbf{g}}^R|\mathbf{y}(0)] = E[\hat{\mathbf{g}}^R] = 0 \; \forall \; \mathbf{y}(0). \tag{12}$$

If $\mathbf{y}(0)$ is included in the set of regressors, the condition in Eq. (12) ensures that there is no omitted variable inconsistency related to $\mathbf{y}(0)$ (see Wooldridge, 2002, pp. 61-63).

3 Empirical Analysis

Our sample is composed by 84 countries.[6]. The estimates of stochastic kernel (SK) with lag equal to 15 years reported in Figure 1 and of the actual distribution (AD) reported in Figure 2 point out to the presence of two peaks in the distribution around 0.3 and 2.7 in 2008.[7]

[6]Country list is in Appendix A.

[7]All estimation are made by R Development Core Team (2012). Codes and dataset are available on Davide Fiaschi's website (`http://dse.ec.unipi.it/~fiaschi/WorkingPapers.html`)

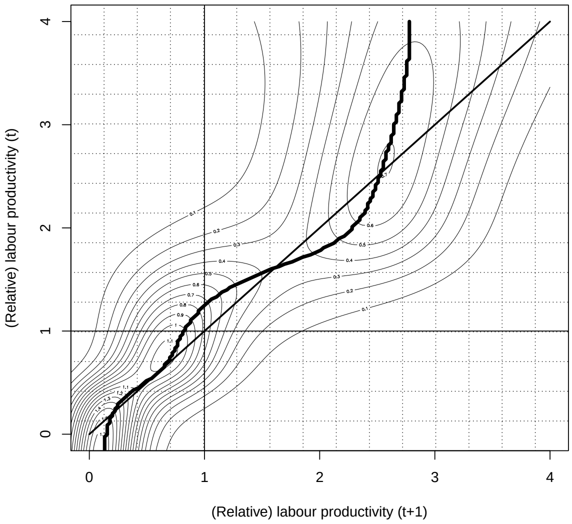

Figure 1: SK of productivity, the median of SK (thick line).

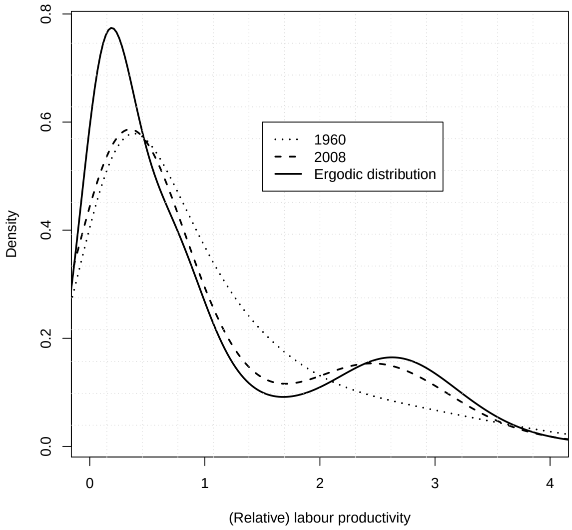

Figure 2: 1960 (dotted line), 2006 (solid line) and ergodic (dashed line) distributions of productivity

The comparison between the AD and the ergodic distribution (ED) suggests that the distribution dynamics has not still reached its equilibrium in 2008 and dispersion and polarization should increase in the future.[8] In each figure displaying the estimate of the stochastic kernel we report also a solid line representing the estimated median value of productivity at $t+\tau$ conditioned on the productivity level at time t and the 45° line. The median provide crucial information on the possible emergence of polarization: it can be observed how the peaks of ED are in corre-

[8]In general, stochastic kernels indicate for each level of productivity in period t its probability distribution in period $t+\tau$, $\tau > 0$. For each stochastic kernel it is possible to estimate the corresponding ergodic distribution following the procedure proposed by Johnson (2005). Specifically, the ergodic distribution solves $f_\infty(z) = \int_0^\infty g_\tau(z|x) f_\infty(x)\,dx$, where z and x are two levels of the variable, $g_\tau(z|x)$ is the density of z, given x, τ periods ahead. The ergodic distribution highlights whether the estimated distribution dynamics over the period of interest has completely exhausted its effects or, otherwise, significant distributional changes are expected in the future. To estimate $g_\tau(z|x) = g(z,x)/f(x)$, the stochastic kernel, we estimated the joint density of z and x, $g(z,x)$, and the marginal density of x, $f(x)$. In the estimation of $g(z,x)$ we followed Johnson (2005), who used the *adaptive kernel estimator* discussed by (Silverman, 1986, p. 100), in which the window of the kernel (Gaussian in our case) increases when the density of observations decreases. Here we adjust the estimate of ergodic distributions for the use of normalized variables (with respect to the average). See D. and Romanelli (2009).

spondence with the points where the median crosses the bisector from below.

Table 1 shows how the dispersion in the distribution of labour productivity has increased from 1960 to 2008 of about 4 base points in terms of Gini index and is doomed to increase in the long run of further 1 base points according to the estimated ergodic distribution reported in Figure 2.

	1960	2008	Ergodic
Gini	0.49.8	0.529	0.539
s.e.	(0.025)	(0.028)	(0.151)
Test of unimodality	0.56	0.0	
Test of bimodality	0.66	0.69	

Table 1: Gini index for 1960, 2008 and for the ergodic distribution (standard errors calculated by 500 bootstraps). P-values of the tests of unimodality and bimodality for 1960 and 2008.

Tests of multimodality reported in Table 1 state that the null hypothesis of unimodality and bimodality cannot be rejected for 1960 at conventional statistical levels, while the null hypothesis of unimodality can be rejected to 0.001 confidence level (the null hypothesis of bimodality in 2008 cannot be rejected).[9] This confirms the visual impression that in 1960 the distribution was unimodal and becomes bimodal in 2008.

Two main results on the distribution dynamics of the labour productivity can be drawn: i) the dispersion increased from 1960 to 2008 as showed by the Gini indices; ii) the polarization also increased from 1960 to 2008 as showed by the estimate of SK and the unimodality tests.

3.1 The Estimate of Growth Model

For each country we consider the average growth rate of labour productivity in the period 1960-2008 ($AV.PROD.GR$), the (log of) labor productivity in 1960 ($LOG.PROD.1960$), the (log of) average investment rates ($LOG.AV.INV.RATE$), the (log of) average growth rate of employment ($LOG.AV.EMPLOYMENT.GR$), the (log of) the share of population aged 15 or over with complete secondary education ($LOG.AV.SEC.EDU$).[10] The

[9] Tests of multimodality follow the bootstrap procedure described in Silverman (1986, p. 146) and are performed using 1000 bootstraps.

[10] All variables are drawn from Penn World Table 7.1 (`http://pwt.econ.upenn.edu/`), except for the share of population aged 15 or over with complete secondary education taken from Cohen and Soto

size of sample is limited by the available data on the stock of human capital. Appendix B contains some descriptive statistics of the variables used in the analysis.

This information set is sufficient to estimate an augmented Solow model as proposed by MRW with the inclusion of dummies ($DUMMIES$) aiming at capturing possible omitted variables (see discussion below for more details). In its semiparametric version the growth regression is given by:

$$AV.PROD.GR_i = CONST + \gamma DUMMIES + \mu_1 (LOG.PROD.1960_i) + \mu_2 (LOG.AV.INV.RATE_i) + \\ + \mu_3 (LOG.AV.SEC.EDU_i) + \mu_4 (LOG.AV.EMPLOYMENT.GR_i) + \epsilon_i, \quad (13)$$

where $CONST$ ia a constant, γ is the vector of coefficients of dummies, μ_k is a smooth component, and ϵ_i is a random component

Durlauf et al. (2005) contain a very detailed derivation of Eq. (13) from the original augumented Solow model via a first-order approximation of growth rate of labor productivity around the steady state; (Durlauf et al., 2005, p. 592) also claim that this approximation is reasonably accurate.[11] Their derivation is easily adapted to include nonlinearities in the explanatory variables and dummies. In particular, i) the presence of nonlinearities can derive from adopting a more general production function with respect to the Cobb-Douglas case which makes the elasticity of the equilibrium level of income with respect to investment rate, the growth rate of employment (augmented by depreciation rate and growth rate of technology), and the stock of human capital not constant (the constant elasticity is exactly the result from assuming a Cobb-Douglas production function); and ii) the presence of dummies can derive from differences across countries in initial levels of technology not

(2007), the best quality dataset on education of countries to our knowledge.

[11]In particular, (Durlauf et al., 2005, p. 578,) derive the following growth equation for country i:

$$\gamma_i = g - \beta \log A_{i,0} - \beta \log y^E_{i,\infty} + \beta \log y_{i,0} + v_i,$$

where γ_i is the growth rate of labor productivity of country i, g the (common) growth rate of technological progress, β a coefficient, $y^E_{i,\infty}$ the equilibrium level of income per efficient units of worker, $A_{i,0}$ the initial level of technology of country i, $y_{i,0}$ the initial level of labor productivity of country i and v_i a stochastic term. Adopting the MRW model as explanation of the equilibrium level of income per (efficient units of) worker, $y^E_{i,\infty} = f_i\left(k^E_{i,\infty}, h^E_{i,\infty}\right)$, where $k^E_{i,\infty}$ and $h^E_{i,\infty}$ are the equilibrium level of physical and human capital per efficient units of labor and f_i the intensive production function. Finally, $k^E_{i,\infty}$ is a positive function of the investment rate and the stock of human capital and a negative function of growth rate of employment of country i, being implicitly defined by $s_{K,i} f_i\left(k^E_{i,\infty}, h^E_{i,\infty}\right) = (n_i + g + \delta)\, k^E_{i,\infty}$, where $s_{K,i}$ is the investment rate and n_i the growth rate of employment of country i and δ the depreciation rate of physical capital.

accounted by their initial levels of labour productivity or in their production function, or all omitted variables (e.g. different institutions and geographical factors) constant over the period.

3.1.1 The Selection of Dummies and the Estimate of Best Growth Model

The estimate of Eq. (13) is conditioned to the choice of dummies. In literature it is very common to include in growth regressions geographical dummies, as for example dummies for sub-saharan countries , for countries in tropical areas, or for oil producers, in order to control for non directly observable characteristics (see Durlauf et al. (2005)) . Here we take a different route. Our aim is to get a parsimonious but at the same time a high goodness-of-fit model. A crucial-side effect of our approach is to have a measure of "our ignorance" in the explanation of countries' growth, which reflects in a measure of unexplained inequality in the cross-country distribution of labour productivity.

The parsimony of the model follows the philosophy of the *Occam's Razor*, i.e. to choice the simplest model in the candidate collection that adequately accommodates the data. From the statistical perspective we should have a model where the impact of each variables is more easily understood and with higher accuracy in the estimation of parameters. However, a too restrict set of regressors makes more likely to incur in a omitted variable bias and therefore the choice of suitable dummies can help to protect against this bias. The selection of appropriate dummies is made by the corrected Akaike Information Criterion, denoted $AICc$ (see Hurvich and Tsai (1989)), which represents an improvement of the usual Akaike Information Criterion for small samples (for a very detailed discussion on model selection for non nested model in small sample see Burnham and Anderson (2002)).

In the estimate we have considered the cases both with two and three dummies (we stopped to the three-dummy model given the excellent goodness-of-fit we have got, see Table 2), corresponding to the case in which there are two or three clusters of countries with the same unobservable characteristics respectively. For each case (e.g. the three-dummy case), we have searched the combination of dummies providing the best goodness of fit measured by AICc (the set of possible combinations for the three-dummy case is 3^8).[12] Model III in Table 2 reports the estimate of the

[12] The search for the optimal combination of dummies was performed by genetic algorithms, an optimization technique well suited for problems where the space of parameters is discrete and local minima are present (see Goldberg (1989)).

best specification for the three-dummy case, while the best combination of the three dummies is reported in Appendix A. The semiparametric estimation is made following the approach describe in Wood (2006) based on penalized regression splines.[13]

The estimated degree of freedom *EDF* is a measure of the nonlinearities of the impact of the variable; an *EDF* equal to one means that the impact is linear (see Wood (2006)). For comparison in Table 2 we report also the standard augumented Solow model (Model I), the standard augumented Solow model estimated by a semiparametric specification (only $LOG.PROD.1960$ appears to have a nonlinear impact on $AV.PROD.GR$).

Dependent variable: $AV.PROD.GR$			
Model	I	II	III
Endogeneity	YES	YES	YES
$CONST$	-0.0312	0.0245	
$DUMMY1$			0.0243***
$DUMMY2$			0.0152***
$DUMMY3$			0.0048***
$LOG.PROD.1960$	-0.0075***	-0.072***	(**3.20**)
$LOG.AV.INV.RATE$	0.0126**	0.0171***	(**7.41**)
$LOG.AV.SEC.EDU$	0.0085***	0.0068***	(**1.00**)
$LOG.AV.EMPLOYMENT.GR$	-0.0319***	-0.0345***	(**6.41**)
Num. obs	84	84	84
$\bar{R}^2$	0.604	0.60	0.971
AIC_c	-555.984	-554.898	-743.074

Table 2: Results of estimation of different growth models. For the Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 0. Coefficients in brackets (and in bold) are the estimated degree of freedom (*EDF*).

All coefficients have the expected sign in all three models, but the goodness of fit measured both by $\bar{R}^2$ and AIC_c greatly improves from the first to the third model. Not reported estimate of semiparametric model with two dummies shows a substantial improvement in adding a dummy (in the model with two dummies $\bar{R}^2 = 0.79$ and

[13]In particular, we use *mgcv* packages in R Development Core Team (2012), with the option "REML" discussed in Wood (2011).

$AIC_c = -610.54$), while no substantial improvement derives from adding a fourth dummy.

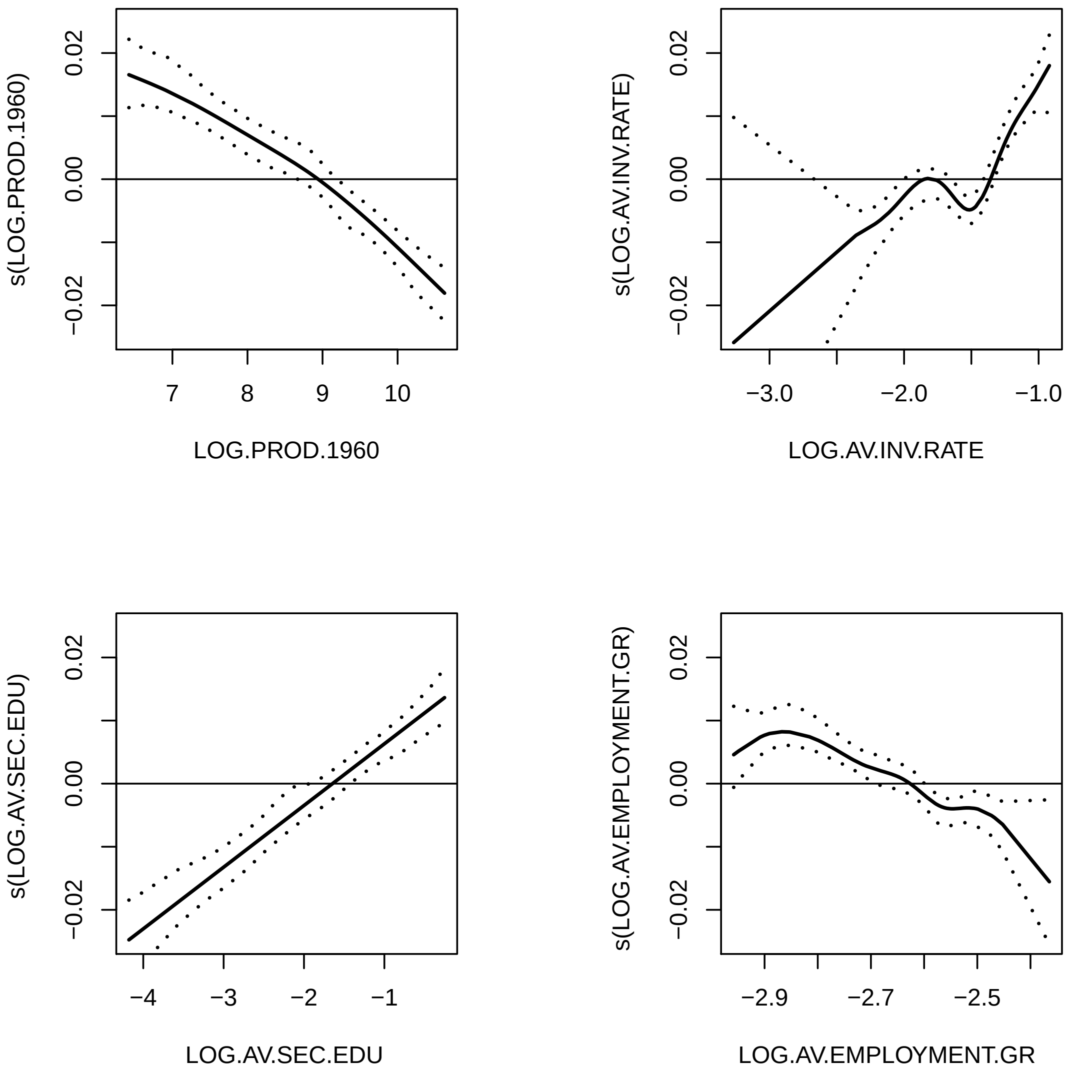

Figure 3: The estimated impact on growth of variables included in Model III in Table 2 and relative confidence bands.

Figure 3 reports the impact on growth of explanatory variables included in Model III and the relative confidence bands. It is evident the strong nonlinearity of the impact that, as we discussed in Section 2, can have a crucial role in the explanation of the distribution of labour productivity. Our findings are in line with the current literature on the presence of parameter heterogeneity in cross-country regressions (see

(Durlauf et al., 2005, p. 616)). For the period 1960-1990 Liu and Stengos (1999) found a similar nonlinear impact on growth of labor productivity in 1960 and of the stock of human capital (compare Figures 1 and 2 in Liu and Stengos (1999) with the estimated impacts of $LOG.PROD.1960$ and $LOG.AV.SEC.EDU$ reported in Figure 3).

We remark how the inverted U-shaped relationship for $LOG.PROD.1960$ is in contrast with the hypothesis of conditional convergence (at least for the increasing part), the investment rate appears particularly effective at high levels, as well as the stock of human capital, while the growth rate of employment seems to exert the higher impact in the middle range.[14]

Dummies do not have not any clear geographical pattern (for example there is not any dummy identifying sub-Saharan countries); moreover, their magnitude is high (about 2.3% of difference in terms of average growth rates) pointing out the very remarkable impact of impact of the omitted variable in the standard augmented Solow model.

Figures 4 and 5 shows how the estimated Model III reproduce almost entirely both the observed growth rates and the distribution in 2008 starting from the distribution in 1960.

[14] It is out of the scope of the paper to test of possible complementaries among explanatory variables, that it can be the case with a CES production function; or to test the presence of parameter heterogeneity related to different regimes, as in Durlauf et al. (2001).

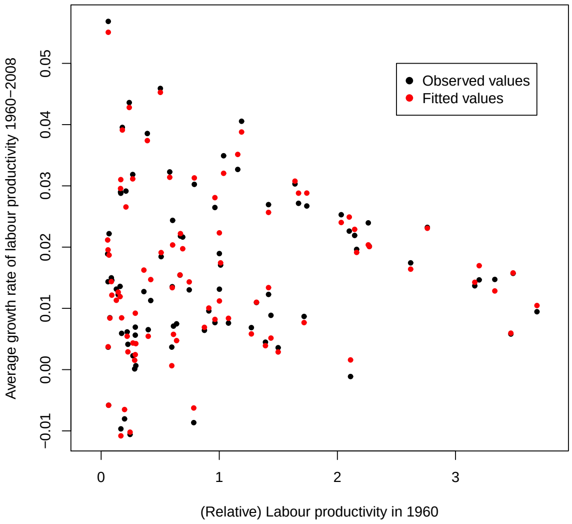

Figure 4: Observed vs fitted growth rate of productivity in 1960-2008 from Model III versus the (relative) labour productivity in 1960.

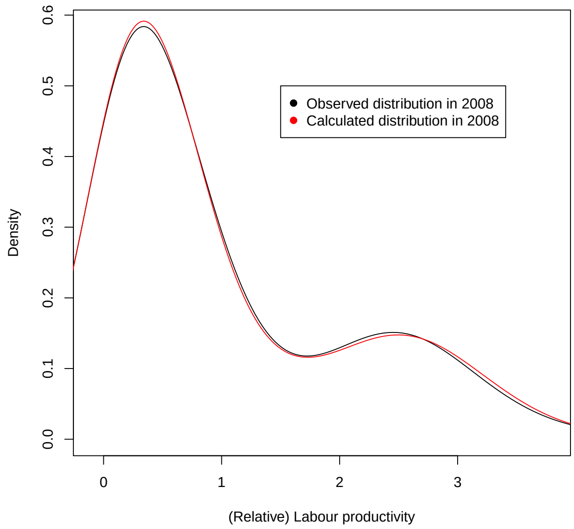

Figure 5: Observed vs calculated distribution of labour productivity in 2008 from Model III

Finally, Figure 6 reports how the estimate of Model III passes the test on the possible presence of distributional effects in the residual growth discussed in Section 2.3. The estimate therefore appears well-specified to our scope.

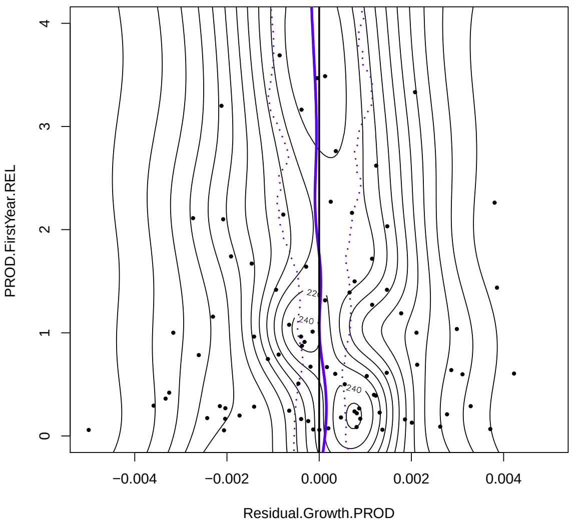

Figure 6: Test for the possible presence of distributional effects in the residual growth. Confidence bands derive from a bootstrap procedure with 300 replications.

3.2 Distributional Impact of Explanatory Variables

Figures 7-26 reports the detailed estimated of marginal growth effect (MGE), counterfactual productivity in 2008 (CD), counterfactual stochastic kernel (CSK) and counterfacutal ergodic distribution (CED) described in Section 2, while and Tables 3-7 reported the relative Gini indexes and tests of multimodality.

3.2.1 Initial labour Productivity

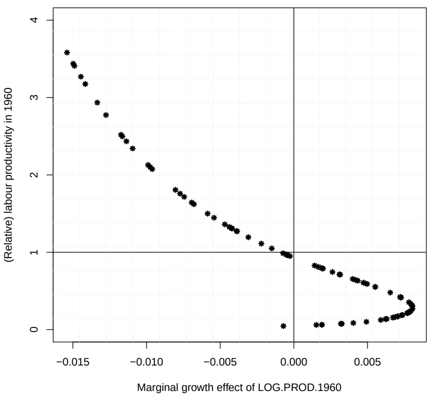

Figure 7: MGE of labour productivity in 1960.

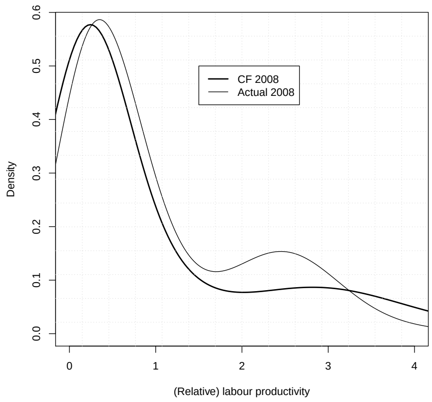

Figure 8: AD in 2008 (thin line), CD in 2008 (thick line). Counterfactual variable: labour productivity in 1960.

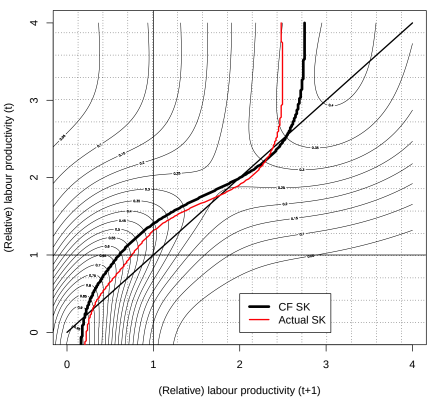

Figure 9: CSK, the median of SK (thick line) and ASK (red line). Counterfactual variable: labour productivity in 1960.

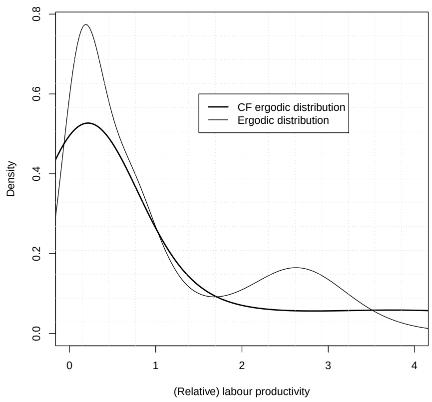

Figure 10: AED (thin line) and CED (thick line). Counterfactual variable: labour productivity in 1960.

	1960	2008	2008 CF	Ergodic	Ergodic.CF
Gini	0.498	0.529	0.628	0.539	0.626
s.e.	(0.025)	(0.028)	(0.027)	(0.151)	(0.615)
Test of unimodality	0.56	0.0	0.01		
Test of bimodality	0.66	0.69	0.33		

Table 3: Gini Indexes (and their standard errors) and tests of multimodality of AD, CD, AED and CED. Counterfactual variable: labour productivity in 1960.

3.2.2 Investment Rate

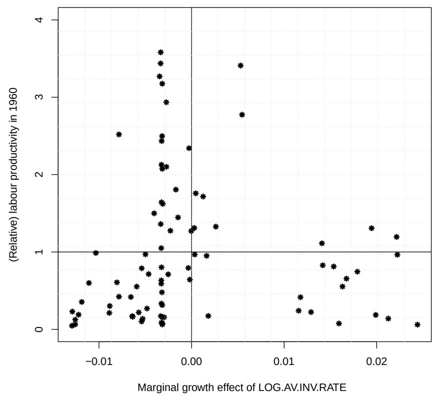

Figure 11: MGE of (log of) average investment rate.

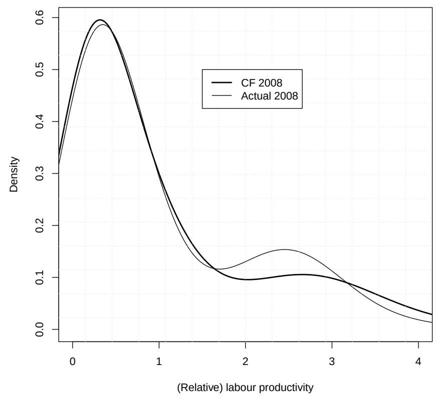

Figure 12: AD in 2008 (thin line), CD in 2008 (thick line). Counterfactual variable: (log of) average investment rate.

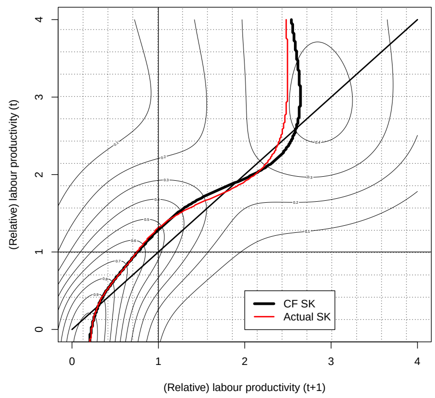

Figure 13: CSK, the median of SK (thick line) and ASK (red line). Counterfactual variable: (log of) average investment rate.

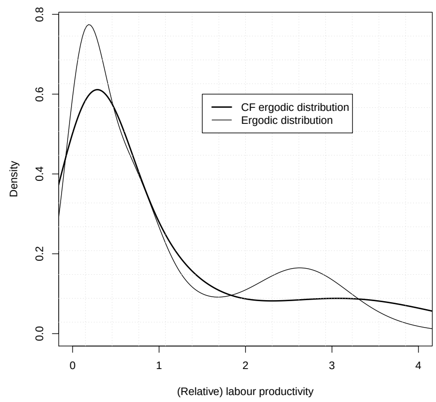

Figure 14: AED (thin line) and CED (thick line). Counterfactual variable:(log of) average investment rate.

	1960	2008	2008 CF	Ergodic	Ergodic.CF
Gini	0.498	0.529	0.551	0.539	0.572
s.e.	(0.025)	(0.028)	(0.025)	(0.151)	(0.467)
Test of unimodality	0.56	0.00	0.01		
Test of bimodality	0.66	0.69	0.80		

Table 4: Gini Indexes (and their standard errors) and tests of multimodality of AD, CD, AED and CED. Counterfactual variable: average investment rate.

3.2.3 Human Capital

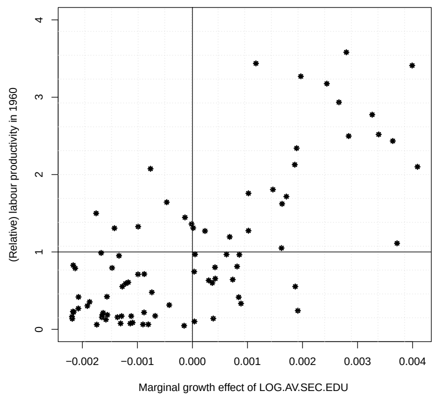

Figure 15: MGE of (log of) average human capital.

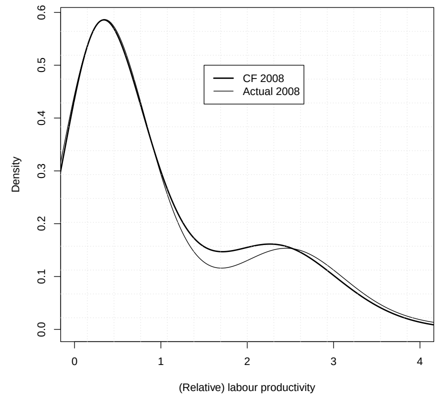

Figure 16: AD in 2008 (thin line), CD in 2008 (thick line). Counterfactual variable: (log of) average human capital.

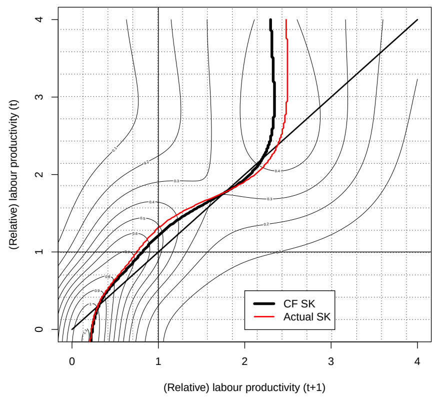

Figure 17: CSK, the median of SK (thick line) and ASK (red line). Counterfactual variable: (log of) average human capital.

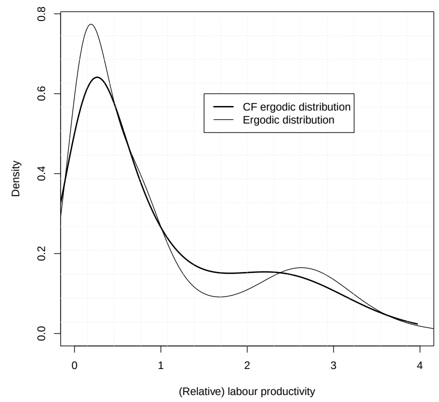

Figure 18: AED (thin line) and CED (thick line). Counterfactual variable:(log of) average human capital.

	1960	2008	2008 CF	Ergodic	Ergodic.CF
Gini	0.498	0.529	0.511	0.539	0.528
s.e.	(0.025)	(0.028)	(0.026)	(0.151)	(0.389)
Test of unimodality	0.56	0.00	0.01		
Test of bimodality	0.66	0.69	0.64		

Table 5: Gini Indexes (and their standard errors) and tests of multimodality of AD, CD, AED and CED. Counterfactual variable: average human capital.

3.2.4 Employment Growth

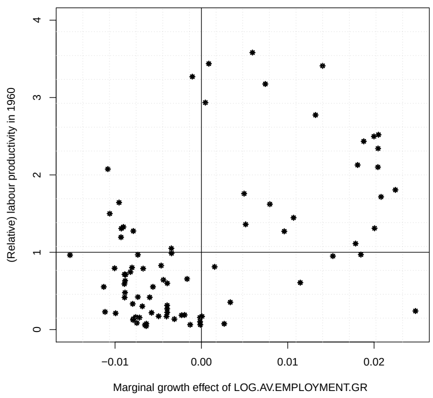

Figure 19: MGE of (log of) average growth rate of employment.

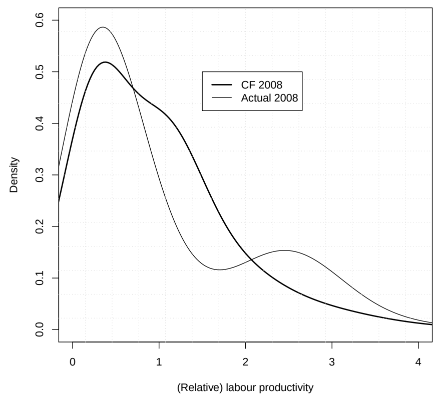

Figure 20: AD in 2008 (thin line), CD in 2008 (thick line). Counterfactual variable: (log of) average growth rate of employment.

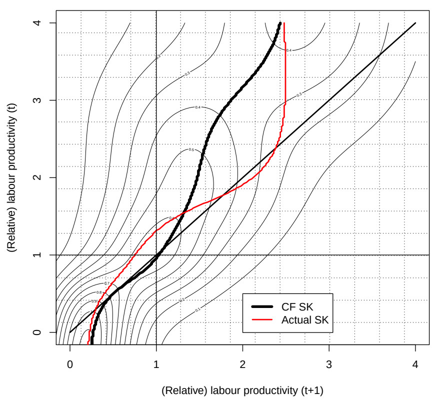

Figure 21: CSK, the median of SK (thick line) and ASK (red line). Counterfactual variable: (log of) average growth rate of employment.

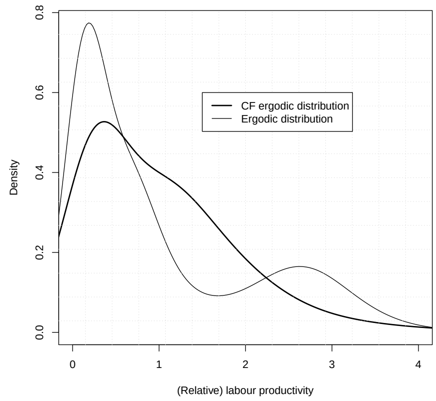

Figure 22: AED (thin line) and CED (thick line). Counterfactual variable:(log of) average growth rate of employment.

	1960	2008	2008 CF	Ergodic	Ergodic.CF
Gini	0.498	0.529	0.469	0.539	0.455
s.e.	(0.025)	(0.028)	(0.032)	(0.151)	(0.196)
Test of unimodality	0.56	0.00	0.03		
Test of bimodality	0.66	0.69	0.23		

Table 6: Gini Indexes (and their standard errors) and tests of multimodality of AD, CD, AED and CED. Counterfactual variable: average growth rate of employment.

3.2.5 Dummies

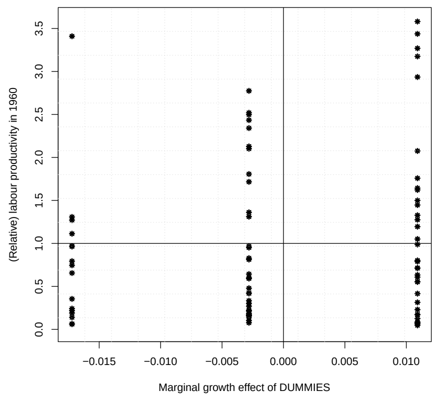

Figure 23: MGE of dummies.

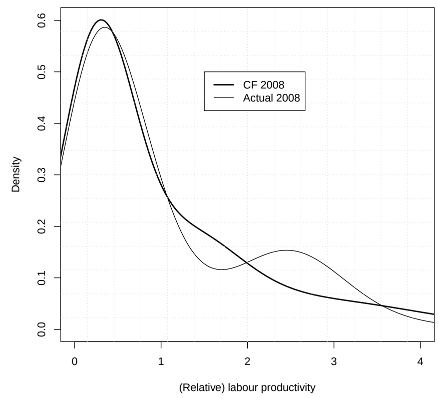

Figure 24: AD in 2008 (thin line), CD in 2008 (thick line). Counterfactual variable: dummies.

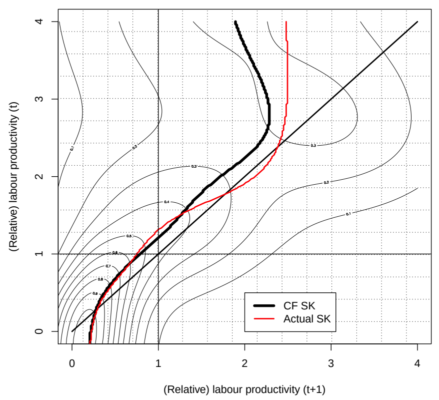

Figure 25: CSK, the median of SK (thick line) and ASK (red line). Counterfactual variable: dummies.

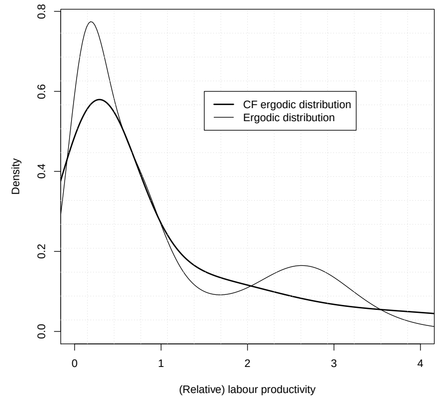

Figure 26: AED (thin line) and CED (thick line). Counterfactual variable: dummies.

	1960	2008	2008 CF	Ergodic	Ergodic.CF
Gini	0.498	0.529	0.556	0.539	0.577
s.e.	(0.025)	(0.028)	(0.03)	(0.151)	(0.400)
Test of unimodality	0.56	0.00	0.237		
Test of bimodality	0.66	0.69	0.187		

Table 7: Gini Indexes (and their standard errors) and tests of multimodality of AD, CD, AED and CED. Counterfactual variable: dummies.

Table 8 summarizes the impact on distribution in terms of dispersion and polarization based on the estimate of Model III reported in Table 2.

	Dispersion		**Polarization**	
	2008	Ergodic	2008	Ergodic
Actual distribution	0.529	0.539	Bimodal	Bimodal
Variable	Δ **Gini**		**Change in polarization**	
$LOG.PROD.1960$	-0.098	-0.087	- -	- -
$LOG.AV.INV.RATE$	-0.021	-0.033	-	-
$LOG.AV.SEC.EDU$	0.019	0.012	+	+
$LOG.AV.EMPLOYMENT.GR$	0.061	0.084	++	++
$DUMMIES$	-0.027	-0.037	+++	+++

Table 8: Summary of the distributional impact of variables based on the estimate of Model III reported in Table 2. Differences between Gini indexes of CD and CED versus AD and ED respectively, and the effect on polarization based on the estimate of CSK and tests of multimodality.

In particular, Table 8 reports the differences between Gini indexes of CD and CED and AD and ED respectively; a negative sign (as in the case of $LOG.PROD.1960$) means that that variable has contributed to reduce dispersion. For polarization the intensity of impact is valued in terms of the number of points (number of peaks) and distance between points (distance between peaks) where the estimated conditional median cross bisector from below in CSK and the test of multimodality for counterfactual distributions; for example two "-" ("+") for $LOG.PROD.1960$ means that the variable has contributed to reduce (increase) to a large extent polarization.

The initial level of labour productivity had the largest negative impact on dispersion both in 2008 and in the long run (9.8 and 8.7 base points) and a moderate negative impact on polarization (the distance between the two crosses from below of estimated conditional is higher in CSK than in SK). At the opposite, the growth of employment has the largest positive impact both in 2008 and in the long run (6.1 and 8.4 base points) and a moderate positive impact on polarization (two crosses from below of estimated conditional median are still present in 2008).

Investment rates and the stock of human capital push in the opposite directions being investing rates source of divergence (both in terms of dispersion and polarization), while the stock of human capital a source of convergence. However, the magnitude of impact for both is lowest with respect to the other explanatory variables. Finally, *DUMMIES* have a low impact on dispersion but a extreme impact on polarization, leading to a CD with just one peak.

4 Concluding Remarks

We find that all (proximate) explanatory variables of the augmented Solow model have a statistically significant impact on labour distribution.

In agreement with Easterly and Levine (2001) (and in contrast with Mankiw et al. (1992) and Henderson and Russell (2005)) investment rates and human capital, two variables strictly related to factor accumulation, play a minor role in terms of both dispersion and polarization. Instead, initial levels of productivity, which in literature is considered a proxy for technological "catching up" and/or for diminishing returns to productive factors, play a substantial role in reducing both dispersion and inequality (a finding in contrast with Henderson and Russell (2005)). On the opposite, the growth rate of employment had the moderate and positive impact on both dispersion and polarization. Finally, DUMMIES, a measure of "our ignorance" of countries'growth, have the maximum positive impact both on polarization and just a minimum reducing role of dispersion; our "ignorance" on countries'growth is therefore not so important for explaining the dispersion of distribution of labour productivity, but, unfortunately, it is crucial to explain its polarization. This finding resemble the conclusion of Easterly and Levine (2001) for TFP (the measure of "our ignorance" in growth accounting), even though DUMMIES and TFP measure complete different characteristics of countries.

Our findings have clear policy implications, but a note of caution is needed in this respect. All the analysis is based on the implicit hypothesis of the absence of any endogenous relationship between the explanatory variables (and the dummies) included in the growth regression. Since the assumption does not hold in general, the policy implications of our finding must be particularly scrutinized. For example a policy aiming at reducing the growth rate of population (employment) should take into account that the level of per capita income could have an impact on fertility choice of individuals. However, still remain that a change in the growth rate of population (employment) represents one of the most effective policy to reduce inequality in countries'productivity according to our findings.

Another crucial note of caution is related to the fact that our results cannot be used to discriminate between alternative growth paradigms, e.g. institutions as fundamental explanation of development against the accumulation of human capital. One may be tempted to identify DUMMIES also as a proxy for the quality of institutions. However, a visual inspection of the different clusters of countries identified by DUMMIES reported in Appendix A makes clear that countries with very different quality of institutions belong to the same cluster (this could rule out also the interpretation of dummies as "residual technology progress"). This is not surprising given that, if the theory of the quality of institutions were right, investment rates, the stock of human capital, the growth rate of population (employment) and, overall, the labor productivity in 1960 would be endogenous variables reflecting the quality of institutions.

References

P. Burnham and D. R Anderson. *Model Selection and Multi-Model Inference*. Springer, Berlin, 2002.

D. Cohen and M. Soto. Growth and human capital: good data, good results. *Journal of Economic Growth*, 12(1):51–77, 2007.

Fiaschi D. and M. Romanelli. Nonlinear dynamics in welfare and the evolution of world inequality. Temi di discussione (Economic working papers) 724, Bank of Italy, Economic Research and International Relations Area, 2009. URL `http://ideas.repec.org/p/bdi/wptemi/td_724_09.html`.

S.N. Durlauf. The rise and fall of cross-country growth regressions. *History of Political Economy*, 41:315–333, 2009.

S.N. Durlauf, A. Kourtellos, and A. Minkin. The local solow growth model. *European Economic Review*, 45:928–940, 2001.

S.N. Durlauf, P. Johnson, and J. Temple. Growth econometrics. In S. N. Durlauf and P. Aghion, editors, *Handbook of Economic Growth*. Elsevier, 2005.

S.N. Durlauf, A. Kourtellos, and C. M. Tan. Are any growth theory robust? *The Economic Journal*, 118:329–346, 2009.

W. Easterly and R. Levine. It's not factor accumulation: Stylized facts and growth models. *World Bank Economic Review*, 15(2):177–219, 2001.

D. Fiaschi, A. M. Lavezzi, and A. Parenti. On the determinants of distribution dynamics: a new method and an application to eu regions. *Mimeo*, 2012a.

D. Fiaschi, A. M. Lavezzi, and A. Parenti. On the determinants of distribution dynamics: a new method and an application to a cross-section of countries. *Mimeo*, 2012b.

D. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, New York, 1989.

W. Härdle, M. Müller, S. Sperlich, and A. Werwatz. *Nonparametric and Semiparametric Models*. Springer, Berlin, 2004.

D. Henderson and R. Russell. Human capital and convergence: a production-frontier approach. *International Economic Review*, 4(46):1167–1205, 2005.

C.M Hurvich and C. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76:297–307, 1989.

P.A. Johnson. A continuous state space approach to convergence by parts. *Economic Letters*, 86:317–321, 2005.

Z. Liu and T. Stengos. Non-linearities in cross-country growth regressions: A semiparametric approach. *Journal of Applied Econometrics*, 14(5):527–538, 1999.

N.G. Mankiw, D. Romer, and D. Weil. A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 107:407–437, 1992.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL `http://www.R-project.org/`. ISBN 3-900051-07-0.

B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.

S. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 1(73):3–36, 2011.

S. N. Wood. *Generalized Additive Models. An Introduction with R*. Chapman and Hall, London, 2006.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data.* MIT press, 2002.

A Country List

CLUSTER 1		CLUSTER 2		CLUSTER 3	
CODE	**COUNTRY**	**CODE**	**COUNTRY**	**CODE**	**COUNTRY**
DZA	Algeria	AUS	Australia	ARG	Argentina
BEL	Belgium	BGD	Bangladesh	AUT	Austria
FRA	France	BDI	Burundi	BEN	Benin
JPN	Japan	CMR	Cameroon	BOL	Bolivia
JOR	Jordan	CAN	Canada	BRA	Brazil
MWI	Malawi	CAF	Central African Republic	BFA	Burkina Faso
MUS	Mauritius	CHL	Chile	CYP	Cyprus
NZL	New Zealand	CHN	China	DNK	Denmark
NIC	Nicaragua	COL	Colombia	ETH	Ethiopia
ROM	Romania	CRI	Costa Rica	FIN	Finland
THA	Thailand	CIV	Cote d'Ivoire	GAB	Gabon
TTO	Trinidad & Tobago	DOM	Dominican Republic	GRC	Greece
URY	Uruguay	ECU	Ecuador	HND	Honduras
ZWE	Zimbabwe	EGY	Egypt	IND	India
		SLV	El Salvador	ITA	Italy
		FJI	Fiji	JAM	Jamaica
		GHA	Ghana	KEN	Kenya
		GTM	Guatemala	MDG	Madagascar
		HTI	Haiti	MLI	Mali
		IDN	Indonesia	MAR	Morocco
		IRN	Iran	NPL	Nepal
		IRL	Ireland	NER	Niger
		KOR	Korea, Republic of	NGA	Nigeria
		MYS	Malaysia	NOR	Norway
		MEX	Mexico	PAN	Panama
		MOZ	Mozambique	PRY	Paraguay
		NLD	Netherlands	PER	Peru
		SEN	Senegal	PHL	Philippines
		SGP	Singapore	PRT	Portugal
		ZAF	South Africa	SWE	Sweden
		ESP	Spain	TUR	Turkey
		CHE	Switzerland	GBR	United Kingdom
		SYR	Syria	USA	United States of America
		TZA	Tanzania	VEN	Venezuela
		UGA	Uganda	ZMB	Zambia

Table 9: Countries in the three clusters

B Descriptive Statistics

	$AV.PROD.GR$	$PROD.1960$	$AV.INV.RATE$	$AV.SEC.EDU$	$AV.EMPL.GR$
Mean	0.02	10895	0.22	0.27	0.02
s.d.	0.01	10276	0.08	0.2	0.01

Table 10: Mean and standard deviation of the explanatory variables

	$PROD.1960$	$AV.INV.RATE$	$AV.SEC.EDU$	$AV.EMPL.GR$
1960	10895	0.21	0.14	0.02
2008	27677	0.24	0.37	0.02

Table 11: The values of the variables in 1960 and 2008

	$D1$	$D2$	$D3$	$LOG.PROD.1960$	$LOG.AV.INV.RATE$	$LOG.AV.SEC.EDU$	$LOG.AV.EMPL.GR$
$D1$	1	-0.38	-0.38	0.05	0.34	0.16	-0.16
$D2$	-0.38	1	-0.71	-0.08	-0.16	-0.08	0.3
$D3$	-0.38	-0.71	1	0.04	-0.1	-0.05	-0.18
$LOG.PROD.1960$	0.05	-0.08	0.04	1	0.21	0.77	-0.34
$LOG.AV.INV.RATE$	0.34	-0.16	-0.1	0.21	1	0.43	-0.05
$LOG.AV.SEC.EDU$	0.16	-0.08	-0.05	0.77	0.43	1	-0.42
$LOG.AV.EMPL.GR$	-0.16	0.3	-0.18	-0.34	-0.05	-0.42	1

Table 12: Correlations between explanatory variables