MPRA
Munich Personal RePEc Archive

Analyzing the Kuznets Relationship using Nonparametric and Semiparametric Methods

Li, Kui-Wai
City University of Hong Kong

January 2012

CIRJE-F-839

Revisiting the Kuznets Relationship using Nonparametric and Semiparametric Methods

Kui-Wai Li
City University of Hong Kong

Xianbo Zhou
Sun Yat-sen University

January 2012

Revisiting the Kuznets Relationship using Nonparametric and Semiparametric Methods

Kui-Wai Li [a, *] and Xianbo Zhou [b]
[a] City University of Hong Kong, Hong Kong
[b] Lingnan College, Sun Yat-sen University, China

Abstract: This paper studies the income inequality and economic development relationship by using unbalanced panel data of OECD and non-OECD countries for the period 1962 - 2003. The nonparametric estimation results show that income inequality in OECD countries are almost on the backside of the inverted-U relationship, while non-OECD countries are approximately on the foreside, except that the relationship in both country groups shows an upturn at a high level of development. Development has an indirect effect on inequality through control variables, but the modes are different in the two country groups. The model specification tests show that the relationship is not necessarily captured by the conventional quadratic function. The cubic and fourth-degree polynomials, respectively, fit the OECD and non-OECD country groups best. Our finding is robust regardless whether the specification uses control variables. Development plays a dominant role in mitigating inequality.

Keywords: Kuznets inverted-U; Nonparametric and semiparametric models; Unbalanced panel data
JEL classification: C14; C33; O11.

---

* Corresponding author: City University of Hong Kong, Tel.: 852 3442 8805, Fax: 852 3442 0195; E-mail: efkwli@cityu.edu.hk.
Acknowledgement: An earlier version of this paper was presented in late 2010 in the Geneva Trade and Development Workshop at the World Trade Organization. The authors are indebted to the comments from the workshop participations, including Marcelo Olarreaga, Frédéric Robert-Nicoud, Roberta Piermartini, and Cėline Carrėre. Research funding from the City University of Hong Kong under the Strategic Research Grant (Project no. 7002433) and the China National Natural Science Foundation (Grant no. 70971143) are gratefully acknowledged. Julian Chow, Helen H. K. Lam and Mansfield Wong have provided useful research inputs. The usual disclaimer applies.

I Introduction

The Kuznets (Kuznets, 1955) inverted-U curve relationship between income inequality and economic development argued that income inequality tends to increase at an initial stage of development, but income inequality will eventually fall as income continues to rise in developing countries. Other studies have been conducted along the line of the inverted-U relationship. For example, Sen (1991, 1992 and 1993) discussed inequality based on the concept of individual capability and functioning. Studies that have concentrated on the causes of income inequality included human capital, technological advancement, job diversity and political stability, while other studies have examined income inequality convergence (Galor and Zeire, 1993; Galor and Moav, 2000; Gould *et al*. 2001; Acemoglu, 2001; Desai *et al*. 2005; Bėnabou, 1996; Ravallion, 2003).

The Kuznets inverted-U relationship has attracted both theoretical and empirical debates, including whether the relationship should be considered as a law or could be improved through appropriate economic policies (Kanbur, 2000). It has been argued that the Kuznets inverted-U relationship has not been fully confirmed and validated in studies using parametric quadratic models (Li *et al*. 1998, Barro, 2000, Bulíř, 2001, Iradian, 2005). For example, by using nonparametric estimation based on a sample of cross-section country data, Mushinski (2001) showed that the quadratic parametric form of the relationship between Gini coefficient and real income per capita is misspecified. Huang (2004) presented a flexible nonlinear framework for a cross-section data of 75 countries and showed evidence of nonlinearity in the inverted-U relationship, Lin *et al*. (2006) confirmed the validity of the inverted-U relationship and presented a partial semiparametric linear investigation with some control variables.

In studying the relationship between inequality and development, the choice is whether control variables are included in the regression model. Some studies have complied with the original work of Kuznets and examined the total effect, instead of the direct effect, of development on inequality by using unconditional models (Mushinski, 2001; Wan, 2002). By using only one regressor of development in the regression model, the inequality-development relationship is considered as a law and the discussion on the impact of economic policy can be minimized. On the contrary, other studies considered the determinants of inequality and examined the impact of policy in affecting inequality, besides development. As such, the regressors in the regression model would include other policy variables and economic indicators, besides development (Li *et al*. 1998; Bulíř, 2001; Wang, 2006; Huang *et al*. 2009). The empirical results from these conditional models do reflect both the direct and indirect (via control variables) effects of development on inequality.

This paper presents a nonparametric (without control variables) and semiparametric (with control variables) investigation on the inequality-development relationship by using unbalanced panel data from a sample of developed OECD (Organization for Economic Co-operation and Development) and a sample of developing non-OECD countries. The unbalanced panel data set can provide observations over several periods of time. Such an analysis can incorporate heterogeneity across different sample of countries. In the panel data model, country-specific effects are specified to be fixed effects that are dependent on the regressors. This can help to obtain consistent estimators in the nonparametric regression function when inequality is regressed on the development variable and/or other control variables. The methodology in Henderson *et al*. (2008) is modified to cater for the nonparametric and semiparametric estimations with unbalanced panel data, and data-driven specification tests are conducted for the selected models.

The empirical results from unconditional and conditional models show that the channel effects of development via the control variables on inequality in both OECD and non-OECD sample countries are different, depending on the level of development in each sample country group. There is, however, much resemblance in the shapes of the nonparametric functions from both nonparametric and semiparametric estimations in each country group, implying that the control variables as a whole do not change the dynamic mode but the degree of inequality. Development still plays a dominant role in mitigating inequality. For the Kuznets' inverted-U hypothesis, our findings supported the cubic and fourth-degree polynomials for the OECD and non-OECD sample countries, respectively, in capturing the nonlinearity suggested by the nonparametric and semiparametric regressions.

Section II discusses the data and model specification and presents a parametric study on the inverted-U relationship. Section III briefly generalizes the methodology to suit the unbalanced panel data. Section IV conducts the nonparametric and semiparametric estimations and tests, while Section V concludes the paper.

II Data and Model Specification

The Gini coefficient for each sample country is used as the inequality proxy, and the dataset that was compiled and adapted from three datasets is obtained from "All the Ginis Database" under the World Bank project "Inequality around the World".[1] The variable chosen from the database is "Giniall" that gives the values of Gini coefficients from

[1] The web reference for the dataset is: http://web.worldbank.org/projects/inequality. The book reference is Milanovic (2005). Details for the dataset description of the December 2006 version are provided in the web. Some recent years' data are also given. The Deininger-Squire dataset that covers the period 1960-1996, the WIDER dataset that covers the period 1950-1998 and the World Income Distribution dataset that covers the period 1985-2000.

household-based surveys for 1,067 country/years.[2] To suit the nonparametric estimation, we have selected the countries with at least two years data. The final dataset used in this study contains 401 observations on 30 OECD countries and 303 observations on 45 non-OECD countries for the period from 1962 to 2003 (summarized in the Appendix).

The real GDP per capita is used as the proxy for the level of economic development. The study on the total effect on the relationship between inequality and development is extended to the study on the direct effect by the use of control variables. Two kinds of controls are considered. The policy control is indicated by the variables of openness (indicated by the percentage trade share of GDP in 2005 constant prices), urbanization (indicated by the percentage of urban population in total population), and investment (indicated by the percentage of investment share in real GDP per capita), denoted as openk, urbanize and ki, respectively. The other control variables that reflect the economic characteristics of the sample country are GDP growth and inflation (indicated by the annual percentage of GDP deflator). These data are obtained from the Penn World Table and World Development Indicators.

Table 1 Basic Statistics

	Gini	gdppc	growth	openk	urbanize	ki	deflator
OECD							
minimum	17.80	2,028.78	-19.29	5.29	2.90	10.41	-2.00
maximum	58.00	63,419.40	12.49	264.14	67.60	48.83	208.00
mean	34.34	18,658.44	2.53	45.59	28.84	26.85	9.35
Standard deviation	7.70	8,009.19	3.80	33.88	11.70	6.07	14.53
Non-OECD							
minimum	20.69	561.52	-21.60	10.32	0.00	2.11	-8.00
maximum	63.66	33,401.80	16.47	399.22	94.94	56.14	4,107.00
mean	44.45	7,167.09	1.73	74.08	46.53	22.78	83.19
Standard deviation	10.45	5,723.07	5.55	61.44	23.53	9.59	376.19

[2] This study chooses the Gini coefficients sample of country/years with "Di =1", where the dummy variable "Di" refers to the welfare concept of the Gini coefficient indicated either by income (=1) or by consumption (=0).

Table 2 Correlations between Variables

	Gini	log(gdppc)	growth(-1)	openk	urbanize	ki	inflation
OECD							
Gini	1.0000	-0.3506	0.1994	-0.3245	-0.0789	-0.1309	0.0394
log(gdppc)	-0.3506	1.0000	-0.0465	0.3689	-0.4740	0.1870	-0.3698
growth(-1)	0.1994	-0.0465	1.0000	-0.0565	0.1285	0.3530	-0.2705
openk	-0.3245	0.3689	-0.0565	1.0000	-0.1844	0.0376	-0.0196
urbanize	-0.0789	-0.4740	0.1285	-0.1844	1.0000	0.3676	0.1828
ki	-0.1309	0.1870	0.3530	0.0376	0.3676	1.0000	-0.2246
inflation	0.0394	-0.3698	-0.2705	-0.0196	0.1828	-0.2246	1.0000
Non-OECD							
Gini	1.0000	0.1266	0.0255	-0.0144	-0.0835	-0.2156	0.1248
Log(gdppc)	0.1266	1.0000	-0.0807	0.4232	-0.7964	0.2183	0.0532
growth(-1)	0.0255	-0.0807	1.0000	0.0810	0.1103	0.2938	-0.1586
openk	-0.0144	0.4232	0.0810	1.0000	-0.3339	0.4074	-0.0652
urbanize	-0.0835	-0.7964	0.1103	-0.3339	1.0000	-0.1201	-0.1577
ki	-0.2156	0.2183	0.2938	0.4074	-0.1201	1.0000	-0.1393
inflation	0.1248	0.0532	-0.1586	-0.0652	-0.1577	-0.1393	1.0000

Table 1 reports the basic statistics of these two types of variables for both OECD and non-OECD sample countries. One observation is that on average non-OECD countries have a larger inequality and variation than OECD countries, while OECD countries have higher level of development with more variations than non-OECD countries. Table 2 shows the correlation statistics between the variables. In OECD countries, the Gini has a negative correlation of -0.3506 with the logarithm of GDP per capita, which is the highest among its correlations with the other variables. In non-OECD countries, such correlation is smaller but positive (0.1266). However, this only provides the correlation between inequality and development for linear parametric regression models, but is invalid in the nonlinear and nonparametric relationships. Inequality and development are also moderately correlated with other variables. For example, the correlations of log GDP per capita with openness and urbanization are generally larger when compared to the other variables. Inequality is correlated moderately with openness in OECD countries and with investment in non-OECD countries. These reciprocal relationships imply that, in addition to the direct effect on inequality, the level of development may have an indirect effect on inequality through other channels.[3]

To study the relationship between inequality and economic development, we first specify the following nonparametric (unconditional) panel data model with fixed effects without control variables:

$$gini_{it} = g(lgdp_{it}) + u_i + v_{it}, \ t = 1, 2, \cdots, m_i; i = 1, 2, \cdots, n, \tag{1}$$

where the functional form of $g(\cdot)$ is not specified and $lgdp_{it}$ is the natural logarithm of real GDP per capita. For every country i, there are m_i observations from year 1 to m_i. The individual effects, u_i, of country i are fixed-effects/random-effects that are correlated/uncorrelated with country *i*'s economic development. For consistent estimation of $g(\cdot)$, we use a nonparametric estimation with fixed-effects model. The error term v_{it} is assumed to be i.i.d. with a zero mean and a finite variance, and is mean-independent of $lgdp_{it}$, namely $E(v_{it} \mid lgdp_{it}) = 0$.

There is no control variable in discussing the relationship between inequality and development in Model (1). This is consistent with the original Kuznets inverted-U relationship that provides a general framework to explain inequality unconditional on other variables other than the level of economic development. However, recent studies on the Kuznets inverted-U relationship have considered the determinants of inequality with control variables as that can provide *ceteris paribus* an analysis on the causality from economic development to inequality. The semiparametric (conditional) counterpart of Model (1) with control variables can be shown as:

[3] The indirect effects via channels can also be found in growth studies (Barro, 2000; Frankel and Rose, 2002).

$$gini_{it} = g(lgdp_{it}) + x_{it}^{'}\beta + u_i + v_{it},\ t = 1, 2, \cdots, m_i; i = 1, 2, \cdots, n, \qquad (2)$$

where x_{it} is the vector of the control variables. We adopt the assumptions in Model (1) and that v_{it} is also mean-independent of x_{it}. The control variable of “growth” is in the lagged form as growth may be endogenous in the inequality model (Huang *et al*. 2009). In Model (2) the indirect effect of development on inequality is controlled by the term $x_{it}^{'}\beta$, and hence $g(\cdot)$ shows the direct effect of the inequality from development.

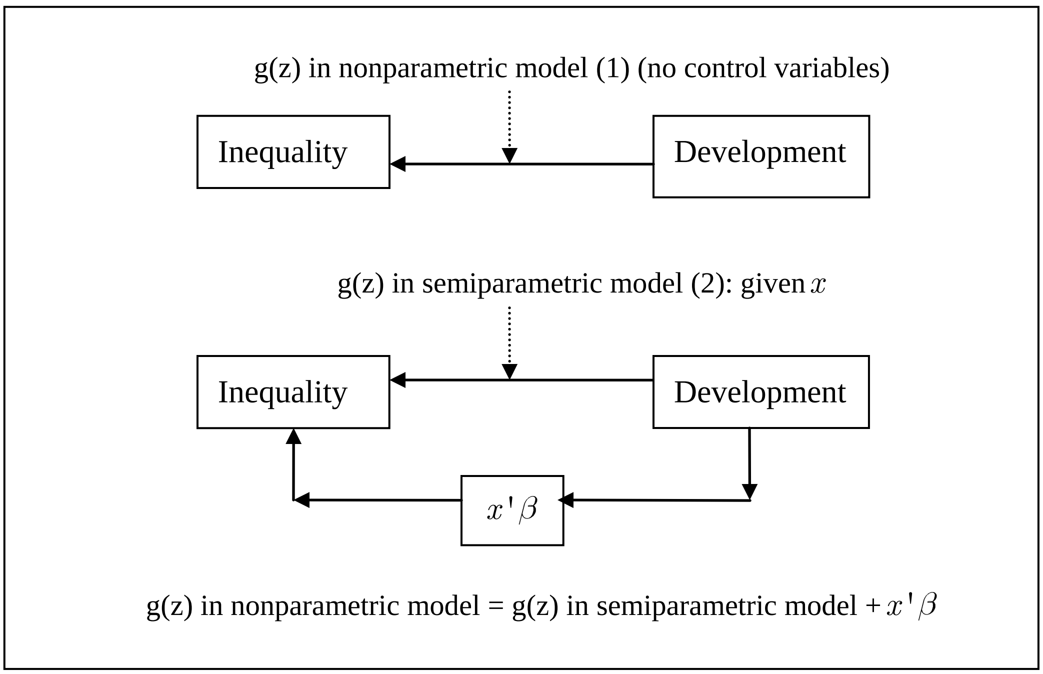

Figure 1 The Mechanism in the Nonparametric and Semiparametric Models

The relationships in Models (1) and (2) are intuitively illustrated in Figure 1. The g(z) in nonparametric Model (1) gives the gross contribution of development to inequality, while the g(z) in semiparametric Model (2) gives the net contribution of development to inequality, given x. The difference between the two g(z) is the indirect contribution of development to inequality via control variables x. When $g(\cdot)$ is specified as a parametric quadratic, cubic or fourth-degree polynomial function of $lgdp_{it}$, Model (1) and Model (2) become parametric unbalanced panel data models with fixed-effects, which can be estimated by the conventional method (Baltagi, 2008). However, in order to keep the approach comparable to the nonparametric counterpart, we use the difference of $y_{it} - y_{i1}$ instead of the transformation of $y_{it} - \overline{y}_{i\cdot}$ or the difference of $y_{it} - y_{i,t-1}$ in removing the fixed effects.

Table 3 contains the parametric estimation results for the developed OECD and developing non-OECD sample countries. The conventional quadratic specification is used to test the Kuznets hypothesis, and the coefficients on the linear and quadratic terms are expected to be positive and negative, respectively. The estimates for the non-OECD countries have the expected signs and are highly significant, while those for the OECD countries do not have the expected signs, regardless whether control variables are added into the model. The estimated models with higher-degree polynomials of the logarithm of GDP per capita are as shown by the "cubic" and "4-th degree" columns in Table 3. For the OECD sample countries, the cubic specification presents significant estimates of the coefficients in both the conditional and unconditional models, while the 4-th degree polynomial specification does not provide significant estimates. For the non-OECD sample countries, the estimates for the models without controls are all ideal while the estimate in the 4-th degree specifications with controls is perfect, although the quadratic estimate is also ideal as an explanation of the inverted-U relationship. These parametric estimation results show that the quadratic specifications do not give a best fit in both country samples, thereby casting doubts on the conventional quadratic specification in the inequality-development relationship.

Table 3 Parametric Estimation Results

	Parametric model						Semiparametric model
	quadratic	cubic	4-th degree	quadratic	cubic	4-th degree	
OECD							
lgdp	-1.372	388.180*	-87.381	-17.801*	262.500*	799.739	
	(5.071)	(72.324)	(875.06)	(7.053)	(69.337)	(803.449)	
$lgdp^2$	-0.098	-42.309*	34.887	0.661**	-29.430*	-116.847	
	(0.267)	(7.823)	(141.78)	(0.351)	(7.414)	(130.456)	
$lgdp^3$		1.517*	-4.029		1.074*	7.366	
		(0.281)	(10.175)		(0.264)	(9.380)	
$lgdp^4$			0.1489			-0.169	
			(0.273)			(0.252)	
growth(-1)				0.117*	0.113*	0.114*	0.121**
				(0.023)	(0.023)	(0.023)	(0.069)
openk				0.037*	0.035*	0.035*	-0.001
				(0.007)	(0.007)	(0.007)	(0.017)
urbanize				0.009	0.031	0.028	-0.171**
				(0.037)	(0.038)	(0.038)	(0.095)
ki				0.220*	0.1973*	0.200*	0.231*
				(0.028)	(0.028)	(0.028)	(0.083)
inflation				-0.025*	-0.029*	-0.029*	-0.030
				(0.007)	(0.007)	(0.007)	(0.020)
Non-OECD							
lgdp	24.775*	147.99*	-1797.5*	32.343*	68.221	-1358.56*	
	(4.297)	(42.84)	(423.62)	(4.287)	(43.601)	(410.51)	
$lgdp^2$	-1.272*	-16.05*	336.356*	-1.710*	-6.037	253.14*	
	(0.256)	(5.118)	(76.508)	(0.258)	(5.240)	(74.33)	
$lgdp^3$		0.583*	-27.540*		0.171	-20.576*	
		(0.202)	(6.095)		(0.207)	(5.939)	
$lgdp^4$			0.834*			0.617*	
			(0.181)			(0.177)	
growth(-1)				0.024	0.025	0.038	0.023
				(0.029)	(0.029)	(0.030)	(0.065)
openk				0.034*	0.033*	0.032*	0.021**
				(0.006)	(0.006)	(0.006)	(0.013)
urbanize				0.037	0.027	0.007	-0.004
				(0.029)	(0.031)	(0.032)	(0.060)
ki				-0.180*	-0.177*	-0.166*	-0.148*
				(0.021)	(0.022)	(0.022)	(0.048)
inflation				0.001*	0.001*	0.001*	0.001*
				(0.000)	(0.000)	(0.000)	(0.000)

Notes: The dependent variable is Gini. The numbers in the parentheses are standard errors of the coefficient estimates. Estimates of the intercepts in parametric models are not reported. * = 5% significance and ** = 10% significance.

III Nonparametric Estimation and Testing Method with Unbalanced Panel Data

Following the notations used in Henderson *et al*. (2008) and by denoting $y = gini$ and $z = lgdp$, Models (1) and (2) can be estimated by the iterative procedures to cater for the unbalanced panel data. To remove the fixed effects in Model (1), we write

$$\tilde{y}_{it} \equiv y_{it} - y_{1t} = g(z_{it}) - g(z_{i1}) + v_{it} - v_{i1} \equiv g(z_{it}) - g(z_{i1}) + \tilde{v}_{it} .$$

Denote $\tilde{y}_i = (\tilde{y}_{i2}, \cdots, \tilde{y}_{im_i})'$, $\tilde{v}_i = (\tilde{v}_{i2}, \cdots, \tilde{v}_{im_i})'$, and $g_i = (g_{i2}, \cdots, g_{im_i})'$, where $g_{it} = g(z_{it})$. The variance-covariance matrix of $\tilde{v}_i$ and its inverse are calculated, respectively, as $\Sigma_i = \sigma_v^2 (I_{m_i-1} + e_{m_i-1} e_{m_i-1}^{'})$ and $\Sigma_i^{-1} = \sigma_v^{-2} (I_{m_i-1} - e_{m_i-1} e_{m_i-1}^{'} / m_i)$, where I_{m_i-1} is an identity matrix of dimension $m_i - 1$ and e_{m_i-1} is a $(m_i - 1) \times 1$ vector of unity. The criterion function is given by

$$\Xi_i (g_i, g_{i1}) = -\frac{1}{2} (\tilde{y}_i - g_i + g_{i1} e_{m_i-1})' \Sigma_i^{-1} (\tilde{y}_i - g_i + g_{i1} e_{m_i-1}), \ i = 1, 2, \cdots, n .$$

Denote the first derivatives of $\Xi_i (g_i, g_{i1})$ with respect to g_{it} as $\Xi_{i,tg} (g_i, g_{i1})$, $t = 1, 2, \cdots m_i$. Then

$$\begin{aligned} \Xi_{i,1g} (g_i, g_{i1}) &= -e_{m_i-1}^{'} \Sigma_i^{-1} (\tilde{y}_i - g_i + g_{i1} e_{m_i-1}), \\ \Xi_{i,tg} (g_i, g_{i1}) &= c_{i,t-1}^{'} \Sigma_i^{-1} (\tilde{y}_i - g_i + g_{i1} e_{m_i-1}), \ t \geq 2, \end{aligned}$$

where $c_{i,t-1}$ is an $(m_i - 1) \times 1$ matrix with $(t-1)^{\text{th}}$ element/other elements being 1/0. Denote $(\alpha_0, \alpha_1)' \equiv \left(g(z), dg(z)/dz \right)'$. It can be estimated by solving the first order conditions of the above criterion function through iteration:

$$\sum_{i=1}^{n} \frac{1}{m_i} \sum_{t=1}^{m_i} K_h (z_{it} - z) G_{it} \Xi_{i,tg} \left(\hat{g}_{[l-1]}(z_{i1}), \cdots, G_{it} (\alpha_0, \alpha_1)', \cdots, \hat{g}_{[l-1]}(z_{im_i}) \right) = 0 ,$$

where the argument $\Xi_{i,tg}$ is $\hat{g}_{[l-1]}(z_{is})$ for $s \neq t$ and $G_{it} (\alpha_0, \alpha_1)'$ when $s = t$, and

$\hat{g}_{[l-1]}(z_{is})$ is the $(l-1)^{\text{th}}$ iterative estimates of $(\alpha_0, \alpha_1)'$. Here $G_{it} \equiv \left(1, (z_{it} - z)/h\right)'$ and $k_h(v) = h^{-1}k(v/h)$, $k(\cdot)$ is the kernel function. The next iterative estimator of $(\alpha_0, \alpha_1)'$ is equal to $\left(\hat{g}_{[l]}(z), \hat{g}_{[l]}(z)\right)' = D_1^{-1}(D_2 + D_3)$, where

$$D_1 = \sum_{i=1}^{n} \frac{1}{m_i} \left(e_{m_i-1}' \Sigma_i^{-1} e_{m_i-1} K_h(z_{i1} - z) G_{i1} G_{i1}' + \sum_{t=2}^{m_i} c_{i,t-1}' \Sigma_i^{-1} c_{i,t-1} K_h(z_{it} - z) G_{it} G_{it}' \right),$$

$$D_2 = \sum_{i=1}^{n} \frac{1}{m_i} \left(e_{m_i-1}' \Sigma_i^{-1} e_{m_i-1} K_h(z_{i1} - z) G_{i1} \hat{g}_{[l-1]}(z_{i1}) + \sum_{t=2}^{m_i} c_{i,t-1}' \Sigma_i^{-1} c_{i,t-1} K_h(z_{it} - z) G_{it} \hat{g}_{[l-1]}(z_{it}) \right),$$

$$D_3 = \sum_{i=1}^{n} \frac{1}{m_i} \left(-K_h(z_{i1} - z) G_{i1} e_{m_i-1}' \Sigma_i^{-1} H_{i,[l-1]} + \sum_{t=2}^{m_i} K_h(z_{it} - z) G_{it} c_{i,t-1}' \Sigma_i^{-1} H_{i,[l-1]} \right),$$

and $H_{i,[l-1]}$ is an $(m_i - 1) \times 1$ vector with elements

$$\left(\tilde{y}_{it} - (\hat{g}_{[l-1]}(z_{it}) - \hat{g}_{[l-1]}(z_{i1}))\right), \ t = 2, \cdots, m_i .$$

The series method is used to obtain the initial estimator for $g(\cdot)$. The convergence criterion for the iteration is set to be

$$\sum_{i=1}^{n} \frac{1}{m_i} \sum_{t=2}^{m_i} \left(\hat{g}_{[l]}(z_{it}) - \hat{g}_{[l-1]}(z_{it})\right)^2 / \sum_{i=1}^{n} \frac{1}{m_i} \sum_{t=2}^{m_i} \hat{g}_{[l-1]}^2(z_{it}) < 0.01.$$

Further, the variance σ_v^2 is estimated by

$$\hat{\sigma}_v^2 = \frac{1}{2n} \sum_{i=1}^{n} \frac{1}{m_i - 1} \sum_{t=2}^{m_i} (y_{it} - y_{i1} - (\hat{g}(z_{it}) - \hat{g}(z_{i1})))^2 .$$

The variance of the iterative estimator $\hat{g}(z)$ is calculated as $\kappa(nh\hat{\Omega}(z))^{-1}$, where $\kappa = \int k^2(v)dv$, and $\hat{\Omega}(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{m_i - 1}{m_i} \sum_{t=2}^{m_i} K_h(z_{it} - z) / \hat{\sigma}_v^2$.

We use the series method to obtain an initial estimator for $\theta(\cdot)$ and then conduct the iteration process. The convergence criterion for the iteration is set to be

$$\sum_{i=1}^{n} \frac{1}{m_i} \sum_{t=2}^{m_i} \left(\hat{g}_{[l]}(z_{it}) - \hat{g}_{[l-1]}(z_{it})\right)^2 / \sum_{i=1}^{n} \frac{1}{m_i} \sum_{t=2}^{m_i} \hat{g}_{[l-1]}^2(z_{it}) < 0.01.$$

Further, the variance σ_v^2 is estimated by

$$\hat{\sigma}_v^2 = \frac{1}{2n}\sum_{i=1}^{n}\frac{1}{m_i-1}\sum_{t=2}^{m_i}(y_{it}-y_{i1}-(\hat{g}(z_{it})-\hat{g}(z_{i1})))^2 .$$

The variance of the iterative estimator $\hat{g}(z)$ is calculated as $\kappa(nh\hat{\Omega}(z))^{-1}$, where $\kappa = \int k^2(v)dv$, and $\hat{\Omega}(z) = \frac{1}{n}\sum_{i=1}^{n}\frac{m_i-1}{m_i}\sum_{t=2}^{m_i}K_h(z_{it}-z)/\hat{\sigma}_v^2$.

For the estimation of semiparametric Model (2), we denote the nonparametric estimator of the regression functions of the dependent variable y and the control variables x, respectively, as $\hat{g}_y(\cdot)$ and $\hat{g}_x(\cdot) = (\hat{g}_{x,1}(\cdot),\cdots,\ \hat{g}_{x,d}(\cdot))'$, where d is the number of controls. Then β is estimated by $\hat{\beta} = \left(\sum_{i=1}^{n}\tilde{x}_{i*}'\Sigma_i^{-1}\tilde{x}_{i*}/m_i\right)^{-1}\left(\sum_{i=1}^{n}\tilde{x}_{i*}'\Sigma_i^{-1}\tilde{y}_{i*}/m_i\right)$, where $\tilde{y}_{i*}$ and $\tilde{x}_{i*}$ are, respectively, $(m_i-1)\times 1$ and $(m_i-1)\times d$ matrices with the t-th row element being $\tilde{y}_{it*} = \tilde{y}_{it} - (\hat{g}_y(z_{it}) - \hat{g}_y(z_{i1}))$ and $\tilde{x}_{it*} = \tilde{x}_{it} - (\hat{g}_x(z_{it}) - \hat{g}_x(z_{i1}))$. The nonparametric function $g(\cdot)$ is estimated by the same method shown above, except that $\tilde{y}_{it}$ is replaced by $\tilde{y}_{it} - x_{it}'\hat{\beta}$ whenever it occurs.

For the selected model to incorporate a data-driven procedure, we further modify the specification tests to suit an unbalanced panel data case. Regardless whether the models have control variables as regressors, we perform the following two specification tests. The first specification test is to choose in Model (1) between parametric and nonparametric models without control variables. The null hypothesis H_0 is parametric model with $g(z) = g_0(z,\gamma)$. For example, $g_0(z,\gamma) = \gamma_0 + \gamma_1 z + \gamma_2 z^2$. The alternative H_1 is that $g(z)$ is nonparametric. The test statistic for testing this null is $I_n^{(1)} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{t=1}^{m_i}(g_0(z_{it},\hat{\gamma}) - \hat{g}(z_{it}))^2$, where $\hat{\gamma}$ is a consistent estimator of the parametric panel data model with fixed effects; $\hat{g}(\cdot)$ is the iterative consistent estimator of Model (1). The second specification test is to choose in Model (2) between parametric and semiparametric models with control variables. The null hypothesis H_0 is parametric model with $g(z) = g_0(z, \gamma)$. The alternative is that $g(z)$ is nonparametric in Model (2). The test statistic for testing this null is $I_n^{(2)} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{t=1}^{m_i}(g_0(z_{it}, \tilde{\gamma}) + x_{it}^{'}\tilde{\beta} - \hat{g}(z_{it}) - x_{it}^{'}\hat{\beta})^2$, where $\tilde{\gamma}$ and $\tilde{\beta}$ are consistent estimators in the parametric panel data model with fixed effects; $\hat{g}(\cdot)$ and $\hat{\beta}$ are the iterative consistent estimator of Model (2).

In the following empirical study, we apply bootstrap procedures in Henderson *et al*. (2008) to approximate the finite sample null distribution of test statistics and obtain the bootstrap probability values for the test statistics.

IV Empirical Results

The kernel in both the estimation and the testing is the Gaussian function and the bandwidth is chosen according to the rule of thumb[4]: $h = 1.06\sigma_z \left(\sum_{i=1}^{n} m_i\right)^{-1/5}$, where σ_z is the sample standard deviation of { z_{it} }. All the bootstrap replications are set to be 400. The last column in Table 3 reports the coefficient estimation for the control variables in the parametric part of semiparametric Model (2). For the OECD countries, with the exception of “openk” and “urbanize”, the coefficient estimates of all other control variables have the same signs and similar values in both parametric and semiparametric models. For countries in the non-OECD sample, with the exception of the “urbanize” variable, the coefficient estimates of all other control variables are highly similar in both

[4] We also slightly change the constant instead of 1.06, and find that the estimation and test results are not significantly affected.

parametric and semiparametric models.

The inconsistency in the coefficient estimates of such variables as “urbanize” and “openk” casts doubts on model specification once again, but this will be tested in the final stage of our analysis. An interesting finding is the different signs in the coefficients of investment and inflation between the two samples. Investment share has a positive effect on inequality in OECD, but a negative effect in non-OECD, implying that investment aggravates inequality in OECD countries, while it alleviates inequality in non-OECD countries. The effect of inflation is exactly the opposite to that of investment between the OECD and non-OECD countries.

In Table 4, the nonparametric function $g(\cdot)$ is estimated at some quantile points of the logarithm of GDP per capita by using nonparametric Model (1) and semiparametric Model (2). For the OECD sample countries, when their development level is at the 2.5 percent quantile from the bottom, the estimate of $g(\cdot)$ from semiparametric Model (2) is larger than that from nonparametric Model (1). The difference is the total contribution by control variables to inequality. But when the development level is one of the other quantiles, the estimate of $g(\cdot)$ from nonparametric Model (1) becomes larger, which implies that the integrated contribution by control variables to inequality becomes positive. In short, policy variables and economic characteristics can indeed play a role in affecting inequality in the higher stage of development.

However, the estimation results in OECD countries are opposite to those in non-OECD countries. Table 4 shows that for the non-OECD sample countries all the estimates of $g(\cdot)$ at each quantile from semiparametric Model (2) are larger than that at the same quantile from nonparametric Model (1). The integrated contribution of control variables to inequality is therefore negative, namely, they totally decrease inequality. We need to discuss and compare the results from the OECD and non-OECD countries.

Table 4 Nonparametric Estimation of $g(\cdot)$ at Different Points of lgdp

Quantile of lgdp		Nonparametric model (1)		Semiparametric model (2)	
	z	m(z)	std. err.	m(z)	std. err.
OECD					
2.5%	8.5937	40.7812	5.4563	42.5567	5.2224
25%	9.4911	35.4923	1.9810	34.5420	1.8960
50%	9.8375	33.4746	1.5039	31.6731	1.4394
75%	10.0647	32.3819	1.5291	30.3893	1.4635
95%	10.3257	31.4750	2.1687	28.6795	2.0757
97.5%	10.3968	30.3402	2.5887	27.2607	2.4777
at sample mean of z					
mean	9.7299	34.2492	1.5917	32.5366	1.5234
Non-OECD					
2.5%	7.0233	40.5527	3.2093	42.6521	3.0549
25%	8.0689	44.9614	2.1619	47.1526	2.0579
50%	8.6832	46.5913	1.5724	48.3060	1.4968
75%	9.0522	45.2422	1.6618	47.0502	1.5819
95%	9.8949	42.3625	3.2022	43.5062	3.0481
97.5%	9.9989	42.9649	3.4717	43.8056	3.3047
at sample mean of z					
mean	8.5836	46.6044	1.6252	48.3800	1.5470

OECD Countries

Figures 2 and 3 illustrate the nonparametric and semiparametric estimations of $g(\cdot)$ in Models (1) and (2), respectively, for the OECD sample countries, and show also the lower and upper bounds of 95 percent confidence intervals for the estimates. The nonparametric estimates are reasonable, though the estimation has some boundary effects. The two curves of $g(\cdot)$ in Figures 2 and 3 look very similar in shape, which implies that the control variables, though having an effect on inequality, have played little role in changing the nonlinear shapes of $g(\cdot)$.[5]

[5] Huang (2004) also shows this finding in his cross-section analysis.

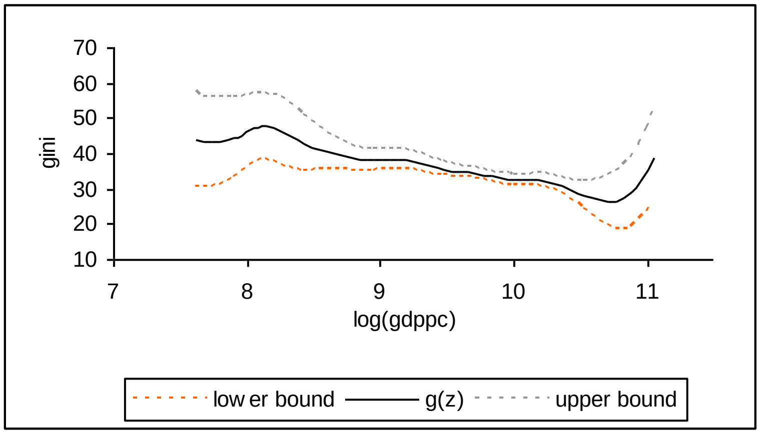

Fig. 2 Nonparametric Estimation in Model 1: OECD

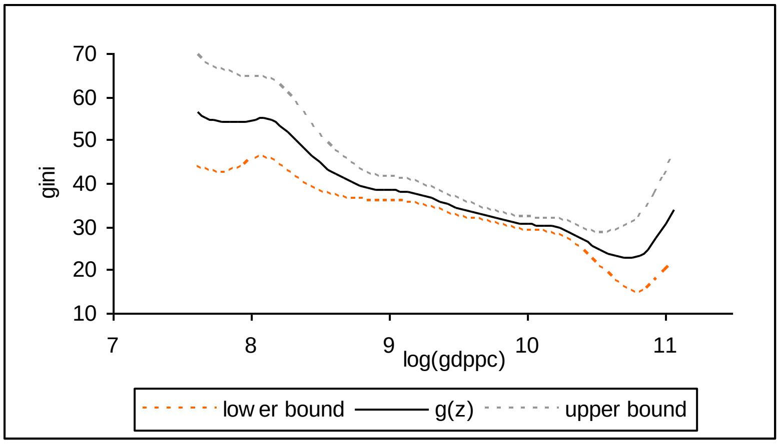

Fig. 3 Semiparametric Estimation in Model 2: OECD

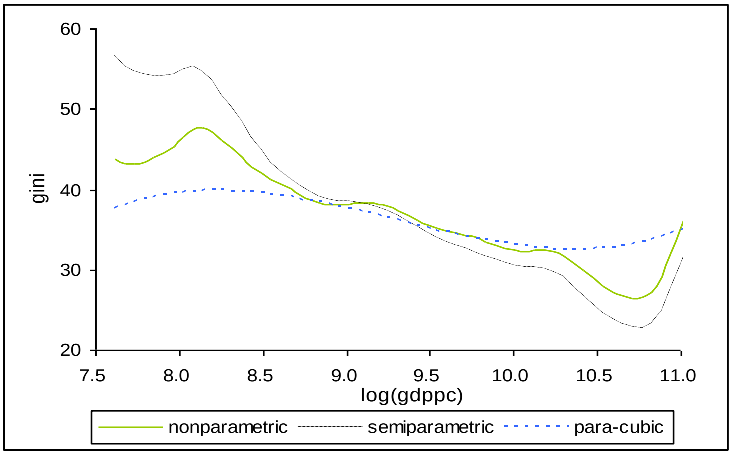

Fig. 4 Comparing Non- and Semi-parametric Estimation: OECD

Our finding shows that the shapes of $g(\cdot)$, whether estimated from semiparametric Model (2) or nonparametric Model (1), are very similar to that in Mushinski (2001) who used a nonparametric estimation model with cross-section data. The results reflect an increasing, albeit short, portion at lower levels of development, a turning point (where the effect of development on inequality changes from positive to negative) at 8.2 (about \$3,640 in 2005 dollars), followed by a longer decreasing portion of the curve. While the results in Mushinski (2001) reflected the predominance of middle-income countries in his dataset, the result from our dataset using nonparametric estimation for the OECD sample countries that represented mainly high-income or upper-middle-income countries reflected the predominance of the high-income and upper-middle-income group. Also, the curves hint an upturn at a higher income level (around 10.8, about \$49,020), which accords with the high-level upturn point from a highly developed economy (Ram, 1991).

The contribution of development to the reduction of inequality via control variables can be seen from the vertical difference between the two "nonparametric" and "semiparametic" curves in Figure 4. The integrated effect of development on inequality via the control variables is negative at lower income levels (lower than 9.2, about \$9,900) since the curve from nonparametric estimation is below that from semiparametric estimation. Control variables can contribute to reduce inequality at or below this level of income. When the development level is above \$9,900, this indirect integrated effect becomes positive, implying that control variables as a channel increase inequality at this higher stage of income level. The original Kuznets hypothesis ignored the channel effect of development on inequality via other determinants. The inclusion of control variables in the semiparametric model can show the indirect effect of development on inequality. The evidence shows that whether the channel effects of development are positive or negative in OECD countries depend on the development level.

The analysis so far shows that nonparametric and semiparametric estimations can provide additional information on the Kuznets hypothesis. However, which specification is most suited to the sample requires further hypothesis tests. The upper portion of Table 5 presents various test results. In the case without control variables, the p-values for the quadratic and cubic parametric specifications against nonparametric specification are, respectively, 0.07 and 0.06. At the 10 percent significant level, the nulls of quadratic and cubic parametric specifications are rejected and the alternative of nonparametric specification is accepted. However, at the 5 percent level, the nulls are accepted. In the case with control variables, the p-values are less than 5 percent. So we reject the nulls of parametric specification at the 5 percent level and accept the semiparametric specification. The test results shown in Table 5 provide further support to our analysis.

Since the cubic parametric model without control variables is accepted at the conventional 5 percent significant level (note that the quadratic curve is not considered here because its estimate is not significant, see Table 3: OECD), we include the curve of the estimated cubic function in Figure 4. This cubic function implies a turning point which is almost the same as those in the curves from the nonparametric and semiparametric estimations, and hence can capture some of the non-concavities suggested by the nonparametric and semiparametric regressions. The F test for the parametric specifications in Table 6 shows that the cubic function of the logarithm of GDP per capita provides a better fit than the quadratic function and cannot be rejected against the alternative fourth-degree specification. Table 3 (OECD) also shows that all terms in the cubic polynomial are significant, and they are better than the estimates in quadratic and fourth-degree polynomials. Hence the tests support the estimation of a cubic function of development for the OECD sample rather than a quadratic or fourth-degree function in the case of no control variables.

Table 5 Nonparametric and Semiparametric Model Specification Tests

Model		Hypotheses	I_n (p-value)	Model selected
OECD				
without control variables	Parametric or Nonparametric ?	H_0: Quadratic parametric H_1: Nonparametric	4.006 (0.070)	Nonparametric (10%) Parametric (5%)
		H_0: Cubic parametric H_1: Nonparametric	3.964 (0.060)	Nonparametric (10%) Parametric (5%)
with control variables	Parametric or Semiparametric ?	H_0: Quadratic parametric H_1: Semiparametric	10.205 (0.013)	Semiparametric
		H_0: Cubic parametric H_1: Semiparametric	11.823 (0.008)	Semiparametric
Non-OECD				
without control variables	Parametric or Nonparametric ?	H_0: Quadratic parametric H_1: Nonparametric	6.748 (0.415)	Quadratic parametric
		H_0: Cubic parametric H_1: Nonparametric	5.212 (0.375)	Cubic parametric
with control variables	Parametric or Semiparametric ?	H_0: Quadratic parametric H_1: Semiparametric	6.796 (0.438)	Quadratic parametric
		H_0: Cubic parametric H_1: Semiparametric	5.942 (0.465)	Cubic parametric

Table 6 Parametric Model Tests for Inclusion of Polynomial Terms

	Without control variables	With control variables
Degree of polynomial	F statistic: OECD / Non-OECD	F statistic: OECD / Non-OECD
H_0:Second vs H_1:Third	11.2043* / 5.3481*	0.3217 / 1.0814
H_0:Third vs H_1:Fourth	1.5146 / 9.2195*	2.8767 / 5.5398*
H_0:Fourth vs H_1:Fifth	0.2849 / 1.0849	2.9167 / 0.0058

Note: * = 5% significance.

In the case with control variables, the tests in Table 6 present no obvious evidences to show which parametric specification is best. One can conclude from Table 3 (OECD) that the cubic form is preferred since all the estimated coefficients of the cubic polynomial statistically prevail over those of the quadratic and fourth-degree counterparts. However, as the test with control variables in Table 5 shows, parametric cubic specification is rejected and semiparametric model is accepted at the 5 percent significant level. Hence, the insignificance of the F tests in the case of control variables for OECD countries is expected since the F tests based on the estimation of parametric model may not be valid for semiparametric models.

Non-OECD Countries

Figures 5 and 6 respectively present the nonparametric estimation of $g(\cdot)$ in Models (1) and (2) for the non-OECD sample countries. The estimates provide a result stronger than those in Figures 2 and 3 since the boundary effect for nonparametric estimation is less significant. There is much resemblance between the shapes of nonlinearity of $g(\cdot)$ in Figures 5 and 6. The results reflect a rapidly increasing, albeit short, portion at lower levels of development and a first turning point at 7 (about $1,100), then another increasing, albeit long and flat, portion at the middle income level of development, then followed by a slowly decreasing portion of the curve with the second turning point at 8.7 (about $6,000). Finally, the curves also hint at an upturn at a higher income level (around 10, about $22,026), similar to the processes shown in OECD countries and the findings in Ram (1991) and Mushinski (2001). The second turning point is higher than the first one and the final upturn occurs at even higher inequality level. This process presents a "roller coaster" mode, albeit flat and long in the middle of the process.[6] In the non-OECD sample countries, if the final upturn is not accounted for, the process approximately accords with the inverted-U hypothesis.

Figure 7 contains both the nonparametric and semiparametic curves estimated for the non-OECD sample countries. The vertical difference reflects the contribution of control variables to inequality. The integrated effect of development on inequality via control variables is always negative, which is different from that in the OECD countries. In short, control variables generally mitigate inequality in non-OECD countries, except when the logarithm of income level is very large (greater than 10.2). This evidence shows that the channel effect of development on inequality via the control variables as a whole is negative in non-OECD countries.

Figure 7 contains also the curve of the estimated fourth-degree polynomial. It shows that the fourth-degree function can capture some of the non-concavities suggested by nonparametric and semiparametric regressions. Although the shape resembles the inverted-U relationship, with the exception when the development reaches a very high level, the conventional quadratic form used to estimate the inequality-development relationship might be misspecified for the non-OECD dataset.

[6] This "roller coaster" mode also appeared in the study by Keane and Prasad (2002) that empirically examined Poland's inequality-development relationship and generalized the mode by using the sample of transitional economies.

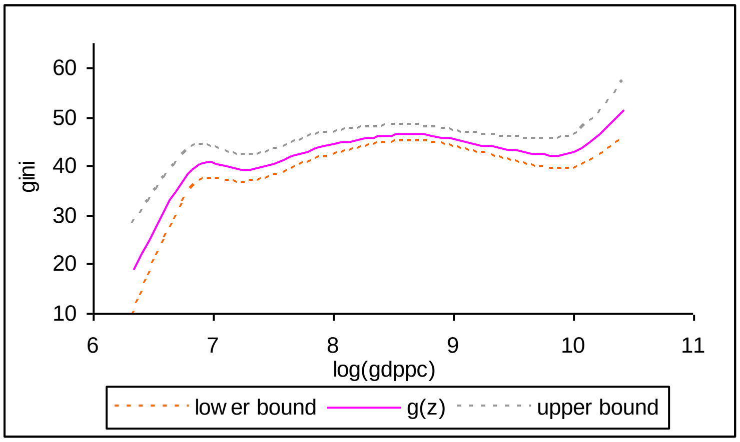

Fig. 5 Nonparametric Estimation in Model 1: Non-OECD

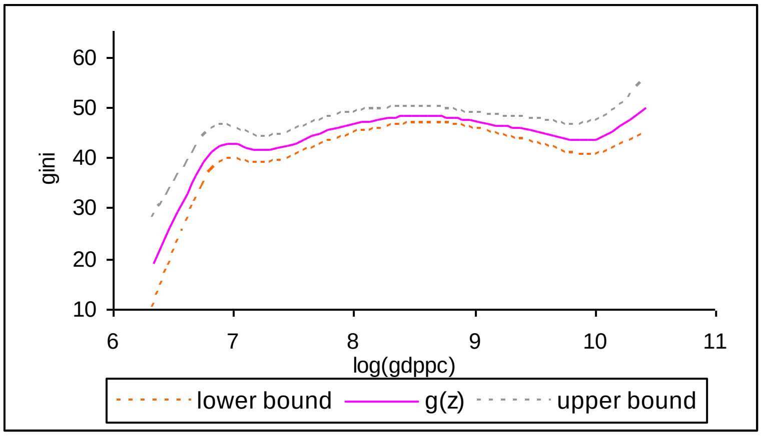

Fig. 6 Semiparametric Estimation in Model 2: Non-OECD

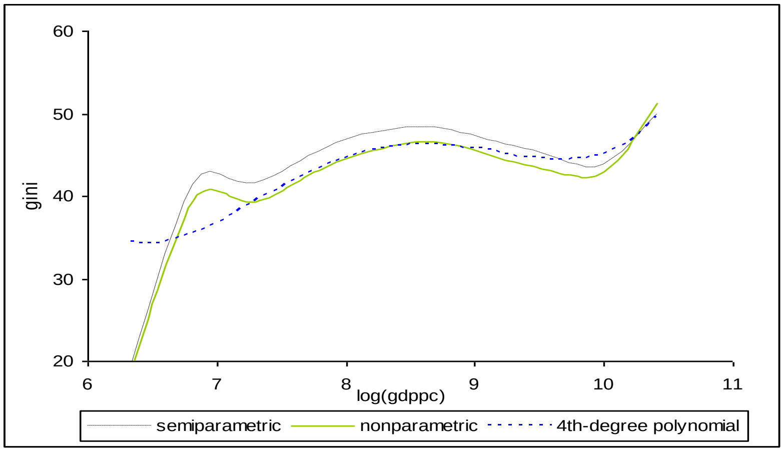

Fig. 7 Comparing Non- and Semi-parametric Estimation: Non-OECD

The lower portion of Table 5 presents the test results for choosing between nonparametric or semiparametric and parametric specifications in the non-OECD sample countries. The quadratic and cubic parametric specifications cannot be rejected at any usual significant level whether or not the model includes control variables as regressors. Among the parametric models, the F tests for parametric specifications in Table 6 show that when compared to the third- and fifth-degree polynomial specifications, the fourth-degree polynomial of the logarithm of GDP per capita gives a sufficiently accurate description of the non-OECD countries, whether or not the control variables are added into the models. Recall in Table 3 (non-OECD) that all coefficient estimates in the fourth-degree polynomials are significant at the 5 percent level. Hence the fourth-degree function specification gives a better description of the data than the quadratic and cubic polynomials. The tests for the non-OECD countries support the estimation of a fourth-degree function of development rather than a quadratic or cubic form.[7]

Comparing OECD and non-OECD Countries

Additional differences between the estimates of the inequality-development relationship for the OECD and non-OECD sample countries can be seen from the nonparametric approach. The nonparametric and semiparametric estimates of the relationship between inequality and development in OECD countries have only one turning point at a low income level of about $3,640 (see Figure 4), while the estimates in non-OECD countries have two turning points, with the first being at a rather low income

[7] Such a support on the use of a fourth-degree polynomial rather than a quadratic form can also be seen in Mushinski (2001).

level (about \$1,100) and the other being at a middle income level (about \$6,000) (see Figure 7). Generally, when the level of economic development is not very high, inequality in OECD/non-OCED countries will decrease/increase with development. However, both curves make an upturn at an advanced level of development, though they act with a different manner: inequality in OECD countries shows an upturn at a lower inequality level than non-OECD countries. If this upturn were not accounted for, the evidence from OECD countries would roughly suggest that the inverted-U relationship between inequality and development is located beyond the turning point and on the backside of the inverted-U relationship, while the non-OECD evidence shows that the relationship roughly accords with a full inverted-U relationship.

Figure 8 presents an intuitive comparison of the nonparametric (left) and semiparametric (right) estimates of the nonparametric function g(.) between OECD and non-OECD countries. On the same portion of development ($lgdp \in [7.6, 10.4]$), the two curves for OECD and non-OECD countries intersected at one point at about 8.4, namely at GDP per capita = \$4,447. It can be seen that OECD countries has higher inequality than non-OECD countries when GDP per capita is less than \$4,447, but has lower inequality when development has exceeded this level. The dynamics in Figure 8 implies that non-OECD countries seem to face higher inequality at middle or high income levels. The difference of the intersection of OECD and non-OECD countries between nonparametric and semiparametric estimates reflects the role of control variables. The effects of policy variables on inequality change the mode of the intersection, but have little effect on the relative location of the two curves from nonparametric and semiparametric estimations.

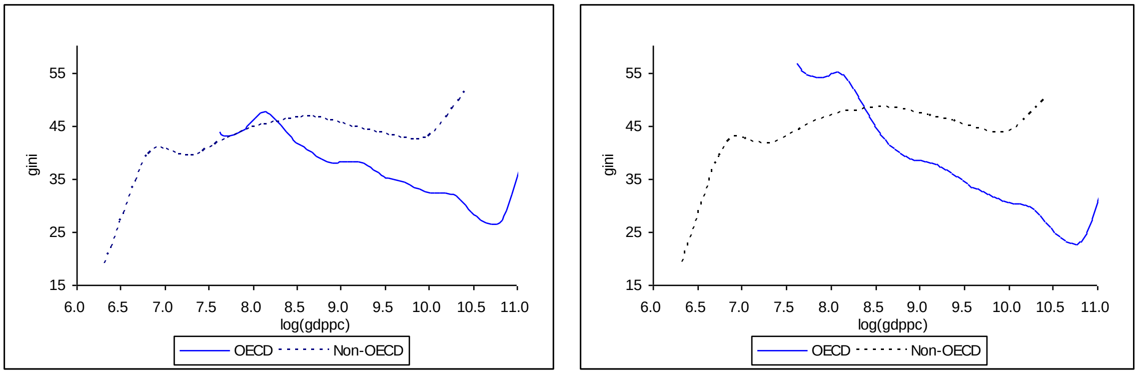

Fig. 8 Nonparametrc (Left) and Semiparametric (Right) Estimation of g(.)

The estimation and test results from nonparametric model without control variables imply that the conventional quadratic concave function may not necessarily capture the relationship between inequality and development in both OECD and non-OECD countries. The tests show that the cubic polynomial of development levels can capture the relationship more accurately than quadratic and fourth-degree polynomials in OECD countries, while the fourth-degree polynomial can give a better description of the data than other polynomials in non-OECD countries, whether or not the control variables are added as regressors. In both sample countries, analysis based on a quadratic specification for the relationship between inequality and development is misleading.

The estimation and test results from the semiparametric model with control variables show that the data-driven model selection for OECD countries requires a semiparametric specification while the non-OECD countries require a fourth-degree polynomial parametric specification. Given the integrated contribution by control variables to inequality shown above, we next study the effects of the control variables on inequality by comparing the estimates in the parametric part of the semiparametrc model for the OECD countries (shown in the last column in Table 3) and the 4-th degree parametric model for the non-OECD countries (shown in the "4-th degree" column in Table 3). The implications of the effects of the control variables are different in the two sample country groups, except the variable “growth (-1)” which has a positive effect on inequality.

Specifically, the effect of openness on inequality in OECD countries is negative, albeit insignificant both economically and statistically, while openness has a positive and significant effect on inequality in non-OECD countries. Openness generally will aggravate income inequality in non-OECD countries, but has an emollient effect on inequality in OECD countries. Although integration to the world market is expected to help non-OECD countries to promote prosperity, increasing opportunities to trade are also likely to affect income distribution. Whether or not increasing openness to trade is accompanied by a reduction or an increase in inequality has strongly been debated (Julien, 2007; Wood, 1997).

The effect of urbanization on inequality is negative (-0.171) and significant for OECD countries, but positive (0.007) and insignificant for non-OECD countries. Urbanization helps to mitigate inequality in OECD countries, but increases inequality in non-OECD countries, albeit insignificantly. According to Anand (1993), the urban-rural difference generally results in larger inequality in total income distribution due to urbanization. Hence, in the process of urbanization, income inequality will first increase and then decrease with urbanization or the migration of rural population to cities. In our case, OECD countries have much higher urbanization than non-OECD countries. Hence the negative effect of urbanization on inequality in OECD and the positive effect in non-OECD accord with this general urbanization-inequality relationship.

The finding that investment share aggravates inequality in OECD countries but reduces inequality in non-OECD countries contrasts with the result in Barro (1999) that showed little overall relationship between income inequality and investment. One explanation is that investment may have potential endogeneity in inequality models.

Inflation has a negative albeit insignificant effect on inequality in OECD countries but has a positive and significant effect on inequality in non-OECD countries. Generally, cross-country evidence on inflation and income inequality suggests that they are positively related. For example, Albanesi (2007) argued that the correlation between inflation and income inequality is the outcome of a distributional conflict underlying the determination of government policies, and inflation is positively related to the degree of inequality as low income households are more vulnerable to inflation. Since non-OECD countries have a high average inflation, their monetary authorities should reduce inflation to alleviate income inequality. However, the impact of inflation on income distribution may be nonlinear (Bulíř, 2001), and the positive and significant effect of inflation on inequality would need to be explained with caution.

V Conclusion

This paper provides evidences on the relationship between inequality and development from the estimations and tests of nonparametric and semiparametric panel data models with fixed effects. Based on an unbalanced panel dataset, this study presents new evidences about the inequality-development relationship in both the developed OECD and the developing non-OECD countries and provides additional information on the mechanics of the effect of development on inequality.

For the OECD countries, inequality generally decreases with development, with the exception of an upturn at a higher income level. The control variables will help reduce income inequality at lower income levels (below about $9,900), but they tend to increase inequality when development exceeds that level. For the non-OECD countries, the inequality-development relationship appears in a “roller coaster” mode with two turning points, and the second upturn appears at a very high income level. When compared to the performance in OECD countries, the effect of development on inequality via the control variables is always negative in non-OECD countries, except after an upturn at a high income level. Non-OECD countries seem to face serious inequality at the middle or high income level.

Nonparametric estimations without control variables suggest that a polynomial of higher-degree might give a sufficiently accurate description of both OECD and non-OECD countries. Our tests support estimating a cubic function of development for the OECD sample and a fourth-degree function of development for the non-OECD sample. Both the parametric estimates capture some of the non-concavities suggested by nonparametric and semiparametric regressions.

Semiparametric estimations and tests with control variables present some implications on policy. In OECD countries, growth and investment share in GDP aggravate inequality, but openness, urbanization and inflation would reduce inequality. In non-OECD countries, growth, openness, urbanization, and inflation generally increase inequality, but investment share helps mitigate inequality. It would be appropriate to argue that investment in non-OECD countries should be geared more to improve the conditions of the low income groups, as that shall reduce inequality overtime.

References

Acemoglu, D., 2001. Good jobs versus bad jobs. Journal of Labor Economics 19 (1): 1-21.

Albanesi, A., 2007. Inflation and inequality. Journal of Monetary Economics 54: 1088-1114.

Anand, S., 1993. The Kuznets process and the inequality-development relationship. Journal of Development Economics 40 (1): 25-52.

Baltagi, Badi H., 2008. Econometric Analysis of Panel Data. Fourth Edition, John Wiley & Sons, Ltd.

Barro, Robert J., 1999. Inequality, Growth and Investment. NBER Working Paper No. 7038.

Barro, Robert J., 2000. Inequality and growth in a panel of countries. Journal of Economic Growth 5: 5-32.

Bėnabou. R., 1996. Inequality and growth. In Bernanke, B. and Rotemberg, J. (Eds.), National Bureau of Economic Research Macroeconomics Annual, MIT Press, 11-74.

Bulíř, Aleš, 2001. Income inequality: Does inflation matter? IMF Staff Papers 48: 139-159.

Chen, Been-Lon, 2003. An inverted-U relationship between inequality and long-run growth. Economics Letters 78: 205-212.

Desai, R. M., Olofsgard, A. and Yousef, T. M., 2005. Inflation and inequality: does political structure matter? Economic letters 87: 41-46.

Frankel, F. and Rose, A., 2002. An estimate of the effect of common currencies on trade and income. Quarterly Journal of Economics 117 (2): 437-466.

Galor, O. and Maov, O., 2000. Ability biased technological transition, wage inequality and economic growth. Quarterly Journal of Economics 115: 469-498.

Galor, O. and Zeira, J., 1993. Income distribution and macroeconomics. Review of Economic Studies 60: 35-52.

Gould, E., Maov, O. and Weinberg, B. A., 2001. Precautionary demand for education, inequality and technological progress. Journal of Economic Growth 6:285-315.

Henderson D. J., Carroll, R. J. and Q. Li, 2008. Nonparametric estimation and testing of fixed effects panel data models. Journal of Econometrics 144: 257-275.

Huang, Ho-Chuan River, 2004. A flexible nonlinear inference to the Kuznets hypothesis. Economics Letters 84: 289-296.

Huang, Ho-Chuan River, Lin, Yi-Chen, and Yeh, Chih-Chuan, 2009. Joint determinations of inequality and growth. Economics Letters 103: 163-166.

Iradian, Garbis, 2005. Inequality, poverty, and growth: cross-country evidence. Working Paper No. 05/28. International Monetary Fund, Washington.

Julien, Gourdon, 2007. Openness and inequality in developing countries: a new look at the evidence. MPRA Paper No. 4176, November.

Kanbur, Ravi, 2000. Income distribution and development. In: Atkinson, Anthony B., Bourguignon, Francious (Eds.), Handbook of Income Distribution. Elsevier, 791-841.

Keane, M.P. and Prasad, E., 2002. Inequality, transfer, and growth: new evidence from the economic transition in Poland. Review of Economics and Statistics 84(2): 324-341.

Kuznets, Simon, 1955. Economic growth and income inequality. American Economic Review 45: 1-28.

Li, Hongyi, Squire, Lyn, and Zou, Heng-Fu, 1998. Explaining international and intertemporal variations in income inequality. Economic Journal 108: 26-43.

Lin, Shu-Chin, Huang, Ho-Chuan, and Weng, Hsiao-Wen, 2006. A semiparametric partially linear investigation of the Kuznets' hypothesis. Journal of Comparative Economics 34, 634-647.

Milanovic, Branko, 2005. Worlds Apart: Measuring International and Global Inequality, Princeton University Press.

Mushinski, David W., 2001. Using non-parametrics to inform parametric tests of Kuznets' Hypothesis. Applied Economics Letters 8: 77-79.

Ram, R., 1991. Kuznets' inverted U hypothesis: evidence from a highly developed country. Southern Economic Journal 57, 1112-23.

Ravallion, M., 2003. Inequality convergence. Economics Letters 80: 351-356.

Sen, Amartya K., 1991. Presidential address: the nature of inequality. In Arrow, Kenneth (Ed.), Issues in Contemporary Economics, Volume 1: Markets and Welfare, MacMillan Academic and Professional Ltd., 3-21.

Sen, Amartya K., 1992. Inequality Reexamined, Harvard University Press.

Sen, Amartya K., 1993. Capability and well-being. In Nussbaum, Martha, and Sen, A. K. (Eds), The Quality of Life, Clarendon Press, 30-53.

Wan, Guang Hua, 2002. Income Inequality and Growth in Transition Economies: Are Nonlinear Models Needed? Discussion Paper No. 2002/104, World Institute for Development Research, United Nations University.

Wang, Xiaolu, 2006. Income Inequality in China and Its Influencing Factors. Research Paper No. 2006/126, World Institute for Development Research, United Nations University.

Wood, Adrian, 1997. Openness and wage inequality in developing countries: the Latin American challenge to East Asian conventional wisdom. World Bank Economic Review 11(1): 33-57.

Appendix

Sample of 30 OECD countries and years:

Australia, 1967-69, 76, 78-79, 81-82, 85-86, 89-90, 94-96, 2002. Austria, 1987, 1991, 1995, 2000. Belgium, 1979, 85, 88, 92, 96, 2000. Canada, 1965, 67, 69, 71, 73-75, 77, 79, 81-88, 91, 94, 97, 2000. Chile, 1968, 71, 80-94, 98, 2000. Czech Republic, 1991-97, 2002. Denmark, 1963, 76, 78-95, 97, 2000. Finland, 1962, 77-84, 87, 91, 95, 2000. France, 1962, 65, 70, 75, 79, 81, 84, 89, 95. Germany, 1973, 75, 78, 80, 81, 83-85, 89, 94, 97, 98, 2000. Hungary, 1972, 77, 82, 87, 89, 91, 93-97, 99. Ireland, 1973, 80, 87, 94, 99, 2000. Italy, 1967-69, 71-84, 86, 87, 89, 91, 93, 95, 98, 2000. Japan, 1962-65, 67-82, 85, 88-90, 93, 98, 2002. South Korea, 1965, 66, 70, 71, 76, 80, 82, 85, 88, 93, 98, 2003. Luxembourg, 1985, 91, 94, 98, 2000. Mexico, 1963, 68, 69, 75, 77, 84, 89, 92, 94, 98, 2002. Netherlands, 1962, 75, 77, 79, 81-83, 85-99. New Zealand, 1973, 75, 77, 78, 80, 82, 83, 85-87, 89-91. Norway, 1962, 63, 67, 73, 76, 79, 82, 84-91, 95, 96, 2000. Poland, 1991-97. Portugal, 1973, 80, 89-91, 94, 97. Slovak Republic, 1988-97, 2005. Slovenia, 1991-93, 97, 2002. Spain, 1965, 73, 75, 94, 2000. Sweden, 1963, 67, 75, 76, 80-96, 2000. Switzerland, 1982, 92, 2002. Turkey, 1968, 73, 87, 94, 2003. United Kingdom, 1964-76, 79, 85, 86, 91, 95, 2002. United States, 1960-91, 94, 97, 2000.

Sample of 45 Non-OECD countries and years:

Argentina, 1989, 92, 98, 2001. Armenia, 1994-1997. Bahamas, The, 1970, 73, 75, 77, 79, 86, 88, 91-93. Bangladesh, 1963, 66, 67, 69, 73, 77, 78, 81, 83, 86. Barbados, 1979, 96. Belarus, 1995-97, 2002. Brazil, 1970, 72, 76, 78-91, 93, 96, 98, 2002. Bulgaria, 1981-97, 2003. China, 1970, 75, 78, 80, 82-99, 2001. Colombia, 1964, 70, 71, 74, 78, 88, 91, 94, 98, 2003. Costa Rica, 1961, 69, 71, 77, 79, 81, 83, 86, 89, 93, 98, 2001. Cyprus, 1990, 96. Dominican Republic, 1976, 84, 89, 92, 96, 97, 2003. Ecuador, 1968, 88, 93, 94, 95, 98, 2003. El Salvador, 1965, 77, 89, 94, 95, 97, 2002. Estonia, 1990-94. Gabon, 1975, 77. Guatemala, 1986-87, 89, 98, 2002. Honduras, 1968, 89-94, 98, 2003. Hong Kong, 1971, 73, 76, 80, 81, 86, 91, 96, 98. Israel, 1986, 92, 97. Jamaica, 1958, 2003. Kazakhstan, 1993, 96, 2002. Latvia, 1995, 96, 98, 2002. Malaysia, 1967, 70, 73, 76, 79, 84, 89, 95, 97. Nepal, 1976, 77, 84. Nicaragua, 1998, 2001. Nigeria, 1959, 81, 82. Pakistan, 1963, 64, 66, 67, 69, 70. Panama, 1969, 70, 79, 80, 89, 95, 97, 2002. Paraguay, 1990, 95, 98, 2001. Peru, 1961, 71, 81, 96, 2002. Philippines, 1961, 65, 71, 75, 85, 88, 91, 94, 97. Puerto Rico, 1963, 69, 79, 89. Romania, 1989-92, 94, 98. Russian Federation, 1990, 93-96, 98. Senegal, 1960, 95. Singapore, 1973, 78, 80, 89, 92, 97, 2003. South Africa, 1990, 93, 95. Sri Lanka, 1963, 69, 73, 79, 80, 81, 86, 87. Thailand, 1962, 68, 69, 71, 75, 81, 86, 88, 90, 92. Trinidad & Tobago, 1971, 76, 81, 88, 94. Uruguay, 1989, 92, 98. Uzbekistan, 1990, 2002. Venezuela, Rep., 1962, 71, 76-79, 81, 87, 89, 90, 93, 99, 2000.