An Empirical Analysis of the Relationship Between Inequality and Innovation in a Schumpeterian Framework

Ozan Hatipoglu*
Boğaziçi University†

20/02/2007

Abstract

I empirically investigate the non-linear relationship between inequality and innovation in a Schumpeterian setup where growth is expressed by the rate of innovations. In this framework income distribution plays a role in determining the dynamic market sizes for innovators and therefore is a major determinant of growth. By using two new cross-country inequality data sets, I find support for an inverted U-shaped relationship between inequality and innovative activities. This result is robust to two common inequality definitions and several parametric and non-parametric estimation procedures.

JEL classification: 015,031

Keywords: inequality, innovation, patents.

*This paper could not have been completed without the support and guidance by William Easterly during my PhD years at NYU. I would like express gratitude to Jess Benhabib for his insight and helpful comments. I also would like to thank seminar participants at New York University, Columbia University and CUNY. All errors are author's own.

†Department of Economics, Natuk Birkan Hall, Bogazici University, Bebek , 34342, Istanbul, Turkey. E-mail: ozan.hatipoglu@boun.edu.tr

1 Introduction

Since Kuznets' (1955) seminal work, one of the widely researched and documented phenomenons of the late 20th century has been the relationship between inequality and growth. In this literature, inequality affects growth through either efficiency losses embedded in redistribution, or through loss of resources and productive capacity due to rent-grabbing behavior in social conflicts induced by inequality. Whereas there are ample empirical studies which test such links, there have been very few works that investigate the effect of inequality on growth through technological progress. This is surprising since the need for such investigation is called for on several grounds.

First, one might argue that accelerated technological progress is the main source of growth in the long run, because it channels resources towards more efficient production and thereby releasing them. On this reason alone, it seems natural to study the effect of inequality on innovation besides income growth to get a better understanding of how income distribution affects growth process. A relevant framework is the Schumpeterian setup where growth is achieved through the arrival of new and more efficiently produced goods. Inequality not only affects market sizes for new goods and profitability of efficient firms, but also income growth through ownership facility in a dynamic general equilibrium context. Second, as recent research has shown, inequality might not be exogenous to technological progress, and technological progress might not be exogeneous to inequality. An increase in the supply of skilled labor, for instance, might direct the technological progress to be skill-biased[1]. Whereas the increase in wage inequality as a result of technological progress has been widely researched and documented, the reverse link has rarely attracted attention and the empirical research has mostly ignored the endogeneity problem. Third, empirical studies which look at the effect of inequality on growth, have provided weak support and more often than not

[1] Whether the information technology revolution directly lead to rising inequality favoring the skilled, or the increased supply of skilled workers induced skill-biased technological change rising inequality, is still a puzzle. Nevertheless,the implications on inequality remain the same. See Acemoglu (98), Galor and Moav (00) and Violante (01) for respective arguments.

conflicting answers[2]. Even though some of this confusion can be attributed to the problems with data, the sign of the relationship is not clear because results are not robust to estimation techniques or inequality indices[3]. And finally, the observation that widening inequality is generally coupled with accelerated technological progress indicates that a Kuznets type of nonlinear relationship might be at hand. This requires a different approach to the estimation problem.

In this paper, I argue that an alternative empirical investigation based on Kuznets' idea and the Schumpeterian hypothesis might explain the incompatible empirical evidence. Assuming technological progress is driven by innovations and innovations are determined by the demand structure based on the underlying income distribution, this paper empirically investigates how inequality affects the innovative activity in a cross-country setting. Using two new data sets on inequality, I estimate several dynamic panel data models, including a non-parametric setup, to test the validity of the hypothesis that innovation and inequality are negatively related at high levels of inequality and positively related at low levels of inequality. The main conclusion is that the relationship between innovative activity and inequality can be described by a Kuznets type inverted U shape. This finding is also consistent with recent theories of inequality and growth[4].

An inequality-determined demand structure can be described by the distribution of demand across goods at a given time as a consequence of the underlying income distribution. Such a demand distribution can be easily found , when one assumes people have hierarchic preferences, which ranks goods in an order in such a way that the highest goods are the most

[2]Among the many conflicting recent reports one can cite Forbes (2000) who argues that the relationship between inequality and subsequent growth is positive at least in the short run. Barro (2000) finds a negative relationship between inequality and growth in developing countries and a positive relationship between inequality and growth in developed countries. Banerjee and Duflo(2003) argue that fitting linear models is inappropriate for explaining this relationship. They argue that any change in inequality will cause subsequent growth to fall.

[3]Deininger and Squire (1997) have provided an extensive data set on inequality used in most of the subsequent studies. Atkinson and Brandolini(2001), Banerjee and Duflo(2003), Forbes(2000), Dollar and Kraay(2001) and Galbraith and Kum (2003) argue that the Gini coefficients in this data set are not fully reliable.

[4]See Murphy, Shleifer and Vishny(1989), Baland and Ray(1991), Zweimuller(2000), Benhabib(2003).

luxurious and the lowest are the most basic ones. In a Schumpeterian setup, Engel's law gives us a first clue as to what the sources of discrepancies between cross-country empirical findings might be. When people have hierarchic preferences, they first spend proportionally more on new and efficiently produced goods as their incomes rise. At the top of hierarchy, however, there are luxuries which are produced by more traditional and inefficient methods because of chronic low demand. In a society, where assets are highly concentrated, a redistribution might increase growth if it increases the total demand for innovators. Suppose the growth rate is determined by the number of new goods or new firms entering into the monopolistic sector. The entry rate into the monopolistic sector, where new goods are produced through R&D, is higher if the firms face a demand increase in the near future as a result of decreasing inequality. One such case is a redistributive scheme which makes poor just rich enough to afford the innovators' product - maybe now maybe in the near future - without making the rich poor enough such that the rich forgoes consumption of the innovators' product today. This is, in effect, a Pareto improving allocation in which resources are freed up to be used in the most efficient sector. Such a scenario is most likely to occur when inequality is already high. On the other hand, if the rich becomes just poor enough such that the demand for innovators' product falls, reducing inequality might further hurt growth. The latter is likely to occur when the inequality is already low. Moreover, if the low inequality is coupled with a high purchasing power, a further decrease in inequality shifts resources away from efficient production to other areas such as luxuries causing inefficiencies in production. This analysis suggests that the relationship between inequality and innovation might be described by an inverted U - shape. A theoretical background for this type of model can be found in Foellmi and Zweimüller (2005) where a higher initial demand for an innovator increases the likelihood that it will innovate.

The argument, that luxury producers are less efficient than the monopolists, deserves some discussion. In general, one can assume that the level of competition among luxury producers is lower which lead to inefficiencies in the production of such items. Moreover, the markets for luxury goods have been traditionally small and the demand rather inelastic, hence there are overall less incentives to innovate. The link between competition and efficient production is highlighted by a recent line of research emphasizing the role of mergers in monopolistic industries in increasing efficiency by reallocating resources to the more efficient R&D sector[5] Since the high end producers have little incentive to form mergers because of brand protection concerns, they allocate less resources to R&D among other factors.

Given the above setup, the relation between innovation and inequality is expected to be negative at high levels of inequality, and positive at low levels of inequality. In addition, since high incomes are generally associated with lower inequality and vice versa[6], we can expect that the inequality - growth relationship to be positive in rich countries, and negative in poor countries as shown by Barro(2000). The implication of the above analysis is that the inequality-growth relationship is inversely U-shaped.

In the light of this discussion it is natural to look at the effect of inequality on the level of innovations before looking at growth, especially if the aim is to test these theories within a Schumpeterian context in which new technologies are embodied in new goods. Sedgley (2006), for instance, finds that innovation is a major factor in explaining the growth of U.S. economy. This approach is rarely taken in the empirical literature where most studies link inequality to the growth of real gross domestic product per capita. The effect of inequality on growth has been both theoretically and empirically studied by previous researchers extensively albeit inconclusively[7]. The effect of inequality on the level innovations has also been theoretically analyzed in the literature as in Murphy, Shleifer and Vishny (1989) and Foellmi and Zweimüller (2005). However, within the set of models where innovation is the source of growth, effect of inequality on innovation has rarely been empirically studied. The only work, that the author is aware of is by Weinhold and Reichert (2005) who look at the effect of the size of the middle class on innovations by controlling for institutional features. This paper differs from Weinhold and Reichert's paper in several respects. First,

[5] See Jovanovic and Rousseau (2001), Carol and Hanan (2000), Faria(2002)

[6] The empirical support for this association is extensive. See Galbraith(2002), for instance.

[7] For a survey of this literature see Aghion,Caroli, and Penalosa (1999)

I particularly focus on non-linearity as predicted by the previous theoretical models on inequality and growth. Second, due to the endogenous nature of inequality I specify the arrival of innovations as a non-parametric Poisson process. Third, I use a different and larger data set which includes several inequality definitions such as a Gini coefficient and Theil index as opposed to just the size of the middle class. In this paper, the demand for innovations is not solely determined by size of the middle class but also how relatively rich the countries are compared to others as well as their innovative capacity.

During 1980s there has been a worldwide increase in inequality which is linked to the information technology revolution by subsequent research. One implication of this phenomenon for empirical study is the fact that skilled-biased technological change will increase both inequality and innovative activities causing a spurious relation between them. This type of endogeneity is not adequately addressed in the previous literature linking inequality to growth. Moreover, each country has its own institutions and innovation policies which would play a role in the amount of innovations. I address these issues by employing several methods. First, by making use of robust panel data techniques, specifically a GMM estimation method by Arellano and Bover(1995) and Blundell and Bond (2000) and a Kernel density estimator by Hausman and Newey (1995), I control for the endogeneity and fixed effects. Second, in line with the traditional modeling of innovations, I introduce a hazard model to estimate arrival of innovations as determined by the underlying income distribution to check the robustness of the empirical model. Finally, I estimate non-linear specifications within the original and the non-parametric setup.

The difficulty with interpreting any demand based model of innovation is to determine whether innovative activity is pro or anti-cyclical, and whether the changes in demand are exogenous to the process of producing and using innovations. This is also important in determining the appropriate lag structure and the expected signs in the econometric model. There are two main theories regarding the source of innovations. In supply push models, innovations are driven by exogenous shocks to scientific knowledge, whereas in demand pull models, changes in profitability stimulates investments in R&D . There is some, albeit not strong, empirical evidence which suggests that innovations are mostly demand driven. For instance, Geroski and Walters(1995) show evidence that variations in economic activity Granger causes changes in innovative activity but the opposite is not necessarily true. Similarly, using Italian data, Piva and Vivarelli (2006) show that firms' research activities respond to sales. In line with the demand pull theory, I use demand variables to control for innovation. Technology push theories emphasize the importance of technological opportunity to innovation, therefore today's innovative activities at least partly determine future level of innovation. Since the empirical support for a demand pull theory is not a strong one, I make use of both aspects of innovative activity, i.e., both market demand variables and lagged values for innovation are included in the empirical modeling.

The plan of the paper is as follows. In the next section I introduce the empirical issues and the data set used. In the third section I present the empirical model and the estimation results based on dynamic generalized method of moments (GMM) and interval location approach. In the fourth section, I introduce nonlinearity to the model by developing a semi-parametric hazard model. Using the hazard model and other non-linear fits to the original model I present the parametric and the non-parametric estimation results. Section six concludes.

2 Empirical Issues and Data

There are many potential problems with cross-country empirical work on inequality and growth. One major problem with inequality growth regressions is the omitted variable bias or setting up a well specified estimable model. If for instance there exists an omitted variable which is negatively related to growth and positively related inequality (or vice versa), then the coefficients are negatively biased (Forbes, 2000). I address this issue by making use of a panel data where individual and time-invariant effects can be accounted and part of this bias can be adjusted for. Another problem is that the measurement error, if present in one or more regressors, leads to inconsistency and least squares attenuation. It is well known that the national reports of inequality generally underreport inequality because of surveying problems. This is more evident in countries where inequality is high. In fact, this been a major concern especially in studies using the Deininger and Squire (1996 and hereafter D&S) data set. D&S has fewer than 700 observations in the high-quality subset with infrequent observations for underdeveloped countries and there are only five countries for which D&S has annual observations over long periods of time.[8] Atkinson and Brandolini(2001) argue that D&S fails to provide an adequate or accurate, longitudinal and cross country coverage. Further criticisms on the Gini coefficient presented in D&S are brought forward by Galbraith and Kum(2003). For instance, Gini coefficient in D&S does not reflect the same unit of account, i.e., some of the samples are based on expenditure data whereas others are based on income data. Moreover, the coverage is neither extensive nor representative. To give an example nearly most of the countries selected belong to OECD. Some regions such as the Sub-Saharan Africa are underrepresented in the sample. This probably has caused researchers to develop empirical models that inherently take into account the biases associated with surveys of inequality either by filling the gaps or restricting the data set to a more balanced panel.[9] Banerjee and Duflo (2003) advanced the models a step further by giving a theoretical role to measurement error in inequality as a statistical determinant of inequality - growth relationship. They argue that statistical agencies are more likely to mismeasure when their societies are under stress and experiencing lower growth rates. One of the original aspects of this article is the use of two new inequality data sets. The first one is based on the Theil index reflecting industrial wage inequality(UTIP-UNIDO 2002) and the second one is based on household income inequality (Galbraith and Kum 2003). Both data sets address many of the problematic issues associated with the D&S "high quality" data set which is frequently used by researchers. They are less plagued by measurement problems[10] and wider in scope

[8] Great Britain, Bulgaria, India , Taiwan and United States

[9] Sala-i Martin(2002) interpolates the data and Forbes(2000) and Banerjee and Duflo(2003) restrict the data set to five-year intervals.

[10] UTIP-UNIDO data set is based on the source data.

than D&S.[11]

Researchers have used patents and R&D as indicators in the analysis of technical change[12] In the firm level, patent numbers and R&D figures are used to study a wide variety of issues such as the productivity effect of innovation, firm size and the nature of spillover[13], whereas in the aggregate level both measures are taken to reflect the technological capacity of industries and countries. There has been a recent improvement in the quality of both measures as a result of the development of measurement standards and computerization of patent offices. Both measures capture different aspects of the innovation process. The number of R&D employees or R&D can be viewed as resources devoted to innovative activity, whereas the number of patents shows the results of innovative activity. The choice of patents in this paper is not arbitrary. First, a patent is more likely to be obtained, if the innovated product faces future competition. This is related to the future market size the firm will be facing for its new product *vis-a-vis* the inequality level; an important aspect of the hypothesis this paper is trying to prove. A patent is not the only method to protect profits, nor does it capture all the innovation output. Nevertheless, given the active effort and trained statisticians required to measure R&D expenditures, patent statistics are less prone to measurement problems, especially in developing countries. Moreover, R&D statistics are not a measure of innovation output but an input to the innovation process. The use of patents as an indicator of innovation is not uncommon. Aghion et al. (2002) use patents as a proxy to innovation in examining the relation between competition in the product market and innovation.

In terms of econometric modeling, the patents as an endogenous variable is also less prone to the problems of the inequality and growth literature. There are fewer variables obviously correlated with both inequality and the level of innovation rather than with inequality and growth, which reduces the possibility of omitted variable bias. One can also expect less serial

[11]Luxembourg Income Study(LIS) and World Income Inequality Data Set(WIID) are well known alternative data sets. LIS is restricted to wealthy western countries and WIID is built on D&S data and therefore they are not used in this study.

[12]See Pavitt (1985) , Griliches (1990).

[13]See Lach (1995)

correlation due to business cycles in patent data than in per capita income growth data.

The main criticism of the use of patents is the fact that not all inventions will be patented. Even though the incremental and imitative innovations represent a large and an increasingly important part of innovation activities, they are not covered by patent statistics. The most obvious shortcoming in this regard is the undercoverage of innovation activities in small firms. The small firms are less likely to engage in research, but if they do, they invest more compared to medium sized firms and less compared to large firms. In addition, other statistics suffer also from the same type of heterogeneity. Moreover, other input variables such as R&D expenditures do not reflect the total cost of innovation. Brouwer and Kleinknecht(1994) find that the total product innovation expenditure to be four times the amount of product-related R&D expenditure.

The patent data in this study are taken from industrial property statistics published by World Intellectual Property Organization. It is the number of patents granted each year. This data is based on direct surveys of the statistical agencies around the world and provides coverage for over 40 years, over 100 countries and has 2504 observations. The US patent data is obtained from Bureau of Labor Statistics and consists of non-medical patents granted both to domestic and foreign applicants.

Inequality data is provided by the University of Texas Inequality Project. The two inequality measures used in this paper are the Theil measure of industrial pay inequality (Theil) by UTIP-UNIDO (2002) and the household income inequality (HCIN) by Galbraith and Kum (2003). The first set is based on the Industrial Statistics database published annually by the United Nations Industrial Development Organization, and it is a set of measures of the dispersion of pay across industrial categories in the manufacturing sector [14]. Overall wage inequality has been widely used as an alternative to income inequality in the literature. Atkinson (1997) indicates that earnings and wage inequality are the main

[14] Note that incomes outside of manufacturing are generally not covered in this data set and transfers and taxes are not covered at all. Therefore any changes in the structure of the employment is likely to bias the Theil statistic.

components of income inequality in US. The second set includes estimates of gross household income inequality, computed from a regression relationship between the Deininger & Squire inequality measures and the UTIP-UNIDO pay inequality measures. By controlling for the source characteristics in the D&S data and for the share of manufacturing in total employment, it provides over 3000 estimates which include a much larger and balanced set for the developing countries than the Deininger and Squire data set. (156 countries, 3194 observations, 1963-1999 time-span).

The data on educational attainment comes from two sources: Barro and Lee(1997) and World Bank. The capital per worker data comes from Easterly and Levine(1999). The final consumption expenditures as a percentage of gdp and foreign direct investment data are taken from World Bank Development Indicators. The price level of investment data which is measured as the purchasing power parity of investment/exchange rate relative to United States is taken from Penn World Table Version 6.1 (Summers, Heston and Aten 2002).

An initial look at the data is provided in Table 1 where I report mean and standard deviation for selected variables for two benchmark years. Since the purpose of this paper is to look at non-linearities, the data is divided into subsamples such as low, middle and high. The choice of these intervals except inequality[15] uses simple and arbitrary ranges where the upper and lower tiles represent high and low[16]. Note that this table is a static picture of the indicators and does not say anything about the link between inequality and innovation or growth. Nevertheless, one should note that innovative activities were higher in 'high income' countries and lower in 'high inequality' countries in both benchmark years. The 'medium inequality' countries had more innovative activities than 'low inequality' countries in 1965 but the opposite is true in 1999 although the difference is not significant. 'High growth' countries had comparably less inequality than 'low growth' countries on the average in 1965 but not so in 1999, whereas the inequality in rich countries is also significantly lower

[15] The choice of low, medium and high inequality intervals is done using the interval location approach explained in the next section.

[16] Even though the choice of intervals is rather arbitrary, the comparative results obtained are robust to the changes in intervals within acceptable distances.

in average than it is in poor countries. The 'high growth' countries had significantly lower innovative activities in 1999 compared to 1965. This might suggest that the 'high growth' countries innovated in the early take off phases whereas they adopted or imitated in the later phases.

INSERT TABLE 1.

3 The Model

Suppose the stock of knowledge in an economy is represented by the number of innovations up to date t, $n(t)$. Then growth of this number is represented by $g = \frac{\dot{n}(t)}{n(t)}$. Modeling a growth rate of this type necessitates the calculation of the stock of knowledge in terms of innovations[17]. In practical terms , this requires the calculation of innovations up to time t for each country. One major difficulty with this approach is the lack of appropriate data starting from a specified date of historical origin. One possible remedy is to assume that each country initially has no stock of knowledge and do the calculations from the beginning of the available data. However, this would be quite unrealistic considering the differences in initial conditions between countries. Nevertheless, since each arrival adds to $n(t)$, yearly arrivals do represent growth of the stock of knowledge, if not the growth rate. In fact, growth regressions on inequality have mostly used the same setup in which only changes between two periods are recorded as growth. One can control then for the differences in stock of knowledge or technological opportunities by using another variable such as income per capita. One drawback with applying the traditional methodology here is that it allows negative growth, which does not reflect the nature of knowledge creation. In a broader view, one might also want to model the 'stochastic' property of innovations (i.e. current innovations being explained at least partly by past innovations). Taking into account the

[17] Other measures of stock of knowledge based on R&D flows and international trade exist in Coe and Helpman(1995) and Keller (2001) which are not applicable within the context of this paper.

nature of innovation process, we can at least control for the stock of the knowledge. In the light of this discussion the empirical model can be written as;

$$y_{i,t} = \beta y_{i,t-a} + \gamma f\left(h_{i,t-a}\right) + \delta X_{i,t-a} + \alpha_i + \eta_t + \varepsilon_{i,t} \tag{1}$$

where $y_{i,t}$ is the level of innovation proxied by the logarithm of patents cited each year. Since the number of available products is assumed to be equal to the sum of all previous innovations, patents cited each year represent the growth of this number. $f\left(h_{i,t-a}\right)$ is a function of the inequality index, $X_{i,t}$ are the control variables, α_i are country dummies, and η_t are time dummies.

The first question is what should be the right set of control variables. The empirical literature on innovations generally uses variables that capture market distortion, assuming innovation is determined by degree of competition. In a cross-country setting, the purchasing power parity of investment goods is such a measure. For the base specification I choose the control variables $X_{t,i}$ to be the level of per capita income GDP, capital per worker (CPW), and the price level of investment (PPPINV). I also include the final consumption expenditures as a percentage of GDP(FCE), the level of foreign direct investment(FDI), and education variables. The capital per worker (CPW) is a measure of the productivity which is correlated with level of innovation based on the assumption that R&D is capital intensive. The expected sign is positive and significant. PPPINV is a degree of market distortion, and along the historical lines of the innovation literature, it is included as an explanatory variable. The link between the degree of market competition and innovations is analyzed by several authors in the literature[18]. Aghion et. al. (2002) find for instance that relationship between product market competition and innovation is an inverted U shape. The fact that rich economies can support large markets for new products despite large differences of wealth within their populations, might cloud any evidence on the link between inequality and growth. To control for such market size effects I use final consumption expenditures as

[18] See for example Fellner (1951), Arrow (1962), Levinn and Reiss(1984), Geroski(1990), and Geroski(1995)

a percentage of GDP. This variable includes both government and household consumption. The innovative capacity is also closely linked to the per capita income level (GDP). Income is a demand variable that is correlated with innovations in a demand based analysis. Per capita income controls for the level effects created by the cyclical nature of innovative activity.

3.1 The Linear Model

In order to facilitate the spline regressions I rewrite the model by letting $f\left(h_{i,t-a}\right)=h_{i,t-a}$.This gives us the linear model,

$$y_{i,t} = \beta y_{i,t-a} + \gamma h_{i,t-a} + \delta X_{i,t-a} + \alpha_i + \eta_t + \varepsilon_{i,t}$$

which can be estimated by the standard panel data techniques such as fixed effects[19] or GMM estimation based on Arellano and Bover(1995).

As mentioned above one of the difficulties in setting up an estimable model when using patents as a proxy for innovation is the cyclical nature of innovative activities. In other words, how does one interpret the signs of demand variables and/or determine the number of lags?

First, there is a time period between a firm foresees a demand jump in the near future due to decreasing inequality until the demand jump actually occurs. This time lag is not easy to distinguish from the period during which a firm's undertake of research leads to a patent and during which direction of the cycle this activity occurs. Moreover, firms generally do not file for patents and innovate simultaneously. If, during a recession, the value of existing profits falls faster than the value to be attained by innovating net of research costs, then firms will turn to R&D during cyclical downturns. This makes the innovative activity countercyclical. This argument fails, however, when there are complementaries in innovation. Foellmi and Zweimüller (2005) argues that more resources are diverted to research when a demand jump

[19] Note that if $\beta \neq 0$ then the OLS, random and fixed effects estimators are biased.

is to occur in the nearer future. This implies that research activity and patenting are more likely to occur in upturns. Moreover, the firms' incentive to utilize full benefits of patenting causes them to file before downturns.

The markets ability to absorb new markets at any time is limited. When a wave of imitative or 'me too' products arrive, the profitability of each of them falls during a recession. Innovative activities will increase only if growth is high enough to create demand expansion making them procyclical. Product innovations are likely to cluster during economic booms which generate enough income to absorb these products. Finally, there are also strong incentives to make investments in organizational capital during recessions which reduce innovation output. Hence a cyclicality between innovations and growth is expected. The empirical literature gives more support to the procyclical nature of innovations (Geroski and Walters 1995). The second discussion involves the determination of the number of lags in the empirical model. In this setup, it should not take too long for market expansions to be exploited by innovating firms. The choice of lag length below is then a compromise between the Akaike information criterion and the modeling criteria imposed by the procyclicality as above. In the light of this discussion, I suggest the following linear model

$$PATENT_{i,t} = \beta_0 PATENT_{i,t-2} + \beta_1 THEIL_{i,t-2} + \beta_2 PPPINV_{i,t-2} + \tag{2}$$

$$\beta_3 CPW_{i,t-2} + \beta_4 FDI_{i,t-2} + \beta_5 FCE_{i,t-2} + \beta_6 GDP_{i,t-2} + \alpha_i + \eta_t + \varepsilon_{i,t}$$

In Table 2, I report the linear estimation results (i.e. $f(h_{i,t}) = h_{i,t}$) using the base specification and adding different control variables such as percentage of labor force with a college education (SETETGR), population (POP), male secondary education (ME),and female secondary education (FE) [20]. The estimations are run for two different inequality

[20] The education variables are taken from Baroo and Lee(1997). Population data is taken from World Bank.

indices, THEIL and HCIN and using several methods.[21]. The coefficient of inequality[22] remains negative and significant with respect to different choices of control variables and estimation reports except in two cases.[23] The estimation is repeated for different choices of lags which do not affect the sign of inequality or other coefficients but income. With fixed effects, the sign of income is positive and significant when I choose a lag of 5 or less years. It is insignificant with GMM except for 3 or 4 years of lags. With pooled OLS and random effects, the sign is positive and significant at all levels. The education variables are insignificant and their inclusion does not have a significant effect on other variables in almost all of the estimations, therefore they are dropped. A Hausman chi-squared test for fixed effects based on Wald statistics with 6 degrees of freedom results in 25.32 which rejects the random effects model.

INSERT TABLE 2.

To facilitate GMM estimation, identification of the model requires restrictions on the serial correlation properties of the error term $\varepsilon_{t,i}$. In these models it is assumed that if the error term was originally autoregressive, the model has been transformed so that coefficients satisfy a set of common factor restrictions. Therefore only serially uncorrelated or moving average errors are explicitly allowed. Generally, the $\varepsilon_{i,t}$ are assumed to be independently and identically distributed across individuals with zero mean, but arbitrary forms of heteroskedasticity across units and time are also possible. The assumption of no serial correlation in $\varepsilon_{i,t}$ is crucial for the consistency of the estimators, since they instrument the lagged dependent variable with further lags of the same variable, therefore they are reported for each GMM estimation.

[21] See Appendix I for a brief discussion of the GMM estimator used in this study.
[22] Other coefficients and their statistics are given in Appendix.
[23] The control variables used in the estimations reported in Table 2 are as follows:
1) PPPINV, CPW, FDI, FCE, GDP
2) PPPINV, CPW, FCE, SETETGR
3) PPPINV, CPW, SETETGR
4) PPPINV, CPW, FE
5) PPPINV, CPW, ME
6) PPPINV, CPW, FE, ME

3.2 Threshold Identification

I address the issue of non-linearity first by identifying thresholds above or below which the sign of the relationship reverses sign. A relevant procedure is to employ spline regressions within the established dynamic panel data methods[24]. In order to facilitate the spline specification, it is useful to identify the locations of the thresholds. I run spline regressions using a dummy variable within the main specification of (1) and I estimate multiple parameters by systematically changing the initial value of the dummy along the range of inequality values. I expect the path of the slopes from these regressions to indicate the possible structural changes on the coefficient of inequality. Moreover, the estimated slopes from these regressions should hint to the location and size of the ranges where inequality and growth are positively or negatively related.

If the inequality and growth are indeed positively related as we move from complete equality to low levels of inequality as the theory predicts, the coefficient on inequality should be positive and statistically significant along the low inequality values. As we move further away from equality, the slope should start to decrease, and after a certain threshold it should become again negative and statistically significant.

I use a second approach in which I run the regressions for each interval separately and contrast the coefficients with the full sample. I check if the coefficients obtained from these regressions are significantly different from the ones obtained for the full sample. For instance, I expect the slopes for the outer regions to be alternating in sign and to be significantly higher in absolute terms than those in the full sample, if the relation between inequality and growth is indeed non-monotonic.

I also check the consistency of the inequality coefficient by systematically changing the size of the identified low, middle and high inequality intervals. I expect to find that as intervals get wider, the coefficient of inequality should start to decrease in absolute terms.

[24] See Chong and Zanforlin(02) for a similar treatment.

Since this methodology might be plagued by a sample selection bias, I also use a restricted approach in which I use the full sample in all regressions but place two dummy variables at the beginning of these conjectured intervals. I further check for consistency again by changing the size of the intervals.

Table 3 reports the estimations using the interval location approach explained above. I place a dummy variable to the benchmark specification (2). The initial values of the dummy are placed along the inequality values starting from 0.022 for THEIL and 20 for HCIN which replaces THEIL in the benchmark specification. Estimation results using fixed effects lend support to the hypothesis that as inequality increases the inequality and growth relation reverses sign from positive to negative. The slope for the range of THEIL values, 0.022 to 0.0325, are positive and significant. The slopes become insignificant and negative for the THEIL values from 0.035 to 0.06 and become negative and significant again. The striking result is that the slopes follow a smooth line monotonically decreasing from positive to negative. If only for sample biases one would expect the slopes to remain positive even though they should decrease. With GMM estimation the reversal of the sign remains, although the threshold above which this change occurs is now higher. The slopes remain both positive and significant until 0.045 and become both negative and significant at 0.075. When HCIN instead of THEIL is used as an alternative inequality index results remain the same. With fixed effects, the slopes remain positive for HCIN values from 20 to 32. For the range, 41 to 56, the slopes are increasingly negative and significant which again supports the hypothesized inverse U shape. When GMM is used, the threshold again shifts upward but slopes follow the same pattern.

INSERT TABLE 3.

The fixed effects and GMM slope patterns for both types of inequality indices are shown in Figures 1-4. Figure 1 shows the slopes above and below conjectured thresholds for THEIL under fixed effects estimation . The lower end coefficients tend to drop sharply as threshold increases and become flat again when the threshold is further increased. The upper threshold smoothly declines and becomes negative. The combined changes in upper end and lower end slope coefficients support a non-monotonic, possibly inversed U-shape pattern. GMM estimation results for THEIL are graphed in Figure 2. They show a slightly different pattern in which the upper end coefficients decline vaguely and are always negative , whereas there is a sharp drop in lower end coefficients.

INSERT FIGURE 1.

INSERT FIGURE 2 .

Similar results are obtained when HCIN is used instead of THEIL. Figure 3 and Figure 4 show that the drops in lower end coefficients are not smooth with occasional jumps at the lower inequality levels. The upper end coefficients become eventually negative in both estimation methods as the threshold is further increased.

Having established these thresholds the main equation is tested using the regression results for each subsample as determined by above analysis against the regression results from the full sample. All slopes are significantly different.

INSERT FIGURE 3

INSERT FIGURE 4 .

I use the identified intervals to further explore the non-linearity using the main specification of (1) . I run both fixed effects and GMM regressions for the subsamples as determined by these intervals and test for the differences in slopes[25]. These results are reported in Table

[25] The lower sample includes observations where THEIL(HCIN) is less than 0.033(33) for fixed effects and THEIL(HCIN) is less than 0.045(31) for GMM. The upper sample has observations where THEIL(HCIN) is greater than 0.06(41) for fixed effects and THEIL(HCIN) is less than 0.075(52) for GMM.

4 and they show again a non-monotonic pattern. For the lower subsample the slopes are positively significant and for the upper subsamples they are negatively significant, except the fixed effects estimation using HCIN. As a simple way to check for local consistency, I allow the subsamples to change in both directions. In this case one expects the coefficients to increase in absolute terms, if thresholds are decreased for the lower subsample or if thresholds are increased for the upper sample. For example, if the THEIL inequality threshold is lowered to 0.028 for the fixed effects estimator, I obtain a coefficient that is higher than the original conjectured threshold of 0.033. Similarly, if I increase the upper THEIL threshold to 0.07 for the fixed effect estimator, I obtain a higher coefficient in absolute terms. This exercise is repeated for GMM estimators and HCIN inequality indicator. The results are similar. The signs of the coefficients are also as expected, i.e., for the low inequality sample, the slopes are positive and significant in both types of estimation and for both inequality indicators. They are also significantly different from the coefficients of high inequality sample.

INSERT TABLE 4.

Since the division of the whole sample to subsamples brings about a sample bias problem, I re-run the spline regressions using the whole sample, but placing dummies on the conjectured thresholds. The results are reported in Table 5. By using the same locations as in the unrestricted model, I estimate three slopes within the identified intervals. [26]This model is restricted in the sense that it is continuous across all intervals, i.e., there is no jump in the constant coefficient. Slope 1 is the estimated coefficient from the lowest value of the inequality index to the first threshold. Slope 2 is the estimated coefficient within the shown interval and Slope 3 is the estimated slope between the second threshold and upper limit. The middle columns of the left and the right hand side of the table are the identified intervals in previous methods. Reading the table from top to bottom, we see that except the fixed effects estimation of HCIN , the slopes within these identified intervals generally support the previous results. Slopes are decreasing across intervals. To see if this methodology is locally

[26]See Appendix III for a formal presentation of the procedure.

consistent, I increase and decrease the intervals by either keeping the lower threshold or the upper threshold fixed. Now, reading the table from left to right it is clear that increasing the higher threshold more often than not produces a lower Slope 3. In the same manner, increasing the lower threshold produces a lower Slope 1 in most cases. For slope 2, the same conclusion can not be fully confirmed. In both types of estimations using HCIN, the middle interval slope is higher, i.e., increasing the upper threshold does not decrease the middle slopes. The right and left columns at both sides of the table are also consistent in the above sense; the slopes are decreasing along the inequality level.

INSERT TABLE 5.

3.3 Non-linear estimation and non-parametric methods

The purpose of this section is to investigate how much inequality can explain the evolution of innovative activity under the hypothesis that firms globally respond directly to changes in inequality. Specifically, the focus is on the direct effect of inequality on the nature of innovative activity rather than the effect on growth within a Schumpeterian context. If there is a relation between inequality and innovation, how can it be described? The empirical modeling of innovations has generally assumed that the data generating process underlying the arrival of patents can be described by a Poisson density. In the framework of this article we can write this process as:

$$\Pr(y = n \mid h) = \frac{e^{f(h)n} - e^{f(h)}}{n!} \tag{3}$$

where y is the number of patents, f(h) is a function of inequality level and $e^{f(h)}$ is the hazard rate. In the panel data setting we can specify the first moment condition as[27]

[27] See Aghion et. al. (2005) for a similar treatment of the arrival of innovations.

$$E[y_{it} \mid h_{it}] = e^{f(h_{it})} \tag{4}$$

Now, if we specify a parametric linear function for $f(h)$, then we obtain the usual loglinear form employed in the literature for poisson estimation by maximum likelihood methods. The above inequality is then simply a conditional moment restriction of the classical model. The problem here is that it is not clear that inequality is exogenous to innovation. Recent literature consistently points out to the inequality creating effect of technological advances. Skilled biased technological progress is shown to create inequality both within and across industries in developed countries by several authors[28] Any increase in the supply of skilled labor might induce skill-biased technological change which then feeds back into employment choices and causes changes in inequality (Acemoglu, 1998). Increases in the supply of skilled labor on the other hand might be the result of sustained skilled biased technological progress which prevents the profitability of education or the skilled wage differential to fall. This endogeneity of the model necessitates a more robust procedure. A convenient way to approach this problem is to let $f(h_{it})$ be a non-parametric specification under which it identifies the innovation hazard. This specification permits dependence of the parameters in $f(h_{it})$ on other unknown functions as well as on unobserved variables.

Given the above setup, the data in hand is treated like a micro data panel, although the individual firm characteristics are absent. It should be noted that the differences among firms are ignored because of the extensive data requirement in a cross-country setup. However, country effects, time effects or other policies can be fully captured by the model. To further avoid spurious correlation, we can also control for time and individual country effects by estimating the following:

$$E[y_{it} \mid h_{it}, x_{it}] = e^{f(h_{it})+x'_{it}\beta} \tag{5}$$

[28] See Aghion(2001) for a version of this argument in a Schumpeterian setup.

where x_{it} represents time and country dummies. This new moment condition is now semi-parametric in the sense that $f(h_{it})$ is unknown and can be estimated in several ways.

To estimate the model, I first implement the Kernel method discussed in Robinson(1988) and applied to estimation of demand curves in Hausman and Newey(1995)[29] .The results of this estimation for THEIL index are shown in Figure 5 and for the HCIN index in Figure 6.[30]. In both cases an inverse U relation is again evident. In Figure 7., I compare a quadratic specification for $f(h_{i,t})$ in the base specification(1) and a Kernel estimation of the original model. Note the similarity of both curves despite the fact that they are based on two different modeling approaches[31]. Figure 8 shows the parametric estimation of the quadratic fit to the original hazard model. The inverted U shape has shifted right with inequality-innovation relationship being now negative at both ends. The coefficient of the squared term is -1.39 with a t statistic of -4.84 . The coefficient of the linear term is -0.05 with a t statistic of 0.04. Both time and country dummies are significant.

INSERT FIGURES 5-8

4 Conclusion

I study empirically the link between inequality and growth within a Schumpeterian framework. This amounts to taking growth as the increase in the stock of knowledge, specifically the number of new products. I find support for the hypothesis that inequality - innovation relationship is inversely U-shaped. The innovative activities of firms which drives the growth process depend on the demand structure vis-a-vis inequality. Departing from previous empirical literature on inequality, two new data sets were used. The overall relation between inequality and innovation is negative and negativity result is robust to definitions of

[29] See Appendix IV.

[30] A bandwidth of 0.05 is used for the Kernel estimator

[31] In the quadratic fit the coefficient of the linear term is 0.136 and has a t-statistic of 2.95. The coefficient of the squared term is -0.123 and has a t statistic of -2.03.

inequality and estimation procedures .The non-linearity results extend to a non-parametric and a semi-parametric setup as well. The findings in this paper overall are consistent with recent theoretical approaches to inequality - growth relationship which generally suggest a non-linear relationship.

I apply a systematic method which identifies the thresholds below or above which innovation and inequality positively or negatively related. The method involves an interval location approach which conjectures such thresholds and verifies them by using spline GMM regressions. I employ several estimation methods and two recent inequality indices that are less prone to statistical problems than the widely used Deininger and Squire Data Set. An unrestricted approach is employed by dividing the whole sample into subsamples and comparing the slopes between subsamples. The results obtained from the unrestricted approach are similar to the spline method. The consistency checks are made by locally expanding and contracting the samples and examining the resulting slopes. Since this approach brings about a sample selection bias, I run the spline regressions for the whole sample by placing two dummies. The results are similar to those obtained from the unrestricted approach.

Finally, I look at the effect of inequality on innovation within the empirical methodology established in the innovation literature. Specifically, I take into consideration that patents are count data and their arrivals can be described by a Poisson process. This hazard model is estimated using both non-parametric and semi-parametric approaches. I compare these results to a non-linear fit of the base model. The results point again to an inverted U shape. By controlling for income and other institutional effects, I again find that in high inequality countries, decreasing inequality causes more resources to be diverted to innovation, but the same is not true for low inequality countries.

Using a cross-country data on patents has some drawbacks. For instance, countries differ greatly with respect to their policies toward innovation, and not all innovations are patented. Since institutions play a role in the amount of patents issued, it might be the case that R&D expenditures is a better choice. However, in a wide cross-country study, R&D

expenditure data may not be entirely reliable, as it requires trained statisticians to collect such data. The author believes that patents capture the innovation output also better than R&D expenditures, and it is overall a better proxy for incentives to innovate in an empirical Schumpeterian setup, where new technologies are embodied in new goods and inequality affects innovative activity through future profits vis-a-vis dynamic demand distributions.

This paper does not answer the more fundamental question if the inequality is really bad for growth. Overall estimations point to a negative relation between inequality and innovative activities and it suggests that reducing inequality is beneficial for innovative activities especially in countries where inequality is high.

To have a better understanding of the effect of income distribution on the innovation process a further study is needed where institutional features can be more accurately captured and controlled for in the empirical model. As the technological gap between the leaders and the followers declines it becomes less costly to imitate and patenting process becomes more prone to the underlying institutional characteristics such as enforceability of property rights. In fact, recent literature on innovations has shown some progress in this direction and to establish a well defined theoretical link between the institutional features of innovation and the effect of inequality on innovation might be useful.

5 References

Acemoglu, D., 1998. Why do technologies complement skills? Direct technical change and wage inequality. Quarterly Journal of Economics 113, 1055 - 1090.

Acemoglu, D., 2002. Technical change, inequality and the labor market. Journal of Economic Literature XL, 7-12.

Aghion,P., 2002. Schumpeterian growth theory and the dynamics of income inequality. Econometrica 70, 855-882.

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica 60, 323 351.

Aghion, P. , Howitt, P., Harris, C., Vickers, J., 2001. Competition, imitation and growth with step-by-step innovation. Review of Economic Studies 28, 467-92.

Aghion, P. , Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: an inverted U relationship. Quarterly Journal of Economics 120, 701-728.

Aghion, P., Caroli, E., Garcia-Penalosa (1999). Inequality and economic growth: the perspective of the new growth theories. Journal of Economic Literature 37, 1615-60.

Ai, C. ,Chen, X., 2003. Efficient estimation of models with conditional moment restrictions containing unknown functions. Econometrica 71, 1795-1843.

Allen, S., 2001. Technology and the wage structure. Journal of Labor Economics 19, 440-483.

Arellano, M., 2003. Panel Data Econometrics, Oxford University Press , New York.

Arellano, M. and Bover,O., 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68, 29-51.

Atkinson, A., 1997. Bringing income distribution from the cold. Economic Journal. 107, 297-321.

Atkinson, A., Brandolini, A., 2001. Promise and pitfalls in the use of secondary data-Sets: income inequality in OECD countries as a case study. Journal of Economic Literature 34, 771-799.

Banerjee A. V., Duflo E., 2003. Inequality and growth: what can the data say? Journal of Economic Growth 8, 267-299.

Barro, R.,2000. Inequality and growth in a panel of countries. Journal of Economic Growth 5, 87–120.

Barro, R. J., and Lee, J. W., 1996. International measures of schooling years and schooling quality. American Economic Review 86, 218-23.

Barro, R. J., Lee, J.W., 1997. Data Set for a Panel of 138 Countries. Data set available on disk from authors.

Baum, C. F., Schaffer, M.E., Stillman, S., 2003. Instrumental variables and GMM: estimation and testing. Stata Journal 3, 1-31.

Benhabib, J., 2003. The trade-off between inequality and growth. Annals of Economics and Finance 4, 329-345.

Blundell, R., Bond, S., 2000. GMM estimation with persistent panel data: an application to production functions. Econometric Reviews, 19, 321–40.

Blundell, R., Bond,S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87,115-43.

Blundell, R., Bond, S., Windmeijer, F., 2000. Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator, in: B. Baltagi (Ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15. JAI Press, Elsevier Science, Amsterdam, pp 53-91.

Brouwer, E., Kleinknecht, A.H., 1996. Determinants of Innovation: A Micro-Econometric Analysis of Three Alternative Innovation Output Indicators" in A.H. Kleinknecht (Ed.),Determinants of Innovation: The Message from New Indicators, Macmillan Press, London, pp 99-124.

Carol, G., Hanan, T.M., 2000. The Demography of Organizations and Industries, Princeton University Press, Princeton N.Y.

Chong, A., Zanforlin, L., 2002. Technology and epidemics. IMF Staff Papers 49, 426-455.

Doornik, J.A., Arellano, M. ,Bond, S., 2002. Panel data estimation using DPD for Ox. http://www.nuff.ox.ac.uk/Users/Doornik.

Easterly, W., Levine,R.,1999. It's not factor accumulation: stylized facts and growth models. World Bank Economic Review, 15, 177-219.

Faria, L.A., 2002. Mergers and the market for organization capital. Working Paper, University of Chicago.

Fellner, W. ,1951. The influence of market structure on technological progress. Quarterly Journal of Economics 65, 556-77.

Foellmi, R., Zweimüller, J., 2005. Income Distribution and Demand-Induced Innovations, CEPR Discussion Papers 4985.

Forbes, K. 2000. A reassessment of the relationship between inequality and growth. American Economic Review 90, 869-887.

Galbraith, J., Kum, H., 2003. Estimating the Inequality of Household Incomes: Filling Gaps and Fixing Problems in Deininger&Squire Data Set. UTIP Working Paper 22.

Geroski, P.A., 1990. Innovation, technological opportunity, and market structure. Oxford

Economic Papers 42, 586-602.

Geroski P.A., Walters, R., 1995. Innovative activity over business cycle. The Economic Journal 105, 916-28.

Hausman, J.A., Newey, W., 1995. Nonparametric estimation of exact consumers surplus and deadweight loss. Econometrica 63, 1445-1476.

Heston, A., Summers, R., Aten, B., 2002. Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania.

Jovanovic B., Rousseau, P. 2001. Mergers and technological change, 1885-1998. Vanderbilt University Department of Economics Working Paper 01-W16.

Keller, W., 2001. International technology diffusion. NBER Working Paper 8573.

Kuznets, S., 1955. Economic Growth and Income Inequality. American Economic Review 45, 1-28.

Levin, R., Reiss, P., 1984. Tests of a schumpeterian model of R&D and market structure in: Z. Griliches (Ed.), R&D, Patents and Productivity, University of Chicago Press, Chicago, pp 175-208

Murphy, K. M., Shleifer, A.,Vishny, R.1989. Income distribution, market size, and industrialization. Quarterly Journal of Economics 103, 537 - 564.

Panizza, U., 2002. Income inequality and economic growth: evidence from American data"Journal of Economic Growth 7, 25-41.

Piva ,M., Vivarelli, M. 2006. Is demand-pulled innovation equally important in different groups of firms? Iza Discussion Papers 1982.

Robinson, J., 1996. Root-n consistent semi-parametric regressions. Econometrica 56 , 931-954.

Statistiques De Proprietre De Industrielle, 2000. World Intellectual Property Organization, CD-ROM.

Sedgley, N. H., 2006. A Time Series Test of Innovation-Driven Endogenous Growth. Economic Inquiry 44, 318-332.

UTIP-UNIDO, University of Texas Inequality Project, http://utip.gov.utexas.edu.

Windmeijer, F.,2000. A finite sample correction for the variance of linear two-step GMM estimators. Institute for Fiscal Studies,Working Paper 00/19.

Weinhold, D. and Nair-Reichert, U. 2006. Innovation, Inequality and Intellectual Property Rights. Working Paper.

Zweimuller, Y., 2000. Schumpeterian entrepreneurs meet Engel's law: the impact of inequality on innovation-driven growth. Journal of Economic Growth 50, 185 - 206.

Appendix A. Estimation procedures

A.1. Gmm estimation for dynamic panel data

This section closely follows Doornik, Arellano and Bond(2002) The dynamic panel model can be written as:

$$y_{it} = \sum_{k=1}^{p} \alpha_k y_{i(t-k)} + X_{it}\beta'(L) + \lambda_i + \eta_t + \varepsilon_{it},\ t = q+1, ..., T; i = 0, ..., N$$

where λ_i and η_t are respectively individual and time special effects and x_{it} is a k vector of explanatory variables. N is the number of cross-section observations. The idea here is if we can find variables which are not correlated with ε_{it} we can use it as instruments for equation in levels regardless of $X's$ being correlated with error term. $\Delta X_{i,t}$ and $\Delta Y_{i,t}$ are candidates for such instruments. There are many such instruments at hand including different lags, different combinations of lags, deviation from the means etc. which substantially increase the information set that can be utilized thus increasing the consistency of estimates. Then $(T_i - q)$ equations for individual i can be written in the form

$$y_i = W_i\delta + \kappa_i\lambda_i + \varepsilon_i$$

where δ is a parameter vector including the α_k's and β's and the time effects,η_t, and W_i is a data matrix containing the series of the lagged dependent variables, the $x's$ and time dummies. κ_i is a $(T_i - q) \times 1$ vector of ones.

$$\overset{\wedge}{\delta} = \left[\left(\sum_i W_i^{*\prime} Z_i\right) A_N \left(\sum_i Z_i W_i^{*\prime}\right)\right]^{-1} \left(\sum_i W_i^{*\prime} Z_i\right) A_N \left(\sum_i Z_i y_i^{*}\right)$$

where

$$A_N = \left(\frac{1}{N}\sum_i Z_i H_i Z_i\right)^{-1}$$

and $W_i^{*\prime}$ and y_i^{*} denote selected transformations of W_i and y_i (e.g. levels, first differences, orthogonal deviations, combinations of first differences (or orthogonal deviations) and levels, deviations from the means.

A.2. Interval Location

The procedure of interval location starts by estimating the following equation

$$y_{i,t} = y_{i,t-a} + \beta h_{i,t-a} + \gamma_1 D_1(h_{1i,t-a} - \rho_1) + \gamma_2 D_2(h_{2i,t-a} - \rho_2) + \delta X_{i,t-a} + \alpha_i + \eta_i + \varepsilon_{t,i}$$

where $h_{1i,t-a} > \rho_1$ and $h_{2i,t-a} > \rho_2$ and $\rho_2 > \rho_1$ by keeping ρ_1 fixed and increasing ρ_2 and vice versa.

A.3. Non-parametric Estimation

The idea here is to find an estimator for $g(h_{it}, x_{it}) = f(h_{it}) + x'_{it}\beta$. First an estimator for parameters β can be found by estimating

$f(h_{it}) = E[\ln y_{it} \mid h_{it}] - E[x_{it} \mid h_{it}]\beta$. The estimator for β. is then given by

$$\hat{\beta}.= [\sum_{i=1}^{n} \sum_{t=1}^{T} \tau_{it}(h_{it} - E(x_i \mid h_{it})(h_{it} - E(x_i \mid h_{it})']^{-1}$$

$$\times \sum_{i=1}^{n} \sum_{t=1}^{T} \tau_{it} \left[(h_{it} - E(x_i \mid h_{it})(\ln y_{it} - E(\ln y_{it} \mid h_{it})\right]$$

where $\tau_{it} = \tau(h_{it})$ is a trimming function which leaves a .95 quantile in the sample by leaving out the symmetric 5% of the outliers. If we want to estimate $g(h_{it}, x_{it})$ non-parametrically we use $\hat{\beta}$ and a Kernel K of bandwidth ρ such that $\hat{g}(h_{it}, x_{it}) = x'_{it}\hat{\beta} + \sum_{j \neq i}[\ln y_{it} - x'_{it}\hat{\beta}]K_\rho(h_i - h_j) / \sum_{j \neq i} K_\rho(h_i - h_j)$. where $K_\rho = \rho^{-k-1}\Re(v/\rho)$ and $\Re(v)$ is a Kernel function with the property $\int \Re(v)dv = 1$. Interested readers can consult to Hausman and Newey (1995) for an application to demand estimation.

Appendix B. Benchmark Coefficients

Other coefficients and their statistics in the benchmark specification are given below.

	Fixed Effects		Random Effects	
	Coefficient	t-statistic	Coefficient	t-statistic
PPINV	-0.078	-0.98	-0.029	-1.03
FCE	-0.007	-1.76	-0.011	-2.01
FDI	0.004	2.87	0.001	1.70
CPW	0.001	0.88	0.002	1.78
GDP	0.074	2.03	0.024	3.36

Appendix C. Countries

Albania
Algeria
Argentina
Australia
Australia
Austria
Bangladesh
Belgium
Bolivia
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Cuba
Czech Rep.
Czechoslovakia
Denmark
Dominican Rep.
Ecuador
Egypt
Ecuador
Fiji
Finland
France
Gabon
Gambia
Germany
Korea
Kuwait
Latvia
Libya
Macau
Madagascar
Malaysia
Malta
Iraq
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kenya
Germany, E
Germany, W
Ghana
Greece
Guatemala
Haiti
Honduras
Hong Kong
Hungary
India
Indonesia
Iran
Maritius
Mexico
Moldova
Mongolia
Morocco
Nepal
Netherlands
New Zealand
Nicaragua
Luxembourg
Nigeria
Norway
Pakistan
Panama
Peru
Philippines
Poland
Portugal
Puerto Ric
Qatar
Singapore
Slovakia
Slovenia
South Afric
Spain
Sri Lanka
Sweden
Syria
Taiwan
Tanzania
Thailand
Tunisia
Turkey
United King
United Stat
Uruguay
Uruguay
USSR
Venezuela
Yemen S.
Yugoslavia
Yugoslavia
Zambia
Zimbabwe

Table 1. Selected Descriptive Statistics

Mean and Standart Deviation		Income[1]			Inequality[2]			Growth[3]		
		Low	*Middle*	*High*	*Low*	*Medium*	*High*	*Low*	*Medium*	*High*
Real GDP Per Capita										
	1965	1098.9	3269.5	9740.0	6178.0	2528.7	2087.2	2984.0	3511.5	3456.1
		(459.4)	(1124.91)	(8365.03)	(2902.18)	(2451.39)	(1791.55)	(1394.6)	(2770.8)	(2944.4)
	1999	1118.5	3359.1	13584.0	13588.6	8641.6	4373.4	5823.2	8244.4	6240.4
		(493.97)	(1086.48)	(4735)	(4524.07)	(6113.45)	(4555.16)	(5576.4)	(6522.8)	(7048.4)
Real GDP Per capita growth										
	1965	.021	.045	.028	.049	.037	.049	-.025	.030	.083
		(.070)	(.056)	(.039)	(.028)	(.072)	(.064)	(0.04)	(.004)	(.044)
	1999	.026	.021	.028	.027	.034	.015	-.005	0.03	.055
		(.033)	(.035)	(.031)	(.027)	(.020)	(.044)	(.028)	(.005)	(.016)
Patent										
	1965[4]	427	936	13018	6096	8040	636	6638	7513	10659
		(1530)	(9343)	(17401)	(7513)	(16448)	(576)	(10788)	(12138)	(19719.4)
	1999	271	622	12281	13627	12611	5661	12456	9464	10458
		(563)	(1021)	(30283)	(15912)	(25380)	(22125)	(33435)	(11515)	(36943)
HCIN										
	1965	43.4	41.0	33.1	31.8	41.3	46.8	42.6	38.3	39.4
		(3.95)	(5.24)	(6.31)	(3.88)	(3.63)	(1.67)	(5.62)	(7.29)	(5.77)
	1999	48.1	43.2	38.8	33.7	39.1	46.3	41.5	41.7	40.8
		(3.02)	(4.08)	(5.87)	(2.28)	(3.38)	(4.08)	(5.67)	(5.83)	(5.65)
Theil										
	1965	.048	.046	.023	.073	.0342	.011	.053	.036	.034
		(.024)	(.0273)	(.029)	(.019)	(.009)	(.0045)	(.035)	(.023)	(.022)
	1999	0.104	0.071	0.043	0.079	0.034	0.016	0.061	0.04	0.06
		(.054)	(.031)	(.026)	(.012)	(.009)	(.0018)	(.027)	(.027)	(.028)
Number of Countries										
	1965	51	86	35	16	26	13	7	25	39
	2000	64	95	33	8	22	21	31	25	26

[1] Economies are divided according to 1999 GNI per capita, calculated using the World Bank Atlas method. The groups are: low income; less than$2000 , middle income; between $2,000 and $6,000, and high income; more than $6,000

[2] Inequality is taken to be low, medium or high when it is respectively less than 0.02, between 0.02 and 0.06, and higher than 0.06

[3] Growth is taken to be low, medium or high when it is respectively less than 0.02 , between 0.02 and 0.04, and higher than 0.04

[4] Japan excluded

Table 2. Overall Relationship Between Income Inequality and Patents Cited

Dependent Variable : Patents | **Coefficient of Inequality**

	Model[2]	*Pooled OLS*	*Fixed Effects*[3]	*Random Effects*	*GMM(Arellano-Bond)*[4]
Theil	1	-14.19(1.60)	-1.82(0.85)	-3.9(0.85)	-4.83(1.27)
	2	-14.75(1.53)	-2.81(0.87)	-3.02(0.86)	-4.96(1.28)
	3	-16.71(3.51)	-2.09(2.27)	-5.20(2.2)	-1.01(0.65)
	4	-11.60(5.09)	-3.37(1.98)	-3.99(1.81)	2.75(2.77)
	5	-12.08(5.17)	-1.35(2.22)	-3.98(1.35)	.47(2.66)
	6	-15.82(3.55)	-1.91(2.22)	-6.25(2.16)	.45(2.68)
HCIN	1	-0.080(0.010)	-0.024(0.008)	-0.036(0.008)	-0.041(0.014)
	2	-0.100(0.010)	-0.016(0.008)	-0.031(0.008)	-0.042(0.014)
	3	-0.111(0.023)	-0.023(0.099)	-0.056(0.018)	-0.02(0.01)
	4	-0.075(0.024)	-0.054(0.037)	-0.059(0.018)	-0.03(0.02)
	5	-0.106(0.036)	-0.028(0.019)	-0.055(-0.028)	-0.03(0.02)
	6	-0.078(0.024)	-0.021(0.019)	-0.065(0.02)	-0.045(0.02)
Number of Observations:		1285	1285	1285	1047[5]

[1] standard errors in parenthesis

[2] See Appendix for the different specifications used in this estimation

[3] Both country and time dummies are included

[4] Both the patent and gdp and their lags are used as instruments.

[5] For 3-6 the available number of observations for this type of estimation ranges from 95 to 340

Table 3. Interval Location

Threshhold	0.022	0.024	0.025	0.026	0.028	0.03	0.0325	0.035	0.0375	0.04
				Fixed Effects						
Dummy THEIL	0.575 (2.679)	0.534 (2.652)	0.496 (2.722)	0.447 (2.574)	0.340 (2.014)	0.198 (2.270)	0.058 (2.082)	-0.063 (-1.908)	-0.118 (-1.624)	-0.151 (-1.772)
Number of Observations	881(74 Countries)			*GMM (Arellano&Bover)* [1]						
Dummy THEIL	1.101 (2.676)	0.988 (2.629)	0.964 (2.637)	0.953 (2.646)	0.863 (2.603)	0.652 (2.462)	0.454 (2.326)	0.241 (2.175)	0.129 (2.103)	0.06 (2.044)
Sargan Test	0.019	0.022	0.010	0.004	0.008	0.025	0.009	0.002	0.005	0.008
Number of Observations	842(73 Countries)									
Threshhold	20	23	25	26	27	28	29	31	32	33
				Fixed Effects						
Dummy HCIN	0.0025 (3.411)	0.0541 (5.534)	0.0431 (4.569)	0.0129 (2.889)	0.0324 (1.838)	0.0161 (1.305)	0.0039 (0.496)	0.0016 (0.307)	0.0004 (0.085)	-0.0002 (-0.047)
Number of Observations	871(74 Countries)			*GMM (Arellano&Bover)* [2]						
Dummy HCIN	0.0104 (5.095)	0.4844 (6.395)	0.1154 (1.294)	0.0408 (0.877)	0.0178 (0.502)	0.0124 (0.481)	0.0078 (0.479)	-0.0054 (-0.726)	-0.0098 (-0.784)	-0.0092 (-0.795)
Sargan Test	0.013	0.004	0.012	0.011	0.005	0.020	0.032	0.004	0.008	0.040
Number of Observations	834 (73 Countries)									

1 Autocorrelation tests of order one range from -0.082 to -0.0012. Second order autocorrelation tests range from 0.021 to 0.342

2 Autocorrelation tests of order one range from -1.173 to -0.079. Second order autocorrelation tests range from 1.021 to 2.045

Table 3. Interval Location *(continued)*

Threshhold	0.0425	0.045	0.0475	0.05	0.055	0.06	0.065	0.07	0.075	0.08	0.085
				Fixed Effects							
Dummy THEIL	-0.181	-0.2076	-0.23815	-0.26527	-0.30531	-0.32078	-0.32399	-0.31878	-0.32482	-0.33115	-0.33712
	(-1.723)	(-1.477)	(-1.623)	(-1.573)	(-1.789)	(-1.969)	(-1.943)	(-1.989)	(-2.014)	(-2.431)	(-2.414)
Number of Observations	881(74 Countries)			*GMM (Arellano&Bover)*							
Dummy THEIL	0.017	0.005	-0.028	-0.084	-0.176	-0.218	-0.202	-0.172	-0.180	-0.193	-0.218
	(2.013)	(2.003)	(-1.978)	(-1.935)	(-1.865)	(-1.832)	(-1.843)	(-1.864)	(-1.858)	(-1.847)	(-1.826)
Sargan Test	0.007	0.039	0.010	0.008	0.012	0.004	0.003	0.006	0.010	0.005	0.006
Number of Observations	842(73 Countries)										

Threshhold	34	35	36	37	39	41	43	48	52	54	56
				Fixed Effects							
Dummy HCIN	-0.0062	-0.0089	-0.0088	-0.0096	-0.0126	-0.0144	-0.0163	-0.0171	-0.0176	-0.0177	-0.0179
	(-1.446)	(-1.616)	(-1.593)	(-1.972)	(-1.913)	(-2.343)	(-3.032)	(-6.502)	(-6.1554)	(-5.719)	(-6.841)
Number of Observations	871(74 Countries)			*GMM (Arellano&Bover)*							
Dummy HCIN	-0.0085	-0.0071	-0.0066	-0.0067	-0.0066	-0.0064	-0.0063	-0.0061	-0.0060	-0.0059	-0.0087
	(-1.801)	(-1.710)	(-1.708)	(-1.761)	(-1.791)	(-1.811)	(-1.834)	(-1.850)	(-1.869)	(-1.886)	(-1.896)
Sargan Test	0.010	0.006	0.006	0.007	0.016	0.006	0.027	0.003	0.001	0.014	0.005
Number of Observations	834 (73 Countries)										

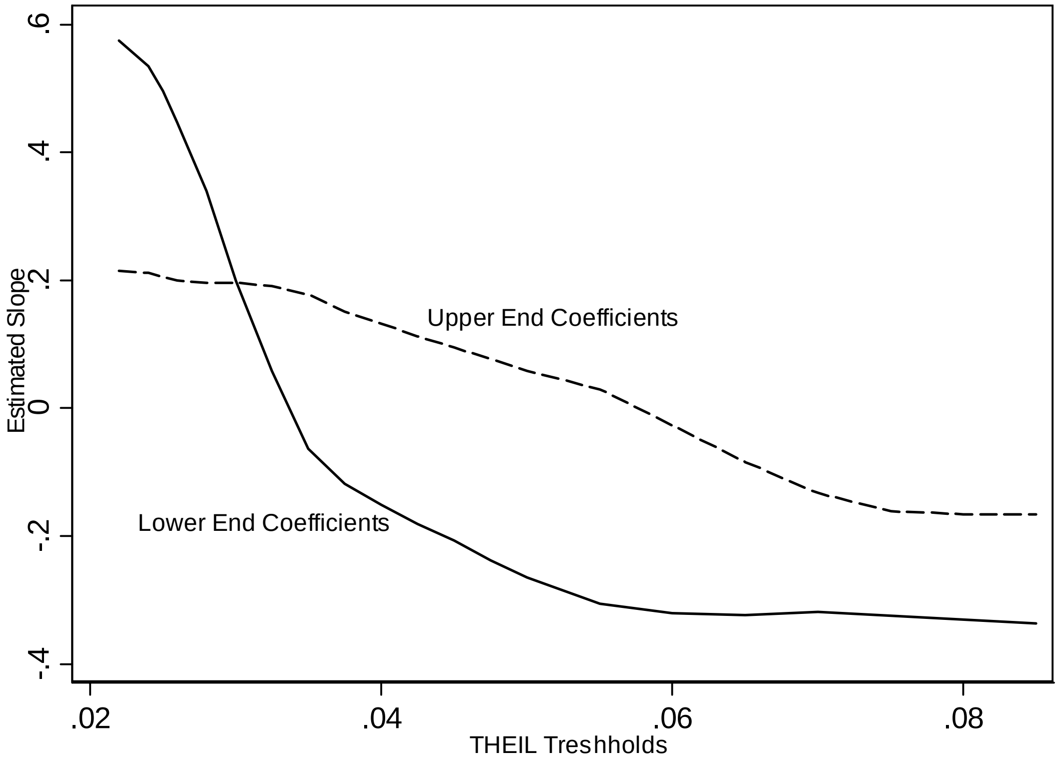

Figure 1. Interval Location: Fixed Effects.

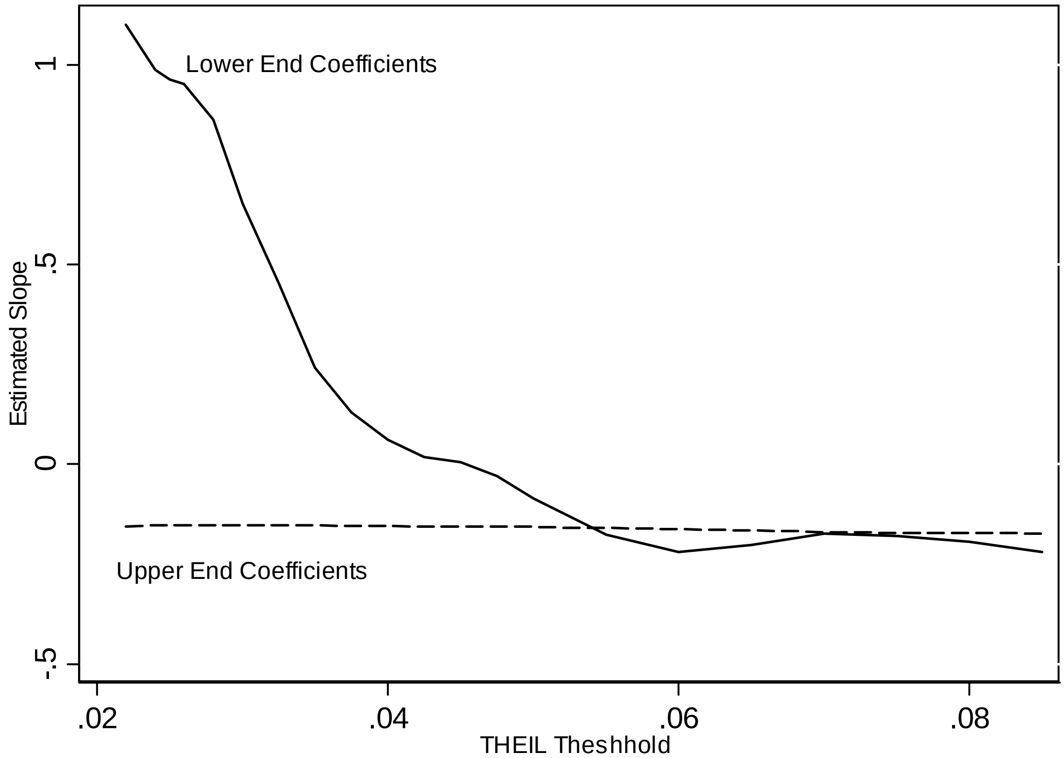

Figure 2. Interval Location: GMM.

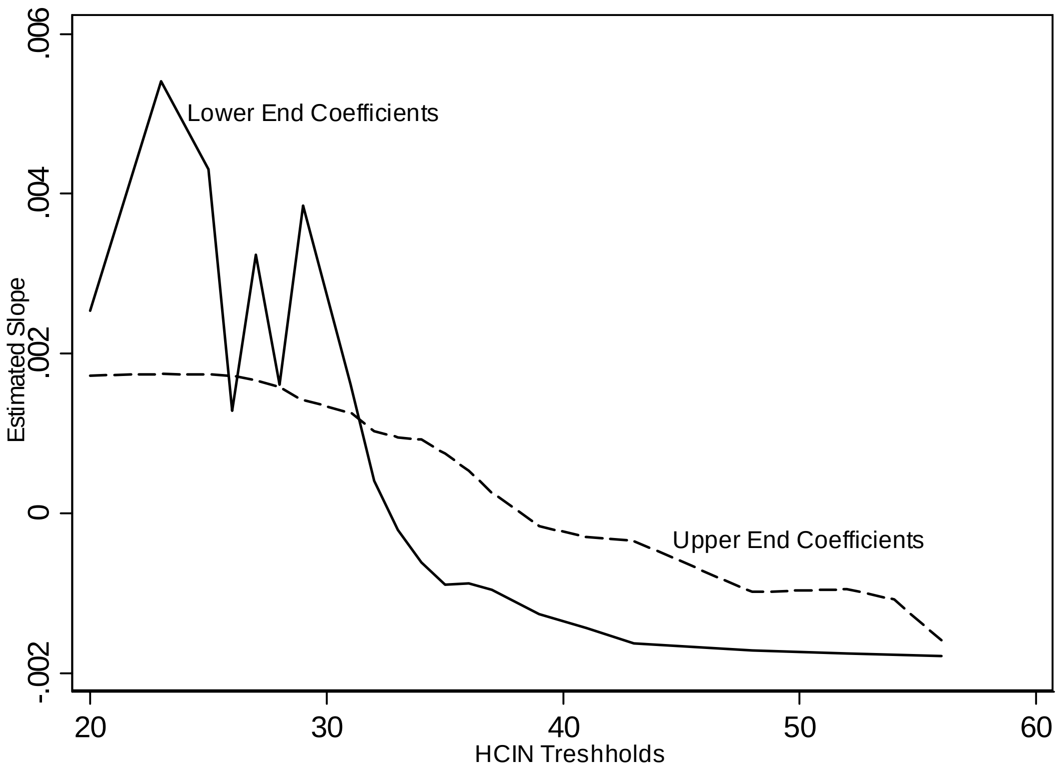

Figure 3. Interval Location: Fixed Effects.

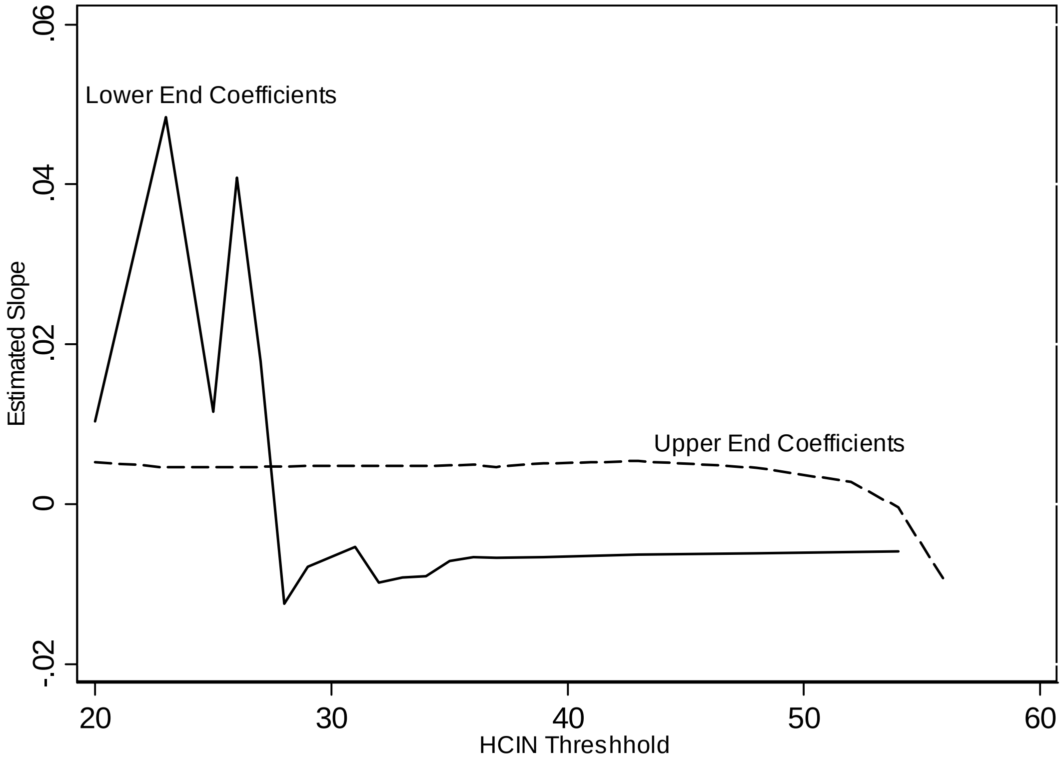

Figure 4. Interval Location: GMM.

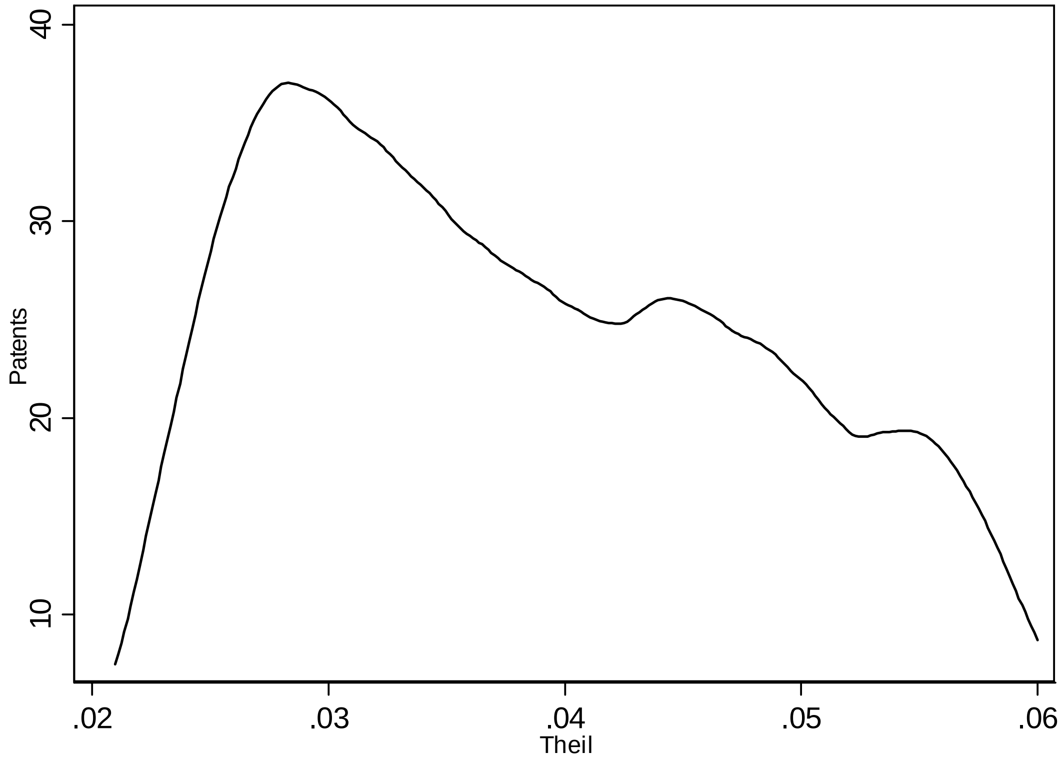

Figure 5. Kernel Regression(Gaussian Weights).

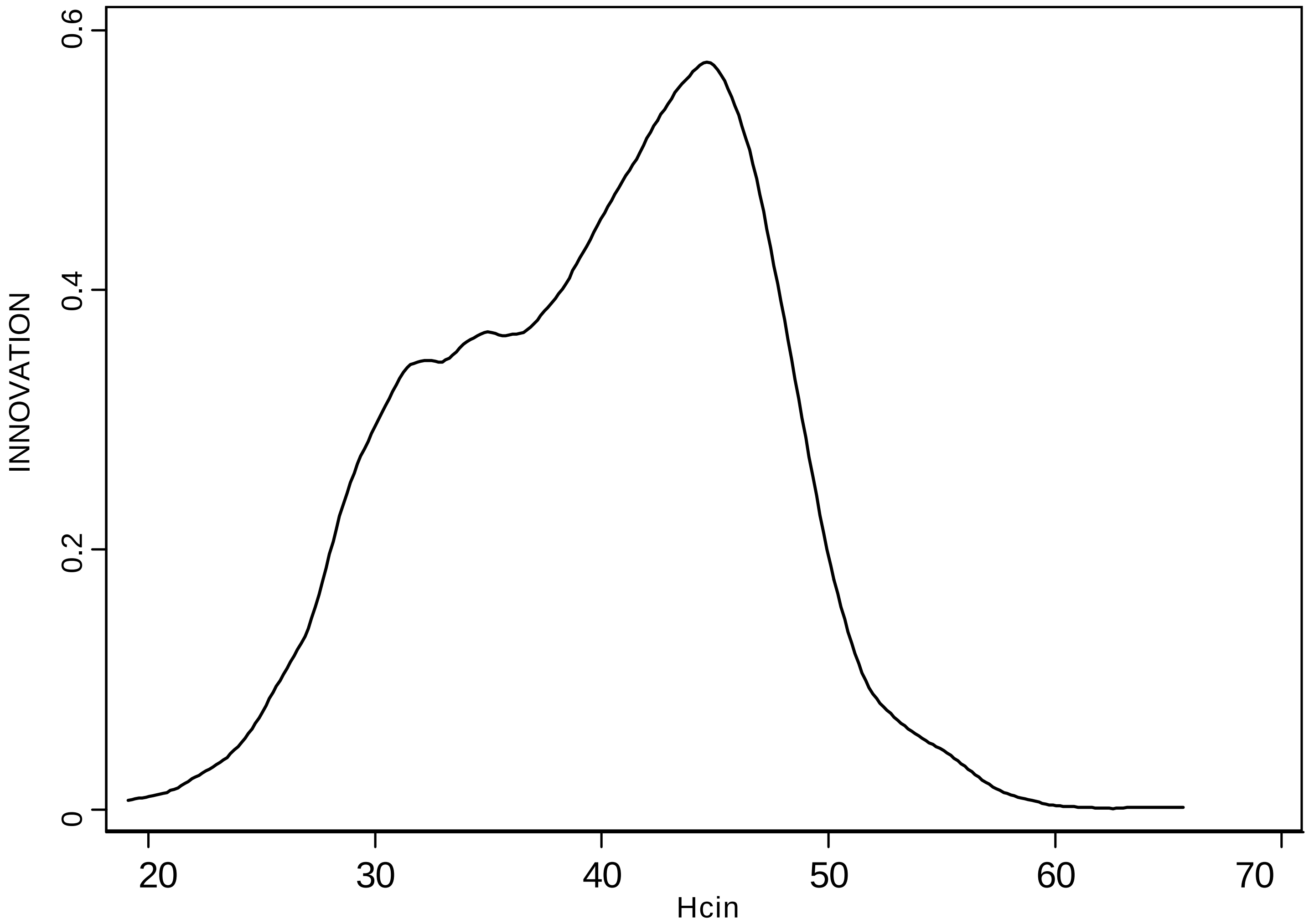

Figure 6. Kernel Regression(Gaussian Weights)

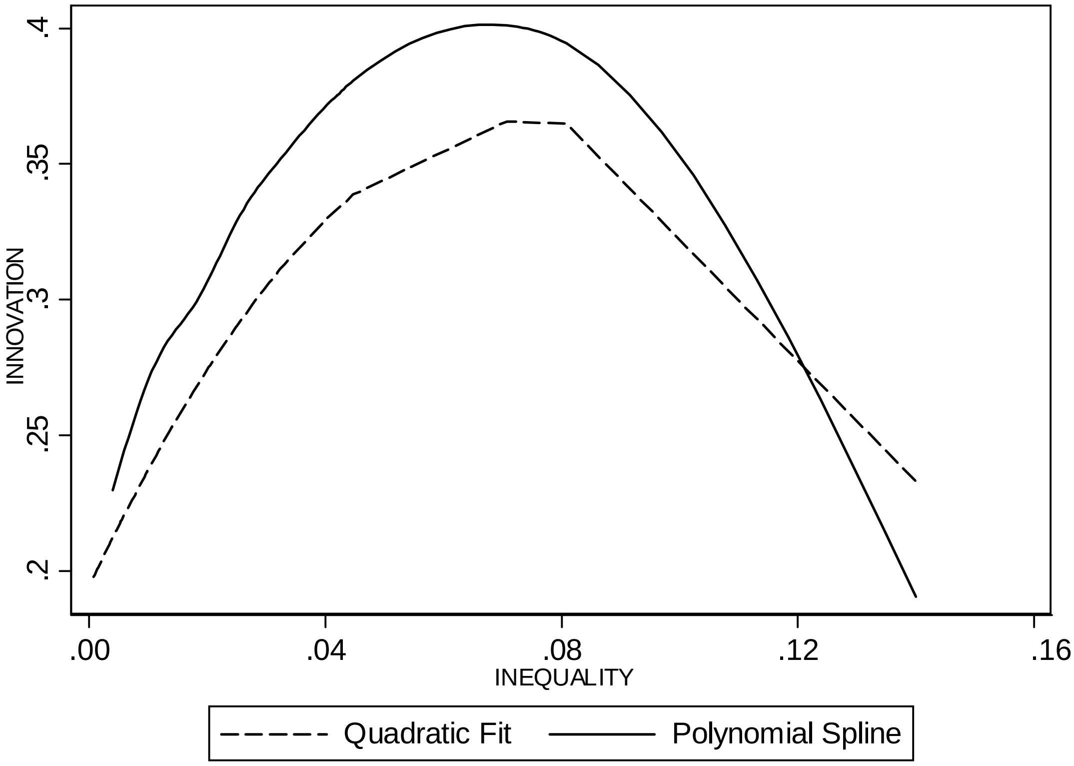

Figure 7. Nonlinear Fit and Semiparametric Estimation with Country and Time Effects.

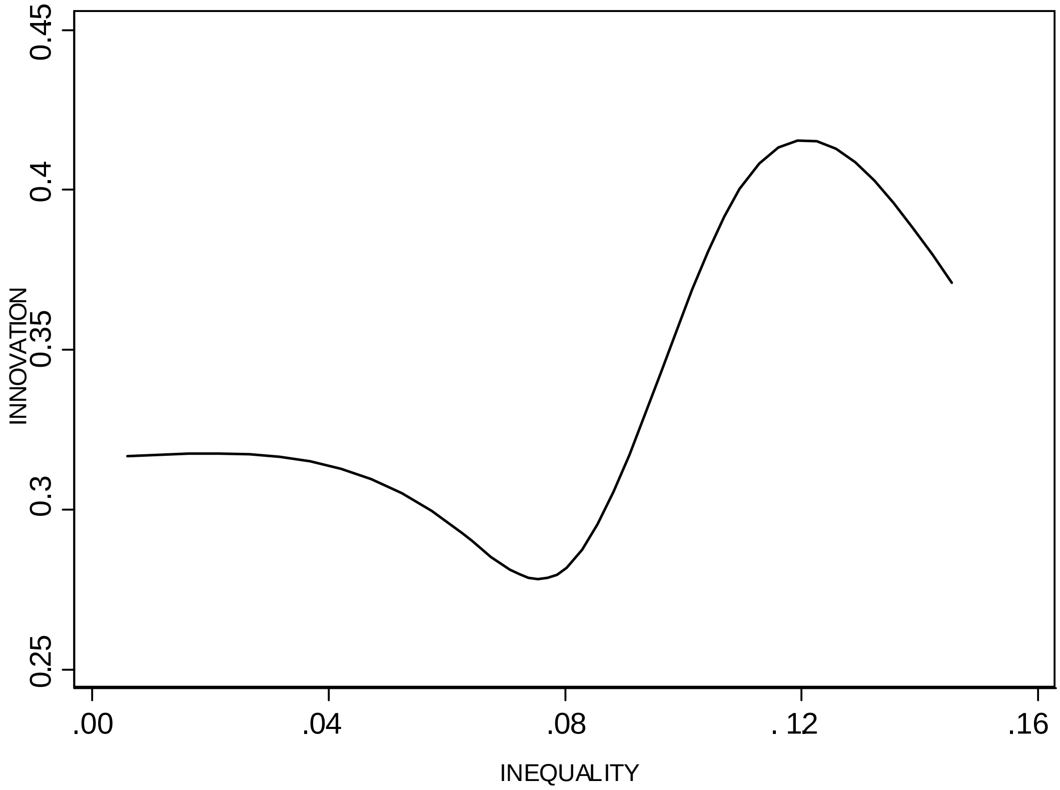

Figure 8. Exponential Quadratic Fit.