An International Comparison of Productivity Change in Agriculture and the Economy as a Whole Using Production Frontiers

A Dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Economics by

Nuno Alexandre Meneses Bastos Moutinho

Supervisors:

Elvira Silva

Faculdade de Economia do Porto, Universidade do Porto

Fernando S. Machado

Faculdade de Ciências Económicas e Empresariais, Universidade Católica Portuguesa

U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

2007

The research program presented in this dissertation was supported by the PRODEP program 5.2.

Biographical Note

Nuno Alexandre Meneses Bastos Moutinho was born on the 13rd of April of 1974, in Porto, Portugal.

In 1996, he graduated in Economics from the *Faculdade de Economia* da *Universidade do Porto* (Portugal), and was awarded a Master's degree in Economics in January 1999 by the same institution. He has been attending the PhD Programme at the *Faculdade de Economia* da *Universidade do Porto* since October 2000.

He is a lecturer at the *Faculdade de Economia* da *Universidade do Porto* since 1996, having taught, within the economics and management undergraduate programmes, Microeconomics (1996-2004) and Microeconomics II (since 2004). He also taught Microeconomics II to the undergraduate students of the management and industrial engineering programme at the *Faculdade de Engenharia* da *Universidade do Porto* (1999/2000) and Principles of Economics to the journalism and communication sciences undergraduate programme at the *Faculdade de Letras* da *Universidade do Porto* (since 2006).

His main scientific areas of research are Agricultural Economics and Efficiency and Productivity Analysis.

Acknowledgments

First, I would like to express my deep respect to my supervisors.

I would also like to thank Devashish Mitra and Donald Larson for granting access to their databases; Mike Tsionas and Subal Kumbhakar for helpful comments on the parametric and semiparametric models; Daniel Henderson for valuable discussions and for providing his software codes; William Greene for the help given in the estimation procedure of the finite mixture model; Maria Durbán and Ciprian Crainiceanu for guidance in the estimation of penalized splines with S-plus and Winbugs, respectively; participants of the 91st European Association of Agricultural Economists Seminar, at the University of Crete and of the VIII European Workshop on Efficiency and Productivity Analysis, at the University of Oviedo, for important comments on earlier drafts of the models.

Finally, special recognition goes out to my family, for their support and patience. This dissertation is dedicated to my lovely wife Raquel, who inspired me and provided constant encouragement during the entire process.

Abstract

Recent evidence on agricultural productivity growth contradicts common and longstanding beliefs that productivity growth in the agriculture sector is lower than in the overall economy and faster in developed than in developing countries. Such contradiction brings new interest to the issue of measuring agricultural productivity growth and of investigating the extent to which agriculture is responsible for the rising cross-country disparities of income and productivity.

First, we evaluate relevant empirical studies identifying the sources of labor productivity growth and the available methods for studying efficiency and productivity. Then, we investigate the role of agriculture in economic growth by using a panel data set for 45 countries and 26 years, for agriculture and for the overall economy. We estimate parametric and semiparametric production frontier models that incorporate heterogeneity across countries. Within each framework, we first determine cross-country distributions of labor productivity, both for the overall economy and for agriculture, and we look at how those distributions have changed over time; second, we try to shed some light on the causes of those changes by investigating the extent to which they are due to catch-up, technical change and factor accumulation.

We find that TFP growth was stronger for agriculture than for the economy as a whole, both for developed and developing countries. Changes in the distribution of labor productivity were mainly caused by capital deepening in the overall economy and by TFP change in agriculture. In this sector, factor accumulation was negative in the developing countries and positive in the developed countries, but the TFP growth rates were higher in the former group of countries than in the latter. Therefore, our results suggest that if disinvestment in this sector had not occurred, agriculture could have been an important engine of per capita income growth for the developing countries.

Table of Contents

List of Tables

List of Figures

1 - Introduction

There are three concepts of world (or inter-national) inequality dealt with in the literature, although they are often confounded. The concept of unweighted inter-national inequality takes each country as the unit of observation: each nation, small or large, counts the same. The weighted inter-national inequality uses the population size of each country as weight. Both approaches assume that within country, distribution is perfectly equal: all residents have the same income. The concept of true world inequality presented by Milanovic (2005) defines each individual as the unit of observation, ignoring country boundaries.

Considering the unweighted inter-national inequality, measured by the gross domestic product (GDP) or gross national product per capita, the literature (e.g., Quah, 1996a, 1997; Kumar and Russell, 2002; Henderson and Russell, 2005) is consensual about the conclusion that it has been rising during the last 30 years, through a phenomenon of bipolar international divergence of labor productivity. With labor productivity growing at rather modest or even negative rates over the last few decades in many developing nations, its distribution across countries evolved from a conventional unimodal shape in the early sixties to a bimodal shape at the end of the last century. Feyrer (2003) and Johnson (2005) have also found evidence of bimodality in the long-run distribution of output per worker. Having in mind that labor productivity is a rough indicator of nations' welfare, this evidence suggests that the world has become bipolarized into the rich and the poor, with the middle-income group of countries nearly disappearing.

A consensus is also obtained but in the opposite direction, when the variable studied is the weighted inter-national inequality, measured by national accounts: it has declined in the same period.

As pointed out by Deaton (2003), the controversy arises when the observation unit is the individual rather than the country, and the welfare is measured by the mean of per capita disposable income (or expenditures) rather than the mean of gross national product per capita. This explains why the contribution of growth to reduce global inequality among countries remains a controversial issue since the early seventies (e.g., Fishlow, 1972; Bardhan, 1973; Chenery *et al.* 1974). According to Milanovic (2005), this acrimonious debate turns out to revolve around the choice of data to measure consumption and hence living standards. In fact, consumption measured from household surveys grows less rapidly than consumption measured in national accounts. Results of studies drawing mainly on national accounts data indicate a reduction in poverty during the eighties and the nineties (e.g., Bhalla, 2002; Dollar and Kraay, 2002; Sala-i-Martin; 2002), called a golden age of capitalism and globalization. Studies relying on household surveys (e.g., Milanovic, 2002; Dikhanov and Ward, 2002; Dowrick and Akmal, 2001) conclude the opposite.

This dissertation does not address this controversial issue. We only deal with the unweighted inter-national concept of inequality and, therefore, assume that it has increased since the seventies. Our purpose is to make a contribution for the explanation of this fact, by investigating the role of agriculture in this process and comparing the determinants of labor productivity growth - factor accumulation or total factor productivity (TFP) growth - in this sector with those observed in the overall economy. Therefore, we intend to use a development accounting technique to determine which factor plays the major role in explaining the differences of labor productivity growth in agriculture and in the overall economy.

Over the last few years, various studies (e.g., Mankiw *et al.,* 1992; Mankiw, 1995; Young, 1995; Klenow and Rodriguez-Clare, 1997; Prescott, 1998; Hall and Jones, 1999; and Kumar and Russell, 2002) that will be analyzed in the next chapter have used a relatively wide range of methodological approaches to pursue this line of research. However, the resulting evidence is mixed and somehow contradictory: while some studies conclude that capital accumulation accounts for most of the increase in output per worker and for the bulk in cross-country growth differences over the last decades, others argue that such differences are mainly due to disparities in TFP growth.

The sectoral decomposition of total output may contribute to explain the increasing gap between the rich and poor countries that occurred in the last decades. Caselli (2005, p. 42) alerts that international productivity differences "could also be the result of variation in the weights in GDP of sectors with different sectoral-level productivity".

Regarding the case of agriculture, most of the empirical evidence appears to suggest that this sector has contributed strongly to the disparities of productivity and income between developed and developing nations. This is so for a number of reasons. Firstly, since the days of Adam Smith and David Ricardo, agriculture has often been regarded as a sector of low productivity growth relatively to the overall economy, due to a more limited scope for the division of labor and also to diminishing returns to land.[1] The confirmation of this argument has major implications for the international distribution of income, given the fact that agriculture still accounts for a significant share of the overall economy in developing countries.

Secondly, some empirical research suggests that TFP growth in agriculture has been higher in developed than in developing countries. In particular, various studies conclude that agricultural TFP has been increasing in rich nations and declining in poor nations (e.g., Fulginiti and Perrin, 1993, 1997, 1998, 1999; Arnade, 1998; Lau and Yotopoulos, 1989; Kawagoe and Hayami, 1985; Kawagoe *et al.* 1985; Trueblood, 1996). This evidence is found even for the "green revolution" Asian countries and for agricultural exporter nations of South America. Therefore, the low levels of productivity growth that arguably characterize the agricultural sector may prove to be even lower (or negative) in the developing world.

Thirdly, various studies conclude that the labor productivity gap between rich and poor nations is much higher in agriculture than in the overall economy. For example, Restuccia *et al.* (2004) find that in 1985, GDP per worker in agriculture for the richest 5 countries in the world is 71 times that of the poorest 5 countries, more than twice than the overall economy. As pointed out by Caselli (2005, p. 49), "if poor countries achieved the same level of agricultural labor productivity as the United States, world income inequality would virtually disappear!"

Finally, some additional factors appear to limit further the potential contribution of agriculture to the process of economic growth and development in developing countries. Given that the demand for agricultural products is rather inelastic with respect to income, there is very limited potential for a developing country to base a process of fast

[1] See Stern (1996) for an extensive theoretical analysis of the role of agriculture in economic growth.

growth on the agricultural sector, unless it manages to exploit its comparative advantages by increasingly supplying foreign markets. However, the highly protective agricultural policies of rich countries and some domestic market distortions strongly constrain this possibility. Indeed, both the low income elasticity of demand for agricultural products and the highly protective agricultural policies of the rich nations of the world appear to constrain severely the possibility of developing countries to base their economic growth process on the agricultural sector.

The dominant role the agriculture sector has in most developing countries appears therefore to be a major source of disadvantage for them. In accordance with this reasoning, some authors argue that the most important role of agriculture in the process of economic development lies on releasing productive resources for the other sectors and reducing its own weight in the economy.

Timmer (1988) presents two stylized facts that are consistent with this view: (i) the share of agriculture in both labor force and total output declines as income per capita increases and (ii) rapid agricultural growth accompanies or precedes general economic growth.

Some further evidence is reported in Gollin *et al.* (2002): a negative relationship between GDP per capita and the share of employment in agriculture and a positive relationship between growth in a country's agricultural productivity and the movement of labor out of agriculture. The latter relationship highlights the important role of agriculture in economic growth.

The arguments and findings mentioned above may have contributed to the strong government policy bias against agriculture and towards manufacturing presented in many developing countries. Such bias is found in Krueger *et al.* (1992), after measuring the income transfers that were induced by price interventions in 18 developing countries over the time period of 1960-83. Krueger *et al.* (1992, p. 1) conclude that “agriculture was clearly the loser, while the big winners were government (net revenue gain), urban consumers (lower food prices), and industry (cheap raw materials and other inputs)”. Restuccia *et al.* (2004) also find evidence of market distortions: estimates for the price of food and price of agricultural products at the farm gate are, respectively, higher and lower in rich than in poor countries.

Recently, some empirical evidence appears however to put into question most of the previous findings. Martin and Mitra (2001) estimate a production model for 50 countries in different development stages over the period 1967-1992. Contrarily to the conventional wisdom, the authors conclude that technical progress has been faster in agriculture than in manufacturing, for both developing and developed countries. In an earlier study, using parametric methods and covering 14 OECD countries over the period 1970-1987, Bernard and Jones (1996) come to the same conclusion, estimating annual TFP growth rates at 2.6 percent for agriculture and 1.2 percent for industry.

The idea of agricultural TFP regression in developing countries has also been recently put into question. Coelli and Rao (2003) examine the growth in agricultural productivity in 93 countries over the period 1980-2000 using data from FAOSTAT (2001). Results show that Asia is the major performer with an annual TFP growth of 2.9 percent and Africa seems to be the weakest performer with only 0.6 percent growth in TFP. Such figures clearly reject a phenomenon of negative productivity trends and technological regression. Nina *et al.* (2003) analyze agricultural productivity growth in developing countries over the period 1961-1994. Their results confirm that measured agricultural TFP in developing countries is generally increasing, with technical change being the main source for this growth. Similarly, the empirical results in Martin and Mitra (2001) indicate a strong positive annual average TFP growth between 1.76% and 2.62%, for developing countries, though lower than for developed countries. Bernard and Jones (1996) go one step further and find evidence of a relatively rapid convergence in agricultural TFP across countries. Coelli and Rao (2003) reach a similar conclusion with a larger and more representative database. In particular, they found that, on average, in the period 1980-2000, TFP has grown at 3.6 percent in countries that were technically inefficient in 1980, a much higher rate than the 1.2 percent found for the group of countries that were efficient in 1980. According to Coelli and Rao (2003, p. 14), "these results indicate a degree of catch-up in productivity levels between high-performing and low-performing countries". Restuccia *et al.* (2004) have also found that productivity differences needed to account for international disparities in economic growth are smaller once the role of agriculture is taken into account.

In short, the idea that agriculture is a sector of low productivity growth and also one with a strong productivity growth bias against the poor countries of the world has been challenged by recent empirical evidence.

Various authors have also pointed out the negative effects of development strategies that rely on a strong support of the manufacturing sector while penalizing agriculture. Krueger *et al.* (1992) find a negative relation between the rate of total taxation of agriculture and GDP growth: policies that depress agriculture's terms of trade below international levels are associated with slower economic growth. Therefore, Krueger *et al.* (1992) prescribe a recipe to governments of developing countries: do not tax agriculture relative to other sectors by protecting industry and maintaining overvalued exchange rates. Similarly, Restuccia *et al.* (2004, p. 20) conclude that market distortions and different types of barriers "reduce the incentives of farmers in poor countries to use modern inputs that are crucial for improving agricultural productivity".

These results suggest that the role of agriculture in the growth process of developing countries may well go beyond releasing labor and helping to create conditions for the development of non-agricultural sectors. Specifically, the empirical results indicate that agriculture can act as an engine of growth in significant parts of the developing world and contribute strongly to reversing the trend for global divergence that has been observed in the last few decades. Martin and Mitra (2001, p. 20) support this view by arguing that their results "suggest that a large agricultural sector need not be a disadvantage, and may be an advantage in terms of growth performance". Accordingly, Martin and Mitra (2001, p. 20) also challenge the validity of "the frequently-advocated policies of discrimination against agriculture on the grounds that it is a stagnant sector".

There seems therefore to be recent conflicting evidence on agricultural productivity growth both in relation to the other sectors of the economy as well as across countries. In our view, such evidence brings new interest to the issues of measuring agricultural productivity growth, investigating further the role of agriculture in economic development and also the extent to which this sector is responsible for the rising cross-country disparities of income and productivity. In particular, the research questions like the following remain largely unanswered:

- What are the trends and relationship between labor productivity growth in agriculture and in the overall economy?

- How has the international distribution of labor productivity evolved in agriculture?

- What is the contribution of the agricultural sector to productivity growth and to the changing distribution of GDP per capita?

- In developing countries, is a "shrinking agriculture" a pre-condition for economic growth and for fast labor productivity growth in agriculture itself?

- What are the main factors affecting the growth rates and changing international distributions of labor productivity growth in agriculture and in the overall economy? In particular, what are the roles of factor accumulation, technical change and catch-up in shaping those distributions?

In order to address those questions, we will try to link the problem of explaining international labor productivity differences across countries to the role played by agriculture in economic growth. More specifically, we will make a development accounting exercise for agriculture and the overall economy, identifying sources of growth in each case. Although estimates of productivity growth for the economy as a whole abound, there are surprisingly very few studies that provide comparisons between productivity in agriculture and the overall economy, particularly in developing countries. We try to make a contribution in filling that gap.

We start by estimating cross-country distributions of labor productivity, both for the economy as a whole and for the agricultural sector, and we look at how those distributions have changed over the period 1967-92. Secondly, we try to shed some light on the causes of those changes by investigating the extent to which they are due to catch-up, technical change and factor accumulation. In particular, we are interested in assessing whether those three factors have played different roles in shaping the distributions of labor productivity of the overall economy and of agriculture through the period under consideration.

Methodologically, we decide to use a frontier production approach, for three main reasons. First, this method allows the decomposition of labor productivity growth into three components (dividing TFP into catch-up and technical change) and none of them is determined residually. Secondly, recent advances in the frontier literature account for heterogeneity of individuals, a desirable characteristic when we are dealing with a sample with countries in different development stages. Last but not least, it is possible to introduce some flexibility in the definition of technology and stochastic noise, avoiding misspecifications problems.

The dissertation extends the literature on frontier-based development accounting models (e.g., Kumar and Russell, 2002; Henderson and Russell, 2005; Kumbhakar and Wang, 2005) in some important directions: it uses a panel data framework; the frontier is not assumed to be common to all countries; the specification of the frontier is flexible. Kumar and Russell (2002) and Henderson and Russell (2005) estimate a single production frontier to cross-country data using a deterministic method. Kumbhakar and Wang (2005) use a translog specification to determine production frontiers of a panel data sample of countries, taking heterogeneity into account. The main conclusion of this study was that ignoring heterogeneity tends to underestimate the catch-up rate and overestimate technical change effect. Therefore, we will try to create stochastic models which account simultaneously for heterogeneity across countries and for the flexibility of the frontier specification.

The rest of the dissertation is organized as follows. Chapter 2 briefly reviews the most relevant empirical studies identifying sources of labor productivity growth in the overall economy and/or in agriculture. After concluding that the results of those studies are largely sensitive to measurement methods they use, we present, in chapter 3 a survey of some empirical methods that have been used in the literature to study efficiency and productivity. The following chapters try to answer the research questions of the dissertation, using panel data of 45 countries and 26 years for agriculture and overall economy. Chapter 4 introduces some labor productivity growth indicators exhibited by the sample. Chapter 5 and 6 use two different methods to perform the decomposition of labor productivity growth, respectively a finite mixture model and a penalized spline approach both for Classical and Bayesian formulations. The last chapter summarizes the main results and conclusions of the dissertation.

2 - Sources of Labor Productivity Growth in the Agricultural Sector and in the Overall Economy: A Literature Review

2.1 - Introduction

Development accounting uses cross-country data on output and inputs to identify the most important source of differences in per-worker incomes: disparities in factor quantities or in the efficiency with which those factors are used, known as TFP. Conceptually, according to Caselli (2005), development accounting can be defined as the task of quantifying the relationship:

$$\textit{Per-capita income}=F(\textit{Factors},\textit{TFP}). \qquad (2.1)$$

In the traditional cross-country growth regression models, the key steps of this assignment are "(1) choosing a functional form for F, and (2) accurately measuring Income and Factors" (Caselli, 2005, p.1). The omission of TFP measurement is explained by the fact that within these models, it is determined as a residual, "a measure of our ignorance on the causes of poverty and under-development" (Caselli, 2005, p. 1).

The baseline equation for this kind of models is developed by Solow (1956), assuming a Cobb-Douglas production function with constant returns to scale, in which total GDP is a function of the economy's labor resources L, its capital stock K, and its TFP level A:

$$Y_t = (K_t)^{\alpha} (L_t)^{1-\alpha} A_t. \qquad (2.2)$$

Using lowercase letters to denote proportional rates of change, we can use equation (2.2) to decompose growth in output per worker into various components:

$$y = \alpha\left(s - \delta - n\right) + \tau, \qquad (2.3)$$

where:

y is the rate of growth of GDP per worker; s is the investment to GDP ratio; δ is the depreciation rate of physical capital stock; n the labor force growth rate; and τ is the TFP growth rate, assumed to be a constant and obtained as a residual, after accounting for factor inputs and for changes in the quality of labor inputs.

We will use the work of Mankiw *et al.* (1992) as an example of a traditional cross-country growth regression model. Mankiw *et al.* (1992) extend Solow's (1956) framework with an augmented production function that incorporates human capital (H):

$$Y_t = (K_t)^{\alpha} (H_t)^{\beta} (L_t)^{1-\alpha-\beta} A_t. \tag{2.4}$$

The decomposition of output-per-worker growth is now:

$$y = \alpha\left(s_k - \delta_k - n\right) + \beta\left(s_h - \delta_h - n\right) + \tau, \tag{2.5}$$

with additional terms arising because growth can be generated both by investments in physical capital and investments in human capital.

All of traditional cross-country growth regression models present two undesirable characteristics: (i) they obtain the effect of TFP growth residually; and (ii) they are heavily model-driven, depending on particular assumptions about the technology, market structure and technological change. Therefore, one potential promising research strategy is to measure directly the effect of TFP on growth and to improve on steps (1) and (2) indicated by Caselli (2005, p.1); i.e., by looking at alternative functional forms and by attempting a more sophisticated measurement tools to perform the development accounting exercise. The frontier production models presented by Kumar and Russell (2002) and Henderson and Russell (2005) represent an important step towards this direction. This method estimates a best practice frontier at each point in time with a deterministic and nonparametric approach known as Data Envelopment Analysis (DEA). It envelops the data in the smallest fitting convex cone, with upper boundary of this set representing the "best practice" production frontier. Then, it calculates measures of TFP growth by means of Malmquist productivity indices presented in the work of Grosskopf (1993) and Färe *et al.* (1994a, 1994b).

To illustrate this decomposition method, we use the subscripts b and c to stand for the base period and the current period, respectively. In period b, x_b units of input per worker are used to produce y_b units of output per worker. However, if the country is fully efficient, it could produce $\overline{y_b}(x_b)$. Therefore, efficiency in period b is measured as:

$$Eff_b = \frac{y_b}{\overline{y_b}(x_b)}. \tag{2.6}$$

Thus, labor productivity in period *b* can be expressed as:

$$y_b = Eff_b \cdot \overline{y_b}(x_b). \tag{2.7}$$

Consider the current period *c*, labor productivity growth is given by the equation:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\overline{y_c}(x_c)}{\overline{y_b}(x_b)}. \tag{2.8}$$

Multiplying the numerator and the denominator of (2.8) by $\overline{y_c}(x_b)$, labor productivity growth can be rewritten as:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\overline{y_c}(x_b)}{\overline{y_b}(x_b)} \cdot \frac{\overline{y_c}(x_c)}{\overline{y_c}(x_b)}. \tag{2.9}$$

The ratio Eff_c to Eff_b is the efficiency change or technological catch-up between the current period and the base period. The ratio of $\overline{y_c}(x_b)$ to $\overline{y_b}(x_b)$ captures the shift in the deterministic frontier caused by technological change, since input quantity per worker does not change. The last term on the right hand side captures the effect of factor accumulation, since it measures the output per worker change along the "deterministic" frontier in period *c*.

Alternatively, (2.8) could be multiplied and divided by $\overline{y_b}(x_c)$ and a different, but valid, decomposition would be obtained. This means that labor productivity growth decomposition is path dependent, forcing the use of geometric averages:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \left[\frac{\overline{y_c}(x_b)}{\overline{y_b}(x_b)} \cdot \frac{\overline{y_c}(x_c)}{\overline{y_b}(x_c)} \right]^{1/2} \cdot \left[\frac{\overline{y_c}(x_c)}{\overline{y_c}(x_b)} \cdot \frac{\overline{y_b}(x_c)}{\overline{y_b}(x_b)} \right]^{1/2}, \tag{2.10}$$

where:

$\left[\frac{\overline{y_c}(x_b)}{\overline{y_b}(x_b)} \cdot \frac{\overline{y_c}(x_c)}{\overline{y_b}(x_c)}\right]^{1/2}$ represents technological change and $\left[\frac{\overline{y_c}(x_c)}{\overline{y_c}(x_b)} \cdot \frac{\overline{y_b}(x_c)}{\overline{y_b}(x_b)}\right]^{1/2}$ indicates factor accumulation and all the other terms are defined as before.

Considering the combined effect of efficiency variation with technological change, we obtain the Malmquist TFP index (e.g., Grosskopf, 1993 and Färe *et al.*, 1994a, 1994b).

Therefore, the specification of a functional form for the technology and the assumptions on technological change and the market structure are not needed to determine the contribution of each component to labor productivity growth. TFP is not obtained residually and it can be decomposed into technological change (shifts in the world production frontier) and technological catch-up (movements towards or away from the frontier as countries adopt "best practice" technologies and reduce or exacerbate technical inefficiency). The other source of labor productivity growth, factor accumulation, is captured through the movements along the frontier.

In this chapter, we briefly review the most relevant empirical studies identifying sources of labor productivity growth in the overall economy and in agriculture with the purpose of answering two questions: are the results sensitive to the methods used? And to the sectors analyzed? To carry on this task, we will also analyze how the literature contextualizes the role of agriculture in the process of economic growth.

2.2 - Sources of Labor Productivity Growth in the Overall Economy

Over the last few years, various studies[2] have used a relatively wide range of methodological approaches to decompose labor productivity growth for the economy as a whole. However, the resulting evidence has been mixed and somehow contradictory: while some studies indicate that capital accumulation accounts for most of the increase in output per worker and for the bulk in cross-country growth differences over the last decades, others suggest that such differences are mainly due to disparities in TFP growth.

---

[2] e.g., Boskin and Lau (1990, 1991, 1992a, b), Kim and Lau (1992a, b, 1994, 1995), Mankiw *et al.* (1992), Mankiw (1995), Young (1995), Lau (1996), Klenow and Rodriguez-Clare (1997), Prescott (1998), Hall and Jones (1999) and Kumar and Russell (2002).

Mankiw *et al.* (1992) are among the first to perform a cross-country analysis of economic growth determinants. Using data for 98 countries and a three-factor Cobb-Douglas production function [equation (2.4)], Mankiw *et al.* (1992) conclude that factor accumulation accounts for 78% of the variation in output per worker between 1965 and 1985. Mankiw (1995, p. 301) argues that for "understanding international experience, the best assumption may be that all countries have access to the same pool of knowledge, but differ by the degree to which they take advantage of this knowledge by investing in physical and human capital."

In a series of papers, Boskin and Lau (1990, 1991, 1992a, 1992b), Kim and Lau (1992a, 1992b, 1994, 1995) and Lau (1996) investigate the sources of economic growth for developed and developing countries using the aggregate meta-production function framework. They find that technical progress for the developed countries and capital accumulation for developing countries are the most important sources of economic growth in the postwar period. When the analysis is extended to take into account the effects of embodied technical progress, capital accumulation is found to be the most important source of economic growth for the developed countries. Furthermore, the hypothesis of no technical progress and no catch-up cannot be rejected for developing countries. Therefore, the gap between TFP levels of developed and developing countries appears to be widening. Lau (1996) concludes that this finding should conduct developing countries to devote greater proportions of their resources to innovative activities, in order to attain a positive rate of growth in productive efficiency and to increase the contribution of technical progress to labor productivity growth.

Young (1995) applies the growth rates accounting approach to the growth miracles of the East Asia in the period 1965-1990 and concludes that the translog index of TFP growth lies between 0 and 2% for these countries, clearly less than previously found by growth accounting studies which attributed 1/3 of growth to TFP. Thus, both studies of Mankiw *et al.* (1992) and Young (1995) are consistent with the idea that factor accumulation is the crucial determinant of growth.

This view was initially questioned by the works of Klenow and Rodriguez-Clare (1997), Prescott (1998) and Hall and Jones (1999), suggesting that disparities in TFP are the main explanation for output per worker differences.

Klenow and Rodriguez-Clare (1997) use a similar method as in Mankiw *et al.* (1992), but question two assumptions made, namely the values chosen for capital shares and the human capital accumulation measure. These modifications change the results, allowing Klenow and Rodriguez-Clare (1997, p. 99) to "call for returning productivity differences to the center of theorizing about international differences in output per worker."

Prescott (1998) also concludes that differences in physical and intangible capital cannot account for international income differences. The most striking result found is that differences in the publicly available stock of technical knowledge do not explain TFP variations across countries. Therefore, Prescott (1998) concludes that a theory of TFP is needed.

Following Prescott (1998), Hall and Jones (1999) show that TFP differences are related to the type of policies and institutions adopted by the countries. Hall and Jones (1999) present a new technique of level accountings instead of growth rates accounting in the decomposition of output per worker into capital intensity, human capital and TFP. Assuming a small capital share coefficient in the Cobb-Douglas production functions, Hall and Jones (1999) conclude that most of the growth gap between any country and the United States of America is due to residual productivity differences, which are primarily related to differences in social infrastructures across countries such as government policies and institutions.

All of economic growth studies reported so far share one important characteristic: TFP is measured residually. According to Easterly and Levine (2001, p. 1), this means that "after accounting for physical and human capital accumulation, 'something else' accounts for the bulk of cross-country growth differences", with the term TFP used "to refer to the 'something else' (besides physical factor accumulation) that accounts for economic growth differences". When the residual is the main variable explaining some phenomenon, it is natural that some authors would try to look for other methods and check if the result holds.

Kumar and Russell (2002) and Henderson and Russell (2005), starting from the decomposition of the Malmquist index presented in the works of Grosskopf (1993) and

Färe *et al.* (1994a, 1994b), develop a method in which none of the decomposed effects is determined residually. Furthermore, catch-up is not measured relatively to a single country, but to a world productiob frontier. DEA is used to estimate a world production frontier for the years of 1965 and 1990 from a large sample of countries. As indicated by Durlauf *et al.* (2005), this is important since "any misspecification of the production function due to the Cobb-Douglas assumption in other studies will tend to increase the apparent variation in TFP relative to that found by Henderson and Russell". In fact, Duffy and Papageorgiou (2000) reject the Cobb-Douglas aggregate production function in favor of the more general constant elasticity of substitution specification for cross-country empirical studies.

Assuming constant returns to scale (CRS), Kumar and Russell (2002) decompose labor-productivity growth into components attributable to physical capital accumulation (movements along the frontier), technological change (shifts in the world production frontier) and technological catch-up (movements towards the frontier). The empirical results suggest that capital deepening, as opposed to technological change or technological catch-up, is the main explaining factor for the international divergence of economies. Furthermore, Kumar and Russell (2002) argue that wealthy countries have benefited more from technological progress than less developed countries and find striking examples of technological regress in low-income countries. Henderson and Russell (2005) extend the approach of Kumar and Russell (2002) by introducing human capital. Henderson and Russell (2005) find that, on average, about 90% of the increase in output per worker over the 1965-1990 period is attributable to the accumulation of human and physical capital.

2.3 - Agriculture and the Overall Economy

2.3.1 - The role of Agriculture in Economic Development

The sectoral decomposition of total output may contribute to explain the increasing gap between the rich and poor countries that occurred in the last decades. The fact that agriculture in developing countries still accounts for a significant share of the overall economy may be important to understand labor productivity growth in those countries. The role of agriculture in economic development and the extent to which the agricultural sector is responsible for the rising cross-country disparities of income and productivity are rather important issues that remain largely unclear.

Gollin *et al.* (2002) decompose growth in per worker GDP over the 1960-1990 period into three components: growth within agriculture, growth within non-agriculture, and growth due to sectoral shifts. On average, Gollin *et al.* (2002, p.164) find that the contribution of agricultural growth, non-agricultural growth, and sectoral shifts are 54 percent, 17 percent and 29 percent respectively, concluding that "agricultural productivity growth, along with the ensuing sectoral shifts in employment, is an important source of economic growth".

Another important finding supporting the crucial role of agriculture in development processes and described in recent studies (e.g., Restuccia *et al.* , 2004; Caselli, 2005) is that cross-country differences in labor productivity for the agricultural sector are much higher than differences in aggregate labor productivity. For instance, in 1985, GDP per worker in the richest 5 countries in the world is 32 times that of the poorest 5 countries and in agriculture, the productivity difference is even larger: GDP per worker in agriculture for the richest countries is 71 times that of the poorest countries.

This idea that growth in agricultural productivity is crucial to the development of poor countries is recurrent in the traditional development literature (e.g., Timmer, 1988) after the publication of Johnston and Mellor's (1961) and Schultz (1964) classical works. Two stylized facts are often mentioned in the literature: agriculture's share in GDP and labor force declines as economies grow (agriculture's direct contribution to economic growth diminishes); for most of poor countries, a rapid, substantial and continued agricultural growth is a necessary condition to induce general economic growth (agricultural indirect contribution must increase to allow the decline of its direct contribution). These facts are empirically confirmed by several studies (e.g., World Development Report, 1992; Gollin *et al.*, 2002) reporting the declining share of agriculture during the process of economic growth and a strong positive relationship between agricultural productivity and the movement of labor out of agriculture. This evidence is valid independently of the time period and countries considered.

Table 1 - Share of Agriculture in GDP

	1965	1990
Low income	41	31
Lower middle income	22	17
Upper middle income	16	9
High income	5	2,5

Source: World Development Report (1992)

Nevertheless, the dominant role of agriculture in developing countries seems to be disadvantageous for two reasons. First, the domestic demand for agricultural products is inelastic and the highly protective agricultural policies of rich countries and some domestic market distortions constrain the possibility of exporting. Second, several empirical studies suggest a decline in agricultural TFP for developing countries (e.g., Fulginiti and Perrin, 1993, 1997, 1998, 1999; Arnade, 1998; Lau and Yotopoulos, 1989; Kawagoe *et al.*, 1986). Each one of these aspects presented in the literature should be analyzed carefully.

Regarding internal and external market distortions, the idea that agriculture was impervious to price incentives and that industry was the engine of growth seems to contribute to strong government policy biases against agriculture and towards manufacturing in many developing countries. According to the World Bank (1992, p. 1), an analysis of these income transfers reveals that "agriculture was clearly the loser, while the big winners were government (net revenue gain), urban consumers (lower food prices), and industry (cheap raw materials and other inputs)".

The study of Krueger *et al.* (1992), focusing on 18 developing countries over the time period of 1960-83, measures the income transfers induced by price interventions among agriculture, government, and the overall economy. Countries with high taxation of agriculture present low rates of growth in the agricultural output (2.7 percent per year, on average) and in GDP (4.2 percent per year on average). Countries with lower taxation of agriculture exhibit higher rates of growth in agricultural output (5.2 percent per year) and in GDP (5.9 percent per year on average). Therefore, Krueger *et al.* (1992) prescribe a recipe to governments of developing countries: if heavy taxation ended, agricultural growth rates would nearly double, inducing a 40% raise in the aggregate growth rate of GDP.

Restuccia *et al.* (2004) also find evidence of market distortions by estimating that the price of food is higher and the price of agricultural products at the farm gate is lower in rich than in poor countries. Thus, Restuccia *et al.* (2004, p. 28) decide to examine direct and indirect barriers in the use of intermediate inputs in agriculture, concluding that "these barriers reduce the incentives of farmers in poor countries to use modern inputs that are crucial for improving agricultural productivity". Schultz (1964) identifies these problems thirty years before, but, as pointed out by Restuccia *et al.* (2004, p. 28), "this quantitative analysis shows that for many countries in the world, barriers to transforming traditional agriculture are still pervasive".

Using a two-sector general equilibrium model with subsistence food requirements, Restuccia *et al.* (2004) emphasize two interesting outcomes: agriculture accounts for 7/8 of the aggregate productivity differences between the rich and poor countries; low labor productivity in agriculture in poor countries can largely be accounted for low levels of economy-wide productivity and barriers to the use of modern intermediate inputs in agricultural production. Restuccia *et al.* (2004) report a vicious circle: a low economy-wide productivity has also an indirect effect on low productivity in agriculture, since the subsistence requirements of food and land act as a fixed factor in agriculture, implying that more labor has to be allocated to this sector in countries with low economy-wide productivity level, which results in a lower land to labor ratio and a further reduction in labor productivity in agriculture. Therefore, the removal of market barriers in agriculture seems crucial in the growth process of developing countries.

2.3.2 - Sources of Labor Productivity Growth in Agriculture

Regarding TFP in developing countries, several empirical studies indicate a prolonged and rapid decline in agricultural TFP: Kawagoe and Hayami (1985), Kawagoe *et al.* (1985), Lau and Yotopoulos (1989), Trueblood (1996), Arnade (1998) and Fulginiti and Perrin (1993, 1997, 1998, 1999). This evidence is found even for the "green revolution" Asian countries and for agricultural exporter nations of South America. If this evidence was true, then any agricultural-oriented development strategies, such as the one suggested by Adelman (1995), would not be followed by policy makers in developing countries and probably more market distortions in favor of non-agricultural sectors would be encouraged.

Kawagoe and Hayami (1985) and Kawagoe *et al.* (1985) analyze data from 22 developing countries and 21 developed countries for 1960, 1970, and 1980, using a Cobb-Douglas specification for the production function with five conventional inputs (land, labor, tractors, livestock, and fertilizer) and two educational variables to adjust for differences in labor quality. Lau and Yotopoulos (1988), using the same data and including first differences to account for fixed country-specific effects, show that results are sensitive to the functional form specification. Nevertheless, the results for TFP growth are similar among these studies: regression in the poor and progress in the rich countries. It is important to note that the studies of Kawagoe and Hayami (1985), Kawagoe *et al.* (1985) and Lau and Yotopoulos (1988) represent cross-section international comparisons.

Fulginiti and Perrin (1993, 1997, 1998, 1999) carry out a panel data study of 18 developing countries considering one output (aggregate agricultural output) and five inputs (land, labor, fertilizer, machinery and livestock) over the period 1961-1985. The majority of countries exhibited negative TFP growth, including nations of the “Green Revolution” (Korea and Philippines) and American agricultural exporters (Argentina, Brazil). Interestingly, this outcome does not change with the method used in the estimation (standard econometric methods with a Cobb-Douglas production function or DEA). Fulginiti and Perrin (1997, 1999), Trueblood (1996) and Arnade (1998) use an output-based Malmquist index to estimate agricultural TFP. Trueblood (1996) applies this method to 115 countries over the period 1961-1991. North America and Western Europe showed high productivity growth throughout the entire period. Asia and Sub-Sahara Africa exhibited negative productivity growth and the Latin American countries experienced positive TFP rates only in the eighties. Arnade (1998) estimates nonparametric Malmquist indices for 70 countries over the years 1961-1993 and finds that 36 of 47 developing countries show negative rates of TFP.

Recent empirical studies contradict this result of regression in TFP for the developing countries. Coelli and Rao (2003) examine agricultural productivity growth in 93 countries over the period 1980-2000 using data from FAOSTAT (2001). Asia is the major performer with an annual TFP growth of 2.9 percent and Africa seems to be the weakest performer with only 0.6 percent growth in TFP. Nevertheless, the phenomenon of negative productivity trends and technological regression is rejected. According to Coelli and Rao (2003, p. 14), "this is most likely a consequence of the use of a different sample period and an expanded group of countries". Nina *et al.* (2003) provide an additional argument: technical regression observed is probably the consequence of biased technical change together with the definition of technology used to estimate the Malmquist index. Nina *et al.* (2003) analyze agricultural productivity growth in developing countries over the period 1961-1994, applying a broader cumulative definition of technology than the one normally used to estimate the Malmquist index. The results confirm that agricultural TFP in developing countries is raising and that technical change is the main source for this growth.

The work of Nina *et al.* (2003) emphasizes the relevance of the definition of technology used in empirical studies. Martin and Mitra (2001) conclude the same within a parametric framework. The Cobb-Douglas production function employed in other studies is empirically rejected, concluding that the constancy of factor shares across countries and over time imposed by this specification is inadequate. Martin and Mitra (2001) adopt the translog form in a study of 50 countries in different development stages over the period 1967-1992 and conclude that the annual average TFP growth rates in developing countries are strong, between 1.76 and 2.62 percent, although lower than developed countries. Furthermore, technical progress has been faster in agriculture than in manufacturing for both developing and developed countries (Martin and Mitra, 2001). Mundlak (2000) concludes that agricultural labor and total factor productivity growth rates, from the 1950s through the late 1980s, have consistently exceeded their non-agricultural counterparts in the majority of countries. Bernard and Jones (1996) obtain similar results for 14 OECD countries over the period 1970-87, using time-series and cross-section parametric methods: the annual TFP growth rates is 2.6 and 1.2 percent, respectively for the agriculture sector and the industry. In addition, there is a tendency for a relatively rapid convergence in agricultural TFP across countries. In a sample of developed countries, this result seems unimportant. However, Martin and Mitra (2001) and Coelli and Rao (2003) conclude the same for a larger database, with countries in all development stages. Martin and Mitra (2001, p. 20) conclude that "at all levels of development, however, technical progress appears to have been faster in agriculture than in manufacturing" and that "there is strong evidence of convergence in levels and growth rates of TFP in agriculture, suggesting relatively rapid international dissemination of innovations". Coelli and Rao (2003, p. 14) find evidence that "countries that were well below the frontier in 1980 have a TFP growth rate of 3.6 percent", contrasting with "a low 1.2 percent growth for the countries that were on the frontier in 1980", clearly a sign of catch-up in productivity levels.

2.4 - Conclusions

The discussion of the previous section shows that there seems to be recent conflicting evidence on agricultural productivity growth both in relation to the other sectors of the economy as well as across countries. In our view, such evidence brings new interest to the issue of measuring agricultural productivity growth and investigating the role of agriculture in economic development. If empirical results such as those of Martin and Mitra (2001) are valid, then the agricultural sector may well have the potential for playing a decisive role in the growth strategies of developing countries and in reversing the trend for global divergence that has been observed in the last few decades. A large agricultural sector may be an advantage in terms of growth performance, weakening "the case for the frequently-advocated policies of discrimination against agriculture on the grounds that it is a stagnant sector" and potentially providing "an explanation for growth convergence at the macroeconomic level where growth rates slow down as the share of the agricultural sector declines" (Martin and Mitra, 2001, p.20).

Furthermore, this overview of existing literature reveals how particular assumptions on the technology, market structure and technological change affect results relative to the growth process, specially the fraction of labor productivity growth attributable to TFP growth and capital deepening. To overcome this dependency on the model specification, Kumar and Russell (2002) adopt a cross-section nonparametric frontier production method. The application of production frontier approaches to international growth-accounting has the main advantage of determining directly the factor accumulation and the TFP growth, decomposed further into the catch-up and technological change effects. None of the components is determined residually and technical change is measured through the shift of the frontier, avoiding estimation problems with the inclusion of a time trend to capture this effect. Nevertheless, although this linkage of economic growth studies to the production frontier literature looks fruitful, the choice of a deterministic and cross-section approach by Kumar and Russell (2002) seems controversial and worthy of improvement. Estimating a non-stochastic production frontier with information of only one year seems dangerous. Furthermore, the assumption of a single world production frontier can be challenged, particularly for samples including a large and heterogeneous set of countries. Kumbhakar and Wang (2005) show that ignoring heterogeneity invalidates the decomposition of TFP growth. Therefore, we will try to create stochastic models which account simultaneously for heterogeneity across countries and for the flexibility of the frontier specification.

If this assumption is not valid, technological differences may be labeled as inefficiency and the decomposition of output per worker is not valid. Therefore, it is important to find in the panel data production frontier literature, stochastic and flexible models that account for heterogeneity among countries in order to use or extend them to the purpose of this dissertation. In the next chapter we review the main methods of production frontier estimation that have been used recently in the literature.

3 – A Survey on Panel Data Production Frontier Methods

3.1 - Introduction

Farrell (1957) was the first to measure productive efficiency empirically. Nevertheless, his work would not be possible without the developments of theoretical literature on growth and productivity in the 1940s and 1950s under the auspices of the Cowles Commission, directed by Koopmans since 1948.

Koopmans was thinking, ever since his wartime days with the Combined Shipping Adjustment Board, about a systematic way to find the optimal routing plan for empty ships when there were fixed tonnages of cargo per month to go from one port to other ports. In seeking this systematic approach, Koopmans hits upon the principles of linear programming and activity analysis of production. Koopmans (1951) provides a definition of technical efficiency: a combination of inputs and outputs is said to be technical efficient when it is impossible to increase the rate of any output without at the same time increasing some input or decreasing some other output. However, only in Debreu (1951) it is possible to find a measure of productive efficiency, designated by the coefficient of resource utilization. Debreu (1951) uses a radial measure of efficiency, focusing on the maximum feasible equiproportionate reduction in all variable inputs or the maximum feasible equiproportionate expansion of all outputs.

Farrell (1957) extends the work initiated by Koopmans (1951) and Debreu (1951) and shows how to decompose economic efficiency into technical efficiency and allocative efficiency. Within a cost efficiency analysis, Farrell (1957) defines technical efficiency as the ability of a firm to minimize input use in the production of a given output vector. Allocative efficiency refers to the ability of a firm to use inputs in optimal proportions given their relative prices and the production technology.

It is interesting that the diffusion of Farrell's work was very slow. In fact, he was a Cambridge professor with no Ph.D. students and the journal he published in was not included in the social science databases of the United States. The first enthusiastic citation came by Amey (1964), an Operations Research scientist, and by Boles (1967), a Berkeley agricultural economist who wrote FORTRAN codes to improve the Farrell estimation method. Farrell (1957) illustrates the cost efficiency decomposition with an empirical application to U.S. agriculture using linear programming techniques. Farrell's work was important for the development of the Stochastic Frontier Analysis (SFA) independently proposed by Aigner *et al.* (1977) and Meeusen and van den Broeck (1977) and the DEA method by Charnes *et al.* (1978). Other important studies for the development of the frontier analysis directly built upon Farrell's work were the deterministic approach of Aigner and Chu (1968), the statistical foundation of frontier estimation developed by Afriat (1972), and a corrected ordinary least squares method introduced by Richmond (1974).

In the last twenty five years a wide range of methods have been developed for studying efficiency and productivity issues through the estimation of production frontiers. In this chapter, we briefly review the main stochastic panel data production frontier models that have recently been proposed in the literature. We classify those methods into five different categories: i) traditional parametric techniques of fixed, random effects and maximum likelihood, according to the taxonomy presented by Kumbhakar and Lovell (2000); ii) confidence interval methods of bootstrapping and Bayesian analysis, according to the taxonomy suggested by Schmidt and Kim (2001, p. 283); iii) semiparametric approaches; iv) nonparametric methods; v) recent parametric approaches that account for heterogeneity, such as the latent class models or the true fixed- and true random-effects models. We dedicate one section to each group of methods. In the final section of the chapter we summarize the discussion of the main strengths and weaknesses of the various models and justify the methodological option that we make in the following chapters.

3.2 - Traditional Parametric Techniques

In this section, we will follow the taxonomy presented by Kumbhakar and Lovell (2000) and try to explain the main features of each model.

3.2.1 - Origins and Advantages Over Cross-Section Data

Stochastic frontier analysis was independently proposed by Meeusen and van den Broeck (1977) and, one month later, by Aigner *et al.* (1977). Their production frontier specifications involve a stochastic composed error term associated with random shocks outside the control of producers that can affect output and a non-negative random variable representing technical inefficiency. This approach advances the deterministic frontier specification of Aigner and Chu (1968) involving only a non-negative random variable associated with technical inefficiency.

The production frontier model in both papers can be expressed as following:

$$y = \beta' x + v - u, \tag{3.1}$$

where:

y designates a scalar output; x a vector of inputs; β a vector of technology parameters; the error component $v \sim N[0, \sigma_v^2]$ captures the effects of statistical noise; and the other error component, $u \geq 0$, represents technical inefficiency.

Meeusen and van den Broeck (1977) assume an exponential distribution for u and Aigner *et al.* (1977) try both the exponential and the half-normal distributions. The specification of these distributional assumptions became one of the main criticisms in the stochastic frontier analysis since there is no a priori justification for the selection of any particular distributional form for the technical inefficiency term. This criticism led to the specification of more general distributional forms. Stevenson (1980) introduces the truncated-normal model; Greene (1980) proposes the two-parameter gamma model; and Lee (1983) suggests the four-parameter Pearson family of distributions. Other criticism made was the impossibility of obtaining the decomposition of individual residuals. Jondrow *et al.* (1982) devise a method of disentangling these effects using the mean of the conditional distribution $[u_i|v_i\text{-}u_i]$ to provide individual estimates of technical inefficiency. According to Jondrow *et al.* (1982, p. 234), "this was Farrell's (1957) original motivation for introducing production frontiers, and the ability to compare levels of efficiency across observations remains the most compelling reason for estimating frontiers".

Until the work of Pitt and Lee (1981), all efficiency measurement studies were cross-sectional. As pointed out by Schmidt and Sickles (1984, p. 367), these models have three problems: "first, the technical inefficiency of a particular observation can be estimated but not consistently. (…) Second, the estimation of the model and the separation of technical inefficiency from statistical noise require specific assumptions about the distribution of technical inefficiency and statistical noise. (…) Third, it may be incorrect to assume that inefficiency is independent of the regressors". A rich panel data can overcome some of these difficulties. In fact, individual technical efficiency can be estimated consistently with a panel data, whereas the estimates of technical inefficiency using the decomposition developed by Jondrow *et al.* (1982) are not consistent in a cross-sectional context. Furthermore, panel data make it possible to control for individual heterogeneity, which can lead to inconsistent estimation due to the correlation problem between the technical inefficiency term and the regressors. Finally, panel data estimation techniques can be adapted to the efficiency measurement problem while not requiring strong distributional or independence assumptions, since repeated observations on a sample of individuals can serve as a substitute for those assumptions.

3.2.2 - Time-invariant Technical Efficiency

a) Maximum Likelihood

Pitt and Lee (1981) are the first authors to analyze a panel data set by extending the cross-sectional Maximum Likelihood technique. The production frontier is estimated as:

$$y_{it} = \beta' x_{it} + v_{it} - u_i , \qquad (3.2)$$

where:

i refers to individuals (*i=1,...,N*); *t* to time periods (*t=1,...,T*); β is a *Kx1* vector of coefficients[3]; $v_{it} \sim N[0,\sigma_v^2]$; $u_i \sim N^+[0,\sigma_u^2]$.

Kumbhakar (1987) and Battese and Coelli (1988) generalize this specification by considering a truncated-normal distribution for the inefficiency term with mean μ and variance σ_u^2. We will present this model, since results for the half-normal specification can be easily obtained by defining μ=0.

The mixture of a normal distribution for the statistical noise with a truncated-normal distribution for the inefficiency term results in the following functional form for the log likelihood function for a sample of *N* individuals, each one observed over *T* periods of time:

$$Log\ L = \text{constant} - \frac{N(T-1)\log\sigma_v^2}{2} - \frac{N\log\left(\sigma_v^2 + T\sigma_u^2\right)}{2} - N\log\left[1-\Phi\left(-\frac{\mu}{\sigma_u}\right)\right] +$$

$$+\sum_{i=1}^{N}\log\left[1-\Phi\left(-\frac{\mu_i^*}{\sigma_i^*}\right)\right] - \frac{1}{2\sigma_v^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^2 - \frac{N}{2}\left(\frac{\mu}{\sigma_u}\right)^2 + \frac{1}{2}\sum_{i=1}^{N}\left(\frac{\mu_i^*}{\sigma_i^*}\right)^2, \qquad (3.3)$$

where:

$$\varepsilon_{it} = y_{it} - \beta' x_{it}\ ;\ \mu_i^* = \frac{\mu\sigma_v^2 - \sigma_u^2\sum_{t=1}^{T}\varepsilon_{it}}{\sigma_v^2 + T\sigma_u^2}\ ;\ \sigma_i^* = \frac{\sigma_u\sigma_v}{\sqrt{\sigma_v^2 + T\sigma_u^2}}.$$

In order to obtain estimates of the parameters, this function is maximized with respect to β. In a second step, the mean of the distribution of u_{it} given ε_{it} can be used as a point estimator of technical efficiency:

$$E\left(u_i \left|\varepsilon_{it}\right.\right) = \mu_i^* + \sigma_i^*\left\{\phi\left(-\frac{\mu_i^*}{\sigma_i^*}\right) \Big/ \left[1-\Phi\left(-\frac{\mu_i^*}{\sigma_i^*}\right)\right]\right\}. \qquad (3.4)$$

[3] from this point, we will omit the meaning of i, t and β in the production frontier specifications presented in this chapter.

Alternatively, producer-specific technical efficiency can be estimated using the minimum squared error estimator:

$$E\left[\exp\left(-u_i\right)\middle|\varepsilon_{it}\right] = \frac{\left[1-\Phi\left(\sigma_i^* - \dfrac{\mu_i^*}{\sigma_i^*}\right)\right]}{\left[1-\Phi\left(-\dfrac{\mu_i^*}{\sigma_i^*}\right)\right]}\exp\left(-\mu_i^* + \frac{\sigma_i^{*2}}{2}\right). \tag{3.5}$$

b) Fixed and Random Effects

Schmidt and Sickles (1984) apply fixed and random effects procedures on a panel towards the estimation of a stochastic production frontier.

Regarding the fixed effects model, its basic framework is given by the equation:

$$y_{it} = \alpha_i + \beta' x_{it} + v_{it}. \tag{3.6}$$

This model can be estimated consistently and efficiently by ordinary least squares. It is reinterpreted by treating α_i as a compound function of the independent term and the firm-specific inefficiency variable:

$$\alpha_i = \alpha_0 - u_i. \tag{3.7}$$

The inefficiency terms are fixed, but possibly correlated with the regressors and no distributional assumption is made on u_i.

To retain the essential characteristics of the frontier model, Schmidt and Sickles (1984) suggest that individuals be compared on the basis of:

$$u_i^* = \max_i \alpha_i - \alpha_i. \tag{3.8}$$

In order to assure that the most efficient individual has a score of one, technical efficiency is calculated as following:

$$TE_i = \exp\left\{-u_i^*\right\}. \tag{3.9}$$

A disadvantage of this model is that estimates of α_i's are only consistent as T tends towards infinity and consistency of the estimates of u_i* requires N and T tending to infinity. Furthermore, the fixed effect does not include only variation in inefficiency across firms, but all kinds of heterogeneity sources. Nevertheless, it has the advantage of simplicity and consistency does not depend on the distribution of u_i* or the independence assumption between u_i* and the regressors. In fact, when the distribution of the inefficiency effects is not known or if one has strong reasons to believe that they are correlated with the regressors, then the fixed-effects model is preferable to other approach.

In the random-effects approach, the u_i*'s* are assumed to be randomly distributed with mean μ and variance σ_u^2 and uncorrelated with v_{it} and the regressors. Thus, the production frontier is expressed by the equation:

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}, \tag{3.10}$$

where:

$\varepsilon_{it} = v_{it} - u_i$.

The only difference relative to the standard panel data models is that the individual effects u_i are one-sided. Schmidt and Sickles (1984) solve this problem using the transformation:

$$\alpha^* = \alpha - \mu ,$$

$$u_i^* = u_i - \mu ,$$

to obtain the following production frontier:

$$y_{it} = \alpha^* + \beta' x_{it} + v_{it} - u_{it}^*. \tag{3.11}$$

Schmidt and Sickles (1984) apply the standard two-step Generalized Least Squares (GLS) estimator to disentangle the individual effects from the residuals and use (3.9) as in the fixed-effects models to recover the technical efficiency index.

The random-effects approach, like the fixed-effects model and contrary to the maximum likelihood method, does not require a specific distribution on the efficiency term. However, it shares with the maximum likelihood technique the imposition of independence between the regressors and the inefficiency term.

3.2.3 - Time-variant Technical Efficiency

The assumption of time-invariant technical efficiency seems to be very unreasonable with large panels. "Particularly if the operating environment is competitive, it is hard to accept the notion that technical inefficiency remains constant through very many time periods." (Kumbhakar and Lovell, 2000, p. 108). It is impossible to eliminate this assumption without additional changes in the specification due to the excess of parameters to be estimated. Cornwell *et al.* (1990) and Kumbhakar (1990) were the first to propose a model with technical efficiency varying with time. The first study uses a fixed-random effects approach and a different pattern of technical efficiency variation for each individual while the second study applies a maximum likelihood technique with the assumption that technical efficiency varies in the same way for all individuals.

a) Common Pattern of Variation

a.1) Maximum Likelihood

This technique presumes independent and known technical inefficiency effects distributions. The production frontier is written as:

$$y_{it} = \beta' x_{it} + v_{it} - u_{it}, \tag{3.12}$$

where:

$u_{it} = \alpha_t \cdot u_i$; $v_{it} \sim N[0, \sigma_v^2]$; $u_i \sim N^+[0, \sigma_u^2]$.

The log-likelihood function for the entire sample is given by the following equation:

$$Log\ L \ = \ \text{constant} - \frac{NT\log\sigma_v^2}{2} - \frac{N\log\sigma_u^2}{2} - \frac{N\log\sigma^{*2}}{2} +$$

$$+\sum\nolimits_{i=1}^{N}\left\{\log\left[1-\Phi\left(-\frac{\mu_i^*}{\sigma^*}\right)\right]\right\} - \frac{1}{2\sigma_v^2}\sum\nolimits_{i=1}^{N}\left[\sum\nolimits_{t=1}^{T}\varepsilon_{it}^2 - \frac{\sigma_u^2\left(\sum\nolimits_{t=1}^{T}\alpha_t\varepsilon_{it}\right)^2}{\sigma_v^2+\sigma_u^2\sum\nolimits_{t=1}^{T}\alpha_t^2}\right], \tag{3.13}$$

where:

$\varepsilon_{it} = y_{it}\ -\ \beta' x_{it}$; $\mu_i^* = \dfrac{\sigma_v^2\sum\nolimits_{t=1}^{T}\alpha_t\varepsilon_{it}}{\sigma_v^2+\sigma_u^2\sum\nolimits_{t=1}^{T}\alpha_t^2}$; $\sigma^* = \dfrac{\sigma_u\sigma_v}{\sqrt{\sigma_v^2+\sigma_u^2\sum\nolimits_{t=1}^{T}\alpha_t^2}}$.

The mean of the distribution of u_i given ε_{it} can be used as a point estimator of technical efficiency:

$$E\left(u_i \left|\varepsilon_{it}\right.\right) = \mu_i^* + \sigma^*\left\{\phi\left(-\frac{\mu_i^*}{\sigma^*}\right) \Big/ \left[1-\Phi\left(-\frac{\mu_i^*}{\sigma^*}\right)\right]\right\}. \tag{3.14}$$

Alternatively, producer-specific technical efficiency can be estimated using the minimum squared error estimator:

$$E\left[\exp\left(-u_{it}\right)\left|\varepsilon_{it}\right.\right] = \frac{\left[1-\Phi\left(\alpha_t\sigma^* - \frac{\mu_i^*}{\sigma^*}\right)\right]}{\left[1-\Phi\left(-\frac{\mu_i^*}{\sigma^*}\right)\right]}\exp\left(-\alpha_t\mu_i^* + \frac{\alpha_t^2\sigma^{*2}}{2}\right) \tag{3.15}$$

There is in the literature two well known time-variant specifications with a common pattern of variation for all individuals. Kumbhakar (1990) proposes the following specification for α_t in (3.12):

$$\alpha_t = \frac{1}{1+\exp\left(\gamma t + \delta t^2\right)}. \tag{3.16}$$

Battese and Coelli (1992) propose a less flexible specification for α_t but convex in *t*:

$$\alpha_t = \exp\left[-\gamma\left(t-T\right)\right]. \tag{3.17}$$

Instead of considering the half-normal distribution for the inefficiency component, Battese and Coelli (1992) use a truncated-normal assumption:

$$u_{it} \sim \mathrm{N}^{+}[\mu, \sigma_\mu^2].$$

In this case, the individual efficiency estimates can be found applying the formula:

$$E\left[\exp\left(-u_{it}\right)\middle|\varepsilon_{it}\right] = \frac{\left[1-\Phi\left(\alpha_t\sigma^* - \frac{\mu_i^*}{\sigma^*}\right)\right]}{\left[1-\Phi\left(-\frac{\mu_i^*}{\sigma^*}\right)\right]}\exp\left(-\alpha_t\mu_i^* + \frac{\alpha_t^2\sigma^{*2}}{2}\right). \tag{3.18}$$

where:

$$\mu_i^* = \frac{\mu\sigma_v^2 - \alpha_t\sigma_u^2 \frac{\sum_{t=1}^{T}\varepsilon_{it}}{T}}{\sigma_v^2 + \alpha_t^2\sigma_u^2} \; ; \; \sigma^* = \frac{\sigma_u\sigma_v}{\sqrt{\sigma_v^2 + \alpha_t^2\sigma_u^2}}.$$

a.2) Fixed and Random Effects

Lee and Schmidt (1993) generalize the model of Schmidt and Sickles (1984) to the case of time-varying technical efficiency, using the following specification:

$$u_{it} = \alpha\left(t\right)\cdot u_i, \tag{3.19}$$

where:

$\alpha\left(t\right)$ is a set of time dummy variables, treated as coefficients of the inefficiency terms.

The temporal pattern of the inefficiency terms is more flexible, since it is not restricted to a specific parametric functional form, implying a more complex estimator of technical inefficiency. However, this formulation imposes a temporal pattern that is invariant across individuals. As in Kiefer (1980a)), Lee and Schmidt (1993) use a concentrated least-squares (CLS) estimator as the solution to an eigenvalue problem.

As in the time-invariant case, the most efficient unit should have a score of one. After estimating all parameters in the fixed- and random-effects models, technical efficiency is calculated as following:

$$TE_{it}=\exp\left\{-\left[\max_i\left(\alpha(t)\cdot u_i\right)-\alpha(t)\cdot u_i\right]\right\}. \quad (3.20)$$

Ahn *et al.* (2001) consider that the CLS estimator presented in Lee and Schmidt (1993) is consistent only if the terms ε_{it} are non-autocorrelated and with constant variance. Therefore, given these strong assumptions, the CLS estimator tends to be inefficient. Ahn *et al.* (2001) show that, for unrestricted $\alpha\left(t\right)$, a Generalized Method of Moments (GMM) estimator that makes use of the first- and second-order moment conditions, implied by the exogeneity of the regressors, non-autocorrelation and homoskedasticity of ε_{it}, dominates the CLS estimator, in the sense of being asymptotically more efficient. Han *et al.* (2005) show that the same results hold for a parametric (therefore, restricted) function, allowing for smoothness.

Greene (2005) argues that by interpreting the firm-specific term as inefficiency, the models of Schmidt and Sickles (1984), Cornwell *et al.* (1990) , Lee and Schmidt (1993) and Han *et al.* (2005) are assuming away any kind of cross-firm heterogeneity. Greene (2005) proposes a “true-fixed effects stochastic frontier model”, in which it is possible to disentangle the firm-specific term from inefficiency.

b) Different Pattern of Variation Across Individuals

b.1) Maximum Likelihood

Kumbhakar (1991) argues that estimates of u_{it} in equation (3.12) are biased, although consistent, since technical efficiency is not decomposed into individual-specific and time-specific effects. The production frontier is defined as:

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}, \qquad (3.21)$$

where:

$\varepsilon_{it} = v_{it} - u_{it}$; $u_{it} = \mu_i + \lambda_t$; v_{it} is the white noise disturbance term; u_{it} is the inefficiency component; μ_i is the firm-specific effect; λ_t is the time-specific effect.

Kumbhakar and Hjalmarsson (1993) propose a two-stage model. In the first stage, fixed- and random-effects models similar to the ones in Schmidt and Sickles (1984) are used with a different specification:

$$y_{it} = \alpha_0 + \beta' x_{it} + v_{it} + u_{it}, \qquad (3.22)$$

where:

$u_{it} = \tau_i + \xi_{it}$; τ_i is the firm-specific effect; ξ_{it} is the technical inefficiency component.

In the first stage, parameters are estimated without any distributional assumptions on the errors. In the second stage, the rest of the parameters are estimated by conditional maximum likelihood, assuming that $v_{it} \sim N[0, \sigma_v^2]$ and $\xi_{it} \sim N^+[0, \sigma_\xi^2]$.

Subjacent to this formulation is the idea that there are two components of technical efficiency: one is persistent and estimable in the same way as in Schmidt and Sickles (1984); the other one is residual and captured by the maximum likelihood estimator of the one-sided error component.

Kumbhakar and Lovell (2000, p. 115) note that “the problem with this approach is that any time-invariant component of technical inefficiency is captured by the fixed effects, rather than by the one-sided error component, where it belongs.” Greene (2002, p. 28) does not agree entirely since “whether those time invariant effects really belong in the inefficiency is debatable. (…) Once again, this is a methodological issue that deserves closer scrutiny."[4]

Some studies try to estimate simultaneously individual efficiencies and provide explanations for those differences with a single-stage estimation procedure. Battese and Coelli (1995) extend the model of Kumbhakar *et al.* (1991) and Reifschneider and Stevenson (1991) to a panel data framework.

The Battese and Coelli (1995) model specification may be expressed as:

$$y_{it} = x_{it}\beta + (v_{it} - u_{it}) , \tag{3.23}$$

where:

u_{it} are assumed to account for technical inefficiency in production and are assumed to be independently distributed as truncations at zero of $N(m_{it}, \sigma_u^2)$; $m_{it} = z_{it}\delta$; z_{it} is a (p×1) vector of variables which may influence the efficiency of an individual; δ is a (1×p) vector of parameters to be estimated.

The main problem of this model is that an incorrect choice of variables explaining differences in predicted individual efficiencies influences all results.

b.2) Fixed and Random Effects

Cornwell *et al.* (1990) propose a time-variant model where no assumptions are made on the distribution of the technical inefficiency effects. Their approach uses the following equation:

$$y_{it} = \alpha_{it} + \beta' x_{it} + v_{it}, \tag{3.24}$$

where:

$\alpha_{it} = \alpha_{0t} - u_{it}$; α_{0t} is the intercept common to all individuals in period t.

The specification adopted is:

[4] Please refer to Heshmati and Kumbhakar (1994) and Kumbhakar and Heshmati (1995) for further discussion.

$$\alpha_{it} = \eta_{i1} + \eta_{i2}t + \eta_{i3}t^2, \tag{3.25}$$

where:

t is time and η are the firm-specific parameters to be estimated.

The fixed- and random-effects approaches can be used to estimate this model. The method used in the fixed-effects approach changes with the size of the ratio $\frac{N}{T}$. If it is relatively small, the technical inefficiency terms are included in the model and η_{i1} parameters are treated as coefficients of dummies and η_{i2} and η_{i3} as coefficients of dummies interacted with the linear and quadratic time trends. As in the time-invariant approach, technical efficiency for each individual in time period t is determined as:

$$TE_{it} = \exp\left[-\left(\max_i \alpha_{it} - \alpha_{it}\right)\right]. \tag{3.26}$$

If the ratio $\frac{N}{T}$ is relatively large, u_{it} is not considered in the frontier equation and the vector β is estimated directly from the residuals. Then, the residuals are regressed on a constant, t and t^2 in order to obtain the individual-specific parameters η_{i1}, η_{i2} and η_{i3}.

Regarding the random-effects model, the approach is similar to the time-invariant case. The u_{it} are assumed to be uncorrelated with v_{it} and the regressors. The authors apply the standard two-step Generalized Least Squares (GLS) estimator to separate the individual effects from the residuals and use (3.26) as in the fixed-effects models to calculate individual technical efficiency. In order to relax the assumption of uncorrelation, Cornwell *et al.* (1990) develop an efficient instrumental variable estimator that is consistent even with correlation between the inefficiency and the regressors.

Through the text, some comments were made on each model. To summarize the main conclusions, the fixed-effects model is the only possible choice with short panels when we suspect there is correlation between inefficiency and the regressors. Nevertheless, as pointed out by Sena (2003, p. 15), “in this case, a lot of effort must be put to make sure that inefficiency is the only source of heterogeneity as picked up by the fixed effects.” If the regressors and inefficiency can be treated as independent and distributional assumptions can be made on technical inefficiency, then a maximum likelihood approach is more efficient than random-effects models. Long panels imply time-variant measures of technical efficiency.

3.3 - Confidence Intervals for Individual Inefficiencies

Schmidt and Kim (2001, p. 283) present a taxonomy of models that create confidence intervals for individual inefficiencies, generated by parametric or nonparametric traditional panel data approaches. Therefore, some of its properties (e.g., flexibility of the production frontier) are highly sensible to the framework adopted. Although these techniques are an extension of models presented in the previous section, the advantages brought to the literature justify its presentation in an autonomous section.

3.3.1 - Bootstrapping

The bootstrap is a computer-intensive non-parametric method introduced by Efron (1979). An unknown distribution is approximated by the empirical distribution of the original sample, and the data are resampled, with replacement, to obtain the bootstrap sample and confidence intervals for the unknown parameters.

Suppose $x_1, x_2, \ldots, x_n$ are independent and identically distributed random variables from a population with unknown cumulative distribution function (cdf) F, and suppose the goal is to draw inference about some parameter θ of the population. Let $\hat{\theta}(x_1, x_2, \ldots, x_n)$ be an estimator of θ and let $\hat{F}$ be the sample cdf, that is, the cdf that assigns mass $\frac{1}{n}$ to each x. The bootstrap approximates the distribution of θ under F by the sampling distribution of $\hat{\theta}$ under $\hat{F}$ (Efron, 1979).

The application of bootstrapping to efficiency analysis is relatively recent. Simar (1992) was the first to apply the percentile bootstrapping method defined by Efron (1979) to the problem of estimating frontier models and inefficiencies within standard econometric methods (fixed- and random-effects) for panel data. This procedure can be summarized as follows:

(1) Calculate the fixed-effects estimates and obtain $\hat{v}_{it} = y_{it} - \hat{\alpha}_i - \hat{\beta}' x_{it}$ from the original data set, where *i=1,...n*; and *t=1,...,T;*

(2) Construct $\hat{F}\left(\hat{v}_{it}\right)$ by associating mass $\frac{1}{nT}$ at each observed residual;

(3) Draw a random sample of size nT with replacement from a smoothed $\hat{F}$;

(4) Independently repeat (3) B times to provide the set of bootstrap estimates $\hat{v}_{it}^{*b}$; $b = 1,...,B$ and the pseudo-data $y_{it}^{*b} = \hat{\alpha}_i + \hat{\beta}' x_{it} + \hat{v}_{it}^{*b}$;

(5) Calculate the bootstrap estimators $\alpha_i^{(b)}, \beta^{(b)}, u_i^{*(b)}, TE_i^{(b)}$ and obtain the confidence interval based on the bootstrap percentile method for each parameter as an approximate $(1-2\alpha)$ central confidence interval of the bootstrap distribution.

Hall *et al.* (1993) show that the bootstrap estimators proposed by Simar (1992) have consistency problems with small samples. This problem has been raised in the literature (Efron,1985), but it gains even more relevance in the context of efficiency analysis due to the calculation of u_i*. The assumption of assymptotic normality is very difficult to accept with the maximum function (necessary to calculate u_i*), introducing an important bias with small to moderate sample sizes. When inconsistency occurs, the bootstrap distribution estimator does not even converge in probability. Hall *et al.* (1995) suggest the use of the iterated bootstrap (also known as the double bootstrap) to correct for the coverage probability of confidence intervals obtained by the percentile method relatively to the maximum of the intercepts in a fixed-effects model.

Efron (1982, 1985) introduces the bootstrap bias-corrected percentile method which also uses the percentiles of the bootstrap distribution, but not exactly the $\alpha\, th$ and $(1-\alpha)\, th$. Instead, it corrects these values for possible bias in the estimation of the partameters. Simar and Wilson (1998, 2000) use this bootstrap bias-corrected percentile approach to efficiency measures obtaineded by DEA or FDH (non-stochastic methods).

These works were possible due to the discussion of the nature of the Data-Genetating Process (DGP) implicit in DEA models. The steps of the algorithm are the following:

1. Calculate the Malmquist Index for each observation and obtain $\hat{M}_i$ for i = 1,...,n;

2. Construct the density function for $\hat{M}$ by putting mass $\frac{1}{n}$ at each observed Malmquist Index;

3. Extract with reposition *n* observations from the original sample in order to build the pseudo-samples;

4. Calculate the *bootstrap* estimator $\hat{M}_i^*$;

5. Repeat B times (3) and (4) to obtain B *bootstrap* estimates $\hat{M}_i^*$;

6. Calculate $Bias = \frac{\sum_{b=1}^{B} \hat{M}_{b,i}^*}{1000} - \hat{M}_i$;

7. Calculate the bias-corrected *bootstrap* estimator $\tilde{M}_i^* = \hat{M}_i - bias = 2\hat{M}_i - \frac{\sum_{b=1}^{B} \hat{M}_{b,i}^*}{1000}$;

8. Obtain confidence intervals $(\tilde{M}_i^{*(\alpha)}, \tilde{M}_i^{*(1-\alpha)})$ where $\tilde{M}_i^{*(\alpha)}$ stands for the 100α percentile of the sampling density function of $\tilde{M}_i^*$.

3.3.2 - Bayesian

Koop *et al.* (1997) create a panel data Bayesian model for making inferences about firm-specific inefficiencies, which are assumed to be constant over time. The approach of Schmidt and Sickles (1984) is extended to the Bayesian framework for both fixed- and random-effects models. These Bayesian tools are extended by Koop *et al.* (1997) to cost frontiers and by Koop and Steel (2001) to the context of production frontiers. Therefore, we will use this latter study as the main reference of this section:

$$y_i = \beta_0 \iota_T + x_i \delta + \upsilon_i - z_i \iota_T, \tag{3.27}$$

where:

$i = 1,\ldots,N$; $t = 1,\ldots,T$; y_i is a T x 1 vector of outputs; β_0 is the intercept coefficient; ι_T is a T x 1 vector of ones; x_i a T x k matrix of inputs; δ is a k x 1 vector of unknown parameters; υ_i is a T x 1 vector of the error term, assumed to be *i.i.d.* with probability density function (pdf) $f_N^T\left(\upsilon_i \middle| 0, h^{-1} I_T\right)$, with h being a parameter to be estimated; z_i refers to the firm-specific time-invariant inefficiency.

In the Bayesian fixed-effects model, the individual effect is given by the expression:

$$\alpha_i = \beta_0 - z_i; \tag{3.28}$$

and the model is rewritten as:

$$y_i = \alpha_i \iota_T + x_i \delta + \upsilon_i. \tag{3.29}$$

The classical (frequentist) version of this model uses firm-specific dummy variables for α_i. The Bayesian way of doing this is to use flat, non-informative priors for the $\alpha_i s$.

Defining $\alpha = \left(\alpha_1^{'} \ldots \alpha_N^{'}\right)^{'}$, Koop and Steel (2001) adopt the prior $p\left(\alpha, \delta, h\right) \propto h^{-1} p\left(\delta\right)$.

This prior seems innocuous but since it is impossible to disentangle z_i from β_0 it implies a rather unusual prior for the relative inefficiency measure, favoring low efficiency:

$$z_i^{rel} = z_i - \min_j\left(z_j\right) = \max_j\left(\alpha_j\right) - \alpha_i. \tag{3.30}$$

Relative efficiency is defined as $r_i^{rel} = e^{\left(-z_i^{rel}\right)}$.

$p(r_i^{rel})$ has a point mass of N^{-1} at full efficiency with the prior $p(r_i^{rel}) \propto 1/r_i^{rel}$ for $r_i^{rel} \in (0,1)$. This is an L-shaped improper prior density which, for an arbitrary small $a \in (0,1)$ puts an infinite mass in $(0, a)$ but only a finite mass in $(a, 1)$. Therefore, the prior favors low efficiency.

For future references, $f_N^k(a|b, C)$ indicates that *a* is a *k*-variate normal with mean *b* and a covariance matrix *C* and $f_G(d|g, l)$ indicates that the density function of *d* is a Gamma distribution with shape parameter *g* and scale *l*.

After some calculations, Koop and Steel (2001) find the marginal posterior for δ:

$$p(\delta|y, x) = f_N^k\left(\delta\left|\hat{\delta}, \hat{h}^{-1}S^{-1}\right.\right)p(\delta), \tag{3.31}$$

where:

$f_N^k\left(\delta\left|\hat{\delta}, \hat{h}^{-1}S^{-1}\right.\right)$ indicates that δ is a k-variate normal with mean $\hat{\delta}$ and a covariance matrix $\hat{h}^{-1}S^{-1}$; $\hat{\delta} = S^{-1}\sum_{i=1}^{N}\left(x_i - \iota_T\overline{x_i}\right)'\left(y_i - \iota_T\overline{y_i}\right)$; $S = \sum_{i=1}^{N} S_i, \overline{x_i} = \frac{1}{T}\iota_T' x_i$; $S_i = (x_i - \iota_T\overline{x_i})'(x_i - \iota_T\overline{x_i})$; $\hat{h}^{-1} = \frac{1}{N(T-1)-k}\sum_{i=1}^{N}\left(y_i - \hat{\alpha}_i\iota_T - x_i\hat{\delta}\right)'\left(y_i - \hat{\alpha}_i\iota_T - x_i\hat{\delta}\right)$; $\hat{\alpha}_i$ is the posterior mean of α_i.

This is the standard within estimator from the panel data literature.

The marginal posterior of α is the *N*-variate normal with means:

$$\hat{\alpha}_i = \overline{y}_i - \overline{x}_i\hat{\delta}; \tag{3.32}$$

and covariances

$$\widehat{\text{cov}}\left(\alpha_i, \alpha_j\right) = \hat{h}^{-1}\left(\frac{\Delta(i, j)}{T} + \overline{x_i}S^{-1}\overline{x_j}'\right), \tag{3.33}$$

where $\Delta(i, j) = 1$ *if* $i = j$; and 0 otherwise.

For relative inefficiencies, r_i^{rel}, analytical expressions for posterior means and standard deviations are not available and, therefore, posterior simulation methods are necessary, such as Gibbs sampling or Monte Carlo integration.

The calculation of the posterior for r_i^{rel} in the cases where the firm is not the most efficient unit results in the expression:

$$p\left(r_i^{rel} \middle| y, x\right) = \sum_{j=1, j \neq i}^{N} p\left(r_i^{rel} \middle| y, x, r_j^{rel} = 1\right) P\left(r_j^{rel} = 1 \middle| y, x\right), \tag{3.34}$$

where:

$P\left(r_i^{rel} = 1 \middle| y, x\right) = P\left(\alpha_i = \max_j\left(\alpha_j\right) \middle| y, x\right)$ is the probability[5] that a given firm *i* is the most efficient, which can be calculated using Monte Carlo integration and $p\left(r_i^{rel} \middle| y, x, r_j^{rel} = 1\right)$ can be obtained from a posterior simulation method.

The Bayesian fixed-effects model is not very attractive due to unreasonable prior assumptions made for the relative inefficiencies. Therefore, Koop and Steel (2001) derive a Bayesian random-effects model by combining the production frontier in (3.27) with the prior:

$$p\left(\beta_0, \delta, h, z, \lambda^{-1}\right) \propto h^{-1} p\left(\delta\right) f_G\left(\lambda^{-1} \middle| 1, -\ln\left(\tau^*\right)\right) \prod_{i=1}^{N} f_G\left(z_i \middle| 1, \lambda^{-1}\right), \tag{3.35}$$

where:

$-\ln\left(\tau^*\right)$ is a non-negative random variable; *h* and β_0 have non-informative priors and inefficiencies are assumed to follow an exponential distribution with mean λ.

[5] This notation in capital letters is used to distinguish from posterior probabilities.

Koop and Steel (2001) set up a Gibbs sampler with data augmentation. Defining $\beta = (\beta_0 \quad \delta')'$ and $X = (\iota_{NT} : x)$, the posterior conditional for the measurement error precision can be written as:

$$p\left(h \middle| y, x, z, \beta, \lambda^{-1}\right) = f_G\left(h \middle| \frac{NT}{2}, \frac{1}{2}\left[y - X\beta + (I_N \otimes \iota_T) z\right]'\left[y - X\beta + (I_N \otimes \iota_T) z\right]\right); \quad (3.36)$$

and

$$p\left(\beta \middle| y, x, z, h, \lambda^{-1}\right) = f_N^{k+1}\left(\beta \middle| \overline{\beta}, h^{-1}(X'X)^{-1} p(\delta)\right), \quad (3.37)$$

where:

$\overline{\beta} = (X'X)^{-1}\left[y + (I_N \otimes \iota_T) z\right]$.

The posterior conditional for the inefficiencies takes the form:

$$p\left(z \middle| y, x, \beta, h, \lambda^{-1}\right) \propto \quad f_N^N\left(z \middle| \left(\iota_N : \tilde{x}\right)\beta - \tilde{y} - (Th\lambda)^{-1}\iota_N, (Th)^{-1} I_N\right) \prod_{i=1}^{N} I(z_i \geq 0), \quad (3.38)$$

where $\tilde{y} = \left(\tilde{y}_1 ... \tilde{y}_N\right)'$; $\tilde{x} = \left(\tilde{x}_1' ... \tilde{x}_N'\right)'$ and $I(.)$ is the indicator function, assuming the value of 1 if the event occurs and 0 otherwise.

Furthermore, the posterior conditional for λ^{-1} is given by:

$$p\left(\lambda^{-1} \middle| y, x, z, \beta, h\right) = f_G\left[\lambda^{-1} \middle| N + 1, z'\iota_N - \ln(\tau^*)\right]. \quad (3.39)$$

Using these results, Bayesian inference can be carried out using a Gibbs sampling algorithm.

The model of Koop *et al.* (1997) presented in this section with the production frontier version of Koop and Steel (2001) was extended in many directions. Koop *et al.* (1999, 2000) relax the assumption of time-invariant inefficiency; Fernández *et al.* (2000, 2005) apply the method to the multiple output case; Koop and Poirier (2004) abandon the assumption of a linear functional form to the frontier.

Studies of Koop *et al.* (1999, 2000) are relevant to this survey for two reasons: they present Bayesian models with time-variant inefficiency and simultaneously analyze the decomposition of output and productivity growth in a panel of countries. The frontier common to all countries is given by the expression:

$$Y_{it} = f_t\left[K_{it}, L_{it}\right]\tau_{it} w_{it}, \tag{3.40}$$

where:

Y_{it}, K_{it}, L_{it} stand for real output, capital stock and labor, respectively, in country *i* in time period *t*; f_t is the production frontier; τ_{it} designates the efficiency term and w_{it} is the error term.

Regarding the production frontier, it is important to note that Koop *et al.* (1999, 2000) assume variations of a translog production frontier:

$$y_{it} = x_{it}^{'}\beta_t + \upsilon_{it} - u_{it}, \tag{3.41}$$

where:

y_{it} stands for the natural log of output; $x_{it} = \begin{pmatrix}1 & k_{it} & l_{it} & k_{it}l_{it} & k_{it}^2 & l_{it}^2\end{pmatrix}$ with $k_{it} = \ln K_{it}$ and $l_{it} = \ln L_{it}$; $u_{it} = -\ln\left(\tau_{it}\right)$ is a non-negative random variable; $\upsilon_{it} = \ln(w_{it})$ and it is assumed to have an symmetric distribution with mean zero.

Changes in productivity of country *i* between time periods *t* and *(t+1)* is measured by the output-oriented Malmquist index:

$$PC_{i,t+1} = TC_{i,t+1} \times EC_{i,t+1}, \tag{3.42}$$

where:

$TC_{i,t+1} = \exp\left[0.5\left(x_{i,t+1} + x_{it}\right)^{'}\left(\beta_{t+1} - \beta_t\right)\right]$ represents technical change and $EC_{i,t+1} = \exp\left[u_{it} - u_{i,t+1}\right] = \dfrac{\tau_{i,t+1}}{\tau_{it}}$ is the efficiency change.

Koop *et al.* (1999, 2000) assume that the production frontiers are independent across time, which is equivalent to estimate T independent cross-sectional stochastic production frontiers. Therefore, the Bayesian model is given by:

$$\prod_{t=1}^{T} f_N^N\left(y_{it}\left|x_t\beta_t - u_t, \sigma_t^2 I_N\right.\right) p\left(\beta_t\right) p\left(\sigma_t^{-2}\right) p\left(\lambda_t\right) \prod_{i=1}^{N} f_G\left(u_{it}\left|1, \lambda_t^{-1}\right.\right). \tag{3.43}$$

Given that the time-variant efficiency model in Koop *et al.* (1999, 2000) requires the estimation of T independent cross-sectional production frontiers, the panel structure of the data is not properly and fully exploited. Therefore, all results for the cross-section models of van den Broeck *et al.* (1994) can be imported into this panel data approach.

Griffin and Steel (2004) show the limitations of the exponential parametric model that is often used in this literature (e.g., Koop *et al.* , 1997). Griffin and Steel (2004, p. 149), conclude that "predicting the efficiency for an unobserved firm on the basis of this parametric model is shown to be totally misleading" and propose a Bayesian semiparametric approach for stochastic frontier models[6]. The stochastic production function proposed in Griffin and Steel (2004) is defined as follows:

$$y_{it} = \alpha + x_{it}^{'}\beta + \upsilon_{it} - u_i, \tag{3.44}$$

where:

x_{it} is a vector of appropriate explanatory variables, υ_{it} represents an i.i.d. error term reflecting measurement and specification errors with $\upsilon_{it} \sim N\left(0, \sigma^2\right)$, and u_i is the one-sided disturbance, representing time-invariant inefficiency, which is the nonparametric component of the model, since it is assumed to be independently distributed as $u_i \sim F$, with F being a random probability measure based on a Dirichlet process.

The model is estimated using a Markov Chain Monte Carlo (MCMC) algorithm, which is a modification of the Gibbs sampler described in Koop *et al.* (1997). The main difference between the Dirichlet process-based approach and the parametric approach described by Koop *et al.* (1997) is the form of the full conditional distribution for u. The

[6] These methods eill be discussed in section 3.4.

inefficiency term is assumed to be the same for some units, grouped in each cluster *k*. The sampler uses a data augmentation scheme where each observation is associated with an element of the Dirichlet process using latent variables. Griffin and Steel (2004) also explore the case in which the inefficiency distributions are allowed to vary with some covariates, using a slight modification of the MCMC sampler.

3.4 - Semiparametric Methods

Semiparametric models involve parametric and nonparametric components. In general, the stochastic production frontier is defined parametrically and the inefficiency distribution is generated in a nonparametric fashion. A clear advantage of these models is the possibility to relax strong parametric assumptions on the distribution of the inefficiency term, mitigating possible specification errors.

Park and Simar (1994) and Park *et al.* (1998) estimate a semiparametric panel data frontier, in the sense that individual effects have an unknown density function. The production frontier is assumed to be linear:

$$y_{it} = \alpha_i + \beta x_{it} + \upsilon_{it}, \tag{3.45}$$

where:

α_i are *i.i.d.* random variables from an unknown density h whose support is bounded above by $B(h)$ and υ_{it} are *i.i.d.* random variables from $N(0, \upsilon^2)$. Therefore, the deterministic production frontier is given by $\beta\mathbf{x}_{it} + B(h)$ and technical efficiency of the i-th observation is calculated as $\alpha_i - B(h)$.

In the model of Park and Simar (1994), the individual effects and regressors are assumed to be independent. In this case, the semiparametric efficient estimator is the within estimator. Park *et al.* (1998) extend this model by allowing two cases of dependency: (i) between the individual effects and a subset of regressors and (ii) between the effects and long-run movements in a subset of regressors.

In the first case, the semiparametric efficient estimator is:

$$\hat{\beta} = \hat{\beta}^{IV} + N^{-1}\sum_{i=1}^{N} \hat{\hat{l}}, \tag{3.46}$$

where:

$\hat{\beta}^{IV}$ is the Hausman-Taylor instrumental variable estimator of β and $\hat{\hat{l}}$ is a nonparametric kernel estimator with (1 + Tq) dimensions, where T is the number of time periods and q is the number of regressors correlated with the individual effects

In the second case, the efficient semiparametric estimator is the same as in (3.46), but now $\hat{\hat{l}}$ is a nonparametric kernel estimator, with (1 + q) dimensions.

Adams *et al.* (1999) apply the model of Park *et al.* (1998) in two distinct ways. First, Adams *et al.* (1999) estimate a stochastic distance frontier, where a subgroup of multiple regressors is correlated with the individual effects. Second, Adams *et al.* (1999) assume independence between the covariates and the individual effects but use a nonparametric regression, in which no functional form is imposed on the distance function.

In the first model, Adams *et al.* (1999) derive the semiparametric efficient estimator for the panel Cobb-Douglas stochastic distance frontier, combining a Hausman-Taylor estimator with a fourth-order normal kernel. In the second model, Adams *et al.* (1999) apply the nonparametric Nadaraya-Watson estimator to find the conditional expectation of the output given the regressors. Both models mitigate possible specification errors that can occur in estimation.

3.5 - Nonparametric Models

In this class of methods, the production frontier is not estimated parametrically, with the purpose of removing any influence of a particular specification on the results. Furthermore, it is not assumed a specific distribution for the inefficiency term[7].

Henderson and Ullah (2004, 2005) apply to a production frontier model the nonparametric method introduced by Cleveland (1979) that accommodates statistical noise using random-effects procedures on a panel data as follows:

[7] For a detailed analysis of nonparametric methods, please refer to Hardle (1990) and Pagan and Ullah (1999).

$$y_{it} = m(x_{it}) + \varepsilon_{it}, \tag{3.47}$$

where:

y_{it} is the endogenous variable; x_{it} is a vector of k exogenous variables; $m(.)$ is an unknown smooth function and ε_{it} is the error component.

In this framework, heterogeneity in the distribution of y_{it} is assumed to impact the density function in the form of a random effect. Assuming that $\beta(x_{it})$ is a varying gradient vector of the function $m(x_{it})$, estimates of these parameters can be obtained by using local polynomial estimation. Cleveland (1979) introduces the form of a local polynomial regression that is referred to as LOWESS. Simply, LOWESS is a locally-weighted regression with the local weights being defined by a kernel function (Hastie and Tibshirani, 1990). The basic idea is to approximate the unknown production frontier $m(x_{it})$ in a neighborhood of x_0 by a polynomial of degree 1 on x_{it}, using the Taylor series expansion:

$$m(x_{it}) \approx m(x_0) + m^{(1)}(x_0)(x_{it} - x_0). \tag{3.48}$$

This function can be fitted locally by the minimization of a weighted least squares (WLS) regression:

$$\min \sum_{i=1}^{N} \sum_{t=1}^{T} \left[y_{it} - \beta_0 - \beta_1 (x_{it} - x_0) \right]^2 K\left(\frac{x_{it} - x_0}{h} \right), \tag{3.49}$$

where:

$K(.)$ represents a kernel function with the bandwidth parameter h that controls the size of the neighborhood.

The minimization problem in (3.49) can be stated in the matrix form:

$$\min_{\beta} \left[Y - X\beta(x_0) \right]' W(x_0) \left[Y - X\beta(x_0) \right], \tag{3.50}$$

where:

$Y = \left(y_{11},...,y_{1T},...,y_{N1},...,y_{NT}\right)'$; $\beta\left(x_0\right) = \left(\beta_0, \beta_1\right)'$;

$X = \begin{bmatrix} 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \\ \left(x_{11}-x_0\right) & \cdots & \left(x_{1T}-x_0\right) & \cdots & \left(x_{N1}-x_0\right) & \cdots & \left(x_{NT}-x_0\right) \end{bmatrix}'$; and $W\left(x_0\right)$ is an $\left(NT \times NT\right)$ matrix of kernel-based weights.

The minimization results in the following solution:

$$\hat{\beta}\left(x_0\right) = \left(X'W\left(x_0\right)X\right)^{-1} X'W\left(x_0\right)Y. \tag{3.51}$$

Henderson and Ullah (2004, 2005) argue that the estimator presented in (3.51) ignores the information contained in the error component. Instead of doing simple WLS fits to the points local to x_0, Henderson and Ullah (2004, 2005) prefer to perform Local Linear Weighted Least Squares regressions, after estimating the covariance matrix **Ω** of the disturbance vector. This approach extends the Local Linear Generalized Least Squares Estimator suggested by Ullah (2001) within a cross-section framework.

Assuming that Ω is a $\left(NT \times NT\right)$ covariance matrix of the disturbance term ε_{it}, the objective function changes to:

$$\min_{\beta}\left[Y - X\beta\left(x_0\right)\right]' \sqrt{W\left(x_0\right)}\Omega^{-1}\sqrt{W\left(x_0\right)}\left[Y - X\beta\left(x_0\right)\right], \tag{3.52}$$

resulting in the solution vector:

$$\hat{\beta}\left(x_0\right) = \left(X'\sqrt{W\left(x_0\right)}\Omega^{-1}\sqrt{W\left(x_0\right)}X\right)^{-1} X'\sqrt{W\left(x_0\right)}\Omega^{-1}\sqrt{W\left(x_0\right)}Y. \tag{3.53}$$

Henderson and Ullah (2004, 2005) consider three different functions for Ω, according to the estimators used in Lin and Carroll (2000) and Ullah and Roy (1998). Henderson (2004), Henderson and Ullah (2004, 2005) divide the error term into inefficiency and stochastic error components:

$$\varepsilon_{it} = v_{it} - u_i, \tag{3.54}$$

where:

v_{it} is the error term with $v_{it} \sim i.i.d.(0, \delta_v^2)$ and u_i is the inefficiency component, with $u_i \sim i.i.d.(\mu, \delta_u^2)$.

The production frontier in (3.47) is modified since the mean of the inefficiency component is non-zero:

$$y_{it} = m(x_{it}) + \varepsilon_{it} \Leftrightarrow y_{it} = m(x_{it}) - \mu + v_{it} + \mu - u_i \Leftrightarrow y_{it} = m^*(x_{it}) + v_{it} + u_i^*, \tag{3.55}$$

where $u_i^* \sim i.i.d.(0, \delta_u^2)$.

Assuming that the efficient component and the shocks are uncorrelated, it is possible to obtain the covariance matrix of the disturbance term:

$$\Omega = \mathrm{E}(\varepsilon_{it}\varepsilon_{it}') = \mathrm{E}(\varepsilon_i\varepsilon_i') \otimes \mathrm{I}_T = \left[\delta_v^2 \mathrm{I}_T + \delta_u^2 i_T i_T'\right] \otimes \mathrm{I}_T, \tag{3.56}$$

where I_T is a $T \times T$ identity matrix and i_T is a $T \times 1$ vector of ones.

The estimation of Ω requires the calculation of $\hat{\delta}_u^2$ and $\hat{\delta}_v^2$. It can be used the spectral decomposition of the covariance matrix, resulting in the following consistent estimators:

$$\hat{\delta}_v^2 = \frac{1}{N(T-1)} \sum_i \sum_t \left\{ [y_{it} - \hat{m}_0(x_{it})] - \frac{1}{T} \sum_t [y_{it} - \hat{m}_0(x_{it})] \right\}^2; \tag{3.57}$$

$$\hat{\delta}_u^2 = \frac{1}{N \cdot T} \sum_i \sum_t [y_{it} - \hat{m}_0(x_{it})]^2 - \frac{1}{T} \hat{\delta}_v^2, \tag{3.58}$$

where:

$\hat{m}_0(x_{it})$ is obtained using a simple WLS estimator.

These estimates can be used to obtain the estimator of Ω necessary to estimate $m(x_{it})$.

Henderson (2004) obtains estimates of u_i^* by maximization of the following objective function with respect to u_i^*:

$$\max_{u_i^*} \frac{1}{\hat{\delta}_v^2} \sum_i \sum_t \left[y_{it} - \hat{m}(x_{it}) - u_i^* \right]^2 + \frac{1}{\hat{\delta}_u^2} \sum_i \sum_t u_i^{*2}. \qquad (3.59)$$

Solving the maximization problem in (3.59) results in the following expression:

$$\hat{u}_i^* = \left(\frac{\hat{\delta}_u^2}{\hat{\delta}_v^2 + T \cdot \hat{\delta}_u^2} \right) \left[y_{it} - \hat{m}(x_{it}) \right]. \qquad (3.60)$$

In order to obtain an individual technical efficiency index, Henderson (2004) proceeds to the usual normalization in the literature:

$$\widehat{TE}_i = \exp\left[-\left(\max_i \hat{u}_i - \hat{u}_i \right) \right]. \qquad (3.61)$$

3.6 - Parametric Methods that Account for Heterogeneity of Production Units

Since the studies of Laird (1978) and Heckman and Singer (1984), the problem of latent heterogeneity has been accounted within panel data models. Nevertheless, only very recently this question has been addressed in the stochastic frontier literature (Greene, 2001a, 2001b, 2003, 2005; Tsionas, 2002; Orea and Kumbhakar, 2004). If all units in the sample face exactly the same production possibilities set and differ only with respect to their degree of inefficiency, the traditional stochastic frontier model is adequate. Nevertheless, "in practice, production possibilities are expected to differ in a cross-section of firms, and a set of different technologies may simultaneously coexist at any given time" (Tsionas, 2002, pp. 128). In these cases, classical approaches can not be used. We will analyze three different approaches addressing this question: the random parameters approach suggested by Tsionas (2002); the latent class model developed by Greene (2001a, 2001b, 2003) and Orea and Kumbhakar (2004) and the true fixed-effects and true random-effects models of Greene (2005).

Tsionas (2002) relaxes the assumption that the frontier is common to all firms by proposing a random coefficient stochastic frontier model where absolute firm-specific efficiency can be separated from technological differentials across firms. Exact finite sample results for parameters as well as latent efficiencies are derived using a Bayesian MCMC method, more specifically the Gibbs sampler with data augmentation. The model structure suggested in Tsionas (2002) is:

$$y_{it} = \alpha + x_{it}'\beta_i + v_{it} - u_{it}, \tag{3.62}$$

where:

v_{it} is the measurement error distributed as i.i.d. $N[0,\sigma_v^2]$ and u_{it} is the inefficiency component with an exponential density;

$$u_{it} \sim \theta \exp(-\theta u_{it}),\ \theta > 0,\ u_{it} \geq 0, \tag{3.63}$$

where:

$E[u_{it}] = 1/\theta$ and $Var[u_{it}] = 1/\theta^2$.

In this model, each firm has its own production frontier, since parameters β_i reflect heterogeneity in the technology. Furthermore, each observation experiences a shock which determines its inefficiency level u_{it} from an exponential distribution.

Parameters β_i are distributed according to a (K-1)-variate normal distribution: $\beta_i \sim N[\overline{\beta}, \Omega]$, where $\overline{\beta}$ is a (K-1)x1 vector of parameter means and Ω is a $(K-1)$x$(K-1)$ positive definite covariance matrix.

This assumption implies that model can be written as:

$$y_{it} = \alpha + x_{it}'\overline{\beta} + e_{it} - u_{it} \Leftrightarrow y_{it} = z_{it}'\delta + e_{it} - u_{it}, \tag{3.64}$$

where:

e_{it} is i.i.d. with $e_{it} \sim N[0,\sigma_v^2 + x_{it}'\Omega x_{it}]$; $z_{it}' = [1\ \ x_{it}']$ and $\delta = [\alpha\ \ \overline{\beta}']'$.

Therefore, the model in (3.64) is similar to the stochastic frontier model with a normal, heteroscedastic[8] measurement error. Tsionas (2002) uses Gibbs sampling to estimate the posterior means and variances of the various quantities of interest in the model: α, $\overline{\beta}$, σ_v, Ω, θ, and u.

Latent class models have been developed in the context of several cross-sectional studies (e.g., Quandt and Ramsey, 1978; Kiefer, 1979, 1980a, 1980b; Poirier and Ruud, 1981). Only after the works of Heckman and Singer (1984), Wedel *et al.* (1993), Nagin and Land (1993) and Wang *et al.* (1998) developed in a panel data framework, latent class models have been used within the stochastic frontier literature. The basic assumption of these models is that there is a latent sorting of the observations in the data set into J latent classes, unobserved by the econometrician. We will try to present briefly the studies of Greene (2001a, 2001b, 2003) and Orea and Kumbhakar (2004) using the single theoretical production frontier model summarized in Greene (2005):

$$y_{it} \mid j \;=\; \alpha_i + \beta_j' x_{it} + v_{it} \mid j - u_{it} \mid j, \tag{3.65}$$

where:

j refers to the class number, with $j=1,\ldots,J$; $v_{it} \mid j = N[0, \sigma_{vj}^2]$; $u_{it} \mid j = N^+[0, \sigma_{uj}^2]$.

For an observation from class j, the model is characterized by the conditional density:

$$P(i,t \mid j) = f(y_{it} \mid x_{it}, \beta_j, \sigma_j, \lambda_j) = \frac{\Phi(\lambda_j \varepsilon_{it|j} / \sigma_j)}{\Phi(0)} \frac{1}{\sigma_j} \phi\left(\frac{\varepsilon_{it|j}}{\sigma_j}\right),\; \varepsilon_{it|j} = y_{it} - x_{it}'\beta_j, \tag{3.66}$$

where:

$\sigma_j = [\sigma_{vj}^{\,2} + \sigma_{uj}^{\,2}]^{1/2}$; $\lambda_j = \sigma_{uj} / \sigma_{vj}$; $\Phi(.)$ refers to the standard normal cumulative density function evaluated at the point; and $\phi(.)$ stands for the standard normal probability density function evaluated at the point.

The log likelihood of the model is given by:

[8] For an overview of stochastic frontier models with heteroscedasticity, in either or both of the two error components, please see Kumbhakar and Lovell (2000).

$$\log L = \sum_{i=1}^{N} \log P(i) = \log\left[\sum_{j=1}^{J} F_{i,j} P(i \mid j)\right] = \log\left[\sum_{j=1}^{J} F_{i,j} \prod_{t=1}^{T} P(i,t \mid j)\right], \qquad (3.67)$$

where *P(i)* is the unconditional likelihood for individual *i*, averaged over the classes; $F_{i,j}$ is the prior probability attached by the econometrician to membership in class *j,* parameterized by the multinomial logit form: $F_{i,j} = \exp(\theta_{ij}) / \Sigma_j \exp(\theta_{ij})$, $\theta_{ij} = \theta_j' z_i$, with θ_j denoting a vector of parameters and z_i a vector of latent variables; $P(i \mid j)$ reflects the contribution of individual *i* to the conditional on class *j* likelihood; $P(i,t|j)$ denote the density for observation *i* at time *t* assuming class *j*.

Using Bayes' theorem, the posterior probability of a particular class membership is given by:

$$P(j \mid i) = \frac{P(i,j)}{P(i)} = \frac{\prod_{t=1}^{T_i} P(i,t \mid j)}{\sum_{j=1}^{J} P(i \mid j) F_{ij}} = \frac{P(i \mid j) F_{ij}}{\sum_{j=1}^{J} P(i \mid j) F_{ij}}, \qquad (3.68)$$

where $P(i,t|j)$ denotes the density for observation *i* at time *t* assuming class *j*.

Using this result, it is possible to obtain the index of the group with the highest posterior probability. The log likelihood function presented in (3.67) can be maximized with respect to all parameters using conventional quasi-Newton methods for unconstrained optimization such as BFGS (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; and Shanno, 1970) and DFP (Davidon, 1959 and Fletcher and Powell, 1963).

Another approach is the EM algorithm (Dempster *et al.*). The EM algorithm is employed simply by iterating back and forth among all optimization problems. The choice of the algorithm is a strictly empirical matter. Standard gradient methods are preferred if there are no convergence problems; otherwise, EM algorithm can be employed. The EM algorithm is slower to converge, but it is more stable.

The other approaches to account for heterogeneity are presented in Greene (2005), the true fixed- and true random-effects models. According to Greene (2005), traditional fixed-effects models have a common shortcoming. By interpreting the firm-specific term as inefficiency, any unmeasured time-invariant cross-firm heterogeneity must be assumed away. Furthermore, the inefficiency must be assumed to be time-invariant which is a very unreasonable assumption. Greene (2005) proposes a true fixed-effects formulation:

$$y_{it} = \alpha_i + \beta' x_{it} + \upsilon_{it} - u_{it}. \qquad (3.69)$$

There remain two problems that must be solved. First, the model may involve too many parameters to be estimated. Second, with small T, many fixed-effects estimators of model parameters are inconsistent and subject to a small sample bias. Greene (2005) solves these questions by noticing that in the linear case, the regression model using group mean deviations sweeps out the fixed-effects. Unlike the estimator of the fixed-effect, the slope estimator is consistent, since it is not a function of the fixed-effects. Therefore, the log likelihood is a function of β that is free of the fixed effects. The log likelihood function for the fixed-effects stochastic frontier model is:

$$\log L = \sum_{i=1}^{N}\sum_{t=1}^{T} \log\left[\frac{1}{\Phi(0)}\Phi\left(-\lambda\left(\frac{y_{it}-\alpha_i-\beta' x_{it}}{\sigma}\right)\right)\phi\left(\frac{y_{it}-\alpha_i-\beta' x_{it}}{\sigma}\right)\right], \qquad (3.70)$$

where:

$\lambda = \sigma_u / \sigma_\upsilon ; \sigma = \sqrt{\sigma_\upsilon^2 + \sigma_u^2}$.

Maximization of the unconditional log likelihood function can be accomplished using Newton's method and some well-known results from matrix algebra. Using these results, it is possible to compute directly both the joint maximizers of the log likelihood and the appropriate submatrix of the inverse of the Hessian for estimating asymptotic standard errors.

Regarding the true random-effects model, Greene (2005) justifies it by noting three shortcomings of the traditional random-effects models. First, the assumption that the effects are not correlated with the regressors. Second, inefficiency is assumed to be the same in every period. The third shortcoming is that u_i incorporates both the inefficiency and any time invariant firm-specific heterogeneity.

Greene (2005) uses the following general form of the random parameters stochastic frontier model:

(1) Stochastic Frontier $\quad y_{it} = \alpha_i + \beta_i' x_{it} + v_{it} - u_{it}$

$$v_{it} \sim N[0,\sigma_v^2]$$

(2) Inefficiency Distribution $\quad u_{it} \sim N^+[\mu_i,\sigma_{ui}^2]$

$$\mu_i = \delta_i' z_i \tag{3.71}$$

$$\sigma_{ui} = \sigma_u \times \exp(\gamma_i' h_i)$$

(3) Parameter heterogeneity $\quad (\alpha_i,\beta_i) = \left(\overline{\alpha},\overline{\beta}\right) + \Delta_{\alpha,\beta} q_i + \Gamma_{\alpha,\beta} w_{\alpha,\beta i}$

$$\delta_i = \delta + \Delta_\delta q_i + \Gamma_\delta w_{\delta i}$$

$$\gamma_i = \gamma + \Delta_\gamma q_i + \Gamma_\gamma w_{\gamma i},$$

where:

z_i and h_i are vectors of firm specific characteristics which affect the mean and variance of the inefficiency term, respectively; (α_i,β_i) is allowed to vary randomly with mean vector $\left(\overline{\alpha},\overline{\beta}\right)$ + $\Delta_\beta q_i$; Δ_j $(j = \beta,\delta,\gamma)$ is a matrix of parameters to be estimated; q_i is a set of related variables which enters the distribution of the random parameters; w_{ji} parameterizes random variation and it is the random vector normally distributed with mean vector zero and known diagonal covariance matrix Σ_j; Γ_j is a free, lower triangular matrix, allowing the generation of an unrestricted covariance matrix.

Greene (2005) derives from (3.71) the following true random-effects specification:

$$y_{it} = \alpha + \beta' x_{it} + w_i + \upsilon_i - u_{it}, \tag{3.72}$$

where:

w_i is the firm-specific random-effect and υ_{it} and u_{it} are, respectively, the symmetric and one-sided components specified earlier.

To avoid identification problems, Greene (2005) uses the compound error formulation:

$$y_{it} = \alpha + \beta' x_{it} + w_i + \varepsilon_{it}, \tag{3.73}$$

where $\varepsilon_{it} = \upsilon_i - u_{it}$.

This equation represents an ordinary random-effects model, albeit one in which the time-varying component has an asymmetric distribution.

The distribution of the compound disturbance is given by the expression:

$$f\left(\varepsilon_{it}\right) = \frac{\Phi\left(-\varepsilon_{it}\lambda / \sigma\right)}{\Phi\left(0\right)} \frac{1}{\sigma} \phi\left(\frac{\varepsilon_{it}}{\sigma}\right). \tag{3.74}$$

Thus, the model specified in (3.65) - (3.66) is actually a random-effects model in which the time-varying component does not have a normal distribution, though w_i may be assumed to follow this distribution. In order to estimate this random-effects model by maximum likelihood, it is necessary to integrate the common term out of the likelihood function. There is no closed form for the density of the compound disturbance in this model. However, the integration can be done either by quadrature or simulation techniques, although the former technique is impractical for models with more than one random parameter.

Note that it is possible to re-write the model in (3.73) as a stochastic frontier model with a firm-specific random constant term:

$$y_{it} = \left(\alpha + w_i\right) + \beta' x_{it} + \upsilon_i - u_{it}. \tag{3.75}$$

This model can be extended to the normal-truncated normal model and to a singled- or doubled-heteroscedastic model with only minor modifications and, therefore, be solved as in Tsionas (2002).

Assuming that, conditioned on the firm-specific w_i, the observations are independent, the conditional log likelihood for the sample is:

$$\log L \left| w_1,...,w_N = \sum_{i=1}^{N}\sum_{t=1}^{T} \log f\left(\Theta_i \left| y_{it}, x_{it}, z_i, h_i, q_i, w_i\right.\right),\right. \tag{3.76}$$

where Θ_i contains all the parameters of the model.

In order to estimate the model parameters, the heterogeneity is integrated out of the log likelihood. The unconditional log likelihood is given by:

$$\log L = \sum_{i=1}^{N}\int_{w_i}\sum_{t=1}^{T} \log f\left(\Theta_i \left| y_{it}, x_{it}, z_i, h_i, q_i, w_i\right.\right) g\left(w_i\right) dw_i, \tag{3.77}$$

where:

$g(w_i)$ is the multivariate density of the random vector w_i.

The unconditional log likelihood function must be maximized with respect to the unknown parameters. However, there is no closed form solution for the integral in the unconditional log likelihood in (3.77). The integral may be satisfactorily approximated by simulation. As long as it is possible to simulate primitive draws from the distribution of w_i, the problem may be solved by maximizing the simulated log likelihood:

$$\log L_S = \sum_{i=1}^{N}\frac{1}{R}\sum_{r=1}^{R}\left[\sum_{t=1}^{T} \log f\left(\Theta_i \left| y_{it}, x_{it}, z_i, h_i, q_i, w_{ir}\right.\right)\right], \tag{3.78}$$

where:

R is the number of replications and w_{ir} is the simulated random.

This function is smooth and twice continuously differentiable in the underlying parameters and can be maximized with conventional techniques.

3.7 - Conclusions

Despite all potential advantages of frontier production models, very few were used to address development accounting problems (e.g., Kumar and Russell, 2002; Henderson and Russell, 2005). Furthermore, the referred studies have important undesirable characteristics such as its deterministic and cross-sectional nature and its assumption of homogeneity across countries. Therefore, in panel data production frontier literature, it was important to look up for stochastic models that account for heterogeneity of countries in different development stages and that are capable of introducing some flexibility in the definition of technology or stochastic noise.

In order to summarize the main advantages and disadvantages of each group of stochastic panel data production frontier models, we characterize them in terms of some desirable features such as flexibility of the frontier, possibility to account for the random nature of the frontier, possibility to account for time-varying inefficiency, possibility to account for latent heterogeneity of production units, requirement of a specific distribution on the efficiency term, flexibility of inefficiency component and requirement of independence between the regressors and inefficiency.

The main results of this analysis are presented in tables 2 and 3. It is possible to conclude that only the three last approaches (random parameters, latent class and true fixed effects and true random effects) account for latent heterogeneity of production units. Nevertheless, all of them present inflexible production frontier and inefficiency component, contrary for example to the semiparametric or Bayesian methods. In chapter 5, we present a finite-mixture model approach, which accounts for heterogeneity, but fails to assure flexibility of all specifications. For this reason, chapter 6 introduces a semiparametric approach, more specifically a penalized spline model both for Classical and Bayesian formulations, which combines the advantages of the three models accounting for heterogeneity across countries with the flexibility of the semiparametric and Bayesian models also presented in table 3.

Table 2 - Summary of Traditional Method's Attributes

	Time-Invariant Technical Inefficiency			***Common Pattern of Time Varying Inefficiency***			***Different Pattern of Time Varying Inefficiency***		
	Random Effects	Fixed Effects	Maximum Likelihood	Random Effects	Fixed Effects	Maximum Likelihood	Random Effects	Fixed Effects	Maximum Likelihood
Works	Schmidt and Sickles (1984)	Schmidt and Sickles (1984)	Pitt and Lee (1981); Kumbhakar (1987); Battese and Coelli (1988).	Lee and Schmidt (1993); Ahn *et al*. (2001); Han et al. (2005)	Lee and Schmidt (1993); Ahn *et al*. (2001); Han et al. (2005)	Kumbhakar (1990) Battese and Coelli (1992)	Cornwell *et al.* (1990)	Cornwell *et al.* (1990)	Kumbhakar (1991); Kumbhakar and Hjalmarsson (1993, 1995); Battese and Coelli (1995)
Flexibility of frontier	Low	Low	Low	Low	Low	Low	Low	Low	Low
Possibility to account for random nature of frontier	No	No	No	No	No	No	No	No	No
Possibility to account for time varying inefficiency	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Possibility to account for latent heterogeneity of production units	No	No	No	No	No	No	No	No	No
Requires a specific distribution on the efficiency term	No	No	Yes	No	No	Yes	No	No	Yes
Flexibility of inefficiency component	Low	Low	Low	Low	Low	Low	Low	Low	Low
Requires independence between variables and inefficiency	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes

Table 3 - Summary of Non-Traditional Method's Attributes

	Bootstrap[9]	**Bayesian**		**Semiparametric**	**Local Linear Weighted Least Squares**	**Local Maximum Likelihood**	**Latent Class**	**True Fixed and True Random Effects**
Works	Simar and Wilson (1998, 2000)	Koop *et al.* (1999, 2000); Griffin and Steel (2004).	Tsionas (2002)	Park *et al.* (1998); Adams et al. (1999)	Henderson and Ullah (2004, 2005)	Kumbhakar *et al.* (2004)	Greene (2001a, 2001b, 2003) and Orea and Kumbhakar (2004);	Greene (2005).
Flexibility of frontier	It depends on the subjacent method.	It depends on the subjacent method.	Low	High	Very High	High	Low	Low
Possibility to account for random nature of frontier	No	Yes	Yes	No	Yes	Yes	No	No
Possibility to account for time varying inefficiency	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Possibility to account for latent heterogeneity of production units	No	No	Yes	No	No	No	Yes	Yes
Requires a specific distribution on the efficiency term	No	No	Yes	No	No	Yes	Yes	No
Flexibility of inefficiency component	High	High	Low	High	High	High	Low	Low
Requires independence between variables and inefficiency	No	No	Yes	No	No	Yes	Yes	Yes

9 Bootstrap is not an autonomous method and it can be used with parametric and nonparametric production frontier approaches.

4 - Growth Patterns of Labor Productivity

This study relies on two cross-country datasets (one for the overall economy and the other for the agricultural sector) covering 45 countries and the period 1967-92, as indicated in table 4. Although a larger number of countries can be included in each dataset, our analysis is deliberately restricted to those countries for which both economy-wide data and agricultural sector data are available.[10] Please refer to the annexes for a description of data sources used.

Table 4 - Countries List by Alphabetical Order

Argentina	Kenya
Australia	Korea, Republic of
Austria	Madagascar
Canada	Malawi
Chile	Morocco
Colombia	Netherlands
Costa Rica	New Zealand
Denmark	Norway
Dominican Republic	Pakistan
Egypt	Peru
El Salvador	Philippines
Finland	Portugal
France	South Africa
Great Britain	Sri Lanka
Greece	Sweden
Guatemala	Syrian Arab Republic
Honduras	Tunisia
India	Turkey
Indonesia	United States of America
Iran	Uruguay
Israel	Venezuela
Italy	Zimbabwe
Japan	

With the purpose of extracting some global indicators about the aggregate labor productivity growth, we will construct the kernel estimator of its probability density function and proceed to a descriptive analysis of some available data.

[10] As it is common in the convergence literature (e.g., Kumar and Russell, 2002), we exclude the two major oil-producing countries (Iran and Venezuela) from the overall economy dataset. Therefore, the sample is of 45 countries to the overall economy and 43 to the agricultural sector.

Following Kumar and Russell (2002), we use a nonparametric kernel density estimator for measuring the probability density functions of labor productivity for the overall economy and for agriculture. Assuming *n* independent observations $x_1, x_2, \ldots, x_n$ from a random variable *X*, the kernel density estimator of the density value *f(x)* at point *x*, $\hat{f}(x)$, is defined as:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x_i - x}{h}\right), \tag{4.1}$$

where:

k(.) denotes a Gaussian kernel function and *h* is the optimal bandwidth (for details, please see Pagan and Ullah, 1999).

The choice of the optimal bandwidth for a kernel density estimate is typically calculated on the basis of the minimization of the mean integrated squared error function:

$$MISE(\hat{f}) = \int E\left[\hat{f}(x) - f(x)\right]^2 dx. \tag{4.2}$$

Under the asymptotic conditions $h \to 0, nh \to \infty$:

$$MISE(\hat{f}) \underset{asymp.}{\approx} \frac{1}{nh}\|k\|_2^2 + \frac{h^4}{4}\left[\mu_2(k)\right]^2 \|f''\|_2^2, \tag{4.3}$$

where:

- $\|k\|_2^2$ and $\left[\mu_2(k)\right]^2$ are parameters depending on the kernel function *k(.)*,

- $\|f''\|_2^2$ is an unknown term, denoting the second derivative of the unknown density *f*.

Minimizing (4.3) with respect to *h*, we obtain the following optimal bandwidth:

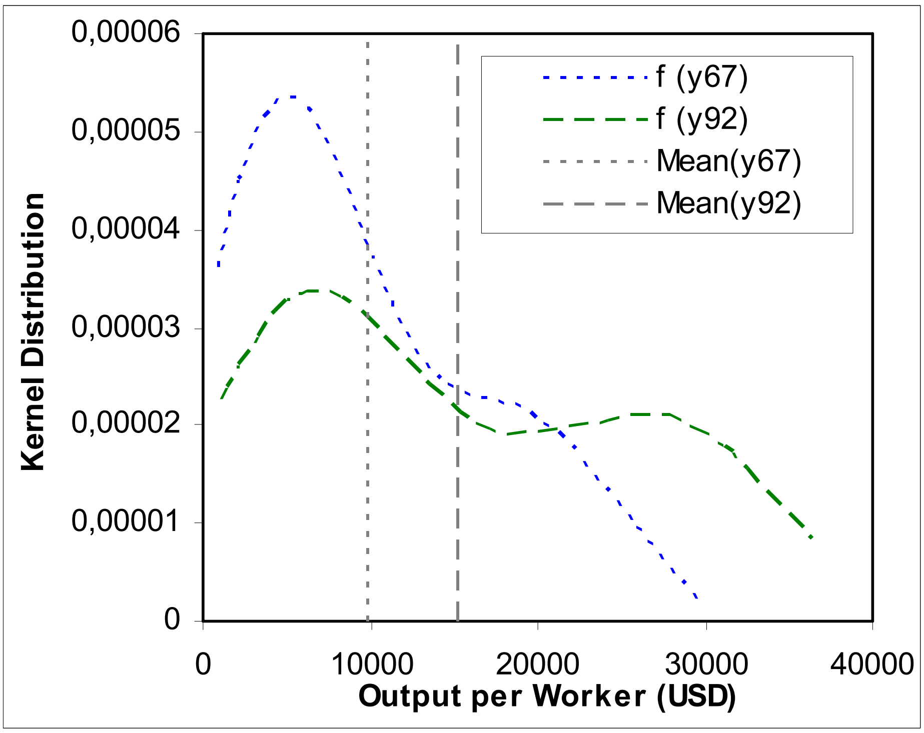
$$h_{opt} = \left\{ \frac{\|k\|_2^2}{\|f''\|_2^2 \left[\mu_2(k)\right]^2 n} \right\}^{1/5}. \tag{4.4}$$

Using the method of Silverman (1986) and assuming a normal distribution $N(\mu, \sigma^2)$ for *f*, the optimal bandwidth for a Gaussian kernel is:

$$\hat{h}_{opt} = 1.06\hat{\sigma}n^{-1/5}. \tag{4.5}$$

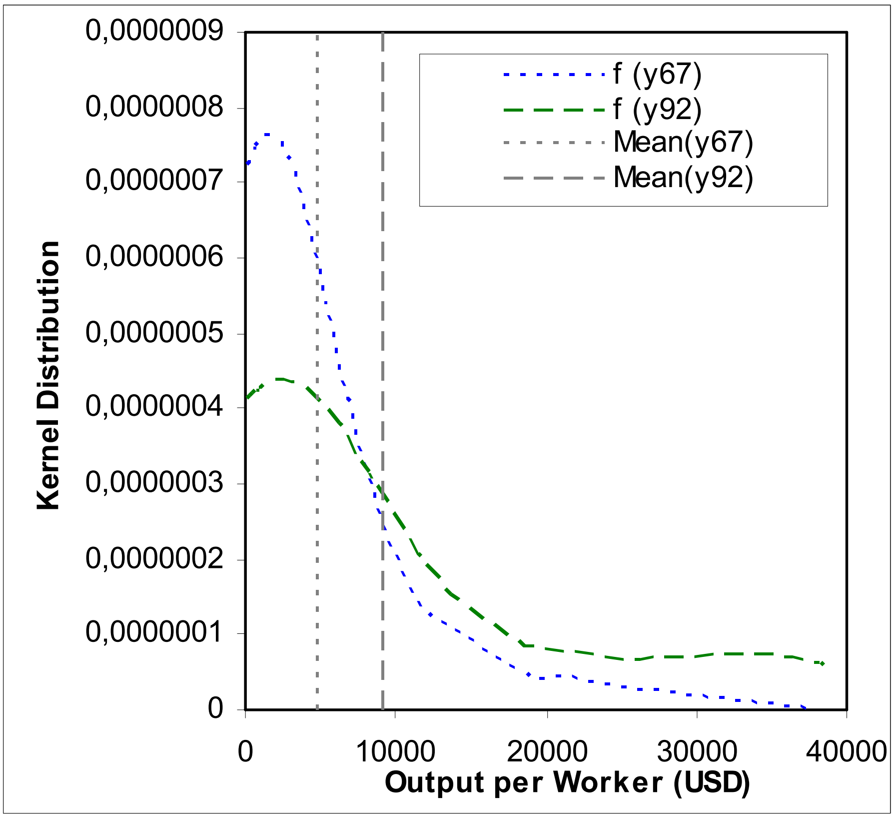
The kernel distributions of labor productivity are presented in figures 1 and 2 for the overall economy and for agriculture, respectively[11]. We focus on mean-preserving distributions; i.e., departures from the productivity mean.

Figure 1 - Gaussian Kernel of Labor Productivity for the Economy



[11] All kernel calculations were carried out using GAUSS (Aptech Systems, Inc.,1999).

Figure 2 - Gaussian Kernel of Labor Productivity for Agriculture



For the economy as a whole, our empirical results are similar to those of Kumar and Russell (2002) and Quah (1996a, 1997). In particular, the labor productivity distribution evolves from a unimodal to a bimodal distribution with a higher mean.

The results for the agricultural sector are substantially different. There is a probability shift from the lower tail toward the rest of the distribution. The increase of density mass for the middle-income countries contradicts the idea of the world becoming polarized into high-productivity and low-productivity (or rich and poor) countries. Consequently, it appears that the agricultural sector has not contributed to the bipolarization phenomenon that has been observed at the aggregate economy level by studies such as Quah (1996a, 1997).

The descriptive analysis of some global indicators related to labor productivity growth is presented in table 5.

Table 5 - Global Indicators for the Sample

	Mean	Std. Deviation	Minimum	Maximum
% change of output per worker agriculture	,89747	,886375	-,331	3,953
% change of output per worker economy	,6483	,77966	-,34	4,23
% variation in weight of agriculture	-,3266	,18455	-,67	,07

Results indicate that, on average, the output per worker increased more in agriculture than in the economy for the period 1967-1992, with a bigger spread. Furthermore, it is evident a global reduction in the weight of agriculture in employment. These global indicators are incapable of determining if countries behave in the same manner. To perform such a task, it is necessary to group nations according to an objective and defendable criterion measuring the development stage and compare outcomes for each set.

The World Bank uses per capita gross national income as the main criterion of classifying countries. The thresholds are defined according to a stable relationship between a summary measure of wellbeing such as poverty incidence and infant mortality on the one hand and economic variables including per capita gross national income[12] on the other. Those limits are updated every year to incorporate the effect of international inflation, in order to assure their constancy in real terms over time. The available thresholds in USD for the last 6 years in the sample are presented in the table 6.

Table 6 - Countries Classification Adopted by the World Bank

	1987	1988	1989	1990	1991	1992
Low income (L)	< 480	< 545	< 580	< 610	< 635	< 675
Lower middle income (LM)	481-1,940	546-2,200	581-2,335	611-2,465	636-2,555	676-2,695
Upper middle income (UM)	1,941-6,000	2,201-6,000	2,336-6,000	2,466-7,620	2,556-7,910	2,696-8,355
High income (H)	> 6,000	> 6,000	> 6,000	> 7,620	> 7,910	> 8,355

[12] To obtain real values with a small impact of exchange rate fluctuations in the cross-country comparison of national incomes, the World Bank uses Atlas conversion factor for any year; i.e., the average of a country's exchange rate for that year and its exchange rates for the two preceding years, adjusted for the differences between the rate of inflation in the country and the G-5 countries (France, Germany, Japan, the United Kingdom, and the United States).

We adopt the criterion used by the World Bank for the last year of our sample, leading to sample classification presented in table 7.

Table 7 - Application of the World Bank's Criterion to the Sample

Developed countries	Developing countries[13] (29)		
Rich (16)	**Upper Middle (6)**	**Lower Middle (13)**	**Poor (10)**
Australia	Argentina	Chile	Egypt
Austria	Greece	Colombia	Honduras
Canada	Rep. Korea	Costa Rica	India
Denmark	South Africa	Dominican Rep.	Indonesia
Finland	Uruguay	El Salvador	Kenya
France	Venezuela	Guatemala	Madagascar
Israel		Iran	Malawi
Italy		Morocco	Pakistan
Japan		Peru	Sri Lanka
Netherlands		Philippines	Zimbabwe
New Zealand		Syria	
Norway		Tunisia	
Portugal		Turkey	
Sweden			
UK			
USA			

A descriptive analysis of data according to the classification is presented in table 8.

Table 8 - Descriptive Analysis of Growth Indicators

		N	Mean	Std. Deviation	Minimum	Maximum
% change of output per worker in agriculture	Rich	16	1,48531	,734183	,009	2,800
	Upper Middle	6	1,40550	1,346254	,471	3,953
	Lower Middle	13	,53492	,325233	-,093	,997
	Poor	10	,12340	,390770	-,331	,931
	Total	45	,89747	,886375	-,331	3,953
% change of output per worker in economy	Rich	16	,6719	,46763	,05	1,68
	Upper Middle	5	1,1956	1,73630	,17	4,23
	Lower Middle	12	,4334	,46987	-,34	1,26
	Poor	10	,5949	,79398	-,34	2,49
	Total	43	,6483	,77966	-,34	4,23
% variation in weight of agriculture	Rich	16	-,4436	,16198	-,65	-,12
	Upper Middle	5	-,3993	,19184	-,67	-,20
	Lower Middle	12	-,2971	,09658	-,45	-,14
	Poor	10	-,1385	,14051	-,42	,07
	Total	43	-,3266	,18455	-,67	,07

[13] The term "developing economies" is used in World Bank reports to denote the set of non-rich economies. Therefore, this concept includes upper middle, lower middle and poor countries.

Output per capita growth in agriculture varies directlty with countries' income. Furthermore, differences among groups are very meaningful. The rich and upper middle countries exhibit strong growth rates, contrasting with moderate rates for the other countries. In the economy, growth rates are stronger for the upper middle nations than the rich ones and for the poor than the lower middle income countries, the group with the worst performance of all. Growth rates are very homogeneous among groups as the ANOVA test presented in table 9 proves.

Table 9 - ANOVA Test: Analysis of Variance for the Classification in 4 Groups

		Sum of Squares	Df	Mean Square	F	Sig.
% change output per worker agriculture	Between Groups	14,778	3	4,926	10,205	,000
	Within Groups	19,791	41	,483		
	Total	34,569	44			
% change output per worker economy	Between Groups	2,089	3	,696	1,159	,338
	Within Groups	23,441	39	,601		
	Total	25,531	42			
% variation in weight of agricult.	Between Groups	,609	3	,203	9,649	,000
	Within Groups	,821	39	,021		
	Total	1,430	42			

It is not possible to reject the hypothesis of the mean being equal among sets. This evidence contrasts with the heterogeneity of the variation in weight of agriculture in employment. The reduction is stronger for the rich countries, declining as the income of the reference group diminishes.

Tables 10 and 11 refer to the case of considering only two kinds of countries (developed and developing).

Table 10 - Descriptive Analysis of Growth Indicators for a Classification of Countries in 2 Groups

		N	Mean	Std. Deviation	Minimum	Maximum
% change output per worker agriculture	Developed	16	1,48531	,734183	,009	2,800
	Developing	29	,57314	,799648	-,331	3,953
	Total	45	,89747	,886375	,132133	3,953
% change output per worker economy	Developed	16	,6719	,46763	,05	1,68
	Developing	27	,6344	,92479	-,34	4,23
	Total	43	,6483	,77966	-,34	4,23
% variation in weight of agriculture	Developed	16	-,4436	,16198	-,65	-,12
	Developing	27	-,2573	,16272	-,67	,07
	Total	43	-,3266	,18455	-,67	,07

Table 11 - ANOVA Test: Analysis of Variance for the Classification in 2 Groups

		Sum of Squares	df	Mean Square	F	Sig.
% change output per worker agriculture	Between Groups	8,579	1	8,579	14,195	,000
	Within Groups	25,990	43	,604		
	Total	34,569	44			
% change output per worker economy	Between Groups	,014	1	,014	,023	,881
	Within Groups	25,516	41	,622		
	Total	25,531	42			
% variation in weight of agriculture	Between Groups	,349	1	,349	13,207	,001
	Within Groups	1,082	41	,026		
	Total	1,430	42			

Conclusions are the same: very high rates for the developed in agriculture, similar rates in the economy and a stronger reduction of agriculture weight of agriculture in employment.

In order to identify the different roles of factor accumulation, technical change and catch-up among countries' groups in shaping labor productivity distributions of the agricultural sector and the overall economy, first it is necessary to perform its decomposition. This task is sensitive to the approach used. In chapter 5 we apply a parametric method and in chapter 6 a semiparametric approach, both accounting for the heterogeneity of countries.

5 - A Stochastic Frontier Finite Mixture Approach for the Decomposition of Labor Productivity Growth

5.1 - The Model

So far, parametric and non-parametric production frontier studies of international productivity growth such as Kumar and Russell (2002) and Martin and Mitra (2001) have normally assumed a common frontier for all countries. A simple panel data stochastic production frontier can be expressed as follows:

$$y_{it} = \beta' x_{it} + v_{it} - u_{it} \qquad i = 1,...,N\,;\ t = 1,...,T\,; \qquad (5.1)$$

where:

'i' indexes countries and *'t'* indexes time periods; y_{it} is the log of the production level in year *t* for the *i*-th country; x_{it} is a 1 × K vector of the log of inputs in year *t* for the *i*-th country; β is a 1 × K vector of coefficients; v_{it} is the measurement error, and u_{it} refers to the inefficiency component.

In this framework, heterogeneity in the distribution of y_{it} is assumed to impact the density function in the simple form of a random effect.

However, the underlying belief that the production technology is common to all countries can be challenged, particularly for samples including a large and heterogeneous set of countries. If this assumption is not valid, technological differences may be labeled as inefficiency and the decomposition of output per worker is imprecisely determined.

One method to solve this problem is based on a two-stage approach: first, countries are classified into several classes, according, for instance, to a cluster analysis applied to the dependent variable; and second, a production frontier is estimated separately for each class (e.g., Kolari and Zardkoohi, 1995; Mester, 1997). However, such a procedure has the disadvantage of estimating the production frontier of a particular class without using information regarding the other classes. This problem may be overcome by using the Stochastic Frontier Finite Mixture Model (SFFMM) approach, which was proposed by

Heckman and Singer (1984), also drawing on important recent developments suggested by Greene (2001a, 2001b). The SFFMM allows for the simultaneous estimation of both the class membership probabilities and the parameters of the mixed frontier functions.

Following this approach, we accommodate the unobserved heterogeneity with a model where the density function is specific to each endogenously determined country class:

$$y_{it} \,|\, j = \beta_j' \, x_{it} + v_{it} \,|\, j - u_{it} \,|\, j \quad ; \quad i = 1,\ldots,N \; ; \; t = 1,\ldots,T \; ; \; j = 1,\ldots, M, \qquad (5.2)$$

where:

j indicates class number.

The observations of the sample arise from M unobserved classes in unknown proportions, $p_1, p_2, \ldots, p_M$, such that:

$$\left(0 \le p_j \le 1\right) \text{ and } \sum_{j=1}^{M} p_j = 1. \qquad (5.3)$$

To ensure conditions in (5.3), a logit parameterization is used as follows:

$$p_j = \frac{\exp\left(c_j\right)}{\sum_{\gamma=1}^{M} \exp\left(c_\gamma\right)} \qquad j = 1,\ldots,M \; ; \; c_M = 0, \qquad (5.4)$$

where:

c_j refers to lower level parameters.

Within each class, the basic form of a half normal specification in (1) applies:

$$v_{it} \,|\, j = N[0, \sigma_{vj}^2] \quad ; \quad u_{it} \,|\, j = N^+[0, \sigma_{uj}^2]. \qquad (5.5)$$

After some algebra work, the distribution of the dependent variable conditional on the j class has the form:

$$f(y_{it} \mid \mathbf{x}_{it}, j) = \frac{\Phi(\lambda_j \varepsilon_{it|j} / \sigma_j)}{\Phi(0)} \frac{1}{\sigma_j} \phi\left(\frac{\varepsilon_{it|j}}{\sigma_j}\right), \tag{5.6}$$

where:

$\varepsilon_{it|j} = y_{it} | j - \beta'_j x_{it}$; $\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$; $\lambda_j = \sigma_{uj} / \sigma_{vj}$; $\Phi(.)$ refers to the standard normal cumulative distribution function ; $\phi(.)$ designates the standard normal probability density function.

The class from which from which a particular observation arises is unknown a priori. Assuming that the T events are independent within each class, the contribution of country i to the likelihood function is:

$$\sum_{j=1}^{M} p_j \left[\prod_{t=1}^{T} f\left(y_{it} \middle| x_{it}, j \right) \right]. \tag{5.7}$$

Thus, the log likelihood function for the sample is given by:

$$\ln L(\alpha) = \sum_{i=1}^{N} \ln \left\{ \sum_{j=1}^{M} p_j \left[\prod_{t=1}^{T} f\left(y_{it} \middle| x_{it}, j \right) \right] \right\}, \tag{5.8}$$

where:

$$\alpha = \left[(\beta_{11}, ..., \beta_{1M}), ..., (\beta_{K1}, ..., \beta_{KM}), (p_1, ..., p_M), (\sigma_{u1}, ..., \sigma_{uM}), (\sigma_{v1}, ..., \sigma_{vM}) \right].$$

The log likelihood can be maximized with respect to α using conventional gradient methods. Once estimates of α are obtained, we can also compute the posterior estimate of the probability of a particular class membership by using the Bayes theorem:

$$\mathrm{P}(j \mid i) = \frac{p_j \times \prod_{t=1}^{T} f\left(y_{it} \middle| x_{it}, j \right)}{\sum_{j=1}^{M} p_j \left[\prod_{t=1}^{T} f\left(y_{it} \middle| x_{it}, j \right) \right]}. \tag{5.9}$$

Using (5.9), we can identify the index of the group with the highest posterior probability and therefore determine which class generates each observation. Furthermore, the posterior probability can be used in the computation of the efficiency estimates. Following Greene (2001a, 2001b), the individual efficiencies are computed as:

$$\ln EF_{it} = \sum_{j=1}^{M} P(j \mid i) \ln EF_{it} \left| j, \right. \qquad (5.10)$$

where:

$EF_{it}|j$ is the estimator of the efficiency of the *i*-th country, calculated applying the Jondrow *et al.* (1982) approach to the production frontier of class *j*.[14]

There remains an unsolved question: how to determine the number of classes, *M*? In fact, *M* is not an estimable parameter and, therefore, it cannot be obtained by maximization of the likelihood function. A model with (*M-1*) classes is nested within a model with *M* classes by imposing restrictions on the parameters. Testing 'up' from (*M-1*) to *M* is not a valid procedure because if there are *M* classes, then estimates based only on (*M-1*) are inconsistent. Conversely, testing 'down' is an acceptable method, as suggested by Greene (2002). Therefore, we would only need to pick a large *M** and test down to the "true" *M* based on likelihood ratio tests. Unfortunately, the latent class model is a little volatile and the estimation of models with larger number of classes or/and restrictions may not be possible, because the estimated variance matrix of estimates can be singular. Furthermore, according to McLachlan (1987) and Feng and McCulloch (1996), Pearson fit, Kolmogorov-Smirnov and likelihood ratio tests do not have a nice distribution for this sort of problems. Thus, some authors (see, for example, Fraley and Raftery, 1998 and Roeder *et al.*, 1999) propose the use of information criteria such as the Akaike Information Criterion (AIC) and the Schwarz Bayesian Information Criterion (SBIC). Both AIC and SBIC take the following form:

$$MSC(h) = -2 \ln \max L(h) + a(n)m(h), \qquad (5.11)$$

where:

[14] In models with a single frontier, it is a standard procedure the application of the Jondrow *et al.* (1982) estimator of individual inefficiencies $E[u_{it}|v_{it}-u_{it}]$ to calculate efficiency $E[exp(-u_{it})|v_{it}-u_{it}]$.

MSC(h) is the value of the criterion for the h-th model (the lower the score, the better); $L(h)$ is the likelihood for the h-th model; $m(h)$ is the number of parameters used in the h-th model; $a(n) = 2$ for AIC and $a(n) = \ln n$ for SBIC; and $h = 1, 2, \ldots, H$ indexes the alternative models.

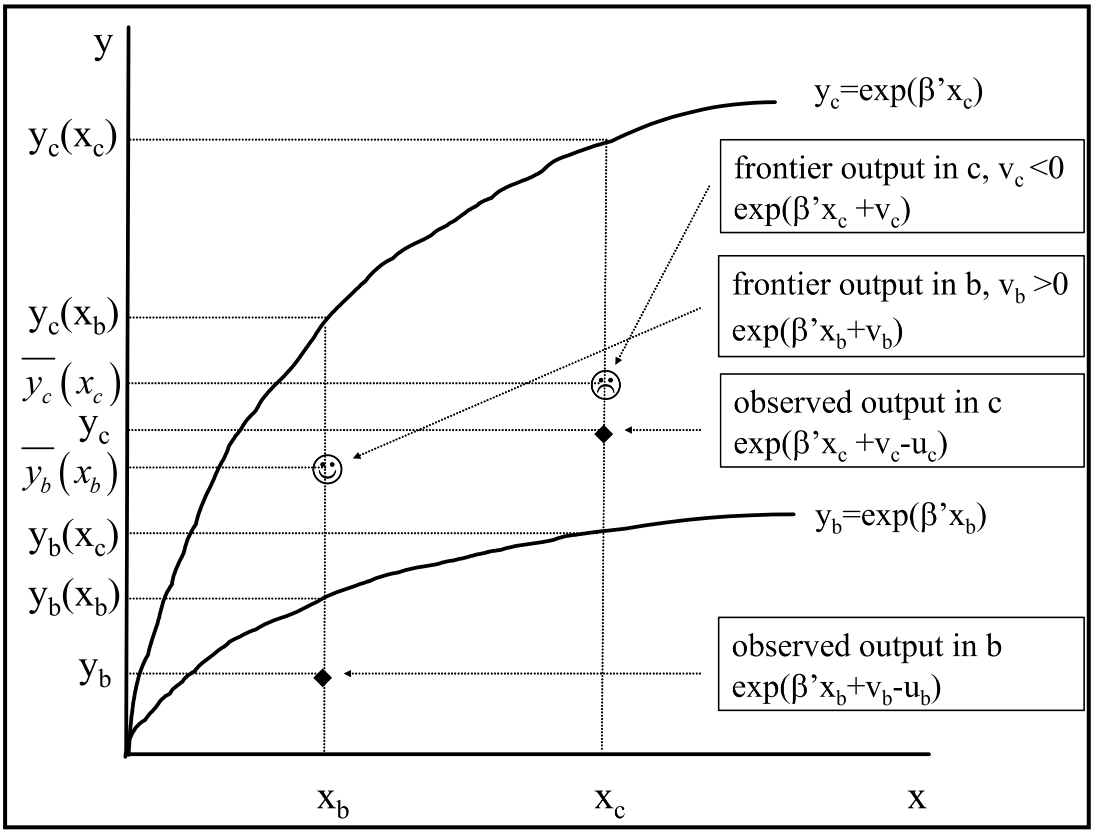
Most statisticians who are involved with the theory and application of model selection criteria prefer SBIC, since it penalizes models with more components heavier than AIC. Moreover, Leroux (1992) concludes that SBIC does not underestimate the number of classes; and Roeder and Wasserman (1997) argue that this method is consistent. On the other hand, Berger and Pericchi (1998) argue that information criteria are valid only for "nice" problems, that is, large sample sizes models with regular asymptotics and models for which the likelihood is not concentrated on the boundary of the parameter space. Furthermore, authors such as Zhang (1997) and Greene (2002) have suggested that an evaluation of the estimation results should play a role in choosing between alternative model specifications. In particular, Greene (2002) eliminates models with parameter estimates not significant for at least one of the classes. Zhang (1997) recommends that a simpler specification that sufficiently approximates the true model might be preferred to a more complex specification, even if information criteria indicators point towards adopting the latter. We adopt the approach of combining SBIC with an evaluation of the estimation results.

5.2 - Decomposition of Labor Productivity Growth

We define a CRS reference technology with one aggregate output, Y, and a K-dimensional vector of inputs, X. The CRS hypothesis allows us to transform the dependent variable in labor productivity, y, and the vector X into the $(K-1)$-dimensional vector of inputs per worker, x. For the economy as a whole, K=2 and X=(labor, capital); and for agriculture, K=3 and X=(labor, land, capital).

Figure 3 illustrates the decomposition of output per worker growth, assuming an aggregate input per labor x. Let b and c stand for the base period and the current period, respectively. For simplicity in the analysis, we suppress the subscript i, and consider only one country.

Figure 3 - Illustration of Labor Decomposition



This decomposition is similar to the one presented by Kumar and Russell (2002) and Henderson and Russell (2005) described in equations (2.6) - (2.10), with the important difference that stochastic shocks are accounted for.

In period *b*, x_b units of input per worker are used to produce y_b units of output per worker. However, the country faces a positive shock v_b in this period and, in reality, it could produce $\overline{y_b}(x_b)$. Therefore, efficiency in period *b* is measured as:

$$Eff_b = \frac{y_b}{\overline{y_b}(x_b)} = \frac{y_b}{\exp(\beta' x_b + v_b)} = \frac{y_b}{\exp(\beta' x_b) \cdot \exp(v_b)} = \frac{y_b}{y_b(x_b) \cdot \exp(v_b)}. \tag{5.12}$$

Thus, labor productivity in period *b* can be expressed as:

$$y_b = Eff_b \cdot \exp(v_b) \cdot y_b\left(x_b\right). \tag{5.13}$$

Mutatis mutandis, labor productivity in period *c* is given by:

$$y_c = Eff_c \cdot \exp(v_c) \cdot y_c\left(x_c\right). \tag{5.14}$$

Dividing (5.14) by (5.13), we obtain labor productivity growth:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\exp(v_c)}{\exp(v_b)} \cdot \frac{y_c\left(x_c\right)}{y_b\left(x_b\right)}. \tag{5.15}$$

Multiplying the numerator and the denominator of equation (5.15) by $y_c\left(x_b\right)$, labor productivity growth can be rewritten as:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\exp(v_c)}{\exp(v_b)} \cdot \frac{y_c\left(x_b\right)}{y_b\left(x_b\right)} \cdot \frac{y_c\left(x_c\right)}{y_c\left(x_b\right)}. \tag{5.16}$$

The ratio Eff_c to Eff_b is the efficiency change or technological catch-up between the current period and the base period. The second component on the right hand side of (5.16), $\frac{\exp(v_c)}{\exp(v_b)}$, represents the stochastic shocks effect. The ratio of $y_c\left(x_b\right)$ to $y_b\left(x_b\right)$ captures the shift in the "deterministic" frontier caused by technological change, since input quantity per worker does not change. The last term on the right hand side captures the effect of factor accumulation, since it measures the output per worker change along the "deterministic" frontier in period *c*.

Alternatively, equation (5.15) could be multiplied and divided by $y_b\left(x_c\right)$ and a different, but valid, decomposition would be obtained:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\exp(v_c)}{\exp(v_b)} \cdot \frac{y_c\left(x_c\right)}{y_b\left(x_c\right)} \cdot \frac{y_b\left(x_c\right)}{y_b\left(x_b\right)}. \tag{5.17}$$

As said in chapter 2, this means that labor productivity decomposition is path dependent[15], forcing the use of the geometric average of equations (5.16) and (5.17):

$$\frac{y_c}{y_b}=\frac{Eff_c}{Eff_b}\cdot\frac{\exp(v_c)}{\exp(v_b)}\cdot\left[\frac{y_c\left(x_b\right)}{y_b\left(x_b\right)}\cdot\frac{y_c\left(x_c\right)}{y_b\left(x_c\right)}\right]^{1/2}\cdot\left[\frac{y_c\left(x_c\right)}{y_c\left(x_b\right)}\cdot\frac{y_b\left(x_c\right)}{y_b\left(x_b\right)}\right]^{1/2}, \tag{5.18}$$

where:

$\left[\frac{y_c\left(x_b\right)}{y_b\left(x_b\right)}\cdot\frac{y_c\left(x_c\right)}{y_b\left(x_c\right)}\right]^{1/2}$ represents technological change and $\left[\frac{y_c\left(x_c\right)}{y_c\left(x_b\right)}\cdot\frac{y_b\left(x_c\right)}{y_b\left(x_b\right)}\right]^{1/2}$ indicates factor accumulation and all the other terms are defined as before.

Considering the combined effect of efficiency variation with technological change, we obtain the Malmquist TFP index (e.g., Grosskopf, 1993 and Färe *et al.*, 1994a, 1994b).

In the stochastic finite mixture model, there is not a unique frontier for the entire sample, but one frontier for each class. Furthermore, one observation does not belong to a single class, it has a probability of class membership. Thus, the decomposition of labor productivity in equation (5.18) must be adjusted to this framework. Following a similar procedure used in the computation of individual inefficiency, the potential output per worker of each country in each year is determined by:

$$y_{it}\left(x_{it}\right)=\sum_{j=1}^{M}P\left(j\ |i\right)y_{it}\left(x_{it}\right)\big|j. \tag{5.19}$$

Using (5.18) and (5.19), labor productivity is decomposed as follows:

[15] In the presence of constant returns to scale and disembodied Hicks neutral technical change, equations (16) and (17) would be identical (e.g., Grosskopf, 1993).

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\exp(v_c)}{\exp(v_b)} \cdot \left[\frac{\sum_{j=1}^{M} P(j \mid i) y_c(x_b) | j}{\sum_{j=1}^{M} P(j \mid i) y_b(x_b) | j} \cdot \frac{\sum_{j=1}^{M} P(j \mid i) y_c(x_c) | j}{\sum_{j=1}^{M} P(j \mid i) y_b(x_c) | j} \right]^{1/2} \cdot$$

$$\cdot \left[\frac{\sum_{j=1}^{M} P(j \mid i) y_c(x_c) | j}{\sum_{j=1}^{M} P(j \mid i) y_c(x_b) | j} \cdot \frac{\sum_{j=1}^{M} P(j \mid i) y_b(x_c) | j}{\sum_{j=1}^{M} P(j \mid i) y_b(x_b) | j} \right]^{1/2} . \qquad (5.20)$$

Using the components of the labor productivity change decomposition given in equation (5.20), it is possible to obtain the corresponding counterfactual distributions. In other words, it is possible to analyze how the distribution of labor productivity would change through time under the influence of a particular effect or a combination of the decomposition effects (catch-up, technical change and factor accumulation). Counterfactual distributions are rather more informative than summary measures of the decomposition effects, such as the mean or the variance (Quah, 1993, 1996a, 1996b, 1997).

5.3 - Frontier Estimates and the Determinants of Labor Productivity Growth Across Countries

5.3.1 - Economy as a Whole

We now turn to investigating which factors are mainly responsible for labor productivity distribution changes. This requires estimating the SFFMM, performing the decomposition analysis described in the previous section and finally using the results for determining the counterfactual distributions of labor productivity. We use a translog specification (Christensen *et al.*, 1971) for the production frontier model. This flexible functional form allows the elasticity of substitution to vary with the type of inputs and the returns to scale and output elasticity to vary with the size of the inputs. The production frontier model (ignoring the *j*-class subscript, for notational ease) can be written as:

$$\ln y_{it} = \beta_0 + \beta_1 \ln k_{it} + \beta_2 \left(\ln k_{it}\right)^2 + v_{it} - u_{it}, \quad (5.21)$$

where:

y_{it} refers to output per worker in year *t* for the *i*-th country; k_{it} designates capital per worker in year *t* for the *i*-th country; β's label coefficients; v_{it} is the measurement error and u_{it} refers to the inefficiency component.

The production frontier (5.21) is estimated separately for the time periods 1967-1979 and 1980-1992[16]. This procedure overcomes the estimation problems when a time trend is included in the specification to capture the technological change. The utilization of relatively large periods is explained by the need of using richer panels for estimating these models.

We start by estimating our model with a large number of classes. As discussed in section 2, the SBIC indicator can be used to help choosing the most appropriate number of classes in this type of models. Table 12 reports the SBIC scores for the 1, 2 and 3 class models. The 4-class model for the economy is over-specified since convergence is not attained.

Table 12 - Score for Schwarz Bayesian Information Criterion (SBIC)

		Number of classes			
		1	**2**	**3**	**4**
Economy as a Whole	1967-1979	-18,749	-349,442	-541,640	-
	1980-1992	-21,104	-400,116	-567,284	-
Agriculture	1967-1979	1040,398	491,191	-	-
	1980-1992	736,106	363,016	-	-

The score values suggest the use of a 3-class model for the overall economy. However, a judgment about the estimation results of each model is also advisable, as discussed in

[16] All estimation results were obtained using LIMDEP (Econometric Software, Inc., 2003).

section 2. The estimation results for the 3-class model in both periods are presented in table 13.

Table 13 - Three Class Model Estimation Results for the Economy as a Whole

a) 1967-1979

	Model parameters for latent class 1			**Model parameters for latent class 2**			**Model parameters for latent class 3**		
Variable	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***
Constant	-0,0213	271298	1,0000	1,4252	0,1289	0,0000	2,3218	0,4848	0,0000
$\ln k_{it}$	1,5186	0,0421	0,0000	1,1556	0,0282	0,0000	0,7225	0,1171	0,0000
$(\ln k_{it})^2$	-0,0547	0,0022	0,0000	-0,0350	0,0016	0,0000	0,0002	0,0069	0,9721
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,1146	0,0025	0,0000	0,1736	0,0094	0,0000	0,2204	0,0097	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	0,0000	2965820	1,0000	1,2826	0,2439	0,0000	46,7975	173,865	0,7878
Prior Probabilities for Class Membership	0,3010	0,0758	0,0001	0,5051	0,0913	0,0000	0,1939	0,0975	0,0466

b) 1980-1992

	Model parameters for latent class 1			**Model parameters for latent class 2**			**Model parameters for latent class 3**		
Variable	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***
Constant	6,5086	0,2879	0,0000	2,7048	0,2778	0,0000	-0,9588	0,3120	0,0021
$\ln k_{it}$	0,0205	0,0671	0,7605	0,7847	0,0650	0,0000	1,7624	0,0655	0,0000
$(\ln k_{it})^2$	0,0290	0,0037	0,0000	-0,0137	0,0039	0,0004	-0,0680	0,0033	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,1842	0,0164	0,0000	0,1579	0,0705	0,0251	0,1577	0,0126	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	1,3392	0,4973	0,0071	0,7111	1,7676	0,6875	3,1117	0,9559	0,0011
Prior Probabilities for Class Membership	0,4189	0,0756	0,0000	0,1628	0,0563	0,0038	0,4183	0,0756	0,0000

At least one of the lambdas is not statistically significant and some of the estimation results are poor for this class. Following a testing down procedure, empirical results are generated for the 2-class model (table 14).

Table 14 - Two Class Model Estimation Results for the Economy as a Whole

a) 1967-1979

	Model parameters for latent class 1			**Model parameters for latent class 2**		
Variable	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***
Constant	0,8585	0,2222	0,0001	-1,3149	0,1741	0,0000
$\ln k_{it}$	1,3241	0,0483	0,0000	1,7275	0,0392	0,0000
$(\ln k_{it})^2$	-0,0431	0,0025	0,0000	-0,0640	0,0022	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,2089	0,0087	0,0000	0,2768	0,0050	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	1,4707	0,2060	0,0000	5,4008	0,7528	0,0000
Prior Probabilities for Class Membership	0,4652	0,0763	0,0000	0,5348	0,0763	0,0000

b) 1980-1992

	Model parameters for latent class 1			**Model parameters for latent class 2**		
Variable	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***
Constant	4,7043	0,0903	0,0000	1,1935	0,1693	0,0000
$\ln k_{it}$	0,5064	0,0197	0,0000	1,1765	0,0372	0,0000
$(\ln k_{it})^2$	-0,0005	0,0011	0,6766	-0,0338	0,0020	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,1898	0,0151	0,0000	0,2936	0,0089	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	1,4959	0,3978	0,0002	7,8311	2,0985	0,0002
Prior Probabilities for Class Membership	0,5375	0,0765	0,0000	0,4625	0,0765	0,0000

The estimation results are very satisfactory indicating the assumption of a common production frontier for all countries does not seem appropriate. Consequently, we adopt the 2-class specification. The grouping of countries generated by the adopted model is reported in table 15.

Table 15 - Countries Classification According to the Stochastic Frontier Finite Mixture Model for the Economy as a Whole

1967- 1979				1980-1992			
Countries	Class	Countries	Class	Countries	Class	Countries	Class
Argentina	1	Austria	2	Australia	1	Argentina	2
Australia	1	Chile	2	Canada	1	Austria	2
Canada	1	Costa Rica	2	Chile	1	Costa Rica	2
Colombia	1	Dominican Rep.	2	Colombia	1	Denmark	2
Denmark	1	Egypt	2	Dominican Rep.	1	El Salvador	2
France	1	El Salvador	2	Egypt	1	Finland	2
Guatemala	1	Finland	2	France	1	Greece	2
Indonesia	1	Greece	2	Guatemala	1	Honduras	2
Israel	1	Honduras	2	Índia	1	Indonesia	2
Italy	1	India	2	Israel	1	Japan	2
Madagascar	1	Japan	2	Italy	1	Kenya	2
Netherlands	1	Kenya	2	Madagascar	1	Korea	2
New Zealand	1	Korea	2	Netherlands	1	Malawi	2
Philippines	1	Malawi	2	New Zealand	1	Morocco	2
Sri Lanka	1	Morocco	2	Pakistan	1	Norway	2
Sweden	1	Norway	2	Philippines	1	Peru	2
Syria	1	Pakistan	2	Portugal	1	South Africa	2
United Kingdom	1	Peru	2	Sri Lanka	1	Tunisia	2
Uruguay	1	Portugal	2	Sweden	1	Turkey	2
USA	1	South Africa	2	Syria	1	Zimbabwe	2
		Tunisia	2	United Kingdom	1		
		Turkey	2	Uruguay	1		
		Zimbabwe	2	USA	1		

This classification is influenced by several factors such as different factor elasticities, efficiency patterns and/or shock effects. After estimating production frontiers, it is possible to perform the decomposition of labor productivity growth[17]. The decomposition results considering the evolution of all components between the first and the last year of the sample are reported in table 16[18].

[17] All decomposition calculations were carried out using GAUSS (Aptech Systems, Inc.,1999).
[18] We also use the mean values of all components in the time periods 1967-1979 and 1980-1992 to evaluate their contribution to the relative change in output per worker. For presentation reasons, we decide not to exhibit these results. Nevertheless, they are available upon request.

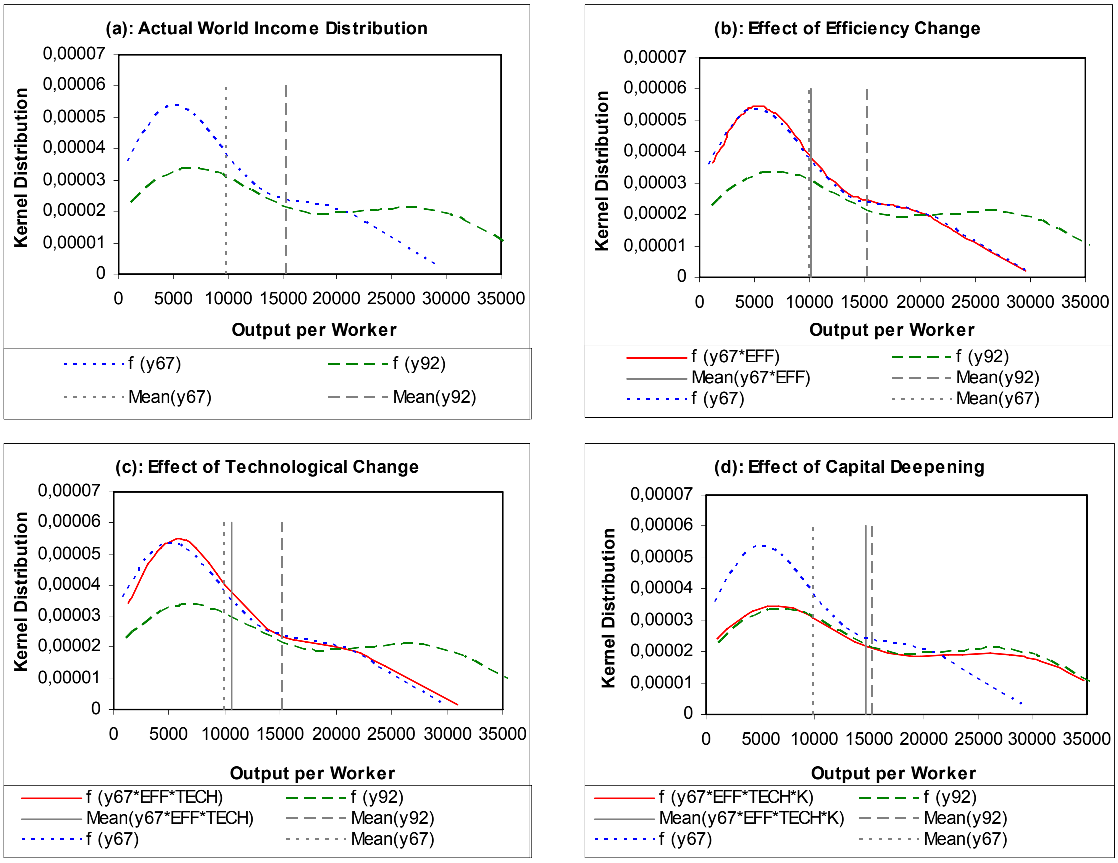
Table 16 - Decomposition of Labor Productivity Growth for the Economy

Country (by ascending order of output per worker in the first year)	Percentage Change in Output per Worker	Contribution to Percentage Change in Output per Worker of			
		Change in Efficiency	Change in Technology	Capital Deepening	Stochastic Shocks
Malawi	22,55%	30,81%	25,81%	-33,93%	12,70%
Kenya	15,74%	-14,93%	13,18%	22,65%	-1,99%
Indonesia	248,48%	2,36%	-20,79%	328,30%	0,35%
India	95,07%	2,90%	79,44%	12,45%	-6,05%
Madagascar	-33,85%	8,17%	70,30%	-73,94%	37,81%
Zimbabwe	-6,22%	9,36%	3,33%	-25,13%	10,83%
Pakistan	66,21%	32,46%	42,82%	-27,04%	20,42%
Sri Lanka	78,95%	10,90%	4,53%	18,77%	29,98%
Korea, Republic of	422,65%	-0,76%	6,21%	419,20%	-4,50%
Philippines	24,41%	10,28%	7,01%	-14,22%	22,90%
Egypt	92,30%	19,45%	32,44%	8,54%	12,00%
Honduras	15,56%	21,23%	2,01%	-14,70%	9,55%
Turkey	107,44%	3,04%	0,91%	94,07%	2,80%
Morocco	53,10%	5,10%	0,70%	39,36%	3,80%
Tunisia	103,62%	29,59%	-0,09%	39,49%	12,74%
Dominican Republic	40,71%	1,93%	24,44%	23,85%	-10,43%
El Salvador	3,22%	2,81%	0,25%	-2,83%	3,06%
Guatemala	23,56%	5,62%	0,12%	-8,23%	27,33%
Colombia	31,98%	8,63%	-3,65%	1,56%	24,17%
Portugal	160,32%	3,48%	22,25%	130,16%	-10,59%
Syrian Arab Republic	126,30%	4,37%	-5,04%	83,33%	24,56%
South Africa	17,02%	-12,17%	1,41%	31,10%	0,21%
Greece	108,70%	13,35%	4,73%	67,81%	4,76%
Costa Rica	15,18%	1,43%	-0,12%	12,54%	1,04%
Japan	167,57%	1,61%	14,67%	118,67%	5,01%
Peru	-33,56%	-28,76%	-0,07%	0,54%	-7,17%
Uruguay	30,30%	6,99%	-6,41%	5,47%	23,37%
Chile	24,22%	9,34%	16,61%	-8,74%	6,75%
Israel	108,47%	13,08%	-0,90%	43,10%	29,99%
Argentina	19,13%	-11,56%	-17,97%	58,71%	3,47%
Finland	71,42%	0,75%	14,38%	44,59%	2,88%
Austria	84,33%	-4,35%	13,63%	69,73%	-0,07%
Italy	95,40%	-0,49%	2,14%	97,50%	-2,66%
United Kingdom	51,35%	0,47%	-0,66%	50,97%	0,45%
Denmark	36,22%	-4,19%	-13,84%	52,90%	7,93%
France	63,81%	1,54%	5,65%	47,81%	3,30%
Norway	67,40%	1,23%	18,38%	37,97%	1,25%
Sweden	26,29%	-3,97%	6,44%	34,24%	-7,97%
Netherlands	38,77%	-2,36%	3,98%	49,99%	-8,87%
Australia	35,76%	1,31%	3,08%	25,41%	3,66%
New Zealand	4,80%	1,41%	-0,32%	-0,57%	4,26%
Canada	39,33%	1,71%	3,15%	24,47%	6,70%
USA	23,76%	0,85%	4,21%	12,50%	4,67%
Mean	**64,83%**	**4,28%**	**8,80%**	**44,15%**	**7,08%**

As we can see in table 16, there is evidence of catch-up for the majority of the countries: rich as well as poor countries have, on average, moved towards the frontier. Capital deepening is, in general, the most important determinant of labor productivity growth for the majority of countries

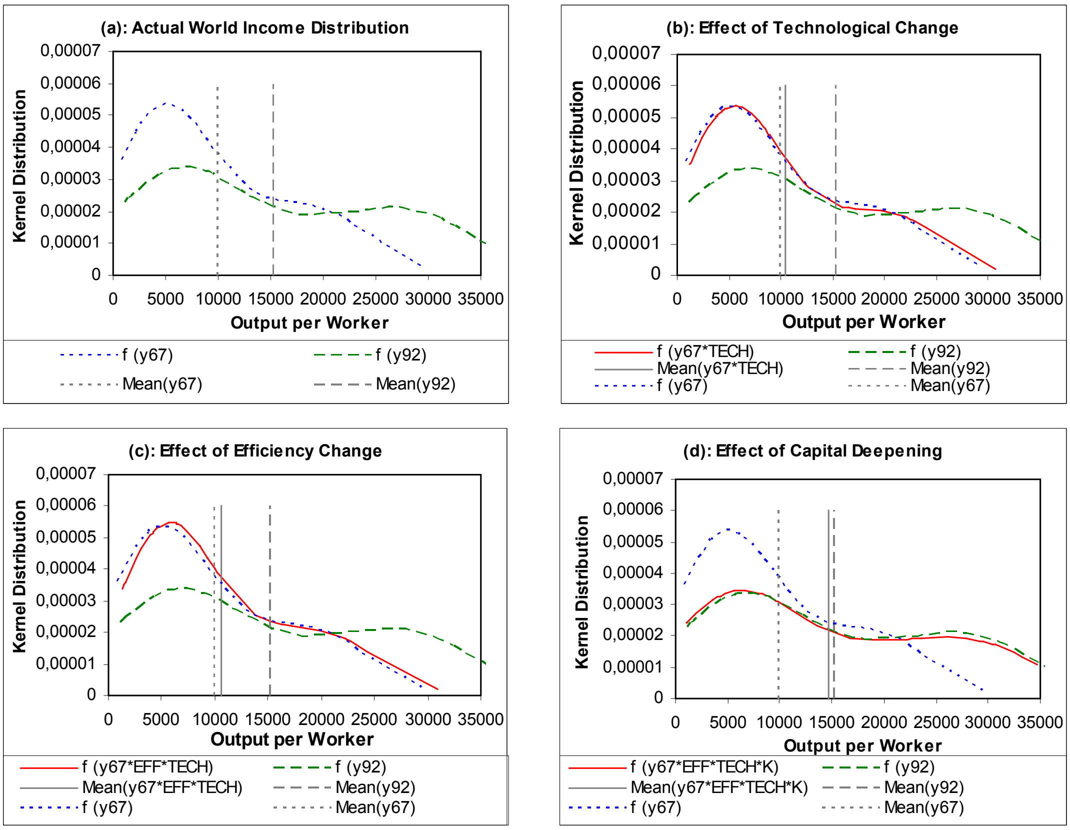
Finally, the results of the decomposition analysis are used for obtaining the counterfactual distributions through the nonparametric kernel density estimator. The estimated counterfactual distributions of labor productivity are presented in figures 4-6.

Figure 4 - Counterfactual Distributions of Output per Worker for the Economy



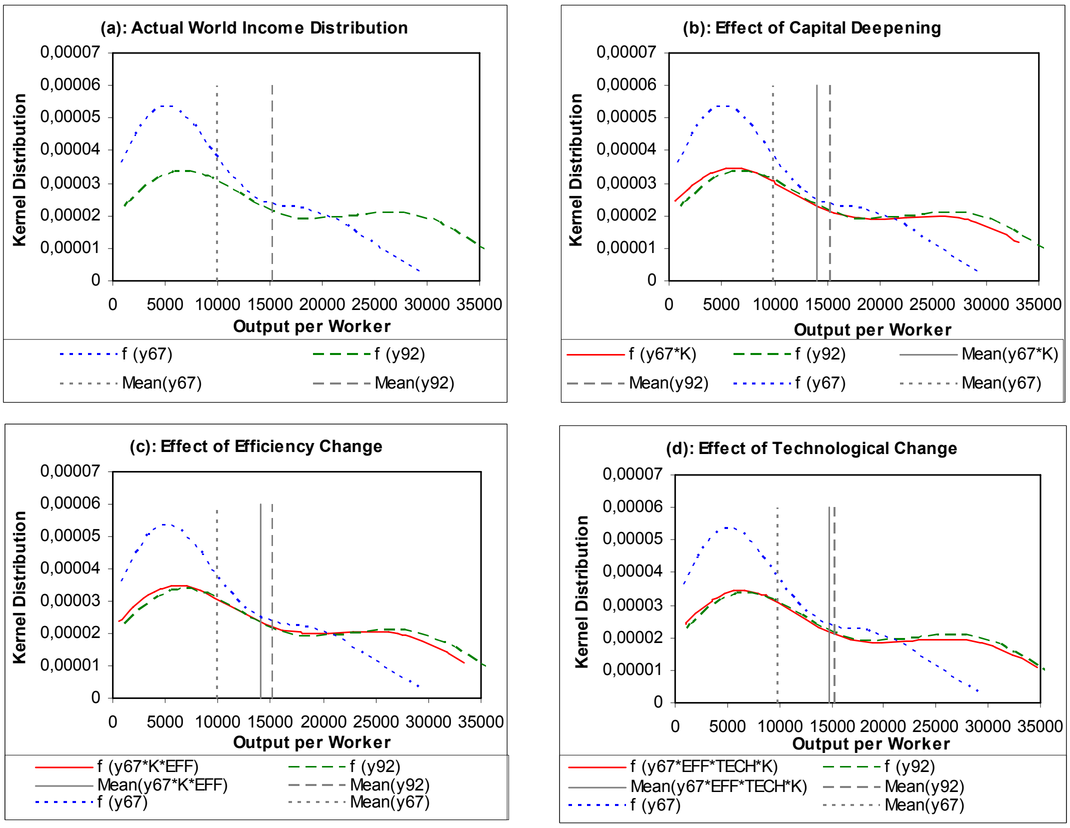
Panel b of figure 4 reveals that the efficiency change has an almost imperceptible effect on the first year labor productivity distribution. There is a very small shift of the density function from the lower and upper tails to the middle, without significant changes in the labor productivity mean.

Figure 5 - Counterfactual Distributions of Output per Worker for the Economy



Panel b of figure 5 shows that technological change is responsible for a small shift of the density function from the lower tail to the low-middle and a more significant transfer from the high-middle to the upper tail of the distribution, with a small rise of the labor productivity mean. This fact confirms the result of Kumar and Russell (2002) that technological change has contributed more to the welfare of richer countries than poorer ones. Nevertheless, as we can observe in table 8, there are some low and low-middle income countries such as Malawi, India, Madagascar, Zimbabwe, Pakistan, Dominican Republican and Chile in which efficiency or technology change is the main contributor to growth. This outcome can also be confirmed by the transfer of mass from the low to the low-middle income countries brought by the conjugated effect of the technological change and catch-up (panel c of figures 4 and 5).

Figure 6 - Counterfactual Distributions of Output per Worker for the Economy



Again following Kumar and Russel (2002), we test for the closeness of each of counterfactual distributions to the labor productivity distribution of 1992, using Li's T-test (Li, 1996).

For any two distributions f(x) and g(x) on the integrated-square-error metric space, $I(f,g)=\int[f(x)-g(x)]^2\,dx$:

$$T=\frac{I\cdot n\cdot\sqrt{h}}{\hat{\sigma}}, \tag{5.22}$$

where:

$$I = \frac{1}{n^2 h} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left[k\left(\frac{x_i - x_j}{h} \right) + k\left(\frac{y_i - y_j}{h} \right) - k\left(\frac{y_i - x_j}{h} \right) - k\left(\frac{x_i - y_j}{h} \right) \right]; \qquad (5.23)$$

$$\hat{\sigma}^2 = \frac{1}{n^2 h \sqrt{\pi}} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left[k\left(\frac{x_i - x_j}{h} \right) + k\left(\frac{y_i - y_j}{h} \right) + 2k\left(\frac{x_i - y_j}{h} \right) \right]; \qquad (5.24)$$

k(.) denotes a Gaussian kernel function and *h* is the optimal bandwidth (for details, please refer to Pagan and Ullah, 1999).

Li (1996) demonstrates that this statistic test is valid for dependent and independent variables. Fan and Ullah (1999) show that the T-statistic goes asymptotically to the standard normal. The results of Li's tests for the closeness of distributions are reported in table 17[19].

Table 17 - Li's Distribution Hypothesis Tests for the Economy

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	2.398	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	2.639	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech)$	2.412	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * FAcc)$	-0.073	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech)$	2.643	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	-0.020	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	-0.026	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	-0.009	**H_0 not rejected**	**H_0 not rejected**

[19] Our calculations for this test were carried out using GAUSS (Aptech Systems, Inc.,1999).

Comparing panel b in figures 4-6, which reports the effect of a single component, we can infer that capital deepening causes the emergence of a bimodal distribution and leads to a significant increase in the mean of labor productivity. The appropriate statistical tests support this conclusion (table 17). At both significance levels, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual distribution assuming only capital deepening. Additionally, when this effect is combined with each of the other components, the null hypothesis of equivalence relatively to the last year distribution cannot be rejected either.

5.3.2 - Agricultural Sector

For agriculture we adopt the following translog production frontier model (ignoring the j-class subscript, for notational ease):

$$\ln y_{it} = \beta_0 + \beta_1 \ln k_{it} + \beta_2 \ln la_{it} + \beta_3 \left(\ln k_{it}\right)^2 + \beta_4 \left(\ln la_{it}\right)^2 + \beta_5 \ln k_{it} \ln la_{it} + v_{it} - u_{it} \quad (5.25)$$

where la_{it} designates land per worker in year t for the i-th country and all the other variables are defined as before.

The model is estimated first with a large number of classes. Table 12 reports the score values of SBIC for 1 and 2 class models. The 3 and 4 class models are over-specified for agriculture since convergence is not attained. The score values suggest a 2-class model, whose estimation results are presented in table 18.

Table 18 - Two Class Model Estimation Results for Agriculture
a) 1967-1979

	Model parameters for latent class 1			**Model parameters for latent class 2**		
Variable	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***
Constant	2,3565	1,2931	0,0684	9,0558	0,2264	0,0000
$\ln k_{it}$	0,7012	0,2254	0,0019	0,4485	0,0479	0,0000
$\ln la_{it}$	1,5487	0,2835	0,0000	-0,3563	0,0655	0,0000
$(\ln k_{it})^2$	0,0127	0,0065	0,0521	-0,0825	0,0041	0,0000
$(\ln la_{it})^2$	-0,0899	0,0419	0,0320	-0,1014	0,0100	0,0000
$\ln k_{it} \cdot \ln la_{it}$	-0,0285	0,0366	0,4371	0,2378	0,0113	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,1864	0,3559	0,6006	0,6742	0,0078	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	0,3477	9,7291	0,9715	6,2148	0,7656	0,0000
Prior Probabilities for Class Membership	0,2221	0,0621	0,0004	0,7779	0,0621	0,0000

b) 1980-1992

	Model parameters for latent class 1			**Model parameters for latent class 2**		
Variable	***Coeff.***	***St.Err.***	***P[|Z|>z]***	***Coeff.***	***St.Err.***	***P[|Z|>z]***
Constant	6,5715	0,4885	0,0000	7,9509	0,1258	0,0000
$\ln k_{it}$	0,4140	0,1657	0,0125	0,5748	0,0735	0,0000
$\ln la_{it}$	0,7919	0,0995	0,0000	-0,0543	0,0793	0,4940
$(\ln k_{it})^2$	-0,0173	0,0143	0,2263	-0,1196	0,0065	0,0000
$(\ln la_{it})^2$	-0,0413	0,0110	0,0002	-0,1037	0,0166	0,0000
$\ln k_{it} \cdot \ln la_{it}$	0,0322	0,0174	0,0640	0,2493	0,0219	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,4413	0,0682	0,0000	0,3561	0,0114	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	0,6658	0,4957	0,1792	4,0321	0,9388	0,0000
Prior Probabilities for Class Membership	0,4693	0,0995	0,0000	0,5307	0,0995	0,0000

Inspection of those outcomes indicates that one of the lambdas and some of the coefficients are not statistically significant in both time periods. Hence, we consider the single class model, the results of which are reported in table 19.

Table 19 - One Class Model Estimation Results for Agriculture

a) 1967-1979

	Model parameters for latent class 1		
Variable	*Coeff.*	*St.Err.*	*P[|Z|>z]*
Constant	7,8479	0,3347	0,0000
$\ln k_{it}$	0,1632	0,0737	0,0268
$\ln la_{it}$	0,4114	0,0793	0,0000
$(\ln k_{it})^2$	-0,0305	0,0046	0,0000
$(\ln la_{it})^2$	-0,1150	0,0110	0,0000
$\ln k_{it} \cdot \ln la_{it}$	0,1584	0,0150	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,7570	0,0132	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	1,5599	0,1306	0,0000

b) 1980-1992

	Model parameters for latent class 1		
Variable	*Coeff.*	*St.Err.*	*P[|Z|>z]*
Constant	8,1360	0,2091	0,0000
$\ln k_{it}$	0,4601	0,0715	0,0000
$\ln la_{it}$	0,0151	0,0647	0,8159
$(\ln k_{it})^2$	-0,0600	0,0068	0,0000
$(\ln la_{it})^2$	-0,0884	0,0096	0,0000
$\ln k_{it} \cdot \ln la_{it}$	0,1773	0,0133	0,0000
$\sigma_j = [\sigma_{vj}^2 + \sigma_{uj}^2]^{1/2}$	0,5380	0,0130	0,0000
$\lambda_j = \sigma_{uj} / \sigma_{vj}$	1,1033	0,1177	0,0000

Based on the outcomes for both time periods, there is evidence supporting the use of a single production frontier for all countries.

As before, we perform the decomposition analysis of labor productivity growth and generate the corresponding counterfactual distributions. The results of the decomposition are presented in table 20; the counterfactual distributions of labor productivity in figures 7-9; the Li's tests in table 21.

Table 20 - Decomposition of Labor Productivity Growth for Agriculture

Country (by ascending order of output per worker in the first year)	Percentage Change in Output per Worker	Contribution to Percentage Change in Output per Worker of: Change in Efficiency	Change in Technology	Factor Accumulation	Stochastic Shocks
Malawi	-30,59%	13,43%	19,75%	-54,84%	13,14%
Indonesia	93,06%	34,52%	28,68%	-39,05%	82,96%
India	42,03%	21,57%	23,87%	-35,34%	45,85%
Kenya	6,92%	19,24%	28,34%	-42,91%	22,37%
Korea, Republic of	395,34%	26,19%	32,55%	81,51%	63,16%
Zimbabwe	-33,14%	4,51%	24,17%	-37,70%	-17,30%
Madagascar	-17,99%	14,56%	26,73%	-55,55%	27,07%
Pakistan	20,61%	11,47%	24,34%	-33,70%	31,25%
Sri Lanka	-13,25%	8,14%	30,42%	-41,93%	5,92%
Guatemala	22,25%	26,79%	29,45%	-53,95%	61,74%
Morocco	10,96%	24,08%	26,20%	-58,73%	71,72%
Philippines	38,81%	14,91%	29,48%	-45,28%	70,50%
Egypt	39,67%	8,10%	37,56%	-23,76%	23,20%
Turkey	63,49%	39,59%	30,35%	-49,57%	78,17%
Iran	74,37%	169,14%	-1,57%	-82,88%	284,50%
El Salvador	40,37%	26,79%	28,99%	-45,92%	58,69%
Japan	258,68%	18,83%	6,68%	109,66%	34,94%
Peru	-9,33%	7,12%	29,73%	-34,48%	-0,41%
Tunisia	96,41%	33,47%	38,93%	-23,81%	39,02%
Dominican Republic	75,75%	9,00%	30,31%	-4,36%	29,38%
Honduras	15,95%	19,97%	31,94%	-41,49%	25,19%
Colombia	65,67%	15,04%	30,76%	-34,45%	68,02%
Portugal	177,78%	-0,90%	32,55%	214,12%	-32,68%
Syrian Arab Republic	71,98%	37,39%	37,62%	-45,49%	66,87%
Venezuela	100,32%	23,80%	33,51%	-28,41%	69,29%
Costa Rica	99,72%	30,50%	23,93%	-37,24%	96,76%
South Africa	57,51%	17,55%	37,16%	-30,78%	41,14%
Greece	183,84%	24,96%	27,73%	-11,46%	100,84%
Chile	44,82%	12,58%	35,28%	-27,27%	30,74%
Finland	138,57%	6,07%	28,24%	118,11%	-19,58%
Italy	197,22%	33,53%	8,28%	29,46%	58,78%
Norway	100,08%	5,68%	18,29%	86,79%	-14,32%
Austria	156,24%	12,71%	19,34%	50,87%	26,27%
Sweden	199,32%	-0,25%	36,99%	198,03%	-26,50%
Israel	173,93%	29,24%	0,80%	13,17%	85,81%
France	279,99%	26,94%	19,24%	46,53%	71,32%
Uruguay	59,25%	12,29%	34,04%	-51,67%	118,91%
Argentina	47,14%	23,16%	43,54%	-59,17%	103,85%
Denmark	151,35%	9,79%	29,41%	44,04%	22,83%
Netherlands	140,19%	5,00%	10,39%	27,62%	62,37%
Canada	192,47%	4,56%	91,98%	57,70%	-7,62%
United Kingdom	78,66%	6,78%	29,02%	-20,52%	63,14%
USA	67,95%	11,24%	65,28%	-37,03%	45,08%
Australia	63,04%	27,59%	88,29%	-63,18%	84,32%
New Zealand	0,95%	0,94%	13,88%	11,51%	-21,25%
Mean	89,74%	**20,61%**	**30,05%**	**-3,62%**	**47,68%**

Figure 7 - Counterfactual Distributions of Output per Worker for Agriculture

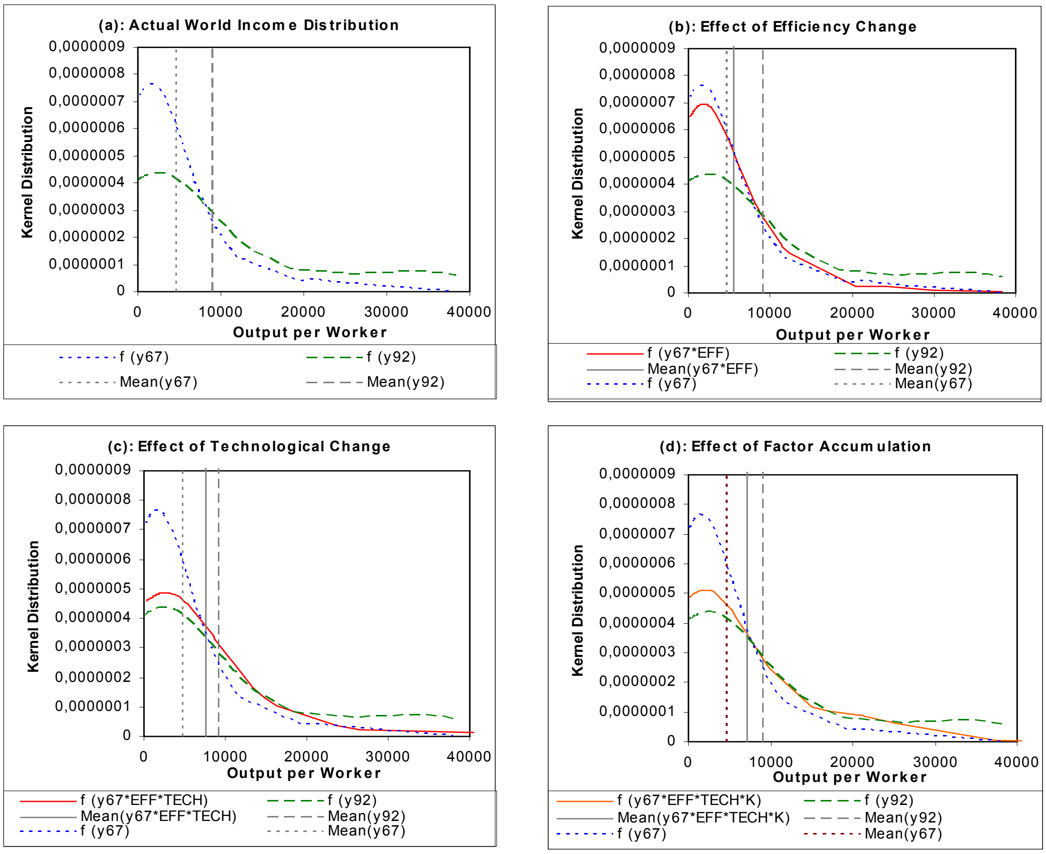


Figure 8 - Counterfactual Distributions of Output per Worker for Agriculture

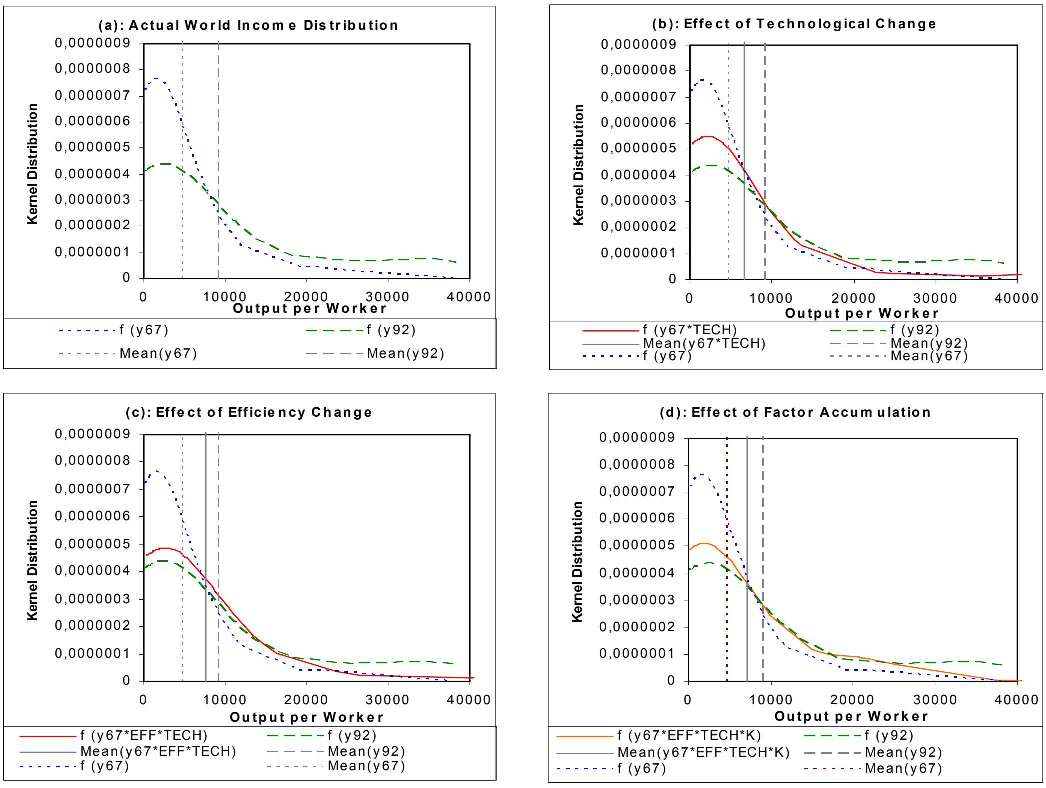


Figure 9 - Counterfactual Distributions of Output per Worker for Agriculture

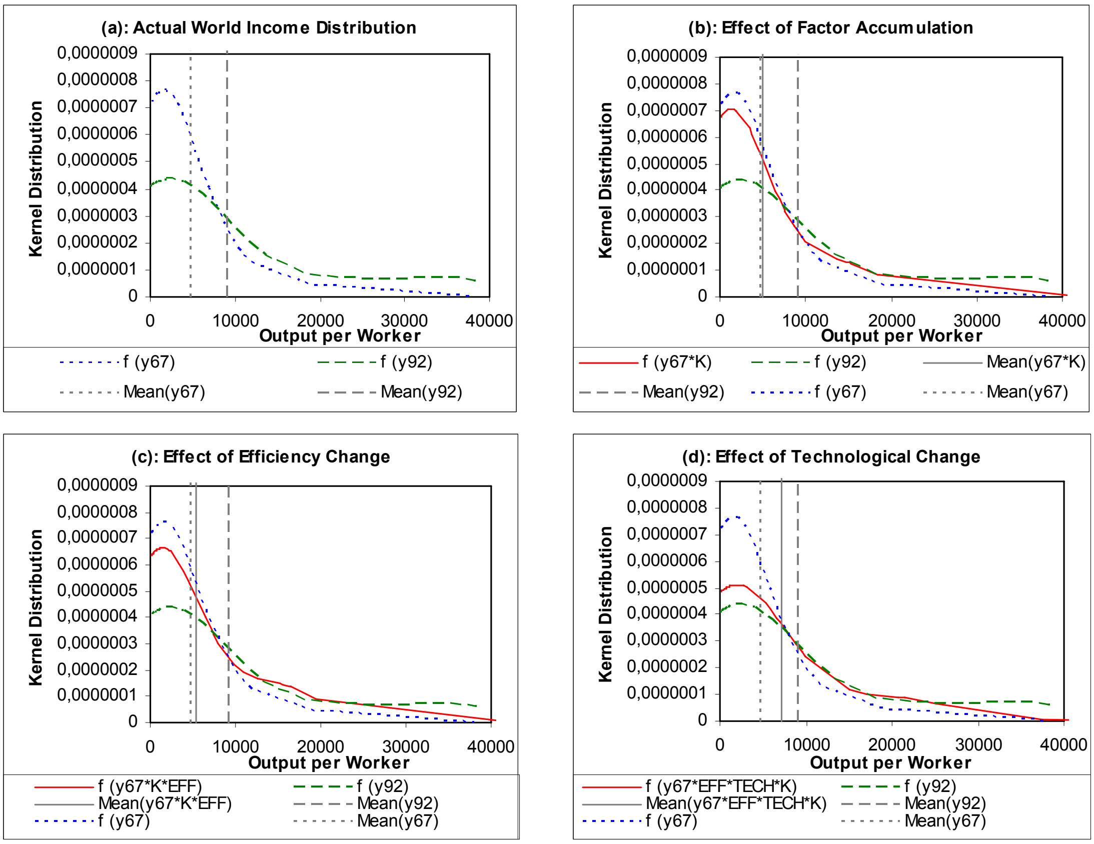


Table 21 - Li's Distribution Hypothesis Tests for Agriculture

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	5.842	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	4.336	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech)$	1.832	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * FAcc)$	4.686	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * Tech)$	0.982	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	3.719	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	1.350	H_0 rejected	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	0.774	**H_0 not rejected**	**H_0 not rejected**

The analysis of those results suggests the following:

- The catch-up effect is stronger for agriculture than for the overall economy, as described in table 20. Panel *b* of figure 7 indicates that efficiency change is responsible for a shift of the density from both tails to the middle; i.e., exactly the opposite of the divergence process described by Quah (1993) as the “pilling up of probability mass in the tails, and a thinning out in the middle”. Thus, although rich as well as poor countries move toward the world production frontier, the technological catch-up seems to help convergence in agriculture.

- The analysis of table 20 reveals that technological change is the most important component for the majority of countries. Panel *b* of figure 8 suggests that technological change is responsible for an important shift of the density from the lower tail to the low-middle and an almost imperceptible mass change from the high-middle to the upper tail of the distribution. This means that, contrary to the overall economy, technological change in agriculture, contributes more to the welfare of poorer countries than richer ones. The combined effect of the catch-up and technological change components on the distribution of labor productivity is presented in panel *c* of figures 7 and 8. The analysis of panel c shows that the combined effect of the two components results in a higher mean of output per worker and in a 1967-distribution closer to the 1992-distribution than the individual effect of each component. This conclusion is supported by the statistic tests of Li (1996) presented in table 21. At both significance levels, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual distribution assuming only TFP change.

- It is notable that many countries experience reductions in factor per worker endowments[20], as we can observe in table 20. Nevertheless, factor accumulation is a very important determinant of growth for some countries, such as two Southeast Asian growth miracles (Korea and Japan) and some European countries (Portugal, Finland, Norway, Austria, Sweden, France, Netherlands). Panel *b* of figure 9 indicates that factor accumulation effect leads to a shift from the lower tail to the rest of distribution.

[20] For instance, although labor has diminished in the period for most of the countries, the reduction of factor endowments was even stronger.

However, this is a small effect, with minor changes on labor productivity distribution and its mean.

5.4 - Differences Among Groups of Countries

In chapter 4, we conclude that, on average, the output per worker increased more in agriculture than in the overall economy for the period 1967-1992 and that the weight of agriculture in employment declined during that time. To study the behavior of each set of countries, we use the criterion defined by the World Bank presented in the table 6, leading to the sample classification accessible in table 7. In the economy, labor productivity presents similar growth rates across sets, as shown in table 8. On the contrary, in agriculture, the rich and upper middle countries exhibit high growth rates, contrasting with moderate rates for the other countries.

Table 22 presents the decomposition of labor productivity growth for each of the 4 groups of countries considered.

Table 22 - Decomposition of Labor Productivity Growth (4 Groups)

		N	Mean	Std. Deviation	Minimum	Maximum
TFP change agriculture	Rich	16	,4352	,28449	,15	1,16
	Upper Middle	6	,5610	,06788	,46	,67
	Lower Middle	13	,6275	,33726	,37	1,68
	Poor	10	,4313	,09980	,29	,63
	Total	45	,5067	,26276	,15	1,68
TFP change economy	Rich	16	,0677	,10011	-,18	,26
	Upper Middle	5	-,0328	,17976	-,30	,18
	Lower Middle	12	,0785	,15796	-,29	,30
	Poor	10	,3757	,35987	-,18	,82
	Total	43	,1307	,24639	-,30	,82
Capital deepening agriculture	Rich	16	,55431	,762177	-,632	2,141
	Upper Middle	6	-,16683	,510632	-,592	,815
	Lower Middle	13	-,41800	,189011	-,829	-,044
	Poor	10	-,40610	,094301	-,555	-,238
	Total	45	-,03616	,665226	-,829	2,141

		N	Mean	Std. Deviation	Minimum	Maximum
Capital deepening economy	Rich	16	,5247	,35968	-,01	1,30
	Upper Middle	5	1,1646	1,70993	,06	4,19
	Lower Middle	12	,2174	,36076	-,14	,94
	Poor	10	,2162	1,11748	-,74	3,28
	Total	43	,4416	,84715	-,74	4,19
Y/L chg no stochastic shocks agriculture	Rich	16	1,1788	1,00153	-,12	3,13
	Upper Middle	6	,3663	,86197	-,28	2,04
	Lower Middle	13	-,0544	,26127	-,55	,41
	Poor	10	-,1230	,16397	-,39	,13
	Total	45	,4249	,89093	-,55	3,13
Y/L chg no stochastic shocks economy	Rich	16	,6430	,49724	,00	1,91
	Upper Middle	5	1,1669	1,88452	,06	4,47
	Lower Middle	12	,3128	,41164	-,29	1,02
	Poor	10	,4675	,83186	-,52	2,47
	Total	43	,5709	,82768	-,52	4,47

For all groups, TFP growth rates are higher in agriculture than in the overall economy. Furthermore, factor accumulation in agriculture is negative for all classes, except for the rich countries. Although output per capita grew faster in the period for agriculture, when we remove the stochastic shocks, this conclusion is altered.

The ANOVA tests presented in table 23 indicate that the null hypothesis that the means of output per capita change without shocks are equal is rejected for the agriculture and not rejected for the economy. The same indication applies to capital deepening and the opposite to TFP. Capital deepening in the overall economy mimics the behavior of output per capita after removing shocks. In agriculture, capital deepening is only positive for rich countries while TFP is similar among groups and stronger than the growth rates observed in the overall economy for all sets of nations.

Table 23 - ANOVA Tests (4 Groups of Countries)

		Sum of Squares	df	Mean Square	F	Sig.
TFP change agriculture	Between Groups	,346	3	,115	1,758	,170
	Within Groups	2,692	41	,066		
	Total	3,038	44			
TFP change economy	Between Groups	,830	3	,277	6,276	,001
	Within Groups	1,720	39	,044		
	Total	2,550	42			
Capital deepening agriculture	Between Groups	8,945	3	2,982	11,614	,000
	Within Groups	10,526	41	,257		
	Total	19,471	44			
Capital deepening economy	Between Groups	3,835	3	1,278	1,895	,146
	Within Groups	26,307	39	,675		
	Total	30,142	42			
Y/L chg no stochastic shocks agriculture	Between Groups	15,103	3	5,034	10,413	,000
	Within Groups	19,822	41	,483		
	Total	34,925	44			
Y/L chg no stochastic shocks economy	Between Groups	2,766	3	,922	1,383	,262
	Within Groups	26,006	39	,667		
	Total	28,772	42			

Same conclusions are obtained when we only consider, in table 24, the classical split up of countries into developed or rich and developing.

Table 24 - Decomposition of Labor Productivity Growth (2 Groups)

		N	Mean	Std. Deviation	Minimum	Maximum
TFP change agriculture	Developed	16	,4352	,28449	,15	1,16
	Developing	29	,5461	,24618	,29	1,68
	Total	45	,5067	,26276	,15	1,68
TFP change economy	Developed	16	,0677	,10011	-,18	,26
	Developing	27	,1680	,29732	-,30	,82
	Total	43	,1307	,24639	-,30	,82
Capital deepening agriculture	Developed	16	,55431	,762177	-,632	2,141
	Developing	29	-,36193	,273941	-,829	,815
	Total	45	-,03616	,665226	-,829	2,141
Capital deepening economy	Developed	16	,5247	,35968	-,01	1,30
	Developing	27	,3924	1,03822	-,74	4,19
	Total	43	,4416	,84715	-,74	4,19
Y/L chg no s. shocks economy	Developed	16	,6430	,49724	,00	1,91
	Developing	27	,5282	,97923	-,52	4,47
	Total	43	,5709	,82768	-,52	4,47
Y/L chg no s. shocks agriculture	Developed	16	1,1788	1,00153	-,12	3,13
	Developing	29	,0090	,45390	-,55	2,04
	Total	45	,4249	,89093	-,55	3,13

In agriculture, although developing countries exhibit a strong positive growth rate for TFP, the disinvestment in this sector shown by the factor accumulation effect causes a null growth of output per capita without shocks. For developed countries, high rates of factor accumulation and TFP originate an even higher growth rates of output per capita with no stochastic shocks. For both sets of countries, TFP growth rates are lower in the overall economy than in agriculture. Furthermore, TFP growth is stronger in developing countries than in developed countries both for agriculture and for the overall economy. ANOVA tests exhibited in table 25 reveal that capital deepening is the only case that the null hypothesis of group means being equal is rejected.

Table 25 - ANOVA Tests (2 Groups of Countries)

		Sum of Squares	df	Mean Square	F	Sig.
TFP change agriculture	Between Groups	,127	1	,127	1,874	,178
	Within Groups	2,911	43	,068		
	Total	3,038	44			
TFP change economy	Between Groups	,101	1	,101	1,691	,201
	Within Groups	2,449	41	,060		
	Total	2,550	42			
Capital deepening agriculture	Between Groups	8,656	1	8,656	34,417	,000
	Within Groups	10,815	43	,252		
	Total	19,471	44			
Capital deepening economy	Between Groups	,176	1	,176	,241	,626
	Within Groups	29,966	41	,731		
	Total	30,142	42			
Y/L chg no s. shocks economy	Between Groups	,132	1	,132	,189	,666
	Within Groups	28,640	41	,699		
	Total	28,772	42			
Y/L chg no s. shocks agriculture	Between Groups	14,110	1	14,110	29,150	,000
	Within Groups	20,815	43	,484		
	Total	34,925	44			

In neoclassical growth models (e.g., Solow, 1956; Cass, 1965 and Koopmans, 1965), per capita output growth rate for a given period is assumed to be inversely related to its starting level. Therefore, countries with different starting levels of per capita income will tend to converge. Barro (1991) tests this hypothesis to a cross section of 98 countries, using average rates of the period 1960-1985. Barro (1991) finds out some empirical regularities about growth, fertility, and investment. Following this study, in addition to the initial level of income, we introduce two agriculture-related indicators as explanatory variables of per capita output growth rates. Our purpose is to find out if a shrinking agricultural sector is a pre-condition for economic growth and for fast labor productivity in agriculture itself.

Table 26 reports information about a linear regression by ordinary least squares in which the dependent variable is the rate of output per worker growth for the overall economy.

Table 26 - Regressions on Output per Worker Growth Rates for the Economy

	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	1,421	,747		1,904	,064
Starting weight of agric. in employment	-1,631	,885	-,547	-1,842	,073
Variation of agricultural weight in employment	-3,789	1,819	-,350	-2,083	,044
Starting output per worker in the economy	-5,21E-005	,000	-,501	-1,536	,133
Stepwise method:					
	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	,200	,179		1,116	,271
Variation of agricultural weight in employment	-4,771	1,520	-,440	-3,140	,003

As the stepwise method reveals, the only significant independent variable at 5% is the variation of agricultural weight in employment. Nevertheless, estimates suggest that, ceteris paribus, output growth rates are higher in countries:

- with a smaller agricultural sector at the beginning of the period;
- in which the agricultural sector has a higher shrinkage.

This evidence apparently indicates that agriculture has a negative role in economic growth, since the reduction of agriculture weight appears to be a condition for development. The same result is obtained in table 27 for output per worker growth in agriculture, with all variables being significant at 5%.

Table 27 - Regressions on Output per Worker Growth Rates for Agriculture

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	1,821	,326		5,594	,000
Starting weight of agric in emp 1974	-2,874	,491	-,828	-5,854	,000
Variation of agricultural weight in employment	-3,770	1,623	-,299	-2,323	,026
Starting output per worker in agriculture	-4,08E-007	,000	-,349	-2,245	,031

We have found previously that only developed countries do not experience reductions in factor accumulation. Therefore, it could be interesting to check if the negative role of agriculture holds for total factor productivity growth rates. In table 28, it is possible to find that only the variation of agricultural weight in employment, is significant, but it is important to notice that the signs of the estimates related to agriculture have changed.

Table 28 - Regressions on TFP Growth Rates for the Economy

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	,028	,242		,116	,908
Starting weight of agriculture in employment	,423	,286	,448	1,477	,148
Variation of agricultural weight in employment	,404	,588	,118	,686	,497
Starting output per worker in the economy	-1,76E-006	,000	-,053	-,160	,874
Stepwise method:					
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-,032	,059		-,538	,593
Variation of agricultural weight in employment	,435	,131	,461	3,330	,002

This evidence indicates that, ceteris paribus, a country with a large and non-declining agricultural sector tends to exhibit stronger TFP rates. Therefore, the reason for the negative role of agriculture to growth in developing countries seems to be the large negative rates of the factor accumulation effect occurring in the agricultural sector, probably due to the market distortions introduced by agricultural policies of the developed nations. If this disinvestment did not occur, since TFP growth rates in agriculture are stronger for developing nations and higher than the ones presented for the economy, agriculture could be the growth engine for developing countries, inducing rises in TFP growth rates for the overall economy and, as a result, in output per labor growth rates.

6 - A Penalized Spline Approach for the Decomposition of Labor Productivity Growth

6.1 - Introduction

Panel data refers to data where cross-section units are observed over several periods of time. The main difference between panel data analysis and cross-sectional or time-series analysis is that it allows for the cross-sectional and time heterogeneity. It is possible to understand population heterogeneity and disaggregate changes over time within individuals from cohort effects, frequently mistaken with changes occurring within subject determinants. On the other hand, dealing with these dynamical relationships introduced by time-varying covariates adds complexity to the statistical model.

The most general approaches for analysis of longitudinal data are the linear mixed models built on the work of Laird and Ware (1982). The linear mixed models are an extension of linear models for which covariance structure is based on random effects and their covariance parameters. As pointed out by Wand (2005, p. 1), this general approach is developed through vigorous research both on analytic results and computational methods. Laird and Ware (1982) define each cross-section unit as a linear sum of a time-dependent population mean (modeled as a fixed effect) and a subject-specific component (modeled as a polynomial with random effects). This approach is quite robust to missing data and irregular spaced measurements, since it is a full-likelihood method making full use of all available information on each cross-section unit. It provides a very flexible framework that can be applied to both continuous and discrete dependent variables. Nevertheless, the model presented by Laird and Ware (1982) imposes a linear relationship between the covariates and the dependent variable.

Generalizations of Laird and Ware (1982)´s approach to nonparametric and semiparametric frameworks were developed: penalized splines (e.g., O´Sullivan, 1986); smoothing splines (e.g., Eubank, 1988; Wahba, 1990; Green and Silverman, 1994); regression splines (e.g., Hastie and Tibshirani, 1990; Friedman, 1991); kernel methods (e.g., Härdle, 1990; Zeger and Diggle, 1994; Wand and Jones, 1995; Fan and Gijbels, 1996). These early studies use several kinds of smoothers to estimate the mean population curve but random effects are modeled by parametric functions. Zhang *et al*. (1998) extend this work by accounting for the within-subject correlation using a Gaussian process, though they do not consider smooth curves for individual subjects. Brumback and Rice (1998) prove the equivalence between smoothing splines and mixed model representations with fixed subject-specific effects. However, they run into computational problems because they assume fixed slopes and intercepts for the subject-specific curves. Verbyla *et al*. (1999) generalize the equivalence proof to the case where mix models include individual random effects. Using the results of Verbyla *et al*. (1999), Rice and Wu (2001) model individual curves as spline functions with random coefficients where model-fitting and inferences are based on standard parametric methods operationally. Guo (2002) introduces a functional mixed effects model using smoothing splines. This model faces computational problems, since it uses as many knots as data points.

Spline regression models are too dependent on the number and position of the knots while smoothing spline models are too computationally intensive with large data sets. Smoothing spline models take all of the distinct time points as knots, overcoming the dependence of spline regression models on the location and number of knots, and use a roughness penalty to control the smoothness of the resulting smoothers. Yet, the model may not be estimated when the number of distinct time points is too large. Penalized spline models combine a penalty approach, introduced by Eilers and Marx (1996), with a low-rank smoother presented by Rice and Wu (2001). The penalty approach alleviates the dependence on the number and location of the knots and the low-rank smoother solves computational problems. More recent penalized spline models combine spline regression with smoothing spline (e.g., Ruppert *et al.*, 2003).

In this chapter, we estimate production and inefficiency functions using the penalized splines (P-splines) approach of Ruppert *et al.* (2003). It is possible to find in the literature some spline models used to estimate production and cost functions (e.g., Humphrey and Vale, 2004; Fox, 1998; Carbo Valverde and Humphrey, 2004;

Peeters and Surry, 2000; Fox and Grafton, 2000). In the particular case of P-splines, applications in economics are more scarce (e.g., Tanggaard, 1997; Huang and Nychka 2000). Our framework will bring all the advantages of the most recent approach for analysis of longitudinal data to the efficiency and productivity field of research. Recent international productivity growth studies combine stochastic models with parametric production functions (e.g., Martin and Mitra, 2001) and more flexible specifications with deterministic approaches (e.g., Kumar and Russell, 2002).

We specify a semiparametric and stochastic panel data model using both the frequentist (or classical) and Bayesian approaches. The classical approach uses a restricted maximum likelihood (REML) method; the Bayesian approach employs a Markov Chain Monte Carlo (MCMC) algorithm. The MCMC algorithm does not rely on normality or asymptotic assumptions allowing to overcome problems when the sample size is small. The model specifications in both approaches are similar with one important exception: all parameters are random in the Bayesian approach, while some of the spline components are fixed in the classical approach. Other stochastic frontier models with random coefficients can be found in Tsionas (2002)[21].

Proceeding in the same way as in the previous chapter, the stochastic semiparametric models are estimated separately for the time periods 1967-1979 and 1980-1992.

This chapter is organized as follows. In section 2, we present the general semiparametric penalized spline model. In section 3 and 4, this model is completely specified for a frequentist and Bayesian approach, respectively, within univariate and multivariate cases. In section 5, we extend these models by considering the division of countries into 2 classes presented in chapter 5.

[21] Please refer to chapter 3 for a comprehensive survey on panel data production frontier methods, including other semiparametric and nonparametric frameworks.

6.2 - A Penalized Spline Model

We propose a semiparametric production frontier model in the following way:

$$y_{it} = m^*(x_{it}) + \varepsilon_{it} \qquad i = 1,...,N\,;\; t = 1,...,T, \tag{6.1}$$

where:

'i' indexes countries and *'t'* indexes time periods; y_{it} is the log of the production level per worker in year t for the i-th country; x_{it} is a $q \times 1$ vector of the log of inputs per worker in year t for the i-th country; $m^*(.)$ is an unknown world production frontier; ε_{it} is the error term.

The error term ε_{it} can be decomposed into statistical noise and inefficiency, resulting in the following frontier model:

$$y_{it} = m^*(x_{it}) + v_{it} + u_i^*(x_{it})\,,\; u_i^*(.) \le 0\,, \tag{6.2}$$

where:

$u_i^*(.)$ is the technical inefficiency component function representing deviations of each country's frontier from the world production frontier and v_{it} is the error term specified as $v_{it} \sim N\left(0, \delta_v^2\right)$.

6.2.1 - The Frequentist Approach

For clarity of exposition, we treat separately the univariate and multivariate cases.

a) Univariate Smoothing

Given that x_{it} is the log of capital per worker, the penalized linear spline model for $m^*(.)$ in (6.1) is:

$$m^*(x_{it}) = \beta_1 x_{it} + \sum_{k=1}^{K} w_k \left(x_{it} - \kappa_k\right)_+ , \tag{6.3}$$

where:

$m^*(x_{it})$ is a piecewise linear function with K number of knots $\kappa_1, \kappa_2, ..., \kappa_K$, w_k is the *k*th knot coefficient, $w_k \sim N(0, \delta_w^2)$, $(x_{it} - \kappa_k)_+$ is the linear spline basis function defined as $(x_{it} - \kappa_k)_+ = \begin{cases} x_{it} - \kappa_k & if \quad x_{it} > \kappa_k, \\ 0 & if \quad x_{it} \le \kappa_k. \end{cases}$.

The linear regression spline in (6.3) has two components: a fixed and linear component and a random deviation from linearity, using truncated lines as the basis for regression. The technical efficiency component in (6.2) is specified as:

$$u_i^*(x_{it}) = \alpha_{i1} x_{it} + \sum_{k=1}^{K} u_{ik} \left(x_{it} - \kappa_k\right)_+ , \tag{6.4}$$

where:

$\alpha_{i1} \sim N^-(0, \delta_\alpha^2)$ and $u_{ik} \sim N^-(0, \delta_u^2)$ with $N^-(.)$ denoting the half-normal distribution function[22].

Compared to $m^*(x_{it})$, $u_i^*(x_{it})$ has also two components, one linear and the other non-linear, yet both of them are random. This specification is innovative in the literature, allowing for subject-specific differences to be more than a random intercept.

Substituting (6.3) and (6.4) in (6.1), the resulting model can be written as:

$$y_{it} = \beta_1 x_{it} + \sum_{k=1}^{K} w_k \left(x_{it} - \kappa_k\right)_+ + \alpha_{i1} x_{it} + \sum_{k=1}^{K} u_{ik} \left(x_{it} - \kappa_k\right)_+ + v_{it} . \tag{6.5}$$

The model in (6.5) has two advantages: it employs a semiparametric approach to the production frontier and uses a relatively flexible inefficiency specification.

[22] We also tried the definition of a modal value different from zero. Nevertheless, estimation was not possible, due to convergence problems.

$m^*(x_{it})$ refers to the world production function, common to all countries. After considering inefficiency, each country has its own subject-specific production frontier. One alternative was to consider, as in Schmidt and Sickles (1984), a simple random effect $U_i \leq 0$, allowing for functions to differ only in the intercept. Lee and Schmidt (1993) generalize the model of Schmidt and Sickles (1984) to the case of time-varying technical efficiency. However, this formulation imposes a temporal pattern that is invariant across individuals. Kumbhakar (1990) and Battese and Coelli (1992) use maximum likelihood models, but propose rigid specifications, with a time trend included to allow for time-variant inefficiency, but unable to distinguish a pattern of variation for each individual. The specification presented in this study is more flexible, treating the differences relatively to the world production frontier as regression splines $u_i^*(x_{it})$ with all parameters being random. Inefficiency of each country is time variant in the sense that it depends on the level of inputs used, a dimension variable to the problem. The model of Battese and Coelli (1995) allows the use of panel data to estimate time variant inefficiency. It consists in the simultaneous estimation of the frontier production function with the variables which may influence the efficiency of a country's production function. In our model, contrary to the approach of Battese and Coelli (1995), it is not necessary to choose variables that affect efficiency, avoiding probable identification problems.

Using matrix notation, the model in (6.5) can be written as:

$$Y = X\beta + Z\omega + v, \tag{6.6}$$

where:

$$Y = \begin{bmatrix} Y_1 \\ . \\ . \\ . \\ Y_N \end{bmatrix}; Y_i = \begin{bmatrix} y_{i1} \\ . \\ . \\ . \\ y_{iT} \end{bmatrix}; \; X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}; \; X_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{iT} \end{bmatrix};$$

$$Z=\begin{bmatrix} Z_1 & X_1 & 0 & \cdots & 0 & Z_1 & 0 & \cdots & 0 \\ Z_2 & 0 & X_2 & \cdots & 0 & 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_N & 0 & 0 & \cdots & X_N & 0 & 0 & \cdots & Z_N \end{bmatrix}; Z_i=\begin{bmatrix} (x_{i1}-\kappa_1)_+ & \cdots & (x_{i1}-\kappa_K)_+ \\ \vdots & \ddots & \vdots \\ (x_{iT}-\kappa_1)_+ & \cdots & (x_{iT}-\kappa_K)_+ \end{bmatrix};$$

$$\omega=\begin{bmatrix} w & \alpha_{i1} & \dots & \alpha_{iN} & u_1 & \dots & u_N \end{bmatrix}^T; w=\begin{bmatrix} w_1 & \dots & w_K \end{bmatrix}; u_i=\begin{bmatrix} u_{i1} & \dots & u_{iK} \end{bmatrix};$$

$$v=\begin{bmatrix} v_{11} \\ \vdots \\ v_{NT} \end{bmatrix}; Cov(\omega)=G=\begin{bmatrix} \delta_w^2 I_K & 0 & 0 \\ 0 & \delta_\alpha^2 I_N & 0 \\ 0 & 0 & \delta_u^2 I_{NK} \end{bmatrix}.$$

Z is a matrix of dimension $NT\times(K+N+NK)$, ω with dimension $(K+N+NK)\times 1$ and G a symmetric matrix of dimension $(K+N+NK)\times(K+N+NK)$. For given values of the covariance components, the estimates of (β,ω) can be obtained by minimizing the penalized least squares function:

$$\min_{\beta,\omega}(Y-X\beta-Z\omega)^T(Y-X\beta-Z\omega)+\frac{\delta_v^2}{\delta_w^2}\omega^T P_1\omega+\frac{\delta_v^2}{\delta_u^2}\omega^T P_2\omega, \tag{6.7}$$

where:

$$P_1=\begin{bmatrix} I_K & 0_{\mathrm{K}\times(\mathrm{N+NK})} \\ 0_{(\mathrm{N+NK})\times\mathrm{K}} & 0_{(\mathrm{N+NK})\times(\mathrm{N+NK})} \end{bmatrix}; P_2=\begin{bmatrix} 0_{(\mathrm{K+N})\times(\mathrm{K+N})} & 0_{(\mathrm{K+N})\times\mathrm{NK}} \\ 0_{\mathrm{NK}\times(\mathrm{K+N})} & I_{\mathrm{NK}} \end{bmatrix};$$

$0_{\mathrm{r}\times\mathrm{s}}$ is a $(r\times s)$ matrix of zeros.

As in Durbán *et al*. (2005), $\frac{\delta_v^2}{\delta_w^2}$ and $\frac{\delta_v^2}{\delta_u^2}$ are smoothing parameters, controlling the amount of smoothing of the production frontier. P_1 and P_2 are designed in a way to penalize the coefficients of the spline basis functions $(x_{it}-\kappa_k)_+$.

Dividing (6.7) by the error variance δ_v^2:

$$\min_{\beta,\omega}\frac{1}{\delta_v^2}(Y-X\beta-Z\omega)^T(Y-X\beta-Z\omega)+\frac{1}{\delta_w^2}\omega^T P_1\omega+\frac{1}{\delta_u^2}\omega^T P_2\omega. \tag{6.8}$$

If we assume that ω is a matrix of random coefficients independent relatively to v, i.e., with $\mathrm{E}(\omega)=0$ and $Cov(\omega)=G$, then solution of (6.8) is equivalent to the BLUP in the linear mixed model representation (Laird and Ware, 1982 and Brumback *et al*., 1999) of a penalized spline. This is a major empirical step since it reveals a correspondence between the penalized spline smoother and the optimal predictor in a mixed model. This fact allows taking advantage of the methodology and software existent for mixed model analysis, and makes possible a simple implementation of otherwise complicated models. More specifically, the smoothing or penalty parameters are playing the role of a ratio of variances in the mixed model which suggests the application of maximum likelihood methodology for estimation (Kauermann, 2004) such as maximum likelihood estimation and restricted maximum likelihood estimation (REML). Both have the same merits of being based on the likelihood principle which leads to useful properties such as consistency, asymptotic normality, and efficiency. As pointed out by Verbeke and Molenberghs (2000, p. 46), the main advantage of the REML approach has been given by Patterson and Thompson (1971): in the absence of information on β, inference can be based only on a set of error contrasts rather than on the vector of dependent variables. The REML estimates for the variance components are identical to classical ANOVA-type estimates obtained from solving within a balanced panel data model the equations which set mean squares equal to their expectations. Furthermore, REML estimates do not rely on any normality assumption, only on moment assumptions. In short, with this approach, optimal minimum variance properties are attained.

Covariance components are obtained with REML of Patterson and Thompson (1971) by maximization of the logarithm of the likelihood function:

$$\log L_R = \frac{1}{2}\log|V| - \frac{1}{2}\log\left|X^T V^{-1} X\right| - \frac{1}{2} Y^T \left[V^{-1} - V^{-1} X \left(X^T V^{-1} X\right)^{-1} X^T V^{-1}\right] Y, \quad (6.9)$$

where:

V is a $(NT \times NT)$ matrix with $V = ZGZ^T + \delta_v^2 I_{NT}$ (6.10)

The REML covariance components can be estimated with the software functions successfully tested such as PROC MIXED (Littell *et al.*, 1996) in SAS (SAS Institute Inc, 2004) or lme(.) (Pinheiro and Bates, 2000) in S-PLUS (Insightful Corporation, 2003).

After minimizing the penalized least squares function, the BLUP of (β, ω) can be written as:

$$\begin{bmatrix} \tilde{\beta} \\ \tilde{\omega} \end{bmatrix} = \begin{bmatrix} \left(X^T \hat{V}^{-1} X \right)^{-1} X^T \hat{V}^{-1} Y \\ \hat{G} Z^T \hat{V}^{-1} \left(Y - X \tilde{\beta} \right) \end{bmatrix}. \tag{6.11}$$

There remains an unsolved question: how to determine the number and the value of the knots to be used in the model? The number of knots (*K*) should be large enough to ensure flexibility over the linear regression methods but smaller than the number of observations to obtain smoothness. The choice of *K* is discussed in Berry *et al.* (2002) and Ruppert (2002). For P-splines, it is argued that the exact value of *K* has little effect on the estimator, provided that *K* is at least a certain minimum value, because the amount of smoothing is determined not by *K* but rather by the penalty parameter. Nevertheless, as Ruppert *et al.* (2003, p. 177) point out, automatic smoothing methods like REML are somewhat erratic when we change the number of knots used. The amount of smoothing depends on the effective degrees of freedom, which varies with the degree of the spline and with the number of knots. Therefore, it must be used the full-search algorithm presented by Ruppert (2002) for *K*=10, *K*=20, *K*=30, *K*=40, *K*=50. The idea is to evaluate REML estimation results with different number of knots. After the application of this method, we decide to use a number of knots equal to 50 for *m*(.)* and for the technical efficiency component.

After determining K, the knot of order k is chosen in the univariate case as the sample quantile of the independent variable with probability $\left(k / K+1 \right)$.

b) Multivariate Smoothing

When x_{it} is a vector of the log of q inputs per worker $(q>1)$, the estimation of splines becomes a more difficult task, more specifically the determination of the knots and the basis functions to smooth the world production frontier and inefficiency. The method for determining the number of the knots is the same, but now the method presented by Ruppert (2002) implies the choice of K=50 for $m^*(.)$ and K=10 for the technical efficiency component. Furthermore, the choice of their value is more challenging, since it requires the use of an efficient space filling algorithm such as the one defined by Nychka and Saltzman (1998) and available on the S-PLUS (Insightful Corporation, 2003) and R (R Development Core Team, 2005) modules FUNFITS (Nychka *et al*., 1998) and FIELDS (Fields Development Team, 2004), respectively.

The purpose of the algorithm is to determine a subset of design points from a larger set of candidate points which not only serve as possible design points but determine the coverage criterion. A very intuitive example can be found in the FIELDS manual:

Suppose that you are charged with locating a fixed number of convenience stores in a city. Given a particular set of store locations each resident will have a store that is closest to their home. Out of all the residents, find the one who is farthest to their nearest store. This is the maximum over the nearest neighbor distances. A good design for the stores makes this criterion as small as possible. In words, one seeks to minimize the distance the residents must travel to their closest convenience store.

Fields Development Team (2004)

This example describes the concept of the minimax distance design. With the purpose of formalizing this concept, we consider the set of all candidate points Φ at which an experiment may be run and a distance function d defined on $\Phi \text{x} \Phi$. A finite number of sites, $|\Phi|$ will be assumed, where $|\Phi|$ denotes the cardinality of Φ. The distance function, d, satisfies the usual properties in order to define (Φ, d) as a metric space. Therefore, the distance between a point $\phi \in \Phi$ and a non-empty subset, $S \subset \Phi$, can be defined as $d(\phi,S) = \min_{s \in S} d(\phi,s)$. The minimax distance space-filling criterion chooses a subset of Φ so that the farthest points from the subset are made as close as possible to the subset:

$$d^* = \min\left\{\max_{\phi\in\Phi} d\left(\phi, S\right) \mid S \subset \Phi \; , \left|S\right| = K\right\}$$

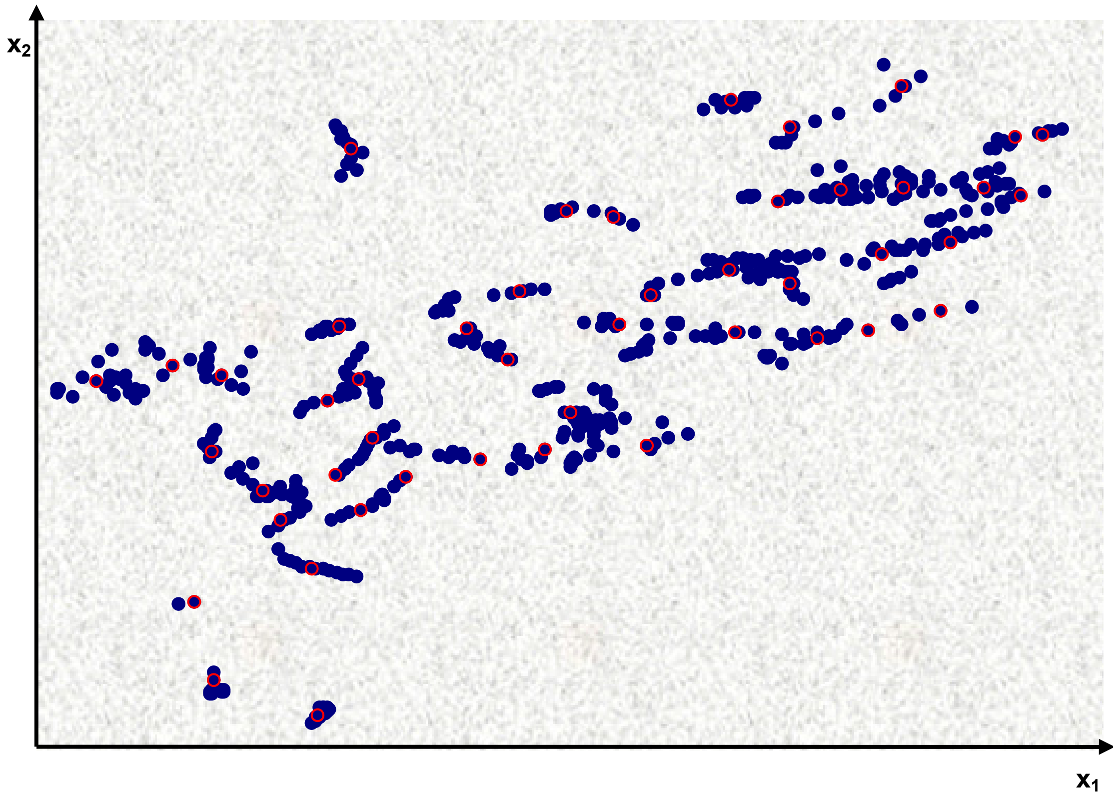
where:

K is fixed with $1 \leq K < |\Phi|$. $S^* \subset \Phi$ is defined to be a minimax distance design of size K provided $\max_{\phi\in\Phi} d\left(\phi, S^*\right) = d^*$ and $|S^*| = K$.

We use a swapping algorithm to solve this minimization problem, in the sense that for each design point, one checks whether a swap with a candidate point will yield a smaller coverage criterion. If this is the case, the substitution is made: the new point is adopted as part of the design and the old design point is moved into the candidate set. This process continues until no more productive swaps can be made or the number of iterations is exceeded.

For illustrative purposes, Figure 10 shows the knots chosen by applying the algorithm to the first panel for the agricultural sector with two inputs. Dots in blue represent all the candidate points. The circled red dots correspond to the 50 knots chosen for *m*(.)* by the minimax distance space-filling criterion.

Figure 10 - Application of the Swapping Algorithm



Regarding the construction of the multivariate basis functions, as pointed out by Wand (2003, p. 234), penalized spline regression can be performed in at least two ways: taking products of one-dimensional splines or using radial basis functions. We avoid the tensor product method for three reasons: dependence on the coordinate axes; improper for the estimation of nonlinear models; convergence problems due to the introduction of product terms between all covariates to both linear and truncated terms of the splines.

In this study, we employ radial basis functions to avoid the problems underlying the tensor product method. First, rotational invariance can be achieved through the utilization of radial functions. Second, as reported by Simpson *et al*. (2001), there is always a drawback when applying polynomial response surfaces to highly nonlinear or irregular models. On the other hand, radial basis function approximations have produced robust fits to arbitrary contours of both deterministic and stochastic response functions (Powell, 1987). Last, but not least, the addition of product terms between all covariates brings some empirical problems to the inference[23]. As pointed out by Wood (2006), all proposals for tensor product smoothing with a single penalty are unsatisfactory since they do not assure both smoothness range and scale invariance.

We decide to use the thin plate spline family of smoothers, which combines kriging with radial basis functions. Assuming $x_{it} = \left(x_{it}^1, ..., x_{it}^q\right) \in \mathbb{R}^q$ and $\kappa_k \in \mathbb{R}^q$, the penalized linear spline model for $m^*(.)$ in (6.1) is:

$$m^*(x_{it}) = \beta_1 x_{it}^1 + ... + \beta_q x_{it}^q + \sum_{k=1}^{K} w_k \left[\varsigma\left(\left\|x_{it} - \kappa_k\right\|\right)\right]\left[\varsigma\left(\left\|\kappa_k - \kappa_{k'}\right\|\right)\right]^{-1/2}, \qquad (6.12)$$

where:

$m^*(x_{it})$ is a piecewise linear function with K number of knots $\kappa_1, \kappa_2, ..., \kappa_K$; w_k is the *k*th knot coefficient, $w_k \sim N(0, \delta_w^2)$; $\varsigma(.)$ is the radial basis function, defined as:

$$\varsigma(r) = \begin{cases} \left\|r\right\|^{2h-q} & \Leftarrow \text{q odd} \\ \left\|r\right\|^{2h-q} \log\left\|r\right\| & \Leftarrow \text{q even} \end{cases}, \qquad (6.13)$$

where *h* is an integer satisfying $2h - q > 0$ that controls the smoothness of $\varsigma(.)$.

Similarly, the technical efficiency component in (6.2) can be specified as:

$$u_i^*(x_{it}) = \alpha_{i1} x_{it}^1 + ... + \alpha_{iq} x_{it}^q + \sum_{k=1}^{K} u_{ik} \left[\varsigma\left(\left\|x_{it} - \kappa_k\right\|\right)\right]\left[\varsigma\left(\left\|\kappa_k - \kappa_{k'}\right\|\right)\right]^{-1/2}, \qquad (6.14)$$

where:

$\alpha_{i1} \sim N^-(0, \delta_{\alpha_1}^2), ..., \alpha_{iq} \sim N^-(0, \delta_{\alpha_q}^2)$; $u_{ik} \sim N^-(0, \delta_u^2)$.

The linear Gaussian mixed model can be defined as:

$$Y = X\beta + Z_R \omega + v, \qquad (6.15)$$

[23] We try to estimate our model with the tensor product method and convergence was not reached.

where:

$$Y=\begin{bmatrix} Y_1 \\ . \\ . \\ . \\ Y_N \end{bmatrix},\ Y_i=\begin{bmatrix} y_{i1} \\ . \\ . \\ . \\ y_{iT} \end{bmatrix};\ X_i=\begin{bmatrix} x_{i1}^1 \cdots x_{i1}^q \\ \vdots \ \ddots \ \vdots \\ x_{iT}^1 \cdots x_{iT}^q \end{bmatrix};\ \beta=\begin{bmatrix} \beta_1 & \dots & \beta_q \end{bmatrix}^T;$$

$$Z_R=\begin{bmatrix} Z_1 & X_1 & 0 & \cdots & 0 & Z_1 & 0 & \cdots & 0 \\ Z_2 & 0 & X_2 & \cdots & 0 & 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_N & 0 & 0 & \cdots & X_N & 0 & 0 & \cdots & Z_N \end{bmatrix};$$

$$Z_i=\begin{bmatrix} \varsigma\left(\left\|x_{i1}-\kappa_1\right\|\right) & \cdots & \varsigma\left(\left\|x_{i1}-\kappa_K\right\|\right) \\ \vdots & \ddots & \vdots \\ \varsigma\left(\left\|x_{iT}-\kappa_1\right\|\right) & \cdots & \varsigma\left(\left\|x_{iT}-\kappa_K\right\|\right) \end{bmatrix}\square\Omega^{-1/2};$$

$$\Omega=\begin{bmatrix} 1 & \varsigma\left(\left\|\kappa_1-\kappa_2\right\|\right) & \cdots & \varsigma\left(\left\|\kappa_1-\kappa_K\right\|\right) \\ \varsigma\left(\left\|\kappa_2-\kappa_1\right\|\right) & 1 & \cdots & \varsigma\left(\left\|\kappa_2-\kappa_K\right\|\right) \\ \vdots & \vdots & \ddots & \vdots \\ \varsigma\left(\left\|\kappa_K-\kappa_1\right\|\right) & \varsigma\left(\left\|\kappa_K-\kappa_2\right\|\right) & \cdots & 1 \end{bmatrix};$$

$$\omega=\begin{bmatrix} w & \alpha_1 & \dots & \alpha_N & u_1 & \dots & u_N \end{bmatrix}^T;w=\begin{bmatrix} w_1 & \dots & w_K \end{bmatrix};\alpha_i=\begin{bmatrix} \alpha_{i1} & \dots & \alpha_{iq} \end{bmatrix};$$

$$u_i=\begin{bmatrix} u_{i1} & \dots & u_{iK} \end{bmatrix};\ v=\begin{bmatrix} v_{11} \\ \vdots \\ v_{NT} \end{bmatrix};$$

$$Cov\left(\omega\right)=G=\begin{bmatrix} \delta_w^2 I_K & 0 & \cdots & 0 & 0 \\ 0 & \delta_{\alpha_1}^2 I_N & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \delta_{\alpha_q}^2 I_N & 0 \\ 0 & 0 & \cdots & 0 & \delta_u^2 I_{NK} \end{bmatrix}.$$

Now, X is a matrix of dimension $NT \times q$, β is a $q \times 1$ vector; Z is a matrix of dimension $NT \times (K + Nq + NK)$; ω is a $(K + Nq + NK) \times 1$ matrix and G a symmetric matrix of dimension $(K + Nq + NK) \times (K + Nq + NK)$

The use of $\Omega^{-1/2}$ in the definition of Z_R is one of the possible ways to assure that $Cov(\omega)$ is a positive definite covariance matrix, allowing the estimation by standard mixed effects methods. This formulation might seem arbitrary and with no justification other than its symmetry, but it can be shown that it corresponds to the thin plate spline family of smoothers presented in Green and Silverman (1994) with parameter *h* controlling smoothness. For more details, please refer to Wood (2003, 2006) and Ruppert *et al.* (2003, pp. 248-254).

The resulting penalized least squares function can be expressed by:

$$\min_{\beta,\omega} \left(Y - X\beta - Z_R\omega\right)^T \left(Y - X\beta - Z_R\omega\right) + \frac{\delta_v^2}{\delta_w^2}\omega^T P_1 \omega + \frac{\delta_v^2}{\delta_u^2}\omega^T P_2 \omega \ ,$$

where:

$$P_1 = \begin{bmatrix} I_K & 0_{\mathrm{K}\times(\mathrm{Nq+NK})} \\ 0_{(\mathrm{Nq+NK})\times \mathrm{K}} & 0_{(\mathrm{Nq+NK})\times(\mathrm{Nq+NK})} \end{bmatrix};\ P_2 = \begin{bmatrix} 0_{(\mathrm{K+Nq})\times(\mathrm{K+Nq})} & 0_{(\mathrm{K+Nq})\times \mathrm{NK}} \\ 0_{\mathrm{NK}\times(\mathrm{K+Nq})} & I_{\mathrm{NK}} \end{bmatrix};$$

$0_{\mathrm{r\times s}}$ is a $(r \times s)$ matrix of zeros; P_1 and P_2 are designed in a way to penalize the coefficients of the radial basis functions.

As in (6.9), covariance components are obtained with REML of Patterson and Thompson (1971) by maximization of the logarithm of the likelihood function:

$$\log L_R = \frac{1}{2}\log\left|V_R\right| - \frac{1}{2}\log\left|X^T V_R^{-1} X\right| - \frac{1}{2} Y^T \left[V_R^{-1} - V_R^{-1} X \left(X^T V_R^{-1} X\right)^{-1} X^T V_R^{-1}\right] Y \quad (6.16)$$

where:

$V_R = Z_R G Z_R^T + \delta_v^2 I_{NT}$.

The resulting BLUP estimates of (β, ω) can be written as:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\omega} \end{bmatrix} = \begin{bmatrix} \left(X^T \hat{V}_R^{-1} X\right)^{-1} X^T \hat{V}_R^{-1} Y \\ \hat{G} Z_R{}^T \hat{V}_R^{-1} \left(Y - X\hat{\beta}\right) \end{bmatrix}.$$

Summarizing, all parameters in the multivariate case can be estimated by carrying out the following steps:

(1) Determine the number of knots using $K = \max\left\{10, \min\left[50, round\left(NT/4\right)\right]\right\}$.

(2) Knots are obtained with the application of an efficient space filling algorithm such as the one defined by Nychka and Saltzman (1998) and available on the S and R modules FUNFITS (Nychka *et al.*, 1998) and FIELDS (Fields Development Team, 2004).

(3) Form the matrices Ω and Z_R. Since for agricultural sector, q=2:

$$\Omega = \begin{bmatrix} 1 & \|\kappa_1 - \kappa_2\|^{2h-2} \log\|\kappa_1 - \kappa_2\| & \cdots & \|\kappa_1 - \kappa_K\|^{2h-2} \log\|\kappa_1 - \kappa_K\| \\ \|\kappa_2 - \kappa_1\|^{2h-2} \log\|\kappa_2 - \kappa_1\| & 1 & \cdots & \|\kappa_2 - \kappa_K\|^{2h-2} \log\|\kappa_2 - \kappa_K\| \\ \vdots & \vdots & \ddots & \vdots \\ \|\kappa_K - \kappa_1\|^{2h-2} \log\|\kappa_K - \kappa_1\| & \|\kappa_K - \kappa_2\|^{2h-2} \log\|\kappa_K - \kappa_2\| & \cdots & 1 \end{bmatrix};$$

$$Z_R = \begin{bmatrix} Z_1 & X_1 & 0 & \cdots & 0 & Z_1 & 0 & \cdots & 0 \\ Z_2 & 0 & X_2 & \cdots & 0 & 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_N & 0 & 0 & \cdots & X_N & 0 & 0 & \cdots & Z_N \end{bmatrix};$$

$$Z_i = \begin{bmatrix} \|x_{i1} - \kappa_1\|^{2h-2} \log\|x_{i1} - \kappa_1\| & \cdots & \|x_{i1} - \kappa_K\|^{2h-2} \log\|x_{i1} - \kappa_K\| \\ \vdots & \ddots & \vdots \\ \|x_{iT} - \kappa_1\|^{2h-2} \log\|x_{iT} - \kappa_1\| & \cdots & \|x_{iT} - \kappa_K\|^{2h-2} \log\|x_{iT} - \kappa_K\| \end{bmatrix} \Omega^{-1/2}.$$

(4) Use mixed model computer functions such as PROC MIXED (Littell *et al.*, 1996) in SAS (SAS Institute Inc, 2004) or lme(.) (Pinheiro and Bates, 2000) in S-

PLUS (Insightful Corporation, 2003) to fit the mixed model $Y = X\beta + Z_R\omega + v$, in which covariance components are chosen via REML.

(5) The resulting BLUP estimates $\begin{bmatrix} \hat{\beta} \\ \hat{\omega} \end{bmatrix}$ can be used to compute $\hat{Y} = X\hat{\beta} + Z_R\hat{\omega}$.

6.2.2 - The Bayesian Approach

a) Introduction

Advances in machine learning and data mining changed the focus of applied statistics from choosing simple tractable frameworks to building computationally demanding models that are expected to present a better fit to the data. One example is the combination of the MCMC algorithm with Bayesian analysis for models that were previously intractable. This technique produces a correlated sample from the joint posterior distribution of parameters given the data, allowing statistical inference about all parameters. For clarity of exposition, we will present some introductory concepts, according to some classical textbook references (Gilks *et al*., 1996; Tanner, 1996; Gamerman, 1997).

A stochastic process is a procedure by which a system moves through a series of well-defined states in a way that exhibits some element of randomness. The set of distinct values assumed by a stochastic process is called the state space. The state space may have countably many or uncountably many members. In the first case, we have a discrete parameter stochastic process, also designated as chain; otherwise, it is a continuous parameter stochastic process.

A discrete parameter stochastic process is a Markov process if it has no memory; i.e., the probability that the system moves into a particular state depends only upon the state it is currently in, and not on the history of the past visitations of states. Therefore, a Markov process can be fully specified via a set of transition probabilities that describe the likelihood with which the system moves into a state given the current state. The Metropolis-Hastings algorithm is a class of Markov chain which new states are selected according to any convenient transition probability matrix. Each state is accepted with a probability that ensures that the overall transition probability is consistent with the desired limiting distribution.

The Gibbs sampler (introduced in the context of image processing by Geman and Geman, 1984) is a special case of Metropolis-Hastings sampling wherein the random value is always accepted. The task remains how to construct a Markov Chain whose values converge to the target distribution. The key to the Gibbs sampler is that only conditional distributions are considered, i.e., distributions where all random variables, but one, are assigned fixed values. Such conditional distributions are far easier to simulate than complex joint distributions and usually have common prior distributions. Thus, one simulates n random variables sequentially from the n univariate conditionals rather than generating a single n-dimensional vector in a single step using the full joint distribution. Repeating this process k times, generates a Gibbs sequence of length k, where a subset of points are taken as our simulated draws from the full joint distribution. To obtain the desired total of m sample points, the chain is sampled after a sufficient burn-in period to remove the effects of the initial sampling values and at every n samples following the burn-in. The Gibbs sequence converges to a stationary distribution that is independent of the starting values, and by construction this stationary distribution is the target distribution we are trying to simulate (Tierney 1994).

We can conclude that, given a model and a sample, MCMC produces a correlated sample from a sequence of distributions having as stationary distribution the joint posterior distribution of the parameters. Using this correlated sample the entire posterior density of one parameter given the data is estimated. In some models, maximum likelihood produces biased estimates or faces convergence difficulties. MCMC overcomes these problems.

b) Bayesian P-splines Using WINBUGS: Univariate Smoothing

Please recall the general mixed model presentation in (6.6) for the univariate case:

$$Y = X\beta + Z\omega + v, \qquad Cov(\omega) = G = \begin{bmatrix} \delta_w^2 I_K & 0 & 0 \\ 0 & \delta_\alpha^2 I_N & 0 \\ 0 & 0 & \delta_u^2 I_{NK} \end{bmatrix}.$$

In a Bayesian model, all parameters are random and prior distributions must be assigned in order to simulate the posterior distribution of them. The parameter vector of the linear Gaussian mixed model includes (β, ω) and the covariance components. The prior on ω is intrinsically specified by the model as a normally distributed function with $\mathrm{E}(\omega) = 0$ and $Cov(\omega) = G$. Regarding β, the usual choice for a proper prior is a normal distribution with an extremely large standard deviation that may cover a very wide range. For example if we take $\beta \sim N(0,1000000)$, then we are allowing a 95% prior credible interval between -2000 and +2000.

Priors for the standard deviation parameters are a bit more difficult. Either we can set a flat prior for the variance, over a fixed range which gives a Pareto prior for the precision or we can use a member of the Gamma family for the precision, which is the conjugate prior for this problem. For more details, please refer to Gelman and Rubin (1992) and Carlin (1992). We assume that the prior distribution for each covariance parameter is the inverse gamma:

$$\delta_v^2 \sim IG(A_v, B_v); \delta_w^2 \sim IG(A_w, B_w); \delta_\alpha^2 \sim IG(A_\alpha, B_\alpha); \delta_u^2 \sim IG(A_u, B_u) \ .$$

Regarding the conditional posterior of (β, ω) given all variance components of G and the error variance δ_v^2, also known as complete conditional, Berry *et al.* (2002) and Ruppert *et al*. (2003) prove that:

$$\left[\beta, \omega \middle| \delta_v^2, \delta_w^2, \delta_\alpha^2, \delta_u^2, Y\right] \sim N\left[\left(C^T C + \delta_v^2 F\right)^{-1} C^T Y, \delta_v^2 \left(C^T C + \delta_v^2 F\right)^{-1}\right], \qquad (6.17)$$

where:

C is a $\left[NT \times (1 + K + N + NK)\right]$ matrix defined as $C = \left[X \quad Z\right]$,

$$F = \begin{bmatrix} 0_{1x1} & 0_{1x(K+N+NK)} \\ 0_{(K+N+NK)x1} & G^{-1} \end{bmatrix},$$

$$Cov(\omega) = \begin{bmatrix} \delta_w^2 I_K & 0 & 0 \\ 0 & \delta_\alpha^2 I_N & 0 \\ 0 & 0 & \delta_u^2 I_{NK} \end{bmatrix}.$$

Berry *et al*. (2002, p. 164) and Ruppert *et al.* (2003, p. 281) calculate the conditional posterior of each variance component given all other data:

$$\delta_w^2 \sim IG\left(A_w + 0.5K, B_w + 0.5|w_1|^2 + \cdots + 0.5|w_K|^2\right). \tag{6.18}$$

$$\delta_\alpha^2 \sim IG\left(A_\alpha + 0.5N, B_\alpha + 0.5|\alpha_{11}|^2 + \cdots + 0.5|\alpha_{N1}|^2\right). \tag{6.19}$$

$$\delta_u^2 \sim IG\left(A_u + 0.5NK, B_u + 0.5|u_{11}|^2 + \cdots + 0.5|u_{NK}|^2\right). \tag{6.20}$$

$$\delta_v^2 \sim IG\left(A_v + 0.5NT, B_v + 0.5\|Y - X\beta - Z\omega\|^2\right), \tag{6.21}$$

where $\|x\|$ refers to the Euclidean norm of a vector and $|x_i|$ to the absolute value of the i-element of the vector.

As pointed out by Ruppert *et al.* (2003, p. 281), as part of the MCMC chain, (β, ω) are generated from the current values of the variance components according to the multivariate normal distribution defined in (6.17). It is possible to iterate the MCMC between sampling the regression coefficients given all variance components and vice-versa (all conditional on the data, *Y*). The MCMC chain needs a starting value for all variance parameters. Since given (β, ω), each variance component is independent from the others, the starting point is indifferent for the algorithm results. Following Ruppert *et al.* (2003, p. 280) and Carroll *et al.* (2006, p. 305), we set up both parameters of the inverse gamma distributions in (6.18)-(6.21) close to zero (0.001), in order to obtain a non-informative, but proper, prior. We define large starting values for the variance components $\left(\delta_v^2 = \delta_w^2 = \delta_\alpha^2 = \delta_u^2 = 100\right)$, although a specific choice of starting values is not important, since the chain converges quickly to the stationary distribution and the burn-in period discards the beginning of the chain. Convergence to the posterior distributions is assessed by using several initial values (including the model parameters for the frequentist version presented in the last section) and visually inspecting several chains corresponding to the model parameters.

The model is estimated with the software WINBUGS (Spiegelhalter *et al.*, 2003). We discard the first 30000 burn-in simulations. For inference we use 50000 simulations and we monitor all parameters of the model. For each one, we report the 95% equal tail probability credible interval and the posterior mean and median. Since the parameters of the spline tend to be weakly identified with poor mixing properties, we use the posterior mean (which tends to be well identified and with good asymptotical properties) as the point estimator. Furthermore, Pérez and Quintana (2003) show that adaptive Bayesian estimators are rarely unbiased and that using the posterior mean as a point estimator also yields better estimates than using the posterior median or mode.

The chain can be summarized in the following way:

(1) Define $\delta_v^2 = \delta_w^2 = \delta_\alpha^2 = \delta_u^2 = 100$ and $A_v = B_v = A_w = B_w = A_\alpha = B_\alpha = A_u = B_u = 0,001.$

(2) Sample (β, ω) from the multivariate normal distribution $N\left[\left(C^T C + \delta_v^2 F\right)^{-1} C^T Y, \delta_v^2 \left(C^T C + \delta_v^2 F\right)^{-1}\right].$

(3) Sample δ_w^2 from $IG\left(A_w + 0.5K, B_w + 0.5|w_1|^2 + \cdots + 0.5|w_K|^2\right).$

(4) Sample δ_α^2 from $IG\left(A_\alpha + 0.5N, B_\alpha + 0.5|\alpha_{11}|^2 + \cdots + 0.5|\alpha_{N1}|^2\right).$

(5) Sample δ_u^2 from $IG\left(A_u + 0.5NK, B_u + 0.5|u_{11}|^2 + \cdots + 0.5|u_{NK}|^2\right).$

(6) Sample δ_v^2 from $IG\left(A_v + 0.5NT, B_v + 0.5\left\|Y - X\beta - Z\omega\right\|^2\right)$.

(7) Return to (2) and iterate.

c) Bayesian P-splines Using WINBUGS: Multivariate Smoothing

Please recall the general mixed model presentation in (6.15) for the multivariate case, with q (>1) inputs:

$$Y = X\beta + Z_R\omega + v,$$

where:

$$\omega = \begin{bmatrix} w & \alpha_1 & \dots & \alpha_N & u_1 & \dots & u_N \end{bmatrix}^T; w = \begin{bmatrix} w_1 & \dots & w_K \end{bmatrix}; \alpha_i = \begin{bmatrix} \alpha_{i1} & \dots & \alpha_{iq} \end{bmatrix};$$

$$u_i = \begin{bmatrix} u_{i1} & \dots & u_{iK} \end{bmatrix};\ Cov(\omega) = G = \begin{bmatrix} \delta_w^2 I_K & 0 & \cdots & 0 & 0 \\ 0 & \delta_{\alpha_1}^2 I_N & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \delta_{\alpha_q}^2 I_N & 0 \\ 0 & 0 & \cdots & 0 & \delta_u^2 I_{NK} \end{bmatrix}.$$

Z_R is a design matrix obtained outside MCMC and entered as data collected by the frequentist version of the model. Compared with the univariate case, now we have *(q-1)* additional covariance parameters, with the same assumption for priors:

$$\delta_v^2 \sim IG(A_v, B_v); \delta_w^2 \sim IG(A_w, B_w); \delta_{\alpha_1}^2 \sim IG\left(A_{\alpha_1}, B_{\alpha_1}\right); \dots; \delta_{\alpha_q}^2 \sim IG\left(A_{\alpha_q}, B_{\alpha_q}\right);$$

$$\delta_u^2 \sim IG(A_u, B_u).$$

The conditional posterior of (β, ω) given all variance components of G and the error variance δ_v^2 is defined in a similar way as in the univariate case. This distribution is defined as follows:

$$\left[\beta, \omega \middle| \delta_v^2, \delta_w^2, \delta_\alpha^2, \delta_u^2, Y\right] \sim N\left[\left(C^T C + \delta_v^2 F\right)^{-1} C^T Y, \delta_v^2 \left(C^T C + \delta_v^2 F\right)^{-1}\right], \quad (6.22)$$

where:

C is a $\left[NT \times \left(q+K+Nq+NK\right)\right]$ matrix defined as $C=\left[X \quad Z_R\right]$,

$$F=\begin{bmatrix} 0_{q\mathrm{x}q} & 0_{q\mathrm{x}\left(\mathrm{K+Nq+NK}\right)} \\ 0_{\left(\mathrm{K+Nq+NK}\right)\mathrm{x}q} & G^{-1} \end{bmatrix}.$$

In the same way, model is estimated with WINBUGS (Spiegelhalter *et al.*, 2003), discarding the first 30000 burn-in simulations. For inference we use 50000 simulations and the posterior mean as the point estimator, for the reason already described.

The chain can be summarized in the following way:

(1) Define $\delta_v^2=\delta_w^2=\delta_{\alpha_1}^2=...=\delta_{\alpha_q}^2=\delta_u^2=100$ and $A_v=B_v=A_w=B_w=A_{\alpha_1}=B_{\alpha_1}=...=A_{\alpha_q}=B_{\alpha_q}=A_u=B_u=0{,}001.$

(2) Sample $\left(\beta,\omega\right)$ from the multivariate normal distribution $N\left[\left(C^TC+\delta_v^2F\right)^{-1}C^TY,\delta_v^2\left(C^TC+\delta_v^2F\right)^{-1}\right]$ with $C=\left[X \quad Z_R\right]$ and

$$F=\begin{bmatrix} 0_{q\mathrm{x}q} & 0_{q\mathrm{x}\left(\mathrm{K+Nq+NK}\right)} \\ 0_{\left(\mathrm{K+Nq+NK}\right)\mathrm{x}q} & G^{-1} \end{bmatrix}.$$

(3) Sample δ_w^2 from $IG\left(A_w+0.5K,B_w+0.5\left|w_1\right|^2+\cdots+0.5\left|w_K\right|^2\right)$.

(4.1) Sample $\delta_{\alpha_1}^2$ from $IG\left(A_{\alpha_1}+0.5N,B_{\alpha_1}+0.5\left|\alpha_{11}\right|^2+\cdots+0.5\left|\alpha_{N1}\right|^2\right)$.

...

(4.q) Sample $\delta_{\alpha_q}^2$ from $IG\left(A_{\alpha_q}+0.5N,B_{\alpha_q}+0.5\left|\alpha_{1q}\right|^2+\cdots+0.5\left|\alpha_{Nq}\right|^2\right)$.

(5) Sample δ_u^2 from $IG\left(A_u+0.5NK,B_u+0.5\left|u_{11}\right|^2+\cdots+0.5\left|u_{NK}\right|^2\right)$.

(6) Sample δ_v^2 from $IG\left(A_v + 0.5NT, B_v + 0.5\|Y - X\beta - Z_R\omega\|^2\right)$.

(7) Return to (2) and iterate.

6.3 - The Determinants of Labor Productivity Growth Across Countries

We define a CRS reference technology with one aggregate output, *Y*, and a *K*-dimensional vector of inputs, *X*. The CRS hypothesis allows us to transform the dependent variable in labor productivity, *y*, and the vector *X* into the *(K-1)*-dimensional vector of inputs per worker, *x*. For the economy as a whole, *K*=2 and *X*=(labor, capital); and for agriculture, *K*=3 and *X*=(labor, land, capital).

Labor productivity decomposition is performed using equation (5.18) and counterfactual distributions by (4.1).

As in chapter 5, each production frontier is estimated separately for the time periods 1967-1979 and 1980-1992.

6.3.1 - Economy as a Whole

Splines models estimate numerous parameters. All of them have the correct signs and small standard errors. We detect small differences between the Classical and the Bayesian approach. As expected, in the latest framework, mean and mode values are very similar. In table 29, we illustrate these findings with estimation results for the fixed part of *m*(.)*.

Table 29.1 - Estimation Results for the Fixed Part of m*(.) Using a Frequentist Approach (Overall Economy).

First Period: 1967-1979			
	Value	Std.Error	p-value
ln k_{it}	0.858	0.005	<.0001
Second Period: 1980-1992			
	Value	Std.Error	p-value
ln k_{it}	0.855	0.005	<.0001

Table 29.2 - Estimation Results for the Fixed Part of m*(.) Using a MCMM Approach (Overall Economy).

First Period: 1967-1979				
node	Mean	Std. Error	MC error	Median
Ln kit	0.857	0.006	0.0003	0.8578
Second Period: 1980-1992				
Node	Mean	Std. Error	MC error	Median
ln kit	0.8495	0.007	0.0004	0.8494

Table 30 presents the decomposition of labor productivity growth for both classical and Bayesian approaches. Figures 11, 12 and 13 refer to the counterfactual distributions of output per worker. Table 31 indicates the Li's distribution hypothesis tests presented earlier.

Table 30.1 - Decomposition of Labor Productivity Growth for the Economy Using a Frequentist Approach

Country (by ascending order of output per worker in the first period)	Percentage Change in Output per Worker	Contribution to Percentage Change in Output per Worker of			
		Change in Efficiency	Change in Technology	Capital Deepening	Stochastic Shocks
Malawi	22,55%	-68,36%	248,55%	-0,04%	11,17%
Kenya	15,74%	-49,44%	141,75%	-0,21%	-5,10%
Indonesia	248,48%	-9,48%	159,26%	15,53%	28,53%
India	95,07%	-40,30%	149,80%	-0,25%	31,12%
Madagascar	-33,85%	-86,81%	372,69%	2,69%	3,32%
Zimbabwe	-6,22%	-41,40%	81,01%	-11,01%	-0,66%
Pakistan	66,21%	-27,57%	92,46%	-9,42%	31,63%
Sri Lanka	78,95%	-21,97%	88,18%	7,72%	13,14%
Korea, Republic of	422,65%	18,35%	40,31%	188,21%	9,21%
Philippines	24,41%	-43,31%	94,19%	-4,31%	18,11%
Egypt	92,30%	-16,49%	76,33%	3,96%	25,62%
Honduras	15,56%	-32,98%	71,10%	-6,86%	8,21%
Turkey	107,44%	-11,61%	30,73%	52,02%	18,09%
Morocco	53,10%	-14,96%	40,10%	21,94%	5,38%
Tunisia	103,62%	21,85%	8,58%	25,94%	22,20%
Dominican Republic	40,71%	-19,42%	40,58%	14,94%	8,07%
El Salvador	3,22%	-35,20%	35,22%	-1,75%	19,90%
Guatemala	23,56%	-33,54%	69,16%	-4,37%	14,92%
Colombia	31,98%	-8,06%	34,05%	1,12%	5,89%
Portugal	160,32%	9,42%	7,41%	108,66%	6,15%
Syrian Arab Republic	126,30%	-6,66%	13,15%	64,85%	29,98%
South Africa	17,02%	-0,52%	-4,62%	30,91%	-5,79%
Greece	108,70%	14,79%	-1,61%	64,62%	12,25%
Costa Rica	15,18%	-3,14%	0,46%	8,64%	8,97%
Japan	167,57%	8,22%	-5,65%	144,81%	7,04%
Peru	-33,56%	-13,25%	-2,21%	0,32%	-21,93%
Uruguay	30,30%	10,32%	-1,17%	3,64%	15,32%
Chile	24,22%	21,40%	-4,36%	-9,77%	18,56%
Israel	108,47%	5,94%	3,11%	39,47%	36,83%
Argentina	19,13%	-19,49%	-1,73%	54,90%	-2,79%
Finland	71,42%	5,45%	1,71%	40,55%	13,72%
Austria	84,33%	0,85%	-2,65%	73,70%	8,09%
Italy	95,40%	4,64%	0,23%	70,31%	9,39%
United Kingdom	51,35%	-7,37%	3,04%	44,25%	9,93%
Denmark	36,22%	-9,78%	4,11%	39,53%	3,94%
France	63,81%	1,84%	3,41%	33,77%	16,28%
Norway	67,40%	8,92%	1,88%	45,97%	3,35%
Sweden	26,29%	-2,00%	5,91%	20,42%	1,05%
Netherlands	38,77%	-6,47%	4,74%	32,98%	6,51%
Australia	35,76%	-2,18%	9,53%	14,78%	10,39%
New Zealand	4,80%	-3,44%	5,74%	-0,61%	3,26%
Canada	39,33%	4,59%	9,69%	13,94%	6,59%
USA	23,76%	-4,25%	11,19%	4,12%	11,64%
Mean	**64,83%**	**-11,69%**	**45,01%**	**28,85%**	**11,10%**

Table 30.2 - Decomposition of Labor Productivity Growth for the Economy Using a MCMC Approach

Country (by ascending order of output per worker in the first period)	Percentage Change in Output per Worker	Contribution to Percentage Change in Output per Worker of			
		Change in Efficiency	Change in Technology	Capital Deepening	Stochastic Shocks
Malawi	22,55%	-64,41%	213,90%	-1,40%	11,25%
Kenya	15,74%	-46,36%	125,91%	0,52%	-4,97%
Indonesia	248,48%	-8,08%	142,18%	21,70%	28,64%
India	95,07%	-36,14%	132,50%	0,25%	31,06%
Madagascar	-33,85%	-84,30%	319,91%	-2,85%	3,30%
Zimbabwe	-6,22%	-39,20%	76,00%	-11,50%	-0,98%
Pakistan	66,21%	-23,97%	85,74%	-10,27%	31,16%
Sri Lanka	78,95%	-19,51%	82,21%	8,00%	12,97%
Korea, Republic of	422,65%	21,26%	36,03%	191,51%	8,70%
Philippines	24,41%	-40,86%	87,28%	-4,83%	18,01%
Egypt	92,30%	-14,80%	71,95%	4,28%	25,87%
Honduras	15,56%	-30,99%	67,28%	-7,38%	8,08%
Turkey	107,44%	-11,84%	29,19%	53,92%	18,34%
Morocco	53,10%	-14,77%	38,30%	23,04%	5,56%
Tunisia	103,62%	21,86%	8,24%	26,30%	22,23%
Dominican Republic	40,71%	-18,98%	38,98%	15,76%	7,94%
El Salvador	3,22%	-34,47%	34,25%	-1,81%	19,49%
Guatemala	23,56%	-31,90%	65,59%	-4,73%	15,02%
Colombia	31,98%	-7,59%	33,17%	1,20%	5,97%
Portugal	160,32%	12,11%	5,73%	104,39%	7,45%
Syrian Arab Republic	126,30%	-6,73%	13,06%	65,37%	29,78%
South Africa	17,02%	0,10%	-4,73%	29,63%	-5,34%
Greece	108,70%	17,90%	-2,81%	60,89%	13,20%
Costa Rica	15,18%	-3,20%	0,51%	8,65%	8,97%
Japan	167,57%	15,67%	-8,39%	134,87%	7,50%
Peru	-33,56%	-13,36%	-2,08%	0,33%	-21,94%
Uruguay	30,30%	10,28%	-1,17%	3,39%	15,63%
Chile	24,22%	20,75%	-3,80%	-9,41%	18,03%
Israel	108,47%	11,45%	0,26%	35,79%	37,40%
Argentina	19,13%	-17,75%	-2,83%	51,33%	-1,50%
Finland	71,42%	12,24%	-2,30%	38,18%	13,14%
Austria	84,33%	7,50%	-5,85%	67,90%	8,47%
Italy	95,40%	11,16%	-2,36%	64,76%	9,26%
United Kingdom	51,35%	-2,27%	0,13%	40,64%	9,97%
Denmark	36,22%	-4,51%	0,63%	36,19%	4,09%
France	63,81%	8,05%	-0,92%	31,85%	16,06%
Norway	67,40%	15,68%	-3,12%	43,83%	3,85%
Sweden	26,29%	3,81%	1,04%	19,09%	1,10%
Netherlands	38,77%	-0,94%	0,86%	30,90%	6,10%
Australia	35,76%	3,68%	4,00%	14,29%	10,17%
New Zealand	4,80%	0,53%	1,35%	-0,65%	3,52%
Canada	39,33%	10,67%	4,02%	13,40%	6,73%
USA	23,76%	1,48%	5,34%	4,68%	10,60%
Mean	**64,83%**	**-8,62%**	**39,19%**	**27,72%**	**11,16%**

Figure 11.1 - Counterfactual Distributions of Output per Worker for the Economy Using a Frequentist Approach

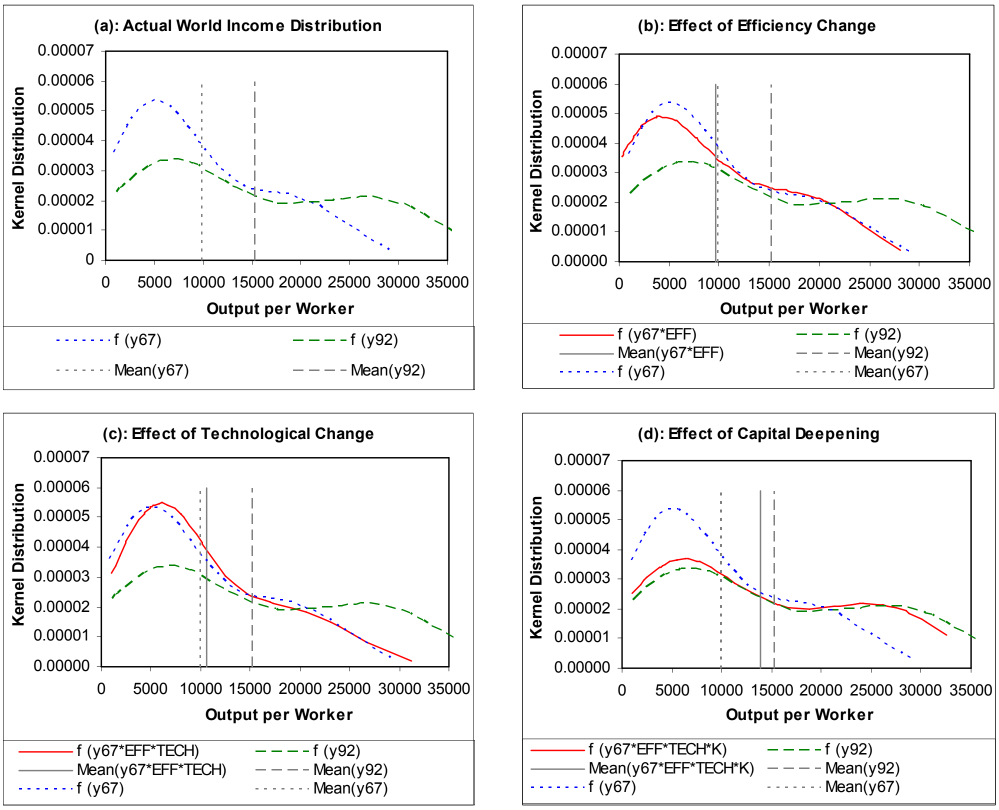


Figure 11.2 - Counterfactual Distributions of Output per Worker for the Economy Using a MCMC Approach

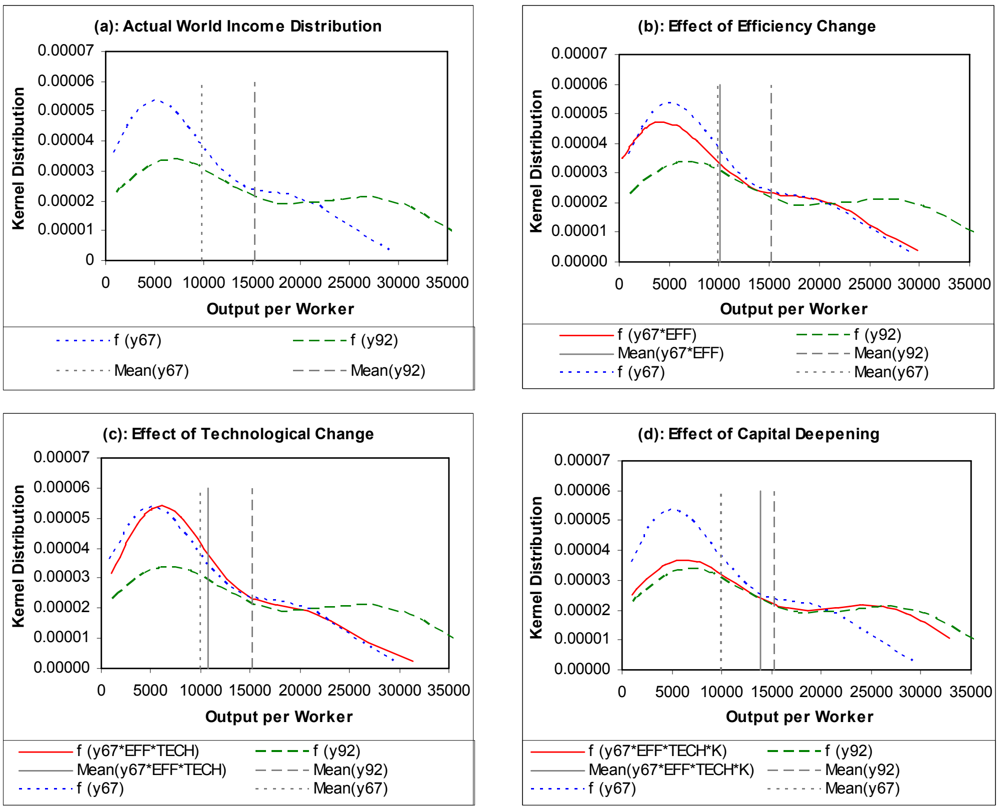


Figure 12.1 - Counterfactual Distributions of Output per Worker for the Economy Using a Frequentist Approach

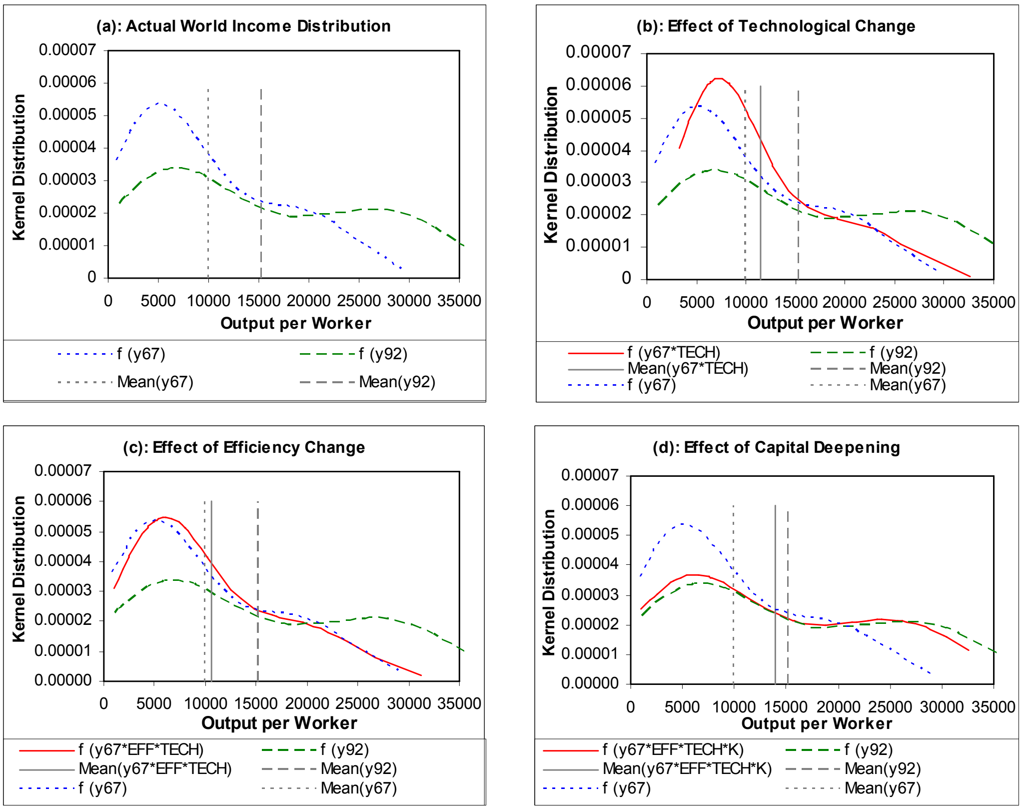


Figure 12.2 - Counterfactual Distributions of Output per Worker for the Economy Using a MCMC Approach

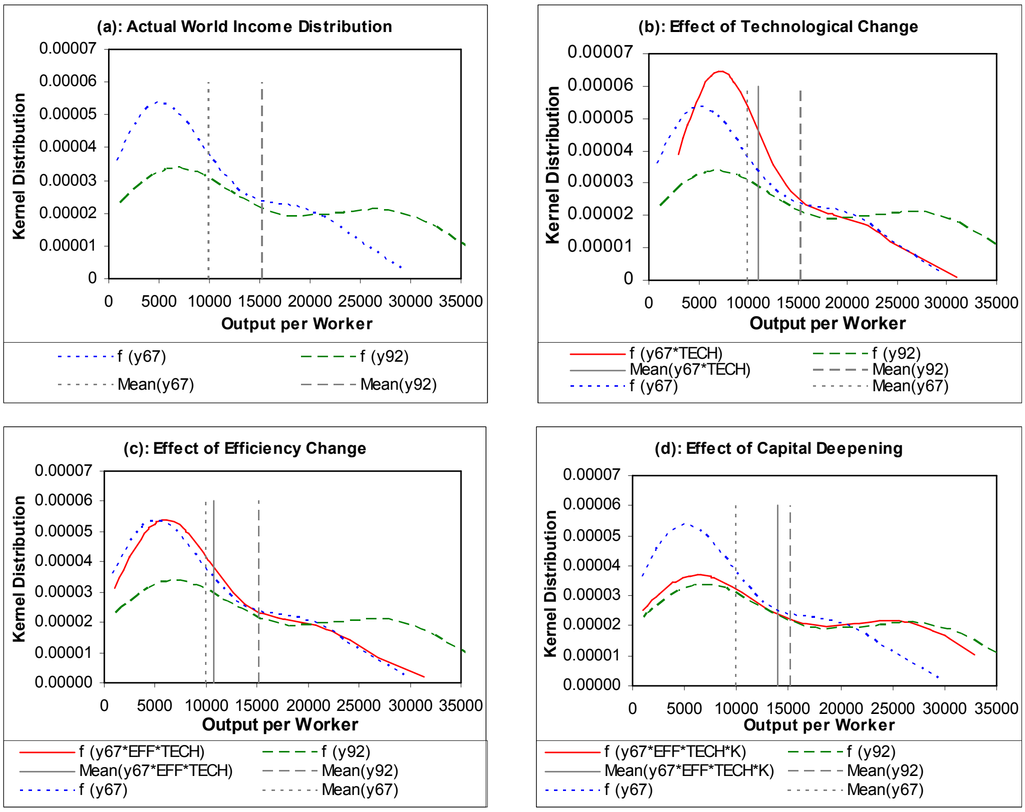


Figure 13.1 - Counterfactual Distributions of Output per Worker for the Economy Using a Frequentist Approach

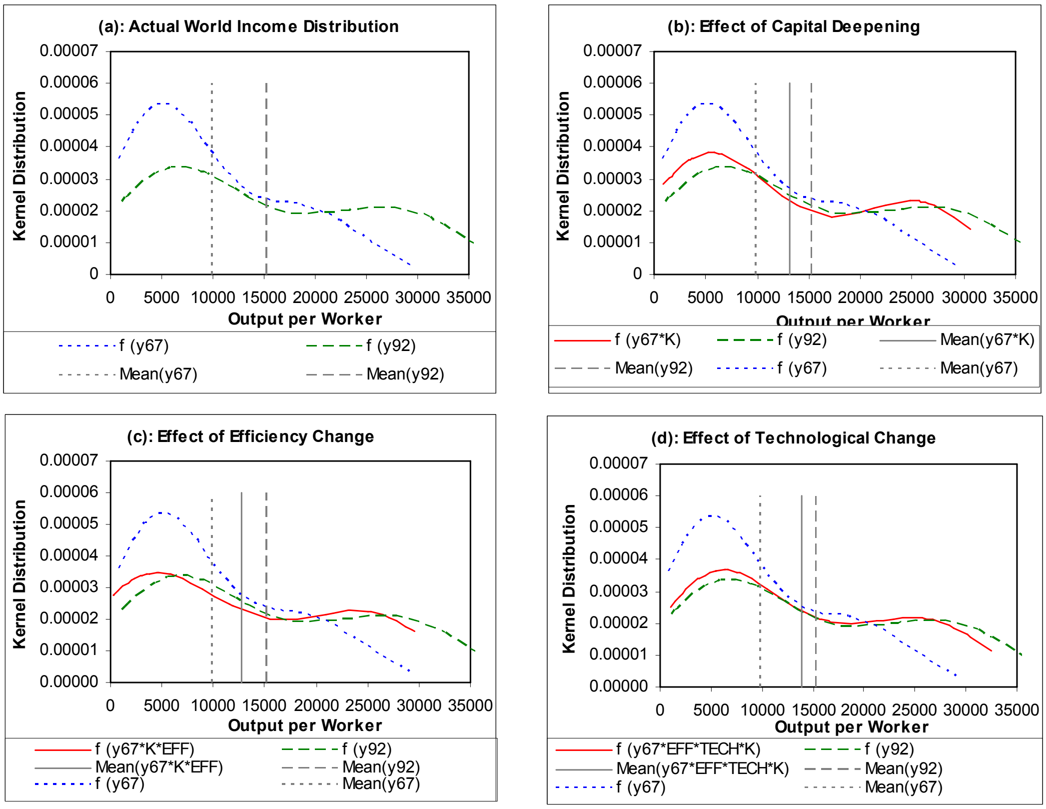


Figure 13.2 - Counterfactual Distributions of Output per Worker for the Economy Using a MCMC Approach

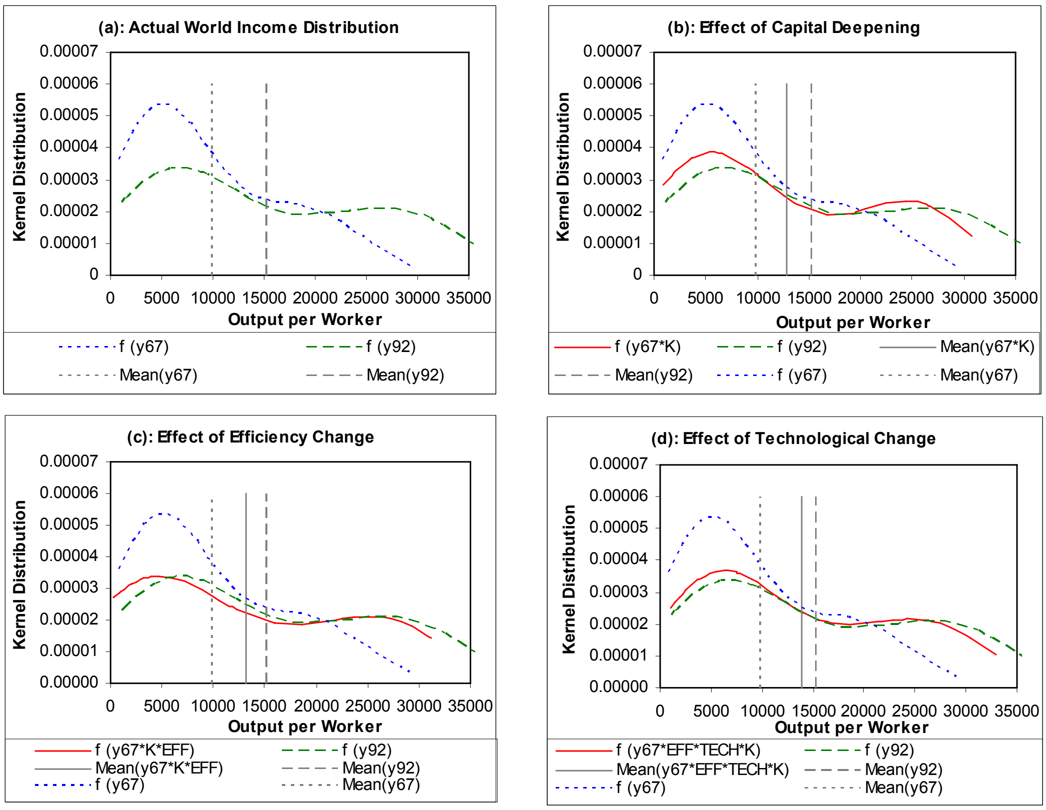


Table 31.1 - Li's Distribution Hypothesis Tests for the Economy Using a Frequentist Approach

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	2.398	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	1.726	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech)$	4.099	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * FAcc)$	0.087	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech)$	2.765	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	-0.027	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	0.529	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	0.068	**H_0 not rejected**	**H_0 not rejected**

Table 31.2 - Li's Distribution Hypothesis Tests for the Economy Using a MCMC Approach

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	2.398	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	1.348	H_0 rejected	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Tech)$	4.440	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * FAcc)$	0.145	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech)$	2.584	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	-0.044	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	0.902	**H_0 not rejected**	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	0.073	**H_0 not rejected**	**H_0 not rejected**

The combination of the results presented allows us to conclude the following:

- As we can see in tables 30.1 and 30.2, there is no sign of catch-up. In fact, observations with stronger movements away of the frontier are from low and low-middle income countries. Furthermore, this effect helps the emergence of a bimodal distribution, as we can observe in figures 11.1 and 11.2. This is more evident in the Bayesian framework, where, as we can see in table 31.2, for 5% of significance level, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual one.

- Panel b of figures 12.1 and 12.2 shows that technological change is responsible for an important shift of the density function from the lower tail to the middle, which means that technological change has an important contribution to the welfare of poorer countries. It also reveals a very small change of mass from the high-middle to the higher tail, meaning that also the welfare of the rich is raised, although in a more reduced way. If we combine the effects of efficiency and technological change, the reading of tables 30.1 and 30.2 reveals that total factor productivity growth is very small for the economy.

- Analyzing panel b in figures 13.1 and 13.2, which reports the effect of a single component, we can infer that capital deepening is the only effect that, for all levels of significance and for both frequentist and Bayesian methods, causes, *per se*, the emergence of a bimodal distribution and leads to a significant increase in the mean of labor productivity. The appropriate statistical tests support this conclusion (tables 31.1 and 31.2). At both significance levels, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual distribution assuming only capital deepening.

6.3.2 - Agricultural Sector

As before, we present parameters estimated for the fixed part of m*(.), the decomposition analysis of labor productivity growth and the corresponding counterfactual distributions. Estimation results are presented in table 32. We can confirm the quality of the estimation and the similarity between Classical and Bayesian versions of the model.

Table 32.1 - Estimation Results for the Fixed Part of m*(.) Using a Frequentist Approach (Agriculture).

First Period: 1967-1979			
	Value	Std.Error	p-value
ln k_{it}	0.545	0.217	0.01
ln la_{it}	1.33	0.285	<.0001
Second Period: 1980-1992			
	Value	Std.Error	p-value
ln k_{it}	0.883	0.162	<.0001
ln la_{it}	1.024	0.202	<.0001

Table 32.2 - Estimation Results for the Fixed Part of m*(.) Using a MCMC Approach (Agriculture).

First Period: 1967-1979				
node	Mean	Std. Error	MC error	Median
ln k_{it}	0.670	0.177	0.01	0.688
ln la_{it}	1.203	0.214	0.01	1.211
Second Period: 1980-1992				
node	Mean	Std. Error	MC error	Median
ln k_{it}	0.8441	0.1809	0.01	0.8576
ln la_{it}	1.125	0.2203	0.01	1.156

The results of the decomposition are presented in tables 33.1 and 33.2; the counterfactual distributions of labor productivity in figures 13-15; the Li's tests in tables 34.1 and 34.2.

Table 33.1 - Decomposition of Labor Productivity Growth for Agriculture Using a Frequentist Approach

Country	Percentage Change in Output per Worker	Contribution to Percentage Change in Output per Worker of			
		Change in Efficiency	Change in Technology	Capital Deepening	Stochastic Shocks
Malawi	-30.59%	278.04%	-79.42%	4.71%	-14.78%
Indonesia	93.06%	149.89%	-33.25%	-3.54%	19.99%
India	42.03%	190.33%	-47.48%	-10.51%	4.09%
Kenya	6.92%	150.10%	-51.65%	-11.00%	-0.65%
Korea, Republic of	395.34%	168.57%	32.70%	39.92%	-0.66%
Zimbabwe	-33.14%	53.44%	-34.84%	-15.72%	-20.66%
Madagascar	-17.99%	52.83%	-44.46%	-2.02%	-1.40%
Pakistan	20.61%	63.12%	-21.47%	-20.17%	17.95%
Sri Lanka	-13.25%	55.47%	-38.75%	-5.59%	-3.50%
Guatemala	22.25%	37.03%	5.53%	-27.98%	17.38%
Morocco	10.96%	45.48%	7.68%	-30.23%	1.52%
Philippines	38.81%	53.23%	-19.16%	2.28%	9.55%
Egypt	39.67%	104.65%	-28.65%	-11.07%	7.55%
Turkey	63.49%	35.74%	56.12%	-26.24%	4.59%
Iran	74.37%	-18.49%	161.03%	-24.90%	9.13%
El Salvador	40.37%	-3.08%	53.28%	-22.75%	22.31%
Japan	258.68%	-32.35%	180.63%	81.43%	4.14%
Peru	-9.33%	-10.90%	15.63%	-10.57%	-1.58%
Tunisia	96.41%	88.99%	-8.27%	-7.01%	21.83%
Dominican Republic	75.75%	0.85%	25.34%	20.63%	15.26%
Honduras	15.95%	-0.35%	45.28%	-27.32%	10.20%
Colombia	65.67%	23.99%	46.63%	-20.15%	14.12%
Portugal	177.78%	-4.16%	4.80%	179.33%	-1.00%
Syrian Arab Republic	71.98%	118.21%	8.91%	-29.72%	2.97%
Venezuela	100.32%	40.24%	24.11%	0.58%	14.43%
Costa Rica	99.72%	-15.60%	107.35%	-23.04%	48.29%
South Africa	57.51%	55.65%	-7.30%	25.52%	-13.03%
Greece	183.84%	16.71%	62.81%	29.58%	15.27%
Chile	44.82%	31.38%	11.21%	-17.04%	19.49%
Finland	138.57%	-41.35%	73.30%	138.94%	-1.77%
Italy	197.22%	-50.67%	182.21%	79.64%	18.84%
Norway	100.08%	-53.14%	132.57%	79.27%	2.40%
Austria	156.24%	-36.84%	134.23%	71.30%	1.11%
Sweden	199.32%	-22.70%	51.57%	171.91%	-6.04%
Israel	173.93%	-54.22%	207.64%	72.08%	13.04%
France	279.99%	-29.96%	115.72%	139.40%	5.05%
Uruguay	59.25%	18.75%	28.24%	-16.44%	25.14%
Argentina	47.14%	27.38%	28.13%	-2.06%	-7.96%
Denmark	151.35%	-15.66%	76.13%	76.37%	-4.07%
Netherlands	140.19%	-46.90%	191.65%	14.25%	35.74%
Canada	192.47%	63.17%	3.33%	47.42%	17.66%
United Kingdom	78.66%	-12.78%	66.16%	11.51%	10.55%
USA	67.95%	5.05%	19.62%	24.01%	7.79%
Australia	63.04%	-22.93%	26.58%	38.85%	20.37%
New Zealand	0.95%	-92.83%	135.42%	443.35%	10.05%
Mean	89.74%	**30.30%**	**42.37%**	**31.72%**	**8.24%**

Table 33.2 - Decomposition of Labor Productivity Growth for Agriculture Using a MCMC Approach

Country	Percentage Change in Output per Worker	Contribution to Percentage Change in Output per Worker of			
		Change in Efficiency	Change in Technology	Capital Deepening	Stochastic Shocks
Malawi	-30.59%	143.98%	-61.36%	-13.39%	-14.99%
Indonesia	93.06%	92.24%	-1.14%	-13.00%	16.75%
India	42.03%	86.62%	-12.98%	-16.46%	4.70%
Kenya	6.92%	69.94%	-22.22%	-19.21%	0.12%
Korea, Republic of	395.34%	163.53%	32.53%	43.48%	-1.15%
Zimbabwe	-33.14%	3.66%	4.20%	-21.87%	-20.77%
Madagascar	-17.99%	21.59%	-13.96%	-20.01%	-2.00%
Pakistan	20.61%	16.67%	24.24%	-27.15%	14.22%
Sri Lanka	-13.25%	15.24%	-10.19%	-13.56%	-3.04%
Guatemala	22.25%	10.88%	44.49%	-35.52%	18.33%
Morocco	10.96%	15.23%	54.41%	-39.27%	2.68%
Philippines	38.81%	25.41%	14.25%	-9.71%	7.29%
Egypt	39.67%	76.21%	-17.81%	-10.74%	8.04%
Turkey	63.49%	28.45%	89.67%	-35.67%	4.32%
Iran	74.37%	-16.38%	97.38%	-2.73%	8.61%
El Salvador	40.37%	-8.27%	76.17%	-28.09%	20.79%
Japan	258.68%	-23.27%	141.12%	85.05%	4.77%
Peru	-9.33%	-30.10%	58.33%	-16.42%	-1.97%
Tunisia	96.41%	88.46%	1.20%	-12.18%	17.26%
Dominican Republic	75.75%	-22.19%	65.28%	18.71%	15.12%
Honduras	15.95%	-7.87%	74.21%	-34.57%	10.42%
Colombia	65.67%	13.93%	79.78%	-28.01%	12.36%
Portugal	177.78%	-37.78%	32.33%	239.09%	-0.50%
Syrian Arab Republic	71.98%	119.94%	14.78%	-34.25%	3.63%
Venezuela	100.32%	32.38%	36.63%	-2.70%	13.83%
Costa Rica	99.72%	-19.77%	105.07%	-17.89%	47.84%
South Africa	57.51%	51.30%	3.89%	14.23%	-12.27%
Greece	183.84%	7.54%	61.25%	39.60%	17.24%
Chile	44.82%	25.82%	25.34%	-21.19%	16.52%
Finland	138.57%	-18.05%	36.09%	129.17%	-6.65%
Italy	197.22%	-37.28%	102.51%	95.63%	19.62%
Norway	100.08%	-33.24%	72.49%	72.27%	0.85%
Austria	156.24%	-21.40%	81.17%	78.34%	0.90%
Sweden	199.32%	6.38%	25.97%	139.16%	-6.61%
Israel	173.93%	-37.12%	107.86%	85.69%	12.88%
France	279.99%	-12.67%	60.11%	159.67%	4.66%
Uruguay	59.25%	27.55%	16.94%	-13.23%	23.06%
Argentina	47.14%	32.05%	7.87%	8.49%	-4.79%
Denmark	151.35%	4.79%	37.65%	78.83%	-2.55%
Netherlands	140.19%	-32.50%	129.23%	11.93%	38.68%
Canada	192.47%	87.46%	2.28%	30.97%	16.46%
United Kingdom	78.66%	2.94%	26.03%	20.64%	14.15%
USA	67.95%	2.79%	6.03%	43.40%	7.45%
Australia	63.04%	-37.53%	29.98%	66.86%	20.32%
New Zealand	0.95%	-92.33%	69.08%	615.60%	8.83%
Mean	89.74%	**17.45%**	**40.18%**	**35.33%**	**7.90%**

Figure 14.1 - Counterfactual Distributions of Output per Worker for Agriculture Using a Frequentist Approach

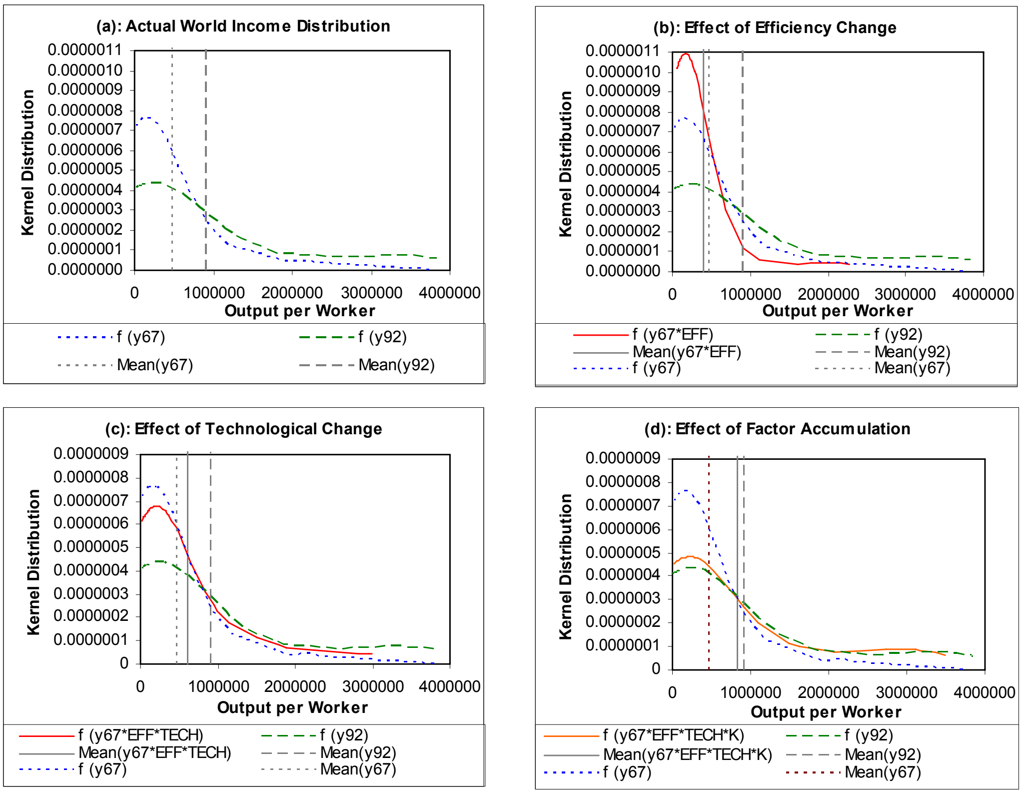


Figure 14.2 - Counterfactual Distributions of Output per Worker for Agriculture Using a MCMC Approach

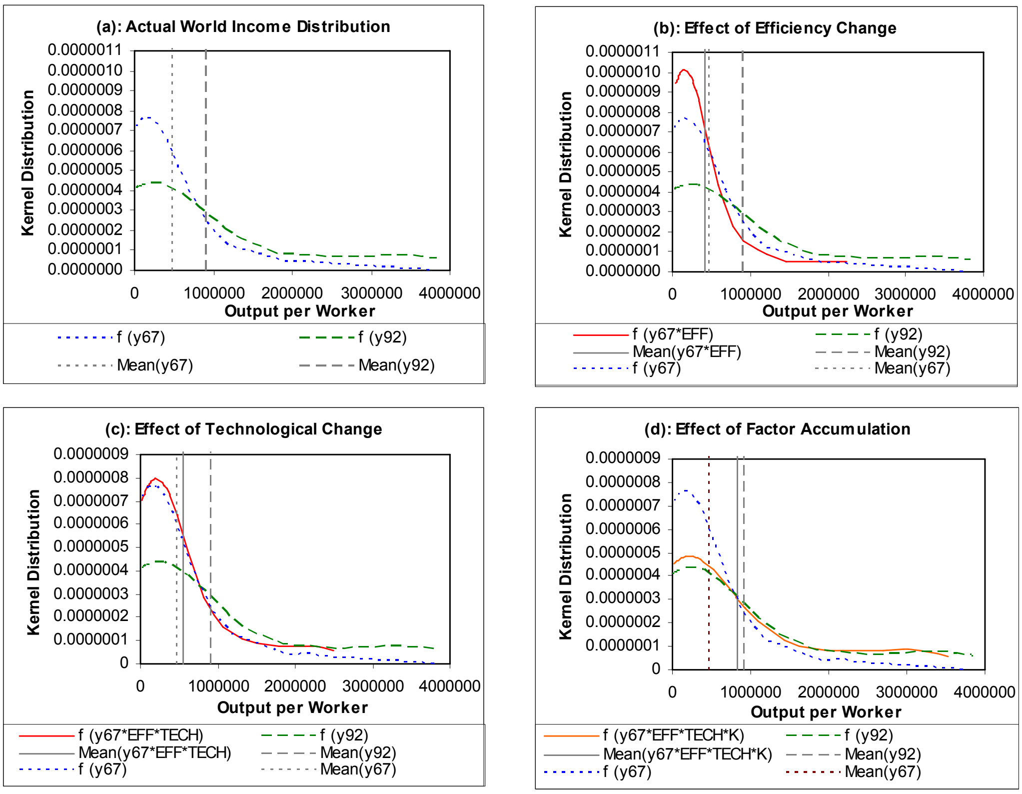


Figure 15.1 - Counterfactual Distributions of Output per Worker for Agriculture Using a Frequentist Approach

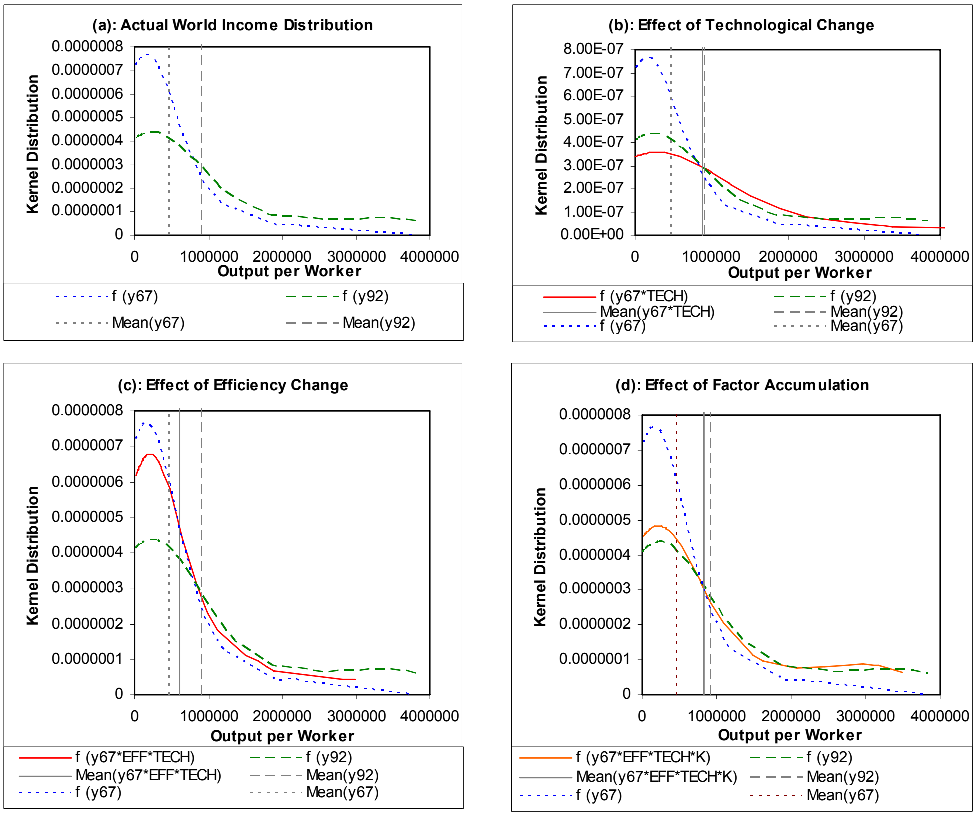


Figure 15.2 - Counterfactual Distributions of Output per Worker for Agriculture Using a MCMC Approach

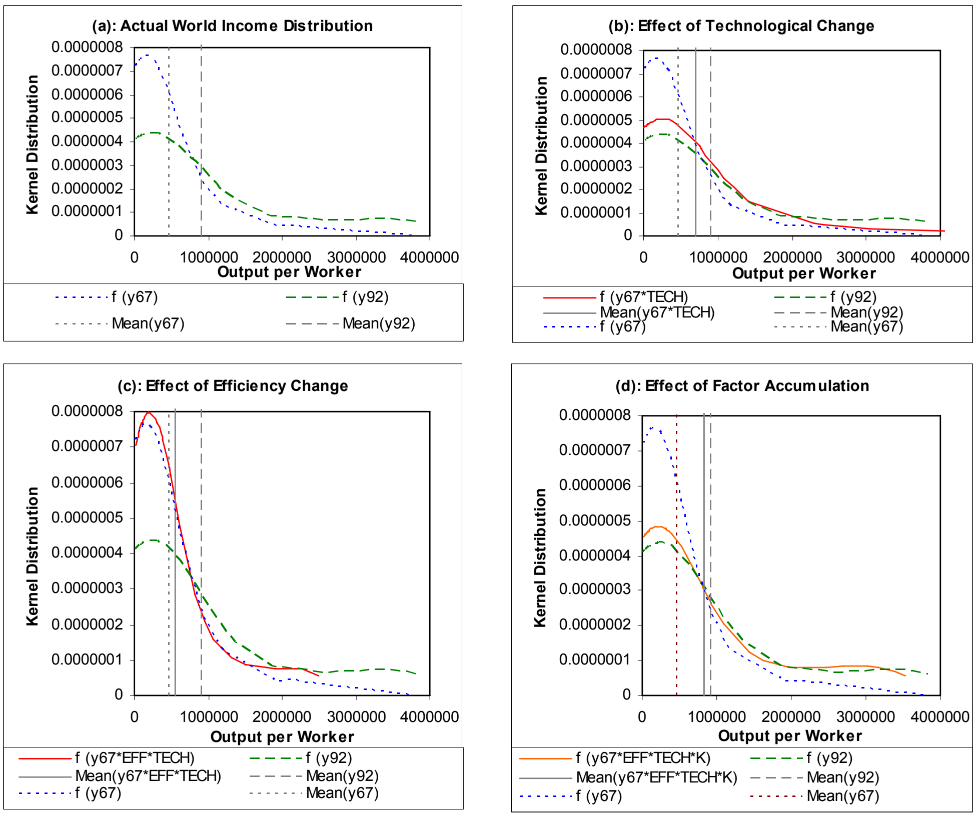


Figure 16.1 - Counterfactual Distributions of Output per Worker for Agriculture Using a Frequentist Approach

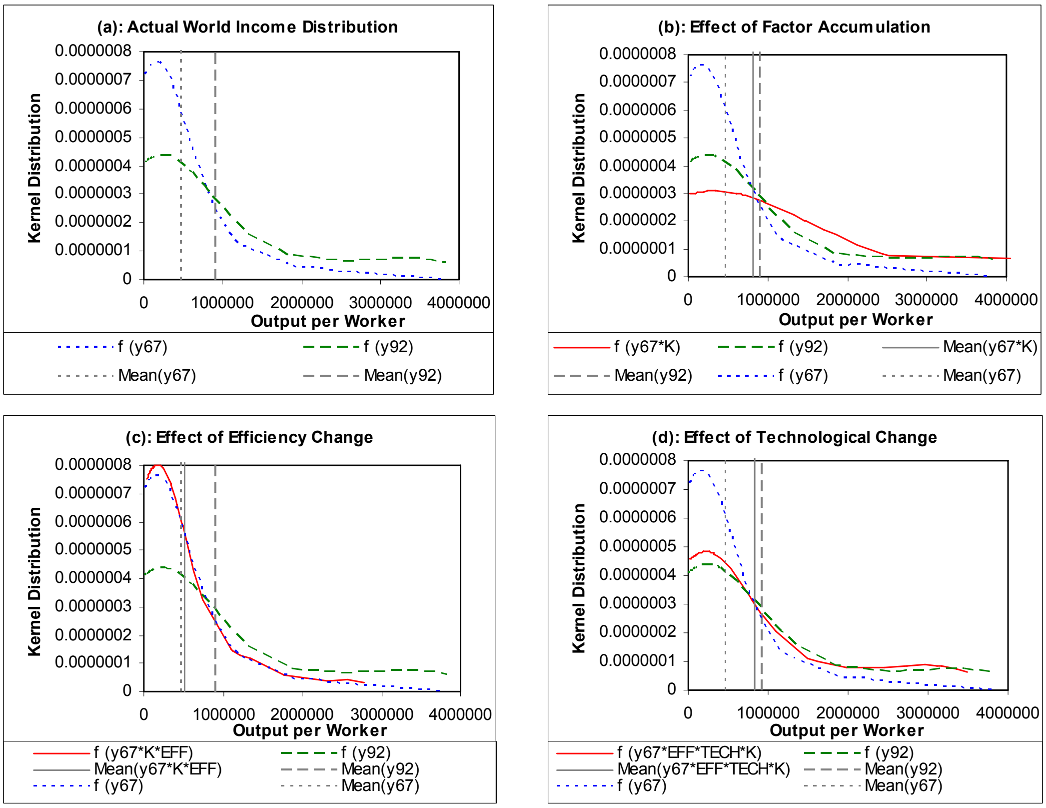


Figure 16.2 - Counterfactual Distributions of Output per Worker for Agriculture Using a MCMC Approach

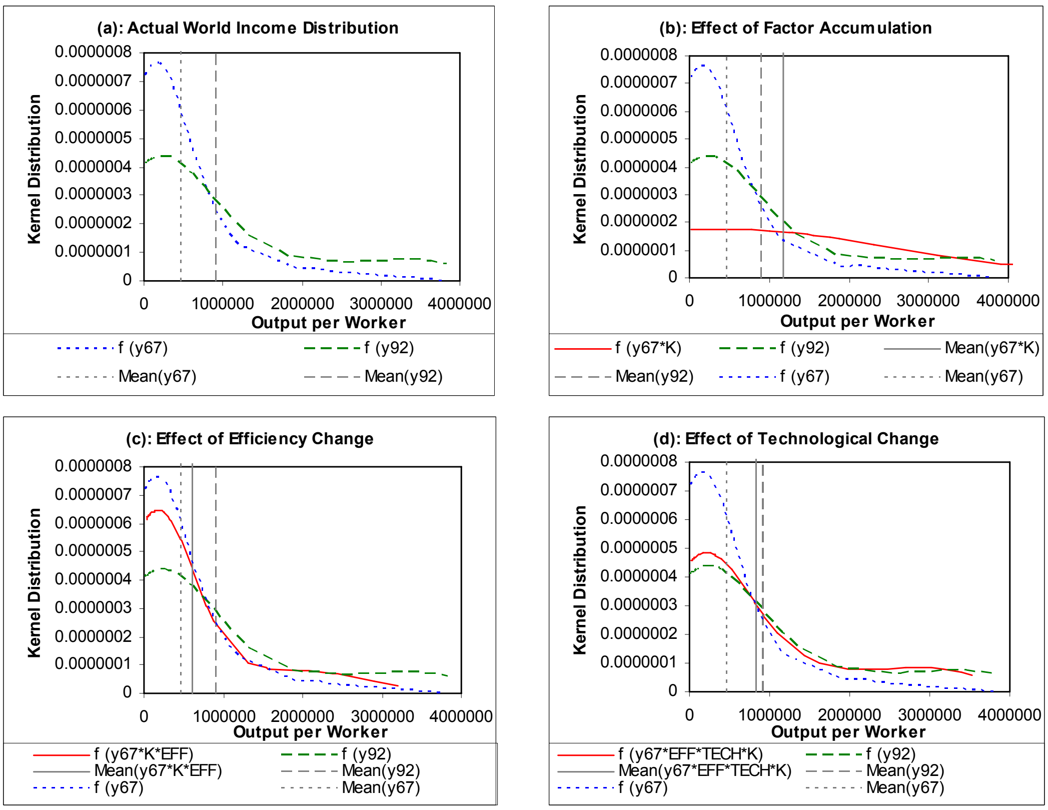


Table 34.1 - Li's Distribution Hypothesis Tests for Agriculture Using a Frequentist Approach

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	5.842	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	11.226	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech)$	0.196	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * FAcc)$	0.683	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech)$	4.504	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	5.051	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	8.652	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	0.152	**H_0 not rejected**	**H_0 not rejected**

Table 34.2 - Li's Distribution Hypothesis Tests for Agriculture Using a MCMC Approach

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	5.842	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	9.555	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech)$	1.044	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * FAcc)$	1.291	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech)$	5.509	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	2.553	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	6.848	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	0.154	**H_0 not rejected**	**H_0 not rejected**

The analysis of results suggests the following:

- The catch-up effect is stronger for agriculture than for the overall economy, as described in tables 33.1 and 33.2. Furthermore, the poorest countries have strong positive effects, contrasting with movements away from the world production frontier for some rich countries. Panel b in figures 14.1 and 14.2 shows a clear unimodal counterfactual distribution, contrasting with the results for the overall economy.

- The analysis of tables 33.1 and 33.2 reveals that technological change is the most important component for the majority of countries. Furthermore, if we combine this result with efficiency change, we can conclude that total factor productivity rates are very high for the agricultural sector. Panel *b* of figures 15.1 and 15.2 suggests that technological change is responsible for an important shift of the density from the lower tail to the rest of the distribution. This means that technological change in agriculture contributes to the welfare of poorer countries. This effect results in a higher mean of output per worker and in a distribution closer to the 1992-distribution. This conclusion is supported by the statistic tests of Li (1996) presented in tables 34.1 and 34.2. At both significance levels and using both methods, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual distribution assuming only technical change.

- It is notable that many countries experience reductions in factor per worker endowments[24], as we can observe in tables 33.1 and 33.2. This effect is similar to the last one, resulting in a shift of the density from the lower tail to the rest of the distribution (panel b in figures 15.1 and 15.2). At both significance levels, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual distribution assuming only this effect (tables 34.1 and 34.2).

6.4 - Differences Among Groups of Countries

In section 4, we conclude that output per capita growth in agriculture grows as the income presented by countries increases. Developed countries present very strong

[24] For instance, although labor has diminished in the period for most of the countries, the reduction of factor endowments was even stronger.

growth rates while developing countries exhibit modest values. In the economy, output growth rates are similar for both groups. In tables 35.1 and 35.2, the decomposition of labor productivity growth for the overall economy of developed and developing countries is presented.

Table 35.1 - Decomposition of Labor Productivity Growth for the Economy Using a Frequentist Approach (2 Groups of Countries)

		N	Mean	Std. Deviation	Minimum	Maximum
TFP change agriculture	Developed	16	,2839	,39444	-,83	,90
	Developing	29	,5651	,53967	-,22	2,56
	Total	45	,4651	,50687	-,83	2,56
TFP change economy	Developed	16	,1354	,55184	-,52	1,97
	Developing	27	,1993	,46551	-,56	1,26
	Total	43	,1756	,49385	-,56	1,97
Capital deepening agriculture	Developed	16	1,04317	1,042233	,115	4,434
	Developing	29	-,08339	,183488	-,302	,399
	Total	45	,31717	,830149	-,302	4,434
Capital deepening economy	Developed	16	,5699	,57462	-,10	1,99
	Developing	27	,3374	,91328	-,49	3,32
	Total	43	,4239	,80448	-,49	3,32
Y/L chg no s. shocks agriculture	Developed	16	1,3260	,75557	-,08	2,62
	Developing	29	,4532	,77187	-,19	3,99
	Total	45	,7635	,86734	-,19	3,99
Y/L chg no s. shocks economy	Developed	16	,6756	,69507	-,19	2,42
	Developing	27	,4010	,71018	-,37	3,26
	Total	43	,5032	,70909	-,37	3,26

Table 35.2 - Decomposition of Labor Productivity Growth for the Economy Using a MCMC Approach (2 Groups of Countries)

		N	Mean	Std. Deviation	Minimum	Maximum
TFP change agriculture	Developed	16	,2447	,42264	-,87	,92
	Developing	29	,6410	,51536	-,06	2,49
	Total	45	,5001	,51643	-,87	2,49
TFP change economy	Developed	16	,1354	,55184	-,52	1,97
	Developing	27	,1993	,46551	-,56	1,26
	Total	43	,1756	,49385	-,56	1,97
Capital deepening agriculture	Developed	16	1,22018	1,434408	,119	6,156
	Developing	29	-,12493	,206581	-,393	,435
	Total	45	,35333	1,073588	-,393	6,156
Capital deepening economy	Developed	16	,5699	,57462	-,10	1,99
	Developing	27	,3374	,91328	-,49	3,32
	Total	43	,4239	,80448	-,49	3,32
Y/L change with no stochastic shocks agriculture	Developed	16	1,3297	,75975	-,07	2,63
	Developing	29	,4584	,77329	-,18	4,01
	Total	45	,7682	,86904	-,18	4,01
Y/L change with no stochastic shocks economy	Developed	16	,6756	,69507	-,19	2,42
	Developing	27	,4010	,71018	-,37	3,26
	Total	43	,5032	,70909	-,37	3,26

It is possible to conclude that the TFP effect is higher for agriculture. Furthermore, in this sector, the TFP growth rate of developing countries is twice the one exhibited by the developed nations. This contrasts with the behavior of TFP growth rates for the overall economy, very similar for the 2 groups. Therefore, in agriculture, TFP does not contribute to the large difference between the rich and the poor. Capital deepening, per se, is responsible for the gap between country groups. Developing countries present evidences of factor disaccumulation, probably due to the market distortions introduced by agricultural policies of the developed nations.

Even if we desegregate developing countries into middle and poor countries in tables 36.1 and 36.2, conclusions are maintained: different capital deepening and TFP rates for agriculture among groups (as ANOVA tests confirm in tables 37.1 and 37.2) and similar rates for economy. It is interesting to notice that middle income countries present the strongest TFP rates for agriculture and negative values for factors accumulation.

Table 36.1 - Decomposition of Labor Productivity Growth for the Economy Using a Frequentist Approach (3 Groups of Countries)

		N	Mean	Std. Deviation	Minimum	Maximum
TFP change agriculture	Rich	16	,2839	,39444	-,83	,90
	Médium	19	,7484	,54995	,03	2,56
	Poor	10	,2170	,30915	-,22	,67
	Total	45	,4651	,50687	-,83	2,56
TFP change economy	Rich	16	,1354	,55184	-,52	1,97
	Médium	17	,1014	,35578	-,56	,93
	Poor	10	,3658	,59336	-,22	1,26
	Total	43	,1756	,49385	-,56	1,97
Capital deepening agriculture	Rich	16	1,04317	1,042233	,115	4,434
	Médium	19	-,07348	,218498	-,302	,399
	Poor	10	-,10222	,093058	-,273	,047
	Total	45	,31717	,830149	-,302	4,434
Capital deepening economy	Rich	16	,5699	,57462	-,10	1,99
	Médium	17	,4467	,92643	-,44	3,32
	Poor	10	,1516	,90711	-,49	2,62
	Total	43	,4239	,80448	-,49	3,32
Y/L chg no s. shocks agriculture	Rich	16	1,3260	,75557	-,08	2,62
	Médium	19	,6490	,87987	-,08	3,99
	Poor	10	,0811	,26539	-,19	,61
	Total	45	,7635	,86734	-,19	3,99
Y/L chg no s. shocks economy	Rich	16	,6756	,69507	-,19	2,42
	Médium	17	,4419	,77415	-,14	3,26
	Poor	10	,3316	,61894	-,37	1,83
	Total	43	,5032	,70909	-,37	3,26

Table 36.2 - Decomposition of Labor Productivity Growth for the Economy Using a MCMC Approach (3 Groups of Countries)

		N	Mean	Std. Deviation	Minimum	Maximum
TFP change agriculture	Rich	16	,2447	,42264	-,87	,92
	Médium	19	,7966	,53794	,11	2,49
	Poor	10	,3453	,31509	-,06	,90
	Total	45	,5001	,51643	-,87	2,49
TFP change economy	Rich	16	,1354	,55184	-,52	1,97
	Médium	17	,1014	,35578	-,56	,93
	Poor	10	,3658	,59336	-,22	1,26
	Total	43	,1756	,49385	-,56	1,97
Capital deepening agriculture	Rich	16	1,22018	1,434408	,119	6,156
	Médium	19	-,09072	,245109	-,393	,435
	Poor	10	-,18995	,073781	-,346	-,107
	Total	45	,35333	1,073588	-,393	6,156
Capital deepening economy	Rich	16	,5699	,57462	-,10	1,99
	Médium	17	,4467	,92643	-,44	3,32
	Poor	10	,1516	,90711	-,49	2,62
	Total	43	,4239	,80448	-,49	3,32
Y/L chg no s. shocks agriculture	Rich	16	1,3297	,75975	-,07	2,63
	Médium	19	,6539	,88086	-,08	4,01
	Poor	10	,0869	,27296	-,18	,65
	Total	45	,7682	,86904	-,18	4,01
Y/L chg no s. shocks economy	Rich	16	,6756	,69507	-,19	2,42
	Médium	17	,4419	,77415	-,14	3,26
	Poor	10	,3316	,61894	-,37	1,83
	Total	43	,5032	,70909	-,37	3,26

Table 37.1 - ANOVA Tests Using a Frequentist Approach (3 Groups of Countries)

		Sum of Squares	Df	Mean Square	F	Sig.
TFP change agriculture	Between Groups	2,666	2	1,333	6,482	,004
	Within Groups	8,638	42	,206		
	Total	11,304	44			
TFP change economy	Between Groups	,481	2	,241	,986	,382
	Within Groups	9,762	40	,244		
	Total	10,243	42			
Capital deepening agriculture	Between Groups	13,091	2	6,546	15,955	,000
	Within Groups	17,231	42	,410		
	Total	30,322	44			
Capital deepening economy	Between Groups	1,091	2	,546	,837	,441
	Within Groups	26,091	40	,652		
	Total	27,182	42			
Y/L chg no s. shocks agriculture	Between Groups	9,968	2	4,984	9,049	,001
	Within Groups	23,132	42	,551		
	Total	33,100	44			
Y/L chg no s. shocks economy	Between Groups	,834	2	,417	,822	,447
	Within Groups	20,284	40	,507		
	Total	21,118	42			

Table 37.2 - ANOVA Tests Using a MCMC Approach (3 Groups of Countries)

		Sum of Squares	Df	Mean Square	F	Sig.
TFP change agriculture	Between Groups	2,953	2	1,477	7,062	,002
	Within Groups	8,782	42	,209		
	Total	11,735	44			
TFP change economy	Between Groups	,481	2	,241	,986	,382
	Within Groups	9,762	40	,244		
	Total	10,243	42			
Capital deepening agriculture	Between Groups	18,721	2	9,360	12,288	,000
	Within Groups	31,993	42	,762		
	Total	50,714	44			
Capital deepening economy	Between Groups	1,091	2	,546	,837	,441
	Within Groups	26,091	40	,652		
	Total	27,182	42			
Y/L chg no s. shocks agriculture	Between Groups	9,935	2	4,967	8,956	,001
	Within Groups	23,295	42	,555		
	Total	33,230	44			
Y/L chg no s. shocks economy	Between Groups	,834	2	,417	,822	,447
	Within Groups	20,284	40	,507		
	Total	21,118	42			

In the spirit of the regressions presented in Barro (1991), we search for empirical regularities about growth and agriculture, trying to answer the question if a shrinking agricultural sector is a pre-condition for economic growth and for fast labor productivity in agriculture itself. Estimates obtained by ordinary least squares in the previous section suggest that, ceteris paribus, output growth rates are higher in countries with a smaller agricultural sector at the beginning of the period and in which the agricultural sector has a higher shrinkage. Nevertheless, when we check if this negative role of agriculture holds for total factor productivity growth rates in tables 38.1 and 38.2, although none of the variables are significant at 5%, the sign of the weight of agriculture in employment changes. This evidence indicates that, ceteris paribus, a country with a large agricultural sector at the start tends to exhibit stronger TFP rates. Nevertheless, the reduction of the weight of agriculture in employment through time appears to be important to raise both TFP and output per worker growth.

Table 38.1 - Regressions on Total Factor Productivity Growth Rates Using a Frequentist Approach

	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	,196	,622		,315	,754
Starting weight of agriculture in employment	,063	,764	,033	,082	,935
% variation in weight of agriculture	-,213	,571	-,079	-,372	,712
Starting output per worker in the economy	-1,14E-005	,000	-,173	-,496	,622

Table 38.2 - Regressions on Total Factor Productivity Growth Rates Using a MCMC Approach

	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	,046	,534		,085	,932
Starting weight of agriculture in employment	,095	,634	,050	,149	,882
% variation in weight of agriculture	-1,273	1,302	-,185	-,978	,334
Starting output per worker in the economy	-2,52E-006	,000	-,038	-,104	,918

Summing up all results, it seems that if this disinvestment expressed by factor disaccumulation did not occur, the agricultural sector, presenting stronger TFP rates than the economy and with a growth rate twice for the poor when we compare with the one of the rich nations, could be the growth engine for developing countries, inducing rises in TFP and, as a result, in output per labor growth rates. Therefore, the results of the parametric model presented in section 5.4 are confirmed within a semiparametric framework.

6.5 - Model Extension

For the overall economy, in the last chapter, using the Schwarz Bayesian Information Criterion (SBIC) within a Stochastic Frontier Finite Mixture Model, we group the countries into 2 classes. If we do not take into account this information, technological differences could be labeled as inefficiency, invalidating the decomposition of output per worker. Therefore, we decide to address this issue within the penalized spline model. Ruppert, Wand and Carroll (2003, pp. 188-190) address a similar problem in determining the influence of four ethnic groups on the spinal bone mineral density. Their decision was to incorporate those differences in the constant of the fixed part of the spline. This solution in our model would be equivalent to assume that the production frontiers for the groups are parallel, differing only on the origin. In this case, the production frontier model would be:

$$y_{it} = m^*(x_{it}) + \beta_0 \cdot class_i + v_{it} + u_i^*(x_{it}), \tag{6.23}$$

where:

$$class_i = \begin{cases} 0 & if \quad \text{i is from class 1} \\ 1 & if \quad \text{i is from class 2} \end{cases}. \tag{6.24}$$

We prefer to extend this model of Ruppert, Wand and Carroll (2003, pp. 188-190) to the case where the frontier changes in all parameters of the production-frontier spline:

$$y_{it} = m^*(x_{it}) + m_c^*(x_{it}) + v_{it} + u_i^*(x_{it}), \tag{6.25}$$

where:

$$m^*(x_{it}) = \beta_1 x_{it} + \sum_{k=1}^{K} w_k \left(x_{it} - \kappa_k\right)_+ ; \qquad (6.26)$$

$$m_c^*(x_{it}) = \gamma_1 x_{it}^{class=2} + \sum_{k=1}^{K} g_k^{class=1} \left(x_{it}^{class=1} - \kappa_k\right)_+ + \sum_{k=1}^{K} g_k^{class=2} \left(x_{it}^{class=2} - \kappa_k\right)_+ . \qquad (6.27)$$

If we use other way of presenting the last equation, it becomes clear that eventual problems of nonidentifiability do not exist:

$$\begin{cases} m_1^*(x_{it}) = \sum_{k=1}^{K} g_k^{class=1} \left(x_{it}^{class=1} - \kappa_k\right)_+ \\ m_2^*(x_{it}) = \gamma_1 x_{it}^{class=2} + \sum_{k=1}^{K} g_k^{class=2} \left(x_{it}^{class=2} - \kappa_k\right)_+ \end{cases} . \qquad (6.28)$$

Ruppert (2004, p. 33) indicates that the use of a non-parallel difference between group functions should be difficult to implement due to the small number of times each individual is usually observed. We try to implement the frequentist version of the extended model, but convergence was not attained. The MCMC version of the model eliminates this kind of problems. In each spline, we use the method already described.

In this model, there is not a unique frontier for the entire sample, but one frontier for each class. Thus, the decomposition of labor productivity in (5.18) must be adjusted to this framework:

$$\frac{y_c}{y_b} = \frac{Eff_c}{Eff_b} \cdot \frac{\exp(v_c)}{\exp(v_b)} \cdot \left[\frac{y_c\left(x_b\right)|j_c}{y_b\left(x_b\right)|j_b} \cdot \frac{y_c\left(x_c\right)|j_c}{y_b\left(x_c\right)|j_b} \right]^{1/2} \cdot \left[\frac{y_c\left(x_c\right)|j_c}{y_c\left(x_b\right)|j_c} \cdot \frac{y_b\left(x_c\right)|j_b}{y_b\left(x_b\right)|j_b} \right]^{1/2} . \qquad (6.29)$$

Using the components of the labor productivity change decomposition given in equation (6.29), it is possible to obtain the corresponding counterfactual distributions.

Reading table 39, it is possible to conclude that the catch-up effect and capital deepening are stronger than the single class model presented in table 30.2. As expected, gains in efficiency are more evident in poorer countries. Technical change represents now a much more reduced effect than the single frontier case.

Table 39 - Decomposition of Labor Productivity Growth for the Economy Using a MCMC Approach With 2 Classes

Country (by ascending order of output per worker in the first period)	**Percentage Change in Output per Worker**	**Contribution to Percentage Change in Output per Worker of**			
		Change in Efficiency	**Change in Technology**	**Capital Deepening**	**Stochastic Shocks**
Malawi	22,55%	-30,75%	20,37%	36,36%	7,81%
Kenya	15,74%	13,38%	-11,41%	18,57%	-2,81%
Indonesia	248,48%	23,22%	-36,61%	262,45%	23,10%
India	95,07%	64,56%	33,72%	-28,48%	23,95%
Madagascar	-33,85%	-32,93%	38,34%	-31,78%	4,50%
Zimbabwe	-6,22%	-1,53%	1,96%	-22,89%	21,14%
Pakistan	66,21%	-11,70%	155,37%	-48,84%	44,06%
Sri Lanka	78,95%	-19,83%	92,13%	6,79%	8,80%
Korea, Republic of	422,65%	6,65%	-7,58%	332,45%	22,61%
Philippines	24,41%	-32,55%	60,52%	-12,62%	31,50%
Egypt	92,30%	-25,61%	176,37%	-30,55%	34,69%
Honduras	15,56%	-3,06%	8,89%	-9,99%	21,64%
Turkey	107,44%	-1,07%	-0,39%	71,71%	22,59%
Morocco	53,10%	5,53%	6,66%	31,18%	3,69%
Tunisia	103,62%	25,72%	-6,21%	21,85%	41,72%
Dominican Republic	40,71%	-31,68%	170,47%	-25,48%	2,18%
El Salvador	3,22%	-22,11%	1,88%	8,44%	19,95%
Guatemala	23,56%	-21,69%	50,62%	-5,65%	11,02%
Colombia	31,98%	13,14%	18,97%	0,44%	-2,38%
Portugal	160,32%	-10,12%	230,97%	15,10%	-23,97%
Syrian Arab Republic	126,30%	15,94%	-15,90%	60,42%	44,68%
South Africa	17,02%	-9,04%	-7,84%	34,11%	4,10%
Greece	108,70%	8,58%	3,91%	55,77%	18,74%
Costa Rica	15,18%	-0,30%	-8,03%	14,42%	9,78%
Japan	167,57%	-14,63%	-16,23%	198,61%	25,30%
Peru	-33,56%	-9,92%	4,71%	-1,59%	-28,42%
Uruguay	30,30%	25,49%	-18,20%	9,91%	15,49%
Chile	24,22%	-1,26%	95,42%	-43,53%	14,00%
Israel	108,47%	10,77%	-7,14%	87,35%	8,18%
Argentina	19,13%	11,99%	-60,69%	207,59%	-12,02%
Finland	71,42%	-21,06%	43,63%	142,80%	-37,73%
Austria	84,33%	3,67%	-12,30%	101,59%	0,58%
Italy	95,40%	5,26%	5,93%	77,03%	-1,01%
United Kingdom	51,35%	5,25%	-29,20%	59,56%	27,29%
Denmark	36,22%	37,93%	-65,16%	102,17%	40,19%
France	63,81%	-0,88%	9,67%	26,00%	19,60%
Norway	67,40%	12,70%	-19,93%	-10,13%	106,41%
Sweden	26,29%	-4,15%	-10,10%	24,96%	17,29%
Netherlands	38,77%	-8,40%	7,16%	13,28%	24,80%
Australia	35,76%	0,39%	29,51%	28,72%	-18,88%
New Zealand	4,80%	5,57%	9,81%	-3,13%	-6,67%
Canada	39,33%	-2,24%	24,82%	24,26%	-8,12%
USA	23,76%	6,40%	45,91%	23,73%	-35,57%
Mean	**64,83%**	**-0,33%**	**23,60%**	**42,39%**	**12,65%**

Capital deepening is the only effect that, per se, causes the emergence of a bimodal distribution, as we can observe in figures 17-19 and in table 40.

Figure 17 - Counterfactual Distributions of Output per Worker for the Economy Using a MCMC Approach With 2 Classes

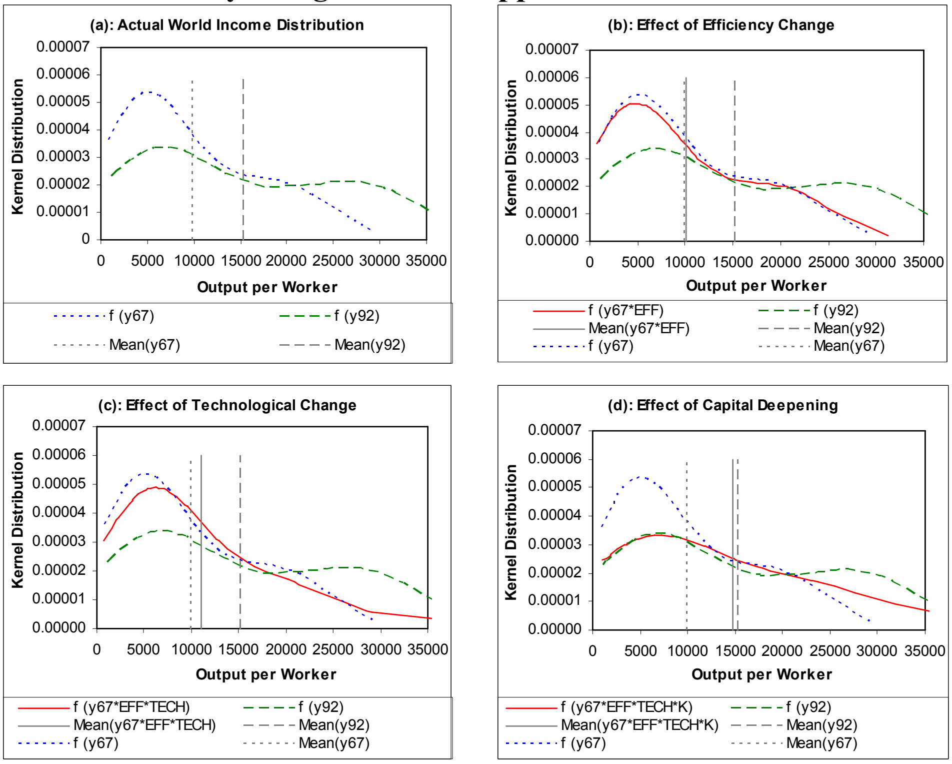


Figure 18 - Counterfactual Distributions of Output per Worker for the Economy Using a MCMC Approach With 2 Classes

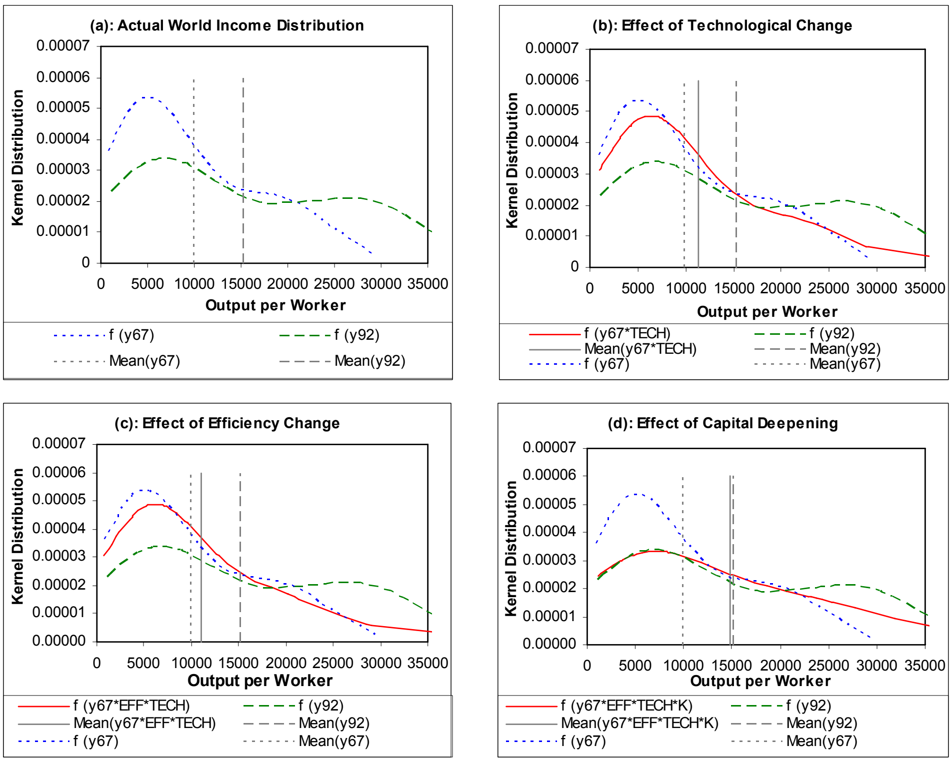


Figure 19 - Counterfactual Distributions of Output per Worker for the Economy Using a MCMC Approach With 2 Classes

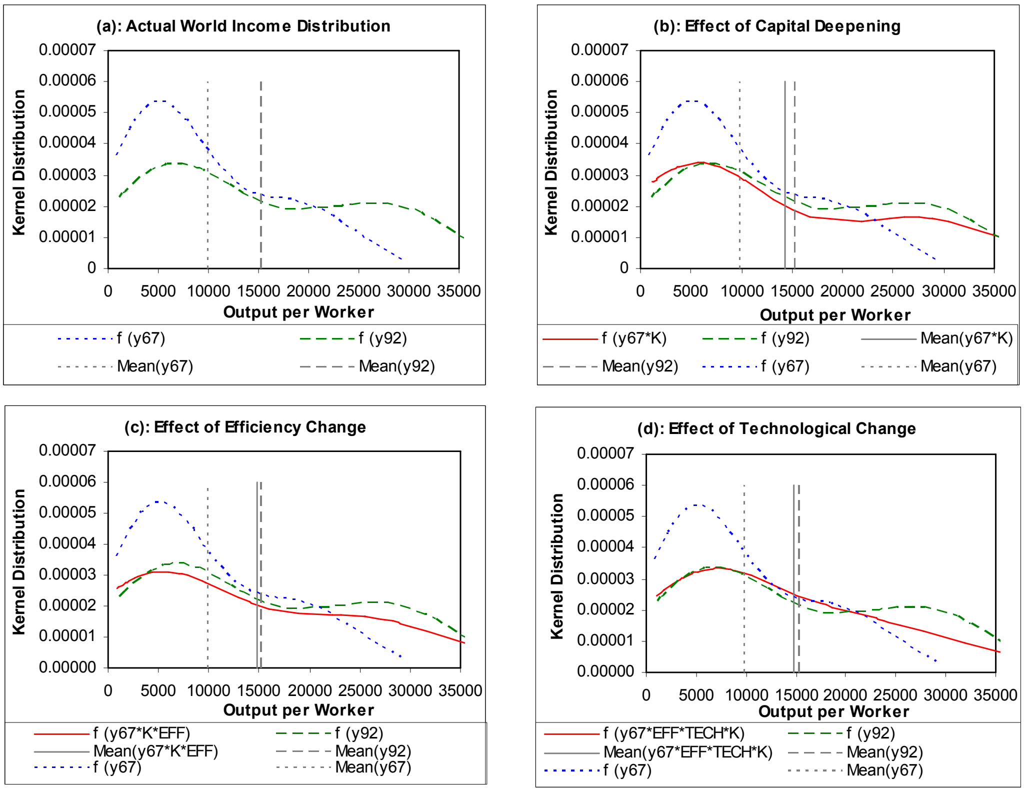


Table 40 - Li's Distribution Hypothesis Tests for the Economy Using a MCMC Approach With 2 Classes

Null Hypothesis (H_0)	T-test	Ten percent significance level (critical value: 1.28)	Five percent significance level (critical value: 1.64)
$f(y_{92}) = g(y_{67})$	2.398	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff)$	1.806	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Tech)$	1.919	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * FAcc)$	0.000	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech)$	2.214	H_0 rejected	H_0 rejected
$f(y_{92}) = g(y_{67} * Eff * FAcc)$	0.001	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Tech * FAcc)$	0.379	**H_0 not rejected**	**H_0 not rejected**
$f(y_{92}) = g(y_{67} * Eff * Tech * FAcc)$	0.271	**H_0 not rejected**	**H_0 not rejected**

Tests of Li in table 40 support this conclusion: at both significance levels, it is not possible to reject the equivalence of the 1992-distribution and the counterfactual distribution assuming only capital deepening. Panel b of figure 18 shows that the efficiency change effect remains very small, with the first year distribution almost unchanged. Panel b of figure 19 reveals that technical change benefits both poor and rich countries. When compared to the single class model, the effect is now smaller for the poorest and higher for the richest.

7 - Conclusion

The main purpose of this dissertation is to solve the conflicting evidence on the role of agriculture in economic growth, especially for developing countries. A quantitative study of this kind with countries in different development stages can only be accurate if technology differences are accounted for. Recent advances in the stochastic frontier literature, such the ones presented in section 3.6, account for heterogeneity of individuals. Furthermore, it is possible to introduce some flexibility in the definition of technology and stochastic noise and to allow the direct decomposition of labor productivity growth into catch-up, technical change and factor accumulation effects. We propose two different ways of accommodating heterogeneity across countries using a production frontier: a parametric known as stochastic frontier finite mixture model and a semiparametric approach with penalized splines, both using a panel data and fully stochastic frameworks.

The stochastic frontier finite mixture model is built on the assumption that there is a latent sorting of the observations in the data set into latent classes, unobserved by the investigator. Empirical results presented in chapter 5 show that, for the overall economy, countries can be grouped into 2 latent classes, each one with a specific parametric frontier, stochastic and inefficiency terms. If this outcome was not taken into account, technological differences would be labeled as inefficiency and the decomposition of labor productivity would not be valid.

In order to obtain more flexibility in the specifications of the frontier and the inefficiency component, we construct a semiparametric model using penalized splines. As we conclude in section 3, traditional semiparametric models involve a stochastic production frontier defined parametrically and an inefficiency distribution generated in a nonparametric fashion. In the model we create, each one is determined semiparametrically, since the linear regression splines used in the specification of the frontier and inefficiency component have two components: a linear one; and a random deviation from linearity using truncated lines as the basis for regression. We estimate both classical and Bayesian versions of the model. In the Bayesian approach, all parameters are random, while some of the spline components are fixed in the classical approach. Additionally, to increase the sensitivity of the model to the heterogeneity of countries, we create the exact semiparametric counterpart version of the parametric stochastic finite mixture model.

Although some country-specific outcomes change with the method used, it is possible to identify several global results common to all models, suggesting that agriculture can act as an engine of growth in the developing countries and contribute to change the trend for global divergence that has been observed in the last few decades.

One important conclusion of our analysis is that labor productivity in the overall economy evolved from a unimodal to a bimodal distribution, with the middle-income countries nearly disappearing. This contrast with the changes occurred in agriculture. The increase of mass in the middle of the labor productivity distribution in the agricultural sector contradicts the idea of the world becoming polarized into rich and poor countries. Furthermore, our results suggest that changes in labor productivity distribution are brought by capital deepening in the overall economy and by TFP change in agriculture.

In agriculture, the catch-up phenomenon is evident, while the opposite occurs for the overall economy. In the later case, factor accumulation and efficiency contributes to the welfare of the rich more than the poor, causing the formation of the twin-peak distribution. For both cases, technical change seems to increase the welfare of low and low-middle income countries, contradicting the results presented in Kumar and Russell (2002). Nevertheless, our remaining outcomes seem to confirm the main conclusions of Kumar and Russell (2002) for the overall economy, namely the reduced importance of total factor productivity to growth and the bipolar international divergence of labor productivity. Furthermore, it also supports the results of Bernard and Jones (1996) and Martin and Mitra (2001) that TFP growth rates are higher in agriculture and there is evidence of catch-up in this sector.

The output per worker increased more in agriculture than in the overall economy and the weight of agriculture in employment declined during the period 1967-1992.

To study the behavior of each set of countries, we use the criterion defined by the World Bank. In the overall economy, labor productivity presents similar growth rates across income groups. It is not possible to reject the hypothesis of the mean being equal among 2 or 4 sets. On the contrary, in agriculture, differences among groups are very meaningful: developed countries exhibit high growth rates, contrasting with moderate rates for the other nations. Regarding the variation in weight of agriculture in employment, the reduction is stronger for the rich countries, declining as the income of the reference group diminishes.

Capital deepening in the overall economy mimics the behavior of output per capita after removing shocks. In agriculture, capital deepening is only positive for the rich countries while TFP is similar among groups and stronger than the growth rates observed in the overall economy for all sets of nations. In agriculture, although developing countries exhibit a strong positive growth rate for TFP, the disinvestment in this sector shown by the factor accumulation effect causes a null growth of output per capita without shocks. For developed countries, strong rates of factor accumulation and TFP originate an even stronger growth of output per capita with no stochastic shocks.

In the spirit of the study of Barro (1991), we try to find some empirical regularities about labor productivity growth and agriculture. In addition to the initial level of income, we introduce two agriculture-related indicators as explainable variables of per capita output growth rates. Our purpose is to answer the question if a shrinking agricultural sector is a pre-condition for economic growth and for fast labor productivity in agriculture itself. Estimation results suggest that, ceteris paribus, output growth rates for the overall economy and agriculture are higher in countries with a smaller agricultural sector at the beginning of the period and in which the agricultural sector has a higher shrinkage. This evidence apparently indicates that agriculture has a negative role in economic growth, confirming that the reduction of agriculture weight is a necessary condition for development. Nevertheless, when we study the overall TFP growth, we conclude that, using a parametric model, a country with a large and rising agricultural sector tends to exhibit stronger TFP rates, ceteris paribus. The semiparametric approach confirms the reversion of the conclusion only for the dimension of the agricultural sector at the start.

Therefore, the reason for the apparently negative contribution of agriculture to productivity growth in developing countries seems to lie mainly on the large negative rates of factor accumulation occurred in the agricultural sector. In those countries, TFP growth was stronger in agriculture than in the overall economy. Furthermore, agricultural TFP growth rates were higher in the developing countries than in the developed ones. Consequently, agriculture could have contributed to higher rates of TFP and income per capita growth in the developing world. Therefore, our results suggest that if such a disinvestment in agriculture had not occurred, if policy makers did not discriminate against agriculture, this sector could have played a much more important role in the economic growth process of the developing countries.

References

Adams, R, A. Berger and R. C. Sickles (1999), “Semiparametric Approaches to Stochastic Panel Frontiers With Applications in Banking Industries”, Journal of Business and Economic Statistics, vol. 17(3), pp. 349-358.

Adelman, I. (1995), Institutions and Development Strategies: The Selected Essays of Irma Adelman, vol. 1, Ashgate, Brookfield.

Afriat, S. (1972), “Efficiency Estimation of Production Functions”, International Economic Review, vol. 13(3), pp. 568-598.

Ahn, S. C., Y. H. Lee, and P. Schmidt (2001), “MM Estimation of Linear Panel Data Models with Time-varying Individual Effects”, Journal of Econometrics, vol. 101, pp. 219-255.

Aigner, D. J. and S.F. Chu (1968), “On Estimating the Industry Production Function”, American Economic Review, vol. 58, pp. 226-239.

Aigner, D. J., K. Lovell and P. Schmidt (1977), “Formulation and Estimation of Stochastic Frontier Function Models”, Journal of Econometrics, vol. 6, pp. 21-37.

Amey, L. R. (1964), “The Allocation and Utilization of Resources”, Operational Research Quarterly, vol. 15(2), pp. 87-100.

Aptech Systems, Inc. (1999), “GAUSS, Version 3.2.38”, Maple Valley, WA.

Arnade, C. (1998), “Using a Programming Approach to Measure International Agricultural Efficiency and Productivity”, Journal of Agricultural Economics, vol. 49, pp. 67- 84.

Bardhan, P. K. (1973) “On the Incidence of Poverty in Rural India in the Sixties.” Economic and Political Weekly, vol. 8, pp. 245-54.

Barro, R. J. (1991), “Economic Growth in a Cross Section of Countries”, Quarterly Journal of Economics, vol. 106(2), pp. 407-443.

Battese, G. and T. Coelli (1988), "Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data", Journal of Econometrics, vol. 38, pp. 387-99.

Battese, G. and T. Coelli (1992), "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", Journal of Productivity Analysis, vol. 3, pp. 153-169.

Battese, G. and T. Coelli (1995), "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data", Empirical Economics, vol. 20, pp. 325-332.

Berger, J. O. and L. R. Pericchi (1998), "On Criticisms and Comparisons of Default Bayes Factors for Model Selection and Hypothesis Testing", in Proceedings of the Workshop on Model Selection, ed. W. Racugno, Bologna: Pitagora, pp. 1-50.

Bernard, A. and C. I. Jones (1996), "Productivity Across Industries and Countries: Time Series Theory and Evidence", Review of Economics and Statistics, vol. 78(1), pp. 135-146.

Berry, S. M., R. J. Carroll and D. Ruppert (2002), "Bayesian Smoothing and Regression Splines for Measurement Error Problems", Journal of the American Statistical Association, vol. 97, pp. 160-169.

Bhalla, S. S. (2002), Imagine There is no Country: Poverty, Inequality, and Growth in the Era of Globalization. Washington, DC, Institute for International Economics.

Boles, J. N. (1967), "Efficiency Squared-Efficient Computation of Efficiency Indexes", Western Farm Economic Association, Proceedings, Pullman, Washington, pp. 137-142.

Boskin, M. J. and L. J. Lau, (1990), "Post-War Economic Growth in the Group-of-Five Countries: A New Analysis", Publication No. 217. Palo Alto, CA: Stanford University, Center for Economic Policy Research

Boskin, M. J. and L. J. Lau, (1991), “Capital Formation and Economic Growth,” in Technology and Economics: A Volume Commemorating Ralph Landau's Service to the National Academy of Engineering, Washington, D.C.: National Academy Press, pp. 47-56.

Boskin, M. J. and L. J. Lau, (1992b), “International and Intertemporal Comparison of Productive Efficiency: An Application of the Meta-Production Function Approach to the Group-of-Five (G-5) Countries”, Economic Studies Quarterly, vol. 43 (4), pp. 298-312.

Boskin, M. J. and. L. J. Lau, (1992a), “Capital, Technology, and Economic Growth”, in Nathan Rosenberg, Ralph Landau, and David C. Mowery, eds. Technology and the Wealth of Nations. Stanford, Calif.: Stanford University Press.

Broyden, C. G. (1970), The Convergence of a Class of Double-Rank Minimization Algorithms. Parts I and II, Institute of Mathematics and its Applications, Malden n.6, pp. 76-90 and pp. 222-231.

Brumback, B. A. and J. A. Rice (1998), “Smoothing Spline Models for the Analysis of Nested and Crossed Samples of Curves (With Discussion)”, Journal of the American Statistical Association, vol. 93, pp. 961-94.

Brumback, B.A., D. Ruppert and M. P. Wand (1999), “Comment on Shively,Kohn, andWood”, Journal of the American Statistical Association, vol. 94, pp. 794-797.

Carbo Valverde, S. and D. B. Humphrey (2004), “Predicted and Actual Costs from Individual Bank Mergers”, Journal of Economics and Business, vol. 56(2), pp. 137-157.

Carlin, J. B. (1992), “Meta-analysis for 2 x 2 Tables: a Bayesian Approach”, Statistics in Medicine, vol. 11, pp. 141-59.

Carroll, R. J., D. Ruppert, L. A. Stefanski and C. Crainiceanu (2006), Measurement Error in Nonlinear Models: A Modern Perspective, Chapman and Hall, London.

Caselli (2005), "The Missing Input: Accounting for Cross-Country Income Differences", in Handbook of Economic Growth, P. Aghion and S. Durlauf (eds.), North-Holland, chapter 8.

Cass, D. (1965), "Optimum Growth in an Aggregative Model of Capital Accumulation", Review of Economic Studies, vol. 32, pp. 233-240.

Charnes, A., W. W. Cooper, and E. Rhodes (1978), "Measuring the Efficiency of Decision-Making Units", European Journal of Operational Research, vol. 2, pp. 429-44.

Chenery, H., S. A. Montek and C. L. G. Bell (1974), Redistribution with Growth, Oxford, Oxford University Press for the World Bank.

Christensen, L. R., D. W. Jorgenson and L. J. Lau (1971), "Conjugate Duality and the Transcendental Logarithmic Production Function", Econometrica, vol. 39, pp. 255-256.

Cleveland, W. S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots", Journal of the American Statistical Association, vol. 74, pp. 829-836.

Coelli, T. and P. Rao (2003), "Total Factor Productivity Growth in Agriculture: A Malmquist Index Analysis of 93 Countries, 1980-2000", International Association of Agricultural Economics (IAAE) Conference, Durban, South Africa.

Cornwell, C., P. Schmidt, and R. Sickles (1990), "Production Frontiers with Cross Sectional and Time-Series Variation in Efficiency Levels", Journal of Econometrics, vol. 46, pp. 185-200.

Crego, A., D. Larson, R. Butzer, and Y. Mundlak (1998), "A New Database on Investment and Capital for Agriculture and Manufacturing", World Bank Working Paper, No. 2013, Washington D.C., The World Bank.

Deaton, A. S. (2003), "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)", NBER Working Paper No. W9822.

Debreu, G. (1951), "The Coefficient of Resource Utilization", Econometrica, vol. 19, pp. 14-22.

Dikhanov, Y. and M. Ward (2002), Evolution of the Global Distribution of Income in 1970-99 (version 0.2; available at http://www.warwick.ac.uk/fac/soc/CSGR/Pdikhanov).

Dollar, D. and A. Kraay (2002), "Growth is Good for the Poor", Journal of Economic Growth, vol. 7, pp.195-225.

Dowrick, S. and M. Akmal (2001), "Contradictory Trends in Global Income Inequality: A Tale of Two Biases", mimeo (available at: http://ecocomm.anu.edu au/economics).

Duffy, J. and C. Papageorgiou (2000): "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification", Journal of Economic Growth, vol. 5, pp. 87-129.

Durbán M., J. Harezlak, M. P. Wand, R. J. Carroll (2005), "Simple fitting of subject-specific curves for longitudinal data", Statistics in Medicine, vol. 24, pp. 1153-1167

Durlauf, S. N., P. Johnson and J. R. W. Temple (2005), "Growth Econometrics" in Handbook of Economic Growth, P. Aghion and S. Durlauf, eds., North-Holland, chapter 7.

Easterly, W. and R. Levine (2001), "It's Not Factor Accumulation: Stylized Facts and Growth Models", The World Bank Economic Review, vol. 15(2), pp. 177-219.

Econometric Software, Inc. (2003), "LIMDEP, Version 8.0," ESI, New York.

Efron, B. (1979): "Bootstrap Methods: Another Look at the Jackknife", Annals of Statistics, vol. 7, pp. 1-16.

Efron, B. (1982), The Jackknife, the Bootstrap and Other Resampling Plans, CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, SIAM, Philadelphia.

Efron, B. (1985), "Bootstrap Confidence Intervals for a Class of Parametric Problems", Biometrika, vol. 72(1), pp. 45-58.

Eilers, P. H. C. and B. D. Marx (1996), "Flexible Smoothing With B-splines and Penalties", Statistical Science, vol. 11, pp. 89-121.

Eubank, R. L. (1988), Spline Smoothing and Nonparametric Regression, Marcel Dekker, New York.

Fan, Y. and A. Ullah (1999), "On Goodness-of-fit Tests for Weekly Dependent Processes Using Kernel Method", Journal of Nonparametric Statistics, vol. 11, pp. 337-360.

Fan, Y. and I. Gijbels (1996), Local Polynomial Modelling and Its Applications, Chapman and Hall, London.

FAOSTAT (2001), FAOSTAT Statistical Databases, Rome, Food and Agriculture Organization.

Färe, R., S. Grosskoff and C.A. Lovell (1994b), Production Frontiers, Cambridge University Press, Cambridge.

Färe, R., S. Grosskopf, M. Norris and Z. Zhang (1994a), "Productivity Growth, Technical Progress and Efficiency Change in Industrialized Countries", American Economic Review, vol. 84(1), pp. 66-83.

Farrell, M. J. (1957), "The Measurement of Productive Efficiency", Journal of the Royal Statistical Society, Series A, General, vol. 120(3), pp. 253-81.

Feng, Z. D. and C. E. McCulloch (1996), "Using Bootstrap Likelihood Ratio in Finite Mixture Models", Journal of the Royal Statistical Society (Series B), vol. 58, pp. 609-617.

Fernandez, C., G. Koop and M. Steel (2000), “A Bayesian Analysis of Multiple-Output Production Frontiers”, Journal of Econometrics, vol. 98, pp. 47-79.

Fernandez, C., G. Koop and M. Steel (2005), “Alternative Efficiency Measures for Multiple-Output Production”, Journal of Econometrics, vol. 126(2), pp. 411-444.

Feyrer, J., (2003), “Convergence By Parts,” mimeo, Dartmouth College (available at http://www.dartmouth.edu/~jfeyrer/parts.pdf).

Fields Development Team (2004), “Fields: Tools for Spatial Data”, National Center for Atmospheric Research, Boulder, CO (available at http://www.cgd.ucar.edu/Software/Fields).

Fishlow, A. (1972) “Brazilian Size Distribution of Income”, American Economic Review, vol. 62(1/2), pp. 391-402.

Fletcher, R. (1970), “A New Approach to Variable Metric Algorithms”, The Computer Journal, vol. 13, pp. 317-322.

Fox, K. J. (1998), “Non-Parametric Estimation of Technical Progress”, Journal of Productivity Analysis, vol. 10, pp. 235-250.

Fox, K. J. and R. Q. Grafton (2000), “Nonparametric Estimation of Returns to Scale: Method and Application”, Canadian Journal of Agricultural Economics, vol. 48, pp. 341-54

Fraley, C. and A. E. Raftery (1998), “How Many Cluster? Which Clustering Method? Answers Via Model-Based Cluster Analysis”, Technical Report No. 329, Department of Statistics, University of Washington.

Friedman, J. H. (1991), “Multivariate Adaptive Regression Splines (with Discussion)”, Annals of Statistics, vol. 19, pp. 1-141.

Fulginiti, L. and R. Perrin (1993), “Prices and Productivity in Agriculture”, Review of Economics and Statistics, vol. 75, pp. 471-482.

Fulginiti, L. and R. Perrin (1997), “Declining Productivity in LDC Agriculture”, mimeographed, Department of Agricultural Economics, University of Nebraska, Lincoln, USA.

Fulginiti, L. and R. Perrin (1998), “Agricultural Productivity in Developing Countries”, Agricultural Economics, vol. 19, pp. 45-51.

Fulginiti, L. and R. Perrin (1999), “Have Price Policies Damaged LDC Agricultural Productivity?”, Contemporary Economic Policy, vol. 17, pp. 469- 475.

Gamerman, D. (1997), Markov Chain Monte Carlo, Stochastic Simulation for Bayesian Inference, Chapman and Hall, London.

Gelman, A. and D. B. Rubin (1992), “Inference From Iterative Simulation Using Multiple Sequences”, Statistical Science, vol. 7, pp. 457-72.

Geman, S. and D. Geman (1984), “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”, vol. 6, chapter 37, pp. 721-741, in Neurocomputing: Foundations of Research, 1984, IEEE Transactions on Pattern Analysis and Machine Intelligence.

Gilks, W. R. , S. Richardson and D. J. Spiegelhalter (1996), Markov Chain Monte Carlo in Practice, CRC Press, London.

Gollin, D., S. Parente and R. Rogerson (2002), “The Role of Agriculture in Development”, American Economic Review, vol. 92(2), pp. 160-164.

Green, P.J. and Silverman, B.W. (1994), Nonparametric Regression and Generalized Linear Models, a Rougness Penalty Approach, Chapman & Hall, London.

Greene, W. (1980), “Maximum Likelihood Estimation of Econometric Frontiers”, Journal of Econometrics, vol. 13, pp. 27-56.

Greene, W. (2001a), “Fixed and Random Effects in Nonlinear Models”, Working Paper EC-01-01, Department of Economics, Stern School of Business, New York University.

Greene, W. (2001b), “New Developments in the Estimation of Stochastic Frontier Models with Panel Data”, Efficiency Series Paper 6/2001, Departmento de Economia, Universidad de Oviedo (available at: http://www.uniovi.es/eficiencia/)

Greene, W. (2002), “Fixed and Random Effects in Stochastic Frontier Models”, Working Paper EC-02-16, Department of Economics, Stern School of Business, New York University.

Greene, W. (2003), “Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Care Systems”, Working Paper EC-03-10, Department of Economics, Stern School of Business, New York University.

Greene, W. (2005), “Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model”, Journal of Econometrics, vol. 126(2), pp. 269-303.

Griffin, J.E. and M. Steel (2004), “Semiparametric Bayesian Inference for Stochastic Frontier Models”, Journal of Econometrics, vol. 123, pp. 121-152.

Grosskopf, S. (1993), “Efficiency and Productivity”, in Fried, Lovell and Schmidt (eds.), The Measurement of Productive Efficiency, pp.3-67, Oxford University Press, New York.

Guo,W. (2002), “Functional Mixed Effects Models”, Biometrics, vol. 58, pp. 121-128.

Hall, P., Härdle, W. and L. Simar (1993), “On the Inconsistency of Bootstrap Distribution Estimators”, Computational Statistics and Data Analysis, vol. 16, pp. 11-18.

Hall, P., Härdle, W. and L. Simar (1995), “Iterated Bootstrap with Application to Frontier Models”, The Journal of Productivity Analysis, vol. 6, pp. 63-76.

Hall, R. E., and C. I. Jones (1999), “Why Do Some Countries Produce So Much More Output Per Worker Than Others?”, Quarterly Journal of Economics, vol. 114, pp. 83-116.

Han, C., L. O. and P. Schmidt (2005), "Estimation of a Panel Data Model with Parametric Temporal Variation in Individual Effects", Journal of Econometrics, vol. 126(29), pp. 241-267.

Hardle, W. (1990), Applied Nonparametric Regression, Econometric Society Monographs, Cambridge University Press.

Hastie, T. J. and R. J. Tibshirani. (1990), Generalized Additive Models, Chapman and Hall, London..

Heckman, J. and B. Singer (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", Econometrica, vol. 52, pp. 271-320.

Henderson D. and A. Ullah (2004), "A Nonparametric Random Effects Estimator", mimeo, State University of New York at Binghamton.

Henderson D., A. Ullah (2005), "A Nonparametric Random Effects Estimator", Economic Letters, vol. 88(3), pp. 403-407.

Henderson, D. (2004), "Nonparametric Kernel Measurement of Technical Efficiency", mimeo, State University of New York at Binghamton.

Henderson, D. and R. Russell (2005), "Human Capital and Convergence: a Production-Frontier Approach", International Economic Review, vol. 46(4), pp. 1167-1205.

Heshmati, A. and S. C. Kumbhakar (1994), "Farm Heterogeneity and Technical Efficiency: Some Results from Swedish Dairy Farms," Journal of Productivity Analysis, vol. 5, pp. 45-61.

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

Huang, J.-C. and D. W. Nychka (2000), "A Nonparametric Multiple Choice Method within the Random Utility Framework", Journal of Econometrics, vol. 97(2), pp. 207-225.

Humphrey, D. B. and B. Vale (2004), "Scale Economies, Bank Mergers, and Electronic Payments: A Spline Function Approach", Journal of Banking and Finance, vol.28(7), pp. 1671-1696.

Insightful Corporation (2003), "S-PLUS 6.2", Reinach, Switzerland.

Johnson, P. (2005), "A Continuous State Space Approach to Convergence by Parts", Economics Letters, vol. 86, pp. 317-321.

Johnston, B.F. and J.W. Mellor (1961), "The Role of Agriculture in Economic Development", American Economic Review, vol. 51, pp. 566-93.

Jondrow, J., I. Materov, K. Lovell and P. Schmidt (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," Journal of Econometrics, vol. 19, 2/3, pp. 233-238.

Kauermann, G. (2004), "A Note on Smoothing Parameter Selection for Penalized Spline Smoothing", Journal of Statistical Planing and Inference, vol. 127, pp. 53-69.

Kawagoe, T. and Y. Hayami (1985), "An Intercountry Comparison of Agricultural Production Efficiency", American Journal of Agricultural Economics, vol. 67, pp. 87-92.

Kawagoe, T., Y. Hayami and V. Ruttan (1985), "The Intercountry Agricultural Production Function and Productivity Differences Among Countries", Journal of Development Economics, vol. 19, pp. 113-132.

Kiefer, N. (1979), "On the Value of Sample Separation Information," Econometrica, vol. 47, pp. 997-1003.

Kiefer, N. (1980a), "A Note on Regime Classification in Disequilibrium Models", Review of Economic Studies, vol. 47(1), pp. 637-639.

Kiefer, N. (1980b), “A Time Series-Cross Section Model with Fixed Effects with an Intertemporal Factor Structure”, Unpublished manuscript, Cornell University.

Kim, J.-I. and L. J. Lau, (1992a), “Human Capital and Aggregate Productivity; Some Empirical Evidence from the Group-of-Five Countries”, CEPR, No. 318.

Kim, J.-I. and L. J. Lau, (1992b), “The Importance of Embodied Technical Progress: Some Empirical Evidence from the Group-of-Five Countries”, Manuscript.

Kim, J.-I. and L. J. Lau, (1994), “The Sources of Economic Growth of the East Asian Newly Industrialized Countries,” Journal of the Japanese and International Economies, vol. 8(3), pp. 235-271.

Kim, J.-I. and L. J. Lau, (1995), “The Role of Human Capital in the Economic Growth of the East Asian Newly Industrialized Countries,” Asia-Pacific Economic Review, vol. 1(3), pp. 3- 22.

Klenow, P. and Rodriguez-Clare (1997), “The Neoclassical Revival in Growth Economics: Has it Gone Too Far?”, NBER macro annual, pp. 73-114.

Kolari, J. and A. Zardkoohi (1995), “Economies of Scale and Scope in Commercial Banks with different output Mixes”, Texas A&M Working Paper.

Koop G., J. Osiewalski and M. Steel (1997), “Bayesian Efficiency Analysis through Individual Effects: Hospital Cost Frontiers”, Journal of Econometrics, vol. 76, pp. 77-105.

Koop G., J. Osiewalski and M. Steel (1999), “The Components of Output Growth: A Stochastic Frontier Analysis”, Oxford Bulletin of Economics and Statistics, vol. 61(4), pp. 455-487.

Koop G., J. Osiewalski and M. Steel (2000), “Modeling the Sources of Output Growth in a Panel of Countries”, Journal of Business and Economic Statistics, vol. 18(3), pp. 284-299.

Koop G.and M. Steel (2001), “Bayesian Analysis of Stochastic Frontier Models”, in B. Baltagi (ed.) A Companion to Theoretical Econometrics, Blackwell, Massachusetts, pp. 520-537.

Koop, G. and D. Poirier (2004), “Bayesian variants of some classical semiparametric regression techniques, Journal of Econometrics, vol. 123, pp. 259-282.

Koopmans, T. (1951), “An Analysis of Production as an Efficient Combination of Activities” in T. Koopmans (ed.), Activity Analysis of Production and Allocation, Wiley, New York.

Koopmans, T. C. (1965), “On the Concept of Optimal Growth”, in T. C. Koopmans (ed.), The Econometric Approach to Development Planning. Rand-McNally, Chicago.

Krueger, A., M. Schiff and A. Valdés, (1992), The Political Economy of Agricultural Pricing Policy: A Synthesis of the Economics of Developing Countries, Johns Hopkins University Press, Baltimore.

Kumar, S., and R. R. Russell (2002), “Technological Change, Technological Catch-Up, and Capital Deepening: Relative Contributions to Growth and Convergence”, American Economic Review, vol. 92, pp. 527-548.

Kumbhakar S. C. and A. Heshmati (1995), “Efficiency Measurement in Swedish Dairy Farms: An Application of Rotating Panel Data, 1976-1988”, American Journal of Agricultural Economics, vol. 77, pp. 660-674.

Kumbhakar, S. C. (1990), “Production Frontiers with Panel Data and Time Varying Technical Inefficiency”, Journal of Econometrics, vol. 46, pp. 201-212.

Kumbhakar, S. C. and K. Lovell (2000), Stochastic Frontier Analysis, Cambridge University Press, Cambridge.

Kumbhakar, S. C. Ghosh, S. Mc. Gukin, T. (1991), "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in US. Dairy Farms", Journal of Business and Economic Statistics, vol. 9(3), pp. 279-86.

Kumbhakar, S.C. (1987), "The Specification of Technical and Allocative Inefficiency in Stochastic Production and Profit Frontiers", Journal of Econometrics, vol. 34, pp. 335-48.

Kumbhakar, S.C. (1991), "Estimation of Technical Inefficiency in Panel Data Models with Firm and Time-Specific Effects", Economics Letters, vol. 36, pp. 43-48.

Kumbhakar, S.C. and L. Hjalmarsson (1993), "Technical Efficiency and Technological Progress in Swedish Dairy Farms", in H.O. Fried, C.A.K. Lovell, and P. Schmidt (eds.), The Measurement of Productive Efficiency Techniques and Applications, Oxford University Press, Oxford, pp. 257-270.

Kumbhakar, S.C., Park, B.U., Simar, L. and E.G. Tsionas (2004), "Nonparametric Stochastic Frontiers: a Local Maximum Likelihood Approach", Technical Report 0418, Interuniversity Attraction Pole (IAP) Statistics Network.

Kumbhakar, S.C. and H.-J. Wang (2005), "Estimation of growth convergence using a stochastic production frontier approach", Economics Letters, Elsevier, vol. 88(3), pp. 300-305.

Laird, N. M. (1978), "Nonparametric Maximum Likelihood Estimation of Mixing Distributions", Journal of the American Statistical Association, vol. 73, pp. 805-811.

Laird, N. M. and J. H. Ware (1982), "Random-Effects Models for Longitudinal Data", Biometrics, vol. 38, pp. 963-974.

Lau, L. J. and P. A. Yotopoulos (1989), "The Meta-Production Function Approach to Technological Change in World Agriculture", Journal of Development Economics, vol. 31, pp. 241-269.

Lau, L. J. (1996) “The Sources of Long-Term Economic Growth: Observations from the Experience of Developed and Developing Countries”, in R. Landau, T. Taylor and G. Wright (eds) The Mosaic of Economic Growth, Stanford University Press.

Lau, L. J. and P. A. Yotopoulos (1988), “The Meta-Production Function Approach to Technological Change in World Agriculture”, Journal of Development Economics, vol. 31, pp. 241-269.

Lee, L. (1983), “Generalized Econometric Models with Selectivity”, Econometrica, vol. 51, pp. 507-512.

Lee, Y. and P. Schmidt (1993), “A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency,” in H. Fried and K. Lovell (eds.), The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press, New York.

Leroux, B.G. (1992), “Consistent Estimation of a Mixing Distribution”, Annals of Statistics, vol. 20, pp.1350-1360.

Li, Q. (1996), “Nonparametric Testing of Closeness between Two Unknown Distribution Functions”, Econometric Reviews, vol. 15, pp. 261-274.

Lin, X. and R. J. Carroll (2000), “Nonparametric Function Estimation for Clustered Data when the Predictor is Measured Without/With Error,” Journal of the American Statistical Association, vol. 95, pp. 520-34.

Littell, R. C., G. A. Milliken, W. W. Stroup and R. D. Wolfinger, (1996), SAS System for Mixed Models, Cary, NC: SAS Institute Inc.

Mankiw, N. G. (1995), “The Growth of Nations”, Brookings Papers on Economic Activity, vol. 1, pp. 275-310.

Mankiw, N. G., D. Romer, and D. N. Weil (1992), “A Contribution to the Empirics of Economic Growth”, Quarterly Journal of Economics, vol. 107(2), pp. 407-437.

Martin, W. and D. Mitra (2001), "Productivity Growth and Convergence in Agriculture versus Manufacturing", Economic Development and Cultural Change, vol. 49(2), pp. 403-422.

McLachlan, G. J. (1987), "On Bootstrapping the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture", Applied Statistics, vol. 36, pp. 318-324.

Meeusen, W and J. van den Broeck (1977), "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error", International Economic Review, vol. 18(2), pp. 435-444.

Mester, L. (1997), "Measuring Efficiency at US banks: Accounting for Heterogeneity is Important", European Journal of Operational Research, vol. 98, pp. 230-424.

Milanovic, B. (2002), "True World Income Distribution, 1988 and 1993: First Calculation Based on Household Surveys Alone", Economic Journal, vol. 112(1), pp.51-92.

Milanovic. B. (2005), Worlds Apart: Measuring Global and International Inequality, Princeton University Press, Princeton.

Mundlak, Y. (2000), Agriculture and Economic Growth: Theory and Measurement, Cambridge, Harvard University Press.

Nagin, D. and K. Land (1993), "Age, Criminal Careers, and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Poisson Model", Criminology, vol. 31(3), pp. 327-362.

Nina, A., C. Arndtb and P. V. Preckelb (2003), "Is Agricultural Productivity in Developing Countries Really Shrinking? New Evidence Using a Modified Nonparametric Approach", Journal of Development Economics, vol. 71, pp. 395-15.

Nychka, D. and N. Saltzman. (1998), “Design of Air Quality Monitoring Networks”, in D. Nychka, L. Cox and W. Piegorsch. (eds.), Case Studies in Environmental Statistics Nychka, Lecture Notes in Statistics, Springer-Verlag, pp. 51-76.

Nychka, D., P. Haaland, M. O'Connell and S. Ellner (1998), “FUNFITS, Data Analysis and Statistical Tools for Estimating Functions” in D. Nychka, Cox, L., Piegorsch, W. (eds.), Case Studies in Environmental Statistics Nychka, Lecture Notes in Statistics, Springer-Verlag, pp. 159-179.

O’Sullivan, F. (1986), “A Statistical Perspective on ill-Posed Inverse Problems (with Discussion)”, Statistical Science, vol. 1, pp. 505-27.

Orea, L. and S. C. Kumbhakar (2004), “Efficiency Measurement using a Latent Class Stochastic Frontier Model”, Empirical Economics, vol. 29, pp. 169-184.

Pagan, A. and A. Ullah (1999), Nonparametric Econometrics, Cambridge, Cambridge University Press.

Park, B. and L. Simar (1994), “Efficient Semiparametric Estimation in a Stochastic Frontier Model”, Journal of the American Statistical Association, vol. 89(427), pp. 929-936.

Park, B., R. C. Sickles and L. Simar (1998), “Stochastic Panel Frontiers: A Semiparametric Approach”, Journal of Econometrics, vol. 84, pp. 273 - 301.

Patterson H. D. and R. Thompson (1971), “Recovery of Inter-Block Information when Block Sizes are Unequal”, Biometrika, vol. 58, pp. 545-554.

Peeters, L. and Y. Surry (2000), “Incorporating Price-Induced Innovation in a Symmetric Generalised McFadden Cost Function with Several Outputs”, Journal of Productivity Analysis, vol. 14(1), pp. 53-70.

Pérez, G. and A. Quintana (2003), “The Dirty Little Secrets of Adaptive Bayesian Estimation”, IX Conferencia Española de Biometría, La Coruña, Seminar Outline (available at http://www.udc.es/dep/mate/biometria2003/Archivos/i03.pdf).

Pinheiro, J. C. and D. M. Bates, (2000), Mixed-Effects Models in S and S-PLUS, New York, Springer.

Pitt, M. and L. Lee (1981), "The Measurement and Sources of Technical Inefficiency in Indonesian Weaving Industry", Journal of Development Economics, vol. 9, pp. 43-64.

Poirier, D. and P. Ruud (1981), "On the Appropriateness of Endogenous Switching", Journal of Econometrics, vol. 16(2), pp. 249-256.

Powell, M. J. D. (1987), "Radial Basis Functions for Multivariable Interpolation: a Review", in J. C. Mason and M. G. Cox (eds.), Algorithms for Approximation, pp. 143-167. Oxford Clarendon Press, London.

Prescott, E., (1998), "Needed: A Theory of Total Factor Productivity", International Economic Review, vol. 39, pp. 525-551.

Quah, D. (1993), "Galton's Fallacy and Tests of the Convergence Hypothesis", Scandinavian Journal of Economics, vol. 95, pp. 427-443.

Quah, D. (1996a), "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics", Economic Journal, vol. 106(6), pp. 1045-55.

Quah, D. (1996b), "Convergence Empirics Across Economies with (Some) Capital Mobility", Journal of Economic Growth, vol. 1, pp. 95-124.

Quah, D. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", Journal of Economic Growth, vol. 2, pp. 27-59.

Quandt, R. and J. Ramsey (1978), "Estimating Mixtures of normal Distributions and Switching Regressions", Journal of the American Statistical Association, vol. 73, pp. 730-738.

R Development Core Team (2005), "R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing", Vienna, Austria (available at http://www.R-project.org/).

Rao D. (1993), Intercountry Comparisons of Agricultural Output and Productivity, FAO Economic and Social Development Paper, No. 112, Rome.

Reifschneider, D. and R. Stevenson (1991), "Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency", International Economic Review, vol. 32, pp. 715-23.

Restuccia, D., D. T. Yang and X. Zhu (2004) "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis", NBER Summer Institute Workshop, Economic Fluctuations and Growth, Cambridge, Massachusetts.

Rice, J. A., and C. O. Wu (2001), "Nonparametric Mixed Effects Models for Unequally Sampled Noisy Curves", Biometrics, vol. 57, pp. 253-259.

Richmond, J. (1974), "Estimating the Efficiency of Production", International Economic Review, vol. 15, pp. 515-521.

Roeder, K. and Wasserman, L. (1997), "Practical Density Estimation Using Mixtures of Normals", Journal of the American Statistical Association, vol. 92, pp. 894-902.

Roeder, K., K. Lynch and D. Nagin (1999), "Modeling Uncertainty in Latent Class Membership: A Case Study in Criminology", Journal of the American Statistical Association, vol. 94, pp. 766-776.

Ruppert D., M. P. Wand, and R. J. Carroll (2003), Semiparametric Regression, Cambridge University Press, Cambridge.

Ruppert, D. (2002), "Selecting the Number of Knots for Penalized Splines", Journal of Computational and Graphical Statistics, vol. 11, pp. 735-757.

Sala-i-Martin, X. (2002), The Disturbing "Rise" of Global Income Inequality. NBER Working Paper 8904. (Cambridge: NBER).

SAS Institute Inc (2004), "SAS/STAT User's Guide", Cary, NC. Version 8.

Schmidt, P. and R. Sickles (1984), “Production Frontiers with Panel Data”, Journal of Business and Economic Statistics, vol. 2, pp. 367-374.

Schmidt, P. and Y. Kim (2001) “Intervalos de Confianza por Bootstrap para Niveles de Eficiencia en Modelos de Frontera Estocástica con Datos de Panel” in A. Pinilla (ed.) La Medición de la Eficiencia y la Productividad, Madrid: Ed. Piramide, pp. 283-298.

Schultz, T. W. (1964), Transforming Traditional Agriculture, New Haven, Yale University Press.

Sena, V. (2003), “The Frontier Approach to the Measurement of Productivity and Technical Efficiency”, Economic Issues, vol. 8(2), pp. 71-97.

Shanno, D. F. (1970), “Conditioning of Quasi-Newton Methods for Function Minimization”, Mathematics of Computation, Boston, vol. 24, pp. 647-657.

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

Simar, L. (1992), “Estimating Efficiencies from Frontier Models with Panel Data : a Comparison of Parametric, Non-parametric and Semi-Parametric Methods with Bootstrapping”, Journal of Productivity Analysis, vol. 3, pp. 167-203.

Simar, L. and P. Wilson (1998), “Sensitivity of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models”, Management Sciences, vol.44 (1), pp. 49-61.

Simar, L. and P. Wilson (2000), “Statistical Inference in Nonparametric Frontier Models: The State of the Art”, Journal of Productivity Analysis, vol. 13, 49-78.

Simpson, T. W., D. K. J. Lin and W. Chen (2001), “Sampling Strategies for Computer Experiments: Design and Analysis”, International Journal of Reliability and Applications, vol. 2(3), pp. 209-240.

Solow, R. M. (1956), “A Contribution to the Theory of Economic Growth”, Quarterly Journal of Economics, vol. 70, pp. 65-94.

Spiegelhalter, D., A. Thomas, N. Best (2003), "WinBUGS Version 1.4 User Manual", Medical Research Council Biostatistics Unit, Cambridge, UK.

Stern, N. H. (1996), Growth Theories, Old and New: and the Role of Agriculture in Economic Development, Food and Agriculture Organization of the United Nations, Rome.

Stevenson, R. (1980), "Likelihood Functions for Generalized Stochastic Frontier Functions", Journal of Econometrics, vol. 13, pp. 57-66.

Tanggaard, C. (1997), "Nonparametric Smoothing of Yield Curves", Review of Quantitative Finance and Accounting, vol. 9(3), pp. 251-67.

Tanner, M. A. (1996), Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions, 3rd ed, Springer-Verlag, NewYork.

Tierney, L. (1994), "Markov Chains for Exploring Posterior Distributions", The Annals of Statistics, vol. 22, pp. 1701-1722.

Timmer, C. P. (1988), "The Agricultural Transformation", in H. Chenery and T. N. Srinivasan (eds), Handbook of development economics, vol. 1, Amsterdam: Elsevier Science Publishers, pp. 275-331.

Trueblood, M.A. (1996), An Intercountry Comparison of Agricultural Efficiency and Productivity, PhD dissertation, University of Minnesota.

Tsionas, M. (2002), "Stochastic Frontier Models with Random Coefficients", Journal of Applied Econometrics, vol. 17, 127-147.

Ullah, A. (2001), "Nonparametric Kernel Methods of Estimation and Hypothesis Testing", in B. Baltagi (ed.), A Companion to Theoretical Econometrics, Blackwell, Oxford, pp. 429-443.

Ullah, A. and N. Roy (1998). "Nonparametric and Semiparametric Econometrics of Panel Data", Handbook of Applied Economics Statistics, pp. 579-604.

van den Broeck, J., G. Koop, J. Osiewalski, and M. Steel (1994),"Stochastic Frontier Models: A Bayesian Perspective", Journal of Econometrics, vol. 61, pp. 273-303.

Verbeke, G. and Molenberghs, G. (2000), Linear Mized Models for Longitudinal Data, Springer-Verlag, New York.

Verbyla, A. P., B. R. Cullis, M. G. Kenward and S. J. Welham (1999), "The Analysis of Designed Experiments and Longitudinal Data by Using Smoothing Splines (with Discussion)", Journal of the Royal Statistical Society, Series C, vol. 48, pp. 269-312.

Wahba, G. (1990), Spline Models for Observational Data, Philadelphia: SIAM.

Wand, M. (2003), Smoothing and Mixed Models, Computational Statistics, vol. 18, pp. 223-249.

Wand, M. (2005), "Information for Generalised Linear Mixed Models", Working paper, Department of Statistics, School of Mathematics, University of New South Wales, Sydney, Australia (available at http://web.maths.unsw.edu.au/~wand/infopap.pdf).

Wand, M. and M. C. Jones (1995), Kernel Smoothing, Chapman & Hall, London.

Wang, P., I. Cockburn, and M. Puterman (1998), "Analysis of Patent Data - A Mixed Poisson Regression Model Approach", Journal of Business and Economic Statistics, vol. 16(1), pp. 27-41.

Wedel, M., W. DeSarbo, J. Bult, and V. Ramaswamy (1993), "A Latent Class Poisson Regression Model for Heterogeneous Count Data", Journal of Applied Econometrics, vol. 8, pp. 397-411.

Wood, S. N. (2003), "Thin-Plate Regression Splines", Journal of the Royal Statistical Society (B), vol. 65(1), pp. 95-114.

Wood, S. N. (2006), "Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models", Biometrics, vol. 62(4), pp. 1025-1036.

World Bank (1992), "The Plundering of Agriculture in LDCs", Development Brief, 3.

World Development Indicators (1998), World Bank Database, Washington D.C., The World Bank.

World Development Report (1992), Development and the Environment, World Bank Group, Oxford University Press, Oxford.

Young, A. (1995), "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience", Quarterly Journal of Economics, vol. 110(3), pp. 641-680.

Zeger S. L. , Diggle P. J. (1994), "Semiparametric Models for Longitudinal Data with Application to CD4 Cell Numbers in HIV Seroconverters", Biometrics, vol. 50, pp. 689-699.

Zhang D., X. Lin, J. Raz, M. F. Sower (1998), "Semiparametric Stochastic Mixed Models for Longitudinal Data", Journal of the American Statistical Association, vol. 93, pp. 710-719.

Zhang, P. (1997), Comment on "An Asymptotic Theory for Linear Model Selection", Statistica Sinica, vol. 7, pp. 254-258.

Annexes – Data Sources

For the economy as a whole, we use the following data sources:

(i) Gross Domestic Product at 1990 constant USD is built from Heston *et al.* (2002);

(ii) Economy-Wide Fixed Capital series at 1990 constant USD is drawn from Crego *et al.* (1998);

(iii) Total Labor Force is obtained from World Development Indicators (WDI).

For agriculture, the following data sources were used:

(i) total output at 1990 constant USD is built from Rao (1993, p. 74) and FAOSTAT (2001);

(ii) capital series measured in 1990 thousands USD, agricultural labor and land data measured in hectares are drawn from Martin and Mitra (2001).