South African Journal of Economics Vol. 76:S2 August 2008

SECTORAL ELASTICITY OF SUBSTITUTION AND RETURNS TO SCALE IN SOUTH AFRICA

TEWODROS G. GEBRESELASIE*

Abstract
Elasticity of substitution and returns to scale are estimated on a sectoral basis for South Africa using panel-based generalised least square procedure. Apart from sectoral differences in terms of elasticity of substitution, the study found that elasticity of substitution is below unity in all of the sectors. Most of the sectors studied are found to have increasing returns to scale in production. The study further explores the implications of elasticity of factor substitution and returns to scale on growth and employment creation. It is argued that a greater number of jobs can be created from growth of sectors with constant or decreasing returns to scale than from the same level of output growth generated by sectors with increasing returns to scale. This is the case when the employment-creating potential of the same amount of additional output is compared in all the sectors examined. By virtue of scale economies, a sector like finance, insurance, real estate and business services generates more output with less proportional increase in inputs, which means growth in this sector may not have the desired impact on job creation. However, given the sector's large share (20%) of the country's total output and employment, it may generate more jobs, even if sectors like utilities and construction experience the same level of output growth. Given its importance for growth and employment, the study recommends further investigation into the reasons why elasticity of substitution is lower in sectors like utilities, mining and trade, catering and accommodation services.
JEL Classification: C23, D24
Keywords: Elasticity of substitution, returns to scale, factor income share, South Africa

1. INTRODUCTION

Many reasons have been put forward to explain why unemployment in South Africa (SA) is high and persistent, including labour market conditions, increases in productivity and skills shortages. Even if not sufficient to dramatically reduce unemployment, formal sector employment has grown robustly relative to GDP growth at least since 2002. The Labour Force Survey (LFS) shows that formal sector employment has grown at 2.9% per annum relative to an average GDP growth rate of 4.2% over the period 2002 to 2006. The ILO (KILM19) shows that the employment elasticity in sub-Saharan Africa is among the highest in the world. Based on the LFS conducted between 2000 and 2005, Altman (2007) shows that employment elasticity in SA is higher in the formal, non-agricultural, private sector. She further says, "This seems surprising in the context of a capital intensive economy like South Africa". Future employment trends will not necessarily mirror historical ones: these will partly depend on the pace and character of economic growth.

---

* Research Specialist, Employment, Growth and Development Initiative, Human Sciences Research Council. Contact detail: tgselasie@yahoo.com. I would like to thank Rob Davies and Miriam Altman for their valuable comments and contributions. I am also indebted to the anonymous referee for his valuable comments on the earlier draft. All errors and omissions remain my own.

A deeper understanding of employment-productivity trade-offs or complementarities are needed. Elasticity of factor substitution and returns to scale, particularly at the sectoral level, may provide some insights into the potential pace of formal employment creation in the future. However, there has been an implicit assumption that there is ease of factor substitution (*e.g.* capital substituting for labour, or unskilled labour substituting for skilled) and constant returns to scale in production. This paper contributes to the debates on the link between employment and economic growth by measuring both elasticities of substitution and returns to scale at a sectoral level.

Economic sectors differ in terms of job-creating potential. Manufacturing's contribution to total employment creation is falling, while services sectors have gained prominence in terms of employment creation. A part of the current trend might be explained by the sectoral elasticities of substitution and returns to scale estimated in this study. Panel econometric techniques are applied in an attempt to estimate both these parameters underlying the technology of each of the nine sectors given in Table 1. The table also highlights each sector's contribution to total value added and total employment.

Elasticity of substitution can be defined briefly as the ease with which factors are substitutable for one another. Some sectors may find it relatively easier to replace capital with labour, while others may not be able to raise their production without more investment into capital equipment. Sectors with greater latitude in terms of replacing their capital with labour or increasing their production level with relatively more labour than capital have huge potential to raise employment levels in SA. These are the sectors that are expected to exhibit higher elasticities of factor substitution. Hence, the differential elasticities observed across sectors might be attributed to a number of factors that either promote or hinder the substitutability of one factor for another. The extent to which firms employ more labour is partly influenced by the degree to which the labour is unionised or the labour market is regulated. The easier it becomes for firms to 'hire and fire', the higher the elasticity of substitution would be. This is believed to enhance the employment creation capability of a firm.

Returns to scale in production refer to the way in which output changes when the inputs into production all change by a uniform proportion. Increasing, constant or decreasing returns depend on whether output increases by a greater, the same or a smaller proportion than inputs. These are traditionally seen as properties of the production function that describes the technology used by the individual production unit. Economists have historically identified a number of sources of increasing returns.

Table 1. Sectoral contribution to total value added and employment of South Africa in 2006 (in percentages)

Sector	Contribution to total value added	Contribution to total employment
Agriculture, forestry & fishing	2.38	8.08
Mining & quarrying	6.42	4.67
Manufacturing	17.96	13.59
Utilities (Electricity, gas & water)	2.36	0.52
Construction	3.30	4.70
TCA (Trade, catering & accommodation services)	15.35	17.15
TSC (Transport, storage & communication)	10.78	3.63
FIRB (Finance, insurance, real estate & business services)	21.61	17.93
CSP (Community, social & personal services)	19.80	29.69

Source: Quantec Research.

"Internal economies" are effects that are internal to the firm: increasing the firm's size may allow it to capture technical and engineering efficiencies that escape smaller firms. "External economies", on the other hand, do not depend on the size of the individual firms but rather on the size of the industry and arise because of the impact on efficiency of interactions among firms. They are sometimes referred to as "pecuniary" economies. Thus, the costs of intermediate inputs may fall as the industry expands because supplying firms are able to capture and pass on internal economies because of increased market size. Economists have also identified scale effects that are external to both the firm and the industry. Some of these may arise because of infrastructure that can be provided more efficiently to larger and denser markets.

In spite of the three alternative representations of scale economies, there is a tendency to assume constant returns to scale in practice. This is mainly due to the perfect competition assumption that has dominated economic theory for a long time. One of the main reasons why economists assume constant returns to scale is the "evolutionary argument". If there were increasing returns to scale, we would expect a firm that is slightly bigger than its competitors to have a cost advantage, which would allow it to out-compete rivals and thus become a monopoly. Even in large markets we would expect imperfect competition more frequently than we see here. Furthermore, decreasing and increasing returns to scale are usually not favoured because of the extreme technology specifications they represent. Firms, however, may have one of the three alternative technology representations, and assuming one while another relatively captures the technology would result in incorrect policy prescriptions. If, for instance, we assume constant returns to scale for most of the industries in SA while these industries actually exhibit increasing returns to scale, we will end up with ineffective policies and objectives that are unattainable within the original timeframe. This can be demonstrated better by using the Accelerated and Shared Growth Initiative – South Africa (ASGISA) objectives regarding growth and unemployment. ASGISA aims at halving unemployment by 2014. In order to attain this objective, the economy must grow at an annual average rate of 4.5% between 2005 and 2009. During the subsequent period until 2014, the required rate of growth must be scaled up to at least 6% per annum.[1] If these objectives were set by assuming constant returns to scale because of convenience and the moderate form of technology they capture, and if the economy, because of a number of factors, exhibits increasing returns to scale, the objective of reducing unemployment from around 27.9%[2] to about 14% could not be achieved, even if the economy grew at the rates initially projected. This would mean that each year real GDP would have to grow at more than the rates indicated above in order to attain the objective of halving unemployment. On the other hand, the required quantum of GDP growth will be lower if the economy actually exhibits decreasing returns to scale and the existing unemployment objective was set assuming constant returns to scale.

Accordingly, using models that are built on constant returns to scale is tantamount to saying that perfect competition characterises the SA economy. As a result of the non-

[1] A Summary: Accelerated and Shared Growth Initiative – South Africa. The Presidency, Republic of South Africa.

[2] According to the latest Labour Force Survey (March 2007), SA's unemployment rate was 27.9% in March 2004.

competitive nature of product and factor markets, non-constant returns to scale are believed to offer a sound analytical framework for incorporating market imperfections. One should remember that the assumption of constant returns to scale might also avoid increasing returns to scale as a possible contributor to growth. There are a number of mechanisms at work that allow the economy to display increasing returns to scale. According to Ros (2003), returns to scale may increase because of a number of factors, even if the firms constituting the entire economy operate at constant returns to scale individually. He indicated that industrial training and quality of management might be considered as some of the possible causes of external scale economies. The former enhances the stock of production experience while the latter increases the number and quality of managers that can boost the productivity of modern technology. Accordingly, since this study is conducted at a more aggregate (sectoral) level, we can reasonably expect increasing returns to scale, at least in some sectors, as the result of scale economies at the firm, industry or sectoral level.

The lack of success of many empirical models may explain in part why increasing returns to scale are gaining more popularity. In the past, many economic models (*e.g.* the ACMS[3] method of Arrow *et al.*, 1961 and the neoclassical growth model of Solow, 1956) were developed with the assumption of constant returns to scale. As these theories fail to sufficiently explain economic realities, economists (*e.g.* Arrow, 1962; List and Zhou, 2007) have started to revisit them by placing increasing returns to scale at the core of their investigation.

As indicated above, estimating sectoral elasticities of substitution and returns to scale is the principal objective of this research endeavour. It also attempts to shed some light on the link between these parameter estimates and employment creation. The study contributes to the literature on SA, and to the best of the author's knowledge, no similar study has been conducted at the same level of disaggregation in the past. It is also expected that the computable general equilibrium models already used in SA would benefit a great deal from the results of these estimations.

This study is organised as follows. Section 2 reviews the theoretical background of the model used in this study to estimate returns to scale and elasticity of substitution for each of the nine sectors in SA. The review also sheds some light on the magnitudes of these parameters in other countries where similar studies have been conducted. Section 3 briefly discusses the variables and the specific empirical model used in the study. The empirical results are presented in section 4. Section 5 identifies the major policy implications of our findings, while section 6 concludes.

2. LITERATURE REVIEW

Decomposition of growth usually involves the use of production functions. Growth in output can be attributed to an increased level of usage of the various factors of production (mainly labour and capital), changes over time of factor shares, increased productivity, or

---

[3] ACMS stands for the first alphabets taken from the surnames of economists (ARROW, J. K., CHENRY, H. B. MINHAS, B. S. and R. M. SOLOW) who originally introduced the ACMS model for estimating elasticity of substitution.

a combination of all three.[4] Elasticity of substitution enters into the growth equation when the income shares of labour and capital are determined. Assuming labour and capital as the only factors of production, the income shares of each of these factors have been assumed to remain constant over the years because of the accordance of the Cobb-Douglas (CD) function with the most prominent stylised facts of long-term economic development (Klump *et al.*, 2007). This might be a very restrictive assumption, since changes in factor prices and technology may induce factor income shares to vary across time (Lee, 1970), making the use of the CD functional form inappropriate. On the contrary, the constant elasticity of substitution (CES), apart from allowing non-unitary elasticity of substitution, provides the mechanism to evaluate the impact of changes in factor income shares and technological bias. Technological changes may be termed capital biased, labour biased or neutral. However, the estimations we carried out in this paper assume neutral technological change.

The motivation for estimating the returns to scale parameter lies in exploring the labour-absorption capacity of the different sectors. The relationship that exists between the returns to scale and labour demand can be seen with the help of a CD functional form, *i.e.* $Q = AK^{\alpha}L^{\beta}$. Q denotes the level of production while K and L represent each of the factors of production capital and labour, respectively. α and β measure the elasticity of output Q to capital and labour, respectively. A is an index of disembodied technical efficiency. Using the conditional labour demand derived from this function, the relationship between the growth in output and the growth in labour demand can be shown as:

$$g_L = \frac{1}{\alpha + \beta} g_Q$$

Where g_L and g_Q represent growth in labour and growth in output, respectively. This means that conditional demand for labour will grow more slowly than, the same as, or faster than output, depending on whether $\alpha + \beta$ is greater than, equal to, or less than unity, *i.e.* whether there are increasing, constant or decreasing returns to scale.

It appears that the bulk of the empirical studies conducted in the past identified non-unitary elasticity of substitution and highlighted differences across sectors in terms of elasticity of substitution. Klump *et al.* (2007) indicated that in nine out of 11 empirical studies reviewed in their study the US economy recorded elasticity of substitution below unity. There is an assumption of Hicks-neutral technology behind the remaining two studies that resulted in higher elasticity estimates (elasticity ≥ 1). The review further included nine more studies carried out on European countries and Australia. A considerable number of the elasticities estimated for these countries lie below unity. Sapir (1980) estimated elasticity of substitution for Yugoslavia's manufacturing sector and found a value of 0.319. Chen (1977), however, obtained elasticity of substitution which is above one for the Hong Kong manufacturing sector. Duffy and Papageorgiou (2000) conducted a cross-sectional study of 82 countries, with their findings suggesting that

---

[4] Given the neoclassical production function X = AF(K,L), growth in output could be expressed as

$$\Delta R_X = s_K \Delta R_K + s_L \Delta R_L + R_K \Delta s_K + R_L \Delta s_L + \Delta R_A + \Delta s_K \Delta R_K + \Delta s_L \Delta R_L$$

Where R represents growth, A is an index for disembodied technical progress, and s_K and s_L are income shares of capital and labour respectively (Sapir, 1980).

countries at a higher level of development exhibit elasticity of substitution greater than one. Based on four sub-samples grouped according to the initial level of capital–labour ratio, they found that rich countries exhibited elasticity of substitution that is significantly greater than unity, whereas the poorest group of countries displayed elasticity of substitution that is significantly lower than unity. All of these studies mainly demonstrate the lack of strong evidence that the CD functional form does indeed capture production technology in most countries.

Some attempts have been made in SA to determine the elasticity of substitution and returns to scale at the higher level of aggregation. Based on estimating a cost function, Du Toit and De Wet (2002) derived an aggregate CD function for SA and concluded that the production technology on the aggregate exhibits decreasing returns to scale. They also found a unitary elasticity of substitution for SA using Kmenta's Taylor approximation of the CES function. However, Fedderke (2002, 2005) in his studies on growth generated total factor productivity by allowing factor shares to vary across time. By doing so he avoided the very restrictive assumption implied by the CD's unitary elasticity of substitution, *i.e.* constancy of factor shares over time. Unlike previous studies, this study aims at estimating sectoral elasticity of factor substitution and returns to scale with the help of the CES production function.

The growth implication of finding an elasticity of substitution different from one is discussed in the literature in detail. Klump *et al.* (2007), for instance, have indicated that assuming an aggregate CES production function with an elasticity of substitution above unity is the easiest way to generate perpetual growth. Scarce labour is substituted by capital, which results in the marginal product of capital remaining above zero in the long run. In the standard neoclassical growth model, a CES function with an elasticity of substitution below unity has been shown to generate multiple growth equilibra and development traps (Azariadis, 1996; Duffy and Papageorgiou, 2000). Overall, considering the elasticity of substitution as "a menu of choice available to entrepreneurs", the higher it becomes, the more favourable it is for growth. It can also be regarded as a measure of efficiency of the productive system (Klump and Preissler, 2000).

3. MODEL SELECTION

The CD production function has been the workhorse of neoclassical growth theory. It mainly gained popularity as one can easily employ ordinary least square for its estimation, provided that the data on all the variables are stationary. Klump *et al.* (2007) indicate another important reason why growth economists use the CD function: it accords with "the most prominent stylised facts of long-term economic development, namely the approximate constancy of factor-income shares during a steady increase in capital intensity and per capita income, independent of the direction of technological change". However, the CD form is riddled with restrictive assumptions. The first and most obvious limitation of the model is related to restricting elasticity of substitution to one. The second limitation, which primarily relies on the first, is that factor shares are kept constant over the years.[5] Restricting factor income shares from varying over the years prohibits one from evaluating the impact of technological bias on growth.

[5] Or with the non-unitary elasticity of substitution, the possibility of constancy of factor income shares must be offered by assuming purely labour-augmenting technical progress. As we know,

The CES production function, on the other hand, appears to surpass the limitations of the CD functional form. There is no restriction on the value of the elasticity of substitution, while it further provides for variation of factor income shares over the years, hence the revival of the CES function in growth theory. Klump *et al.* (2007) indicate that the importance of the CES function in growth theory is intimately linked to "normalisation", *i.e.* the choice of baseline values for the level of production inputs and factor-income shares. They further argue that normalisation paved the way for the elasticity of substitution to be considered as an important determinant of growth, while the normalisation property also enables one to evaluate the other important growth determinant – biases in technological change.

It is because of its implication for growth and the growing evidence of non-unitary elasticity of substitution in empirical research that we mainly aim at measuring the sectoral elasticity of substitution in SA. As indicated earlier, elasticity of substitution varies across sectors. Sectors with higher elasticity of substitution are destined to grow faster due to the enhanced efficiency with which factors are substitutable for one another. Klump and Preissler (2000) indicate three different effects of the elasticity of substitution on growth. These are a threshold effect, an effect on the level of the steady state and an effect on the speed of convergence. The threshold effect is related to the existence and stability of the steady state and the possible emergence of permanent growth. They further provide mathematical proof that a higher elasticity of substitution makes permanent growth more probable and increases the possibility that a permanent decline is prevented.

Against this background, one can easily see that estimating the elasticity of substitution is possible with the help of the CES production function. This functional form provides for the variation of factor shares across time and allows us to evaluate the direction of bias in technology. Nevertheless, direct estimation of both parameters of interest from the CES function is difficult, as these parameters enter into the model in non-linear form. As indicated earlier, this is the main reason why most empirical studies use the relatively simpler CD function to capture production technologies. A number of estimating methods have been developed in the past, as briefly discussed below.

The ACMS method of Arrow *et al.* (1961) represents one of these models and allows the estimation of elasticities from the relation between average labour productivity and real wages. They developed the estimating equation from one of the necessary first-order optimising conditions of the CES production function. The specific CES function they use is given as:

$$V = \gamma[\delta K^{-\rho} + (1-\delta)L^{-\rho}]^{-1/\rho} \tag{1}$$

In equation (1), *V, K and L* represent value added, capital and labour, respectively. γ, δ and ρ are important parameters of the model that denote efficiency, distribution and substitution parameters, respectively. The fact that the degree of homogeneity does not appear in (1) suggests that the returns to scale parameter is assumed to be one. However,

technological progress can be neutral, capital or labour augmenting. If we assume a Hicks-neutral technological progress, the factors governing the change in factor shares are the elasticity of substitution, the direction of change in the capital–labour ratio and the extent of non-constant returns to scale.

this method encounters criticism because it makes implausible assumptions, such as constant returns to scale, equality of factor prices to the values of their marginal products and exogeneity of wage rates (Feldstein, 1967).

Because of the difficulties encountered in estimating the CES function, Taylor's series expansion of the CES function (also known as the Kmenta approach) has been applied in the past to retrieve the parameters of interest (*e.g.* Feldstein, 1967; Chen, 1977). However, this approach also has important limitations. Besides being an approximation of the CES function, Feldstein (1967) argues that the elasticities estimated using this method are unstable because of correlation of the third variable (the square of capital-labour ratio) with the other explanatory variables, labour and capital. Instead of relying on the results from this method, Sapir (1980) decided to complement the non-linear least square method used in his study by the Kmenta approach.

The non-linear estimation procedures represent another group of estimating methods. At least two approaches can be cited that employ non-linear estimation techniques. One of these non-linear estimation techniques is discussed in detail in Bodkin and Klein (1967) and the other appears in Feldstein (1967). Although parameter estimates in these non-linear estimation techniques involve linear procedures, both techniques use different parameter selection criteria from the ones of linear least-square estimation technique, and the parameter estimates do not have the minimum variance unbiased properties. One must remember that these procedures generate only approximations to the least-square values. Moreover, the issue of simultaneous equation bias as the result of endogeneity of capital and labour should be mentioned as an important limitation of these non-linear estimation techniques.

Comparing the limitations of the estimating models discussed above, it seems relatively easy to avoid the main criticisms of the ACMS model. It is possible to introduce to the ACMS model non-constant returns to scale and market imperfections in the product and input markets. The resulting model is an improvement on the ACMS method and builds on the CES function that is attributed to Brown and de Cani (1963). Brown and de Cani's specification of the CES function, as given below, does not make any assumption about constant returns to scale or equality of factor prices and the values of their marginal products.

$$V = \gamma[\delta K^{-\rho} + (1-\delta)L^{-\rho}]^{-\upsilon/\rho} \tag{2}$$

Equation (2) is similar to (1) except for υ that measures the degree of homogeneity or returns to scale. The estimating model is derived by equating each of the partial derivatives of the CES function with respect to the inputs and their respective partial derivatives of the profit function. Accordingly, we can employ one of the following expressions to estimate elasticity and returns to scale parameters. Ideally, each of these expressions should provide similar parameter estimates.

$$\ln L = \beta_0 + \beta_1 \ln w + \beta_2 \ln\left[\frac{1+\in_{wL}}{1+\in_{PY}}\right] + \beta_3 \ln V \tag{3}$$

$$\ln K = \varphi_0 + \varphi_1 \ln r + \varphi_2 \ln\left[\frac{1+\in_{rK}}{1+\in_{PY}}\right] + \varphi_3 \ln V \tag{4}$$

Combining (3) and (4) provides the following expression to estimate the relevant parameters in terms of the capital–labour ratio.

$$\ln(K/L) = \alpha_0 + \alpha_1 \ln(w/r) + \alpha_2 \ln\left[\frac{1+\in_{wL}}{1+\in_{rK}}\right] \quad (5)$$

Where,

$$\beta_0 = \sigma \ln[\upsilon(1-\delta)\gamma^{-\rho/\upsilon}]$$

$$\varphi_0 = \sigma \ln[\upsilon\delta\gamma^{-\rho/\upsilon}]$$

$$\alpha_0 = \sigma \ln\left[\frac{\delta}{(1-\delta)}\right]$$

$$\beta_1 = \varphi_1 = \beta_2 = \varphi_2 = -\alpha_1 = -\alpha_2 = -\sigma$$

$$\beta_3 = \varphi_3 = \frac{1+\sigma(\upsilon-1)}{\upsilon}. \quad (6)$$

ln w, *ln r* and *ln V* represent the natural log of real wages, real capital return and value added, respectively. σ denotes elasticity of substitution. $\in_{wL}$ represents the elasticity of the wage rate with respect to labour employed and $\in_{rK}$ is the elasticity of return on capital with respect to capital employed. Similarly, $\in_{PY}$ is the elasticity of price with respect to output. The average nominal wage rate and interest rate in each industry must be deflated by the product price index prevailing in the specific industries or sectors for which the estimation is made. υ as it appears in (6) measures the degree of homogeneity or returns to scale prevailing in each of the industries.

Equation (3) will be estimated for each of the nine sectors given in Table 1. Time series and cross-sectional studies have been widely used in empirical studies of this kind. In this paper, we use a panel econometric technique for each of the sectors, with their constituent industries as cross-sections. By doing so, we benefit from the higher degree of freedom provided by panel-based estimation techniques. As indicated above, the elasticity of substitution is estimated directly from the model and is equivalent to $-\beta_1$. Direct estimation is not possible for the returns to scale parameter. It is rather calculated using (6)[6] once β_1 and β_3 are known from equation (3).

6 *Alternatively, the returns to scale parameter,* υ, is calculated using:

$$\upsilon = \frac{1+\beta_1}{\beta_1+\beta_3}.$$

4. DATA AND UNIVARIATE ANALYSIS

(a) Data and Non-stationarity Issues

Data on the number of employees, capital, remuneration per employee and value added have been collected from Quantec Research for the period 1970 to 2005. Employment is the average number of employees per year, while all the other variables are in constant 2000 prices.

Panel unit root tests are applied to deal with nonstationarity of the variables included in the estimating equation. Some of these tests provide for independent autoregressive terms while others assume common autoregressive terms for all the cross-sections. We estimate the Im, Pesaran and Shin (2003) test, hereafter IPS, from the former class of tests and the Levin, Lin and Chu (2002) test, hereafter LLC, from the latter group. However, both these tests have a null hypothesis of "no unit root". It is decided to apply one more test – the Hadri (2000) test – with the null hypothesis of "unit root".

The test results are presented in Appendix 1. A variable is regarded stationary if it passes the test of unit root by at least one of the three tests described in the last paragraph. This criterion is used in conjunction with the rule of rejection of the null hypothesis at the 10% level of significance.

Most of the panel variables are found to be stationary, *i.e.* integrated of order zero. This makes it possible to apply generalised least squares in estimating the elasticity of factor substitution and returns to scale for the nine sectoral Standard Industrial Classifications. Moreover, it is emphasised that parameter estimation is possible as long as two non-stationary variables exist one on each side of the equation.

(b) Estimation Results

The results in Table 2 are obtained by estimating (3) using generalised least square for all the sectors except for agriculture, forestry and fishing. A panel estimation procedure has not been followed for agriculture because of the lack of disaggregated data. We control for heterogeneity of the various industries contained in each sectoral classification (for instance, 28 industries under the manufacturing sector) and the potential for correlation among cross sections and across time.

It is the availability of disaggregated data that determines panel size across sectors. In estimating for the manufacturing sector, the panel employs data for 28 manufacturing industries, with observations running from 1970 to 2005, *i.e.* a total of 1,008 observations. In the case of mining and CSP (community, social and personal services), the panel includes three industrial classifications for the same time period, providing 108 observations in total. For the rest of the sectors, the size of the panel involves only 72 observations.

Coefficients in the first column represent the elasticity of substitution for each of the nine sectors. Returns to scale are given in column five.

The returns to scale (RTS) parameters do not have standard errors since they are not estimated directly from the model. They are derived by applying equation (6).

As shown in Table 2, all sectoral elasticities of substitution, except for agriculture, forestry and fishing, are statistically significant at 1% level and lie below unity. The statistical insignificance of the parameter estimates for agriculture is not surprising, given the lack of disaggregated data necessary for panel estimation. However, all sectoral elasticities of substitution have economically consistent signs, including agriculture, forestry and fishing.

Table 2. Estimation results

Column	1	2	3	4	5	6
Sectors	**ln_rwage**	**ln_vadd**	**Time**	**Constant**	**RTS**	**Wald Chi-sq (3)**[7]
Agriculture, forestry & fishing♠	−0.0657	0.0221	−0.0066***	14.2407***		
	(−0.0458)	(−0.0480)	(−0.0021)	(0.4323)		
Mining	−0.3746***	0.7195***	−0.0105***	9.1787***	1.2564	233.59
	(0.0467)	(0.0542)	(0.0039)	(0.6691)		
Manufacturing	−0.6837***	0.8446***	−0.0042***	11.1213***	1.1017	77778.18
	(0.0037)	(0.0031)	(0.0004)	(0.0378)		
Utilities (Water, gas & electricity)	−0.1252***	0.7363***	−0.0284***	5.9332***	1.3062	279.53
	(0.0262)	(0.0447)	(0.0039)	(0.4065)		
Construction	−0.9158***	0.9363***	−0.0137***	13.4786***	1.0344	2052.79
	(0.0224)	(0.0269)	(0.0026)	(0.2665)		
TCA (Trade, catering & accommodation services)	−0.5464***	0.7540***	0.0007	10.9558***	1.1891	702.66
	(0.0754)	(0.0491)	(0.0020)	(0.3776)		
TSC (Transport, storage & communication)	−0.8770***	0.8990***	−0.0221***	13.3533***	1.0569	1399.76
	(0.0574)	(0.0536)	(0.0045)	(0.9883)		
FIRB (Finance, insurance, real estate & business services)	−0.7286***	0.4153***	0.0202***	15.7235***	1.5112	2304.39
	(0.0243)	(0.0518)	(0.0029)	(0.6501)		
CSP (Community, social & personal services)	−0.7896***	1.3277***	−0.0244***	8.2590***	0.8452	563.22
	(0.0456)	(0.0830)	(0.0030)	(0.7972)		

Notes: Figures in brackets are standard errors.
*** signifies statistical significance at 99% confidence interval.
♠ Estimation results for this sector are based on ordinary least square. Panel estimations could not be carried out because of lack of data at sub-sector or industries level. The results for the other sectors are based on generalised least square (GLS) method.

It is worth mentioning at this point that the specific CES functional form that we employed in this study assumes a Hicks-neutral technology. The CES production function that gives rise to equation (3) does not take into account factor-augmenting technological change.

Regarding the relationship between the elasticity of substitution and efficiency, Klump and Preissler (2000) refer to de la Grandville, who regards the elasticity of substitution as "a measure of the efficiency of the productive system" and relates it to the economic miracle in East Asia. They further argue that factor reallocations, as well as the application of new methods of production in one sector, are not only technically determined, but are also strongly influenced by the institutional framework. In this context, the relatively lower elasticity of substitution obtained for sectors like mining, utilities and TCA could be partly explained by institutional arrangements such as labour unions. Unions in these sectors might be very powerful and the costs of hiring and firing of labour might be prohibitively high. However, the reasons why these sectors have low elasticities of substitution require further investigation. Maki and Meredith (1987), for instance, studied Canadian industries and concluded that a high degree of unionisation goes hand in hand with low sectoral elasticities of substitution.

Based on the estimated returns to scale parameter given in column 5, all except CSP exhibit increasing returns to scale. In spite of the difficulty to determine its statistical

[7] The Wald chi-square test is more appropriate than the adjusted r-square since the GLS does not break down in the same way as the OLS does the total sum of squares into the sum of the model sum of squares and the residual sum of squares.

significance, it appears that the RTS parameter for construction and TSC is close to 1. This implies that production in these sectors may take place under constant returns to scale. On the other hand, it appears that the CSP sector produces under decreasing returns to scale. These findings are not surprising, given the nature of most jobs within this sector. This is a sector with a lot of semi-skilled and unskilled workers who lack the human capital that drives productivity and increasing returns to scale. According to the 2006 employment data obtained from Quantec Research, semi-skilled and unskilled workers make up roughly 71% of all the employees in this sector.

5. POLICY IMPLICATIONS

Important information about the sectoral employment-generating potential could be obtained from the returns to scale parameters, the sectoral contribution to total GDP or the share of total employment. The five sectors identified above that are producing under increasing returns to scale contributed 65.3% of the total value added in 2006. The rest is accounted for by the three sectors that exhibited constant or decreasing returns to scale. In terms of total non-agricultural employment, the sectors with increasing return to scale account for 58.6%. The other three sectors with constant or decreasing returns to scale employ 41.4% of the total non-agricultural labour force. The results show that the sectors with increasing returns to scale in total make up the highest percentage of the country's non-agricultural value added and total employment.

If the country's growth originates mainly from these sectors, the number of jobs generated would not be as high as the number that can be generated from sectors with constant or decreasing returns to scale. This means that the number of jobs that will be generated by the 4.5% and 6% envisaged GDP growth by the ASGISA project would not be sufficient to halve the unemployment rate if the growth happens to come only from those sectors with increasing returns to scale. This would be true particularly if the sectors with increasing returns to scale are large and capital intensive compared to those with constant and decreasing returns to scale. By the same token, the employment that would be generated will reduce the unemployment rate by more than half if the same level of growth is recorded only in those sectors with decreasing and constant returns to scale.

However, the above simple policy implication does not take into account skill and quality differences among jobs in the economy. Recognising such differences would particularly be helpful in identifying the sources of technology and economic growth. Regardless of the explicit use of capital and labour as the only factors of production in our model, this study could further be extended in the spirit of the new growth theory as detailed out, for instance, in Fedderke (2005). Increasing return to scale that we identified for some of the sectors may have its origin inside or outside of the firms. It is internal, if it is the result of intentional investment in R&D and innovation as characterised by endogenous growth theory. As was indicated in earlier sections, learning-by-doing that spills over to all labour in the economy can be a good source of increasing returns to scale. But the resulting knowledge is more like a public good and can not be a source of imperfect market structure. On the contrary, the knowledge that is generated as a result of intentional investment in R&D can be protected by property rights and may be a source of imperfect competition. The moment we see knowledge production as intentional private endeavour pursuant to endogenous growth theory (*e.g.* Romer, 1990; Fedderke, 2005), the relevant policy consideration starts to shift from the one that encourages investment in capital goods to those that focus on human capital development. It might be possible that the sectors with increasing returns to scale have a considerably higher level of human capital in production. We further gain a different insight on the differential role of human capital or education in countries at different stages of development. There is a body of literature (*e.g.* Eijffinger and Rossi, 2005) that reminds us of the risk that countries may run into if they pursue with developing human capital without considering their proximity to the technological frontier. If they are far from the technological frontier, they stand to gain more from playing the catch-up strategy or "imitation" as opposed to innovation. It is argued that investment in human capital is destined to benefit those countries that are close to the technological frontier. These studies highlight the need for adopting different policies based on the stage of development that the country finds itself against the technology leader.

We may further ask whether the policy implication that is discussed above in the context of the ASGISA project is valid with a substantial foreign trade component. International trade may promote growth through any of the indirect channels (*e.g.* technology transfer, scale economies associated with bigger markets, increased efficiency from cheaper imports of intermediate goods, etc.) discussed in Oliphant (2007). The literature is unclear about the theoretical channel by which trade directly affects economic growth. Though Oliphant duly recognises the endogeneity problem inflicting trade-growth analysis, the greater part of the literature focuses mainly on the relationship that flows only in one direction, *i.e.* from trade to output growth. Apparently trade bears a greater potential to be the source of scale economies. Krugman (1979), on the other hand, argued and built a theoretical model of trade based on scale economies internal to firms. According to this theory, trade originates from increasing returns to scale not the other way around. It predicts that sectors with increasing returns to scale would have a better chance of raising their exports and output, thereby create the opportunity for hiring more labour. But, we need to remind us of the assumption that is implicitly made in paragraph 2 of this section regarding the level of growth when we state the policy implication. We considered the same level of GDP growth, either 4.5 or 6%, in the SA economy except that in the first scenario the economy is dominated by sectors with scale economies and in the second it is dominated by sectors with constant or decreasing returns to scale. Without any predetermined growth level, an economy with many sectors enjoying increasing returns is expected to grow faster than the same economy dominated by sectors with constant or decreasing returns to scale. For instance, GDP growth of R100 in sectors with decreasing and constant returns to scale might generate more jobs in comparison to the same amount of GDP growth in sectors with scale economies. But the sectors with increasing returns might make up for this disadvantage by growing faster than the sectors with decreasing or constant returns to scale. But evaluating sectoral job generation within a given range of output growth, as we did in this study, eliminates the important aspect that gives sectors with scale economies the ability to compensate for the deficiency. This means the policy implication we identified above will be valid by virtue of the implicit assumption regarding the ultimate GDP growth whether we factor in international trade or not in the analysis.

But there is one important fact to keep in mind about the growth impact of trade. It is easier to attain the level of GDP growth projected in the ASGISA project with sectors enjoying scale economies than with sectors having constant or decreasing returns to scale.

6. CONCLUSION

For decades, neoclassical growth theory has relied on the CD production function. However, the assumption of unitary elasticity of substitution embodied in this functional form inhibits to some extent a detailed investigation into the contribution of the various sources of growth. Elasticity of substitution influences growth via the factor income shares that appear to be a determinant of growth. Restricting the elasticity of substitution to one necessitates the factor income share to remain constant over time. This introduces bias in respect of the values of the parameters that enter into a growth equation. Another parameter of the production function, returns to scale, plays a very important role in the demand for factors of production. Robust estimation of this variable at a sectoral level enables us to evaluate the differential job-creating potential of the various sectors of the economy.

Based on the panel estimation technique employed in this paper, we established that the elasticity of substitution remains below unity for all sectors except for agriculture, forestry and fishing. While mining and utilities exhibit lower elasticity of substitution, construction and TSC have higher elasticity of substitution. FIRB, utilities and mining are at the top of the list enjoying increasing returns to scale. Construction and TSC appear to have constant returns to scale production while CSP is the only sector exhibiting decreasing returns to scale.

Given the information on sectoral returns to scale, elasticity of substitution, employment and output shares, more jobs are expected to be created from growth in sectors with constant or decreasing returns to scale compared to the same growth in sectors with increasing returns to scale.

This analysis has important policy implications. Given SA's high unemployment rate, the focus on promoting growth in certain sectors has to take into account the differential in job-creating potential of the various sectors, as discussed above. It is also suggested that the possible causes for sectors such as mining and utilities to exhibit relatively lower elasticity of substitution should be investigated further.

Any policy intervention that the SA government considers must take into account the sectoral elasticity of factor substitution, returns to scale, capital intensity and sectoral growth. In 2006, the construction, TSC and CSP sectors jointly employed about 41.5% of the country's non-agricultural labour force. Value added growth of R1 million in these three sectors would generate more jobs than the same amount of growth in other sectors of the economy. In spite of the lack of empirical investigation into the reasons underlying the differential sectoral elasticity of substitution, we know that these sectors have a higher elasticity of substitution. This does not, however, suggest that growth in the other economic sectors is not employment generating. Rather, the objective of promoting growth in those sectors enjoying increasing returns to scale must be supported by policies that enhance the possibility of factor substitution.

APPENDIX 1

Summary of Panel Unit Root Tests

Variable	**Null: Unit Root (Homogeneous)** *LLC t-stat*	**Null: Unit Root (Heterogeneous)** *IPS w-stat*	**Null: No Unit Root (Homogenous)** *Hadri*
ln_wage			
Agriculture, forestry & fishing			
Mining	−1.468*	−0.623	4.564
	(0.07)	(0.26)	(0.00)
Manufacturing	0.615	−0.341	7.457***
	(0.73)	(0.36)	(0.00)
Utilities (Water, gas & electricity)	−2.896***	−1.787**	2.073**
	(0.00)	(0.03)	(0.02)
Construction	−0.024	1.04	2.866***
	(0.49)	(0.85)	(0.00)
TCA (Trade, catering & accommodation services)	−0.062	−0.159	1.875***
	(0.47)	(0.43)	(0.03)
TSC (Transport, storage & communication)	−1.449*	0.834	4.229***
	(0.07)	(0.79)	(0.00)
FIRB (Finance, insurance, real estate & business services)	−0.004	0.635	1.822**
	(0.49)	(0.73)	(0.03)
CSP (Community, social & personal services)	−0.801	0.015	2.803***
	(0.21)	(0.50)	(0.00)
ln_lab			
Agriculture, forestry & fishing			
Mining	−1.303*	2.141	2.837***
	(0.09)	(0.98)	(0.00)
Manufacturing	−3.755***	−0.011	13.235***
	(0.00)	(0.49)	(0.00)
Utilities (Water, gas & electricity)	−2.474***	−2.362***	2.094**
	0.00)	(0.00)	(0.01)
Construction	0.057	−0.715	2.69***
	(0.52)	(0.23)	(0.00)
TCA (Trade, catering & accommodation services)	0.053	1.103	2.929***
	(0.52)	(0.86)	(0.00)
TSC (Transport, storage & communication)	−1.106	0.853	1.985**
	(0.13)	(0.80)	(0.02)
FIRB (Finance, insurance, real estate & business services)	−0.826	1.301	2.588***
	(0.20)	(0.90)	(0.00)
CSP (Community, social & personal services)	−2.375***	−2.238**	2.491***
	(0.00)	(0.01)	(0.00)
Ln_vadd			
Agriculture, forestry & fishing			
Mining	−1.268*	−0.827	2.720***
	(0.10)	(0.20)	(0.00)
Manufacturing	−2.802***	−2.738***	9.861***
	(0.00)	(0.00)	(0.00)
Utilities (Water, gas & electricity)	−4.279***	−3.227***	3.211***
	(0.00)	(0.00)	(0.00)
Construction	−0.035	−1.412*	0.417
	(0.48)	(0.07)	(0.33)
TCA (Trade, catering & accommodation services)	0.450	−0.058	0.863
	(0.67)	(0.47)	(0.19)
TSC (Transport, storage & communication)	−0.486	0.219	2.775***
	(0.31)	(0.58)	(0.00)
FIRB (Finance, insurance, real estate & business services)	0.667	0.214	1.220
	(0.74)	(0.58)	(0.111)
CSP (Community, social & personal services)	−1.211	−1.055	2.640***
	(0.112)	(0.145)	(0.004)

Note: *, ** and *** denote rejection of null at 10%, 5% and 1% significance levels, respectively.

REFERENCES

ALTMAN, M. (2007). Employment scenarios to 2024. HSRC Working Paper Series – Document 2506.

ARROW, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29: 155-173.

———, CHENERY, H. B., MINHAS, B. S. and SOLOW, R. M. (1961). Capital-labour substitution and economic efficiency. *The Review of Economics and Statistics*, XLIII(3): pp. 225-250.

AZARIADIS, C. (1996). The economics of poverty traps – Part one: Complete markets. *Journal of Economic Growth*, 1(4): 449-486.
BODKIN, R. E. and KLEIN, L. R. (1967). Nonlinear estimation of aggregate production functions. *Review of Economics and Statistics*, 49: 28-45.
BROWN, M. and DE CANI, J. (1963). Technological change and the distribution of income. *International Economic Review*, 4(3).
CHEN, E. K. Y. (1977). Economies of scale and capital-labour substitution in Hong Kong manufacturing. *Hong Kong Economic Papers*, (11): 42-49.
DUFFY, J. and PAPAGEORGIOU, C. (2000). A cross-country empirical investigation of the aggregate production function specification. *Journal of Economic Growth*, 5: 87-120.
DUPUY, A. and DE GRIP, A. (2006). Elasticity of substitution and productivity, capital and skill intensity differences across firms. *Economics Letters*, 90:340-347.
DU TOIT, C. B. and DE WET, T. J. (2002). A neoclassical production function for the South African economy: Theory and evidence. *South African Journal of Economic and Management Sciences*, SS(4).
EIJFFINGER, S. C. W. and ROSSI, A. (2005). Structural reforms and growth: Product and labour market deregulations. DP 5988, London: Centre for Economic Policy Research.
FEDDERKE, J. (2002). The structure of growth in the South African economy: Factor accumulation and total factor productivity growth 1970-97. *South African Journal of Economics*, 70(4): 611-646.
——— (2005). Technology, human capital and growth: Evidence from a middle income country case study applying dynamic heterogeneous panel analysis. *Royal Economic Society annual Conference, Number 71*.
FELDSTEIN, M. S. (1967). Alternative methods of estimating a CES production function for Britain. *Economica*, New series, 34(136): 384-394.
HADRI, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometric Journal*, 3(2): 148-161.
IM, K. S., PESARAN, M. H. and SHIN, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115: 53-74.
KATZ, J. M. (1968). Verdoorn effects, returns to scale, and the elasticity of factor substitution. *Oxford Economic Papers*, New series, 20(3): 342-352.
KLUMP, R. and PREISSLER, H. (2000). CES production functions and economic growth. *Scandinavian Journal of Economics*, 102(1): 41-56.
———, MCADAM, P. and WILLIAM, A. (2007). The long-term success of the neoclassical growth model. *Oxford Review of Economic Policy*, 23(1): 94-114.
KRUGMAN, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, 9: 469-479.
LEE, J. W. (1970). Determinants of the changes in the relative factor shares. *The Review of Economics and Statistics*, 52(3): 331-336.
LEVIN, A., LIN, C. F. and CHU, C. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108: 1-24.
LIST, J. A. and ZHOU, H. (2007). Internal increasing returns to scale and economic growth. *NBER Working Paper Series*, WP 12999.
MAKI, D. R. and MEREDITH, L. N. (1987). A note on unionization and the elasticity of substitution. *Canadian Journal of Economics*, 20: 792-801.
OLIPHANT, P. (2007). International trade and economic growth: The case of South Africa. Masters Dissertation, School of Economics, University of Cape Town, Cape Town.
ROMER, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98: S71-S101.
ROS, J. (2003). *Development Theory & the Economics of Growth*. Ann Arbor: The University of Michigan Press.
SAPIR, A. (1980) Economic growth and factor substitution: What happened to the Yugoslav Miracle? *The Economic Journal*, 90(358): 294-313.
SOLOW, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1): 65-94.