CSAE WPS/2002-07

Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data[$]

Måns Söderbom[+] and Francis Teal*
Centre for the Study of African Economies
Department of Economics
University of Oxford

May 2002

Abstract

Three dimensions of the performance of firms in Ghana's manufacturing sector are investigated in this paper: their technology and the importance of technical and allocative efficiency. We show that the diversity of factor choices in not due to a non-homothetic technology. Observable skills are not quantitatively important as determinants of productivity. Technical inefficiency is not lower in firms with foreign ownership or older firms and its dispersion across firms is similar to that found in other economies. Large firms face far higher relative labour costs than small firms. If these factor price differentials could be levelled out, substantial gains thorough improvements in allocative efficiency would be possible.

JEL Classification: O14, D24.

Keywords: African manufacturing, productivity, efficiency, human capital, firm size.

[$] The data used in this paper were collected by a team from the Centre for the Study of African Economies, Oxford, the University of Ghana, Legon and the Ghana Statistical Office (GSO), Accra over a period from 1992 to 1998. We are greatly indebted to staff from the GSO for their assistance. The surveys from 1992 to 1994 were part of the Regional Program on Enterprise Development (RPED) organised by the World Bank. The questionnaire was designed by a team from the World Bank. The surveys have been funded by the Department for International Development of the UK government and the CSAE is funded by the Economic and Social Research Council of the UK. We would like to thank two anonymous referees for several constructive comments on an earlier version of the paper.

[+] Måns Söderbom is the UNIDO Fellow at the Centre for the Study of African Economies (CSAE), University of Oxford.

* Corresponding author. Tel.: +44 (0)1865 271077; fax: +44 (0)1865 281447; e-mail: Francis.Teal@economics.oxford.ac.uk.

1. Introduction

Three issues have been prominent in policy discussions of the problems facing firms in developing countries. One has been that technology differences may be important in explaining factor choices across firms of differing size, Pack (1976, 1982). Small firms have come to be identified with more labour intensive technologies so that promoting small-scale enterprises is seen as a means of creating jobs. A second issue has been that firms in developing countries lack the technical capacity to perform well. "Without an increase in proficiency, the responsiveness of output to even the best designed structural adjustment program is likely to be limited. Prices are one-half of a scissor, the other being technical skill", Pack (1993, p. 1). The third has been that larger firms are more capital intensive than smaller ones due to factor prices differing by firm size. A comprehensive discussion of these issues in the context of Indian industrial policy can be found in Little, Mazumdar and Page (1987).

While much of the analysis of efficiency and size has used macro or sectoral data there are now numerous studies based on plant-level data sets, Pitt and Lee (1981), Martin and Page (1983), Corbo and de Melo (1986), Chen and Tang (1987), Tybout (1991), Tybout and Westbrook (1995) Clerides, Lach and Tybout (1998), and Lundvall and Battese (2000). Two of these studies use data for sub-Saharan Africa where it has been argued that the problems of technical inefficiency are likely to be most important. In this paper we propose to investigate these issues of technology choice, the importance of skills, both observed and unobserved, and the relative importance of technical and allocative inefficiency, drawing on a seven year panel of plant-level data from Ghana's manufacturing sector.

A wide range of technology choices is made across firms of differing size in our sample. In this paper we investigate if this is caused by technological differences or by factor price differentials, across firms of differing size. The existence of panel data enables us to extend work in this area for sub-Saharan Africa in several respects. The ability to estimate a production function with fixed effects ensures that skills which are unobserved, but time-invariant, do not bias the parameter estimates. We can then ask if these unobserved aspects of the technology, measured through the fixed effects, do depend on ownership or firm age two factors which have been widely argued to measure access to better managerial skills and the importance of firm learning. Thus in measuring technical inefficiency the existence of panel data means that we do not need to assume, as in the case with cross-section data, a distribution for the residual to identify the component of inefficiency.

Our data contain measures of the human capital in the firm. We can therefore address directly the Pack contention that such skills play an important role in the efficiency with which firms perform. We also propose to assess the potential importance of allocative inefficiency across firms of differing size. The finding of a correlation between factor choices and firm size may, of course, reflect differences in underlying technology related to firm size or sector, not differences in factor choice based on different prices. Distinguishing between these alternative hypotheses is one of our objectives.

The relationship between technology and efficiency measures is discussed in the next section. In Section 3 we set out the production function and in Section 4 provide an overview of the data. The results of estimating both value-added and gross output production functions are presented in Section 5. In Section 6 and 7 the importance of technical and allocative efficiency is investigated. A final section concludes.

2. Technology, Factor Choice and Efficiency

The analysis in this paper will be based on standard production theory, assuming that the relationship between inputs and output can be approximated by a production function that is known to the firm. We write the production function in general notation as

$$Y_{it} = A_{it}F(Z_{it}) \tag{1}$$

where Y_{it} is the level of output, A_{it} is total factor productivity (TFP) and Z_{it} is an n order vector of inputs, where we assume that $F:\Re_{+}^{n} \rightarrow \Re_{+}$ is continuous, strictly increasing and quasi-concave.[1] The latter two assumptions are readily testable and we shall come back to this in the empirical analysis.

[1] These are standard assumptions: continuity ensures that small changes in the vector of inputs lead to small changes in output; F(.) being strictly increasing ensures that employing strictly more of every input yields strictly more output; and quasi-concavity implies that the isoquants are convex towards the origin (see e.g. Jehle and Reny, 1998, p. 220).

Technology and Factor Choice

Factor demand is linked to the firm's technology and factor prices. Empirical studies based on firm-level data, both from developed and developing countries, typically report substantial variation in factor intensities across firms. Assuming there are two inputs, capital (K) and labour (L), we illustrate in Figure 1 two possible mechanisms generating such differentials which have been extensively discussed in the literature (e.g. Little, Mazumdar and Page, 1987). In the top panel of the figure we keep technology constant and let the relative cost of capital decrease with size. In this case large firms choose more capital per employee than small ones because capital is relatively cheaper. In the bottom panel we keep factor prices constant and consider a non-homothetic technology. In this particular case the marginal rate of technical substitution decreases in (absolute) magnitude for a given capital-labour ratio. Again large firms have higher capital-labour ratios, but here this is caused by the non-homothetic technology rather than by heterogeneous factor prices.

Technical Efficiency

Equation (1) represents the 'frontier', or 'best practice', production function in that it defines the maximum output attainable for firm *i* at time *t*, given the technology A(it) and the input set Z(it). Firms that use A(it) and Z(it) inefficiently, however, will not achieve their maximum potential output. The ratio between actual and potential output is conventionally defined as the level of *technical inefficiency*, where firms that use A(it) and Z(it) efficiently will have an inefficiency score of unity, and inefficient firms will have scores in the (0, 1) interval.[2] There has been a continuing development of methods over the past 50 years to compute inefficiency scores, with the two principal methods being stochastic frontiers, which is based on econometric methods, and data envelopment analysis (DEA), relying on mathematical programming. While DEA is attractive in that it does not require any parametric assumptions or assumptions about the functional relationship between inputs and output, a significant disadvantage of this procedure is that the computed inefficiency scores are very sensitive to measurement errors, either in output or the input variables. Therefore, in our view, DEA is not very

[2] Modern efficiency measurement begins with Farrell (1957).

well suited to survey data sets and will not be used in this paper.[3] Stochastic frontiers accommodate statistical noise in the dependent variable by means of introducing a residual, while typically treating inefficiency as a random parameter. A general class of such models, which specialises to several in the literature, is presented in Battese and Coelli (1992), another general form is that proposed by Battese and Coelli (1993). One unattractive feature of these random effects models is that the inefficiency term typically is assumed to be uncorrelated with the explanatory variables in the frontier production function. If the inefficiency terms are in fact correlated with firm attributes, the estimated parameters and the inefficiency scores from such models will be biased (Tybout, 1992).

Given that both the inefficiency term and the residual are unobservable, there are substantive identification issues that need to be addressed. With cross-section data it is not possible to separate the residual from inefficiency without making parametric assumptions about the distribution of the residual and the inefficiency term, which is unattractive.[4] If panel data are available, and if it is reasonable to assume that inefficiency is approximately constant over the time-span during which the firm is observed, then we can model inefficiency as a time invariant firm specific effect. This is the route we will take in this paper. Contrary to most papers in the area we shall make no assumptions about the distribution of inefficiency, and we shall also allow the inefficiency term to be freely correlated with the arguments of the production function. Defining the inefficiency term as $U_i = \exp(-m_i)$ and the residual as ε_{it} we rewrite the production function as

$$Y_{it} = A_{it} F(Z_{it}) \cdot U_i \cdot e^{e_{it}} . \quad (2)$$

In the empirical analysis we will allow for correlation between ε_{it} and the arguments of *F*, caused by measurement errors in the explanatory variables.

---

[3] Lundvall (1999) uses both DEA and stochastic frontiers on Kenyan survey data, and reports an average inefficiency score of 0.38 based on DEA and 0.77 based on a stochastic frontier. The substantially lower score yielded by DEA is consistent with presence of measurement errors in the dependent variable.

[4] Even if one is prepared to make such distributional assumptions, measuring technical inefficiency with cross-section data is difficult. As all inputs determining output which are omitted from the production function will give rise to the appearance of technical inefficiency clearly measured technical inefficiency may simply reflect how imperfect are the measures of inputs rather than how poorly managers transform inputs into outputs, Tybout (1992).

Allocative Efficiency

The second dimension of efficiency with which we are concerned is that of allocative efficiency. In production theory, allocative efficiency conventionally reflects the ability of the firm to use optimal factor combinations, given their respective prices. That is, in the context of the graphs in Figure 1, the firm would be allocatively inefficient if it were to choose a point on the isoquant at which the isocost line is not tangential. A related but distinct form of allocative inefficiency occurs when as a result of price differentials firms of differing size select different factor combinations.[5] Such differentials may be due, for example, to non-competitive factor markets or differential taxation on firms of differing size. There is a view common in the development literature that larger firms are more capital intensive than smaller ones and that such factor choices are inappropriate for the factor endowment of poor countries. We propose to investigate the potential importance of such allocative inefficiency by examining how large are the differences in technology choice that can be attributed to differences in factor prices. Our data enable us to provide an estimate of the cost reduction possible if factor price differentials could be removed.

3 The Production Function

In order to translate (2) into an expression suitable for econometric analysis, we need to adopt an explicit functional form of the production function *F* which provides a reasonably close approximation of the real technology. One flexible form which has been used extensively in studies estimating cost and production functions is the second-order transcendental logarithmic (‘translog’) production function (Christensen et al, 1971; Berndt and Christensen, 1973), which we write as

$$\ln Y_{it} = \sum_j b_j \ln X_{jit} + 1/2 \sum_k \sum_m b_{km} \ln X_{kit} \ln X_{mit} \,, \qquad b_{rs} = b_{sr} \text{ for all } s, r, \qquad (3)$$

where X_j is the jth input in the production process, j=1,2,..., J, and b denotes parameters to be estimated. The translog specification is attractive because it nests or approximates a number of popular models in the literature, and for our purposes it is especially useful because output and substitution

[5] In particular, notice that allocative inefficiency defined like this is not the result of optimisation errors made by the firm.

elasticities are allowed to vary with the levels of the inputs, hence homotheticity is not imposed.[6] The following shows three key parameter restrictions on (3) which we shall test for in the empirical analysis:

$\sum_k b_{km} = 0 \quad (m = 1,2,\ldots,J)$ (homotheticity); (4a)

$\sum_k b_{km} = 0 \quad (m = 1,2,\ldots,J)$, (4b)

$\sum_k b_k = 1$ (constant returns to scale);

$b_{km} = 0 \quad (k = 1,2,\ldots,J;\ m = 1,2,\ldots,J)$ (Cobb-Douglas form). (4c)

In our empirical analysis we will use two definitions of the dependent variable in the production function, namely gross output and value-added. Value-added production functions appear to be more common in the literature, however research by Basu and Fernald (1995) shows that adopting a value-added production function can yield misleading results if there is imperfect competition or increasing returns to scale. We will therefore present results for both value-added and gross output production functions in the results below. In the output production function we use four inputs: labour, denoted L, physical capital, K, raw material inputs, M, and indirect inputs, I. Because value-added is defined as output minus costs for raw material and indirect inputs, we use only labour and capital as inputs in the value-added specification.

The productive quality of labour is unlikely to be constant across individuals and firms. To allow explicitly for the role of human capital in the production process, we follow Hall and Jones (1999) and Bils and Klenow (2000) and augment the logarithmic production function (3) with ah_{it}, where a is a parameter vector to be estimated and h_{it} is a vector of human capital variables. We let this vector consist of firm-level averages of employees' years of education, tenure, age and age squared. This specification is similar to that of (stylised) Mincerian earnings functions, which are commonly adopted in the literature, and which typically yield positive coefficients on tenure and education and an

[6] The cost of such flexibility is that it is necessary to test for monotonicity and quasi-concavity at each data point. It is a common result in the empirical literature to find that not all observations comply with these conditions. Typically in such cases the production function is nevertheless considered well-behaved '…if the conditions are met for a sufficient number of the observed levels of inputs and outputs…' (Little, Mazumdar and Page, 1987, p. 162).

inverse u-shaped age-earnings profile. Hence, to the extent that earnings reflect productivity, a notion supported by Jones (2001) for Ghana, we expect similar findings based on the production function. Of course, in addition to observed heterogeneity in labour quality, there is potentially an unobserved dimension. If unobserved labour quality is firm specific and time invariant this will not present us with a problem provided that we control for firm fixed effects. However, if there is time varying unobserved quality which is correlated with inputs, then we need to control for the resulting simultaneity to prevent bias in the estimated production function parameters.

Taking the above into account, we write our empirical specification as follows:

$$\ln Y_{it} = \sum_j b_j \ln X_{jit} + 1/2 \sum_k \sum_m b_{km} \ln X_{kit} \ln X_{mit} + a h_{it} + m_i + d_t + e_{it}, \tag{5}$$

where we have added a time effect d_t to capture common shocks to the firms over time. In estimating (5) we need to deal with the fact that the explanatory variables are likely to be correlated both with the equation error and with the firm specific effect. For instance, the regressors will be correlated with the equation error if managers alter their inputs in response to demand shocks. Failure to allow for such simultaneity is expected to yield an upward bias in the estimated coefficients. It is also likely that explanatory variables are measured with error, as discussed above, which, if ignored, is expected to cause a downward bias in the estimated coefficients. Finally, it is possible that unobserved labour quality varies over time and is correlated with inputs. Sparks (1986), for instance, formulates an efficiency wage model where in equilibrium worker effort depends on the outside option available to the worker should he get fired. Hence to the extent that the unobserved outside option varies over time and across individuals (perhaps being determined by the local unemployment rate, labour demand in the agricultural sector, etc.) there would be time varying effort in this model. Self-selection of high-quality workers into high-performance firms would be another potential mechanism yielding variable labour quality over time (see e.g. Oi and Idson, 1999).

To address the problems posed by the regressors potentially being correlated with the residual we will use an instrumental variables approach, where we exploit the panel dimension of the data and use lagged values of the explanatory variables as instruments.[7] As this approach renders the within transformation impractical (see e.g. Griliches and Hausman, 1986), we will take first differences to wipe out the firm effects. In the case of highly persistent data, lagged variables in levels are likely to be weak instruments for contemporaneous differences, potentially giving rise to finite sample bias and poor precision of the estimates (Blundell and Bond, 1998). Most of our measured inputs exhibit fairly strong persistence, and we will therefore follow Blundell and Bond and combine the differenced equation with a levels equation to form a system generalised method of moments (GMM) estimator. Within this framework we shall use lagged levels as instruments for contemporaneous differences and lagged differences as instruments for contemporaneous levels.[8] Naturally, the legitimacy of this procedure hinges on the instruments being valid, which will be tested for. We provide a brief discussion of the system GMM estimator in Appendix 1.

4 Data

This study uses panel data on manufacturing firms in Ghana, collected in face-to-face interviews with the firms' management during five survey rounds in the 1990s. These data are annual and cover the 1991 to 1997 period. At the same time as the firms were surveyed a sample of workers and apprentices was chosen from each firm designed to cover the full range of personnel employed by the firms. The objective was to have up to ten workers and ten apprentices from each firm where firm size allowed. As a result of this survey design it is possible to use the responses from workers in the firm to create firm-level averages of worker characteristics. During the course of the surveys a sub-set of 143 firms have provided data on the components of value-added and sufficient information that the capital stock, employment and the human capital stock of the firm could be calculated for at least three consecutive

---

[7] Clearly, this is one important benefit of panel data. In the case where the researcher has cross-section data only, purging explanatory variables from simultaneity typically requires extraneous information of the kind often unavailable in practice.

[8] In highly persistent time series, lagged levels will be poor instruments for contemporaneous differences but lagged differences may still be good instruments for contemporaneous levels. For instance if X follows a random walk, $X_t = X_{t-1} + e_t$, implying $\Delta X_t = e_t$, then X_{t-1} will be uncorrelated with ΔX_t, but ΔX_{t-1} will nevertheless be correlated with X_t. Blundell and Bond (1998) present results from a Monte Carlo experiment indicating that the system GMM estimator performs substantially better than the standard differenced GMM

years.[9] In the regression we lag the physical capital stock by one year so the maximum period over which we can observe the firms is six years. The resulting unbalanced panel contains 676 observations. The three major additions to the primary data are the derivation of physical capital stocks from investment flows, the calculation of firm-level human capital stocks based on worker information and the construction of firm specific price indices for outputs and material inputs. These prices, which differ for output and inputs, are used to deflate output and inputs into constant price (1991) domestic currency prices. All references in the text and tables to output, input and physical capital refer to these deflated values. The average size of firm, measured by employment across the seven rounds of the data, is 67 employees and the standard deviation is 113, so the range of enterprises covered by the survey is large. Firms range in size from 2 to 841 employees.

Table 1 shows descriptive statistics for the key variables in our empirical analysis. Four size categories are identified: the micro which is firms with less than 6 employees, small those with from 6 to 30, medium those with from 31 to 99, and large those with 100, or more, employees. The upper panel of the table (under the heading ‘Conditional Means’) shows mean values for the monetary and human capital variables, purged of sectoral and time effects as explained in the notes to the table.[10] The fact, shown in Table 1, that the capital labour ratio differs substantially across firms of differing size when the data are purged of sectoral effects is important for establishing that it is not differences in technology related to sector which explains the dispersion. For all the variables shown in the upper part of the table the mean values increase monotonically over the size range, indicating that large firms have higher labour productivity and capital intensity, more skilled workers and higher labour costs. For the log of output, value-added and capital, per employee, the increases are substantial,

---

estimator when the data are highly persistent. Recent papers adopting the system GMM approach are Blundell and Bond (2000), Blundell, Bond and Windmeijer (2000) and Windmeijer (2000).

[9] To obtain a measure of the firm’s human capital it was necessary to merge the worker with the firm level information. In aggregating from the worker to the firm level we used weights to ensure that we can move from individual data to firm based averages. To do this we weighted the human capital variables by the proportion of workers in a given occupational class within the firm. Eight common occupational groups across the rounds of the survey were identified. These occupational categories for the worker level data are matched with the occupational categories given in the firm level data. A data appendix explaining the details of this procedure is available on request from the authors.

[10] Because the panel is unbalanced, the sample composition is not constant over time. To ensure that the summary statistics are not driven by changes in the sample composition during the course of the surveys, we purge the variables of time effects.

corresponding to about 150 per cent, 230 per cent and 900 per cent, respectively, in levels. These descriptive statistics confirm that a wide range of input combinations are adopted by firms in our sample. In the following sections we investigate the reasons for these choices.

5 Estimates of Value-Added and Gross Output Production Functions

In addressing the issue of technology and productivity, two issues are central to our investigation: whether technology is non-homothetic and how human capital impacts on productivity performance. In Table 2 we report various production function results, based on OLS and the within estimator. We begin by looking at the left part of the table, where we model value-added.

The OLS results for the translog model, Column [1], clearly suggest that the data comply very well with the regularity conditions, as both quasi-concavity and monotonicity is fulfilled at each observation.[11] There is no strong evidence of non-homotheticity or variable returns to scale, and the Cobb-Douglas specification can easily be accepted given the translog functional form. Column [3] reports OLS results for the Cobb-Douglas specification, and the results are very similar to those obtained from the translog model. The estimated coefficient on employment is 0.89, and that on physical capital is 0.18, and both are significant at the one per cent level. The human capital coefficients all have the anticipated signs, however only the age effect, which is a quadratic, is significant.

These OLS results will be biased if, which seems likely, the explanatory variables are correlated with the unobserved firm specific effects. When we introduce controls for firm fixed effects, however, the production function disintegrates, Columns [2] and [4]. Implausibly, physical capital now appears to have a negative, although insignificant, effect on value-added, and monotonicity is hence violated at each observation. The standard explanation why controlling for fixed effects in the production function tends to give unsatisfactory results is that the differencing procedure exacerbates the bias caused by measurement errors in the explanatory variables, Griliches and

[11] Monotonicity requires that each input has a positive marginal product, and quasi-concavity requires that the bordered Hessian matrix of first and second partial derivatives of the production function are negative semi-definite. In the translog specification the marginal products and the partial derivatives depend both on the

Mairesse (1997).[12] To address this issue we will utilise the panel nature of the data to search for valid instruments.

We now turn to the right-hand part of the table, where we model gross output. Column [5] reports OLS results for the translog model. Again we cannot reject homotheticity and there is no evidence for variable returns to scale, however it now seems we should decisively reject the simpler Cobb-Douglas model in favour of the translog specification. Only 44 per cent of the observations fulfil quasi-concavity and 69 per cent comply with monotonicity, which is not satisfactory. Introducing controls for fixed effects, Column [6], only makes matters worse in this respect, presumably for the reasons discussed above. Estimating the Cobb-Douglas model using OLS, Column [7], gives positive and significant coefficients on all inputs. We can compare the point estimates to those obtained in the value-added model [3] if we assume that the cost of raw materials and indirect inputs is a constant fraction of output. Under this assumption, the coefficient on employment of 0.14 in the output model translates into 0.82 in a value-added equation, while the capital coefficient of 0.03 corresponds to 0.18 in a value-added equation. Hence there is no evidence that the two models give radically different results. As for the human capital variables, we now obtain a significant but small coefficient on education, and a significant inverse u-shaped age effect.

None of the models reported in Table 2 will yield consistent results in the presence of endogeneity or measurement errors, so we need to probe the data further before we can draw conclusions about the properties of the production function. In Table 3 we report system GMM results controlling for fixed effects and using lags of explanatory variables in levels as instruments for contemporaneous differences, and lags of the explanatory variables expressed in first differences as instruments for contemporaneous levels. All results in this table are two-step GMM estimates, where the *t*-statistics are based on robust, finite sample corrected standard errors (see Windmeijer, 2000).[13]

---

values of the inputs and on the estimated parameters, and we therefore need to investigate if monotonicity and quasi-concavity hold at each data point.

[12] If the sample is 'large' and the regression has only one regressor, measurement errors alone cannot be the reason why differencing changes the sign of the estimated coefficient (see Deaton, 1997, p. 109). This simple rule does not hold, however, if there are several explanatory variables or the sample size is finite.

[13] It is well known that the asymptotic standard errors in two-step GMM estimators can be severely downward biased in finite samples (e.g. Arellano and Bond, 1991). As a consequence, researchers often draw

Column [1] summarises the translog results for the value-added specification, and because we easily can accept the Cobb-Douglas model within the translog framework, we focus on the Cobb-Douglas results reported in Column [2]. The results appear to be reasonable. The estimated employment coefficient is 0.73, that on capital is 0.31, and both coefficients are significant at the one per cent level. There is no evidence of variable returns to scale, or, by implication of the Cobb-Douglas specification, non-homotheticity. Among the human capital variables only age has a significant impact on productivity. The point estimates on the age terms indicate an inverse *u*-shaped effect of workers' age on productivity, which peaks at the age of approximately 32.[14]

Column [3] in Table 3 reports system GMM estimates of the output translog production function. Like for all other models previously reported, there is no evidence for non-homotheticity or variable returns to scale. Further, in contrast to the OLS and within models reported in Table 2 for the output specification, we can now comfortably accept the Cobb-Douglas specification, reported in Column [4], as a result of using instrumental variable techniques. We consequently focus on the results in Column [4]. The estimated coefficient on employment is equal to 0.17, and significant at the one per cent level, which can be compared with 0.14 in the OLS model and 0.10 in the within specification. More dramatically, the estimated capital coefficient is 0.09, hence substantially higher than in the OLS (0.03) and within (-0.06) specifications, and significant at the five per cent level. The input elasticities sum to 1.00, indicating constant returns to scale. The coefficient on tenure is now significant at the ten per cent level, although its magnitude is quite small. As in all other models, we

---

inference based on one-step GMM estimators, which are less efficient than the two-step estimators. However, Windmeijer (2000) shows how the asymptotic two-step standard errors can be corrected when the sample size is finite. Monte Carlo evidence reported by Bond and Windmeijer (2001) indicates that this procedure yields a much more reliable basis for inference than relying on the asymptotic standard errors.

[14] Jones (2001) uses data on Ghanaian manufacturing firms observed during 1992-93 and reports regressions indicating that years of experience has an inverse u-shaped effect on productivity, peaking at about 28 years of experience (see Table 2 in her paper). Translating this result into one directly comparable to our age effect is not straightforward as experience is correlated with age but also depends on years of education and time spent in unemployment. Further, our regressions control for years of tenure, which is correlated with age, while those reported by Jones do not. Nevertheless, to assess if our results are broadly in line with those reported by Jones, consider a scenario where a set of individuals begin school at the age of six, spend 11 years in school (this is the sample mean reported in Table 1) and then join a firm immediately (i.e. at the age of 17). Assume that there is no labour turnover after these individuals have joined the firm, so that average tenure increases by one year each year. Based on the point estimates reported in Table 3, Column [2] this firm would have its productivity peak when workers reach 21 years of experience.

obtain a quadratic age effect, indicating that productivity increases with the average age of the workforce up until the age of 32 years, and then decreases. The education coefficient is insignificant.

In summary the Cobb-Douglas models in Table 3 appear to be well specified. The Sargan-Hansen tests indicate that the instruments are valid. The fact that the capital coefficient is higher than in the OLS and within specifications is consistent with this variable being measured with error, but also with the argument advanced by Olley and Pakes (1996) that the capital coefficient will be downward biased if unobserved productivity impacts on the investment decision and the exit decision simultaneously. For all estimators we can easily accept the hypothesis that technology is homothetic, thus suggesting that technology is *not* the reason why we observe differing factor intensities over the size range. There is no evidence for variable returns to scale, except in some of the within regressions which we believe to be biased downwards for reasons already discussed.[15]

6 Technical Inefficiency

In reviewing studies which measure the dispersion of productivity in developing countries, Tybout (2000, p. 25) argues that "they are not very informative. Most of them are based on outdated methodologies. With a few exceptions they rely on cross-sectional data, and hence must infer efficiency dispersion from the skewness of the production function residuals. Further, because they measure output as real revenue, they misattribute cross-plant mark-up differences to productivity dispersion. Finally, for lack of data, they typically attribute high productivity with superior performance, ignoring many of the costs that firms incur to enhance their technical efficiency". Our analysis has addressed these concerns. We have used recently developed econometric techniques that enable us to control for simultaneity (e.g. in the form of measurement errors in inputs or feedback

---

[15] Cost minimisation with a Cobb-Douglas technology implies constant factor shares across firms. In investigating if this is consistent with the data, we confine attention to the factor share for labour since we do not have direct data on capital costs. Using the first order condition for capital, the labour factor share reduces to the first-order condition for labour, $\alpha_L(Y/L) = w$. To test whether this is constant across firms we compute a 'residual', $\text{epsilon}_{it} = \{ \log \alpha_L + \log(Y/L)_{it} - \log w_{it} \}$, and then regress this on a constant and firm size. The null hypothesis is that the constant and the coefficient on size are equal to zero. We do this based both on the value-added function and the output function reported in Table 3. In neither case is there any evidence that epsilon varies systematically with firm size. Further, the constant terms are not significantly different from zero. These results are thus consistent with the technology being Cobb-Douglas. The regression results are available on request.

from output to factor choices) and unobserved heterogeneity (in the form of firm fixed effects, freely correlated with the inputs). We have used panel rather than cross-section data and we have used firm-level information on input and output prices to construct firm specific deflators. Finally, we have used detailed data on human capital to directly measure skills. What are the implications of these procedures for measures of technical efficiency?

We estimate the fixed effects based on the output production function reported in Column [4], Table 3, and use these as our measures of technical inefficiency.[16] To be able to compare the dispersion of technical inefficiency with previous studies, we normalise so that the efficiency index, te_i, is bounded in the (0 ,1] interval: $te_i = e^{-(\hat{m}_{\max} - \hat{m}_i)}$, where $\hat{m}_{\max}$ is the sample maximum of the estimated fixed effects, and $\hat{m}_i$ is the estimated fixed effect for firm *i*. Hence an efficiency score equal to one indicates efficiency or 'frontier' technology, and scores less than one indicate inefficiency. Figure 2 shows the distribution of this efficiency index. The sample mean of this fairly skewed distribution is 0.53. However, as can be seen from the graph, the fairly low average inefficiency score is partly caused by a few outliers, and eliminating the five largest fixed effects from the sample would increase the average inefficiency score to 0.60.

Tybout (2000, Tables 2 and 3) provides an extensive summary of inefficiency studies for developing countries with a comparison, from Caves et al. (1992), of efficiency measures from Australia, Japan, Korea, UK and US. The latter reports country averages of the efficiency index ranging between 0.67 and 0.70. As Tybout notes there is little evidence that the dispersion of efficiency in developing countries is very different from that in developed. The average across all the studies cited for developing countries is 0.6, identical to the finding above and only slightly below that for OECD countries.[17]

In Table 4 we regress the fixed effects estimates on time invariant variables, to see if we can detect any systematic variation in efficiency across sectors, ownership structures, locations and other

---

[16] We obtain estimates of the fixed effects by regressing the levels residuals on firm dummies.

[17] While the dispersion of technical inefficiency thus appears to be similar to that in other countries, the firms in our sample may still be very different in terms of 'absolute' efficiency, or total factor productivity, to firms in other regions. To investigate this issue, it is clearly necessary to use cross-country data.

firm characteristics. Few of the regressors have strong explanatory power. With the metal sector as the omitted category, we find inefficiency to be about 24 per cent lower in the wood sector, eleven per cent lower in the food sector and 13 per cent higher in the textiles sector. These effects are individually significant at the ten per cent level or lower, and all the sector dummies are jointly significant at the one per cent level. Neither firm age nor ownership are associated with significant efficiency differentials. The R-squared indicates that our regressors manage to explain only 16 per cent of the variation in the fixed effects.

7 Allocative Inefficiency

In the last section we showed that the dispersion of productivity in Ghanaian manufacturing differed little from those observed in other economies, so there is no empirical support for the notion that there are atypically large inefficiencies in the sector. There is also no evidence that the technology is non-homothetic, so it is not differences in technology which explain differences in factor choice across firms of differing size. We know from Table 1 that the capital labour ratio does increase with size. In this section we explore the potential size of allocative inefficiency if these differences do reflect differences in factor prices.

Factor Intensity, Relative Factor Prices and Firm Size

In measuring capital intensity it is potentially important to recognise the role of worker quality. In our sample large firms tend to employ individuals with more observed human capital (Table 1), and it is also possible that unobserved labour quality is higher in large than in small firms. If the relevant labour input in the production process is total labour quality, rather than total number of employees, then measures of factor intensity not adjusted for quality, such as the capital labour ratio, may be quite misleading. In particular, the denominator (ie labour) would probably be underestimated for large firms, in which case the relation between size and capital intensity would be upward biased. To adjust the labour stock for quality differences, consider the Cobb-Douglas value-added production function

$$\ln Y_{it} = b_1 \ln L_{it} + b_2 \ln K_{it} + a h_{it} + m_i + d_t + e_{it} .$$

The human capital vector h reflects observed labour quality. Assuming that the fixed effect m_i only reflects unobserved worker quality, which is likely to exaggerate the role of unobserved labour quality and is as such a conservative assumption, we can rewrite the production function as

$$\ln Y_{it} = b_1 \ln LQ_{it} + b_2 \ln K_{it} + d_t + e_{it},$$

where LQ is quality-adjusted labour, defined as $LQ_{it} = \exp[(1/b_1)(ah_{it} + m_i)] \cdot L_{it}$. Using the parameter estimates in Table 3, Column [2], we calculate the corresponding quality-adjusted capital labour ratio as[18]

$$\ln k_{it} \equiv \ln(K/LQ)_{it} = \ln K_{it} - \left((1/\hat{b}_1)(\hat{a}h_{it} + \hat{m}_i) + \ln L_{it}\right). \quad (6)$$

To investigate the relation between capital intensity and firm size, we consider a partial linear model of the form

$$\ln k_{it} = y_s + l_t + f(\ln L_{it}) + n_{it}, \quad (7)$$

where y_s is a sector-specific effect equal to one if firm i belongs to industry s and zero otherwise, l_t is a time effect and $f(.)$ is an unknown but smooth function.[19] Our task is to estimate the function $f(.)$, and to this end we use the semiparametric approach outlined in Yatchew (1998). This is flexible as it puts few restrictions on the shape of $f(.)$. The idea of this estimator is that, because $f(\ln L) \approx f(\ln L + o)$ provided that o is 'small', differencing the data sorted by $\ln L$ will effectively remove $f(.)$ provided that the adjacent $\ln L$ values are sufficiently close. Once the data have been sorted and differenced in this way, we can estimate the time and sector effects using OLS. Equipped with these estimates, we compute $\ln k^*_{it} = \ln k_{it} - (\hat{y}_k + \hat{l}_t)$ using the pre-sorted, pre-differenced form of the data, and estimate the regression $\ln k^*_{it} = f(\ln L_{it}) + n_{it}$ using a standard nonparametric kernel smoother technique.

[18] We are very grateful to a referee for this suggestion.

[19] For technical reasons we assume that the first derivative of $f(.)$ is bounded by some constant (Yatchew, 1998).

Figure 3 shows the estimated function $\hat{f}(\ln L)$ and pointwise 95 per cent confidence bands, obtained through bootstrapping.[20] We also show the regression line obtained when the labour variable is not quality adjusted, represented here by the dashed line. While the regression shows that the capital intensity adjusted for quality increases with size, the pattern is non-linear. The positive correlation between size and capital intensity is strongest for firms between ten and 90 employees, outside this range the regression function is relatively flat and the confidence bands wide. Within the (10, 90) range, the average slope of the regression line is about 0.8, indicating that a one per cent increase in the labour force is associated with a 0.8 per cent increase in the capital labour ratio. The dashed regression line diverges from the solid line after ln L = 4.5, indicating that while adjusting for labour quality matters, it does so only for firms with more than 90 employees.

Based on these regressions we use the predicted values of capital intensity, and calculate the implied capital stock values for given levels of labour. This yields the expansion path of factor combinations. We plot this in Figure 4 along with isoquants based on a Cobb-Douglas value-added production function with a capital coefficient of 0.3 and a labour coefficient of 0.7, thus very similar to the regression reported in Column [2], Table 3. Clearly the expansion path bends off sharply towards the capital axis, and because technology is homothetic the slope of the isoquants will be steeper for higher capital labour ratios. The figure implies that firms with 200 employees have up to 5 times higher relative labour to capital costs than firms with 21 employees. These differences in relative factor prices imply a substantial dispersion of factor choices.

Labour and Capital Cost Differentials

The results reported in the previous sub-section show how relative factor prices differ over the size range. In this sub-section we investigate how much of this may be due to labour cost differentials on the one hand and capital cost differentials on the other. While our data set does not contain direct information on the cost of capital, detailed data on labour costs are available. We know from Table 1 that the average labour cost increases with firm size, but it is of course very likely that this reflects, at

[20] In the first stage underlying these calculations we use a weighted tenth order differencing procedure,

least to some extent, the higher levels of human capital in large firms.[21] To investigate if large firms have higher labour costs than small firms conditional on worker quality, we hypothesise that the average labour cost, denoted *w*, depends on human capital measured as above, firm size, time effects t_t, and unobserved heterogeneity in the form of firm fixed effects, w_i:

$$\ln w_{it} = gh_{it} + q_1 \ln L_{it} + q_2 (\ln L_{it})^2 + t_t + w_i + u_{it}, \qquad (8)$$

where g, q_1, q_2 are parameters to be estimated and u_{it} is an error term.[22] We include a squared term of employment as a simple means to allow for non-linear size effects. The specification is similar to the Mincerian earnings function, however our results should not be interpreted as Mincerian earnings function estimates since (8) is being estimated based on firm level data, not individual level data.[23]

In Table 5 a series of tests are conducted using OLS, within and system GMM estimators. The OLS results, shown in Column [1], suggest a non-linear, convex size effect on earnings, which is significant at the five per cent level. All the human capital coefficients are significant at the five per cent level or lower, and have the same signs as in the OLS production functions reported in Table 2. In Column [2] we control for firm fixed effects, which reduces the precision with which the coefficients are estimated, in some cases drastically. Most importantly, the size coefficients are insignificant.

It is quite probable that the OLS and within estimates of the size effect are downward biased, partly due to measurement errors in the explanatory variables, and partly because any effect from size onto earnings will induce the firm to economise on labour. To address this issue we report in Column [3] system GMM estimates based on a specification in which firm size and the human capital variables are treated as endogenous. While the size coefficients have very low *t*-values, they are nevertheless jointly significant at the ten per cent level. There is very little curvature in the implied size-wage

---

where the weights are optimal from an efficiency point of view. See Yatchew (1998) for technical details.

[21] A number of recent studies based on data on African manufacturing firms have shown that wages are positively correlated with firm size, conditional on standard human capital variables (e.g. Valenchik, 1997; Strobl and Thornton, 2001; Manda, 2002; Mazumdar with Mazaheri, 2002).

[22] The literature discusses numerous reasons why wages are positively correlated with firm size. One frequently made argument is that firm size is correlated with omitted worker quality (e.g. Hamermesh, 1993; Dunne and Schmitz, 1995; Kremer, 1993; Kremer and Maskin, 1996). The specification (8) controls for observed human capital and unobserved worker quality in the form of firm fixed effects.

[23] Given that we want to shed light on the relation between firm size and average labour costs, the firm level is the appropriate level of aggregation.

profile, however, and when we drop the squared term from the regression on the grounds that it appears irrelevant, the resulting estimate of the remaining size coefficient is equal to 0.15 and significant at the five per cent level, Column [4]. The coefficients on tenure and the age terms are significant and similar in magnitude to the OLS coefficients. The coefficient on education, however, is far from significant, suggesting that education is correlated with firm specific time invariant factors.

The regression in Column [4] includes controls for time invariant worker quality and observed human capital. Provided that this is a sufficient set of controls, the positive coefficient on size is evidence that firms of differing size face different costs for a given amount of productive labour input.[24] The point estimate of 0.15 implies that labour costs will rise by 40 per cent as firms expand from small to large ie from 21 to 200 employees. This implies that most of the factor intensity differential documented in the previous sub-section is due to differing capital costs.[25]

Finally, we are in a position to calculate the cost of capital. Maintaining the assumption that the value-added production function is two-factor Cobb-Douglas with a capital coefficient of 0.3 and a labour coefficient of 0.7, we use the formula $r = (0.3/0.7)\cdot(w/(K/L))$ obtained from the first-order conditions for conditional cost minimisation.[26] Focussing on how the cost of capital varies with size, we use the estimated labour cost parameters reported in Table 5, Column [4], and the nonparametric capital-labour regression shown in Figure 3, and write r as a function of size such that

$$\hat{r}(L) = (0.3/0.7)\cdot \exp(0.15\log L)\cdot \Gamma / \exp(\hat{f}(\ln L)), \tag{9}$$

[24] We cannot control directly for unobserved time varying worker quality in our regressions. Could such an omission be driving the size effect? While possible in principle, this seems to us rather unlikely. It would be necessary for the changes in unobserved quality to be strongly correlated with changes in firm size, in which case we nevertheless would expect our instruments to correct for the resulting simultaneity. If the instruments failed to do so, we would expect the Sargan-Hansen test to indicate that the model is misspecified. There is no evidence that this is the case. In addition, the observed variables combined with the firm fixed effects already explain more than 86 per cent of the variation in labour costs, hence there is not much variation left to explain by any potentially omitted variable.

[25] Of course, to the extent that the size coefficient nevertheless is overstated, the capital cost differentials are even larger than what is indicated by these calculations.

[26] In these calculations factor prices are assumed fixed from the point of view of the firm.

where Γ is a constant.[27] Figure 5 shows how $\hat{r}(L)$, interpretable as the annual rental price of one unit of capital, varies with firm size. After an initial increase for firms with less than seven employees, the capital price falls sharply with size. Micro firms (firms with at most five employees) face average capital prices between 0.34 and 0.52, while large firms (with more than 99 employees) face prices between 0.09 and 0.14. The figure also shows the marginal product of capital inferred from the output and value-added Cobb-Douglas production functions reported in Table 3, for the four size categories identified in Table 1. The value-added and output based estimates of the marginal product of capital are very similar and well in line with the predicted cost of capital.

One clear implication of our findings is that substantial cost reductions and a more efficient use of capital are possible. If the large firm output were produced using small firm factor proportions the amount of labour would be 64 per cent higher while the capital required would be about a third that actually used, see Figure 4. If large firms faced the labour costs of small firms our findings imply that their unit costs would fall by 20-25 per cent.[28] The extent of the welfare gain from this cost fall would depend on the shape of the demand function.

8 Summary and conclusions

The view that firms in Africa are inefficient has been widespread. In this paper we have used panel data to investigate both technical and allocative inefficiency. We have shown that a very simple functional form, the Cobb-Douglas, adequately represents the production technology. The implication of the acceptance of the Cobb-Douglas functional form is that the technology is homothetic so that differences if factor choices over firms of differing size is not a reflection of differences in technology.

By including controls for workers' education, age and tenure, we have acknowledged that differences in observed human capital may map into productivity differences. We have also

---

[27] Calculated as $\exp(\overline{\log w} - 0.15 \cdot \overline{\log L})$, where $\overline{\log w}$ and $\overline{\log L}$ are sample averages.

[28] These figures can be simply derived from the Cobb-Douglas production function set out in the notes to Figure 4. At point A we have assumed wages of Cedis 357,000 and a capital cost of 8 per cent. At point B the wages are Cedis 245,000 and the capital cost is 29 percent. The unit cost reduction is based on giving the firm at A the relative factor prices of the small firm at B. Clearly these figures are intended to be indicative of orders of magnitude and are not to be interpreted as exact.

recognised that there may be an unobserved dimension of labour quality that varies across firms. To the extent that this is time invariant this will be captured by the firm fixed effect. Our regressions provide some evidence that age and tenure impact on productivity but the quantitative effect is rather small. Further, while the OLS results suggest that education has a significant impact on productivity, this effect vanishes once controls for firm fixed effects and endogeneity are introduced. Hence, taken together, our measures of human capital appear not to be quantitatively very important in determining productivity.

We have shown that the dispersion in technical efficiency over firms in our sample is similar to that found in other countries. There is no evidence that firms in Africa are particularly inefficient in this sense. We have investigated if either ownership, or firm age, are correlated with the unobserved components of skills, as measured by the fixed effect. Neither significantly affects this measure of the efficiency with which firms operate. In contrast we have found evidence that there is substantial allocative inefficiency. Large firms facing higher relative labour costs than smaller firms use a much more capital intensive technology and operate with costs 20-25 per cent higher than those which would occur if factor prices differentials across firms of differing sizes could be eliminated. While that might be the case has been part of the folk wisdom of development economics for a very long time our data have enabled us to document the size of this effect.

References

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277-297.

Basu, S., Fernald, J. G., 1995. Are apparent productive spillovers a figment of specification error? Journal of Monetary Economics 36, 165-188.

Battese, G.E., Coelli, T.J., 1992. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. Journal of Productivity Analysis 3, 159-169.

Battese, G.E., Coelli, T.J., 1993. A stochastic frontier production function incorporating a model of technical inefficiency effects. Working Papers in Econometrics and Applied Statistics 69. Department of Econometrics, University of New England, Armidale.

Berndt, E., Christensen, L., 1972. The translog function and the substitution of equipment, structures, and labor in U.S. manufacturing, 1929-1968. Journal of Econometrics 1, 81-114.

Bils, M., Klenow P., 2000. Does schooling cause growth? American Economic Review 90, 1160-1183.

Blundell R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115-143.

Blundell R., Bond, S., 2000. GMM estimation with persistent panel data: an application to production functions. Econometric Reviews 19, 321-340.

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

Bond, S., Windmeijer, F., 2001. Finite sample inference for GMM estimators in linear panel data models. Mimeo. The Institute for Fiscal Studies, London.

Caves, R., Bailey, S.D., et al., 1992. Industrial efficiency in six nations. MIT Press, Cambridge and London.

Chen, T., Tang D., 1987. Comparing technical efficiency between import-substitution-oriented and export-oriented foreign firms in a developing economy. Journal of Development Economics 26, 277-289.

Christensen, L. R., Jorgensen, D. W., Lau, L. J., 1971. Transcendental logarithmic production frontiers. The Review of Economics and Statistics 55, 28-45.

Clerides, S.K., Lach, S., and Tybout, J. R., 1998. Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico and Morocco. Quarterly Journal of Economics 113, 903-947.

Corbo, V., de Melo, J., 1986. Measuring technical efficiency: A comparison of alternative methodologies with census data, in: Dogramaci, A. (Ed.), Measurement Issues and Behavior of Productivity Variables. Kluwer Nijhoff, Boston.

Deaton, A., 1997. The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Published for the World Bank by The Johns Hopkins University Press, Baltimore and London.

Dunne, T., Schmitz, J. A., Jr. 1995. Wages, employment structure and employer size-sage premia: Their relationship to advanced-technology usage at US manufacturing establishments. Economica 62, 89-107.

Farrell, M. J., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society A120, 253-290.

Griliches, Z., Hausman, J.A., 1986. Errors in variables in panel data. Journal of Econometrics 31, 93-118.

Griliches, Z., Mairesse, J., 1997. Production functions: the search for identification, in: Strom, S. (Ed.), Essays in Honour of Ragner Frisch. Econometric Society Monograph Series. Cambridge University Press.

Hall, R. E., Jones, C.I., 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114, 83-116.

Hamermesh, D. S. (1993). Labor Demand. Princeton University Press, Princeton, NJ.

Hansen, L. P., 1982. Large Sample Properties of Generalized Method of Moments Estimators. Econometrica 50, 1029-1054.

Jehle, G. A., Reny, P. J., 1998. Advanced Microeconomic Theory. Addison-Wesley, Reading, Massachusetts.

Jones, P., 2001. Are educated workers really more productive? Journal of Development Economics 64, 57-79.

Kremer, M., 1993. The O-Ring Theory of Economic Development. Quarterly Journal of Economics 108, 551-75.

Kremer, M., Maskin, E., 1996. Wage Inequality and Segregation by Skill. Working Paper 5718, National Bureau of Economic Research.

Little, I. M. D., Mazumdar, D., Page Jr, J. M., 1987. Small manufacturing enterprises: A comparative analysis of Indian and other economies. Published for the World Bank by Oxford University Press, Washington.

Lundvall, K., 1999. Essays on Manufacturing Production in a Developing Economy: Kenya 1992-94. Ekonomiska Studier 93. PhD Thesis at Department of Economics, Göteborg University.

Lundvall, K., Battese, G., 2000. Firm size, age and efficiency: evidence from Kenyan manufacturing firms. Journal of Development Studies 36, 146-163.

Manda, D. K., 2002. Wage determination in Kenyan manufacturing. Chapter 5 in Structure and Performance of Manufacturing in Kenya, edited by A. Bigsten and P. Kimuyu. Oxford: Palgrave.

Martin, J.P., Page Jr., J., 1983. The impact of subsidies on X-efficiency in LDC industry: theory and an empirical test. Review of Economics and Statistics 65, 608-617.

Mazumdar, D., Mazaheri, A., 2002. Wages and Employment in Africa. Ashgate Publishing Group.

Oi W. Y., Idson, T. L., 1999. Firm size and wages. Chapter 33 in Handbook of Labor Economics, Volume 3, edited by O. Ashenfelter and D. Card, North-Holland.

Olley, G. S., Pakes, A., 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. Econometrica 64, 1263-1297.

Pack, H., 1976. The substitution of labour for capital in Kenyan manufacturing. Economic Journal 86, 45-58.

Pack, H., 1982. Aggregate implications of factor substitution in industrial processes. Journal of Development Economics 11, 1-38.

Pack, H., 1993. Productivity and industrial development in sub-Saharan Africa. World Development 21, 1-16.

Pitt, M.M., Lee, L. F., 1981. The measurement and sources of technical inefficiency in the Indonesian weaving industry. Journal of Development Economics 9, 43-64.

Sparks, R., 1986. A Model of Involuntary Unemployment and Wage Rigidity: Worker Incentives and the Threat of Dismissal. Journal of Labor Economics 4, 560-581.

Strobl, E. Thornton, R., 2001. Do Large Employers Pay More in Developing Countries? The Case of Five African Countries. Mimeo. University College Dublin and Lehigh University.

Tybout, J. R., 1991. The effects of trade reforms on scale and technical efficiency: new evidence from Chile. Journal of International Economics 31, 231-250.

Tybout, J. R., 1992. Making noisy data sing: Estimating production technologies in developing countries. Journal of Econometrics 53, 25-44.

Tybout, J. R., 2000. Manufacturing firms in developing countries: how well do they do and why? Journal of Economic Literature 38, 11-44.

Tybout, J. R., Westbrook, M. D., 1995. Trade liberalisation and the dimensions of efficiency change in Mexican manufacturing industries. Journal of International Economics 39, 53-78.

Valenchik, A. D., 1997. Government intervention, efficiency wages and the employer size effect in Zimbabwe. Journal of Development Economics 53, 305-338.

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

Yatchew, A., 1998. Nonparametric regression techniques in economics. Journal of Economic Literature 36, 669-721.

TABLE 1

SUMMARY STATISTICS

	[1] Micro	[2] Small	[3] Medium	[4] Large	[5] All
CONDITIONAL MEANS*					
Log Output per Employee	13.75	13.91	14.08	14.68	14.08
	(0.12)	(0.06)	(0.07)	(0.10)	(0.04)
Log Value-Added per Employee	12.48	12.67	12.92	13.68	12.90
	(0.15)	(0.07)	(0.09)	(0.12)	(0.05)
Log Capital per Employee	12.39	12.48	13.73	14.72	13.22
	(0.19)	(0.09)	(0.12)	(0.16)	(0.06)
Ratio of Value-Added to Output	0.34	0.36	0.38	0.39	0.37
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Average Education in Years	9.58	9.81	10.28	11.27	10.18
	(0.27)	(0.13)	(0.16)	(0.22)	(0.09)
Average Age in Years	26.52	28.69	33.43	35.55	31.02
	(0.78)	(0.38)	(0.47)	(0.64)	(0.26)
Average Tenure in Years	3.38	4.22	6.90	7.21	5.42
	(0.48)	(0.24)	(0.29)	(0.39)	(0.16)
Log Average Labour Cost	11.94	12.25	12.67	13.05	12.48
	(0.10)	(0.05)	(0.06)	(0.08)	(0.03)
UNCONDITIONAL MEANS					
Firm Age (years)	15.17	14.45	19.28	22.87	17.39
Any Foreign Ownership (proportion)	0.14	0.07	0.23	0.59	0.22
Any State Ownership (proportion)	0.00	0.02	0.09	0.06	0.05
Industry (proportions)					
Food and Bakery	0.32	0.22	0.18	0.25	0.22
Wood	0.00	0.03	0.06	0.29	0.08
Furniture	0.13	0.28	0.27	0.20	0.25
Textiles and Garments	0.33	0.16	0.20	0.00	0.16
Metal and Machinery	0.22	0.31	0.28	0.25	0.28
Location (proportions)					
Accra (capital city)	0.32	0.60	0.63	0.55	0.57
Cape	0.06	0.02	0.03	0.03	0.03
Kumasi	0.49	0.35	0.28	0.13	0.30
Takoradi	0.14	0.03	0.06	0.29	0.10
Number of Observations	72	292	193	119	677
Number of Firms	15	61	40	27	143

Note: The size of the firm is its total number of employees when first observed in the sample, where a micro firm has less than six employees, a small firm has from 6 to 29, a medium firm has from 30 to 99, while a large firm has 100, or more, employees. The figures in ( ) are standard errors. All monetary variables have been deflated using firm-level price indices.

Table notes continue on the next page

Table notes, continued

* The numbers reported in Columns [1]-[4] are predictions based on OLS regressions in which the regressors are sector, time and size dummies. The predicted values are calculated using sample means of the sector and time dummies, i.e. of the form $\hat{y}_{size\_i} = \hat{a} \cdot size\_i + \hat{b} \cdot \bar{x}$, where *size_i* indicates the *i*th size category, $\bar{x}$ is the vector of mean values of the sector and time dummies, and $\hat{a}$ and $\hat{b}$ are estimated coefficients. The numbers reported in the fifth column are predictions based on sample means of all regressors, hence they are effectively unconditional means.

TABLE 2: OLS AND WITHIN ESTIMATES OF PRODUCTION FUNCTION PARAMETERS

	Dependent variable: log Value-Added				Dependent variable: log Output			
	Translog		Cobb-Douglas		Translog		Cobb-Douglas	
	[1] OLS	[2] Within	[3] OLS	[4] Within	[5] OLS	[6] Within	[7] OLS	[8] Within
MARGINAL EFFECTS[1]								
log Employment	0.84	0.30	0.89	0.34	0.10	0.03	0.14	0.08
	(8.79)**	(1.31)	(9.74)**	(1.59)	(3.33)**	(0.46)	(3.78)**	(1.18)
log Capital	0.20	-0.25	0.18	-0.22	0.02	-0.06	0.03	-0.06
	(3.57)**	(0.65)	(3.43)**	(0.59)	(1.68)$^{+}$	(0.46)	(2.12)*	(0.53)
log Raw Materials					0.72	0.69	0.71	0.65
					(37.49)**	(22.78)**	(29.12)**	(17.12)**
log Indirect Costs					0.15	0.11	0.12	0.10
					(8.07)**	(4.51)**	(5.19)**	(3.47)**
HUMAN CAPITAL COEFFICIENTS								
Education	0.04	0.02	0.04	0.02	0.01	0.00	0.01	0.00
	(1.63)	(0.57)	(1.59)	(0.53)	(2.53)*	(0.35)	(2.11)*	(0.35)
Age	0.13	0.21	0.13	0.20	0.05	0.07	0.04	0.08
	(2.19)*	(3.08)**	(2.04)*	(3.09)**	(3.04)**	(2.59)**	(2.53)*	(3.01)**
Age^2 / 100	-0.20	-0.33	-0.19	-0.33	-0.07	-0.11	-0.06	-0.13
	(2.26)*	(3.46)**	(2.09)*	(3.49)**	(2.96)**	(2.82)**	(2.58)**	(3.33)**
Tenure	0.03	0.05	0.03	0.05	0.004	0.01	0.01	0.02
	(1.70)$^{+}$	(2.46)*	(1.43)	(2.56)*	(1.08)	(2.55)*	(1.35)	(3.47)**
DIAGNOSTICS & TESTS								
R^2	0.74	0.10	0.74	0.09	0.98	0.82	0.97	0.80
Quasi-concavity (proportion)	1.00	0.61			0.44	0.04		
Monotonicity (proportion)	1.00	0.00			0.69	0.04		
Homotheticity[2] (*p*-value)	0.30	0.64			0.25	0.68		
Homotheticity and CRS[3] (*p*-value)	0.36	0.21	0.39	0.04	0.31	0.49	0.86	0.06
Cobb-Douglas (*p*-value)[4]	0.49	0.82			0.00	0.00		

Table notes on the next page

Table notes on the next page

Note: Time dummies are included in all regressions. The OLS regressions include controls for the age of the firm, industry, ownership structure and location. The numbers in ( ) are *t*-statistics based on standard errors robust to heteroskedasticity. Significance at the one per cent, five per cent and ten per cent level is indicated by *, ** and + respectively.

(1) For the translog specification the marginal effects is a function of the inputs, and have therefore been evaluated at the sample means. For the Cobb-Douglas specification, the marginal effects are equal to the estimated coefficients.

(2) H_0: $\sum_k b_{km} = 0$, $m = 1,2,\ldots,J$ (see eq. [3]).

(3) For translog specifications, H_0: $\sum_k b_{km} = 0$, $m = 1,2,\ldots,J$, and $\sum_k b_k = 1$. For Cobb-Douglas specifications, H_0: $\sum_k b_k = 1$ (see eq. [3]).

(4) H_0: $b_{km} = 0$, $k = 1,2,\ldots,J$; $m = 1,2,\ldots,J$ (see eq. [3]).

TABLE 3

SYSTEM GMM ESTIMATES OF PRODUCTION FUNCTION PARAMETERS

	Dependent variable: log Value-Added[a]		Dependent variable: log Output[b]	
	[1] Translog	[2] Cobb-Douglas	[3] Translog	[4] Cobb-Douglas
MARGINAL EFFECTS[1]				
log Employment	0.88	0.73	0.10	0.17
	(4.20)**	(3.25)**	(0.93)	(2.37)*
log Capital	0.25	0.31	0.08	0.09
	(2.49)*	(3.58)**	(1.61)	(2.06)*
log Raw Materials			0.68	0.68
			(12.58)**	(14.02)**
log Indirect Costs			0.13	0.06
			(2.35)*	(1.14)
HUMAN CAPITAL COEFFICIENTS				
Education	-0.01	-0.003	0.01	0.006
	(0.82)	(0.07)	(0.65)	(0.35)
Age	0.24	0.26	0.04	0.07
	(2.40)*	(3.11)**	(1.42)	(2.61)**
Age^2 / 100	-0.38	-0.41	-0.08	-0.11
	(2.60)**	(3.35)**	(2.07)*	(2.88)**
Tenure	0.04	0.05	0.01	0.02
	(0.76)	(1.25)	(1.48)	$(1.89)^+$
DIAGNOSTICS & TESTS				
Quasi-concavity (proportion)	0.52		0.44	
Monotonicity (proportion)	0.87		0.59	
Homotheticity[2] (*p*-value)	0.63		0.46	
Constant returns to scale[2] (*p*-value)	0.74	0.85	0.59	0.92
Cobb-Douglas (*p*-value) [2]	0.66		0.37	
m1 (*p*-value) [3]	0.00	0.00	0.01	0.00
m2 (*p*-value) [4]	1.00	0.93	0.81	0.15
Sargan-Hansen (*p*-value) [5]	0.57	0.79	0.42	0.39

Note: Time dummies are included in all regressions. The numbers in ( ) are *t*-statistics. Significance at the one per cent, five per cent and ten per cent level is indicated by *, ** and $^+$ respectively. Hypothesis tests are based on robust, finite sample corrected standard errors (see footnote 13) calculated using the method proposed by Windmeijer (2000).

[a] The instrument set for the differenced equation consists of the levels of employment, physical and human capital, in periods t-2 and t-3. The instrument set for the levels equation consists of employment, physical and human capital, differenced, in period t-1, a constant and year dummies.

[b] The instrument set for the differenced equation consists of the levels of employment, raw material, indirect costs and physical and human capital, in periods t-2 and t-3. The instrument set for the levels equation consists of employment, raw material, indirect costs and physical and human capital, differenced, in period t-1, a constant and year dummies.

Table notes continue on the next page

Table notes, continued

(1) For the translog specification the marginal effects is a function of the inputs, and have therefore been evaluated at the sample means. For the Cobb-Douglas specification, the marginal effects are equal to the estimated coefficients.

(2) See Table 2.

(3) Tests the null hypothesis that the differenced residuals in periods *t* and *t*-1 are uncorrelated.

(4) Tests the null hypothesis that the differenced residuals in periods *t* and *t*-2 are uncorrelated.

(5) Tests for the validity of the overidentifying restrictions.

TABLE 4

CORRELATES WITH THE FIXED EFFECTS IN THE PRODUCTION FUNCTION

	Fixed Effects from Column [4], Table 3	
	Coefficient	*t*-value
SECTORS[(a)]		
Food	0.108	1.89+
Textiles or Garments	-0.126	1.83+
Wood	0.242	2.43*
Furniture	-0.084	1.31
LOCATION[(b)]		
Accra (Capital City)	0.072	1.02
Kumasi	0.022	0.29
Cape Coast	-0.056	0.44
OTHER COMPANY CHARACTERISTICS		
Firm Age/100	0.055	0.29
Any Foreign Ownership	-0.011	0.21
Ghanaian State Ownership	0.032	0.34
R^2	0.16	
F-test slope coefficients		3.03**
Number of firms	143	

Note: The dependent variable is the estimated fixed effect from Column [4], Table 3. The estimation method is OLS. *t*-values are based on robust standard errors. Significance at the ten per cent and five per cent level is indicated by + and *, respectively.

(a) Omitted category: Metal.

(b) Omitted category: Takoradi.

TABLE 5

LABOUR COST REGRESSIONS

	Dependent variable: log Labour Cost			
	[1] OLS	[2] Within	[3] SYS GMM[(a)]	[4] SYS GMM[(a)]
COEFFICIENTS				
log Employment	-0.15	0.09	0.13	0.15
	(1.31)	(0.38)	(0.35)	(2.13)*
[log Employment]^2	0.03	-0.004	0.002	
	(1.98)*	(0.11)	(0.05)	
Education	0.06	0.03	-0.01	-0.01
	(3.53)**	(1.22)	(0.36)	(0.36)
Age	0.26	0.13	0.20	0.20
	(5.18)**	(2.03)*	(3.90)**	(3.93)**
Age^2 / 100	-0.32	-0.15	-0.23	-0.23
	(4.30)**	(1.60)	(3.31)**	(3.37)**
Tenure	0.04	0.04	0.04	0.04
	(2.53)*	(1.92)$^+$	(2.12)*	(2.13)*
EMPLOYMENT ELASTICITY				
Evaluated at employment = 20	0.03	0.07	0.14	0.15
Evaluated at employment = 100	0.13	0.06	0.15	0.15
Evaluated at employment = 200	0.17	0.05	0.15	0.15
DIAGNOSTICS & TESTS				
R^2	0.64	0.21		
Employment coefficients=0 (*p*-value) [(1)]	0.01	0.72	0.095	0.03
m1 (*p*-value) [(2)]			0.00	0.00
m2 (*p*-value) [(3)]			0.13	0.13
Sargan-Hansen (*p*-value) [(4)]			0.28	0.31

Note: Time dummies are included in all regressions. The OLS regressions include controls for the age of the firm, industry, ownership structure and location. The numbers in ( ) are *t*-statistics based on standard errors robust to heteroskedasticity. Significance at the one per cent, five per cent and ten per cent level is indicated by *, ** and $^+$ respectively. Columns [3] and [4] report finite sample corrected standard errors (see footnote 13), calculated using the method proposed by Windmeijer (2000).

[a)] The instrument set for the differenced equation consists of the log of employment, the squared log of employment, education, age, age squared and tenure, in levels, in periods t-2 and t-3. The instrument set for the levels equation consists of employment, employment squared, age, age squared and tenure, differenced, in period t-1, a constant and year dummies.

[(1)] Tests the null hypothesis the coefficients on log Employment and its squared term are jointly equal to zero.

[(2)] Tests the null hypothesis that the differenced residuals in periods *t* and *t*-1 are uncorrelated.

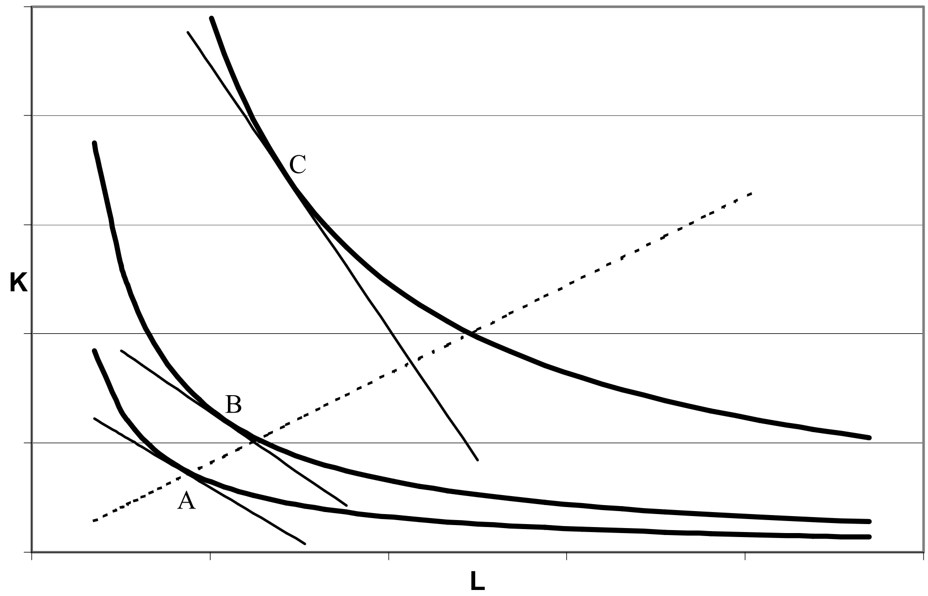
[(3)] Tests the null hypothesis that the differenced residuals in periods *t* and *t*-2 are uncorrelated.

[(4)] Tests for the validity of the overidentifying restrictions.

FIGURE 1

TECHNOLOGY AND FACTOR PRICES

A. Homothetic Technology, Variable Factor Price Ratio



B. Non-Homothetic Technology, Constant Factor Price Ratio

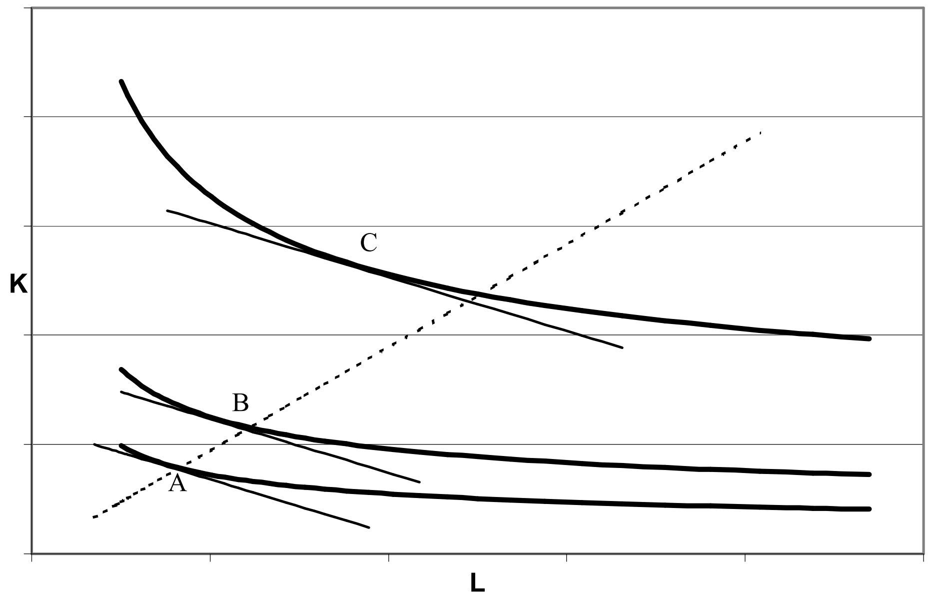
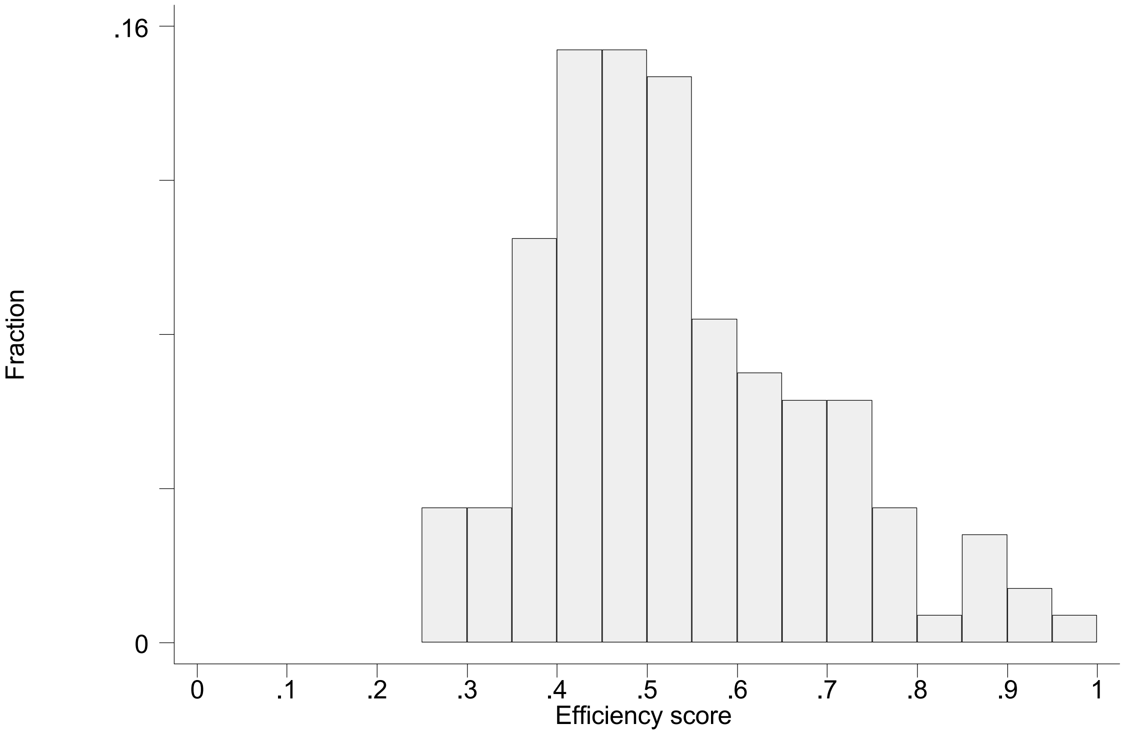


FIGURE 2

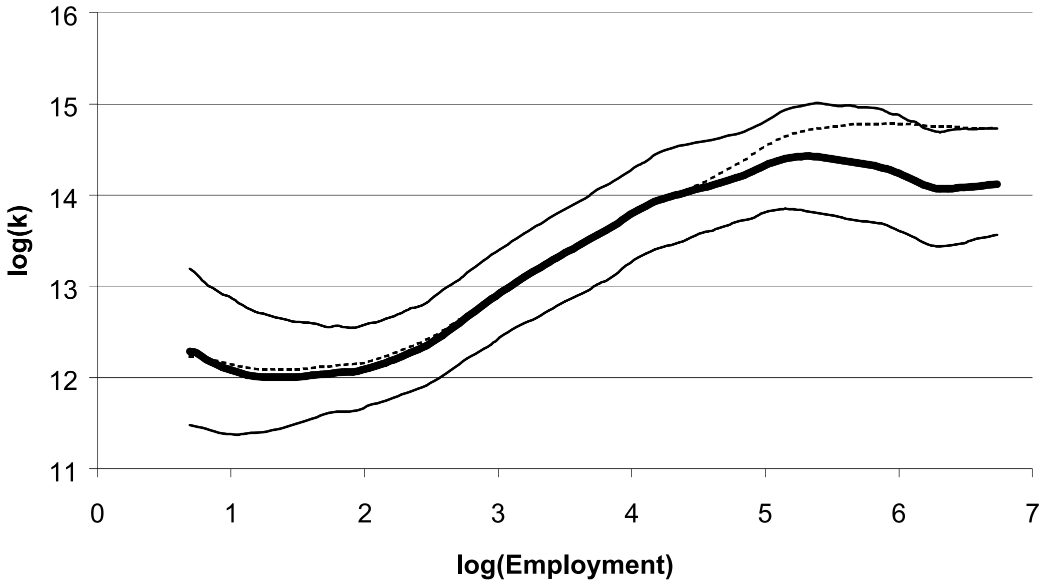
THE DISTRIBUTION OF TECHNICAL EFFICIENCY



Note: Technical efficiency is defined as $te_i = e^{-(\hat{m}_{\max} - \hat{m}_i)}$, where $\hat{m}_{\max}$ is the sample maximum of the estimated fixed effects, and $\hat{m}_i$ is the estimated fixed effect for firm i. This is interpretable as a measure of the dispersion across firms of productivity, conditional on inputs and human capital. The estimated mean of technical efficiency is 0.53 and the estimated standard deviation is 0.15. The fixed effects estimates are based on the regression in Column [4], Table 3.

FIGURE 3

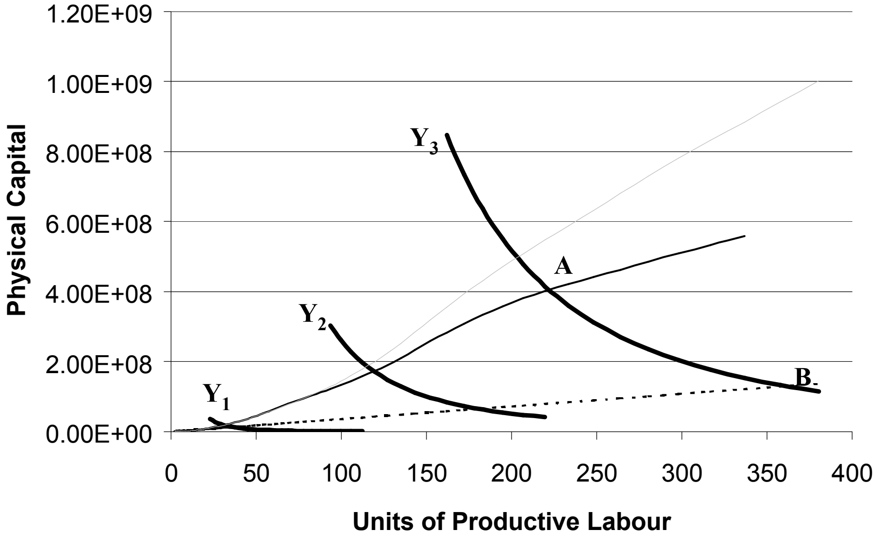
CAPITAL INTENSITY AND FIRM SIZE



Note: The solid line shows the regression line of ln(*k*) on ln L. The kernel is Epanechnikov and the bandwidth is equal to 1.20. The thin lines indicate pointwise 95 per cent confidence bands, calculated from 800 bootstrapped replications. To take the panel nature of the data into account we bootstrapped from the firms rather than from the observations, which is a similar procedure to that used by Deaton (1997, pp 216-218) for clustered data. The dashed line shows the regression line of ln(*k*) on ln L when *k* is not adjusted for worker quality heterogeneity as explained in the text.

FIGURE 4

TECHNOLOGY AND FACTOR CHOICE

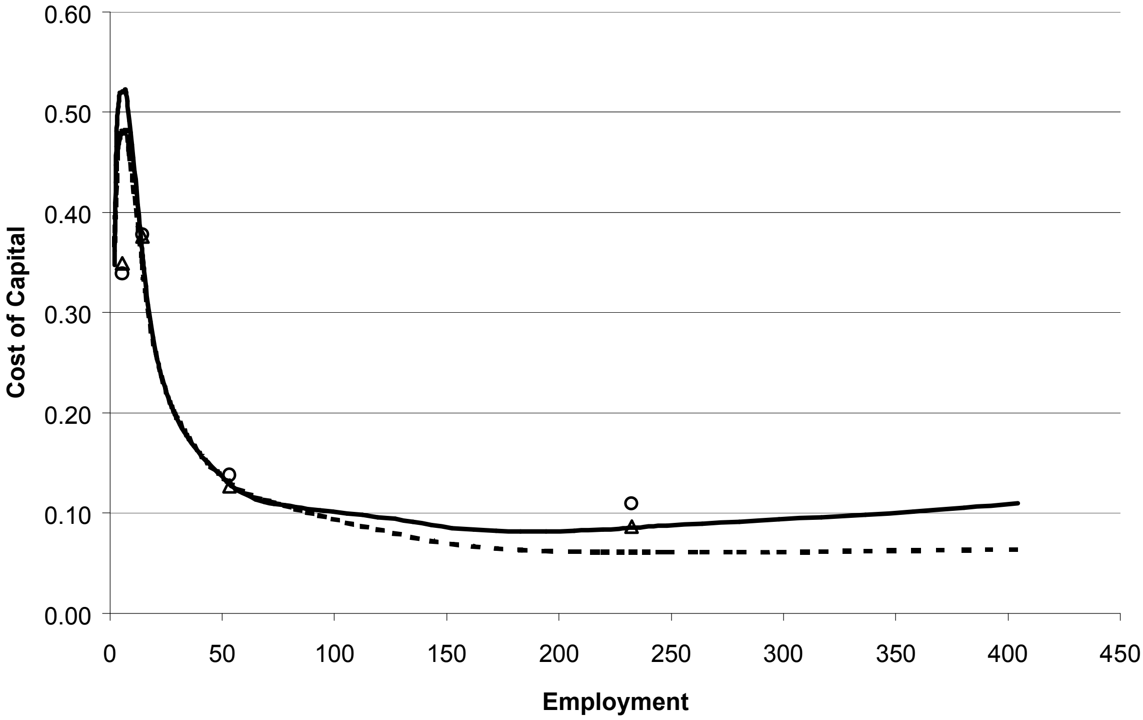


Capital-labour ratio at **A**: 1.83 million Cedis.
Capital-labour ratio at **B**: 0.36 million Cedis.

Notes: The underlying production function is ln $Y = \ln A + 0.3 \ln K + 0.7 \ln L$, where A is a constant. Y_1, Y_2 and Y_3 depict three isoquants under this technology, with $Y_2 = 5Y_1$ and $Y_3 = 10Y_1$. The solid line starting in the origin is the empirical expansion path, derived from the non-parametric regression shown in Figure 3. The shaded line is the expansion path following from the capital intensity regression not adjusted for labour quality. The dashed line starting in the origin is a hypothetical expansion path for relative factor prices constant at the level observed for a firm with 18 employees.

FIGURE 5

THE COST OF CAPITAL AND FIRM SIZE



6   0.09(Output / Capital)   *   0.31(Value-Added / Capital)

Notes: The solid line shows the cost of capital based on the semiparametric regression of quality adjusted capital intensity on size. The dashed line shows the cost of capital based on the same regression without quality adjustment. The points indicated by 6 and ∗ indicate the marginal productivity of capital using the appropriate capital coefficient, evaluated at the mean values of log(Output / Capital) and log(Value-Added / Capital) reported in Table 1, by size category. The points are positioned horizontally at the mean values of employment for each size category.

Appendix 1: The system GMM estimator

This appendix provides a brief description of the system GMM estimator. For more details see Blundell and Bond (1998).

Consider

(A1) $$y_{it} = x'_{it}b + m_i + e_{it}\,, \qquad t = 1,2,\ldots,T,$$

where i and t are firm and time indices, y_{it} is the dependent variable, x_{it} is a row vector of order k of explanatory variables possibly including lags of the dependent variable, b is a column vector of parameters of order k, m_i is a fixed effect potentially correlated with x_{it} and e_{it} is a residual potentially correlated with x_{it}. To eliminate the fixed effect we take first differences:

(A2) $$\Delta y_{it} = \Delta x'_{it}b + \Delta e_{it}\,, \qquad t = 2,3,\ldots,T.$$

If Δx_{it} is correlated with the differenced residual, the standard OLS estimator will be biased and inconsistent. However, assume that there exists a set of instruments that enable us to form a vector of moment conditions of order q, defined as

(A3) $$E\left(z'_{it}\Delta e_{it}\right) = 0\,.$$

Provided $q \geq k$, we can obtain a consistent GMM estimator of b by minimising the quadratic

(A4) $$J\left(\hat{b}_{GMM}\right) = \bar{g}\left(\hat{b}_{GMM}\right)' W_N^{-1}\bar{g}\left(\hat{b}_{GMM}\right),$$

where $\bar{g}(\cdot)$ is the sum over the sample moment conditions of the form in (A3) and W_N^{-1} is a weight matrix (Hansen, 1982). A common procedure is to use lags of x_{it} as instruments for the differenced equation (A2), and because more instruments become available for higher t, we can form a matrix of instrument as

(A5) $$\mathbf{z_i} = \begin{array}{c} \begin{bmatrix} x_{i1} & 0 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & x_{i1} & x_{i2} & \ldots & 0 & \ldots & 0 \\ . & . & . & \ldots & . & \ldots & . \\ 0 & 0 & 0 & \ldots & x_{i1} & \ldots & x_{i,T-l} \end{bmatrix} \end{array} \begin{array}{l} t = 1 + l \\ t = 2 + l \\ \\ t = T \end{array}$$

where l is the lag length in use. The resulting differenced GMM estimator often performs poorly in practice due to the problem of weak instruments. Blundell and Bond (1998) proposed combining the differenced equation (A2) with the levels equation (A1), for which lagged *differences* of the explanatory variables may serve as valid instruments. The vector of moment conditions is then defined as

(A6) $$E\left(\mathbf{z_i^{+}}'\mathbf{u}\right) = 0\,,$$

where

(A7) $$\mathbf{u}_i = \begin{bmatrix} ?\,\mathbf{e_i} \\ \mathbf{e_i} \end{bmatrix}$$

and

(A8) $$\mathbf{z_i^{+}} = \begin{bmatrix} \mathbf{z_i} & 0 & 0 & \ldots & 0 \\ 0 & \Delta x_{i1} & 0 & \ldots & 0 \\ 0 & 0 & \Delta x_{i2} & \ldots & 0 \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & \Delta x_{i,T-l} \end{bmatrix} \begin{array}{l} \\ t = 1 + l \\ \\ \\ t = T \end{array}$$

The system GMM estimates are then obtained by minimising (A4), where $\bar{g}(\cdot)$ is the sum over the sample moments of the form in (A6).