UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Tel: (+43-1) 26026-0 · www.unido.org · unido@unido.org

Does firm size really affect earnings?

December 2002

Måns Söderbom[£,+], Francis Teal[£] and Anthony Wambugu[$]

Abstract

In this paper we investigate the implications of labour and capital market imperfections for the relationship between firm size and earnings. To establish that such a question is of interest we need to show that the firm size-wage effect cannot be explained by either the observed or unobserved skills of the workforce or the characteristics of the workplace. To do that we require data where controls are possible for observable time-varying firm and worker characteristics, as well as the unobservable characteristics of both the firm and its workers. Our data is a sample of workers matched with firms over time so such controls are possible. Changes in wages are shown to respond to changes both to profits per employee and the size of the firm. It is argued that these empirical results clearly reject the hypothesis that the firm-size relationship can be explained by the skills of the workers. They can be shown to be consistent with some forms of non-competitive theories of bargaining and efficiency wages.

JEL classification: J3, L25.
Keywords: Skills, efficiency wages, bargaining, firm size, earnings, African manufacturing.

[£] Centre for the Study of African Economies, Department of Economics, University of Oxford.
[$] Department of Economics, Göteborg University.
[+] Corresponding author: Mans.Soderbom@economics.ox.ac.uk

* The Ghanaian data used in this paper were collected by a team from the Centre for the Study of African Economies (CSAE), 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 Kenyan data for 1994 were collected by a team from the Economics Departments in Göteborg and Nairobi, as part of the Regional Program on Enterprise Development (RPED) organised by the World Bank. The Kenyan RPED surveys were funded by SIDA. The Kenyan data for 1999 were collected by a team from the CSAE as part of a research programme funded by UNIDO. The CSAE is funded by the Economic and Social Research Council of the UK.

1. Introduction

The human capital model, in which earnings reflect skill differentials in perfect factor markets, has dominated the interpretation of earnings functions. The finding that earnings rise with firm size has been widely interpreted in this framework. The human capital explanation is that the vector of relevant productive skills is partially unobserved, and that the significance of firm characteristics in earnings regressions essentially reflects unobserved labour quality, Oi and Idson (1999). If large firms hire more able individuals than do small firms, for instance, and ability is partially unobserved to the econometrician, then the result that firm size is positively correlated with earnings is entirely consistent with the standard human capital model and competitive labour markets.

In recent years there has been a rapid development of non-competitive models of the labour market, specifically models of efficiency wages (see Shapiro and Stiglitz, 1984; Weiss 1980; Akerlof and Yellen, 1986, for theoretical rationales for the payment of efficiency wages and Dickens and Katz, 1987; Raff and Summers, 1987; Krueger and Summers, 1987, 1988; Katz and Summers, 1989; Wadhwani and Wall, 1991; Levine 1992; Moll 1993; and Huang et al 1998, for tests) and bargaining (see Slichter, 1950; Van Reenen, 1996; and Blanchflower et al. 1996, for developed countries and Teal 1996; Valenchik 1997; and Azam 2001, for developing countries). These studies document inter-industry or inter-firm wage differentials that are seemingly inconsistent with the competitive model, and some of the studies proceed by investigating whether these differentials can be linked to observable firm characteristics, notably profitability and monitoring costs.

In parallel with this development of models of wage determination has been an extensive empirical investigation of how wages link to firm size (see Mellow, 1982; Brown and Medhoff, 1989; Troske 1999; and Bayard and Troske, 1999, for analyses based on U.S. data; see Mazumdar 1983; Valenchik 1997; Strobl and Thornton, 2001; Manda, 2002; and Mazumdar and Mazaheri, 2002, for evidence from developing countries). In the most recent and comprehensive analysis on US data, Troske (1999) and Bayard and Troske (1999) use matched employer-employee data to investigate which aspects of the theories of human capital, efficiency wages and bargaining can explain the firm size-wage relationship. Troske (1999) concludes that once as comprehensive an allowance as possible is made for the factors suggested by these theories "there still remains a large, significant and unexplained size-wage premium" (p.25).

In this paper we explore further the implications for the size-wage premium of bargaining and efficiency wage models of wage determination. We extend these models to include capital constraints and show that such constraints can increase the size-wage premium. We draw on data from two African countries – Kenya and Ghana - where capital market constraints are known to be pervasive (see Bigsten et al., 1999). To establish a causal link from size to wages we need to show that the scale effect cannot be removed either by factors that may be related to the unobserved quality of the workforce or by other aspects of the firms' performance such as profitability and monitoring costs.

Section 2 summarises the numerous hypotheses that have been advanced to explain the fact that wages rise with firm size and outlines how we propose to use our matched panel of workers with firms to test these alternative theories. Section 3 covers details about the data including summary statistics. In section 4 we investigate the effects of firm and worker characteristics on the firm size wage relationship allowing for fixed effects by differencing but confining our attention to OLS estimates. The objective in section 5 is to assess if size, and other aspects of the firm's performance, can be given a causal interpretation by finding instruments that allow for the endogeneity of size and firm outcomes. In section 6 we ask whether either bargaining, or efficiency wage, models can explain the results. Section 7 provides conclusions.

2. Analytical Framework and Empirical Approach

Our point of departure is the standard Mincerian framework stating that differences in individual log earnings are driven exclusively by differences in human capital, $\ln w = \psi h + \eta$, where h is a vector of observed human capital variables and η denotes a dimension of labour quality that is observable to the firm but unobservable to the econometrician. This will be an appropriate specification if the labour market is competitive, so that firms are wage-takers. We assume that the vector h consists of years of education, tenure, age and age squared, and model η as an individual specific, time invariant, effect. We augment the Mincerian earnings function with a range of observable firm-level variables, summarised by a vector f, and a firm specific fixed effect τ which is unobserved. Adding a time effect θ_t and a residual ν_{ijt} we hence write our baseline earnings function as

[1] $$y_{ijt} = \psi \cdot h_i + \gamma \cdot f_{jt} + \theta_t + \eta_i + \tau_j + \nu_{ijt},$$

where y denotes the logarithm of the wage, ψ, γ and are parameters to be estimated and i, j, t denote employee, firm and time respectively.

Much recent research on firm characteristics and wages has focussed on the role of firm size, seeking to explain why large firms pay higher wages than small firms. Numerous explanations have been suggested in various strands of the labour economics literature. The most influential theory has been the human capital model, where the common factor in many of the arguments drawn is that firm size will be correlated with some dimension of worker quality. Hamermesh (1980, 1993) argues that if physical and human capital are complements in the production process, then the most skilled workers will be employed by the largest firms. Kremer (1993) and Kremer and Maskin (1996) propose that there are advantages to matching high-skill workers with other high-skill workers, and that there are fixed costs (i.e. decreasing average costs) to hiring skilled workers. Because large firms can absorb the fixed costs they are more likely to match high-skill workers. In a similar vein, Dunne and Schmitz (1992) argue that there is a complementarity between the degree of sophistication of physical capital and the skill of workers, and that large firms have larger amounts of output over which to amortize the fixed costs associated with adopting sophisticated capital. Brown and Medoff (1989) suggest that firms that pay their workers more are more likely to survive and grow - such workers presumably being better motivated. All these hypotheses have in common that some aspect of the skills or quality of the workforce is not adequately controlled for in the regression.

Other theories have predicted a size-wage relationship resulting not from omitted skills, but for reasons to do with the working conditions inside the firm or the way the firm is managed. Efficiency wage models suggest that because monitoring is more expensive in large than in small firms, large firms pay higher wages in order to motivate their workers not to shirk, Bulow and Summers (1986). Doeringer and Piore (1971) put forward a theory of internal labour markets, where as internal recruitment is less costly than hiring outsiders, large firms are willing to pay wage premiums to workers at low levels in the hierarchy in order to retain a sufficiently large pool of potential workers to consider for promotion. Masters (1969) argue that there are compensating wage differentials. Working conditions in larger firms are worse than in smaller ones, so workers must be compensated.

Bargaining models have been less concerned with explaining the relationship between firm size and wages and have instead focused on showing that firm profits enter the wage equation. However as larger firms may be more profitable it is clearly possible that the firm size effect is proxying profits.

These potential explanations provide us with a set of firm variables whose omission might be the reason for the observed size-earnings relation. The firm-level variables - the arguments of the *f* vector of the Mincerian earnings function – that we propose to include in the regression are the log of the capital labour ratio, in order to control for the capital-skill complementarity hypothesis; the log of labour productivity, to control for unobserved productivity of workers (and possibly rent sharing); profits per employee, the variable most directly implied by the rent sharing hypothesis; firm age, as suggested by Brown and Medoff (1989); the average education in the firm, to allow for matching of skilled workers in large firms; and, finally, the proportion of managers and supervisors in the work place, the variables related to monitoring costs as hypothesised by the efficiency wage models.

We remove the unobserved time invariant terms η and τ by differencing equation [1], which yields

$$[2] \qquad \tilde{y}^s_{ijt} = \gamma \cdot \tilde{f}^s_{jt} + \tilde{\theta}^s_t + \tilde{v}^s_{ijt},$$

where $\tilde{f}^s_{jt} = \Delta_s f_{jt}$, $\tilde{v}^s_{ijt} = \Delta_s v_{ijt}$, $\tilde{y}^s_{ijt} = \Delta_s \tilde{y}_{ijt}$, Δ_s indicates differences of the order *s* and $\tilde{\theta}^s_t$ is a time effect that varies by the order of differencing *s*.[1] Notice that the differencing wipes out all time-invariant observable variables, e.g. education, and that to the extent that there are age or tenure effects, these will now be absorbed in the time effect. Although the differencing procedure eliminates all time invariant unobserved factors that potentially affect earnings, e.g. time invariant cognitive skills or

[1] That is, $\Delta_1 X_t = X_t - X_{t-1}$; $\Delta_2 X_t = X_t - X_{t-2}$; and so on.

personality factors, it is likely that there remains a correlation between the regressors and the residual. We distinguish between four types of bias, and discuss these next.

The first potential problem is that posed by attrition, caused in this context by the fact that some employees leave their employers during the period spanned by the panel. It turns out that for our data set, there is considerable attrition (see Section 3), which may create a sample selectivity problem if individuals drop out of the sample for reasons that are not entirely random. More precisely, if there are unobservable factors determining the probability of attrition that are correlated with unobservable factors driving change in earnings, then failure to account for the sample selectivity problem will result in biased and inconsistent results.[2] Following Wooldridge (2002) we attempt to test and correct for selectivity bias by using a two-stage approach, which is closely related to the model developed by Heckman (1976). This approach involves estimating a probit model determining the likelihood that an individual observed in the base period will be observed again in future periods, calculating the selectivity variable (the inverse Mill's ratio) and including this as an additional regressor in the wage equation. Provided that the model is correctly specified, this approach will give consistent estimates of γ in the presence of selectivity. Further details of the model are provided in Appendix 1, Part A.

The second possible source of bias is that explanatory variables are almost certainly measured with some degree of error. If ignored, measurement errors are expected to cause a downward bias in the estimated coefficients. Indeed, if the measurement errors are serially uncorrelated while the true but unobserved values of the explanatory variables are slow changing, taking time differences is likely to aggravate the measurement error bias. Griliches and Hausman (1986) note that in such a case estimators based on 'long' differences will be less severely biased than 'short' differenced results, and we will shortly discuss how we intend to draw on this insight.[3]

The third possible problem is that there are factors unobserved to the econometrician that impact both on explanatory variables and on the wage variable. For instance, managers may respond

[2] Sample selection may also occur if workers choose which firm size to work in. Idson and Feaster (1990) consider this in a cross-sectional setting.

[3] The reason is the signal-to-noise ratio will increase with the length of differencing.

to an unobserved demand shock by raising wages and by investing in physical capital, in which case the OLS estimate of the capital coefficient would be upward biased.

Fourth, it is likely that there is reverse causality in the form of feedback from wages onto both employment and capital through the factor demand functions. Clearly, in a standard neoclassical model an exogenous positive shock to wages will have a negative effect on employment, holding everything else constant, and to the extent that such a shock is unobserved the OLS coefficient on employment in the wage equation will be downward biased. By the envelope theorem the positive wage shock may have a negative effect on the demand for capital as the marginal profitability of capital will be lower for lower levels of employment, see e.g. Denny and Nickell (1992) for a derivation of an investment equation in which the coefficient on the wage variable is negative.

We intend to correct our estimates for the second, third and fourth sources of bias by using instrumental variable techniques. To illustrate the approach we begin by generalising [2] distinguishing between $\Gamma = (T\text{-}1) + (T\text{-}2) + ... + (T\text{-}S)$ equations and allowing for different coefficients across the equations:

Second period, first differenced: $\tilde{y}^1_{ij2} = \gamma_{12} \cdot \tilde{f}^1_{j2} + \tilde{\theta}^1_2 + \tilde{v}^1_{ij2}$,

Third period, first differenced: $\tilde{y}^1_{ij3} = \gamma_{13} \cdot \tilde{f}^1_{j3} + \tilde{\theta}^1_3 + \tilde{v}^1_{ij3}$,

(...)

[3] T^{th} period, first differenced: $\tilde{y}^1_{ijT} = \gamma_{1T} \cdot \tilde{f}^1_{jT} + \tilde{\theta}^1_T + \tilde{v}^1_{ijT}$,

Third period, second differenced: $\tilde{y}^2_{ij3} = \gamma_{23} \cdot \tilde{f}^2_{j3} + \tilde{\theta}^2_3 + \tilde{v}^2_{ij3}$,

(...)

(...)

T^{th} period, S^{th} differenced: $\tilde{y}^S_{ijT} = \gamma_{ST} \cdot \tilde{f}^S_{jT} + \tilde{\theta}^S_T + \tilde{v}^S_{ijT}$.

That is, if $T = 3$ we have two first differenced equations, and one second differenced; if $T = 4$, there will be three first differenced equations, two second differenced equations and one third differenced; and so on. One advantage to generalising the model like this is that it becomes straightforward to test the restriction inherent in [2] that $\gamma_{12} = \gamma_{13} = ... = \gamma_{1T} = ... = \gamma_{ST}$. If the estimates of different lengths differ significantly this suggests that measurement errors are present, particularly if the point estimates tend to increase in absolute magnitude with the length of differencing, Griliches and Hausman (1986).

A second advantage is that we can use different instrument sets for different equations in [3]. It is likely that for some of the equations in [3] only a small number of instruments are available while for other equations the instrument set is richer, and clearly the more instruments that can be exploited, the more efficient is the resulting estimator (see e.g. the literature on estimation of dynamic panel data models, Arellano and Bond, 1992; Arellano and Bover, 1995; Blundell and Bond, 1998). To estimate the parameters we adopt a generalised method of moments (GMM; Hansen, 1982) framework, within which we formulate estimators that assume that the residuals are uncorrelated with the regressors, as well as estimators that do not rely on this assumption. We outline the GMM framework next.

Assume that q_{st} moment conditions of the form

[4] $$E(z_{ijt}^{s}\tilde{v}_{ijt}^{s}) = 0,$$

where z_{ijt}^{s} is a vector of order q_{st}, are available for the relevant equation (*s*, *t*). Define z_{ij} as a $(\Gamma \times q)$ block diagonal matrix with z_{ijt}^{s} in the appropriate block and $\tilde{v}_{ij} = \left(\tilde{v}_{ij2}^{1}, \tilde{v}_{ij3}^{1}, ..., \tilde{v}_{ijT}^{1}, \tilde{v}_{ij3}^{2}, ..., \tilde{v}_{ijT}^{S}\right)'$ as a stacked vector of residuals for individual *i* in firm *j*, where $q = \sum_{s}\sum_{t} q_{st}$ denotes the total number of instruments across all equations. Because the differenced residuals in [3] almost certainly are correlated with each other, it is efficient to allow for cross-equation correlation of the residuals by estimating all equations in [3] simultaneously. Assumed that $q > k$, where *k* is the number of parameters to be estimated, the GMM estimates obtained from simultaneous estimation of [3] are given by

[5] $$(\gamma, \theta)' = \left(\left(Z'\tilde{F}\right)' W^{-1} Z'\tilde{F}\right)^{-1} \left(Z'\tilde{F}\right)' W^{-1} Z'\tilde{Y},$$

where $Z = (z_{11}', z_{21}', ..., z_{IJ}')'$ is a matrix stacking the individual instrument matrices, $\tilde{F}$ and $\tilde{Y}$ are similar matrices stacking the individual explanatory variables, and the dependent variable, respectively, and *W* is a positive semi-definite (*q* x *q*) weight matrix.[4]

[4] All the standard regularity conditions (see e.g. Hansen, 1982) are assumed to hold.

For panel data, potentially valid instruments can be found in the set of contemporaneous, lagged and lead values of f. We consider moment conditions of the form

[6a] $E(\tilde{f}^s_{ij,t+\tau}\tilde{v}^s_{ijt}) = 0$,

[6b] $E(f_{ij,t+\tau}\tilde{v}^s_{ijt}) = 0$,

for certain values of the integer τ. For OLS estimates of [2] or [3] to be consistent, we require [6a] to hold for $\tau = 0$. This will not be true if $\tilde{f}^s_{ijt}$ and $\tilde{v}^s_{ijt}$ are correlated, e.g. for reasons discussed above. To illustrate this, consider a case where f_{jt} is correlated with the contemporaneous residual v_{ijt}, but uncorrelated with lags and leads of the residual. It then follows that $\tilde{f}^s_{jt}$ will be correlated with v_{ijt} and $v_{ij,t-s}$, hence the moment conditions

[7a] $E(\tilde{f}^s_{ijt}\tilde{v}^s_{ijt}) = 0$,

[7b] $E(f_{ijt}\tilde{v}^s_{ijt}) = 0$, $E(f_{ij,t-s}\tilde{v}^s_{ijt}) = 0$,

will not hold. However [6b] contains some moment conditions that will hold, e.g. $E(f_{ij,t-(s+1)}\tilde{v}^s_{ijt}) = 0$ or, for $s > 1$, $E(f_{ij,t-(s-1)}\tilde{v}^s_{ijt}) = 0$. What moment conditions may hold in practice depends crucially on the time series properties of the residuals, and is an empirical question. It is possible that no lagged or lead values of f are valid instruments, in which case use of external instruments are required for consistent estimation.

Based on the model outlined above, we adopt a research strategy similar to that suggested by Griliches and Hausman (1986), p. 114. We begin by estimating [2] for s = 1, 2,..., S, where S is the longest difference available in the data, using OLS. If the estimates of different lengths differ significantly this suggests that measurement errors are present, particularly if the point estimates tend to increase in absolute magnitude with the length of differencing. To formally investigate whether the estimates differ, we estimate [3] imposing $\gamma_{12} = \gamma_{13} = ... = \gamma_{1T} = ... = \gamma_{ST}$, and test for the validity of this restriction. We then estimate [3] omitting the moment conditions in [7b], and using as instruments lagged and lead values of the explanatory variables. We impose $\gamma_{12} = ... = \gamma_{ST}$, and carry out the standard Sargan-Hansen test for the overidentifying restrictions. If the cross-equation restrictions are too strong, or if instruments are not valid, then some or all of the sample moments will be significantly different from zero, hence signalling misspecification.

The data which will be summarised in the next section is a panel of workers who were observed for a maximum of six years within a firm. We can therefore use the methodology we have summarised in this section we estimate equations of different orders of differencing and to test if the effect of firm size on earnings does differ depending on how long is the period of differencing of the equation. We can then proceed to use the methods outlined above to generate valid instruments. We turn next to describing the data.

3. Data

This study uses survey data on manufacturing firms in Ghana and Kenya, collected in face-to-face interviews with the firms' management.[5] The surveys used very similar survey instruments, enabling us to carry out comparisons across the countries. Four manufacturing sub-sectors were covered, namely food processing, textiles and garments, wood and furniture, and metal-working including machinery. These sub-sectors comprise the bulk of manufacturing employment in both countries. Four geographical areas in each country were surveyed, and large as well as small firms, including informal ones, were included in the sample. At the same time as the firms were surveyed a sample of workers and, where applicable, apprentices was chosen from each firm designed to cover the full range of personnel employed by the firms.[6] Hence we have matched employer-employee data. For Ghana we have data over six years, 1995-2000, while for Kenya we only have two rounds of data, covering 1995 and 2000.

---

[5] One advantage in using data from private manufacturing firms is that wages might better reflect productivity unlike the public sector firms where wages may be distorted. This provides a better setting to examine role of firm characteristics.

[6] The objective was to have up to 10 workers and 10 apprentices from each firm where firm size allowed. To increase the informational content of the data, the worker sample was stratified according to occupational status.

After deleting observations with missing values on key variables, a small number of gross outliers and apprentices, we obtain a sample of 4,695 observations on Ghanaian employees and 1,910 observations on Kenyan ones, for which we have complete information on wages and a set of firm-level variables.[7] The panel, which is unbalanced, is shown in Table A1 in Appendix 2. The table shows, for instance, that 70 of the 117 Ghanaian firms first observed in 1995 were observed in 2000; similarly, 52 of the 685 Ghanaian employees observed for the first time in 1995 were observed again in 2000. Hence there is significant attrition from the sample, both for firms and employees, so the sub-samples for which we can estimate the earnings equation using higher order differences are quite small. Indeed, fifth differences can only be taken for 52 Ghanaian employees and 78 Kenyan employees. These observations are possibly the most valuable ones however, as earnings and firm characteristics tend to change slowly.

In our data set the individual time series data will end whenever an individual leaves a firm. That is, the data set contains time series data solely on 'stayers', as distinct from 'movers'. While it is essential to use data on movers in some applications, e.g. in order to identify a tenure effect while controlling for individual fixed effects (see Topel, 1991), this is not the case in our context. In fact, we would argue that in order to analyse if a firm variable impacts causally on earnings, there is a case for not including data on movers even if such data were available. The reason is that the individual's productivity may differ across firms, and to the extent that this is unobservable and correlated with firm variables, the firm coefficients will be biased. Consider a case for instance where the quality of a 'match' between a firm and an individual varies across firms for a given individual. If matching quality is observed to the employer and remunerated accordingly, then moving from one firm to another will alter the wage if the individual's matching quality differs across the two firms. This, of course, is not a causal effect, but as matching quality typically is unobserved to the econometrician, this may bias the coefficients on the firm variables if these are correlated with the quality of the match.

[7] A data appendix that provides details on how the variables were constructed, is available on request from the authors.

It is reasonable to assume matching quality to be constant over time within firms, in which case this will be absorbed by the fixed effect for stayers.

Table 1 shows summary statistics for various variables that we will use in the empirical analysis, for the Ghanaian and Kenyan sub-samples. We identify two size categories: small, which is firms with up to 30 employees, and large those with 31 or more employees. A minor number of observations are incomplete in that information on firm age and the individual human capital variables is missing. This will only affect the OLS levels regressions, as none of these variables are used in the differenced regressions. In both countries, the average level of monthly earnings in small firms is about USD 50, while in large firms it is more than USD 100, suggesting substantial firm size differences in earnings. Looking only at production workers, the average wage in large firms is about 50 per cent higher than in small firms. It is perfectly possible that this earnings differential reflects differences in human capital over the size range, as the average years of education, tenure and age, are higher in large than in small firms. Looking at the firm variables it is clear that large firms tend to be more capital intensive, older and have a higher labour productivity, than small firms. Kenyan firms appear to have a higher capital intensity and labour productivity than their Ghanaian counterparts.

The central issue with which we are concerned is how the large dispersion of earnings across firms of different sizes documented in Table 1 is to be explained. To shed light on this we turn to regression analysis, where we intend to use the data in the following way. In the next section we carry out two sets of OLS regressions. First, we estimate the levels earnings equation [1] using OLS. This is based on the entire sample. Second, we estimate the differenced earnings equation [2] taking first, second, third, fourth and fifth differences of the data. As this requires multiple observations on each individual, individuals observed only once cannot be included. For Kenya, we can only obtain fifth differenced estimates. In section 5 we will implement the instrumental variable GMM estimator that combines different orders of differencing. As this estimator requires continuous series of observations over time, only the Ghana data are used in this part of the analysis.

4. Firm Characteristics as Determinants of Earnings: OLS results

A number of recent studies based on data on African manufacturing firms have shown that individual earnings are positively correlated with firm size (e.g. Valenchik, 1997; Strobl and Thornton, 2001; Manda, 2002) and that the size effect is larger than that found in developed country data sets (Brown and Medoff 1989, Troske 1999 and Bayard and Troske, 1999). While these analyses for African manufacturing firms do control for observed heterogeneity in individual human capital, typically measured by education, experience and age, none of them control for unobserved ability in the form of individual fixed effects. In view of the 'quality-of-labour' arguments summarised in Section 2 this is a potentially serious omission. While our data set does enable us to control for individual fixed effects we begin our empirical analysis by reporting earnings equations of the form often used in the literature, i.e. using OLS and controlling for observable human capital. This facilitates comparison with earlier studies and provides us with a benchmark.

Table 2 presents the results from estimating [1] for Ghana and Kenya, and separating out production workers from others, all under the assumption that any unobserved heterogeneity is captured by a residual uncorrelated with regressors. Except for Ghanaian production workers, education has a non-linear, convex, effect on earnings, manifesting itself through the significance of the squared term. The estimated returns to tenure are low, always less than one per cent, and in all cases insignificantly different from zero. The tenure variable is highly correlated with age so it is possible that these low tenure effects are partly driven by collinearity. There is some evidence that for both Ghana and Kenya, conditional on observable characteristics, women are paid less than men among production workers, which may reflect gender bias or productivity differentials. The time dummy is positive and statistically significant in all cases, indicating that real earnings have risen over the five year period.

We now focus on the role of firm variables in the earnings equation. Four results are noted. First, the coefficient on labour productivity is positive and significant in all cases. Second, firm age and the capital labour ratio are insignificant. Brown and Medoff (2001) also find that the higher wages that older firms pay may be fully explained by their workers characteristics. Third, the coefficient on average education in the firm ranges between 0.01 and 0.03 and has *t*-values around one, except in one case where it is significant at the five per cent level. Fourth, and most important for our purposes, the coefficient on firm size is positive and highly significant. For Kenya it is 0.10, while for Ghana the point estimate is 0.16, for all occupations. The point estimates on the size variable is reduced, particularly for Kenya, if the sample is confined to production workers. The size effect documented in Table 2 is not due to the omission of the other firm characteristics which it has been suggested the firm size variable might be proxying.

The manner of proceeding in Table 2 is similar to the method adopted by Troske (1999) and Bayard and Troske (1999). We have asked if firm characteristics, which may be related to the skills of the workforce, can explain the size effect. Like them we find they cannot. There remains the possibility that the size effect is due to the unobserved ability of workers or the unobserved characteristics of the workplace. Troske (1999, p. 25) notes that "one possible explanation that is consistent with the results reported in this paper is that large employers hire better workers and that both large employers and their employees are more likely to invest in firm-specific human capital". It is noted that between 47 and 70 per cent of the variation in earnings remains unexplained, and it seems very likely that a substantial share of this variation is due to unobserved factors rather than simply measurement errors in the earnings variable. We therefore continue by controlling for individual fixed effects by differencing the data. In doing so we now control for *all* the time invariant characteristics of the worker *and* the firm.

Table 3 reports the results from estimating differenced earnings equations. In Columns [1] – [5] we show for Ghana the estimates using first to fifth differences. In Column [6] we show the result for fifth differences of the Kenya data. In Column [7] we report a pooled estimate for the fifth differences of both Ghana and Kenya, and in Column [8] we include as an additional regressor a selectivity correction term based on a probit modelling attrition. The equation in the top part of the table simply reports the results of regressing the change of log earnings on the change of log employment. The second equation includes the same firm regressors as were used in Table 2. For reasons already discussed we believe the point estimates of the coefficients will rise as the orders of differencing increases. As we wish to compare Ghana and Kenya we focus in this section simply on the fifth difference on the grounds that this is where we expect measurement error bias to be least severe, and where a direct comparison is possible across the two countries. In the next section we will consider the additional uses to which the Ghana data can be put.

We consider first the equation which models the change in earnings as a function of the change in the log of employment for the fifth differences for Ghana and Kenya [Table 3, Columns (5) and (6)]. The point estimate for the coefficient on size is virtually identical for the two countries at 0.11 and 0.10 respectively. Neither is significant at the 10 per cent level reflecting the small sample size that we have for fifth differences. In Table 3 Column [7] the two countries are pooled (pooling is clearly accepted) and the estimated coefficient of 0.10 is now significant at the five per cent level. Remarkably, this estimated coefficient is only slightly below the simple average for the two countries from the Table 2 results.

We turn now to consider the general specification of the differenced equations reported in the bottom part of Table 3. We augment the differenced earnings function with the additional firm variables considered in Table 2, i.e. the average level of education in the firm, the log of the capital labour ratio, the log of output per employee, the proportion of managers in the firm, the proportion of supervisors in the firm, and profits per employee, all differenced.[8] The estimated coefficient on labour productivity is equal to 0.03, which is much smaller than in the regressions reported in Table 2, and insignificant, which is consistent with the notion that unobserved ability of individuals is positively correlated with productivity, Bayard and Troske (1999). That is, once we control for unobserved ability in the form of individual fixed effects the earnings-productivity relation vanishes. Average education in the firm, however, has a positive and significant effect on earnings. The point estimate is 0.03, indicating that a one-year increase in the average level of education in the firm is associated with a rise in individual earnings of about 3 per cent. Just as in the OLS regressions in Table 2, the inclusion of the additional variables does not reduce the coefficient on the size variable. In fact, the estimated size coefficient increases to 0.24, which appears to be driven by a relatively large, but imprecise estimate of the capital labour ratio coefficient. Equally striking is that for the Ghana only

[8] Obviously firm age cannot be included once we difference the equation as there will be no cross-section variation in the variable.

data reported in columns [1]-[4], with lower levels of differencing, the point estimate now varies from 0.10 to 0.26, and using either second or third differences is significant at the one per cent level. The results in Table 3 also show that the point estimate on employment is robust to the inclusion of a range of other firm level variables. It appears that the size effect is not being driven by either the unobservable skills of the workers or the unobserved characteristics of the work place, nor is it a proxy for a productivity or capital intensity effect.

We noted above that one of the disadvantages of the five year differenced model is that attrition will be significant. We now address that problem following the approach outlined in Section 2. We assume that the probability that an individual will not drop out is a function of education, years of tenure, age (allowing for a quadratic effect), gender, firm size, firm age, profit per employee, the log of the capital labour ratio, the log of output per employee and the average level of education in the firm. Experimentation with the data indicated that the some of the coefficients differed across the two countries, so we allow for country specific effects by interacting the country dummy with each of the regressors. The resulting probit regression (not reported) is highly significant, which is important in order to obtain precise estimates of the coefficients in the differenced earnings equation.[9] Based on this, Column [8] reports the selectivity corrected differenced earnings equation. The coefficient on the inverse Mill's ratio (lambda) is positive and significant, indicating that the attrition if ignored will lead to selectivity bias.[10] The result may imply that unobserved factor(s) that increase the probability of exiting the firm are also positively correlated with change in earnings. The estimated firm size coefficient is 0.12 in the bivariate specification, hence marginally higher than previously, and is now significant at the five per cent level. In the full specification the point estimate on employment is unaffected by the inclusion of the lamda term.

---

[9] Because we are fortunate to have numerous variables in our probit model that do not enter the second stage regression, the estimated Mill's ratio and the regressors in the earnings function are only weakly correlated. It is well known that when they are highly correlated, which typically happens when the exclusion restrictions are too few or inadequate, the parameter estimates in the selectivity corrected equation are likely to be very imprecise (see e.g. Leung and Yu, 1996).

[10] The estimate of the correlation coefficient ρ is about 0.60.

5. Simultaneity and Fixed Effects: GMM Results

So far we have followed a common path in the literature in investigating the firm-size earnings relationship. Like many others we have found that the size variable cannot be eliminated by controls for other aspects of the firm. In the last section we went further than is possible in most other studies and showed that the firm size effect survives if we allow for unobserved, time invariant, heterogeneity in the workers and the workplace. In this section we test whether we can identify an effect from size onto earnings while allowing the explanatory variables to be correlated with the residual, potentially as a result of measurement errors or endogeneity. Our method is to use the Ghana data where, as we have six years of annual data, we can create instruments by exploiting the moment conditions implied by the different orders of differencing the data as discussed in Section 2.

Table 4 provides our tests where we combine all the possible levels of differencing from our data set to allow for all individual and firm fixed characteristics and to provide us with a set of instruments. Because we have up to 6 time periods, our system consists of 15 equations across which we impose common coefficients. In Columns [1] and [2] we report one-step GMM results based on the moment conditions in [7], valid only if the explanatory variables are *not* correlated with the residual.[11] These specifications can hence be viewed as restricted versions of the models reported in Table 3, as the moment conditions are the same as in Table 3 but more cross-equation restrictions are imposed here.

Column [1] shows the results from a simple specification where earnings depends only on employment. The estimated employment coefficient is equal to 0.04, and insignificant at conventional levels. There is strong evidence that the model is misspecified, as indicated by the general Sargan-

[11] To obtain one-step results we follow the suggestion of Arellano and Bond (1991) and define the weight matrix W to reflect the correlation of the differenced residuals across equations, see Appendix 1, Part B. We focus on the one-step results in view of the well-known problem with the two-step estimator that the resulting asymptotic standard errors typically will be downward biased in finite samples, potentially giving rise to misleading inference (see Arellano and Bond, 1991). The two-step results, available on request from the authors, are very much the same as the one-step results except that the standard errors are rather much lower.

Hansen test. This is not surprising as we know from Table 3 that the point estimates vary considerably with the length of differencing, hence imposing a common coefficient across the 15 equations is bound to be invalid. As expected there is evidence that the results based on the first differenced equations are significantly different from those of longer differences.

Column [2] shows the full specification used in Table 3, again estimated under the assumption that explanatory variables are not correlated with the residual. The Sargan-Hansen specification test suggests that we can accept the specification, however the narrower test for pooling of first and higher order differenced equations indicates that we can reject pooling at the five per cent level. The estimated coefficient on the capital-labour ratio is 0.07, rather higher than what we obtained in the levels equations estimated by OLS (see Table 2). With a *t*-statistic of 1.56 the coefficient is not all that far from significant at the 10 per cent level. The point estimate of the employment coefficient increases to 0.11, but this is entirely driven by the inclusion of the capital-labour ratio. The marginal effect of employment, holding everything else constant, is still about 0.04 and insignificantly different from zero. Contrary to the finding in Table 2 the coefficient on profit per employee is positive and significant, hence it seems the inclusion of fixed effects reduces the magnitude of the downward bias somewhat. Average education is positive and quite close to significant at the 10 per cent level. Our proxy variables for the monitoring technology, i.e. the proportion of managers, and supervisors, of the total workforce, are both significant at the ten per cent level. The coefficient on supervisors, however, is positive which is at odds with the theoretical prediction that firms substitute more monitoring for higher wages in order to motivate their workers. The coefficient on the proportion of managers is negative.

Perhaps the main finding in Columns [1] and [2] is that a pooled specification based on the assumption that the regressors are uncorrelated with the residuals is rejected by the data. In particular there is evidence that coefficients are not stable across first and higher order differenced equations, which is consistent with explanatory variables being measured with errors, Griliches and Hausman (1986). In the remaining columns of the table all regressors are treated as endogenous, hence we do not use the moment conditions in [7a]-[7b] instead we use lagged and lead values as instruments. This procedure addresses not only the problem posed by measurement errors but also the endogeneity problem in general. The former problem is likely to be most serious for the capital stock series. Details of the instrument sets are shown in the table notes.

In Column [3] we take earnings to depend only on employment. The resulting coefficient on employment is equal to 0.17 and significant at the five per cent level. This estimate is higher than that reported in Column [1], which is to be expected if employment is measured with error or there is reverse causality in the form of an exogenous positive wage shock impacting negatively on employment. To shed light on the role of additional variables, we report the full specification in Column [4]. Compared to previous models we obtain a dramatic increase in the point estimate of the coefficient on the capital-labour ratio, now equal to 0.24 and significant at the five per cent level. The estimated coefficient on employment is 0.36, and, given the coefficients on the capital-labour ratio and output per employee, the marginal effect of employment is equal to (0.36 – 0.24 + 0.01) = 0.12, which is lower than in Column [3], and significant at the ten per cent level (test not reported). We note that the coefficient on output per employee is close to zero and far from significant.

Compared to Column [2] we obtain an increase in the coefficient on profit per employee, now significant at the one per cent level. The point estimate of 0.03 corresponds to an elasticity of wages with respect to profit per employee of about 0.04, evaluated at the sample mean of profits per employee. This is somewhat lower than what has been found in studies of the US labour market, Blanchflower et al. (1996), and much lower than what Teal (1996) reports for Ghana 1991-93 based on instrumented regressions. The coefficient on the proportion of managers is negative and significant, consistent with efficiency wage theory. The point estimate of –0.71 is interpretable as a semi-elasticity of wages with respect to the proportion of managers: an increase by 0.01 in the manager-employee ratio is expected to decrease wages by 0.71 per cent. The coefficients on average education and the proportion of supervisors are both positive but insignificant at conventional levels. Finally, both the general Sargan-Hansen test and the test for pooling of first and higher order differenced equations indicate that we can accept the model specification.[12]

[12] It is well known that the Sargan-Hansen has low power when the number of overidentifying restrictions is high (Bowsher (2000)). This is reflected in Column [4] by the p-value tending to unity. We can distinguish

In Column [5] we report a parsimonious version of the general specification, where we exclude output per employee, education and the proportion of supervisors from the model on the grounds that the associated coefficients are insignificant in Column [4]. The instrument set is unchanged. The results do not change much. The estimated capital coefficient increases marginally to 0.27, and is now significant at the one per cent level. Employment has an estimated coefficient equal to 0.39, implying a marginal effect of 0.12 which is significant at the ten per cent level (test not reported). The coefficient associated with the proportion of managers is somewhat closer to zero than previously, and is no longer significant at the ten per cent level. There is no change in the coefficient on profit per employee, which is still significant at the one per cent level. In Column [6] we investigate if the results are robust to an alternative instrument set in which output per employee, education and the proportion of supervisors have been excluded. The coefficients on employment, capital and profit per employee are identical to those shown in Column [5], however the manager ratio has a smaller

---

between two sources of restrictions imposed on the model: on the one hand the cross-equation restrictions; on the other hand the exclusion restrictions imposed on the instruments i.e. that lagged and lead values of the explanatory variables do not enter the structural equation. The validity of all these restrictions are tested for at the same time by the general test, and because the total number of restrictions is high the power of the test is likely to be low. Our test for pooling of first and higher order differenced equations, which has only seven degrees of freedom and is thus unlikely to suffer from low power, suggests that the cross-equation restrictions are not overly restrictive (the p-value is 0.75). In Column [4] 98 of the overidentifying restrictions result from the cross-equation restrictions. To assess whether the exclusion restrictions imposed on the instruments are valid, consider generalising the model so that no cross-equation restrictions are imposed. In this case there would be 203 – 98 = 105 overidentifying restrictions all of which result from the exclusion restrictions imposed on the instruments. Clearly an upper bound on the Sargan-Hansen J-statistic in such a model would be the value reported in Column [4], i.e. 96.85. With 105 degrees of freedom we would still comfortably accept the validity of the overidentifying restrictions. The second test for overidentifying restrictions shown in Table 4 (indicated with a superscript (b)) reports the Sargan-Hansen J-statistic obtained in a model where 70 of the 98 cross-equation restrictions are relaxed. For the specification in Column [4] the J-statistic decreases to 37.6, which is well below the critical value at any level of significance with 133 degrees of freedom. This suggests that the exclusion restrictions imposed on the instruments are valid.

effect than previously and the *t*-value is just above one so the coefficient is not significant. There is no evidence from the specification tests that the models in Columns [4] and [5] are misspecified.

We next investigate the possibility that our results are being biased by attrition. In Table 3 we obtained a significant coefficient on the selectivity variable, but found that this had a very small effect on the point estimates of interest. Because attrition bias technically is a form of omitted variable bias, we would expect our instrumental variable results shown in Table 4 to be robust to such problems. To assess whether this indeed is the case, we take a closer look at the equations that can only be estimated for the arguably atypical sub-sample for which we have a full set of six observations over time. Specifically, we carry out Wald tests of the null hypothesis that the coefficients in these equations are the same as the coefficients in the other equations. Test results are shown at the bottom of Table 4. In no case can we reject common coefficients across these equations. We conclude that attrition bias is not a significant problem.

The only dimension of skills for which we have not so far controlled is unobserved time varying worker quality. While possible in principle, it seems to us rather unlikely such an effect is driving our results. Even if the changes in unobserved quality were correlated with changes in employment, capital and profits, we would expect our instruments to correct for the resulting simultaneity. If the instruments failed to do so, we would expect our tests to indicate that the model is misspecified. There is no evidence that this is the case.

The main finding documented in Table 4 is that both size and profits per employee appear to have a *causal* effect on earnings. There is some, but weaker, evidence that the monitoring technology as modelled by the proportion of mangers in the firm also affect earnings. Of course all these variables are either choice variables (employment, capital and the number of managers) or functions of choice variables (profits), so the implication of our results is that a change in an underlying (exogenous) factor affecting these endogenous variables will feed into a causal effect on earnings. That is, if a firm (for one reason or another) chooses to increase, say, its capital stock by one per cent, then this will increase individual earnings by about 0.25 per cent, everything else equal. This result contrasts significantly with the interpretation of the size-wage relation in the human capital model, namely that workers in large firms are paid more because they are more skilled. In the next section we discuss whether efficiency wage and bargaining models can be formulated to be consistent with these empirical results.

6. Models of Non-Competitive Factors Markets and Firm Size

We have argued that the firm size effect is not due to the unobserved quality of the workers nor is it proxying aspects of firm performance such as productivity, the capital labour ratio or firm profitability. We now consider if the theories developed to investigate non-competitive labour markets can predict a causal relationship from size and profits per employee onto wages. Our model draws on two broad classes of theories, the first those which have modelled labour market outcomes as the result of bargaining between firms and their employees (Manning, 1987; Blanchflower et al., 1996; van Reenen, 1996), the second being the efficiency wage theory predicting that firms choose higher wages as a means of motivating their employees. To illustrate the model, define net profits as

$$\pi = AF(K, eL) - wL - rK,$$

where A is total factor productivity, F is the production function, K is physical capital, e is labour effort, L is labour, w is the unit price of labour and r is the unit price of capital. The firm and the employees bargain over w and L such that the solution is obtained by maximising Ω:

$$[8] \qquad \Omega = \max_{L,K,w} \phi \log L(w - \overline{w}) + (1-\phi)\log \pi,$$

where ϕ is the relative bargaining power of the employees. Provided that workers have some bargaining power, i.e. $\phi > 0$, the first order condition with respect to w can be written

$$[9] \qquad w = \overline{w} + \frac{\phi}{1-\phi} \cdot \frac{\pi^G - rK}{-\pi_w},$$

where $\pi^G = AF(K, eL) - wL$ is gross profit, and π_w is the partial derivative of π with respect to w.[13]

Efficiency wages implies that w will impact positively on labour effort, hence

$$-\pi_w = L - AF_{eL}Le_w \equiv L(1-g).$$

[13] If the employees have no bargaining power, so that $\phi = 0$, then the optimal wage will satisfy $\pi_w = 0$.

Substituting this into [9] and linearising the resulting equation by taking a first order Taylor expansion about $g = \overline{g}$ yields

$$w = \overline{w} + \frac{\phi}{1-\phi} \cdot \left(\frac{1}{1-\overline{g}} \frac{\pi^G - rK}{L} + \frac{\pi/L}{(1-\overline{g})^2} \left(AF_{eL} e_w - \overline{g} \right) \right).$$

We can look at the link between the relevant variables and wages by computing log differentials:

$$\text{[10]} \qquad d\log w = \frac{\overline{w}}{w} d\log \overline{w} + \frac{\phi}{w(1-\phi)} \left\{ \Gamma \left(d\frac{\pi_G}{L} - r\frac{K}{L} \left(d\log r + d\log K - d\log L \right) \right) + \frac{\pi/L}{(1-\overline{g})^2} dg \right\}$$

where

$$\Gamma = \frac{1}{1-\overline{g}} + \frac{1}{(1-\overline{g})^2} (g - \overline{g}),$$

$$dg = \frac{1}{2} \frac{\alpha_L}{e} \left(e_w A \frac{F}{L} \left(d\log A + \alpha_K d\log K + (\alpha_L - 1) d\log L \right) + A\frac{F}{L} de_w \right),$$

and α_K and α_L are the output elasticities of capital and labour, respectively.

Interpreting [10] is not as straightforward as it seems due to the presence of endogenous variables. We argued above that there is a causal effect from endogenous variables onto wages, ultimately driven by changes in exogenous factors. We now consider the role of two such exogenous variables in the wage determination process, namely total factor productivity, *A*, and the unit price of capital, *r*. It is often argued that these factors are key determinants of the performance of the private sector in developing countries.[14] Further, there is typically considerable heterogeneity across firms with respect to variables potentially related to *A* and *r*, e.g. gross profits per employee, labour productivity and capital intensity (see Table 1). We now ask if variation in *A* or *r* will feed into changes in the endogenous variables of the model that impact on wages, within the framework of the bargaining cum efficiency wage model outlined above. Even under strongly simplifying structural assumptions it is difficult to solve the model analytically, and we therefore use numerical analysis.

---

[14] Pack (1993), for example, stresses the importance of technical capacity in understanding differential firm performance.

Table 5 shows how changes in r and A impact on the endogenous variables of interest, in five situations. In each case we consider the effect of changing log r and log A by 0.01, and we assume throughout the production function to be Cobb-Douglas, with $\alpha_K = 0.2$ and $\alpha_L = 0.6$.[15] Remaining parameter values are listed in the table notes. Model [1] assumes a relative bargaining power of workers equal to 0.4 and that effort is given by $e = ((w - \overline{w})/1.6)^{1/2}$, which follows Sparks (1986).[16] While an increase in productivity or a decrease in the price of capital increases the size of the firm, there is no effect on wages in this model. Hence this form of the non-competitive model, which we can take as a benchmark case, does not predict the relationship we observe between firm size and earnings.[17]

In Models [2] and [3] we adopt a different effort function, derived by Ringuede (1998). Ringuede assumes that the probability that a worker will be monitored can be written (b/L), $b < L$, where b is a monitoring efficiency parameter indicating the number of controls that the firm can make at a given point in time. Hence workers in large firms are less likely to be monitored than workers in small firms (cf. Bulow and Summers, 1986).[18] The effort function, derived from the solution to the employee's utility maximisation problem, is $e = (b(w - \overline{w})/(0.6L + b))^{1/2}$. To maintain a given level of effort as size increases the firm needs to increase the wage, holding b constant. In Model [2] workers

---

[15] It is clear from [10] that wages potentially depend on the form of the production function. We adopt the Cobb-Douglas form here because previous research has shown that this model adequately approximates the nature of the technology in Ghanaian manufacturing (Söderbom and Teal, 2002).

[16] The effort function in Sparks is $e = ((w - \overline{w})/(2i + 1))^{1/2}$, where i is the discount rate of the representative employee. Hence 1.6 in the denominator corresponds to a discount rate of 30 per cent.

[17] It is well known that when the production function is Cobb-Douglas, revenues per employee will not depend on either r or A (MacDonald and Solow, 1981; Abowd and Lemieux, 1993), hence for A to impact on wages, such an effect will have to be transmitted through a changing capital-labour ratio. Similarly, for r to affect wages, $d\log(K/L)/d\log r \neq -1$ is required.

[18] The model can also be derived under the assumption that the firm can alter b, for instance by employing more supervisors, see Fafchamps and Söderbom (2002).

have no bargaining power, $\phi = 0$, while in Model [3] we set $\phi = 0.4$. In contrast to the benchmark model in [1], changes in capital, labour and profit per employee, driven by changes *r* and *A*, now transmit into changes of wages. In the model without bargaining the elasticity of wages with respect to the capital-labour ratio is approximately 0.09. Introducing bargaining increases the elasticity to about 0.13. This form of the model predicts a relationship between size and earnings similar to that we observe in the data.

Finally in Models [4] and [5] we look at the role of credit constraints. To keep the analysis simple we assume that the firm cannot incur capital expenditures in excess of some constant θ, so that the maximisation problem [8] is subject to the inequality constraint $rK \leq \theta$. Model [4] is identical to Model [1] except for the introduction of the credit constraint, which is assumed to bind.[19] The results show that shocks to *r* and *A* do transmit into changes in wages. It is noted that a decrease in *r* results in higher capital intensity, while an increase in *A* leads to lower capital intensity. A similar result is obtained without efficiency wages. In Model [5] we use Ringuede's (1998) efficiency wage model, where effort depends on size. As expected the wage effects are larger in this model. When *r* changes, the resulting elasticity of the wage with respect to the capital-labour ratio is about 0.20.

7. Conclusions

In this paper we have used data from the manufacturing sectors in Ghana and Kenya to investigate if firm size affects individual earnings, while controlling for unobserved labour quality in the form of employee fixed effects as well as a number of firm characteristics. We begin by noting that the results are entirely consistent with the human capital model in that skills are rewarded in the labour market and the fixed effects, interpretable as time invariant skills, are almost always significant in the regression reported in Table 2 (tests not reported). The results are not, however, consistent with the size effect observed in the cross-section reflecting unobserved ability of the worker or unobserved characteristics of the firm. The empirical results in Table 3 indicate that firm characteristics in both Kenya and Ghana are important determinants of earnings, when there are controls for fixed effects.

[19] That is, the Lagrangian multiplier associated with the inequality constraint is strictly larger than zero.

For the Ghana data we can go one step further and control for the endogeneity of the firm factors affecting earnings. The results suggest that firm size, profits per employee and (possibly) the proportion of managers causally determine earnings.

We have developed a model of non-competitive labour markets containing elements drawn from both bargaining and efficiency wage models. We have used numerical simulation to show that in such a model there is a relationship from firm size to earnings. Further if capital constraints are imposed in such a model then the firm-size wage effect increases. Our data is drawn from sub-Saharan Africa where such phenomenon are most likely to be found. Whatever the interpretation of the effect of size on wages that is advanced we would argue that it is an important determinant of wages and cannot be explained by time-invariant characteristics of either the worker or firm, as has been the presumption based on the competitive factor market model which underlies the usual interpretation of earnings functions.

References

Abowd, J.A. and T. Lemieux (1993) "The effects of product market competition on collective bargaining agreements: the case of foreign competition in Canada, *Quarterly Journal of Economics*, 108, 4, November, pp. 983-1014.

Akerlof, G. and J. Yellen (1986). *Efficiency Wage Models of the Labour Market*. Cambridge: Cambridge University Press.

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

Azam, P. and C. Ris (2001) "Rent sharing, hold up and manufacturing sector wages in Cote d'Ivoire," presented at the CSAE and UNIDO International Forum *New Industrial Realities and Firm Behaviour in Africa*, held at St Edmund Hall, Oxford, 22-23 September 2001.

Bayard, K. and K. R. Troske (1999). "Examining the Employer-Size Wage Premium in the Manufacturing, Retail Trade, and Service Industries Using Employer-Employee Matched Data," *American Economic Review* 89: 99-103.

Bigsten, A., P. Collier, S. Dercon, B. Gauthier, J.W. gunning, A. isaksson, A. Oduro, R. Oostendorp, C. Pattilo, M. Söderbom, M. Sylvain, F. Teal and A. Zeufack (1999) "Investment in Africa's manufacturing sector: a four country panel analysis", *Oxford Bulletin of Economics and Statistics*, 61, 4, :489-512.

Blanchflower, D. G., A. J. Oswald and P. Sanfey (1996) "Wages, profits and rent-sharing," *The Quarterly Journal of Economics* 111: 227-251.

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

Bowsher, C.G. (2000) "On testing overidentifying restrictions in dynamic panel data models", mimeo, Nuffield College, University of Oxford.

Brown, C. and Medoff, J. (1989) "The employer size-wage effect," *The Journal of Political Economy* 97: 1027-1059.

Brown, C. and Medoff, J. (2001) “Firm age and wages.” Working Paper 8552, National Bureau of Economic Research.

Bulow, J. and L. Summers (1986), “A Theory of Dual Labor Markets, with Applications to Industrial Policy, Discrimination, and Keynesian Unemployment,” *Journal of Labor Economics* 4: 376-414.

Denny, K. and S. J. Nickell, (1992). “Unions and Investment in British Industry,” *The Economic Journal* 102: 874-887.

Dickens W. T. and L. F. Katz (1987), “Inter-industry wage differences and industry characteristics,” in K. Lang and J. Leonard (eds.) *Unemployment and the Structure of the Labour Markets*. Oxford, UK: Basil Blackwell.

Doeringer, P. and M. Piore (1971), *Internal Labour Markets and Manpower Analysis*. D. C. Heath, Lexington.

Dunne, T. and J. Schmitz Jr. (1995) “Wages employment structure, and employer size wage premia: their relationship to advanced technology usage at U.S. manufacturing establishments”, *Economica* 62: 89-105.

Fafchamps, M. and M. Söderbom (2002) “Wages and worker supervision in African manufacturing”, Mimeo. Oxford: Centre for the Study of African Economies, Department of Economics, University of Oxford.

Griliches, Z. and J.A. Hausman (1986) “Errors in variables in panel data”, *Journal of Econometrics,* 31, pp.93-118.

Hamermesh, D. S. (1980) “Commentary,” in J. J. Siegfried (ed.) *The Economics of Firm Size, Market Structure and Social Performance.* Washington D.C.: Federal Trade Commission.

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

Hansen, L. P. (1982), “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica* 50: 1029-1054.

Heckman, J. (1976) "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement* 5: 475-592.

Huang Tzu-Ling, A. Hallam, P.F.Orazem and E.M. Paterno (1998) "Empirical tests of efficiency wages", *Economica*, 65, pp. 125-43.

Idson, T. L. and D. J. Feaster (1990) "A Selectivity Model of Employer-Size Wage Differentials," *Journal of Labor Economics* 8: 99-122.

Katz, L. R. and L. H. Summers (1989) "Industry rents: evidence and implications," *Brookings Papers on Economic Activity (Microeconomics)*, 209-275.

Kremer, M. (1993) "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics* 108: 551-75.

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

Krueger, A. B. and L. H. Summers (1987) "Reflections on the inter-industry wage structure," in K. Lang and J. Leonard (eds.) *Unemployment and the Structure of the Labour Markets*. Oxford, UK: Basil Blackwell.

Leung, S. F., and S. Yu (1996) "On the Choice Between Sample Selection and Two-part Models," *Journal of Econometrics* 72: 197-229.

Levine, D. (1992) "Can wage increases pay for themselves? Tests with a production function", *The Economic Journal*, 102 (September), 1102-1115.

MacDonald, I.M. and R.M. Solow (1981) "Wage bargaining and employment", *American Economic Review*, LXXI: 886-908.

Manning, A. (1987) "An integration of trade union models in a sequential bargaining framework", *The Economic Journal*, 97, March, pp. 121-139.

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.

Mazumdar, D. (1983) "The rural-urban wage gap migration and the working of urban labour markets: an interpretation based on a study of the workers in Bombay city", *Indian Economic Review*, Vol 18, No. 2.

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

Masters, S. H. (1969) "Wages and Plant Size: an Interindustry Analysis," *The Review of Economics and Statistics* 51: 341-345.

Mellow, W. (1982) "Employer size and wages," *The Review of Economics and Statistics* 64: 495-501.

Moll, P. G. (1993) "Industry wage differentials and efficiency wages: a dissenting view with South African evidence", *Journal of Development Economics*, Vol. 41, pp. 213-46.

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

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

Raff, D. and L. Summers (1987) "Did Henry Ford pay efficiency wages?" Journal of Labor Economics", vol. 5, part 2, pp. 557-86.

Ringuede, S. (1998) "An efficiency wage model for small firms: firm size and wages", *Economic Letters*, 59: 263-8.

Shapiro, C. and J.E. Stiglitz (1984) "Equilibrium unemployment as a worker discipline device", American Economic Review, 74: 433-444.

Slichter, S. (1950) "Notes on the structures of wages," *Review of Economics and Statistics* 32: 80-91.

Söderbom, M. and F. Teal (2002). "Size and efficiency in African manufacturing firms: evidence from firm-level panel data". Mimeo. Oxford: Centre for the Study of African Economies, Department of Economics, University of Oxford.

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

Teal, F. (1996) "The size and sources of economic rents in a developing country manufacturing labour market," *The Economic Journal* 106: 963-976.

Topel, Robert H (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," Journal of Political Economy. 99(1): 145-76.

Troske, K. R. (1999). "Evidence on the Employer Size-Wage Premium From Worker-Establishment Matched Data," *Review of Economics and Statistics* 81: 15-26.

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

Van Reenen, J. (1996) "The creation and capture of rents: wages and innovation in a panel of U.K. companies", *The Quarterly Journal of Economics* 111: 195-225.

Wadhwani, S. B. and M. Wall (1991) "A direct test of the efficiency wage model using UK micro data", *Oxford Economic Papers,* 43:529-548.

Weiss, A. (1980) "Job queues and layoffs in labour markets with flexible wages," *The Journal of Political Economy* 88: 526-538.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Massachusetts and London, England: MIT Press.

TABLE 1: SUMMARY STATISTICS

	Small firms (employment ≤ 30)				Large firms (employment > 30)			
	N	Mean	Median	Std. Dev	N	Mean	Median	Std. Dev
A. GHANA								
Earnings (all occupations)	1294	53.48	44.06	42.4	3401	121.68	77.57	142.3
Earnings (production workers)	571	41.24	37.85	21.3	1185	70.14	57.58	52.7
Employment	1294	16.47	16.00	7.3	3401	177.55	93.00	224.6
Log [Capital / Labour]	1294	7.36	7.40	1.7	3401	8.60	8.74	1.3
Log [Output / Labour]	1294	8.19	8.27	0.9	3401	8.72	8.69	1.1
Average education in firm	1294	10.04	10.39	2.4	3401	11.00	11.21	2.1
Proportion managers	1294	0.04	0.00	0.1	3401	0.03	0.03	0.0
Proportion supervisors	1294	0.05	0.00	0.1	3401	0.05	0.04	0.0
Profit per employee / 1000	1294	0.65	0.38	1.1	3401	1.32	0.86	1.9
Firm age	1291	18.87	19.00	11.6	3155	21.84	19.00	13.4
Years of education	1291	10.04	10.00	4.3	3155	12.15	11.00	4.1
Age	1291	33.91	31.00	11.7	3155	38.54	37.00	10.9
Years of tenure / 10	1291	0.70	0.40	0.8	3155	0.83	0.60	0.8
Male proportion	1291	0.76	1.00	0.4	3155	0.87	1.00	0.3
B. KENYA								
Earnings (all occupations)	664	54.16	45.57	40.6	1246	106.61	64.25	127.7
Earnings (production workers)	468	51.22	45.57	29.0	842	76.20	57.02	59.9
Employment	664	13.34	12.00	8.1	1246	216.40	90.00	355.5
Log [Capital / Labour]	664	8.23	8.55	1.7	1246	9.40	9.53	1.1
Log [Output / Labour]	664	8.51	8.55	1.1	1246	9.32	9.26	1.0
Average education in firm	664	8.38	8.37	1.4	1246	9.56	9.46	1.6
Proportion managers	664	0.17	0.14	0.1	1246	0.06	0.05	0.0
Proportion supervisors	664	0.03	0.00	0.1	1246	0.04	0.03	0.0
Profit per employee / 1000	664	1.48	0.59	3.5	1246	2.72	1.92	3.4
Firm age	629	21.16	20.00	15.4	1157	24.25	22.00	13.1
Years of education	629	8.95	8.00	2.8	1157	10.40	11.00	2.8
Age	629	32.22	30.00	10.1	1157	34.99	34.00	8.8
Years of tenure / 10	629	0.68	0.40	0.7	1157	0.86	0.60	0.7
Male proportion	629	0.85	1.00	0.4	1157	0.83	1.00	0.4

Note: All financial variables are measured in USD.

TABLE 2: OLS EARNINGS FUNCTION ESTIMATES WITH FIRM VARIABLES

	A. All Occupations		B. Production Workers	
	[1] Ghana	[2] Kenya	[3] Ghana	[4] Kenya
Years of Education	0.001 (0.13)	-0.10 (4.84)**	0.01 (0.53)	-0.04 (1.74)
Education ^{2} / 100	0.27 (5.51)**	0.97 (7.57)**	0.04 (0.57)	0.41 (3.05)**
Years of Tenure / 10	0.02 (0.69)	-0.01 (0.36)	0.04 (1.33)	0.03 (1.06)
Age (years)	0.04 (4.41)**	0.03 (3.09)**	0.02 (2.96)**	0.02 (1.98)*
Age^2 / 100	-0.03 (2.42)*	-0.02 (1.50)	-0.02 (2.07)*	-0.01 (0.97)
Male	0.13 (2.71)**	0.08 (1.61)	0.15 (3.11)**	0.17 (2.86)**
Ln Employment	0.16 (7.16)**	0.10 (4.09)**	0.14 (4.78)**	0.05 (2.21)*
Ln (Capital / Employment)	-0.02 (1.00)	-0.02 (0.82)	0.02 (0.98)	-0.02 (1.13)
Ln (Output / Employment)	0.17 (6.47)**	0.07 (2.51)*	0.13 (4.08)**	0.07 (2.28)*
Ln Firm Age	-0.01 (0.22)	0.05 (1.26)	-0.02 (0.37)	0.02 (0.48)
Average Education in Firm	0.02 (1.64)	0.01 (0.69)	0.03 (2.11)*	0.01 (1.00)
Proportion Managers	-0.25 (0.59)	0.44 (1.76)	-1.22 (2.26)*	0.01 (0.03)
Proportion Supervisors	-0.29 (0.74)	-0.24 (0.57)	-0.29 (0.59)	0.07 (0.15)
Profit per Employee / 1000	-0.007 (0.46)	0.005 (0.60)	-0.01 (0.66)	0.002 (0.26)
Marginal return of education at education = 6	0.03	0.02	0.01	0.01
Marginal return of education at education = 12	0.07	0.13	0.02	0.06
R^2	0.53	0.41	0.47	0.30
Number of Observations	4446	1786	1671	1214

Note: The dependent variable is the logarithm of monthly earnings, expressed in USD. Dummy variables for sector, time and location are included in all regressions. The numbers in ( ) are *t*-statistics based on standard errors robust to heteroskedasticity. Significance at the 1 per cent, 5 per cent and 10 per cent level is indicated by *, ** and $^+$ respectively.

TABLE 3: DIFFERENCED EARNINGS EQUATIONS

	[1] Ghana	[2] Ghana	[3] Ghana	[4] Ghana	[5] Ghana	[6] Kenya	[7] Pooled	[8] Pooled
Order of differencing:	1	2	3	4	5	5	5	5
					A. BIVARIATE SPECIFICATION			
Δ Ln Employment	-0.02	0.10	0.17	0.08	0.10	0.11	0.10	0.12
	(1.04)	(2.76)**	(3.74)**	(1.00)	(1.06)	(1.73)$^{+}$	(1.96)$^{+}$	(2.26)*
Kenya							0.12	0.14
							(2.01)*	(2.39)*
Lambda								0.18
								(1.67)$^{+}$
R-squared	0.19	0.07	0.09	0.06	0.02	0.04	0.06	
					B. FULL SPECIFICATION			
Δ Ln Employment	0.10	0.20	0.27	0.15	0.04	0.23	0.25	0.25
	(1.24)	(2.25)*	(3.44)**	(0.78)	(0.24)	(3.01)**	(2.86)**	(3.01)**
Δ Ln (Capital / Employment)	0.12	0.10	0.12	-0.04	-0.18	0.20	0.12	0.11
	(1.87)$^{+}$	(1.37)	(1.65)$^{+}$	(0.18)	(1.14)	(4.04)**	(1.97)*	(1.85)$^{+}$
Δ Ln (Output / Employment)	0.01	-0.03	-0.06	0.12	0.21	-0.07	0.04	0.02
	(0.43)	(0.89)	(1.58)	(1.20)	(2.08)*	(1.92)$^{+}$	(0.78)	(0.52)
Δ Average Education in Firm	0.00	0.01	0.01	0.03	0.05	0.05	0.03	0.04
	(0.48)	(0.58)	(0.47)	(1.52)	(3.32)**	(1.82)$^{+}$	(2.25)*	(2.51)*
Δ Proportion Managers	-0.73	-0.47	-0.62	-0.61	-0.22	0.39	0.50	0.51
	(1.82)$^{+}$	(1.21)	(1.75)$^{+}$	(0.35)	(0.10)	(0.69)	(0.86)	(0.91)
Δ Proportion Supervisors	0.48	0.19	0.20	-0.22	0.34	-0.66	0.09	0.10
	(1.53)	(0.58)	(0.74)	(0.20)	(0.34)	(1.33)	(0.15)	(0.19)
Δ Profit per Employee / 1000	0.01	0.03	0.03	0.02	-0.03	0.01	0.01	0.01
	(1.37)	(2.58)*	(2.54)*	(0.39)	(0.47)	(1.68)$^{+}$	(0.53)	(0.76)
Kenya							0.14	0.17
							(2.39)*	(2.81)**
Lambda								0.20
								(1.86)$^{+}$
R-squared	0.20	0.09	0.12	0.13	0.21	0.18	0.13	
Observations	2493	845	459	105	52	78	130	130

Note: The dependent variable is change in logarithm of earnings. The numbers in ( ) are absolute values of *t*-statistics. Significance at the 1 per cent, 5 per cent and 10 per cent level is indicated by **, * and $^{+}$ respectively.

33

TABLE 4: GMM ESTIMATES OF DIFFERENCED EARNINGS EQUATIONS

	[1]	[2]	[3]	[4]	[5]	[6]
	No allowance for endogeneity		Allowing for endogeneity			
Δ Ln Employment	0.04	0.11	0.17	0.36	0.39	0.39
	(1.03)	(1.98)$^{+}$	(2.38)*	(3.84)**	(3.98)**	(3.87)**
Δ Ln (Capital / Employment)		0.07		0.24	0.27	0.27
		(1.56)		(2.43)*	(2.72)**	(2.61)**
Δ Ln (Output / Employment)		0.01		-0.01		
		(0.48)		(0.24)		
Δ Average Education in Firm		0.01		0.01		
		(1.56)		(1.20)		
Δ Proportion Managers		-0.54		-0.71	-0.64	-0.47
		(1.68)$^{+}$		(1.77)$^{+}$	(1.55)	(1.09)
Δ Proportion Supervisors		0.47		0.21		
		(2.01)*		(0.74)		
Δ Profit per Employee / 1000		0.01		0.03	0.03	0.03
		(2.33)*		(3.25)**	(3.17)**	(2.76)**
SPECIFICATION TESTS						
Overidentifying restrictions[(a)]: *J*-value	36.50	107.22	26.19	96.85	95.27	66.48
p-value	0.00	0.25	0.62	1.00	1.00	1.00
Number of overidentifying restrictions	14	98	29	203	206	116
Overidentifying restrictions[(b)]: *J*-value	4.76	31.31	19.94	37.58	62.63	20.13
p-value	4	0.30	0.40	1.00	1.00	1.00
Number of overidentifying restrictions	0.31	28	19	133	166	76
First & higher order differences pool: *p*-value	0.00	0.04	0.53	0.75	0.90	0.92
Attrition: *p*-value	0.72	0.24	0.40	0.84	0.91	0.74

Tables notes on following page.

34

TABLE 4: GMM ESTIMATES OF COMBINED DIFFERENCED EARNINGS EQUATIONS

Note: The dependent variable is change in logarithm of earnings. The numbers in ( ) are absolute values of t-statistics. Significance at the 1 per cent, 5 per cent and 10 per cent level is indicated by **, * and + respectively. Time dummies are included in all equations and in all regressions.

The instrument set for equation $\{s, t\}$ in [1]-[2] is $\tilde{f}^{s}_{ijt}$ and time dummies.

The instrument set for [3]-[6] is as follows:

$s = 1$ (first differences): $f_{ij,t-2}$, $f_{ij,t-3}$ and time dummies.

$s = 2$ (second differences): $f_{ij,t-1}$, $f_{ij,t-3}$ and time dummies.

$s = 3$ (third differences): $f_{ij,t-1}$, $f_{ij,t-2}$ and time dummies.

$s = 4$ (fourth differences): $f_{ij,t-1}$, $f_{ij,t-2}$, $f_{ij,t-3}$ and time dummies.

$s = 5$ (fifth differences): $f_{ij,t-1}$, $f_{ij,t-2}$, $f_{ij,t-3}$, $f_{ij,t-4}$ and time dummies.

TABLE 5:

EFFECTS OF CHANGES IN UNIT PRICE OF CAPITAL AND TOTAL FACTOR PRODUCTIVITY: NUMERICAL ANALYSIS

		$\Delta\ln K$ x 100	$\Delta\ln L$ x 100	$\Delta\ln K/L$ x 100	$\Delta(\pi/L)$ x 100	$\Delta\ln w$ x 100
[1]	$\Delta\ln r$ = -0.01	2.00	1.00	1.00	0.00	0.00
	$\Delta\ln A$ = 0.01	5.00	5.00	0.00	0.00	0.00
[2]	$\Delta\ln r$ = -0.01	1.51	0.34	1.16	0.53	0.11
	$\Delta\ln A$ = 0.01	2.52	1.70	0.83	2.68	0.55
[3]	$\Delta\ln r$ = -0.01	1.51	0.33	1.18	0.38	0.15
	$\Delta\ln A$ = 0.01	2.53	1.65	0.88	1.89	0.75
[4]	$\Delta\ln r$ = -0.01	1.00	0.53	0.47	0.08	0.02
	$\Delta\ln A$ = 0.01	0.00	2.65	-2.65	0.42	0.12
[5]	$\Delta\ln r$ = -0.01	1.00	0.24	0.76	0.52	0.15
	$\Delta\ln A$ = 0.01	0.00	1.19	-1.19	2.61	0.72

[1]: $e = \left((w-\overline{w})/1.6\right)^{1/2}$, $\phi = 0.4$, $r = 0.15$, $A = 4$, $\overline{w} = 1$.
Solution: $K = 34.6$, $L = 7.79$, $\pi/L = 0.40$, $w = 2.27$.

[2]: $e = \left(b(w-\overline{w})/(0.6L+b)\right)^{1/2}$, $\phi = 0$, $r = 0.15$, $A = 4$, $\overline{w} = 1$, $b = 0.5$.
Solution: $K = 6.74$, $L = 0.78$, $\pi/L = 2.23$, $w = 2.93$.

[3]: $e = \left(b(w-\overline{w})/(0.6L+b)\right)^{1/2}$, $\phi = 0.4$, $r = 0.15$, $A = 4$, $\overline{w} = 1$, $b = 0.5$.
Solution: $K = 9.03$, $L = 0.95$, $\pi/L = 1.54$, $w = 4.17$.

[4]: $e = \left((w-\overline{w})/1.6\right)^{1/2}$, $\phi = 0.4$, $r = 0.15$, $A = 4$, $\overline{w} = 1$, $rK \le 1$.
Solution: $K = 6.67$, $L = 3.84$, $\pi/L = 0.64$, $w = 2.43$.

[5]: $e = \left(b(w-\overline{w})/(0.6L+b)\right)^{1/2}$, $\phi = 0.4$, $r = 0.15$, $A = 4$, $\overline{w} = 1$, $rK \le 1$.
Solution: $K = 6.67$, $L = 0.91$, $\pi/L = 1.67$, $w = 4.20$.

Appendix 1

A. The Two-Period Attrition Model

Assume that the probability that individual i in firm j, observed at time 1, will be observed again at time 1+s can be modelled using a probit model. Letting S_{ij} denote the selection indicator, we write the selection equation for time κ+1 as

[A1] $$S_{ij} = 1[z_{ij}\delta + u_{ij} > 0],$$

where 1[a] is an indicator function equal to 1 if the event a is true and zero otherwise, z is a vector of variables determining attrition, δ is a vector of coefficients to be estimated and u is a normally distributed residual with mean zero and variance equal to one. Sample selectivity bias arises if u is correlated with the differenced residual in equation [2] in the main text. To allow for this possibility we assume that u and the differenced residual follow a joint normal distribution where the correlation coefficient is denoted ρ. Under these assumptions, and provided that f_{jt} is exogenous and selection does not depend on $\tilde{f}^{s}_{j,1+s}$, it follows that

[A2] $$E\left(\tilde{y}^{s}_{ij,1+s}\right) = \gamma \cdot \tilde{f}^{s}_{j,1+s} + \tilde{\theta}^{s}_{1+s} + \rho\lambda\left(z_{ij}\delta\right),$$

where $\lambda\left(z_{ij}\delta\right) = \phi\left(z_{ij}\delta\right)/\Phi\left(z_{ij}\delta\right)$ is the inverse Mill's ratio.[20] Because $\lambda\left(z_{ij}\delta\right)$ is unobserved we use a two-stage procedure, first estimating the probit model in order to obtain estimates of $\lambda\left(z_{ij}\delta\right)$, denoted $\hat{\lambda}_{ij} = \lambda\left(z_{ij}\hat{\delta}\right)$, and then regressing $\tilde{y}^{s}_{ij,1+s}$ on $\tilde{f}^{s}_{j,1+s}$, $\tilde{\theta}^{s}_{1+s}$ and $\hat{\lambda}_{ij}$. It is noted that the attrition model becomes more complicated if the number of periods exceeds two. While we have up to six periods of data, we never estimate the attrition model based on data over more than two periods (see Table 3).

[20] ϕ(.) and Φ(.) denote the density function, and the cumulative density function, respectively, of the standard normal distribution.

B. The Weight Matrix in the One-Step GMM Estimator

Alternative choices for the weights in the matrix W give rise to a set of GMM estimators based on the moment conditions, all of which are consistent. In general the optimal weights are such that W is equal to the inverse of the asymptotic variance of the moment conditions, which typically has to be estimated (see e.g. Wooldridge, 2002, Chapter 14). Unfortunately the asymptotic standard errors associated with the resulting two-step estimator are typically downward biased in finite samples, potentially giving rise to misleading inference. For this reason we focus on a GMM estimator that can be obtained in one step. Arellano and Bond (1991) suggest a one-step GMM estimator in which the weight matrix reflects the moving average process of first differenced residuals, given that the residuals in levels are serially uncorrelated and homoskedastic. The resulting weight matrix takes the form

$$W = \left(N^{-1} \sum_{i=1}^{N} Z_i' H Z_i \right),$$

where H is a square matrix which has twos in the main diagonal, minus ones in the first subdiagonals and zeroes otherwise, and whose dimension conforms to that of the instrument matrix Z_i. Our estimator, which combines different orders of differencing for t=1,2,...,6, uses the following definition of the (symmetric) H matrix:

$$
\begin{array}{ll}
 & \text{(t, s)} \\
H = \left| \begin{array}{ccccccccccccccc}
2 & & & & & & & & & & & & & & \\
-1 & 2 & & & & & & & & & & & & & \\
0 & -1 & 2 & & & & & & & & & & & & \\
0 & 0 & -1 & 2 & & & & & & & & & & & \\
0 & 0 & 0 & -1 & 2 & & & & & & & & & & \\
1 & 1 & -1 & 0 & 0 & 2 & & & & & & & & & \\
-1 & 1 & 1 & -1 & 0 & 0 & 2 & & & & & & & & \\
0 & -1 & 1 & 1 & -1 & -1 & 0 & 2 & & & & & & & \\
0 & 0 & -1 & 1 & 1 & 0 & -1 & 0 & 2 & & & & & & \\
1 & 0 & 1 & -1 & 0 & 1 & 1 & 0 & -1 & 2 & & & & & \\
-1 & 1 & 0 & 1 & -1 & 0 & 1 & 1 & 0 & 0 & 2 & & & & \\
0 & -1 & 1 & 0 & 1 & -1 & 0 & 1 & 1 & 0 & 0 & 2 & & & \\
1 & 0 & 0 & 1 & -1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & & \\
-1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & \\
1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 2
\end{array} \right|
&
\begin{array}{c}
2, 1 \\ 3, 1 \\ 4, 1 \\ 5, 1 \\ 6, 1 \\ 3, 2 \\ 4, 2 \\ 5, 2 \\ 6, 2 \\ 4, 3 \\ 5, 3 \\ 6, 3 \\ 5, 4 \\ 6, 4 \\ 6, 5
\end{array}
\end{array}
$$

where (t, s) denotes the time period, and the order of differencing, respectively. Hence this is a straightforward extension of the matrix proposed by Arellano and Bond, taking into account the different orders of differencing and the resulting correlation of the differenced residuals.

Appendix 2: The Structure of the Panel

TABLE A1: SAMPLE STRUCTURE

	Observed:	1995	1996	1997	1998	1999	2000	Total
Initial observation								
A. GHANA								
1995	Firms	117	107	83	83	65	70	525
	Employees	*685*	*626*	*116*	*110*	*50*	*52*	*1639*
1996	Firms		9	5	7	3	5	29
	Employees		*114*	*18*	*17*	*7*	*6*	*162*
1997	Firms			37	35	22	21	115
	Employees			*727*	*685*	*240*	*245*	*1897*
1998	Firms				7	0	0	7
	Employees				*156*	*20*	*31*	*207*
1999	Firms					5	5	10
	Employees					*354*	*333*	*687*
2000	Firms						2	2
	Employees						*103*	*103*
B. KENYA								
1995	Firms	159					60	219
	Employees	*1010*					*78*	*1088*
2000	Firms						106	106
	Employees						*822*	*822*

Note: The table shows the number of firms and employees observed for the first time in the years indicated in column 1, and how many observations there are on these firms and employees in subsequent years.