

Unobserved Heterogeneity and the Relation between Earnings and Firm Size: Evidence from Two Developing Countries

March 2004

Måns Söderbom^{£,+}, Francis Teaf[£] and Anthony Wambugu[§]

Abstract

It has been argued that the most likely explanation for the result that earnings rise with firm size is that large firms employ high-ability individuals. In this paper we use matched employer-employee panel data from Ghana and Kenya and test for firm size effects in earnings regressions whilst controlling for unobserved ability in the form of worker fixed effects. For both countries we obtain a size effect that is both economically and statistically significant.

JEL classification: J3, L25.

Keywords: Earnings, firm size, unobserved heterogeneity, Africa.

[£] Centre for the Study of African Economies, Department of Economics, University of Oxford.

[§] Department of Economics, Kenyatta University, Kenya.

⁺ Corresponding author: Mans.Soderbom@economics.ox.ac.uk

1. Introduction

The human capital model, in which earnings reflect skill differentials in perfect factor markets, has dominated the interpretation of earnings functions in both developed and developing countries. The result that earnings tend to rise with firm size has been widely interpreted within this framework, e.g. Oi and Idson (1999). If large firms hire more productive individuals than small firms and individual ability is partially unobserved to the econometrician, then the size-wage result is consistent with the standard human capital model and competitive labour markets. In this case, with appropriate controls for unobserved skills, the size effect will be insignificant. In contrast, if there are “structural” reasons for the size-wage relation, e.g. higher monitoring costs in large firms (Bulow and Summers, 1986; Ringuede, 1998) or job matching with search costs (Bertola and Garibaldi, 2001), there will be a significant size effect conditional on unobserved skills.

Testing for a size effect on earnings whilst conditioning on the full set of labour skills is difficult, especially with cross-section data.¹ Several studies report a size-wage premium conditional on a wide range of observable factors, however it is possible that this result is driven by the failure to control for unobserved skills, as conjectured by Oi and Idson.² Progress on the matter can be made with longitudinal data on employees and firms, as this enables the researcher to control for all time invariant unobserved characteristics. A limited number of studies exploiting such data sets have now been produced. Abowd, Kramarz and Margolis (1999) find that individual effects explain about 75 per cent of the firm-size wage effect in France. Arai (2003) shows that, based on Swedish data, the size effect falls as a result of controlling for unobserved worker characteristics; Criscuolo (2000) obtains a similar result for Germany. These studies thus support the notion that unobserved heterogeneity is important.

As far as we know, this issue has not been empirically investigated for developing countries. In this paper we test for the first time whether there is a size effect on earnings in

¹ An early and very influential study in this area is that by Brown and Medoff (1989).

² See e.g. Troske (1999) and Bayard and Troske (1999).

Africa, conditional on individual fixed effects. Labour markets in Africa are much more segmented than in developed countries, which, if the size effect is structural, may result in a larger size effect in Africa.³ Recent empirical studies suggest the correlation between size and earnings is relatively strong in Africa (e.g. Velenchik, 1997; Strobl and Thornton, 2001; Manda, 2002; Söderbom and Teal, 2004), however because none of these control for unobserved heterogeneity it cannot be established whether or not this is driven by omitted factors.

2. Data and Empirical Analysis

Our data have been collected as part of field surveys of manufacturing firms in Ghana and Kenya. These data cover the period 1995-2000 for Ghana, and 1995, 1999 and 2000 for Kenya.⁴ About 200 firms were included in each survey. At the same time as the firms were surveyed up to ten workers from each firm were interviewed, yielding matched employer-employee data and there is a panel dimension in the firm data as well as in the employee data. It should be noted that individuals that leave firms are not traced, and so we only have panel data on 'stayers'. This means that, unlike Abowd et al. (1999) and others, we cannot identify firm and individual effects separately.

Table 1 shows summary statistics for the variables we use in the empirical analysis. We identify two size categories: small, which is firms with up to 30 employees, and large those with 31 or more employees. Our full sample consists of 5,095 observations on Ghanaian employees and 2,741 observations on Kenyan ones. In both countries, the average level of monthly earnings in small firms is about USD 55-60, while in large firms it is more than USD

³ For a recent overview of African labour markets, see Mazumdar and Mazaheri (2002).

⁴ The Kenyan survey in 1995 was part of the World Bank's Regional Programme on Enterprise Development (RPED). The Ghanaian data, and the Kenyan data for 1998 and 1999, were collected by teams from the Centre for the Study of African Economies, Oxford University. The surveys cover seven industrial sub-sectors that comprise approximately 70 per cent of employment and value-added in these economies. The structure of the survey questionnaire was designed to make it comparable over time and across countries. See Söderbom (2001) for details about the Kenyan data; see Rankin, Söderbom and Teal (2002) for details about the Ghanaian data.

100. The average years of education, tenure and age, are higher in large than in small firms so clearly this earnings differential may reflect differences in human capital.

We proceed by regressing the logarithm of earnings on education, age, age squared, tenure, a set of control variables, and the logarithm of firm size.⁵ Table 2 shows OLS results without controls for unobserved heterogeneity. In both countries firm size has positive, large and highly significant coefficient. The human capital coefficients are largely as one would expect, with the possible exception of tenure which is insignificant in all four specifications (tenure is strongly correlated with age, however, which is highly significant). Less than half of the variation in earnings is explained by the models shown in Table 2, suggesting an important role for unobserved heterogeneity.

We now control for unobserved worker heterogeneity in the form of individual fixed effects.⁶ Education and gender are collinear with the fixed effects and so cannot be included in the model. Similarly, the age and tenure effects are absorbed by the time dummies. Results are shown in Table 3. The first two columns show results from the within, or the “fixed effects”, estimator. For both countries, the size coefficient is smaller than what we get with OLS, which is consistent with the hypothesis that firm size is positively correlated with unobserved skills. For Ghana the fall in the point estimate of the coefficient is dramatic: from 0.21 based on OLS to 0.05 with controls for fixed effects. For Kenya the fall is more modest, from 0.10 to 0.08. In both cases the size coefficient is significant at the five per cent level or better. *Thus, conditional on all time invariant factors, there is a size effect on earnings.*

One common explanation why fixed effects estimates tend to be lower than their OLS counterparts is that “going within” exacerbates measurement error bias (Griliches and Hausman, 1986). Under classical measurement errors estimating the model in first differences gives more severely biased results than based on the within estimator, which in turn is more biased than OLS. If measurement errors are absent the within estimator has the same probability limit as the first differenced estimator. Thus comparing the results from these two

⁵ The vector of control variables include gender, location and time dummies.

⁶ We drop individuals that are only observed once.

estimators is informative of whether measurement error bias is a problem. Columns 3-4 in Table 3 show first differenced results. The size coefficient is now much closer to zero, and for neither country is the size effect significantly different from zero.

The measurement error problem appears more severe for Ghana than for Kenya (why is unclear). The continuity of the Ghanaian panel (six years of consecutive observations) enables us correct for measurement errors using the approach proposed by Griliches and Hausman (1986). This method exploits the result that, under serially uncorrelated measurement errors, firm size dated s is a valid instrument for the difference $(f_{ijt} - f_{ij,t-t})$ provided $s \neq t$ and $s \neq (t-t)$. By forming a system of equations with varying length of differencing and different instruments for each equation, the parameters of interest can then be consistently estimated using a generalized method of moments estimator (GMM; Hansen, 1982). Spatial constraints prevent us from going into further details about Griliches and Hausman estimator here; for a recent exposition we refer the reader to Arellano (2003), pp. 51-4. One-step GMM results are shown in column 5 of Table 3 and the instruments are indicated in the table notes.⁷ The estimated size coefficient is 0.15 i.e. rather much higher than what is obtained by the within or the first differenced estimator, but still somewhat lower than the OLS result. The coefficient is significant at the five per cent level, and the Sargan-Hansen test indicates that the overidentifying restrictions can be accepted. The latter result suggests that the assumption of non-autocorrelated measurement errors is not overly restrictive.

3. Conclusions

Our data is consistent with other studies which have shown that the size effect on wages is higher in developing than developed countries. To date the studies for developing countries have had to rely on controls for observable skills. We have extended that analysis by using panel estimators that control for all time invariant unobservable skills of the workforce. Panel estimators, while allowing for fixed effects, can produce seriously biased estimates if

⁷ Two-step GMM results are much the same except that the standard errors are lower. We attribute the latter result to the downward bias in finite samples documented by Arellano and Bond (1991).

measurement errors are present. We have used an estimator developed by Griliches and Hausman to show how important is the presence of measurement error for the Ghanaian results. Our preferred estimates of the size effect are smaller than the OLS results, confirming that unobserved skills are important. However the size effect remains significant and is much higher than that found in developed country data, suggesting that there may well be structural differences between poor and rich economies in the operation of their labour markets.

References

- Abowd J., M. Kramarz and D.N. Margolis (1999). "High Wage Workers and High Wage Firms," *Econometrica* 67: 251-333.
- Arai M. (2003). "Wages, Profits, and Capital Intensity: Evidence from Matched Worker-Firm Data," *Journal of Labor Economics* 21: 593-618.
- Arellano, M. (2003). *Panel Data Econometrics*. Oxford: Oxford 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.
- 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.
- Bertola G. and P. Garibaldi (2001). "Wages and the Size of Firms in Dynamic Matching Models," *Review of Economic Dynamics* 4: 335-368.
- Brown, C. and Medoff, J. (1989) "The employer size-wage effect," *The Journal of Political Economy* 97: 1027-1059.
- 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.

- Criscuolo, C. (2000). "Employer Size-Wage Effect: A Critical Review and an Econometric Analysis". Working Paper 277, Dipartimento di Economia Politica, Università degli Studi di Siena.
- Griliches, Z. and J.A. Hausman (1986) "Errors in variables in panel data", *Journal of Econometrics*, 31, pp.93-118.
- Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50: 1029-1054.
- 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. and Mazaheri, A. (2002). *Wages and Employment in Africa*. Ashgate Publishing Group.
- 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.
- Rankin, N., M. Söderbom and F. Teal (2002) The Ghanaian Manufacturing Enterprise Survey 2000. REP/2002-05 at <http://www.csae.ox.ac.uk/reports/main.html>
- Ringuedé, S. (1998) "An efficiency wage model for small firms: firm size and wages", *Economic Letters*, 59: 263-8.
- 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.
- Söderbom, M. (2001) Constraints and Opportunities in Kenyan Manufacturing: Report on the Kenyan Manufacturing Enterprise Survey 2000 REP/2001-03: <http://www.csae.ox.ac.uk/reports/main.html>.
- Söderbom, M. and F. Teal (2004) "Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data," *Journal of Development Economics* 73, 369-394.

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

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

TABLE 1: SUMMARY STATISTICS

	Small firms (employment ≤ 30)		Large firms (employment > 30)	
	Mean	Median	Mean	Median
A. GHANA ⁽¹⁾				
Monthly earnings (USD)	55.7	44.3	124.0	78.6
Years of education	10.3	10.0	12.2	12.0
Age	33.9	31.0	38.7	37.0
Years of tenure / 10	7.0	0.4	8.5	0.6
Male proportion	0.76	1.0	0.86	1.0
B. KENYA ⁽²⁾				
Monthly earnings (USD)	57.3	46.8	111.8	68.4
Years of education	9.0	8.0	10.3	11.0
Age	32.0	30.0	35.3	34.0
Years of tenure / 10	6.6	0.4	8.9	0.7
Male proportion	0.85	1.0	0.83	1.0

⁽¹⁾ Number of firms: 136. Number of individuals: 2330. Number of observations: 5095.

⁽²⁾ Number of firms: 236. Number of individuals: 1906. Number of observations: 2741.

TABLE 2
EARNINGS FUNCTION ESTIMATES: NO CONTROLS FOR FIXED EFFECTS

	Ghana		Kenya	
	(1)	(2)	(3)	(4)
Education (years)	0.083 (16.64)**	0.067 (15.57)**	0.104 (11.60)**	0.086 (9.86)**
Tenure (years)	0.031 (0.94)	0.015 (0.57)	0.001 (0.02)	-0.024 (0.71)
Age	0.066 (7.54)**	0.040 (4.88)**	0.047 (4.08)**	0.030 (2.85)**
Age ² / 100	-0.056 (5.15)**	-0.030 (2.89)**	-0.030 (2.06)*	-0.011 (0.78)
Male	0.193 (3.86)**	0.155 (3.45)**	0.087 (1.77) ⁺	0.090 (1.91) ⁺
ln Employment		0.212 (9.88)**		0.101 (4.87)**
R-squared	0.38	0.47	0.32	0.35

Note: The dependent variable is the logarithm of monthly earnings, expressed in USD. The estimation method is OLS. A constant and dummy variables for time and location in the capital city are included in all regressions. The numbers in () are *t*-statistics based on standard errors robust to heteroskedasticity and intra-firm correlation. Significance at the 1 per cent, 5 per cent and 10 per cent level is indicated by *, ** and ⁺ respectively.

TABLE 3
EARNINGS REGRESSIONS CONTROLLING FOR UNOBSERVED HETEROGENEITY

	Fixed Effects		First Differences		Griliches-Hausman GMM ⁽¹⁾
	(1) Ghana	(2) Kenya	(3) Ghana	(4) Kenya	(5) Ghana
In Employment	0.049 (2.43)*	0.079 (4.59)**	0.006 (0.15)	0.038 (1.59)	0.149 (2.13)*
Individuals	1878	759	1878	759	1878
R-squared	0.22	0.41	0.19	0.01	
J-test (p-value) ⁽²⁾					0.60

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 regressions.

⁽¹⁾ These are one-step GMM estimates. The instruments for $(f_{ijt} - f_{ij,t-t})$, $t = 1, 2, \dots, 5$ are as follows:

t	Instruments
1	$f_{ij,t-2}$ and time dummies
2	$f_{ij,t-1}$, $f_{ij,t-3}$ and time dummies
3	$f_{ij,t-1}$, $f_{ij,t-4}$ and time dummies
4	$f_{ij,t-1}$ and time dummies
5	$f_{ij,t-1}$ and time dummies

⁽²⁾ This is the Sargan-Hansen test for the validity of the overidentifying restrictions. The test statistic is equal to 15.85 and there are 18 degrees of freedom.