Job Market Paper

Linking Firms and Workers: Heterogeneous Labor and Returns to Education

Garth Frazer[1]

Yale University

First Version: June, 2001; This Version November, 2001

Abstract

Accurate measurement of returns to education is particularly important in African countries where government resources are extremely limited, and education's competition for budget dollars is intense. This paper attempts to measure consistently the returns to education in Ghana. The chief problem to be handled in the measurement of returns to education in Mincer-based wage regressions is ability bias. This paper attempts to address this problem from a completely new perspective--that of the firm. It begins by specifying a production function that is consistent with the Mincerian wage equation. The production function is estimated consistently, taking into account the simultaneity of labor demand and output decisions. A result of including worker schooling, experience and ability in the production function specification is that a measure of worker ability, defined as the worker's contribution to firm product after controlling for schooling and experience, is acquired. This measure of worker ability can be used in a wage equation in order to control directly for worker ability to obtain consistent estimates of the returns to schooling. The model is estimated using data from the manufacturing sector in Ghana, but this technique is replicable to other countries and datasets, given the recent increase in available linked employer-employee datasets.

[1] I would like to thank Chris Udry, Steve Berry, and T.N. Srinivasan for their advice, support and comments during this research, as well as Patrick Bayer, Paul Schultz, Ann Huff Stevens, Philip Levy and seminar participants at Yale University for helpful comments on this paper. I would also like to thank Francis Teal for the opportunity to participate in the Ghanaian survey, and to Oxford University and the Ghana Statistical Service for the use of the data from the Ghanaian Manufacturing Enterprise Survey. Permission from the Ghana Statistical Service to use the Ghana Living Standards Survey - Wave 4 data is gratefully acknowledged. All remaining errors in this paper are my own. Correspondence may be sent to Garth Frazer, Department of Economics, Yale University, P. O. Box 208268, New Haven, CT 06520-8268, phone 203-782-9326, email garth.frazer@yale.edu.

1 Introduction

What is the role of education in increasing incomes and reducing poverty? This has been a central question for development economists and policy makers, as accurate measurement of the wage returns to education is particularly important in contexts where poverty is common, government resources are extremely limited, and education's competition for budget dollars is intense.[1] While consistent estimation of the returns to education has received considerable attention of labor economists in industrialized country contexts, as well as development economists in some countries, comparatively little "best practice" research has been done for African countries (Behrman, 1996; Appleton, Hoddinott, and Mackinnon, 1996). The primary problem in consistent measurement of the wage returns to schooling is that of ability bias. A worker's ability is typically part of the residual, and yet is correlated with an individual's schooling, resulting in biased parameter estimates. This study develops a new technique for controlling for this ability bias, using information from the firm where workers work, and uses this technique to obtain consistent estimates of the returns to education in Ghana.

Over the past twenty-five years, the model that has dominated the estimation of returns to education is Mincer's human capital model (1974), which predicts that a worker's wage should be a function of schooling and work experience. Mincer acknowledged the ability bias problem: "It is widely believed that the omission of ability from the earnings function creates a specification bias: leaving out a variable which is positively correlated with earnings and investment[2] biases the coefficient of investment (average rate of return) upward." (p. 139) The vast majority of studies have attempted to address the issue of ability bias in the context of the Mincer model using "quasi-experiments" and instrumental variable techniques, where some trait, such as distance to the nearest school, level of parental education, different schooling laws, or a twin's level of education, are assumed to exogenously affect a person's level of education. While these studies have considerably advanced our understanding of returns to education, each method has its limitations, as will be summarized in Section 2.

This study attempts to examine the question from a different perspective entirely–in particular, that of the firm. It takes a result of the Mincer model, that an individual's productivity is a function of her schooling, experience, and ability, and develops the unique labor term of the production function that is consistent with this model.[3] Typically, in production

[1] Some studies which have examined the returns to schooling in the context of developing countries include: Behrman and Deolalikar (1991) - Indonesia ; Behrman, Foster, Rosenzweig, and Vashishta (1999) - India; Godana (1997) - Zimbabwe; Li and Zhang (1998) - China; Mwabu and Schultz (1996) - South Africa, Nielsen and Westergård-Nielsen (2001) - Zambia; Ram and Singh (1988) - Burkina Faso; Siphambe (2000) - Botswana, Smith and Metzger (1998) - Mexico.

[2] i.e. schooling investment

[3] The inclusion of schooling and experience into a production function is not new to this study, but was first introduced by the working paper version of Bils and Klenow (2000), and has been used with firm data in Jones (2001), Bigsten et. al. (2000), and Hellerstein, Neumark and Troske (1999). The inclusion of ability,

function studies, the labor term is simply an hours measure, or at best a number of skilled and unskilled workers, so this new specification more accurately depicts the heterogeneity of labor at a firm. With this Mincer-consistent production function, a measure of worker ability can be obtained, where worker ability is defined as the contribution of workers to firm product after controlling for schooling and experience. The thesis of this paper is that this is exactly the definition of ability that matters. Under the Mincer-consistent specification, worker ability is shown to form one component of firm productivity, where firm productivity is defined as in the current standard of the industrial organization literature as that portion of the residual which can be seen and acted upon by the firm manager. To separate worker ability from that portion of firm productivity which is due to management talent or to technological change, management talent is assumed to be fixed for a given manager, and technological change is assumed to be stationary about a trend. The production function is estimated consistently, taking into account the simultaneity of labor demand and output decisions, with a technique that follows from Levinsohn and Petrin (2000) and Olley and Pakes (1996). The measure of worker ability obtained is an average at the firm level, and can be used in a wage regression, which is grouped at the firm level, to control for ability, and to obtain a consistent estimate of the returns to education.

The procedure used in this study can be generalized. The assumptions of any wage equation in which the log(wage) is a linear function (in the broadest definition of this term) of various factors can be incorporated into a production function. The productive contribution of these factors can be compared to their relative remuneration. Assuming within-period profit maximization, the coefficients of the factors in the production function should equal the corresponding coefficients in the wage equation (Section 3.1 and Appendix 1). This technique should allow greater exploitation of the linked employer-employee data which is now available for a number of countries.

The outline of this paper is as follows. Section 2 outlines methods which have been used to date to attempt to control for ability bias. Section 3 outlines the incorporation of the Mincer assumptions into the production function, and the procedure for its estimation. Section 4 describes the wage equation estimation. Section 5 describes the data used for estimation, with the results in Section 6, and a brief conclusion in Section 7.

2 The Problem of Ability Bias

Mincer's widely-used human capital model (1974) predicts that the following relationship between wages, schooling and experience should hold:

$$\log W_i = \lambda_0 + \lambda_S S_i + \lambda_X g(X_i) + \chi_i \tag{1}$$

and the incorporation of the Mincer model into the production function is new to this study.

The work experience variable, X, is designed to capture the effects of on-the-job training which people receive over their worklife, and in Mincer's model was a quadratic. Given the typical difficulty in measuring this exactly, a variable measuring potential experience was proposed by Mincer, and is typically used. A person's potential experience is the length of his post-schooling life, i.e. X=G-S-6, where G is a person's age, and the number six removes a person's pre-school years. Now, the chief problem with this regression is the fact that the residual, χ_i, includes a person's ability, which could both independently affect a person's wage and be correlated with one's schooling. The complete equation would read:

$$\log W_i = \lambda_0 + \lambda_S S_i + \lambda_X g(X_i) + \lambda_A A_i + \upsilon_i \tag{2}$$

The potential of individual ability to confound the returns to schooling has been recognized at least since Becker (1964). The problem in the above specification is that, if ability is omitted from the regression, the schooling coefficient will be biased. To consider the likely direction of the bias (given that ability is not as correlated with experience as it is with schooling), consider the plim for the OLS coefficient in a regression without experience:

$$p\lim \lambda_{S\_OLS} = \lambda_S + \lambda_A \frac{cov(A_i, S_i)}{var S_i} \tag{3}$$

Specifically, λ_S will be biased upward provided that: i) ability affects the wage, independent of its effect on schooling, and ii) schooling and ability are positively correlated. A variety of approaches have been used to address this problem, with the major approaches being summarized in a survey article by Card (1999). As Griliches (1977, p. 5) notes, "the simplest way of dealing with this problem is to find a measure of 'ability' and include it in such an equation," and such an approach was common in the earlier literature (Griliches and Mason, 1972; Griliches, 1976; Griliches, 1977). However, as Griliches (1977) notes in the same article, the value of using IQ-measures as ability controls was controversial: "Two polar views are possible. 'Ability' is IQ, or something close to it, and the only problem is that our measures of it are subject to possibly large (test-retest) errors...The alternative view is that 'ability', in the sense of being able to earn higher wages, other things equal, has little to do with IQ. (p. 7)" In the subsequent 25 years, IQ or other test-related measures have been rarely used to control for ability bias,[4] and Card does not even discuss this method of estimation in his survey article. The current study is consistent with this evaluation of the IQ-ability measures. This study uses a direct control for ability, but not one which is obtained from an IQ-type test, but rather one which is revealed through an individual's productivity after controlling for schooling and experience. Before delineating the procedure employed, a review of the techniques used to date to handle the problem of ability bias is warranted.

[4] Some studies that have used test measures include: Boissiere, Knight and Sabot, 1985; Glewwe, 1999; Knight and Sabot, 1990; Glewwe, 1996; Blackburn and Neumark, 1995.

2.1 Instrumental Variable Techniques

A variety of studies have attempted to exploit institutional features of the school system across states or regions, times, or individuals, as "quasi-experiments" where people have been subjected to differing levels of treatment in terms of the costs or benefits of schooling. These institutional features of the school system have served as instruments for the schooling variable, to purge it of its correlation with unobserved ability.[5] As evident in the results in Appendix Table 1, the instrumental variable estimates are virtually always higher than the OLS estimates and at times considerably so. However, as discussed previously, the theory predicts that the OLS estimates should be upward-biased. Griliches (1977) provides a partial explanation to the puzzle in that the attenuation bias of measurement error could be biasing the OLS coefficients downward, offsetting the ability bias. However, building on the fact that the reliability ratio of self-reported schooling in U.S. datasets is about 90%[6], Card (1999, p. 1841) notes, "since measurement error bias by itself can only explain a 10% gap between OLS and IV, however, it seems unlikely that so many studies would find large positive gaps between their IV and OLS estimates simply because of measurement error."

Card (1995) suggests another explanation for these high IV estimates, namely heterogeneity in the marginal rate of return to schooling in the population. If such heterogeneity exists, then the IV estimates measure the returns to education for the subgroup of the population affected by the instrument. For example, in the studies using minimum school-leaving age as an instrument, the IV estimate provides the return to education for the subgroup of the population for whom the compulsory school-leaving age matters. Card (1995, 2001) argues that the subgroups affected by the institutional innovations in question typically will have higher marginal rates of return than other subgroups of the population, thus explaining why the IV estimates are typically larger than those of OLS.[7]

At least as common in the literature as the institutional variation studies is the use of family background variables as instruments. Here, the assumption is that parental schooling, for example, affects an individual's schooling level, but does not independently affect the wage. However, the fact that many studies use these family background variables to control for individual ability (e.g. Ashenfelter and Zimmerman, 1997; Ashenfelter and Rouse, 1998) suggests that parental variables, if other controls for ability are not used, may remain

[5] Angrist and Krueger (1991) use variations in American compulsory schooling laws, combined with an individual's quarter of birth as identifying instruments. Harmon and Walker (1995) use changes in the minimum school-leaving age in the United Kingdom. Kane and Rouse (1993) use public tuition and the distance to the closest 2-year and 4-year colleges as instruments.

[6] These reliability ratios have been calculated using the datasets primarily used by U.S. and European investigators, but these are primarily the results being presented. Reliability ratios for developing country datasets such as Maluccio's could be different.

[7] As Ichino and Winter-Ebmer (1999) note, this conclusion is consistent with the local average treatment effects (LATE) interpretation of IV estimates (Imbens and Angrist, 1994), namely that IV identifies the average treatment of those who comply with the assignment-to-treatment mechanism implied by the instrument.

correlated with the unobserved ability of the regression (in the residual), and therefore may not be valid instruments. A summary of studies that have used family background variables both as additional controls, and as instruments is found in Appendix Table 2.

Another set of studies that can be interpreted as instrumental-variable studies are studies of identical (monozygotic) twins, with the twin's schooling instrumenting for own-schooling. The assumption in the twins studies is that twins have identical abilities, because of their identical genetics, and therefore the difference in schooling levels between twins can be treated as a natural experiment. Regressing the differences in wages between twins on the differences in their schoolings should remove the effect of ability (which is presumed identical between them, and therefore differenced out). Studies of twins date back to Behrman and Taubman (1976). However, the question of whether identical twins do have identical ability has been raised since Griliches (1979),[8] and has been strongly echoed in recent papers (Rosenzweig and Wolpin, 2000; Bound and Solon, 1999; Neumark, 1999).[9] Still, whether a twin's education is an appropriate instrument is not even relevant for the context of Ghana, or other developing countries, where datasets on twins do not exist. A summary of twins' studies is provided in Appendix Table 3.

2.2 Developing Country Studies

Few studies of returns to schooling in developing countries use the best-practice studies of controlling for ability bias which have been delineated. In the case of twins, the historic absence of such datasets for developing countries is the obvious cause; in other cases, the cause is lack of appropriate questions in the surveys. Still, a number of studies have used the aforementioned techniques, and some of these will be reviewed here.

In the past fifteen years or so, the research activity using direct test-based controls for ability bias has actually been stronger in developing countries than elsewhere. The test that has been used to identify ability in the developing country context has been Raven's Progressive Matrices test (Boissiere, Knight and Sabot, 1985; Glewwe, 1999; Knight and Sabot, 1990; Glewwe, 1996). Reading and math tests have also been conducted, but are

[8] A further criticism of the twins' studies raised by Griliches (1979) is that measurement error in schooling may be exacerbated by differencing across twins, resulting in an even larger attenuation bias than the least-squares estimates. This concern has been addressed in recent twins' studies, which follow Ashenfelter and Krueger (1994) in using the other twin's report of the first twin's schooling to instrument for the measurement error.

[9] For example, in their sample of twins, Behrman, Rosenzweig and Taubman (1994) find that virtually half of the twins had birth weight differences of at least 8 ounces. Bound and Solon (1999) list a series of medical studies documenting the association between differences in birth weight and differences in IQ between twins.. Even if twins' birth weights are identical, if one twin possesses a stronger "work-ethic" (a component of ability) than another, that twin might obtain a higher schooling level (because of its lower psychic cost), and also be rewarded for his/her determination independently in the labor market. As Bound and Solon note (1999), citing work in the psychological literature, another plausible source of variation between twins is their psychological need to differentiate themselves from each other.

generally interpreted as cognitive skills resulting from education. The Raven test has been insignificant in all wage regressions of which I am aware (Boissiere, Knight and Sabot, 1985; Glewwe, 1999; Knight and Sabot, 1990; Glewwe, 1996). Therefore, while ability test scores have been used in developing countries, either they have not measured ability accurately, or ability bias does not exist.

Regarding institutional innovations, given the general lack of strong enforcement of compulsory schooling laws, when they do exist in developing countries, these have not been used as instruments. However, Duflo (2000) uses an institutional innovation in the form of a massive school construction initiative in Indonesia to identify the returns to schooling, and finds a coefficient between 0.0675(0.0280) and 0.106(0.0222), in comparison to her precisely estimated OLS coefficient of 0.077. Another instrument used in the industrialized country literature is distance to the nearest college. The comparably relevant variable in developing countries is distance to the nearest secondary school. Maluccio (1998) uses this measure, distance to the nearest secondary school, as well as whether a private secondary school is located in the nearest town in his study of the Philippines. His results are comparable to those of the distance-to-college instruments, with the IV estimate being considerably higher than the OLS estimate (IV: 0.145 (.041) vs. OLS: 0.0730(.114)), using a standard specification, plus gender and rural area controls.

Using family background variables as control variables (Heckman and Hotz, 1986; Armitage and Sabot, 1987) or as instruments (Schultz, 1995) is a bit more common in the returns to education literature in developing countries. As in the developed country literature, including family background variables as controls typically reduces the OLS estimates, while using them as instrumental variables increases the estimates.[10]

Nevertheless, the vast majority of developing country studies are simple OLS regressions, without correcting for ability bias.[11] A further bias which is considered in some developing country regressions is the issue of selection (Schultz, 1988). In developed countries, the issue of selection bias resulting from participation in the labor market has historically been strong in the case of women. Frequently, this issue has been avoided by using samples of men, for

[10] For example, Heckman and Hotz in a regression for male heads of household in Panama, find that the coefficient reduces from 0.1187(.0069) to 0.0856(.0074) once the education of both parents is included a standard Mincerian regression, with an indicator for technical training as an additional control. Schultz (1995), in a comparison study of Côte d'Ivoire and Ghana, finds that the use of parental education and occupation, local health infrastructure and food prices as instruments increases the schooling parameter from 0.124(.007) (OLS) to 0.165(.040) (IV) in Côte d'Ivoire and decreases it from 0.0393(.004) (OLS) to 0.0214(.024) in Ghana, although the estimation in Ghana is imprecise. Both these regressions also include controls for migration, body mass index and height, but not experience.

[11] Glewwe (1999) presents the national returns to education for Ghana which follow a Mincer regression, and therefore are most comparable to our results, although the data was collected in 1988-89 (Ghana Living Standards Survey (GLSS) - Round 2). He finds that the OLS measure of the returns to education is 8.5%. Using the GLSS data which is closest to the time period of the data under investigation here, the GLSS Round 4 data for 1998-99, we find that the OLS estimate has changed little, and is now 8.8% (author calculation).

whom the bias is not nearly as strong. In the case of developing countries, the bias exists for both genders, as the issue of selection into the labor force is coupled with the issue of selection into wage versus self-employment. Some studies have carefully handled this issue as well (Schafgans, 2000; Mwabu and Schultz, 2000; Lanzona, 1998; Vijverberg, 1993).

In summary, no consensus exists among labor economists on the best practice for handling ability bias. Developing country studies, in general, lack the attention to this issue paid in developed country studies (partly due to the absence of appropriate data, such as that of twins). In this context, this paper attempts to present a very different, and very direct way of handling the issue of ability bias when linked employer-employee data is available. The next section outlines the methodology of the current study.

3 Modeling the Production of the Firm

3.1 Incorporating the Heterogeneity of Labor in the Production Function

A frequent specification for production function estimation is the Cobb-Douglas form, which, written in logarithmic form, using lower cased variables, is:[12]

$$y_{ft} = \beta_0 + \beta_l l_{ft} + \beta_k k_{ft} + \varepsilon_{ft} \tag{4}$$

where y is the value-added, l the labor, and k the capital of firm f in period t. Estimation of equation (4) by least-squares raises two concerns. The first, the problem of simultaneity bias, has been understood in the literature at least since Marschak and Andrews (1944), although truly satisfying solutions to this problem have only arisen recently (Olley and Pakes, 1996; Levinsohn and Petrin, 2000). The second is the assumption of homogeneous labor, in that the variable l is typically specified as the number of employees or the number of worker-hours at a firm, or at best split into two types at the firm level. Considerable effort, therefore, will be given in this paper to incorporating the heterogeneity of labor into the production function. First, however, let us briefly examine the problem of simultaneity, which will also be addressed in this paper.

The simultaneity problem arises because the error term includes firm productivity, which is seen by the firm manager and will very likely be correlated with this period's labor input, which is typically considered to be freely variable. It may not be correlated with this period's capital stock, which is generally considered a quasi-fixed variable. Under standard assumptions, this will result in an upward bias on the labor coefficient, and possibly a downward bias on the capital coefficient (Levinsohn and Petrin, 2000). The current state-of-the-art for handling this problem, as echoed in a survey article on the problem by Griliches

[12] Throughout this paper, upper-cased Roman letters will refer to the standard form of variables, and lower cased letters their natural logarithms.

and Mairesse (1995), comes from the work of Olley and Pakes (1996). In short, they separate the error term into firm productivity, ω_{ft}, which is seen by the firm manager, and η_{ft}, a mean-zero component which is not. The production function then becomes:

$$y_{ft} = \beta_0 + \beta_l l_{ft} + \beta_k k_{ft} + \omega_{ft} + \eta_{ft} \tag{5}$$

In the Olley and Pakes model, the productivity term is derived, in the context of a dynamic model, to be a function of investment and the firm's capital stock, and is calculated as a nonparametric function of these two variables. Then, equation (5) can be estimated for observations where investment is non-zero. While restricting to observations with non-zero investment forces Olley and Pakes to lose 8 percent of their observations, in other datasets (such as the current one) this can force deletion of a much larger fraction, and sometimes the majority of observations (56% of observations in the Ghanaian dataset under examination). To overcome this limitation, Levinsohn and Petrin (2000) modify the Olley and Pakes procedure to use intermediate inputs, instead of investment, in the estimation of firm productivity. This study modifies the Levinsohn and Petrin procedure by including human capital variables in the production function, and the details of the procedure will be described in Section 3.2.

As mentioned, the other problem of simple estimation of equation (4) is the assumption of homogeneity of labor. This simplification reflects the limitations of the data which has typically been available for estimation. Fortunately, with the increasing availability of linked employer-employee datasets, which provide data on at least a sub-sample of, a firm's employees, we can now ask the question of what is a more sensible specification for labor's contribution to firm product.

A sensible place to begin is naturally in the labor literature. Given its success at explaining variation in wages across individuals, the human capital model of Mincer (1974) has dominated the estimation of earnings equations over the past twenty-five years. In a Mincer human capital model, an individual invests in schooling until the net present value of that investment is zero, that is the foregone present wage is equal to the discounted value of the increased future wage resulting from an additional year of schooling. As mentioned, the equation of estimation that results from the Mincer model is that of equation (2), where the log(wage) is a function of the years of schooling, years of experience (or age) and ability of individual i. If an individual is rewarded according to her productivity, then the factors in (2) (or whatever factors are included in the wage equation being estimated) should also be included in the firm's production function. Moreover, given the success of (2) in explaining wages, it is worth paying attention to the implications of this functional form for the firm. Equation (2) tells us that a firm's wage bill is the following:

$$\sum_{\iota} (e^{\lambda_0 + \lambda_S S_i + \lambda_X X_i + \lambda_A A_i + \xi_i}) \tag{6}$$

Note that each individual's remuneration is a convex function of schooling, experience, and ability. To consider the implications of this formulation of the firm's wage bill on the firm's production function, consider an individual factor of the above function, say the overall level of schooling in the firm. In particular, note that schooling is not linearly substitutable between individuals. The cost to the firm of substituting a current member of the workforce with a new worker, identical in every respect, except with an additional year of schooling depends on which individual worker is replaced, according to the functional form of (6). The same is true of other worker characteristics. In fact, if the firm is optimizing (given that the firm is a price-taker in wages), the productive contributions of workers' characteristics should reflect the relative costs of these characteristics. In order for this to be the case, the labor term in the production function should have the same form as (6). That the labor term in the production function should have exactly the same form as the wage bill for a profit-maximizing firm is proven in Appendix 1. Therefore, the production function which is consistent with a Mincerian wage equation is $F[\sum L_j e^{\lambda_0+\lambda_S S_j+\lambda_X X_j+\lambda_A A_j}, K, Q, M]$. Here, managerial ability is M, a further component of firm productivity (to be discussed later) is Q, and the number of workers of type j with characteristics (S_j, X_j, A_j) is L_j (so that $\sum L_j = L$ is the total number of workers at the firm, and $L_j = 0$ for types that are unused by the firm). Therefore, defining the price of output as p, the profit function for the firm is:

$$\Pi = pF[\sum L_j e^{\lambda_0+\lambda_S S_j+\lambda_X X_j+\lambda_A A_j}, K, Q, M] - \sum L_j e^{\lambda_0+\lambda_S S_j+\lambda_X X_j+\lambda_A A_j+\xi_j} - rK \qquad (7)$$

The return to capital is simply r. The choice variables for the firm are the values for each of the L_j's, as well as K, with the first term of the production function labelled as the quality of labor, or effective labor.[13] While, by Appendix 1, the coefficients on an individual's schooling in the production function must equal that on the wage equation in the above profit function, this restriction will be tested, rather than imposed on the data. Therefore, the profit function that will be considered is of the form:

$$\Pi = pF[\sum L_j e^{\beta_1+\beta_S S_j+\beta_X X_j+\beta_A A_j}, K, Q, M] - \sum L_j e^{\lambda_0+\lambda_S S_j+\lambda_X X_j+\lambda_A A_j+\xi_j} - rK \qquad (8)$$

After estimating the production function, and the wage equation separately, the equalities $\beta_S = \lambda_S$, and $\beta_X = \lambda_X$ will be tested. Following the literature standard, the technology for F will be Cobb-Douglas, and therefore the production function to be estimated is:

$$Y = e^{\beta_0}(\sum L_j e^{\beta_1+\beta_S S_j+\beta_X X_j+\beta_A A_j})^{\beta_H} K^{\beta_K} (e^M e^Q)^{\beta_\omega} e^{\varepsilon} \qquad (9)$$

[13] To be precise, it is assumed that the labor force is dense enough that the firm is able to choose the value of L_j for each of the possible three-tuples (S_j, X_j, A_j) in the space of the domain. For the purposes of the empirical investigation, S_j and X_j will be measured in integral values from 0 to say 30 and 90, respectively. For consistency of exposition here, think of A_j as taking on values between 0 and 100. The actual domain definitions are inconsequential for the analysis that follows, but are given for completeness. All of the arguments that follow carry through to the continuous case.

In equation (9), for the summation to calculate labor quality, the only types that matter are those chosen in positive quanity ($L_j > 0$), so that this term can be re-indexed by the $i = 1, ..., L$ workers chosen at a firm to get:

$$Y = e^{\beta_0}(\sum_{i=1}^{L} e^{\beta_1+\beta_S S_i+\beta_X X_i+\beta_A A_i})^{\beta_H} K^{\beta_K}(e^M e^Q)^{\beta_\omega} e^{\varepsilon} \tag{10}$$

The subscripts for firm f and time t remain omitted for this discussion of functional form, but will be introduced subsequently. If we use the logarithmic Cobb-Douglas form, by taking logarithms of both sides of the equation, and representing each variable's natural logarithm using its lower-cased letter, then:

$$y = \beta_0 + \log(\sum_{i=1}^{L} e^{\beta_1+\beta_S S_i+\beta_X X_i+\beta_A A_i})^{\beta_H} + \beta_K k + \beta_\omega(M+Q) + \varepsilon \tag{11}$$

Now, if we observed measures for each of the variables in the estimation procedure then estimation of (11) by non-linear least squares (for example) would provide consistent estimation of the parameters. Such estimation would require knowledge of the distribution of workers at the firm (or at least a sample estimate of the distribution), and their schooling, experience, and ability. Unfortunately, M, Q, and each of the A_i's is unobserved. Before being able to use the most recent techniques from the literature on estimating production functions for estimating equation (11), some further work is needed. First note that the factor e^{β_1} is common across all workers, and can be factored, resulting in:

$$y = \beta_0 + \beta_1\beta_H + \log(\sum_{i=1}^{L} e^{\beta_S S_i+\beta_X X_i+\beta_A A_i})^{\beta_H} + \beta_K k + \beta_\omega(M+Q) + \varepsilon \tag{12}$$

To simplify the labor term, define $f(S_1, ...S_L, X_1, ...X_L, A_1, ...A_L) = \log(\sum_{i=1}^{L} e^{\beta_S S_i+\beta_X X_i+\beta_A A_i})$. A first-order Taylor approximation to f is the following:

$$\begin{aligned} f(S_1, ... \quad & S_L, X_1, ...X_L, A_1, ...A_L) \quad = \\ & f(\underbrace{0, 0, ...0}_{3L\ 0's}) + \sum_{i=1}^{L} S_i \left(\frac{\partial f}{\partial S_i}|_{(0,...0)}\right) + \sum_{i=1}^{L} X_i \left(\frac{\partial f}{\partial X_i}|_{(0,...0)}\right) + \sum_{i=1}^{L} A_i \left(\frac{\partial f}{\partial A_i}|_{(0,...0)}\right) \end{aligned} \tag{13}$$

Note that the first term, $f(0, ...0) = \log(Le^0) = \log L$. Therefore, the labor term which appeared to be missing from the specification of (11) is actually there, at least when considering the Taylor expansion. Now, examine the second term:

$$\begin{aligned} \sum_{i=1}^{L} S_i \left(\frac{\partial f}{\partial S_i}|_{(0,...0)}\right) &= \sum_{i=1}^{L} S_i \left[\left(\sum_{i=1}^{L} e^{\beta_S S_i+\beta_X X_i+\beta_A A_i}\right)^{-1} \left(e^{\beta_S S_i+\beta_X X_i+\beta_A A_i}\right)\beta_S\right]_{(0,0,.....0)} \\ &= \sum_{i=1}^{L} S_i \left(Le^0\right)^{-1}\left(e^0\right)\beta_S \\ &= \beta_S \bar{S}, \text{ where } \bar{S} \text{ is the average level of schooling of the } L \text{ workers.} \end{aligned}$$

The third and fourth terms of (13) are comparable, so that the production function of (12), using this first-order Taylor expansion becomes:

$$y = \beta_0 + \beta_1\beta_H + \beta_H \log L + \beta_H\beta_S\bar{S} + \beta_H\beta_X\bar{X} + \beta_H\beta_A\bar{A} + \beta_K k + \beta_\omega(M + Q) + \varepsilon \quad (14)$$

Note that in the production function estimation method of Olley and Pakes (1996) and Levinsohn and Petrin (2000), the simultaneity in the production function is handled by controlling for that portion of the productivity which is seen by the firm manager and revealed through the firm's investment or the firm's use of intermediate inputs, conditional on the firm's capital stock. By that definition, the ability of workers at a firm is clearly part of the firm's productivity. Therefore, we can reorganize equation (14) and define firm productivity as $\omega = \beta_H\beta_A\bar{A} + \beta_\omega(M + Q)$. Then equation (14) can now be expressed in a form that makes clear how the production function estimation will proceed (letting $\beta_0^* = \beta_0 + \beta_1\beta_H$):

$$y_{ft} = \beta_0^* + \beta_H \log L_{ft} + \beta_H\beta_S\bar{S}_{ft} + \beta_H\beta_X\bar{X}_{ft} + \beta_K k + \omega_{ft} + \varepsilon_{ft} \quad (15)$$

Once the estimation of (15) is complete, the estimates of ω_{ft} can be decomposed (details later) in order to achieve an estimate of the average worker ability at the firm, in order to control for ability in the wage equation estimation. Note that the complicated term which we have defined as f can also be approximated by a second- or higher-order Taylor expansion. Details of these expansions, and a discussion of their estimation are provided in Appendix 2. In general, these higher-order Taylor expansions can easily be used for production function estimation. However, in our case, we wish to not only estimate the production function, but obtain a measure of worker ability from this estimation. The first-order Taylor expansion is the only order for which worker ability can be separated from the other components of firm productivity, and therefore equation (15) is used for estimation.

3.2 Estimating the Production Function

The key component of the production function estimation is the handling of firm productivity, ω. Levinsohn and Petrin (2000) advocate replacing the use of investment in the estimation procedure with intermediate inputs. To consider the estimation procedure, begin by considering a firm's intermediate input function, $i_{ft} = i(\omega_{ft}, k_{ft})$. It should be noted that simply writing this function, $i(\omega_{ft}, k_{ft})$, assumes that firms face the same input prices.[14] While these prices affect the input demand functions, given that the prices are common across firms, we can estimate the function using only the two state variables

[14] When Olley and Pakes (1996) define this function, they use the index t, to explicitly reflect the fact that the input prices are the same in a given period, i.e. $i_t(\omega_{ft}, k_{ft})$. The lengths of the time periods that they use are either three or four years. Given that the entire length of this study's dataset is four years, we are using a single period, and therefore drop the t subscript at the outset, for clarity.

noted.[15] If intermediate input use is monotonically increasing in productivity, conditional on the level of the capital stock, and Levinsohn and Petrin (2000) provide sufficient conditions under which this is true, then the intermediate input function can be inverted to obtain an expression for productivity: $\omega_{ft} = \omega(i_{ft}, k_{ft})$.

Equation (15) can then be rewritten as the following:

$$\begin{aligned} y_{ft} &= \beta_H l_{ft} + \beta_S \beta_H \bar{S}_{ft} + \beta_X \beta_H \bar{X}_{ft} + \phi(i_{ft}, k_{ft}) + \eta_{ft} \\ \phi(i_{ft}, k_{ft}) &= \beta_0 + \beta_k k_{ft} + \omega(i_{ft}, k_{ft}) \end{aligned} \tag{16}$$

Writing the production function in the form of equation (16) not only separates those things that depend on the intermediate input and the capital stock from those that do not, it also separates the freely variable inputs from the productivity and capital stock. Note that schooling and experience in this formulation are treated as freely-variable inputs, just as the labor variable input traditionally is. Given that the schooling and experience levels of the firm change with the hiring and firing of employees, it seems more natural to treat these variables as freely variable rather than quasi-fixed. While this specification does not allow for the tenure of employees to play a role, this is primarily to keep consistency with the determinants of worker productivity in the Mincer human capital model.

Equation (16) is a partially linear model, which can be estimated semiparametrically using a variety of methods to get consistent estimates for the variable-input coefficients. Following the method of Robinson (1988), taking the expectation of equation (16) conditional on i_t, k_t yields:

$$E[y_{ft}|i_{ft}, k_{ft}] = E[l_{ft}|i_{ft}, k_{ft}]\beta_H + E[\bar{S}_{ft}|i_{ft}, k_{ft}]\beta_S\beta_H + E[\bar{X}_{ft}|i_{ft}, k_{ft}]\beta_X\beta_H + \phi(i_{ft}, k_{ft}) \tag{17}$$

since i) $E[\eta_{ft}|i_{ft}, k_{ft}] = 0$, and ii) $E[\phi_t(i_{ft}, k_{ft})|i_{ft}, k_{ft}] = \phi(i_{ft}, k_{ft})$. In the estimation procedure, these conditional expectations are calculated using kernel density estimation with

[15] While this result follows directly from Olley and Pakes (1996) and Levinsohn and Petrin (2000), given the way in which we have decomposed the productivity term, ω, a further note is in order. In so far as the intermediate input use is a function of the state variables, ω and k, which evolve over time, and worker ability forms one component of ω, worker ability also evolves over time, even for a given set of workers. Note that this can happen in one of two ways. First, the ability of workers can improve over time at a firm. It does not need to be fixed, although ability bias arises because it does contain a component which is correlated with schooling, which is fixed for a given worker. Second, the true "ability" of the workers may remain constant, but this ability may be slowly revealed to the firm manager. Firm productivity is something which is acted upon by the firm manager, and revealed by the firm's use of intermediate inputs, conditional on the level of capital stock. If the firm manager is learning over time about the ability of his workers, then the ability that will be both measured in firm productivity and remunerated by the manager will be that portion of worker ability seen and acted upon by the manager. Again, this is the definition of worker ability that matters, as the manager will remunerate what the manager sees. Note, in contrast, that while the manager learns about the ability of his workers over time, the schooling of the workers is revealed before hiring, and does not change for a given set of workers.

a normal (Gaussian) kernel. Subtracting equation (17) from equation (16) gives:

$$y_{ft} - E[y_{ft}|i_{ft}, k_{ft}] = \tag{18}$$
$$(l_{ft} - E[l_{ft}|i_{ft}, k_{ft}])\beta_H + (\bar{S}_{ft} - E[\bar{S}_{ft}|i_{ft}, k_{ft}])\beta_S\beta_H + (\bar{X}_{ft} - E[\bar{X}_{ft}|i_{ft}, k_{ft}])\beta_X\beta_H + \eta_{ft}$$

As a result, running no-intercept OLS with the modified variables of equation (18) will provide a consistent estimate of β_H, $\beta_S\beta_H$, and $\beta_X\beta_H$. Once these parameters have been obtained, the contribution of the variable inputs can be subtracted from equation (16), giving a new dependent variable, y^*:

$$y^* = y_{ft} - \beta_H l_{ft} - \beta_S\beta_H\bar{S}_{ft} - \beta_X\beta_H\bar{X}_{ft} = \beta_0 + \beta_k k_{ft} + \omega_{ft} + \eta_{ft} \tag{19}$$

Therefore, in this first stage of the estimation procedure, the coefficients on the freely-variable inputs are obtained. In order to proceed, some minimal assumption is required on firm productivity, and following Olley and Pakes (1996) and Levinsohn and Petrin (2000), I assume that it follows a first-order Markov process,

$$\omega_{ft} = E[\omega_{ft}|\omega_{ft-1}] + \xi_{ft} \tag{20}$$

where ξ_{ft} is the mean zero innovation in ω_{ft}. The expectation term and the intercept, β_0 can be collected together into the function:

$$g(\omega_{ft-1}) = \beta_0 + E[\omega_{ft}|\omega_{ft-1}] \tag{21}$$

so that equation (19) can be rewritten as:

$$y^*_{ft} = \beta_k k_{ft} + g(\omega_{ft-1}) + (\xi_{ft} + \eta_{ft}) \tag{22}$$

Fortunately, using the coefficients obtained from the first stage of estimation (equation (18)) will provide an estimate of $g(\omega_{ft-1})$, as shall become evident below, and this can then be used in the estimation of equation (22). The restriction made for identification is common to the literature, namely that capital is slow to adjust to the innovations in productivity that are unexpected:

$$E[\xi_{ft} + \eta_{ft}|k_{ft}] = 0 \tag{23}$$

This restriction is not terribly restrictive once the definitions are understood. Given that ξ_{ft} is the *surprise* in productivity this period (from (20)) and η_{ft} is the noise component of the residual, each of these components should be uncorrelated with the capital stock this period. This moment condition is used to estimate the parameter β_k, using Generalized Method of Moments, following these steps:

First, start by choosing a starting value for β_k, labelled β^*_k.

Substituting equations (20) and (21) into equation (19) yields, after substituting the estimates of the variable input coefficients:

$$y_{ft} - \widehat{\beta}_l l_{ft} - \widehat{\beta}_S \bar{S}_{ft} - \widehat{\beta}_X \bar{X}_{ft} - \beta_k^* k_{ft} - g(\omega_{ft-1}) = \xi_{ft} + \eta_{ft} \tag{24}$$

The information for all the left-hand side variables is available at this point of the procedure, as either data or already estimated coefficients, with the exception of $g(\omega_{ft-1})$, and specifically $E[\omega_{ft}|\omega_{ft-1}]$. In order to get an estimate of this, first note that, since $E[\eta_{ft}|i_{ft}, k_{ft}] = 0$, we find that $E[\omega_{ft}|\omega_{ft-1}] = E[\omega_{ft} + \eta_{ft}|\omega_{ft-1}] = E[\widehat{\omega_{ft} + \eta_{ft}}|\widehat{\omega_{ft-1}}]$. We have estimates of these components, as $\widehat{\omega_{ft} + \eta_{ft}} = y_{ft} - \widehat{\beta}_l l_{ft} - \beta_k^* k_{ft}$, and $\widehat{\omega}_{ft=1} = \widehat{\phi}_{ft-1} - \widehat{\beta}_k^* k_{ft-1}$. So, regress, nonparametrically, $\widehat{\omega_{ft} + \eta_{ft}}$ on $\widehat{\omega}_{ft-1}$ to estimate $E[\omega_{ft}|\omega_{ft-1}]$. This provides all the information required to use the moment condition of (23) to estimate β_k.

To summarize to this point, the production function, incorporating the heterogeneity of labor, has been estimated. The problem of simultaneity in the production function has been carefully controlled by allowing the use of intermediate inputs, conditional on a firm's capital stock, to reveal to the econometrician the firm's productivity each period. The additional assumption required was that firm productivity follows a Markov process. With this estimate of firm productivity in hand, one component of which is worker ability, the control for ability bias in the wage equation can be achieved.

4 Controlling for Ability Bias in the Wage Equation

To this point, a firm production function has been estimated. In the process of including the Mincerian components (schooling, experience, and ability) in the production function, the ability term has, through the estimation procedure, formed one component of the firm productivity term. Recall the expression for productivity in our firm production function:

$$\omega_{ft} = \beta_A \beta_H \bar{A}_{ft} + \beta_\omega M_{ft} + \beta_\omega Q_{ft} \tag{25}$$

Therefore, in order to use this measure of ability, it needs to be separated from the managerial component and the other component of firm productivity, which is the task of this section.

To maintain consistency with the production function estimation, recall that firm productivity follows a Markov process: $\omega_{ft} = E[\omega_{ft}|\omega_{ft-1}] + \xi_{ft}$. Substituting and assuming that the conditional expectations are separable implies: $\omega_{ft} = \beta_A \beta_H E(\bar{A}_{ft}|\bar{A}_{ft-1}) + \beta_\omega E(M_{ft}|M_{ft-1}) + \beta_\omega E(Q_{ft}|Q_{ft-1}) + \xi_{ft}$.[16]

Then, to solve for ability explicitly, with $\bar{A}_{ft} = E(\bar{A}_{ft}|\bar{A}_{ft-1}) + \psi_{ft}$ gives:

$$\bar{A}_{ft} = \frac{1}{\beta_A \beta_H} \left[\omega_{ft} - \beta_A \beta_H \psi_{ft} - \beta_\omega E(M_{ft}|M_{ft-1}) - \beta_\omega E(Q_{ft}|Q_{ft-1}) - \xi_{ft}\right] \tag{26}$$

[16] As discussed in the previous section, worker ability, or equivalently, the manager's perception of it, evolves over time, even for a given set of workers.

This provides an expression for the average ability level of workers in a firm in a given year. However, recall that the wage equation from the human capital model is an individual-level equation, specifically:

$$\log W_{ift} = \lambda_0 + \lambda_S \bar{S}_{ift} + \lambda_X \bar{X}_{ift} + \lambda_A \bar{A}_{ift} + \chi_{ift} \tag{27}$$

In order to be able to use the expression for ability from equation (26) in the earnings equation, equation (27) needs to be estimated as a grouped regression, where the grouping is at the level of the firm. The grouped regression will still provide consistent, if inefficient, estimation of the model's parameters, in particular the returns to education, λ_S.[17] Averaging across workers at a firm to run a grouped regression, and defining $w_{ft} = \frac{\sum_{i=1}^{L_{ft}} (\log W_{ift})}{L_{ft}}$ gives:

$$w_{ft} = \lambda_0 + \lambda_S \bar{S}_{ft} + \lambda_X \bar{X}_{ft} + \lambda_A \bar{A}_{ft} + \chi_{ft} \tag{28}$$

Equation (28) is now in a format that can use the measure of worker ability captured from the firm production function. The simplest way to isolate the worker ability is simply to assume that the dominant component of firm productivity (from (25)) is worker ability, and therefore to use the estimate of firm productivity as the control for worker ability in the wage regression, as in:

$$w_{ft} = \lambda_0 + \lambda_S \bar{S}_{ft} + \lambda_X \bar{X}_{ft} + \lambda_\omega \omega_{ft} + \chi_{ft} \tag{29}$$

If the worker ability component dominates equation (25), then this will provide consistent estimates of the parameters. However, if this is not the case, then it is worth attempting to analyze how worker ability might be separated from the other components of firm productivity, if they matter. Then substituting equation (26) into equation (28) gives:

$$\begin{aligned} w_{ft} &= \lambda_0 + \lambda_S \bar{S}_{ft} + \lambda_X \bar{X}_{ft} \\ &+ \frac{\lambda_A}{\beta_A \beta_H} \left[\omega_{ft} - \beta_A \beta_H \psi_{ft} - \beta_\omega E(M_{ft}|M_{ft-1}) - \beta_\omega E(Q_{ft}|Q_{ft-1}) - \xi_{ft} \right] + \chi_{ft} \end{aligned} \tag{30}$$

Now, define a composite error term, $\tau_{ft} = -\lambda_A \beta_A \beta_H (\psi_{ft} - \psi_{ft-1}) - \lambda_A (\xi_{ft} - \xi_{ft-1}) + \chi_{ft} - \chi_{ft-1}$, and take first differences to get:

$$\begin{aligned} & w_{ft} - w_{ft-1} = \lambda_S (\bar{S}_{ft} - \bar{S}_{ft-1}) + \lambda_X (\bar{X}_{ft} - \bar{X}_{ft-1}) + \frac{\lambda_A}{\beta_A \beta_H} (\omega_{ft} - \omega_{ft-1}) \\ & -\beta_\omega \lambda_A (E[M_{ft}|M_{ft-1}] - E[M_{ft-1}|M_{ft-2}]) - \beta_\omega \lambda_A (E[Q_{ft}|Q_{ft-1}] - E[Q_{ft-1}|Q_{ft-2}]) + \tau_{ft} \end{aligned} \tag{31}$$

Before estimating the above expression, a further assumption is required, in particular on the third component of productivity, Q_{ft}. Recall that this term is designed to capture that

[17] This assumes that the χ_i is not dependent on worker characteristics, as the unobservable idiosyncratic component for a given worker is captured in the ability term.

portion of firm productivity which is not captured by either worker ability or management ability. It includes changes in technology, as well as idiosyncratic productivity shocks, such as fire or theft. This process is assumed to be trend stationary (this includes the case of i.i.d. shocks as a special case). Therefore, $E[Q_{ft}|Q_{ft-1}] - E[Q_{ft-1}|Q_{ft-2}] = q$, a constant. Let $\lambda_0^* = -q\beta_\omega\lambda_A$, and then equation (31) becomes:

$$\begin{aligned} w_{ft} - w_{ft-1} &= \lambda_0^* + \lambda_S(\bar{S}_{ft} - \bar{S}_{ft-1}) + \lambda_X(\bar{X}_{ft} - \bar{X}_{ft-1}) \\ &\quad + \frac{\lambda_A}{\beta_A\beta_H}(\omega_{ft} - \omega_{ft-1}) - \beta_\omega\lambda_A(E[M_{ft}|M_{ft-1}] - E[M_{ft-1}|M_{ft-2}]) + \tau_{ft} \end{aligned} \tag{32}$$

At this stage, equation (32), which appears quite straightforward can be estimated in one of two ways. The first is to estimate directly the management component, M_{ft}, using a fairly rich set of observable factors on the firm managers in the data, and calculate the conditional expectations required for estimation in (32) nonparametrically. However, since the firm manager rarely changes in the dataset used, a more reasonable approach in this case is simply to assume that the management ability, or management talent as it is sometimes referred to, is fixed for a given firm manager. Given that the managers rarely change in the dataset, this basically amounts to assuming that the management talent is fixed for a given firm. Then, equation (32) simplifies to:

$$w_{ft} - w_{ft-1} = \lambda_0^* + \lambda_S(\bar{S}_{ft} - \bar{S}_{ft-1}) + \lambda_X(\bar{X}_{ft} - \bar{X}_{ft-1}) + \frac{\lambda_A}{\beta_A\beta_H}(\omega_{ft} - \omega_{ft-1}) + \tau_{ft} \tag{33}$$

This equation, and equation (29) are taken to data to estimate the rate of return to education.

5 Data

The data for the study is the Ghana Manufacturing Enterprise Survey (GMES), a survey of manufacturing firms in Ghana, conducted by Oxford University in partnership with the Ghana Statistical Service, covering the years from 1994 to 1997. The four major manufacturing industries in Ghana, namely woodworking, metalworking, food processing, and textiles and garments, are included in the sample, with roughly 200 firms sampled each year. Some exit and attrition from the sample occurred, and these firms were replaced with firms randomly selected from firms of similar characteristics, so that near (but not exactly) 200 firms were sampled in each year. Overall then, the dataset is an unbalanced panel. Moreover, a sample of 10 workers and 10 apprentices at each firm was surveyed (or all of the firm's workers/apprentices, in the case of less than 10 employees/apprentices), thus creating a linked employer-employee dataset, allowing the use of the techniques described in the methodology section. Information on firm output prices and input costs are used to construct firm-level price and input price deflators. Earnings are deflated using the consumer price index.

In addition to the GMES, the fourth round of the Ghana Living Standards Survey (GLSS 4), a national household survey in Ghana, is also used in this study. Initially, the returns to education calculated through the procedure described in the earlier sections must be interpreted as the returns to education within the manufacturing sector in Ghana. In order to calculate the returns to education for the Ghanaian workforce as a whole, the selection of individuals into the manufacturing sector must be accounted for, with the GLSS data used to handle this issue.

Some summary statistics for both the workers and the firms are provided in Table 1. While the workers of the GMES have similar average characteristics to wage workers from the GLSS, as evident in the Table 1, both of these samples may suffer from selection problems when interpreting the results for Ghana as a whole. Even though the GLSS is a nationally representative household survey, part of the returns to education may be captured in the selection into wage work. The method for handling this will be discussed in the next section.

6 Results

The first step in our procedure to estimate consistently the returns to education involves estimating the production function. One of the key assumptions in this estimation procedure is that the raw materials use at the firm is a strictly increasing function of firm productivity, conditional on the level of the capital stock at the firm. While this assumption is required for the procedure to be consistent, the estimation procedure does not impose this assumption directly at any point. Therefore, we can in fact examine whether this assumption holds in our dataset. To do this, a nonparametric, kernel density (using a normal kernel) regression of materials on firm productivity and capital stock is performed. The result of this regression is presented in Figure 1, where it appears that indeed, for a given level of the capital stock, materials is virtually always increasing in the level of productivity across the range of our data, providing some support for this crucial assumption.

The production function estimates are presented in Table 2. The results are generally precisely estimated, with the coefficients changing magnitude in accord with the presumed effect of the simultaneity biases mentioned earlier.[18] The first column presents the OLS results, and the second presents the results of the GMM estimation of Section 3.2. Given that the average levels of education and experience of the workers of the firm is measured as an average of the sub-sample of workers interviewed at the firm, each of these variables is expected to be measured with error, and therefore may provide biased estimates of the

[18] While constant returns to scale cannot be rejected at any conventional level of significance under a standard t-test, the coefficients on capital and labor do sum to less than 1. Other studies that have also had the coefficients on capital and labor summing to values less than 1 when calculating manufacturing sector production functions using Ghanaian data include Jones (2001), Teal (1995a), Teal (1995b), ISA Group (1998). In one of these studies, constant returns to scale could be rejected statistically (ISA Group, 1998), but not in the others.

desired coefficients. The standard procedure for handling measurement error is to obtain an instrument for each of the variables measured with error. Provided that the *measurement error* for the (averaged) education and experience are uncorrelated between periods, then the lagged values of education and experience can be used as instruments for these respective variables (Hsiao, 1986, p. 64). Note that in the estimation procedure in use, these instruments are not the means for handling the simultaneity problem, that is for handling the correlation between education and firm productivity, or experience and firm productvity. The procedure outlined in Section 3.2 is designed expressly for that purpose. Rather the instruments are simply used to overcome the measurement error.

The task of incorporating instrumental variables estimation with the procedure of Section 3.2 must be done carefully in order to obtain consistent parameter estimates. Fortunately, Li and Stengos (1996) provide a kernel-density based IV-estimator for semiparametric partially linear panel data models (such as equation (16)) and prove its consistency in obtaining the parametric coefficients of the model, which in our case are the freely-variable inputs of equation (16). For the IV results, we use their estimator for the first stage of estimation, to obtain the coefficients on labor, education and experience. The consistency for the second stage of each of the procedures in obtaining the coefficient on capital is provided by Pakes and Olley (1995). The results of this instrumental-variable version of the modified Levinsohn-Petrin procedure are provided in column (3). The only coefficient that changes significantly is the education variable which is substantially higher in the IV-GMM results.

Still, the primary purpose of the production function estimation is to obtain measures of firm productivity which can be used in the specifications of equations (29) and (33) to control for worker ability in the wage regressions. The results of the wage regressions of equations (29) and (33) are presented in Table 3 and Table 4. As previously mentioned, in order to make use of the measures of worker ability obtained from the firm production functions, the regressions are all grouped regressions, which do provide consistent (although inefficient) estimates of all parameters. Column (1) of Table 3 presents the least-squares estimates, that is the grouped regression version of equation (1), using a linear experience term. Column (2) uses equation (29), which assumes that firm productivity is predominantly a measure of worker ability. Inclusion of the firm productivity in column (2) reduces the least-squares coefficient on education by 8%. Column (3) has less restrictive assumptions, allowing for other components of firm productivity (including management ability), under the restrictions described in Section (4). Under these assumptions, the education coefficient estimate falls slightly further, to 0.104, 13% less than the least-squares estimate. While this may be indicative of a degree of ability bias,[19] it is worth examining the coefficient on experience as well, but before that a couple of notes are in order. Typically, the bias on the education coefficient resulting from omission of ability from the wage equation is

[19] The word indicative is carefully chosen here, as the education coefficients of column (1) and (3) are not statistically significantly different.

assumed to be an upward bias (as it was outlined to be in Section 2 in the bivariate case). However, given that the wage equation is a multivariate regression, the actual sign on the bias on the schooling or the experience coefficients resulting from the omission of ability is unknown. Therefore, the fact that the coefficient on experience is also falling from column (1) to column (3) could be a result of the removal of ability bias. It could also be related to the specification of experience. The typical specification for experience in the Mincer model is a quadratic. In this procedure, because the observations are grouped before estimation, and because the quadratic is a concave function involving two terms of the same variable (X and X^2), the problem (using concave functions) of calculating the function value of an average, as opposed to the average of the function values arises. To avoid this problem, a linear experience term was used. Therefore, the fall in the function value for experience from column (1) to column (3) could also reflect this limit on the specification. However, the falling experience coefficient could equally be the result of measurement error in the education and experience variables. Since equation (33) is a differenced specification, any measurement error bias is exacerbated relative to the least-squares estimates.[20] To address this problem, instrumental variable estimates are used, with lagged values of education and experience as the instruments, as they were for the production function estimation. As in the production function case, the instruments here are addressing the measurement error problem, and not the problem of ability bias, which is handled through the controls of equation (29) and (33).

Before examining the instrumental variable estimates, however, a further limitation of the data should be noted. The returns to education being measured in columns (1) to (3) are the returns found within the manufacturing sector. In order to interpret these returns more broadly, we must control for the selection bias of the sample, relative to the Ghanaian population as a whole. The technique used for this purpose is the two-step procedure of Heckman (1979). The dataset representative of the population as a whole is the Ghanaian Living Standards Survey (GLSS) Wave 4 (April 1998-March 1999). Given that this period is a year after the end of our panel, this may introduce some bias into the results. Unfortunately, the previous household survey in Ghana, the GLSS 3 was collected in 1991-1992, again not within the timeframe of our sample, and therefore is not superior. The exclusion restriction required for this selection procedure is that father's occupation and mother's education affect the selection of an individual into the manufacturing sector, but do not affect the wage directly. While family background variables are sometimes included in a wage regression, they are typically used as proxies for the worker's unmeasured ability. Given that direct controls for ability are included in the equation, it is more reasonable to exclude them from the wage equation in this context. The selection-corrected results

[20] See Griliches (1979) for an early discussion of the problem of exacerbating measurement error when variables are differenced to remove a fixed effect, in his case differencing between siblings to remove the family fixed effect.

are presented in the final three columns of Table 3, which are identical to the first three columns, except for the additional of the inverse-Mills ratio (λ) as the selection control. The test for significance of the λ variable is also a test for selection. The significance of the λ variable in columns (4) and (5) suggests that selection is a factor in the sample under examination. In these selection-corrected results, the coefficient on education is very similar to the uncorrected estimates. On the other hand, the experience coefficient is somewhat higher in columns (4) and (5).

All of the regressions of Table 3 are repeated in Table 4, using instrumental variables to address the problem of measurement error in the education and experience variables. With one small exception (the experience coefficient in column (4)), the education and experience coefficients are at least as high in the instrumented version, with the education and experience coefficients remaining significant in all columns except (3) and (6), the differenced regressions. Measuring returns to education in the manufacturing sector, without controlling for selection into the manufacturing sector, yields an estimate of 0.104 for the returns to education in both the uninstrumented and instrumented versions of Tables 3 and 4. Controlling for selection into the manufacturing sector yields point estimates of the returns to education of 0.100 and 0.108 in Tables 3 and 4, although the result is only significant in the uninstrumented version. The experience coefficient is not precisely estimated, but does not appear to be attenuated by the differencing, as we presumed might have been the case in the other differenced regressions.

The coefficient on firm productivity follows a similar pattern throughout the regressions. While it is significant in each of the undifferenced regressions (equation (29)), it falls to essentially zero in all of the differenced regressions (equation (33)). It is worth recalling the assumptions behind each of these regressions. In equation (29), firm productivity is assumed to be dominated by worker ability, without any other significant component. Equation (33) is more general, allowing for a firm-specific component to productivity, as well as a technology trend, and idiosyncratic shocks, but the differencing may remove an important part of the variable's signal. If worker ability varies considerably between firms, but does not vary much within a firm over time, then the differencing could remove most of the variation in ability, and make it difficult to obtain a signal from the remaining variation. In the specification of equation (29), both the portion of worker ability that is fixed at a firm (that does not vary over time), and the time-varying portion make up the productivity variable, whereas in equation (33), all that remains is the time-varying portion. The fairly limited number of observations in this dataset makes it difficult to measure the signal that remains after the differencing.

In addition to addressing the question of returns to education, the procedure outlined in this paper allows us to address other questions as well. In particular, if the firm is within-period profit maximizing, then the wage and productive returns to education and schooling should be the same ($\beta_S = \lambda_S$ and $\beta_X = \lambda_X$). Note that while λ_S and λ_X are the coefficients on education and experience in the wage equations, and therefore are the estimates presented in Tables 3 and 4, the coefficients on education and experience in the production function estimates are $\beta_H\beta_S$ and $\beta_H\beta_X$, respectively. In order to isolate β_S and β_X, we need to divide these coefficients by the coefficient on labor, β_H. The standard error of this estimate is bootstrapped, just as the standard errors of β_H and $\beta_S\beta_H$ are bootstrapped, using a block bootstrap which assumes independence across firms, but not within firms across time. Unfortunately, the resulting standard errors on the estimates of β_S and β_X are too large too make any conclusions regarding the equality of β_S and λ_S, or the equality of β_X and λ_X. Their equality cannot be rejected using the estimates achieved, but the standard errors on β_S and β_X are large enough to not reject equality with a very large range of values.

7 Conclusion

This paper has outlined a new technique for controlling for ability bias in measuring returns to education. The consistent measurement of returns to education is clearly an important economic question, particularly so in a developing country like Ghana, and so hopefully this is an important advance. The use of information about the firms where workers work to acquire information about the workers themselves, and in this case their ability, is a novel approach to the issue of ability bias. Moreover, in the process, we have developed the unique form for the labor term of the production function that is consistent with the Mincerian human capital earnings function. The preferred results from this estimation procedure, at least in terms of the least restrictive assumptions imposed, suggest a point estimate for the returns to an additional year of schooling of roughly 10%, both within the manufacturing sector, and once selection into the manufacturing sector is addressed. Unfortunately, the standard errors prevent any firm conclusions regarding the comparisons between the productive contribution of an additional year of schooling or experience and its remuneration.

Fortunately, the recent expansion in the availability of linked employer-employee datasets has increased the potential to use firm information to answer a wide variety of questions related to workers. In fact, labor economists have been the first to make extensive use of these linked datasets to examine questions of interest to them. However, linked datasets are also of use to those studying the firm. This paper has demonstrated the means by which the wage equation resulting from the dominant model of the labor literature, the Mincer human capital model, can be made consistent with, and incorporated into the production function of the firm. While this technique should expand the potential areas of inquiry, hopefully it will spark further research into the best use of information about a firm's workers, in terms of its contribution to the understanding of the firm. In the best case, linked employer-employee datasets will not only provide new insights into old questions (as they already have done in numerous recent papers exploiting their potential), but will also allow us to examine new questions for which the lack of appropriate data has to this point forced empirical research to be silent.

References

[1] Angrist, Joshua D. and Alan B. Krueger,"Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106(4), November 1991, pp. 979-1014.

[2] Appleton, Simon, John Hoddinott, and John Mackinnon, "Education and Health in Sub-Saharan Africa," *Journal of International Development*, 8(3), 1996, pp. 307-339.

[3] Armitage, J. and R. Sabot, "Socioeconomic Background and the Returns to Schooling in Two Low-Income Countries," *Economica*, 54, February 1987, pp. 103-108.

[4] Ashenfelter, Orley and Alan Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review*, 84(5), Dec. 1994, pp. 1157-1173.

[5] Ashenfelter, Orley and Cecilia Rouse, "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins," *Quarterly Journal of Economics*, 113, February 1998, pp. 253-284.

[6] Ashenfelter, Orley and David Zimmerman, "Estimates of the Return to Schooling from Sibling Data: Fathers, Sons and Brothers," *Review of Economics and Statistics*, 79, 1997, pp. 1-9.

[7] Becker, Gary S. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press, 1964.

[8] Behrman, Jere R., "Human Capital Formation, Returns and Policies: Analytical Approaches and Research Questions," *Journal of International Development*, 8(3), 1996, p. 341-373.

[9] Behrman, Jere R., Andrew D. Foster, Mark R. Rosenzweig, Prem Vashishtha, "Women's Schooling, Home Teaching, and Economic Growth," *Journal of Political Economy*, 107(4), August 1999, pp. 682-714.

[10] Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman, "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment," *Journal of Political Economy*, 102(6), Dec. 1994, pp. 1131-1174.

[11] Behrman, Jere R. and Anil B. Deolalikar, "School Repetition, Dropouts, and the Rates of Return to Schooling: the Case of Indonesia," *Oxford Bulletin of Economics and Statistics*, 53(4), 1991, pp. 467-480.

[12] Behrman, Jere and Paul Taubman, "Intergenerational Transmissions of Income and Wealth," *American Economic Review - Papers and Proceedings*, 66(2), May 1976, pp. 436-440.

[13] Bigsten, Arne, Paul Collier, Stefan Dercon, Marcel Fafchamps, Bernard Gauthier, Jan Willem Gunning, Anders Isaksson, Abena Oduro, Remco Oostendorp, Cathy Patillo, Mans Soderbom, Francis Teal, Albert Zeufack, "Contract Flexibility and Dispute Resolution in African Manufacturing", *Journal of Development Studies*, 36(4): 1-17, April 2000.

[14] Blackburn, McKinley L., and David Neumark, "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look," *Review of Economics and Statistics*, 77(2), May, 1995, pp. 217-230.

[15] Boissiere, M., J. B. Knight, and R. H. Sabot, "Earnings, Schooling, Ability, and Cognitive Skills," *American Economic Review*, 75(5), December 1985, pp. 1016-1030.

[16] Bound, John and Gary Solon, "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling," *Economics of Education Review*, 18, 1999, pp. 169-182.

[17] Brunello, Giorgio, and Raffaele Miniaci, "The Economic Returns to Schooling for Italian Men. An Evaluation Based on Instrumental Variables," *Labour Economics*, 6, 1999, pp. 509-519.

[18] Callan, Tim and Colm Harmon, "The Economic Return to Schooling in Ireland," *Labour Economics*, 6, 1999, pp. 543-550.

[19] Card, David, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), September 2001, pp. 1127-1160.

[20] Card, David, "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. 3. New York: Elsevier Science, 1999.

[21] Glewwe, Paul. *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana*. Houndmills, U.K.: Macmillan Press, 1999.

[22] Glewwe, Paul, "The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment," *Journal of Development Economics*, 51, 1996, pp. 267-290.

[23] Godana, Tekaligne, "Returns to Education in the Manufacturing Sector in Zimbabwe: Some Empirical Evidence," *South African Journal of Economics*, 65(1), March 1997, pp. 99-113.

[24] Griliches, Zvi, "Sibling Models and Data in Economics: Beginnings of a Survey," *Journal of Political Economy* (Supplement), 87, 1979, pp. S37-S64.

[25] Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45, 1977, pp. 1-22.

[26] Griliches, Zvi, "Wages of Very Young Men," *Journal of Political Economy*, 84(4), Part 2, August 1976, pp. S69-S86.

[27] Griliches, Zvi and Jacques Mairesse, "Production Functions: The Search for Identification," National Bureau of Economic Research Working Paper No. 5067, March 1995.

[28] Griliches, Zvi and William M. Mason, "Education, Income and Ability," *Journal of Political Economy*, 80(3), Part 2, 1972, pp. S74-S103.

[29] Harmon, Colm and Ian Walker, "Estimating the Economic Return to Schooling for the United Kingdom," *American Economic Review*, 85, 1995, p. 1278-1286.

[30] Heckman, James J. and V. Joseph Hotz, "An Investigation into the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality," *Journal of Human Resources*, 21(4), 1986, pp. 507-542.

[31] Heckman, James J., "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 1979, pp. 153-161.

[32] Hellerstein, Judith K., David Neumark, and Kenneth R. Troske, "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Functions," *Journal of Labor Economics*, 1999, 17(3), 409-446.

[33] Hsiao, Cheng. *Analysis of Panel Data*. New York: Cambridge University Press, 1986.

[34] Ichino, Andrea and Rudolf Winter-Ebmer, "Lower and Upper bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments," *European Economic Review*, 43, 1999, pp. 889-901.

[35] Imbens, G. and J.D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 1994, pp. 467-475.

[36] ISA Group, "Exports, Firm-level Efficiency and Sources of Comparative Advantage in the African Manufacturing Sector," RPED Discussion Paper No. 97, World Bank, May 1998.

[37] Jones, Patricia, "Are Educated Workers Really More Productive?" *Journal of Development Economics*, 64, 2001, pp. 57-79.

[38] Kane, Thomas J. and Cecilia Elena Rouse, "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?" NBER Working Paper #4268, January 1993, pp. 46.

[39] Knight, John B. and Richard H. Sabot. *Education, Productivity, and Inequality: The East African Natural Experiment*. New York: Oxford University Press, 1990.

[40] Lanzona, Leonardo A., "Migration, Self-Selection and Earnings in Philippine Rural Communities," *Journal of Development Economics*, 56, 1998, pp. 27-50.

[41] Levin, Jesse and Erik J.S. Plug, "Instrumenting Education and the Returns to Schooling in the Netherlands," *Labour Economics*, 6, 1999, pp. 521-534.

[42] Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," NBER Working Paper 7819, 2000.

[43] Li, Qi and thanasis Stengos, "Semiparametric Estimation of Partially Linear Panel Data Models," *Journal of Econometrics*, 71, 1996, pp. 389-397.

[44] Li, Tianyou and Junsen Zhang, "Returns to Education under Collective and Household Farming in China," *Journal of Development Economics*," 56, 1998, pp. 307-335.

[45] Maluccio, John, "Endogeneity of Schooling in the Wage Function: Evidence from the Rural Philippines," IFPRI Food Consumption and Nutrition Division Discussion Paper No. 54, November 1998, pp. 53.

[46] Marschak, J. and W. H. Andrews, "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 1944, pp. 143-205.

[47] Mwabu, Germano. and T. Paul Schultz, "Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa," *American Economic Association Papers and Proceedings*, 86(2), 1996, pp. 335-339.

[48] Mwabu, Germano. and T. Paul Schultz, "Wage Premiums for Education and Location of South African Workers, by Gender and Race," *Economic Development and Cultural Change*, 48(2), January 2000, pp. 307-334.

[49] Neumark, David, "Biases in Twin Estimates of the Return to Schooling," *Economics of Education Review*, 18, 199, pp. 143-148.

[50] Nielsen, Helen Skyt and Niels Westergård-Nielsen, "Returns to Schooling in Less Developed Countries: New Evidence from Zambia," *Economic Development and Cultural Change*, January 2001, pp. 364-394.

[51] Olley, Steve and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, 64(6), 1263-97.

[52] Pakes, Ariel and Steve Olley, "A Limit Theorem for a Smooth Class of Semiparametric Estimators," *Journal of Econometrics*, 65, 1995, pp. 295-332.

[53] Ram, Rati and Ram D. Singh, "Farm Households in Rural Burkina Faso: Some Evidence on Allocative and Direct Return to Schooling, and Male-Female Labor Productivity Differentials," *World Development*, 16(3), 1988, pp. 419-424.

[54] Rosenzweig, Mark R. and Kenneth I. Wolpin, "Natural 'Natural Experiments' in Economics," *Journal of Economic Literature*, 38(4), December 2000, pp. 827-874.

[55] Rouse, Cecilia Elena, "Further Estimates of the Economic Return to Schooling from a New Sample of Twins," *Economics of Education Review*, 18, 1999, pp. 149-157.

[56] Schafgans, Marcia M.A., "Gender Wage Differences in Malaysia: Parametric and Semparametric Estimation," *Journal of Development Economics*, 63, 2000, pp. 351-378.

[57] Schultz, T. Paul. 1995. "Human Capital and Economic Development," in G.H. Peters and Douglas D. Hedley, eds., *Agricultural Competitiveness: Market Forces and Policy Choice–Proceedings of the Twenty-Second International Conference of Agricultural Economists*. Aldershot, U.K.: Dartmouth Publishing Company, p. 523-539.

[58] Schultz, T. Paul. 1988. "Education Investment and Returns," in H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics*, vol. 1. Amsterdam: North Holland Pub. Co.

[59] Siphambe, Happy Kufigwa, "Rates of Return to Education in Botswana," *Economics of Education Review*, 19, 2000, pp. 291-300.

[60] Smith, Paul A. and Michael R. Metzger, "The Return to Education: Street Vendors in Mexico," *World Development*, 26(2), 1998, pp. 289-296.

[61] Strauss, John and Duncan Thomas, "Human Resources: Empirical Modeling of Household and Family Decisions," in Jere R. Behrman and T. N. Srinivisan, eds., *Handbook of Development Economics*, vol. 3. New York: Elsevier Science, 1995.

[62] Teal, Francis, "Efficiency Wages and Rent Sharing: A Note and Some Empirical Evidence," RPED Discussion Paper No. 59, World Bank, 1995a.

[63] Teal, Francis, "Does 'Getting Prices Right' Work? Micro Evidence from Ghana," RPED Discussion Paper No. 58, World Bank, 1995b.

[64] Temple, Jonathan R., "Generalizations that Aren't? Evidence on Education and Growth," *European Economic Review*, 45, 2001, pp. 905-918.

[65] Uusitalo, Roope, "Return to Education in Finland," *Labour Economics*, 6, 1999, pp. 569-580.

[66] Vijverberg, Wim P. M., "Educational Investments and Returns for Women and Men in Côte d'Ivoire," *Journal of Human Resources*, 28(4), Fall 1993, pp. 933-974.

Appendix 1 - Developing the Unique Mincer-consistent Production Function

Each individual worker belongs to a type j that possess schooling S_j, experience X_j, and ability A_j. Individual types have a productive capacity according to their schooling, experience, and ability, as defined by the function $H(S_j, X_j, A_j)$. Wages for a given type j are given by the Mincerian wage formula (with a linear experience term in this case), so that $w_j = e^{\lambda_0 + \lambda_S S_j + \lambda_X X_j + \lambda_A A_j}$. Then, the firm's single-period profit-maximization problem is $(K, Q, M$ defined in the text):

$$\underset{\{\{L_j\},K\}}{Max} F[\sum L_j H(S_j, X_j, A_j), K, Q, M] - \sum L_j w_j - rK$$

The first-order condition for the number of each type chosen by the firm is the following: $F_1 H(S_j, X_j, A_j) = w_j$, if $L_j > 0$.

If j and k are two types chosen by the firm in positive quantity, then this first-order condition implies: $\dfrac{H(S_j, X_j, A_j)}{H(S_k, X_k, A_k)} = \dfrac{w_j}{w_k}$, for all j, k, for which $L_j > 0$, $L_k > 0$.

Consider two types j and k, which both have experience $\widetilde{X}$ and ability $\widetilde{A}$, and differ only by their level of schooling. Then $\dfrac{H(S_j, \widetilde{X}, \widetilde{A})}{H(S_k, \widetilde{X}, \widetilde{A})} = \dfrac{e^{\lambda_0 + \lambda_S S_j + \lambda_X \widetilde{X} + \lambda_A \widetilde{A}}}{e^{\lambda_0 + \lambda_S S_k + \lambda_X \widetilde{X} + \lambda_A \widetilde{A}}} = e^{\lambda_1 (S_j - S_k)}$.

Rewriting this equation gives: $H(S_j, \widetilde{X}, \widetilde{A}) = H(S_k, \widetilde{X}, \widetilde{A}) e^{\lambda_S (S_j - S_k)}$, which again, is true for all j, k for which $L_j > 0$, $L_k > 0$.

Differentiating with respect to S_k gives: $0 = \dfrac{\partial H(S_k, \widetilde{X}, \widetilde{A})}{\partial S_k} e^{\lambda_S (S_j - S_k)} - \lambda_S H(S_k, \widetilde{X}, \widetilde{A})$ $e^{\lambda_S (S_j - S_k)}$.

This simplifies to: $\dfrac{\partial H(S_k, \widetilde{X}, \widetilde{A})}{\partial S_k} = \lambda_S H(S_k, \widetilde{X}, \widetilde{A})$. This is a well-known differential equation, whose unique solution is $H(S_k, \widetilde{X}, \widetilde{A}) = C e^{\lambda_S S_k} f(\widetilde{X}, \widetilde{A})$, where C is a constant.

Following a similar argument for each of experience and ability will lead to:

$H(S_k, X_k, A_k) = C e^{\lambda_0 + \lambda_S S_k + \lambda_X X_k + \lambda_A A_k} = e^{\lambda_0^* + \lambda_S S_k + \lambda_X X_k + \lambda_A A_k}$, for all types k that the firm uses.

Appendix 2 - Higher-order Taylor Expansions of the Labor Term

The term $f(S_1, ...S_L, X_1, ...X_L, A_1, ...A_L) = \log(\sum_{i=1}^{L} e^{\beta_S S_i + \beta_X X_i + \beta_A A_i})$ can also be approximated by a second-order Taylor expansion. Consider such an expansion in the simpler case where workers possess only schooling and experience, but no separate ability. Then $f(S_1, ...S_L, X_1, ...X_L) = \log(\sum_{i=1}^{L} e^{\beta_S S_i + \beta_X X_i})$. Now a second-order Taylor expansion of f would be:

$$\begin{aligned} f(\quad & S_1, ...S_L, X_1, ...X_L) = f(\underbrace{0, 0, ...0}_{2L\ 0's}) + \sum_{i=1}^{L} S_i \left(\frac{\partial f}{\partial S_i}|_{(0,...0)} \right) + \sum_{i=1}^{L} X_i \left(\frac{\partial f}{\partial X_i}|_{(0,...0)} \right) \quad (34) \\ & + \frac{1}{2} \left[S_1 \tfrac{\partial}{\partial S_1} + S_2 \tfrac{\partial}{\partial S_2} + ... + S_L \tfrac{\partial}{\partial S_L} + X_1 \tfrac{\partial}{\partial X_1} + X_2 \tfrac{\partial}{\partial X_2} + ... + X_L \tfrac{\partial}{\partial X_L} \right]^2 |_{(0,...0)} f \end{aligned}$$

The first, second, and third terms are identical to their counterparts in the previous equation. The final term has been written in a shorthand notation. That is,

$$\begin{aligned} & \left[S_1 \tfrac{\partial}{\partial S_1} + S_2 \tfrac{\partial}{\partial S_2} + ... + S_L \tfrac{\partial}{\partial S_L} + X_1 \tfrac{\partial}{\partial X_1} + X_2 \tfrac{\partial}{\partial X_2} + ... + X_L \tfrac{\partial}{\partial X_L} \right]^2 |_{(0,...0)} f \quad (35) \\ = \ & \sum_{i=1}^{L} \sum_{j=1}^{L} S_i S_j \left(\tfrac{\partial^2 f}{\partial S_i \partial S_j}|_{(0,....0)} \right) + 2 \sum_{i=1}^{L} \sum_{j=1}^{L} S_i X_j \left(\tfrac{\partial^2 f}{\partial S_i \partial X_j}|_{(0,....0)} \right) + \sum_{i=1}^{L} \sum_{j=1}^{L} X_i X_j \left(\tfrac{\partial^2 f}{\partial X_i \partial X_j}|_{(0,....0)} \right) \end{aligned}$$

To evaluate this expression, consider first a term of the form $\frac{\partial}{\partial S_i} \frac{\partial f}{\partial S_j}|_{(0,....0)}$.

$$\begin{aligned} \frac{\partial^2 f}{\partial S_i \partial S_j}|_{(0,....0)} \quad & = \frac{\partial}{\partial S_i} \left[\left(\sum_{m=1}^{L} e^{\beta_S S_m + \beta_X X_m} \right)^{-1} \left(e^{\beta_S S_j + \beta_X X_j} \right) \beta_S \right]_{(0,...,0)} \\ & = \beta_S \left(e^{\beta_S S_j + \beta_X X_j} \right) (-1) \left(\sum_{m=1}^{L} e^{\beta_S S_m + \beta_X X_m} \right)^{-2} \left(e^{\beta_S S_i + \beta_X X_i} \right) \beta_S |_{(0,...,0)} \\ & = -\beta_S^2 \left(e^0 \right)^2 \left(L e^0 \right)^{-2} \\ & = -\frac{\beta_S^2}{L^2} \end{aligned}$$

Fortunately, therefore, this second partial derivative is independent of the schooling indices i and j, and can be factored out of the double summation. Similarly, one can show that $\frac{\partial^2 f}{\partial S_i \partial X_j}|_{(0,....0)} = -\frac{\beta_S \beta_X}{L^2}$ and $\frac{\partial^2 f}{\partial X_i \partial X_j}|_{(0,....0)} = -\frac{\beta_X^2}{L^2}$. Therefore, substituting these results into (35), into (34), into a modified version of (12) (removing ability), and redefining the constant as $\beta_0^* = \beta_0 + \beta_1 \beta_H$ gives:

$$\begin{aligned} & y = \beta_0^* + \beta_H \log L + \beta_H \beta_S \bar{S} + \beta_H \beta_X \bar{X} \quad (36) \\ & -\frac{\beta_H \beta_S^2}{2L^2} \sum_{i=1}^{L} \sum_{j=1}^{L} S_i S_j - \frac{\beta_H \beta_S \beta_X}{L^2} \sum_{i=1}^{L} \sum_{j=1}^{L} S_i X_j - \frac{\beta_H \beta_X^2}{2L^2} \sum_{i=1}^{L} \sum_{j=1}^{L} X_i X_j + \beta_K k + \beta_\omega (M + Q) + \varepsilon \end{aligned}$$

Furthermore, the double summations in (36) can be further simplified. In particular, $\sum_{i=1}^{L} \sum_{j=1}^{L} S_i S_j = L^2 \bar{S}^2$, $\sum_{i=1}^{L} \sum_{j=1}^{L} S_i X_j = L^2 \bar{S} \bar{X}$, and $\sum_{i=1}^{L} \sum_{j=1}^{L} X_i X_j = L^2 \bar{X}^2$. Therefore, equation

(36) can be further simplified to:

$$\begin{aligned} y &= \beta_0^* + \beta_H \log L + \beta_H \beta_S \bar{S} + \beta_H \beta_X \bar{X} - \frac{\beta_H \beta_S^2}{2} \bar{S}^2 - \beta_H \beta_S \beta_X \bar{S} \bar{X} - \frac{\beta_H \beta_X^2}{2} \bar{X}^2 \\ &\quad + \beta_K k + \beta_\omega \omega + \varepsilon \end{aligned} \tag{37}$$

Clearly, the second-order Taylor approximation of equation (37) can be estimated as easily as the first-order Taylor approximation of equation (15). However, ability is missing from the above estimation. A full second-order Taylor approximation including schooling, experience, and ability would be, after simplifying the summations:

$$\begin{aligned} y &= \beta_0^* + \beta_H \log L + \beta_H \beta_S \bar{S} + \beta_H \beta_X \bar{X} + \beta_H \beta_S \bar{A} - \frac{\beta_H \beta_S^2}{2} \bar{S}^2 - \frac{\beta_H \beta_X^2}{2} \bar{X}^2 - \frac{\beta_H \beta_A^2}{2} \bar{A}^2 \\ &\quad - \beta_H \beta_S \beta_X \bar{S} \bar{X} - \beta_H \beta_S \beta_A \bar{S} \bar{A} - \beta_H \beta_X \beta_A \bar{X} \bar{A} + \beta_K k + \beta_\omega (M + Q) + \varepsilon \end{aligned} \tag{38}$$

While it is clear how (37) could be estimated, it is not quite as clear how (38) could be estimated, given that ability is unobserved. However, recall our definition of productivity, specifically that portion of the residual which is seen and acted upon by the firm manager, but not viewed, at least initially, by the econometrician. Given that the distribution of the A_i's (and therefore $\bar{A}$) is unobserved, the interaction terms $\bar{S}\bar{A}$ and $\bar{X}\bar{A}$, as well as $\bar{A}^2$, are unobserved, and therefore will be part of the productivity term in the usual procedure, provided that they are viewed by the firm manager. Therefore, even with unobserved ability, the production function can be estimated with a specification similar to equation (37).

Third and higher orders of Taylor expansion follow naturally. It is possible to test for the appropriate order of expansion. Furthermore, note that increasing the order of the Taylor expansion does not increase the number of coefficients to be estimated (at least when they are estimated in the restricted, as opposed to the reduced form), making possible the computation of higher-order Taylor expansions on relatively small datasets.

The technique outlined above is a fully generalizable method for making the production function consistent with the wage equation being estimated. Provided that the dependent variable in the wage equation is in logarithmic form, and the right-hand side of the wage equation is linear in the broadest sense, then a Taylor expansion of whatever order is desired can be used to calculate the productive contribution of the factors of the wage equation. Moreover, there will be a one-to-one correspondence between the coefficients of any specific variable in the wage equation and its partner coefficient in the production function. The number of parameters in the production function is therefore exactly increased by the number of parameters in the wage equation (not including the constant, which cannot be separately identified from the original constant in the production function). In fact, each of the production function coefficients can be compared to the coefficients of the wage equation to test whether that factor is being remunerated according to its contribution to firm product. This technique can therefore be applied in a variety of settings.

Appendix 3 - The Data

First, note that all of the monetary values have been converted into constant 1991 cedis (unit of Ghanaian currency). The value-added was deflated using firm-specific price and cost deflators, calculated using information in the survey. The capital stock was deflated using an investment deflator, calculated as a weighted average of the urban CPI (0.25) and the U.S. Dollar exchange rate (0.75). The wage values are deflated by the urban CPI.

Value-added is calculated as the value of sales less material input costs less indirect costs.

The capital stock is the replacement value of plant and equipment.

Worker Averages:

These averages are calculated as in Bigsten et. al. (2000). The averages of the worker variables (by firm) are calculated from the individual-level data. The occupational composition of the workforce was available from the firm surveys. From interviews with the employees, the years of education, tenure, and age were determined by occupational classification. The weighted average of the worker variables is then calculated from this information, with the weights being the proportions of the workforce in each occupation. If there is no worker-level information for an occupation that exists for the firm, we use the averages for that occupational classification to fill in the missing observations.

GAUSS Thu Sep 13 13:20:45 2001

Figure 1 - Materials Function for Manufacturing Industry

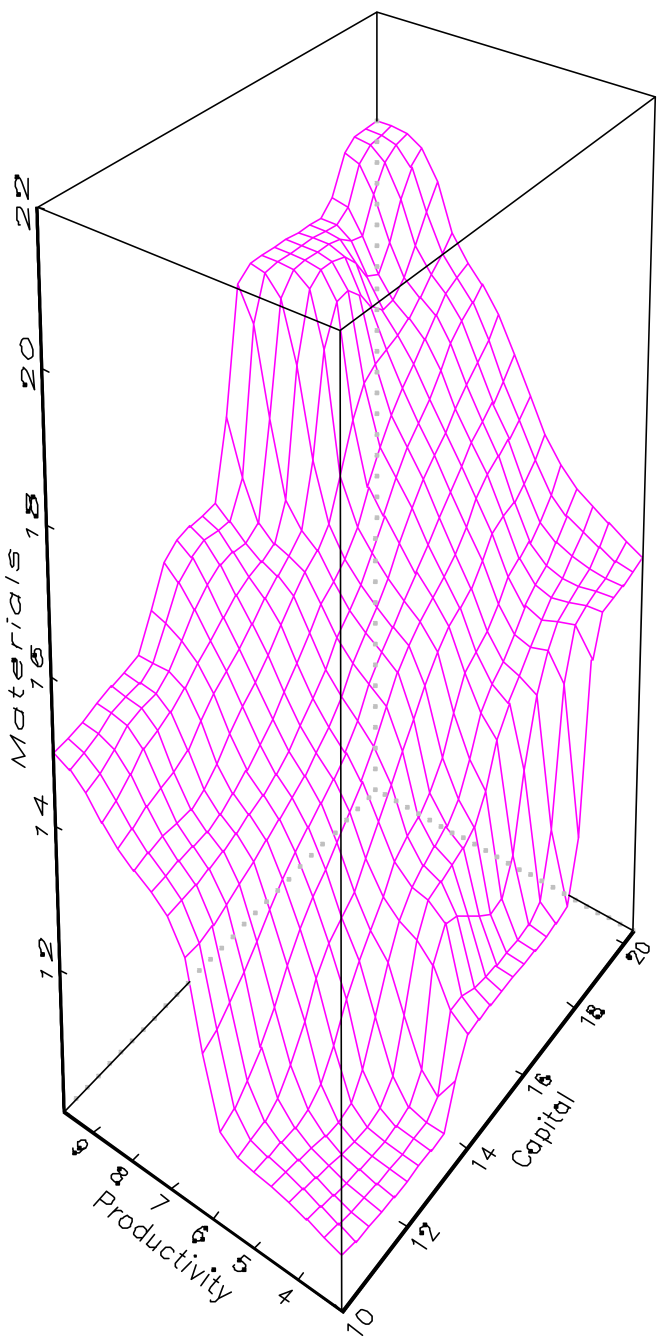

Table 1 **Data Means**

Worker Variables	GMES Sample	GLSS Survey (Employees)
Hourly Wage (Constant 1991 cedis)	167	
Years of Education	10.49	10.68
Years of Experience	18.77	21.93
Age	35.1	38.6
Hours Worked (Weekly)	45.0	44.8
% Apprentices	6.0	1.9
Firm Variables		
Value-Added (millions 1991 cedis)	110	
Capital stock (millions 1991 cedis)	553	
Number of worker-hours (weekly)	3359	
Size of Firm (%):		
1-5 employees	13.1	
6-29 employees	35.0	
30-99 employees	32.9	
100 or more employees	18.9	
Industry (%)		
Wood	28.9	
Metal	23.7	
Textile and Garments	21.8	
Food Processing	25.0	

Table 2

Production Function Estimation - Ghana Manufacturing Sector

N=627 Dependent Variable - log(Value-Added)

	(1) OLS	(2) GMM	(3) IV-GMM
log(employment) (β_H)	0.840 (.058)	0.324 (.089)	0.309 (.087)
Education (β_H*β_S)	0.046 (.024)	0.059 (.027)	0.095 (.054)
Experience (β_H*β_X)	-0.010 (.01)	-0.003 (.012)	-0.013 (.019)
log(Capital)	0.139 (.03)	0.406 (.153)	0.416 (.183)

Note: Standard errors are in parentheses. The standard errors are bootstrapped (b=100). Controls for industry sector, as well as the share of employees who are apprentices were included in all of the estimation procedures. Lagged values of education and experience are used as instruments for education and experience in the column (3) results. Further details in text.

Table 3

Returns to Education in the Ghanaian Manufacturing Sector[1]

Dependent Variable: log(Wage)

				With Selection			
	(1) OLS	(2) Eqn (29)	(3) Eqn (33)	(4) OLS	(5) Eqn (29)	(6) Eqn (33)	(7) Recall Prod. Fn. GMM
Education	0.119	0.109	0.104	0.117	0.110	0.100	β_S=0.183
	(.018)	(.017)	(.024)	(.018)	(.017)	(.021)	(.147)
Experience	0.031	0.024	0.014	0.050	0.038	0.023	β_X=-0.008
	(.008)	(.007)	(.007)	(.011)	(.01)	(.008)	(.045)
Firm Productivity		0.194	0.022		0.164	0.028	
		(.04)	(.047)		(.037)	(.046)	
λ				-1.640	-1.097	-0.581	
				(.496)	(.405)	(.383)	
N	597	597	409	597	597	401	
R^2	0.883	0.896	0.358	0.890	0.899	0.352	

[1] Note: Standard errors are in parentheses. In column (7), these standard errors are bootstrapped (b=100). All regressions include indicator controls for whether an employee is an apprentice. Columns (4), (5), and (6) use the Heckman two-step procedure to correct for selection into the manufacturing workforce using father's occupation and mother's education as identifying variables.

Table 4

Returns to Education in the Ghanaian Manufacturing Sector[1]
Instrumenting for Measurement Error
Dep: log(Wage)

	Instrumenting for Measurement Error						
				With Selection			
	(1)	(2) Eqn (29)	(3) Eqn (33)	(4)	(5) Eqn (29)	(6) Eqn (33)	(7) From Prod. Fn. (IV-GMM)
Education	0.130	0.116	0.104	0.124	0.113	0.108	β_S=0.307
	(.031)	(.03)	(.081)	(.031)	(.03)	(.078)	(0.260)
Experience	0.035	0.027	0.022	0.048	0.038	0.051	β_X=-0.041
	(.014)	(.012)	(.034)	(.02)	(.017)	(.065)	(.114)
Firm Productivity		0.176	0.005		0.162	0.017	
		(.044)	(.062)		(.039)	(.055)	
λ				-1.410	-1.072	-1.376	
				(.853)	(.69)	(2.045)	
	552	552	344	552	552	344	
	0.894	0.905	0.356	0.898	0.907	0.334	

[1] Note: Standard errors are in parentheses. In column (7), these standard errors are bootstrapped (b=100). All regressions include indicator controls for whether an employee is an apprentice. Columns (4), (5), and (6) use the Heckman two-step procedure to correct for selection into the manufacturing workforce using father's occupation and mother's education as identifying variables. Columns (1), (2), (4), and (5) use lagged levels of education and experience as instruments, while columns (3) and (6) use lagged differences as instruments for the differences used in these specifications.

Appendix Table 1 - Estimates of the Returns to Education Using Instrumental Variables

Author	Sample and instruments used		OLS Coefficient	IV Coefficient
1. Angrist and Krueger (1991)	1970 and 1980 Census Data, Men. Instruments are quarter of birth interacted with year of birth. Controls include year of birth and a quadratic in age.	1920-29 cohort in 1970	.0802 (.0004)	.1310 (.0334)
		1930-39 cohort in 1980	.0711 (.0003)	.0760 (.0290)
		1940-49 cohort in 1980	.0573 (.0003)	.0958 (.0223)
2. Harmon and Walker (1995)	British Family Expenditure Survey (1978-86). Instruments are indicators for changes in the minimum school leaving age in 1947 and 1973. Controls include region, year, and a quadratic in age.		.0613 (.001)	.1525 (.015)
3. Kane and Rouse (1993)	National Longitudinal Study of the High School Class of 1972 and the NLSY. Instruments are public state college tuition for and distances to 2- and 4-year colleges. Controls included race, gender, number of years with kids under 6, an interaction between gender and years with kids, city size, region, part-time status, and a dummy for living in California in 1979	Using just public tuition	.080 (.005)	.113 (.035)
		Using tuition and distances to 2- and 4-year colleges	.080 (.005)	.091 (.033)

Appendix Table 2 - Estimates of the Returns to Education Using Family Background Variables as Instruments

Author	Sample and instruments used		OLS		IV
			No Controls	Controls	
1. Ashenfelter and Zimmerman (1997)	NLS Young Men (1966 Cohort) merged with NLS Older Men	Quadratic in age	0.057 (0.009)	0.049 (0.009)	0.109 (0.025)
2. Callan and Harmon (1999)	Survey of Income Distribution, Poverty and Usage of State Services, conducted by the Economic and Social Research Institute (Ireland) 1987. Instruments include mother's and father's education and family's social class. Controls include a quadratic in age or experience and indicators for urban, Dublin, marital status, and an occupation-specific unemployment rate.	Quadratic in age	0.074 (0.005)		0.101 (0.013)
		Quadratic in actual experience	0.082 (0.005)		0.118 (0.013)
3. Uusitalo (1999)	Finnish male constructed dataset linking tax data, census data, and (compulsory) military service records. Instruments include father's education and occupation. Controls include indicators for Helsinki, urban and private sector, ability measures, and a quadratic in potential experience.		0.074 (0.006)		0.129 (0.018)
4. Brunello and Miniaci (1999)	Bank of Italy survey on the income and wealth of Italian households (1993, 1995). Instruments include mother's and father's education and occupation and an indicator for school attendance after a 1969 reform making college more accessible. Controls include region and year of survey and age (linear).		0.048 (0.019)		0.057 (0.022)
5. Levin and Plug (1999)	Brabant Survey (1983) and OSA Panel Survey (1985-1994), Netherlands. Instruments include parental education and occupation. Controls include marital status and quadratics in age or experience and, in the Brabant Survey, IQ.	Brabant Survey - quadratic in potential experience	0.029 (0.004)		0.045 (0.010)
		OSA Survey - quadratic in age	0.036 (0.002)		0.050 (0.006)

Appendix Table 3 - Estimates of the Returns to Education Using Studies of Twins

Author	Sample		Cross-sectional OLS	Differenced[d] LS	IV	IV-CME
1. Ashenfelter and Krueger (1994)	Twinsburg, 1991	Controls for gender, race, and a quadratic in age in X-section OLS	0.084 (0.014)	0.092 (0.024)	0.167 (0.043)	0.129 (0.030)
2. Ashenfelter and Rouse (1998)	Twinsburg, 1991-93	Controls for gender, race, and a quadratic in age in X-section OLS	0.102 (0.010)	0.070 (0.019)	0.088 (0.025)	
3. Rouse (1999)	Twinsburg 1991-93, 95	Controls for gender, race, and a quadratic in age in X-section OLS	0.105 (0.008)	0.075 (0.017)	0.095 (0.027)	0.119[a] (0.021)
4. Behrman, Rosenzweig, and Taubman (1994)	Minnesota 1993		0.071 (0.002)	0.035 (0.005)		
5. Behrman and Rosenzweig (1999)	Minnesota 1993	Controls for actual work experience, gender, and tenure in all specifications	0.118[b] (0.005)		0.104 (0.017)	0.104[c] (0.017)
6. Miller, Mulvey and Martin (1995)	Australia 1980-82,88-89	Controls for marital status, gender, and age	0.064 (0.002)		0.025 (0.005)	0.048 (0.010)

[a] This specification includes controls for union status, marital status, and gender, while the others do not.
[b] They do not have a simple OLS specification, and so the GLS specification is reported here.
[c] No, this is not a typo.
[d] The first column represents the regular between-twin, or twin fixed effect regression. The second regression use's the co-twin's report of the individual's schooling as an instrument. The third column allows for correlated measurement error in each twin's schooling reports.