Learning the Master's Trade: Apprenticeship and Human Capital in Ghana

Garth Frazer

Centre for Industrial Relations, University of Toronto, 121 St. George Street, Toronto, ON, Canada M5S 2E8
School of Management, University of Toronto, 105 St. George Street, Toronto, ON, Canada M5S 3E6

Abstract

This paper explores the institution of apprenticeship in Ghana. A model is presented where apprenticeship training is idiosyncratic, increasing an individual's productivity in the current firm, but not in any other firm. Still, individuals are willing to fund apprenticeships as they can reap the returns to the specific training of apprenticeship if they manage to acquire the capital required to start their own firms, and replicate the technology and business practice of the apprenticeship firm. Predictions of the model for the productivity and remuneration of different workers are developed and tested using both a linked employer-employee survey of manufacturing firms, and a national household survey.

JEL Codes: O12, J23, J24

keywords: apprenticeship, Africa, human capital

§ This work has benefited from the very useful comments and suggestions of two anonymous referees and the Editor, as well as Christopher Udry, David McKenzie, Patrick Bayer, and seminar participants at the University of British Columbia, Northwestern University, Yale University, Stanford University, and Carleton University. Financial support from the Connaught Fund is gratefully acknowledged. Permission from the Ghana Statistical Service to use the Ghanaian Living Standard Survey data is most certainly appreciated. The author owes a particular debt to Francis Teal for the opportunity to participate in the Ghanaian Manufacturing Enterprise Survey, and for the use of its data.

† Tel: +1 416 978 5692, Fax: +1 416 978 5433.
E-mail address: gfrazer@chass.utoronto.ca
1 Introduction

The formulation of an appropriate education and training policy for the manufacturing sector in Africa should begin with an understanding of the training currently in place. Given the importance of apprenticeships in Africa,\(^1\) where apprentices learn one of the trades used in the manufacturing sector, any analysis should begin by examining this institution. Nevertheless, very few articles exist which use economic analysis to explore apprenticeships in Africa (Velenchik, 1995). This paper is a step towards addressing this dearth, and contends that the apprenticeship institution is best understood in the larger context of the specificity of firm training.

Apprenticeships are periods of roughly three years in length during which an apprentice learns a trade, such as metal-working or carpentry from a master of that trade. At the end of the apprenticeship, the apprentice may end up being hired by the firm where the apprenticeship occurred, begin working at another firm, or the apprentice may start a new firm and become self-employed. Apprenticeships occur most often, but not exclusively, in smaller firms, and the master is often the owner of the firm.

Becker (1964) showed that in the human capital model with perfect labor markets, workers (and not firms) always pay for general training, which increases an individual’s productivity with all employers, as the value of their outside wage offer increases with general training. However, firms cannot credibly commit, without some formal contracting or other mechanism, to compensate workers for firm-specific training once it has occurred, as it does not increase their outside wage offers. This dichotomy has provided a framework

\(^1\)Apprenticeships are particularly important in West Africa, but are significant institutions throughout Africa. Their significance is documented in Boehm (1997) and Bas (1989).
for analyzing the nature of training in a number of studies. Acemoglu and Pischke (1998) treat German apprenticeships as general training, and present a model that explains why firms are willing to invest in this training. Acemoglu (1997) and Acemoglu and Pischke (1999) provide distinct models which explain why firms might invest in general training.

Other papers have examined worker investment in firm-specific training. Prendergast (1993) describes the potential for promotion rules to induce worker investments in firm-specific human capital, with the firm’s ability to commit to future wage increases through long-term contracts playing a significant role in the model’s mechanism. Scoones and Bernhardt (1998) remove the need for long-term contracts to achieve worker investment in firm-specific human capital by introducing asymmetric information. In their model, workers invest in firm-specific human capital, in order to be promoted by the firm, as promotion reveals their higher ability to other firms. In a similar model, Scoones (2000) finds that workers invest in firm-specific capital because efficient turnover transforms former employers into outside options. Other papers (e.g. Felli and Harris, 1994) examine firm-specific human capital as something exogenous to worker decisions, but we are interested in examining investment in human capital that is discretionary from the worker perspective.

At first glance, apprenticeships might appear to be general training, that is applicable at least in all of the firms within an industry, if not more generally. As noted, this is how they have been understood in Germany. However, the nature of the training acquired in Ghana is far more specific—specific to the firm providing the training, reflecting the firm’s technology and business practice. In this context, the model of apprenticeship to be developed shares the spirit of the model of Jovanovic and Nyarko (1995). In their
overlapping generations model, each old agent understands an idiosyncratic (which is unique to the firm owned by the old agent) technology which is passed on to a young agent who is his apprentice. This interpretation of apprenticeships as training unique to the firm is consistent with the apprenticeships under examination in this paper.\(^2\) Apprentices are taught by a master how to work the master’s craft, but the way in which that craft is carried out varies highly from firm to firm. In the process of this research investigation, I visited and interviewed dozens of the manufacturing firms in the dataset. A number of the firms interviewed were involved in the manufacture of a single item. When apprentices apprentice at these firms, they may only learn how to manufacture one, or a small number of items. While apprentices at other firms did learn how to manufacture a wider variety of items, product homogeneity, using unique technology, is quite common among the practice of firms with apprentices. The training given by the master, both in terms of technology and business practice, is that of the master, and typically idiosyncratic.\(^3\)

Therefore, the apprenticeship training cannot be applied to work in other firms, which have their own technology and business practice. However, once the apprenticeship is complete, the former apprentice can use this knowledge to start his own firm and pursue self-employment, replicating the apprenticeship firm. In fact, apprentices appear highly motivated to pursue self-employment. Of the Ghanaian manufacturing workers, who had completed apprenticeships, 77% stated that they would prefer to be self-employed rather

---

\(^2\)The apprenticeships under examination in this paper include all of the non-tailoring apprenticeships in Ghana, in particular those within the Ghanaian manufacturing sector. Tailoring apprenticeships are to be considered in a separate paper by the same author in order to account for their idiosyncratic nature.

\(^3\)The idiosyncratic nature of the technologies used by apprenticeship in Ghana relates to a larger literature on the technological capabilities within African countries. Lall and Pietrobelli (2002) explore the roots of the limited technological capabilities within Africa in much greater depth than is possible here. Lall et. al. (1994) focus on technological capabilities in manufacturing (and the limitations thereof) in the particular context of Ghana.
than working in their current job. Naturally, however, apprenticeship knowledge is not enough—capital is required to start a firm. Those former apprentices who get access to capital can start a new firm, while others cannot. The potential to reap the returns of the apprenticeship, this specific training, by creating a new firm and replicating the technology and business practice of the apprenticeship firm motivates individuals to acquire the human capital of apprenticeship.

This apprenticeship human capital is not transferable to any other firm outside the current firm, but it can be used when the apprentice replicates the technology and business practice of the current firm in self-employment. Therefore, while the apprenticeship certainly does not provide general human capital (which is useful in all firms), the human capital is also not ‘firm-specific’ in the pure definition of the term. The human capital can be used outside of the current firm, but only when the apprentice replicates the apprenticeship firm. Perhaps a term such as ‘technology-specific’ might better describe the nature of the apprenticeship human capital. However, this technology should be understood as unique to the firm providing the apprenticeship training. The exact labelling of the human capital is not as important as understanding its nature. Therefore, it shall be called simply ‘specific human capital’ in this paper. Workers are willing to invest in this type of specific human capital, which is tightly tied to the firm providing apprenticeship training, not because of asymmetric information, or any other labor market imperfection, or the firm’s ability to commit to future wage increases, but rather through a worker’s potential ability to reap the benefits of the apprenticeship through self-employment.

Of course, once the apprenticeship is finished, the apprentice can also continue to
work in the master’s firm. While the apprenticeship knowledge is certainly applicable in this context, the master, and not the apprentice, reaps the returns to this knowledge. The reason for this is the fact that while the former apprentice is more productive in the master’s firm, he is not more productive in other firms, and so his outside wage option is low. The master, therefore, only needs to pay the former apprentice marginally above his outside wage option in order to retain him in the firm. Only in self-employment will the former apprentice receive his full marginal product (including the apprenticeship-enhanced productivity) as a wage.

Fortunately, this model, which will be formalized in Section 3, creates predictions which can be tested using the data. Although the model explains worker investment in specific training, the mechanics of the model resemble Acemoglu and Pischke (1998). The predictions of the model include the following. Former apprentices who apprenticed in the current firm should be more productive than former apprentices from other firms. However, they do not need to be paid more than former apprentices from other firms. For this reason, former apprentices will seek self-employment, and the returns to apprenticeship should be found in self-employment. This model shares similarities and differences with Becker (1964)’s original work in this area. In Becker’s model, workers and firms share the cost of firm-specific training, although the relative proportions contributed may vary. Here, we will see that either workers or firms or both may cover the cost of the training, depending on the market for apprenticeship. In Becker’s model, after the training, firms pay these specifically-trained workers less than their marginal product. This fact reflects the firm’s ownership of the worker’s specific capital, and allows the firm to recoup its portion of the training cost. Here, also, specifically-trained workers receive less than their
marginal products in the training firms as a wage. As in Becker’s model, workers with specific training in a different firm will receive their marginal product in the current firm, but this amount is less than what their marginal product would have been in their training firm. While Becker (1964)’s discussion of self-employment is implicit rather than explicit, an application of the original Becker framework to include self-employment (without credit constraints, as in the original Becker model) would also see returns to specific training in self-employment, as the worker becomes owner of his or her specific capital. In fact, one purpose of this paper is to explicitly and formally make this extrapolation—to examine the case of specific capital and self-employment, in the context of apprenticeship in Ghana.

An important characteristic of the model is the fact that firms are not allowed to make time-inconsistent wage offers. That is, in any given period, it will only prove optimal for firms to pay workers their outside option. As will be outlined in the paper, this reflects the difficulties in contract enforcement in Africa, through firm reputation or other means. For comparison purposes, the model for the case of completely enforceable long-term wage contracts will also be provided. Firms make less profit in the absence of full contract enforcement, and a measure of this loss is calculated in the context of this model. Therefore, this paper also provides a specific example of the costs that African firms incur as a result of the difficulties of contract enforcement there.

This paper will both formalize the model described above, and test it in the context of Ghana, and Ghanaian manufacturing firms. Fortunately, the predictions of the model can be tested both using a survey of manufacturing firms and their employees, as well as using a national household survey for Ghana. In particular, the linked employer-employee dataset enables us to separate a worker’s compensation from his productivity, and therefore
measure separately the impact of apprenticeship on worker compensation and productivity. The two datasets used are described in Section 4, with the results and concluding thoughts in Section 5.

2 Background

Apprenticeships are an important form of training for the manufacturing sector in Africa. While apprenticeships are prevalent throughout Africa (Birks, Fluitman, Ouding and Sinclair, 1994), the apprenticeship system is most widespread in Western Africa, including Ghana (Boehm, 1997, p. 320). Callaway (1964) is one of the few studies of apprenticeship to place it in a historical context, describing apprenticeship in Nigeria, where he states, "this vast apprenticeship training system began as part of a wider education process in which the indigenous societies of Nigeria passed on their cultural heritage from one generation to the next." (p. 63) Still, this transmission has been modified from its historic origins, where the apprenticeships largely occurred within families to the current version of apprenticeship with a formalized contract, where apprenticeships only occasionally occur within families. The lack of historical context to the few existing studies of apprenticeship make it difficult to explain the stronger prevalence of the institution in West Africa than in other parts of Africa, although part of the answer surely lies in the cultural histories of each region. A further answer is alluded to by King (1977) who studies artisanal trade in Kenya, and the training for these trades, including apprenticeship. According to King,

[4]Boehm, citing a number of other, primarily unpublished studies, reports rates of apprenticeship among small and medium enterprise entrepreneurs at 84 percent in Ghana and Francophone Africa, 76 percent or 89 percent in Nigeria, 58 percent in Zambia, and 21 percent in Zimbabwe. Since the original studies have restricted access, it is difficult to discern the comparability of these statistics across countries. Still, the numbers suggest that even in Zimbabwe, which likely has one of the lowest rates of apprenticeship, it remains a significant method of training.
"the major irruption of specialised craft communities to Kenya was associated with the waves of Indian immigration," who brought skills in "building tin, wood, steel, car repair, and many other skills," and "there is little doubt that they did restrict their craft expertise as far as possible to their own community." (p. 52) However, later, in the 1950s and 1960s, it was not merely the Indians themselves that were responsible for restricting the spread of skilled trades through apprenticeship, but also the colonial government. The colonial government used labour inspectors to impose fines on firms that had used fee-paying apprentices, as these apprentices were clearly classified by these inspectors as illegally cheap labour, rather than students acquiring a skill (King, p. 53). It is also difficult to disentangle the potential other motives of the colonial government, who may have wanted to avoid replicating the British case of apprenticeship were apprenticeship premiums discriminated against the lowest income levels (King, 1977, p. 53, 61). Still, according to King, the efforts of the colonial government restricted apprenticeship in Kenya, as opposed to West Africa: "The very hostility of the colonial government, however, meant that the system [of apprenticeship] has not really come into the open as in West Africa, even though it expanded very rapidly in the post-colonial period." Of course, King is describing the Kenyan situation, and it is not clear the degree to which similar stories could be told for Southern or Central African countries.

While few comparative cross-country surveys exist to compare the prevalence of apprenticeship across African countries, the Ghanaian Manufacturing Enterprise Survey (GMES) data used in this paper was one of nine surveys conducted in African countries in the 1990s as part of the World Bank’s Regional Program on Enterprise Development (RPED) surveys. The results of these surveys for seven countries are summarized in
Mazumdar and Mazaheri (2003). The percentage of manufacturing sector entrepreneurs with apprenticeship training in Cameroon, Côte d’Ivoire, Ghana, Kenya, Tanzania, Zambia, and Zimbabwe, respectively are 18.4, 59.1, 55.6, 21.9, 31.3, 13.9, and 9.4 percent. While this obviously does not represent the proportion of the population who are currently in, or have completed, apprenticeships, with the exception of Cameroon, it is consistent with the anecdotal reports of the prevalence of apprenticeship in West Africa.

While apprenticeships have remained important throughout Africa, there has also been a rise in the number of technical and vocational schools, as well as polytechnics, particularly since independence in African countries, including in Ghana. These schools were typically developed as part of general educational expansion in these countries after independence, and were intended to provide skills particularly for the formal industrial sector. While the primary means of training for the informal sector has been through apprenticeship, technical and vocational school graduates can also be found in the informal sector. The nature and value of all forms of training is naturally of interest given Africa as a region tends have a relative shortage of skilled labour. While much work has been done evaluating the overall returns to education (sometimes separately for primary and secondary) within African countries (Appleton, et. al., 1996; Bennell, 1996; Boissiere, et. al., 1985; Glewwe, 1999; Glick and Sahn, 1997; Hazlewood et. al., 1989; Jones, 2001; Knight and Sabot, 1990, Mwabu and Schultz, 1996; Nielsen, and Westergård-Nielsen, 2001; Siphambe, 2000; Vijverberg, 1993), little work has been done specifically on technical training or on apprenticeship (Bas, 1989; Boehm, 1997; Velenchik, 1995). That which has been done has been largely descriptive. When formal sector technical and vocational education is evaluated, it is typically criticized for being too theoretical, and therefore not
sufficiently applied to the problems of the workplace, or responsive to changes in the labor market. On the other hand, informal methods of training such as apprenticeship are typically seen as lacking theoretical foundation, and as insufficient for the modern technical demands of the formal sector. An advantage of apprenticeship from the government’s perspective is that the government does not fund this training—the cost of training is borne by the apprentice and/or firm. There is also a widespread belief that apprenticeship provides an option for youth who would otherwise be unemployed. In addition to both the formal sector training and the institution of apprenticeship, other manufacturing sector training is carried out within firms in the form of short-term and long-term on-the-job training. Unfortunately, while a variety of positions are held by various policymakers on the usefulness of these various types of formal and informal training, there has been little economic research in this area. This paper, therefore, seeks to take a single step in advancing this research, and also seeks to encourage further research in this area.

Regarding apprenticeship, Velenchik (1996) focuses on the importance of the apprenticeship institution as a source of finance for some firms. She finds that firms that charge apprenticeship fees are more likely to use other forms of informal finance, and interprets this as evidence that these firms are likely constrained in their ability to access finance, and therefore use the apprenticeship fees as a source of funds. While she does not establish a full, formal model, she discusses the implications of the specificity of the firm training for the size of the fees. She assumes that firms will tend to pay more for firm-specific apprenticeship training, and workers pay more for general apprenticeship training, and that the cost of the training can be borne by fees at the beginning, during, or end of the apprenticeship period. She describes the cases where more of the apprenticeship
fees are paid earlier as apprentice-financing, and the cases where more of the apprenticeship fees are paid later as employer-financing. Unfortunately, obtaining an exact measure of the apprenticeship fee is extremely difficult, given that it represents a net transfer, where part of the transfer from the apprentice to the firm owner comes in the form of contribution to firm product, and arguably even to the general welfare of the firm owner. As Berry (1985) describes in the context of motor mechanic apprenticeship in Nigeria, "Apprentices are bound labor for the term of their training. They may be asked to run errands and perform household tasks...; they are at their masters’ beck and call at all times" (p. 142) Moreover, as I discovered in interviewing these apprentices, the apprentices frequently do not know the terms of their contract, a contract that has been worked out between the firm owner and the apprentices’ parents or senior relatives. As Berry (1985) notes, "my informants often knew little about the terms of their own apprenticeships, which had invariably been arranged between the master and the apprentice’s senior relatives." (p. 142) Oyeneye (1981) states that this is not surprising. "Only in very rare cases would a master accept a prospective candidate on his own standing without the customary introduction by an elderly or responsible person." As Oyeneye continues, "the preponderance of parental decisions in the sample is an indication of the economic aspirations which parents have for their children. It also reflects the authority structure within families." (p. 19)

---

5Part of the reason that the contract is with the apprentice’s parents or other senior relatives, as opposed to the apprentice himself, is to increase the degree of commitment to the apprenticeship contract. If the apprentice should abscond from the apprenticeship, that apprentice would damage not only his reputation, but that of his entire family, and the family has a variety of means of punishing that type of behaviour not available to the apprentice’s master.

6The considerable degree of involvement and authority of parents and the extended family in the decisions regarding the education and training of African children is a general characteristic of African culture. That being said, the graduation ceremony, at the end of the apprenticeship, reflects the apprentices’ freedom not only from the terms of apprenticeship, but also a greater degree of freedom from their parents’ authority (although this is never complete freedom). Former apprentices who either obtain work or start their own business have the same access to the earnings from their labour as any other workers in Ghanaian society.
A principal constraint for former apprentices is obtaining the finance in order to start up their own business. In fact, access to credit typically remains a binding constraint even among those small businesses that manage to acquire start-up capital (Levy(1993), Steel and Webster(1991), Aryeetey et. al. (1994)). In a summary of these studies, Nissanke and Aryeetey (1998) find that "there is evidence that 'lack of access to finance represents the binding constraint on expansion for small- and medium-scale enterprises’ (SMEs)." (p. 33) In a dynamic, structural model of manufacturing firms in Ghana, Schündeln (2004) also finds evidence of financing constraints. It is worth examining the source of funds for those small firms that do obtain start-up capital. In a survey of small and medium-sized enterprises in Ghana, Aryeetey et. al. (1994) find that the most important source of start-up capital is owners’ savings, the main source in 67 percent of firms, followed by assistance from relatives and friends. Only a small fraction of these firms could gain access to bank loans. Moneylenders are also present in Ghana, but these are not a significant source of start-up capital.

In fact, the widespread perception that potential small-scale entrepreneurs lack access to sufficient capital in order to take advantage of their profitable opportunities is part of what has led to the recent massive expansion of microfinance programs in a number of developing countries (Morduch, 1999). While there has been considerable debate about whether microfinance programs are the most cost-effective way of reducing poverty (Pitt and Khandker, 1998; Morduch, 1999), these programs have been found to have a significant effect on the rate of, and profits from, self-employment activities, particularly in non-farm self-employment (McKernan, 2002; Khandker et. al., 1998). These results are consistent with the idea that the microfinance programs are doing what the organizers intend: help
small-scale entrepreneurs to overcome credit constraints that limit the creation and
expansion of small-scale enterprises.

3 The Model

The following model is very simple, but is designed to focus on the institution of interest,
namely apprenticeship. In attempting to reflect reality, the returns to apprenticeship in
this model are in self-employment, and access to capital is what determines the probability
of obtaining self-employment following an apprenticeship. Therefore, in the simplest
version of the model, workers are differentiated only along this characteristic, the
probability of obtaining capital or credit.

3.1 The Basic Model

The world lasts for two periods. There are a large number of workers. Firms can hire
workers at the beginning of either period. In addition, workers can apply for an
apprenticeship at a firm at the beginning of the first period. The apprenticeship increases
the productivity of workers within that firm, but not within other firms, unless the worker
starts a new firm using the knowledge gained in the apprenticeship. Workers and firms are
risk neutral.\footnote{Assuming that workers are risk neutral focuses on the differences in mean earnings, as opposed to the variance of these earnings. In some contexts, such a focus has been criticized because self-employment earnings have been found to have a considerably higher variance in earnings than employee earnings. However, in our dataset (see Table 1), we see that for the labor force as a whole, the standard deviation of earnings is 14% of the mean for both self-employed and employees. For the manufacturing sector, the standard deviation of earnings is 21% for employees and 20% for the self-employed, again essentially the same for both categories. Therefore, in our context, the differences in mean levels of earnings appears to be more important than the variation in the earnings.} There is no discounting between the two periods. All production occurs in
the second period, in firms with a technology which is linear in the number of workers that
they hire. Workers are differentiated in terms of their access to credit. Specifically, each worker will have a probability, $\rho$, known to them and to everyone else, of receiving credit in the second period. For simplicity, $\rho \sim U(0, 1)$. Workers with credit will have the option to start their own firms in the second period.\footnote{Note that the probability will depend on both the likelihood of receiving credit, as well as the what the actual cost of entry into self-employment is. These are jointly summarized by $\rho$.} Workers working at a firm are of two types. Workers that have apprenticed within the firm can produce $m$ products per period within that firm. Workers that have not apprenticed within that firm can produce $n$ products per period, with $m > n$. In order to focus on the specific nature of the apprenticeship rather than asymmetric information, all workers and all firms have perfect information about the productivity of all other workers (as well as perfect information about their credit opportunities). Given the very simple production patterns, knowing a worker’s productive capacity amounts to knowing whether and where a worker apprenticed.

In this economy, there are also reasons for quits. Workers will not get along with their employer, or will need to move for a variety of other reasons. In order to capture this reality, with probability $\lambda$, each worker will receive a disutility shock, $\tau$. For simplicity, we will take $\tau > m - n$. The worker can avoid this disutility by taking a different job. This, therefore, is equivalent to assuming that workers are separated from their employer for some exogenous reason, with probability $\lambda$. The relative importance of this assumption will be captured in the size of $\lambda$.

The exact sequence of events in the economy is as follows:

1. a) Firms set an apprentice fee, $c$, which they charge to all apprentices at the firm.\footnote{Naturally, the apprenticeship is a series of pecuniary and non-pecuniary transfers back and forth between} No restrictions are placed, a priori, on the sign of $c$, so that the firm can actually
pay the apprentices to apprentice at the firm, if this is optimal \((c < 0)\). With this definition, it is always optimal for firms to accept apprentices, as the level of \(c\) can be adjusted to ensure that this holds.

1.b) Workers decide whether or not to apply for an apprenticeship at a firm. The firm accepts the apprentice, given the apprentice fee (posted in 1a), \(c\). Apprentices then complete their apprenticeship. The cost to the firm of training the apprentices (in terms of materials, time of the master, and other costs) is \(t\).

2.a) At the beginning of period 2, firms set a wage for those who have apprenticed in this firm, as well as a wage for workers from outside the firm.

2.b) Next, workers learn whether or not they have received the capital to start their own firm. They then decide whether to become self-employed, or to remain as a worker.

2.c) Next, workers receive their disutility shock, and then decide whether or not the apprentice and the firm owner. Apprentices pay fees at the beginning and end of the apprenticeship, but may also receive a small allowance over the course of the apprenticeship. Apprentices also contribute to firm product. The apprenticeship fee, \(c\), should therefore be thought of as the net value of the transfer (pecuniary and non-pecuniary) from the apprentice to the firm owner. The firm owner has control over the size of \(c\). Even if the firm owner does not control each of the components of \(c\), the firm owner is aware of the value of the components, and does control the pecuniary transfers, and therefore controls the overall size of \(c\).

\(^{10}\)Note that the apprenticeship fee, \(c\), is the same for all apprentices at a given firm. This fact is reflected in the data, where 77 and 72 percent of the variation in apprenticeship fees (initial and final, respectively) can be explained by a regression on firm dummies, and the year that the apprenticeship started (details available from the author upon request). This probably reflects the worker morale difficulties with charging differential fees. It is also consistent with the fact that differential school fees (for different students in the same school) are rare in developing country settings.

Still, consider the case where firms are allowed to charge differential apprenticeship fees, conditional on a person’s probability of receiving credit, \(c(\rho)\). In this model, the details of which were in an earlier version of this paper, everyone apprentices, as the firm charges each individual the fee for which he is just indifferent to apprenticeship. Nevertheless, the major predictions of the model, to be outlined later, follow through in this case.

\(^{11}\)This cost reflects the costs to the firm of performing the training of apprenticeship. The apprentice does not benefit from these costs (other than through the increased human capital resulting from the apprenticeship). Therefore, while the apprentice fee, \(c\), captures the net transfer from the apprentice to the firm, and therefore is a benefit to the firm and a cost to the apprentice, the training cost \(t\) is simply a cost to the firm, and is only captured by the apprentice through the value of the apprenticeship.
they will work at this firm, or quit.

2.d) Outside firms make wage offers to the workers in the secondhand market. These workers will include those who did not engage in an apprenticeship, as well as those workers who did an apprenticeship, and then left because of their disutility shock. Whether the outside firm can distinguish between these two types is irrelevant, because they are equally productive when employed outside of the apprenticeship firm.

Finally, we impose that in period 2, there is free entry of firms, so that no firms earn a positive profit in period 2. We also consider both the cases of free entry in period 1, and no free entry in period 1. The case of no free entry in period 1 describes the case where some firms have some ability to train apprentices, and are able to extract a rent from that, while other firms cannot. Whether or not there is free entry in period 1 will affect the apprentice fee charged to apprentices, and therefore firm profits (clearly), but will not affect the other predictions of the model in the simplest case.

3.2 Equilibrium

In order to demonstrate that the model does not rely on any information asymmetries, we have assumed perfect information. As a result, the equilibrium concept which we will use is a subgame perfect equilibrium, although we find that the equilibrium obtained is also trembling hand perfect.\textsuperscript{12} Therefore, the equilibrium is a set of numbers and functions.

\textsuperscript{12}This equilibrium concept is most useful in this particular extensive form game with perfect information. The game is indeed an extensive form game, as the firm does first set the cost of apprenticeship, followed by the potential apprentice choosing whether or not to apprentice. The wage for the apprentice is then chosen after the apprenticeship is completed. At each of these important stages, the decision of agents at all previous steps is entirely known. This makes the extensive form game with perfect information a more realistic description of the world than in some other game-theoretic situations.

Given that every finite extensive form game has a trembling-hand perfect (in the agent normal form) Nash equilibrium, so does the equilibrium specified here. Moreover, given that this is an extensive form game with perfect and complete information, backwards induction will provide the unique subgame perfect
(A(\rho), c, W_F, W_O, W_n, q(\tau), S, w_O, w_n), such that:\textsuperscript{13}

1) All workers make the optimal decision on apprenticeship, $A(\rho)$, given the cost to the worker of the apprenticeship, $c$. Workers also make the optimal decision on whether to quit, $q(\tau)$, given the realization of their shock, $\tau$.

2) All firms with apprentices choose the optimal (net) apprenticeship fee, $c$. These firms also choose the optimal wage, $W_F$, $W_O$, and $W_n$ for workers that have apprenticed within the firm, apprenticed outside the firm, or not apprenticed, respectively.

3) Workers that have received the capital to start a new firm, decide whether or not to become self-employed, $S$.

4) Outside firms (firms entering in the second period) offer wages for those workers that have had ($w_O$), and have not had ($w_n$) apprenticeships somewhere else.

The next step is to characterize the equilibrium. First, a worker’s decision on whether to apprentice will depend on $\rho$, his probability of receiving capital to start a firm in the second period. In particular, a worker will apprentice ($A(\rho) = 1$) if the expected net earnings through apprenticeship exceed the expected earnings from non-apprenticeship:

$$-c + \rho m + (1 - \rho)(\lambda q(\tau)W_O + (1 - q(\tau))W_F) + (1 - \lambda)W_F > \rho n + (1 - \rho)W_n$$ \textsuperscript{(1)}

The expected net earnings of apprenticeship includes paying the apprentice fee, $c$. With probability $\rho$, the apprentice will receive the capital for self-employment, and therefore earn $m$.\textsuperscript{14} Otherwise, the apprentice will receive the disutility shock with

\textsuperscript{13}Note that these following points are not ordered temporally.

\textsuperscript{14}For clarity, in the above expression, the worker that receives capital decides to enter self-employment
probability \( \lambda \), and then choose to separate with probability \( q(\tau) \). On the other hand, if the worker does not apprentice, with probability \( \rho \) he will receive the capital to become self-employed, and earn his productivity \( n \). Otherwise, he will work and receive wage \( W_n \).

In this base model, \( q(\tau) = 1 \), since the size of the shock is larger than the difference \( m - n \). Therefore, this expression simplifies immediately to:

\[
-c + \rho m + (1 - \rho)(\lambda W_O + (1 - \lambda)W_F) > \rho n + (1 - \rho)W_n
\]  

(2)

If the worker does not apprentice, then he will receive the outside offer of a wage for non-apprentices, \( W_n \). Note that the labels \( W_n \) and \( W_O \) have been used in the above inequalities, instead of \( w_n \) and \( w_O \). In equilibrium, the wage offered to both non-apprentices, and those who have apprenticed in a different firm will need to be the same for both inside firms, and outside firms who enter in the second period (\( W_n = w_n \), \( W_O = w_O \)). Otherwise, of course, workers would choose the firms offering the higher wages. Expression (2) leads to a critical value of \( \rho \), say \( \hat{\rho} \), with a worker choosing to apprentice if \( \rho > \hat{\rho} \).

In this case, firms that have apprentices maximize expected profits in the following way:

\[
\max_{c,W_F,W_O,w_n} \Pi = \int_{\hat{\rho}(c,W_F)}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - W_F))dF(\rho)
\]

\[
+ \lambda \int_{\hat{\rho}(c,W_F)}^{1} (1 - \rho)(n - W_O)dF(\rho) + \int_{0}^{\hat{\rho}(c,W_F)} (1 - \rho)(n - W_n)dF(\rho)
\]  

(3)

Therefore, the firm’s profit, if any, arises from the difference between the worker’s

with probability 1, and therefore receives return \( m \). Technically, the \( \rho m \) in the above expression should be replaced by \( \rho(1m + 0(\lambda W_O + (1 - \lambda)W_F)) \), with 1 (and 0) representing the probabilities that the self-employment wage \( m \) is higher (lower) than the alternative wage. The shorter expression above is naturally clearer.
marginal product and his wage, with the expected profits calculated as the expected surplus for each type of worker. The first integral represents the firm’s profits from those who have apprenticed within the firm, and therefore includes those who have a value of $\rho$ which is greater than the critical value, $\tilde{\rho}$. The second integral includes the profit from all of those workers that apprenticed at another firm, did not get capital, but did get a disutility shock, and quit their original firm.\(^{15}\) The third integral includes the profit from all of those workers that did not apprentice, and did not receive the capital to become self-employed.

Consider the first integral. The first term reflects the fact that the firm receives the apprentice fee ($c$) and has to pay the cost of training ($t$) for all apprentices at the firm, regardless of whether or not they stay in the firm. The second part of the integrand reflects the fact that the firm only reaps a surplus on those workers who do not become self-employed ($1 - \rho$), and do not quit the firm ($1 - \lambda$).\(^{16}\) To understand the impact of the final two terms, we need to proceed to solve the model using backward induction.

Because of free entry in the second period, the workers who have not apprenticed, and those who have left their apprenticeship firm will be offered their marginal products as wages by outside firms ($w^*_b = n, w^*_n = n$).\(^{17}\) Note that while firms can identify whether or not workers have apprenticed, this is not relevant for their marginal product outside the

\(^{15}\)Note that under this formulation, workers that quit their apprenticeship firm will show up in the profit functions of two firms. First, they will show up in the profit function of the apprenticing firm as they paid their apprenticeship fee to that firm. However, as they are working and producing in the new firm, they will also show up in that firm’s profit function.

\(^{16}\)Note that this implicitly assumes, as suggested earlier, that those who receive capital do choose to become self-employed. This would be equivalent to assuming that one’s return in self-employment is marginally larger than working in the firm (say $m + \varepsilon$), so that it is impossible for the firm to compensate a worker at the self-employment rate. In the model, this is treated as breaking the tie on the side of self-employment, for simplicity. This is fully consistent with the stated preference for workers to be self-employed, as mentioned earlier.

\(^{17}\)Given the nature of the technology, the firm is indifferent between hiring and not hiring these non-apprenticed or other-apprenticed workers. Whether the non-apprenticed workers are hired by the firm, as the profit function suggests, or by the outside firms is not of consequence for the main objects of investigation.
firm, as the apprenticeship builds specific ability. Therefore, in equilibrium, given the wages of the outside firms, the inside firms will also pay these workers their marginal products, and no more, so that \( W_O^* = n \), and \( W_n^* = n \). Therefore, by backward induction, the profit function for the firm is really just:

\[
Max \Pi = \int_{\hat{\rho}}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - W_F))dF(\rho) \tag{4}
\]

Now, at this stage (after the workers have completed their apprenticeships), the firm takes the apprenticeship decisions of the workers as given \((\hat{\rho})\), as well as the apprenticeship fees that it set at the outset. Therefore, the firm is really in this period just choosing the optimal wage for those who have apprenticed within the firm:

\[
Max \Pi = \int_{\hat{\rho}}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - W_F))dF(\rho) \tag{5}
\]

subject to the constraint that \( W_F \geq n \). Writing the firm’s problem in this way separates the two parts of the firm’s problem to reflect the two different points in time that the firm takes these decisions, on the apprenticeship fee, and the wage of former apprentices. Separating the problem in this way basically amounts to assuming that the firm does not pay the wage for the workers in advance, and also cannot commit to the level of such a wage before the apprenticeship. In general, the moves by the agents in this economy are sequential, and in order for the equilibrium to be subgame perfect, the actions (and beliefs) of all agents are consistent with optimal play on every possible continuation of the game. Firms cannot pre-commit to wage offers that are time-inconsistent (this assumption will be relaxed in Appendix C). Continuing the backward induction argument, the firm solves the problem of (5) when choosing the wage for former apprentices, subject to the condition that the apprentice receives at least \( W_O = n \), which is what the apprentice
would receive if he left the firm. Therefore, in fact, we see that the profit-maximizing wage
for the firm is to offer $W_F = n$, which is just sufficient to maintain the worker's utility
equivalent to that which he would get outside of the firm.\textsuperscript{18}

Again, working through backward induction, the worker then solves his problem\textsuperscript{19},
in equation (2), which can now be simplified to:

$$-c + \rho m + (1 - \rho)(\lambda n + (1 - \lambda)n) > \rho n + (1 - \rho)n$$

or simply:

$$\rho > \frac{c}{m - n}$$

Therefore, $\hat{\rho} = \frac{c}{m - n}$. That is, when $\rho > \frac{c}{m - n}$, (or when the expected benefit from
self-employment, $\rho(m - n)$, exceeds the cost of apprenticeship, $c$), the worker’s probability
of obtaining self-employment is sufficiently high that it is worthwhile for the worker to do
an apprenticeship. Otherwise, it is not. As expected, the higher the cost of apprenticeship,
the higher the cutoff probability, so that fewer workers apprentice. Conversely, the greater
the returns to apprenticeship $(m - n)$, the lower the cutoff probability. The single cutoff
value reflects the fact that the benefits are increasing, and the costs are constant in $\rho$.

In the final stage of the backward induction, the firms choose the cost of
apprenticeship, given the apprentice’s optimal decision rule above. As noted earlier, there
are two cases to consider. If certain firms are capable of offering apprenticeships while

\textsuperscript{18}While this statement is fairly obvious, the formal mathematical proof is in the appendix.

\textsuperscript{19}In the simplest version of the problem, the worker decisions regarding quitting and joining self-
employment are trivial. If the worker receives a disutility shock, he quits, and if he receives the funding for
self-employment, he starts his own firm. The former is a mathematical simplification that simply affects
the probability of quitting. The latter accurately reflects the reality of apprentices in Ghana, and reduces
the self-employment decision to a single dimension.
other firms are not, then there is no free entry in the first period (although free entry is still allowed in the production period). In this case, firms may earn positive profits, and the apprenticeship fee will be optimally set to maximize profits. In this case, the firm’s problem becomes:

\[
Max \Pi = \int_{m-n}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - n)) dF(\rho)
\]  

(8)

The first-order condition for an interior solution to this optimization problem is:  

\[
[1 - F\left(\frac{c}{m-n}\right)](1) - (c - t)f\left(\frac{c}{m-n}\right) \frac{1}{m-n} - (1 - \lambda)(m - n)(1 - \left(\frac{c}{m-n}\right))f\left(\frac{c}{m-n}\right) \frac{1}{m-n} = 0
\]  

(9)

where \(f\) is the density function associated with distribution function \(F\). In the firm’s optimization problem, it is trading off the different impacts of increasing the apprenticeship fee. The first term reflects the firm’s receipts from apprenticeship fees. The second term reflects the fact that as these fees are raised, the net benefit to the firm, \(c - t\), is collected over fewer apprentices. The third term reflects the fact that the larger the apprenticeship fee, the smaller the fraction of workers over which the firm has monopsony power in the second period. These are the trade-offs that the firm is making, the solution of which, under the assumption of uniform density, is:

\[
c = \frac{t + \lambda(m - n)}{1 + \lambda}
\]  

(10)

Firm profits in this case are \(\Pi = \frac{(m-n-t)^2}{4\lambda(m-n)}\).

Of course, in the case where free-entry is allowed, the apprenticeship fee will be

\textsuperscript{20}That this interior solution is the solution for the profit-maximizing firm is proved in Appendix A.
determined by the zero-profit condition. In that case, we get:

\[ 0 = \int_{\hat{\rho}}^1 (c - t + (1 - \rho)(1 - \lambda)(m - n))dF(\rho) \]  

(11)

where \( \hat{\rho} = \min\{\max\{\frac{c}{m-n}, 0\}, 1\} \). Define the right-hand side of (11) as \( h(c) \).

Simplifying this expression reveals that \( h(c) \) is a piecewise-defined continuous function which is quadratic for \( c > 0 \), and linear for \( c < 0 \) (see Appendix A).

The roots of this quadratic function, which Appendix A demonstrates is concave in \( c \), are:

\[ c = m - n, t - \frac{(1 - \lambda)(m - n)}{2} \]  

(12)

The larger of these two roots, \( m - n \), is clearly positive. At this level of the apprenticeship fee, the profits are zero because no individual chooses to apprentice. The benefits from apprenticeship, which may not be captured, are the same as the costs, which must be paid. Therefore, this is not an equilibrium root. Clearly, firms will compete with each other, charging lower fees, until the zero-profit condition is again satisfied at the lower root, \( t - \frac{(1 - \lambda)(m - n)}{2} \), which may be positive or negative, depending on parameter values.\(^{21}\)

As expected, in the case of free entry, the apprenticeship fee is lower than the monopolist case, and may even be negative (see Appendix A). If the fee, \( c \), is negative, this means that the firm is willing to pay apprentices for the opportunity to potentially have monopsony power over them in the second period. This will occur if the potential benefit of apprenticeship \( (1 - \lambda)(m - n) \) is large relative to the firm’s training cost, \( t \). If firms pay apprentices for the apprenticeships, then this would induce all individuals to engage in apprenticeship. Whether the apprenticeship fee is positive or negative then is an empirical

\(^{21}\)Note that in this solution, the firms have already induced all workers to apprentice as soon as \( c = 0 \). However, at this level of \( c \), profits are still positive, and so \( c \) falls to become negative for profits to fall to 0.
question.

**Proposition 1:**

*An equilibrium exists, both in the case of free-entry, and for a monopolist firm (in the market for apprentices), and is described by either the solution to (12) or (10), respectively, as well as the solution to (7), the quitting and self-employment decisions, and the wage decisions of inside and outside firms, as outlined in the text.*

The proof of the proposition is in Appendix A. The characteristics of this equilibrium that are of interest regard the apprenticeships. In particular, former apprentices who apprenticed in this firm are more productive in this firm than workers who apprenticed elsewhere (reflecting the specific human capital assumption), but are not paid more than other workers. This fact reflects the monopsony power that apprenticeship firms have over their former apprentices. In this model, this monopsony power arises not because of any information asymmetry, but simply because of the specific nature of apprenticeship knowledge. This monopsony power in the second period motivates the firm to take on apprenticeships. For workers, the returns to apprenticeship lie in self-employment. Apprentices who remain in the firm, or move to another firm, after the apprenticeship are not rewarded for their apprenticeship. The only reward comes from starting one’s own firm and replicating the technology and business practice of the apprenticeship firm. Therefore, those workers who have a sufficiently high probability of obtaining the capital required for self-employment will engage in apprenticeship. These predictions can be taken to the data to test the validity of this model.

While this model reasonably describes the Ghanaian situation, it is also interesting
to explore the equilibrium that would result if the wage offers did not need to be
time-consistent, but could be enforced by some contractual mechanism, or through
reputation incentives in a repeated game context. In such contexts, the firm might be able
to commit, ahead of time, to a wage that is higher than the apprenticed worker's outside
option. Before analyzing the mathematics, it is worth noting some theoretical difficulties
with a reputation-enforced contract in the African context. Fafchamps (2004) explains in
considerable detail the reasons why reputation effects do not work well in Africa, reasons
that lead him to conclude that "a reputation mechanism alone is likely to be insufficient for
enforcing contracts." (p. 15) This is partly related to the high transaction costs and
considerable uncertainty in Africa, that make it difficult to determine whether the breach
of a contract was the result of what Fafchamps refers to as an "excusable breach" or not,
that is "whether breach of contract was due to unanticipated events or to carelessness and
incompetence." (p. 33). His analysis suggests that we should be quite surprised if we find
evidence of reputation-supported contracts in the data. With that caveat in mind, the
mathematics of such a case is outlined in Appendix C. In this case, the firm sets the wage
for former apprentices, $W_F$, at the same time as the apprenticeship fee, $c$. As above, the
profit-maximizing and free-entry equilibria are both analyzed. For the profit-maximizing
case that allows for time-inconsistent offers, the firm will set the wage for former
apprentices at $W_F = m$, with $c = \frac{m - n + t}{2}$. Therefore, workers who have apprenticed in the
current firm should both be more productive and also make higher wages than workers who
have apprenticed elsewhere. In this version of the model, if workers receive higher
remuneration in self-employment, that remuneration is only marginally higher. We might
assume that workers would seek self-employment for additional psychic benefits it might
offer, but the workers would not receive significant additional remuneration. Therefore, this case, for time-inconsistent wage offers, has different predictions from the benchmark version of the model. Naturally, the empirical results will partly speak to which version of the model is preferred, with the theoretical difficulties with reputation effects in Africa already noted.

In fact, if the results of this paper support the model presented here (as we will see they will), we can interpret the difference between the time-consistent model of Appendix A and the time-inconsistent (reputation-enforced) model of Appendix C in the context of Fafchamps’s analysis. Firm profits are higher in the time-inconsistent (reputation-enforced) model \( (\Pi = \frac{(m-n-t)^2}{4\lambda(m-n)}) \) than in the time-consistent model \( (\Pi = \frac{(m-n-t)^2}{2(m-n)(1+\lambda)}) \), by a (positive) amount of \( \frac{(1-\lambda)(m-n-t)^2}{4\lambda(1+\lambda)(m-n)^2} \). For Fafchamps, firms’ inability to use reputation in order to enforce contracts is a major burden on firms in Africa. In this model, the loss of profits in the time-consistent (actual) model, as opposed to the reputation-enforced (desired) model can be interpreted as a measure of exactly that burden.

In the free-entry, time-inconsistent (reputation-enforced) equilibrium, the solution set of \( (c, W_F) \) choices consist of a segment of a hyperbola in the \( (c, W_F) \) space, as demonstrated in Appendix C. That segment joins the endpoints \( A \left( \frac{2t-(1-\lambda)(m-n)}{1+\lambda}, n \right) \) and \( B(t, m) \). Point \( A \) is the equilibrium point for zero-profits in the time-consistent case described above. Therefore, the entire set of \( (c, W_F) \) solution values lie in the ranges: \( c \in \left[ \frac{2t-(1-\lambda)(m-n)}{1+\lambda}, t \right] \), and \( W_F \in [n, m] \). That is, the wage to former apprentices may lie anywhere between \( m \) and \( n \), with the apprenticeship fee adjusted accordingly. For zero profits, if \( W_F \) is raised, the apprenticeship fee is raised as well. In this zero profit case, workers will still always get their highest wage in self-employment, but that may also be
equal to their wage offer from the firm (if the equilibrium is at the $B$ endpoint). Therefore, workers that apprenticed in the current firm will remain more productive than apprentices from elsewhere, and they may also be paid more than apprentices from elsewhere (but they also may not be).

A further extension to the model would be to consider the competitive impact to the master of training apprentices. If the idiosyncratic technology described in this model produces an idiosyncratic output, then by training apprentices, the master is increasing his competition, particularly if his apprentices remain local. This case is considered in Appendix B. In the case where apprenticeship occurs in this model, the central predictions of the benchmark model remain the same.

4 Data

Two different data sets are used to explore the questions of this paper. The first dataset is the Ghana Living Standards Survey (GLSS), which is the nationally representative household survey for Ghana. The version of the GLSS used for this paper is Wave 4, which was conducted between April 1998 and March 1999. This survey enables an examination of the apprenticeship institution in the overall context of Ghana. The second dataset is a panel dataset of manufacturing firms in Ghana. This survey was initially conducted as part of the Regional Program on Enterprise Development (RPED) surveys of manufacturing in African countries, coordinated by the World Bank, and organized in Ghana in conjunction with Oxford University and the University of Ghana (Legon). After the RPED program ended, the survey was continued as the Ghanaian Manufacturing Enterprise Survey (GMES), organized by Oxford University in conjunction with the Ghana
Statistical Service.\textsuperscript{22} This dataset is particularly rich for examining the apprenticeship institution, as information was collected not only on the firms themselves, but also on a subsample of the firm’s workers, and a subsample of the apprentices at the firms. This allows a comparison of those who apprenticed within this firm with those who apprenticed in a different firm, both from the perspective of their productive contribution and their remuneration. The details of this approach will be outlined in the following section. This manufacturing survey dataset has been used by other authors (e.g. Teal, 1996) to explore labour market issues. Jones (2001) used this dataset to explore the flexibility of the labor market in Ghana, and found evidence of a flexible, competitive labor market, as productive characteristics such as education and experience were compensated according to their productive contributions. This study confirmed a previous result from Beaudry and Sowa (1994), who also found evidence of a flexible, competitive labor market in Ghana. While the manufacturing dataset is more detailed for some of the questions we wish to explore in this paper, the GLSS is nationally representative of households, and therefore will frequently serve as a reference point. The education variable is taken as the total number of years of schooling, and the experience variable is calculated as:

\[\text{age} - \text{education} - \text{years of apprenticeship} - 6.\]

The earnings variable is taken as the log of the hourly wage, measured as a weighted average of the hourly wage in each of the worker’s employment activities (if the worker is engaged in more than one).\textsuperscript{23}

\textsuperscript{22}Unfortunately, while a rich set of questions were asked in the survey, the questions regarding apprenticeship were not asked in all years of the survey. Therefore, the survey years available for use for the purposes of this paper are 1991, 1992, 1993, 1998 and 1999.

\textsuperscript{23}Specifically, for workers that are self-employed in their primary activity, the hourly wage is calculated as the weighted (by hours) average of the hourly wage in all of the worker’s self-employment activities. Conversely, for workers that work in employee wage work in their primary activity, the hourly wage is calculated as the weighted (by hours) average of the hourly wage in all of the worker’s wage work activities. Therefore, all employment activities are included for completeness, but the main results remain the same.
The significance of the apprenticeship institution is clear when examining the summary statistics for the GLSS in Table 1. Overall, 12% of all adults\textsuperscript{24} have been an apprentice in a field other than tailoring, while 8% have apprenticed in tailoring. These percentages do not include the 5% of individuals over the age of 15 who are currently engaged in an apprenticeship. Table 1b limits the summary statistics to those who are either working in non-farm self-employment, or working for a wage. We shall refer to this group as the labor force. The labor force is more educated than the overall population, and earns more on average as a result, and the apprenticeship institution is even more significant within the labor force. In the model presented in the previous section, the more significant returns to apprenticeship come in self-employment. The data for the labor force as a whole are suggestive of this fact, as self-employed former apprentices have average annual earnings that are roughly 21% higher than former apprentices working as employees, even though they are both younger and less educated on average than their employee counterparts.

Similar patterns are evident when the sample is restricted to examine just those workers within the non-textile manufacturing sector, as seen in the GLSS data in Table 1c.\textsuperscript{25} This sub-sample is broken down for illustrative purposes to compare it to the

\textsuperscript{24}Throughout this paper, the term adult shall actually refer to all those who are at or over the age of 15. While this is a younger age than the legal definition, the data demonstrate that it is the most sensible cutoff age for gainful employment.

\textsuperscript{25}The manufacturing sectors which are sampled in the GMES survey include wood-working, metal-working, textiles and garments and food processing. This paper will not examine the textiles and garments sectors or tailoring apprenticeships, as these will be the subject of a separate paper by this author. Therefore, when the manufacturing sector is referred to in this paper, it shall describe the sectors listed above other than textiles and garments.

While a bijection does not exist between the manufacturing sector classified by the GLSS and the manufacturing sector of the GMES survey (see Data Appendix for details), the correspondence is close enough for some useful comparisons. It will be evident from the text whether the manufacturing sector of the GLSS or the GMES is being referenced.
manufacturing sector of the GMES survey. First, it is worth noting that manufacturing workers are less educated, and earn less than the average member of the labor force.\textsuperscript{26} However, they are even more likely to have apprenticed than other workers, with roughly 37\% of workers in non-textile manufacturing having apprenticed. Here, the returns to apprenticeship appear to be seen in self-employment. Self-employed former apprentices earn about 49\% more per year than those serving as employees (compare 480546 cedis to 321465 cedis). Again, this is true despite the fact that they are slightly less educated and slightly younger than their employee counterparts. While this provides some very preliminary evidence that is consistent with the theoretical model, these are merely summary relationships of means, and should therefore be explored further.

The summary statistics for workers in the GMES survey are provided in Table 2. Also in Table 2 are listed the comparable statistics for manufacturing workers within the GLSS. Given that the GMES worker sample under consideration does not include self-employed workers or apprentices, these have been removed from the GLSS sample for appropriate comparison. We find that the GMES sample is slightly more educated than those in the GLSS sample. On the other hand, more workers have done apprenticeships in the GLSS sample (40\%) than in the GMES sample (25\%). The higher level of education in the GMES sample leads to slightly higher earnings in that the average annual earnings of GMES workers is about 420 000 (in 1991 cedis), while the annual earnings of GLSS workers is approximately 406 000 (1991 cedis).

Further information was captured in the GMES survey about worker opinions and

\textsuperscript{26}The years of schooling must be interpreted in the Ghanaian context. In Ghana, there are 6 years of primary school. Until the 1990s, this was followed by 4 years of middle school, but changed to 3 years of middle school in the 1990s. Therefore, 10 years of schooling reflects a middle-school education (for which no entrance examination is required at any level).
attitudes, which can usefully address some aspects of our model. An important prediction of the model is the fact that individuals who have apprenticed will make a higher wage in self-employment than they would working either in their current firm or in another firm. Of the apprenticed workers (workers who have completed an apprenticeship) in the GMES survey, 77% stated that they would prefer to be self-employed. When those were questioned as to why they would prefer to work for their “own account”, 64% said that it would give a higher income. Furthermore, when this group was asked why they have not started up a business as of yet, 99.54% responded that it was because of a lack of capital.27 On the other hand, among those who would rather not work for their own account, the primary reason given was that self-employment would lower their job security.28 Therefore, in terms of their own explanations for their actions, the preference for self-employment among former apprentices, as well as the constraint on this self-employment (access to capital) are consistent with their treatment in the model presented. While we do not have longitudinal data on former apprentices to be able to capture exactly what their activities are at the completion of the apprenticeship, in the GMES Survey we asked the firm owners what were the activities of former apprentices who had apprenticed within the firm. Specifically, of those employees who had finished their apprenticeships last year, we asked how many of them were engaged in various activities. The results in Table 3 present the overall percentages of apprentices engaged in different activities, according to their former masters.29 The largest fraction, 38%, continued working within the firm, while 29% had already started their own business. Another 18%

27 This is also consistent with the literature on this subject, as discussed in the previous section.
28 This response attracted 25% of the answers; the responses to this question were more evenly spread.
29 This, of course, is not the same as the average of the fractions reported by firm, although in reality these numbers are virtually the same.
were working for another firm somewhere. Given that the firm owners did not know about the activities of 18% of the apprentices, these fractions should in all cases (except the case of apprentices continuing to work in the firm) be seen as a lower bound.

In addition to information about workers in the GMES survey, information was also captured about the apprentices at these firms, which can highlight some of the details of this institution. The length of the apprenticeship is typically about 3 years (with the average being just over 3 years in length in the GMES data). Many apprentices pay fees at the beginning (70%, with the average value being 28700 in 1991 cedis) and at the end (56%, with the average value being 29600 cedis) of the apprenticeship.30 These fees are roughly the equivalent of a month’s salary for an average manufacturing worker. While apprentices receive a small allowance (sometimes in the form of food, clothing, or housing, as well as pocket money) over the period of their apprenticeship, the average value for this is 1300 cedis a month. Therefore, according to those apprentices aware of their fee structure, an average apprentice will receive allowances that cover the cost of his apprentice fees. Still, the ‘apprenticeship fee’, as outlined in the model of the previous section, is the net transfer from the apprentice to the firm owner once all of the in-cash and in-kind contributions are taken into account. As outlined in the previous model, the sign of the apprenticeship fee may be positive or negative, without affecting the main predictions being tested in the data. On average, apprentices are 22 years old, and are slightly less educated than their worker counterparts (compare 9.86 years to 11.2 years).

30 These values are calculated based on those apprentices who actually knew the value of their entry or exit fees, as the previous discussion suggests. That caveat should be noted before placing too much weight on these numbers.
5 Results

The model outlined in Section 2 is a model of apprenticeship as specific human capital, where workers are rewarded for this apprenticeship if they become self-employed. As outlined in the previous section, the fact that self-employed former apprentices earn more than employee former apprentices, despite their slightly lower education levels, is consistent with this model. This section will seek to explore the explanation of individual earnings a bit more carefully.

The basic model that will be used for exploring wages is the standard human capital model of Mincer (1974), which predicts the following expression for the wage, supplemented with an apprenticeship variable:

$$\log E = \beta_0 + \beta_1 S + \beta_2 X + \beta_3 X^2 + \beta_4 \times App + \varepsilon$$  \hspace{1cm} (13)

Here, \(\log E\) is a function of the years of schooling of the individual, as well as a quadratic in his potential experience and a dummy variable for whether the individual apprenticed, \(App\). Here, the returns to apprenticeship are a function of the coefficient estimate \(\beta_4\), with the increase in a person’s wage resulting from apprenticeship being \(e^{\beta_4} - 1\). Running a simple least-squares regression on the above equation runs into the problem of ability bias. Ability is an omitted variable in the above specification, and therefore included in the error term at the same time that it is likely to be correlated with an individual’s schooling level. While another paper by this author develops a methodology for handling this ability bias (Frazer, 2003), it only works when one has linked employer-employee data, which is not the case within the GLSS. Therefore, unfortunately, we are not able to address this potential difficulty.
A further issue for this paper is that we expect the returns to a person’s characteristics to differ by sector. In particular, the model presented in this paper predicts that the returns to apprenticeship will be different for self-employment than for employee work. In fact, we also should not restrict the other coefficients of the wage equation to be the same across self-employed and waged workers. Given that wage determinants such as the returns to education are allowed to differ across sectors, individuals may prefer wage or self-employment, depending on their expected wage for their characteristics. That is, people may not be distributed randomly across wage work or self-employment, resulting in a selection problem. The process of selection into self-employment employed here follows the methodology of Heckman (1979) and in particular Lee (1978), who examines the switching between union and non-union jobs. In the model, individuals choose to participate in the sector that brings the greatest return to their bundle of characteristics. The first stage estimation uses a probit regression to estimate the probability of being self-employed. The results of this stage are used to compute selectivity correction factors (inverse Mills ratios) for inclusion in the second stage. The inclusion of these selection correction terms in the second stage controls for the issue of selection between self-employment or wage work. Moreover, the significance of the selection correction (λ)

31 A further selection issue is the issue of selection is selection into the labour force, in our case the non-agricultural labour force (not including farmers). Although we have data for the above variables for both employees and self-employed individuals, we need to account for the fact that some individuals choose to either work in home production or in farming, or are unemployed, and will not be included in this sample. To handle this, we used a standard Heckman two-step procedure for the selection into the labour force. A variety of sets of standard identifying variables were used to attempt to capture selection into the overall labour force, including non-labour income, land area owned by the household, as well as the number of children in different age groups. In none of these specifications was there evidence of selection into the workforce (i.e. the coefficient on the inverse Mills ratio was never significant), and the standard errors of the variables increased significantly upon handling selection. For this reason, the selection into the overall labour force is not handled. While technically this will require an interpretation of the variable coefficients (e.g. the returns to education), as being the returns within the labour force, the lack of evidence of selection is suggestive that these coefficients can be interpreted more broadly.
term is a test for the existence of selection into a particular sector. While technically the procedure is identified from the assumption of normality in the probit regression, this is a fairly strong assumption, and so it is generally preferable to also have an exclusion restriction. That is, the model is most cleanly identified if there are one or more variables that should affect the probability of being self-employed without affecting the wage independently of the choice on self-employment. While the model predicts that household assets should be such a variable, this variable did not satisfy the over-identification restrictions. Therefore, the variables that are used only in the selection equation are father’s education and occupation. Father’s education and occupation are likely to affect whether an individual becomes self-employed (as well as very likely their level of education), but these variables are typically omitted from wage equations. In addition to F-stat and over-identification restriction tests on these variables, the test for normality will also be performed in this case.

The results from this procedure are presented in Table 4. The first three columns do not handle the selection issue, and use least-squares estimation. The first column suggests that the return to apprenticeship is 0.184 \((e^{\beta_4} - 1)\). However, part of that is capturing a gender effect, as once we include the gender of the individual, the return drops to 0.071 and is no longer significant. Our theory suggests that the wage returns to apprenticeship should occur within the self-employed sector, and not necessarily within the employed sector. It should be noted that according to the model, a higher return to apprenticeship within the self-employed sector does not imply necessarily that former apprentices are more productive within self-employment. Rather, this higher return to apprenticeship in self-employment is mainly because former apprentices become owners of the firm’s portion
of their specific capital when they become self-employed.\textsuperscript{32} The theory appears to be confirmed once we restrict the estimation to those who are self-employed. Here, the return to apprenticeship is estimated to be 0.132 (p-value 0.022) for the full self-employed sample (column (4)), and 0.108 (p-value 0.084) on the restricted self-employed sample (column (5)) (restricted to those for which there is data on father’s education and occupation). Column (6) presents the results once the issue of selection is handled. The first-stage probit results are presented in Appendix Table A1. While the F-stat on the excluded father variables is significant with a p-value of .014, it is relatively small at 2.72. Therefore, we also test the assumption of normality. Davidson and MacKinnon (1993) suggest the method of Pagan and Vella (1989) for testing the normality assumption.\textsuperscript{33} Based on their test, we cannot reject the normality assumption, as reported in the table (p-value 0.138). The excluded variables also pass the usual over-identification test (Davidson and MacKinnon, 1993, p. 236), and a F-test on their exclusion from the second stage, at the 5% level. The coefficient on the selection term is a test for selection into self-employment. This coefficient is significant (p-value 0.022) suggesting that it is important to control for this selection. Once the selection is controlled for, the returns to apprenticeship rise to 0.189. Nevertheless, while the formal tests for identification (exclusion and over-identification restrictions and tests for normality) were passed in the selection model, they were not passed with sufficient strength to make a strong statement about 0.189 as an exact

\textsuperscript{32}In fact, formally in the model, the former apprentices in self-employment will be more productive than those in wage employment because the wage employees include a mix of those productively using their apprenticeship skills working within the same firm, and those less productively using them in a different firm.

\textsuperscript{33}This method involves including higher-order powers of the predicted values from the first-stage equation, weighted by the Mills ratio in the second stage. Departures from normality will result in these terms being significant under an F-test when included in the second stage.
measure of the returns to apprenticeship. Therefore, overall, the regressions for the self-employed should be taken as presenting evidence of returns to apprenticeship within self-employment, without taking a strong position on the exact size of those returns.

The similar regressions for wage workers are performed in columns (7) to (9). Here, again, the normality and over-identification conditions are both passed. The lack of significance of the selection correction term in column (9) suggests that selection is less of an issue in wage employment, mirrored by the fact that the coefficients change less once selection is controlled for. In each of columns (7) through (9), the coefficient on apprenticeship is small and not significant, suggesting that the returns to apprenticeship clearly lie in self-employment. On the other hand, while the returns to education in wage employment are only marginally higher than in the full sample of columns (1) through (3), the evidence suggests that these returns are considerably higher than in self-employment. Those who are highly educated will be better off in wage work, while those who have apprenticeship training will receive higher returns in self-employment.

The results in Table 4 presented the overall wage returns to apprenticeship in Ghana,\textsuperscript{34} controlling for selection into self-employment. The next stage is to handle the return to apprenticeship within the manufacturing sector. These are provided in Table 5. The sample here merely includes manufacturing workers from the GMES survey. Here, apprentices are in fact paid less than non-apprentices among the manufacturing workers. In the theoretical model outlined in this paper, apprentices and non-apprentices are paid the same in wage employment. If higher ability people achieve higher schooling, and schooling and apprenticeship are substitutes, then the apprenticeship variable in the wage

\textsuperscript{34}at least within the labour force (see the earlier discussion).
equation may partly be capturing the fact that the former apprentices in the manufacturing sector (and in particular, those who remain in the manufacturing wage-work and have not yet become self-employed) may also have lower ability, or be less productive than the more highly educated workers in the sector. Even though education is controlled for in the regression, if this schooling variable is not completely capturing this schooling/ability dimension, the apprenticeship variable may partly be capturing this.

Examining the production function estimates later in this section will shed further light on this issue. In this specification, the gender coefficient is insignificant, likely because 87% of the sample is male. It is possible to separate those apprentices who have apprenticed within this firm, compared to those who have apprenticed elsewhere. This is done in the second column, and the coefficient is similar for both of these groups, as the model predicts. In an effort to better understand the reasons for these negative coefficients, as well as to test the productive implications of the model, the productive contribution of apprentices can be measured using production function estimation. This is the next step.

With linked employer-employee data, we can examine not only the remuneration of the characteristics of employees, but also understand the productivity of these characteristics. While Frazer(2003) provides the proofs for the methodology for including the productive characteristics of the wage equation into the production function, and shows that there is a one-to-one correspondence between the coefficients of the wage equation and those of the production function proposed, the important parts of this methodology are fully described below. Before examining the production function to be used in this case, the wage equation needs to be fully specified. The wage equations discussed in the specifications and GMES and GLSS results above is for workers (not apprentices) working
in these firms. At each of these firms, both apprentices and workers can be found working within the firms. The degree to which the apprentices are contributing to firm product is in fact a matter of empirical investigation. Some firms have more apprentices than they have workers, and therefore their productive contribution cannot be ignored. In order to allow for the apprentice characteristics to impact firm production in a manner different from that of regular workers, consider the following wage equation:

\[
\log E_i = \beta_0 + \beta_1 W_i \times S_i + \beta_2 W_i \times X_i + \beta_3 W_i \times X_i^2 + \beta_4 W_i \times T_i + \beta_5 W_i \times App_i (14)
\]

\[
+ \beta_6 A_i + \beta_7 A_i \times S_i + \beta_8 A_i \times X_i + \beta_9 A_i \times X_i^2 + \beta_{10} A_i \times T_i + \varepsilon_i
\]

where \( W_i \) is a dummy variable for whether an employee\(^{35} \) at a firm is a worker, \( A_i \) is a similar dummy for apprentices (\( A_i = 1 - W_i \)), and \( T_i \) represents tenure at the firm. Other variables have the same definition as previously. Note, in particular, that while \( A_i \) delineates whether an employee is an apprentice, \( App_i \) delineates whether a worker did an apprenticeship. Clearly, such a wage equation allows for completely separate sets of coefficients to be operative for workers and apprentices. Still, given that apprentices are typically in their early twenties in age, and have little if any work experience, we would not expect the experience variable to be important for them. On the other hand, the tenure variable for apprentices is kept, as it reflects the change in productivity of an apprentice over the course of the apprenticeship. The tighter version of the equation is then:

\[
\log E_i = \beta_0 + \beta_1 W_i \times S_i + \beta_2 W_i \times X_i + \beta_3 W_i \times X_i^2 + \beta_4 W_i \times T_i + \beta_5 W_i \times App_i (15)
\]

\[
+ \beta_6 A_i + \beta_7 A_i \times S_i + \beta_8 A_i \times X_i + \beta_{10} A_i \times T_i + \varepsilon_i
\]

\(^{35}\)The term employee is being used broadly here to include apprentices.
The appropriate production function that is consistent with this form of the wage equation is developed by first examining the firm’s wage bill, which is:

\[
\sum_{i=1}^{L} e^{\beta_0^*_i + \beta_1 W_i \times S_i + \beta_2 W_i \times X_i + \beta_3 W_i \times X_i^2 + \beta_4 W_i \times T_i + \beta_5 W_i \times App_i + \beta_6 A_i + \beta_7 A_i \times S_i + \beta_8 A_i \times T_i + \epsilon_i}
\] (16)

Under the assumption that individual wages are given by the Mincer wage formula, Frazer (2003) shows that the labor term in the production function for a profit-maximizing firm should look like the wage bill for the firm (differing only in the constant term). Therefore, the Cobb-Douglas production function which is consistent with the above wage bill is the following (after factoring the constant, and with \(\lambda_0^* = \lambda_0 + \beta_0^* \beta_H\)):

\[
y = \lambda_0^* + \log(\sum_{i=1}^{L} e^{\beta_1 W_i \times S_i + \beta_2 W_i \times X_i + \beta_3 W_i \times X_i^2 + \beta_4 W_i \times T_i + \beta_5 W_i \times App_i + \beta_6 A_i + \beta_7 A_i \times S_i + \beta_8 A_i \times T_i} \beta_H)
\] + \beta_k k + \omega + \varepsilon
\] (17)

where \(y\) is the log of a firm’s value added, \(L\) is the total number of workers and apprentices at the firm, \(k\) is the log of the capital stock of the firm, and \(\omega\) and \(\varepsilon\) are the firm’s productivity and the error term.\(^{36}\) In order to develop a usable form for the labor term of the production function, a first-order Taylor expansion to the labor term is taken, which yields:

\[
\beta_H \log L + \beta_1 N_W L S_W + \beta_2 N_W L X_W + \beta_3 N_W L X_W^2 + \beta_4 N_W L T_W + \beta_5 N_W L App + \beta_6 N_W L A + \beta_7 N_W L S + \beta_8 N_W L T
\] (18)

where \(N_A\) and \(N_W\), are the number of apprentices and the number of regular workers at a firm, respectively, and \(\bar{Z}_Y = \frac{\sum_{i=1}^{N_Y} Z_i}{N_Y}\) is the average level of variable \(Z\) for

\(^{36}\)The production function estimation follows closely the procedure of Levinsohn and Petrin (2003) which builds on the work of Olley and Pakes (1996). For a helpful description of the production function estimation and the productivity term in these papers, see Griliches and Mairesse (1998).
apprentices if \( Y = A \), and for workers if \( Y = W \). Overall then, the production function which is estimated is:

\[
y = \lambda^* + \beta_H \log L + \beta_H \beta_1 \frac{N_W}{L} \bar{S}_W + \beta_H \beta_2 \frac{N_W}{L} \bar{X}_W + \beta_H \beta_3 \frac{N_W}{L} \bar{X}^2_W + \beta_H \beta_4 \frac{N_W}{L} \bar{T}_W (19)
\]

\[
+ \beta_H \beta_5 \frac{N_W}{L} \bar{A}_{ppw} + \beta_H \beta_6 \frac{N_A}{L} + \beta_H \beta_7 \frac{N_A}{L} \bar{S}_A + \beta_H \beta_8 \frac{N_A}{L} \bar{T}_A + \beta_k k + \omega + \varepsilon
\]

The methodology used for estimation, in order to handle the issue of the simultaneity of firm productivity, and the choice of the variable inputs, is a slightly modified version of the Levinsohn and Petrin (2003) procedure, and is described in Appendix D. The average levels of the worker and apprentice variables are calculated based on the sub-sample of workers and apprentices at each firm.\(^{37}\) The employment variable is a measure of the total number of workers at the firm. The capital stock is a measure of the firm’s plant and equipment. The firm’s value-added is deflated by a firm-level price deflator. The production function coefficients are provided in Table 6, with the preferred specifications (that handle simultaneity using the Levinsohn and Petrin procedure) in columns (2) and (4).\(^{38}\) The coefficient on the variable controlling for whether a worker completed an apprenticeship is negative, suggesting that such workers are less productive than other workers, but the coefficient is not significant at any of the conventional levels. While this result might be a bit surprising, it is consistent with the wage results that former apprentices are paid less. When the “worker has apprenticed” variable is separated into those who apprenticed in this firm, versus those who apprenticed in a different firm, it

---

\(^{37}\) For example, the average schooling is calculated as a measure of the average level of schooling of those in the sub-sample interviewed at the firm. The sub-sample was designed to cover all occupational categories, and the average is weighted by the number of workers in each occupational category working at the firm.

\(^{38}\) Note that the worker tenure variable is not included in the final specification. When the worker tenure variable was included, the standard errors of all variables grew from multicollinearity. Given this fact, and because worker tenure is rarely used in the wage equation specification, it was excluded from the analysis.
is found that workers who apprenticed in this firm are no different in productivity from other workers. However, those who apprenticed in another firm are significantly less productive than those who apprenticed in this firm. This is consistent with the predictions of the model; the specific human capital acquired in apprenticeship is not transferable to a new firm. The firm’s ability to extract a surplus from those workers that it has apprenticed is consistent with the desire of employers to hire their former apprentices.

The other variables are generally of the expected sign. More educated workers are more productive. Similarly, the sign on the apprentice education variable is positive, but not significant. The sign on the fraction of apprentices indicates that apprentices may be less productive than workers, but not significantly so. The coefficient on the tenure of the apprentices is negative, but not significant. While the negative sign of this coefficient might be surprising, there are ways in which this is consistent with the apprenticeship institution. In the initial period of the apprenticeship, the apprentice is often doing odd jobs, and is very eager to be obedient in order to gain favour with the master.39 During this period, the apprentice may be as productive as an average unskilled worker in the enterprise.40 Later in the apprenticeship, when the apprentice is using the firm’s materials to learn the craft, and occupying more of the master’s time in actual training, the

39Daniel Bas (1989, p. 487) describes this situation in West Africa. “The general view, particularly in West Africa, is that it [apprenticeship] tends to be too long, and responsibility for this is usually laid at the door of the master craftsmen who are reluctant to carry out systematic intensive training and thereby deprive themselves too soon of cheap labour or create potential competition. The traditional cultural models also include a long period of ‘socialisation’ during which the apprentice learns to win the acceptance of the master craftsman and his family by demonstrating his respect and loyalty, thereby progressively meriting the right to learn.” As Ulrich Boehm describes it, “Apprenticeship usually starts with simple labour including cleaning, washing, fetching water or subsistence agriculture. 62 per cent of the masters ask their apprentices for this kind of service beside the work in their trade. Later the apprentice starts learning to know the different tools and work processes of the trade.”

40Bas (1989) continues, “During the [initial] socialisation stage the apprentice works as a messenger boy, doing odd jobs sometimes having no connection with the trade, such as running errands.” This messenger boy activity is more productive for the firm than the period of active training of the apprentice.
contribution of the apprentice actually decreases. Still, this variable is not significant.

The reason then why individuals do apprenticeships is for the potential returns that can accrue to apprenticed workers that become self-employed (specifically, an increased wage of roughly 19%).\footnote{A more careful analysis of the net present discounted value of apprenticeship is provided in Appendix E.} On the other hand, why do the employers train apprentices? Our model suggests that the firm’s monopsony power over its former apprentices motivates such training. A further related reason, as suggested in the other economics article on apprenticeships in Africa, is that apprenticeship fees can be a significant form of finance for some firms (Velenchik, 1995). The evidence regarding how much less productive apprentices are in comparison to other workers is not clear, with the coefficient of interest not significant in the preferred regression. This fact, coupled with the fact that apprentices are paid significantly less than other workers, suggests that while the productive returns of apprentices are less than other workers, they appear to be on net positive for the firm, although the evidence is weak. Given the errors of the estimates, and the various costs and benefits of apprenticeship, as outlined earlier, the net benefit for the firm cannot definitely be signed.

However, another question posed to the firms is of interest on this fact. Firm owners who trained apprentices were asked if they had ever refused an apprentice. Fifty-eight percent responded affirmatively. In other words, forty-two percent of those firms that trained apprentices had never refused an apprentice applicant. Naturally, there is some endogeneity in the answer, in that potential applicants may not apply for apprenticeship positions that they know they will not get. Moreover, there may be some selection issues involved in that the queues may be longer at places that offer ‘better’ apprenticeships.
Still, these responses are concordant with the model where firms are able to set the apprenticeship fees at a level that makes them at worst indifferent to apprenticeships.

Finally, it is worth returning to the question of whether firms are able to enforce the optimal long-term wage contract, through reputation effects or other means. As outlined in the model section, an alternative model is that where long-term contracts are fully enforceable, and is provided in Appendix C. In this alternative model, those who have apprenticed within the current firm are more productive and paid more than apprentices from other firms. The empirical result that apprentices from the current firm are not paid more than apprentices from other firms is not consistent with this model. Therefore, the paper also provides some evidence against full contract enforceability in the Ghanaian context (by reputation or other means), and therefore supports the analysis of Fafchamps (2004) that enforcing contracts is difficult in the African environment. Moreover, this paper provides a very concrete example where such difficulty can matter.

In summary, the results are consistent with the model of apprenticeship as training in specific human capital. Apprenticed workers remaining within the firm are more productive than apprenticed workers from other firms. However, because their outside option is no better than apprenticed workers from elsewhere, their wages are not higher than apprenticed workers from other firms. For this reason, the returns to apprenticeship are in self-employment, where a former apprentice basically replicates both the technology and business practice of the apprenticeship firm. Indeed, apprentices are very eager to enter self-employment, and are constrained only by capital from becoming apprenticed entrepreneurs.
6 Appendix A

Proof of Proposition 1:

The equilibrium concept in use is that of a subgame perfect Nash equilibrium. To prove the equilibrium, what remains is to show the optimality of the decisions at each step. That the optimal choice of wage for former apprentices is $W_F = n$ in the first order condition for equation (5) is fairly obvious, but for completeness, consider the Lagrangian of the optimization problem of the firm’s optimal choice of $W_F$:

$$L = \int_{\hat{\rho}}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - W_F))dF(\rho) + \delta(W_F - n) \quad (20)$$

First, note that $\frac{\partial W_F}{\partial W_F} = 1$ everywhere and so the constraint qualification clearly holds. Note also that $\frac{\partial L}{\partial W_F} = 0$, since $W_F > 0$. Therefore,

$$\frac{\partial L}{\partial W_F} = (1 - \lambda) \int_{\hat{\rho}}^{1} (1 - \rho)(-1)dF(\rho) + \delta = 0.$$

This gives that $-(1 - \lambda)(\rho - \frac{\nu^2}{\tau})\hat{\rho} + \delta = 0$, since $\rho \sim U(0, 1)$. So, $\delta = (1 - \lambda)(\frac{1}{\tau} - \hat{\rho} + \frac{\nu^2}{\tau}) > 0$. Therefore, the constraint binds and $W_F^* = n$. The optimality of the wage decisions for outside firms, as well as the quit and self-employment decisions of workers has already been outlined in the text.

What remains is to show that the firm’s choice of apprenticeship fee is consistent with optimizing behavior. There are two cases to consider. In the first case, free entry into providing apprenticeships is allowed. In the second, opposite case, there are restrictions against entry into providing apprenticeships, and so the firm has monopoly power for setting the apprenticeship fee.

First, consider the firm’s profit function:

$$\Pi = \int_{\frac{c}{m-n}}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - n))dF(\rho) \quad (21)$$
Since $\rho \sim U(0,1)$ (by assumption), as noted in the text, the profit function can be simplified to:

$$\Pi = -\frac{1 + \lambda}{2(m - n)} c^2 + \frac{\lambda(m - n) + t}{(m - n)} c - t + \frac{(1 - \lambda)(m - n)}{2}$$

(22)

for $c \geq 0$, and:

$$\Pi = c - t + \frac{1}{2}(1 - \lambda)(m - n)$$

(23)

for $c \leq 0$ (as the lower bound of the integral cannot fall below 0). Since both pieces (22) and (23) of the $\Pi$ function have the same value at $c = 0$, the profit function $\Pi$ is continuous, and is a quadratic for $c > 0$, and is linear for $c < 0$. For $c > 0$, the function is strictly concave in $c$, as $\Pi''(c) = \frac{-(1 + \lambda)}{(m - n)} < 0$, since $\lambda \geq 0$, and $m > n$ (assuming that apprenticeships are at least marginally productive). For $c < 0$, $\Pi'(c) = 1 > 0$, and $\Pi$ is linear. Overall then, the piecewise-defined function (22) and (23) is globally concave (but not strictly concave), and so it is maximized when the first-order condition, $\Pi'(c) = 0$, i.e. $-(1 + \lambda)c + \lambda(m - n) + t = 0$. Therefore, the $\Pi$ function has a single maximum in the first quadrant, where $c > 0$.

Consider then the first case, that of a profit-maximizing monopolist. The unique choice of apprenticeship fee is:

$$c^*_{\text{monop}} = \frac{t + \lambda(m - n)}{1 + \lambda}$$

(24)

Since the function is globally concave, the second-order conditions are satisfied, and this is a unique maximum. Recall that in this equilibrium, $\hat{\rho} = \frac{c}{m - n}$. To confirm that this is a valid probability ($0 < \hat{\rho} < 1$), note that $\frac{c^*}{m - n} = \frac{t + \lambda(m - n)}{(1 + \lambda)(m - n)} > 0$ always and $\frac{t + \lambda(m - n)}{(1 + \lambda)(m - n)} < 1 \iff t < m - n$, which is assumed true (that the cost of training to the firm
is less than the increase in productivity resulting from the training). Note that, since \( m > n \), that \( c > 0 \), and the apprentice pays the fee to the firm. The profit in the profit-maximizing case is \( \Pi_{\text{monop}} = \frac{(m-n-t)^2}{2(m-n)(1+\lambda)} \).

Now, consider the second case, where there is free-entry, resulting in zero profits. The profit function defined by pieces (22) and (23) has two roots, specifically \( c = m - n, t - \frac{(1-\lambda)(m-n)}{2} \). Firms have no incentive to set the apprenticeship fee to the level of the larger root, \( m - n \). Here, profits are zero because no individual has an incentive to apprentice as the apprenticeship fee is as large as the (uncertain) returns to apprenticeship. At the larger root, firms have an incentive to reduce the apprenticeship fee and increase profits (moving up the right arm of the quadratic). However, the lower root represents a stable equilibrium. Here, if the firm reduces the apprenticeship fee, it will earn negative profits. However, it can’t raise the apprenticeship fee without losing all of its apprentices to competing firms. Therefore, the unique choice of apprenticeship fee for the free-entry equilibrium is:

\[
  c_{FE}^* = t - \frac{(1-\lambda)(m-n)}{2} \tag{25}
\]

First note that \( c_{FE}^* \) may be positive or negative, depending on parameter values. While the fact that \( c_{FE}^* < c_{\text{monop}}^* \) is clear from the sketching of the profit function, it can also be shown algebraically, using the fact that \( t < m - n \), and showing algebraically that \( c_{FE}^* < c_{\text{monop}}^* \) is equivalent to \( t < \frac{1+2\lambda-\lambda^2}{2\lambda}(m-n) \). Since \( 1 + 2\lambda - \lambda^2 > 2\lambda \), as \( 0 < \lambda < 1 \), this result will always hold.
7 Appendix B

In this appendix, we consider the case where the firm takes into account the competitive implications of providing apprenticeships. That is, in this model, providing apprenticeships has the potential to increase the firm’s competition, specifically from those former apprentices who become self-employed. This is handled in the model by including this competitive cost in the firm’s profit function. It should be noted that such a cost is not already included in the firm’s cost of training the apprentices, $t$. The training cost $t$ applies to all apprentices, while the competitive cost to the firm of training apprentices will apply only to those former apprentices that become self-employed. Therefore, this case needs to be handled separately. The firm’s new profit function then is:

$$\Pi = \int_{\hat{\rho}}^{1} (c - t + (1 - \rho)(1 - \lambda)(m - W_F) - \rho q)dF(\rho)$$

subject to the constraints that $\hat{\rho} \geq 0$, and $\hat{\rho} \leq 1$, and $W_F \geq n$. As in the model outlined in the main paper and Appendix A, firms will again choose $W_F = n$, and for apprenticeship, the equilibrium cut-off for $\rho$ will be $\hat{\rho} = \frac{c}{m-n}$. Therefore $c \geq 0$, and $c \leq m-n$. The competitive cost, $\rho q$, for each apprentice, is a function of that apprentice’s probability of becoming self-employed. Workers with low probability of becoming self-employed impose little competitive cost, a priori, on the firm. This profit function simplifies to the following, again under the uniform density assumption, as before:

$$\Pi = c^2 \left[ \frac{q - (1 + \lambda)(m-n)}{2(m-n)^2} \right] + c \left[ \frac{t + \lambda(m-n)}{m-n} \right] + \frac{1}{2}(1 - \lambda)(m-n) - \frac{q}{2} - t$$

(27)

In this case the first-order condition for an interior solution is:

$$c \left[ \frac{q - (1 + \lambda)(m-n)}{(m-n)^2} \right] + t + \lambda(m-n) = 0$$

(28)
This simplifies to:

\[
    c = \frac{\lambda (m - n) + t}{1 + \lambda - \frac{q}{m-n}}
\]  

(29)

Therefore, the solution is similar to the previous profit-maximizing solution, with the additional consideration of the competitive cost in the term, \( \frac{q}{m-n} \). As this competitive cost, \( q \), increases, the cost charged by firms will increase, in order to reduce the number of apprentices that they train, and reduce the negative competitive impact. Mathematically, this follows from the following: 

\[
    \frac{\partial c}{\partial q} = \frac{(m-n)\lambda(m-n)+t}{[(1+\lambda)(m-n)-q]^2} > 0.
\]

Consider the second-order condition for the profit function: 

\[
    \frac{\partial^2 \Pi}{\partial c^2} = \frac{q-(1+\lambda)(m-n)}{(m-n)^2}.
\]

Consider the extreme case of this second-order condition first. If \( q \geq (1 + \lambda)(m - n) \), then the profit function is everywhere convex. The only extremum is a minimum, and it occurs when \( c < 0 \). Therefore, since the second-derivative is positive, profit will be maximized at the upper-bound for \( c \), namely, \( c = m - n \). Here, \( \hat{\rho} = 1 \), and no one apprentices. Therefore, the maximum profit is also a zero-profit equilibrium, and it involves no apprenticeship. This equilibrium will only occur if the competitive cost is sufficiently high, specifically, if the competitive cost, \( q \), per worker that gains self-employment is significantly larger than the productivity boost from apprenticeship, \( m - n \). If this competitive cost is sufficiently high, then firms would be forced to charge an apprenticeship fee that is prohibitive for potential apprentices.

On the other hand, if \( q < (1 + \lambda)(m - n) \), then \( \frac{\partial^2 \Pi}{\partial c^2} < 0 \), and the profit function is strictly concave. There are two cases to consider on this region. If \( q \geq m - n - t \) (and \( q < (1 + \lambda)(m - n) \)), then the constraint \( c = m - n \) will bind for profit maximization. That is to say that the best that firms can do is to make zero profits by inducing no apprenticeship, as in the case above where \( q \geq (1 + \lambda)(m - n) \). Here, while the competitive cost \( q \) is not as large as before, it is larger than the productive benefit of apprenticeship.
\[(m - n), \text{ less the costs of training} \ (t), \text{ and so apprenticeship does not make economic sense.} \]

On the other hand, if \[q < m - n - t\], then the optimal choice of apprenticeship fee will occur at an interior solution defined by \(c\) in equation \((29)\). Here, the second-order condition is also satisfied. Firm profits will be: \[\Pi = \frac{(m - n - t - q)^2}{2(1 + \lambda)(m - n) - q}\]. What is clear from the profit function is that the base model of Appendix A is a special case of this model when \(q = 0\).

We can also consider the case of the competitive equilibrium, where free entry results in zero profits for the firm. Consider the two cases. In the first case, if \[q \geq (1 + \lambda)(m - n)\], the profit function is convex, and as already noted, the zero profit equilibrium that will be chosen is also the profit maximizing equilibrium, where \(c = m - n\), and no one apprentices. Consider the more interesting case of the concave profit function where \(q < (1 + \lambda)(m - n)\). Therefore, for profits to be zero, consider the two roots of the above equation \((27)\): \[c = \frac{2t - (1 - \lambda)(m - n) + q}{1 + \lambda - \frac{m - n}{m - n}}, \quad c = m - n\]. In the case where \(q \geq m - n - t\), \(c = m - n\) is the lower of the two roots, but is also an upper bound on \(c\). As noted above, this profit-maximizing case will also be the zero-profit equilibrium for \(m - n - t < q \leq (1 + \lambda)(m - n)\). Now, consider the more interesting case where \(q < m - n - t\). This is the case where the competitive cost per worker that becomes self-employed is less than the productive increase in apprenticeship \((m - n)\) less the training cost \((t)\). In this case, the lower root will be \[c = \frac{2t - (1 - \lambda)(m - n) + q}{1 + \lambda - \frac{m - n}{m - n}}\), provided that \(\hat{\rho}\) remains positive. This will occur if \(q \geq (1 - \lambda)(m - n) - 2t\). Again, the upper root \(c = m - n\) will not be an equilibrium. Rather, as in the case in Appendix A, firms will compete with each other on the apprenticeship fee, lowering it until it reaches \[c = \frac{2t - (1 - \lambda)(m - n) + q}{1 + \lambda - \frac{m - n}{m - n}}, \text{ the other zero profit equilibrium.} \]

Now, consider the case where the \(\hat{\rho} = 0\) constraint binds, which will occur when \(q < (1 - \lambda)(m - n) - 2t\). Here, firm profits
are $\Pi = c - t - \frac{q}{2} + \frac{1}{2}(1 - \lambda)(m - n)$. Therefore, the zero-profit solution on this region has $c = t + \frac{q}{2} - \frac{1}{2}(1 - \lambda)(m - n)$.

Therefore, to summarize, the profit-maximizing and zero profit equilibrium is the same for $q \geq m - n - t$. In this case, the equilibrium apprenticeship fee is given by $c = m - n$, and profits are zero in this case, as no one apprentices. If $q < m - n - t$, then for the profit-maximizing equilibrium, the apprenticeship fee is given by $c = \lambda \frac{(m-n)+t}{1+\lambda-\frac{m-n}{m-n}}$, with $\hat{\rho} = \frac{\lambda(m-n)+t}{(1+\lambda)(m-n)-q}$, and $\Pi = \frac{(m-n-t-q)^2}{2[(1+\lambda)(m-n)-q]}$. For the zero-profit equilibrium on $q < m - n - t$, there are two cases. For $(1 - \lambda)(m - n) - 2t \leq q < m - n - t$, the apprenticeship fee is given by $c = \frac{2t-(1-\lambda)(m-n)+q}{1+\lambda-\frac{m-n}{m-n}}$, and $\hat{\rho} = \frac{2t+q-(1-\lambda)(m-n)}{(1+\lambda)(m-n)-q}$. For $q < (1 - \lambda)(m - n) - 2t$, the apprenticeship fee is given by $c = t + \frac{q}{2} - \frac{1}{2}(1 - \lambda)(m - n)$, and $\hat{\rho} = 0$. In any case, the main predictions remain as before when apprenticeship occurs.

Since $W_F = n$, those who have apprenticed within the current firm are more productive than apprentices from elsewhere, but need not be paid more than other apprentices, and therefore, the returns to apprenticeship continue to lie in self-employment.
8 Appendix C

In this appendix, we abstract from the problems of time-inconsistency in the firm’s wage offer. If we abstract from these problems, then it could be possible for the firm to commit, ahead of time, to a wage that is higher than the worker’s outside offer. More specifically, suppose that the firm is able to write a binding (and enforceable, in the Ghanaian context) contract at the time of the apprenticeship agreement that commits itself to a particular wage for the apprenticed worker should that worker remain in the firm upon completion of the apprenticeship. While the binding nature of such a contract is probably not realistic for the Ghanaian situation (see Fafchamps (2004) and the discussion in the paper), we can compare this equilibrium where full contracting is allowed to the more realistic case of Appendix A in terms of predictions, and implications for firm profits. These theoretical predictions can be examined on their own terms, and can also be empirically tested in the data, to see whether one model is preferred. In the case of this appendix, where the within firm wage contract is enforceable ex ante, the analysis follows exactly as in the main paper up to (4):

\[
\max_{c,W_F} \Pi = \int_{\hat{\rho}(c,W_F)}^1 (c - t + (1 - \rho)(1 - \lambda)(m - W_F))dF(\rho)
\]  

(30)

The firm will maximize profits subject to the three constraints:  

\( i) W_F \geq n \) (the firm can pay no less than the worker’s outside option),  
\( ii) \hat{\rho} \geq 0 \), and  
\( iii) \hat{\rho} \leq 1 \). Two further constraints will be added below, namely that  
\( iv) W_F \leq m \), and  
\( v) c \leq m - n \). This is to say that the firm cannot commit in the aforementioned contract to pay its apprenticed workers more than their productive contribution to output \( (W_F \leq m) \), and the firm cannot charge an apprenticeship fee that is larger than the increase in productivity resulting from
the apprenticeship \((c \leq m - n)\). Note also that in the case described in this appendix, the firm can jointly decide (and commit to) both the apprenticeship fee, \(c\), and the wage offered to its former apprentices within the firm, \(W_F\). Constraint \(ii\)) can be written more explicitly. Note that if we substitute \(W_O = n\) and \(W_n = n\) into equation (2), and simplify, we will get the result that:

\[
\hat{\rho} = \frac{c + (1 - \lambda)(n - W_F)}{m - \lambda n - (1 - \lambda)W_F}
\] (31)

Since the denominator in this is always strictly positive anywhere, as a result of constraint \(iv\)), we can see that constraint \(ii\)) can be rewritten as \(c + (1 - \lambda)(n - W_F) \geq 0\).

Furthermore, given the other constraints, constraint \(v\)) can be shown to be equivalent to constraint \(iii\)), and is therefore dropped, and so constraint \(iii\)) can be shown to be written as: \(m - n - c \geq 0\). Now, the Lagrangian for the firm’s profit-maximization problem is:

\[
L = \int_1^{\hat{\rho}(c,W_F)} (c - t + (1 - \rho)(1 - \lambda)(m - W_F))dF(\rho)
\]

\[
+ \alpha_1(W_F - n) + \alpha_2(n + \frac{c}{1 - \lambda} - W_F) + \alpha_3(m - n - c) + \alpha_4(m - W_F)
\]

where \(\hat{\rho}(c,W_F)\) is as given in equation (31). I will consider separately the cases of an interior solution, and the boundary solution to find the profit maximization point(s).

First, find the critical points of the profit function, by setting \(\frac{\partial L}{\partial c} = 0\) and \(\frac{\partial L}{\partial W_F} = 0\). First, for \(\frac{\partial L}{\partial c} = 0\), noting from equation (31) that \(\frac{\partial \hat{\rho}}{\partial c} = \frac{1}{m - \lambda n - (1 - \lambda)W_F}\):

\[
\int_{\frac{c + (1 - \lambda)(s - W_F)}{m - \lambda n - (1 - \lambda)W_F}}^{1} (1)d\rho
\]

\[
= \frac{1}{m - \lambda n - (1 - \lambda)W_F}[c - t + (1 - \frac{c + (1 - \lambda)(n - W_F)}{m - \lambda n - (1 - \lambda)W_F}(1 - \lambda)(m - W_F))] = 0
\]

53
Some algebraic work will show that this will give an explicit expression for $c$:

$$c = \frac{(1 - \lambda)tW_F - \lambda(m - n)^2 - tm + \lambda n}{(1 - \lambda)W_F - (1 + \lambda)m + 2\lambda n} \quad (34)$$

Now, consider the first-order condition with respect to $W_F$. First note that, from equation (31) and some algebraic manipulation, we can get that

$$\frac{\partial \hat{\rho}}{\partial W_F} = (1 - \lambda)(n - m + c) \frac{(m - \lambda n - (1 - \lambda)W_F)^2}{(m - \lambda n - (1 - \lambda)W_F)^2}.$$  

Therefore for $\frac{\partial \Pi}{\partial W_F} = 0$, we get:

$$- \int_{\frac{c + (1 - \lambda)(n - W_F)}{m - \lambda n - (1 - \lambda)W_F}}^{1} (1 - \lambda)(1 - \rho)d\rho$$

$$\begin{align*}
&= (1 - \lambda)(n - m + c) \frac{(m - \lambda n - (1 - \lambda)W_F)^2}{(m - \lambda n - (1 - \lambda)W_F)^2} \left[ (c - t) + (1 - \lambda)(m - W_F)(1 - \frac{c + (1 - \lambda)(n - W_F)}{m - \lambda n - (1 - \lambda)W_F}) \right] = 0
\end{align*}$$

This will simplify to an implicit expression for $W_F$:

$$(1 - \lambda)W_F(m - n - 2t) + (1 - \lambda)cW_F + (3\lambda n - (1 + 2\lambda)m)c + (2\lambda - 1)m^2 + (1 - 3\lambda)mn + \lambda n^2 + 2tm - 2\lambda tn = 0 \quad (35)$$

Therefore, the solution to equation (36) and equation (34) will yield the set of all critical points for the profit function. Substituting the expression for $c$ from equation (34) into equation (36) gives:

$$(1 - \lambda)W_F(m - n - 2t) + (1 - \lambda)\left(\frac{(1 - \lambda)tW_F - \lambda(m - n)^2 - tm + \lambda n}{(1 - \lambda)W_F - (1 + \lambda)m + 2\lambda n}\right)W_F \quad (37)$$

$$\begin{align*}
&+ (3\lambda n - (1 + 2\lambda)m)(\frac{(1 - \lambda)tW_F - \lambda(m - n)^2 - tm + \lambda n}{(1 - \lambda)W_F - (1 + \lambda)m + 2\lambda n}) + (2\lambda - 1)m^2 + (1 - 3\lambda)mn + \lambda n^2 + 2tm - 2\lambda tn = 0
\end{align*}$$

After considerable manipulation, this equation in $W_F$ simplifies to:

$$(1 - \lambda)^2(m - n - t)W_F^2 - 2(1 - \lambda)(m - n - t)(m - \lambda n)W_F + (m - n - t)(m - \lambda n)^2 = 0 \quad (38)$$
This quadratic equation has a unique solution (the discriminant is 0):

\[ W_F = \frac{m - \lambda n}{1 - \lambda} \] (39)

Substituting this expression for \( W_F \) back into equation (34) yields:

\[ c = m - n \] (40)

However, this is merely constraint \( iii \). Moreover, since \( m > n \), we can show from equation (39) that \( W_F > m \), which violates constraint \( iv \). Therefore, the unique critical point for the profit function lies outside of our constrained set. Therefore, in order to determine the profit-maximizing solution, we need only consider the boundary cases, which we will now do.

We can consider the cases where any combination of the following four constraints bind:  

- \( i \) \( W_F \geq n \),  
- \( ii \) \( c + (1 - \lambda)(n - W_F) \geq 0 \),  
- \( iii \) \( c \leq m - n \),  
- \( iv \) \( W_F \leq m \).  

Clearly, \( i \) and \( iv \) will not bind at the same time. Suppose that \( i \) binds (\( W_F = n \)). Then, \( \hat{\rho} = \frac{c}{m-n} \).

In this case, firm profits reduce to:

\[ \Pi = \frac{1}{2(m-n)}[-c^2(1 - \lambda) + 2c(\lambda(m - n) + t) - 2(m - n)t + (1 - \lambda)(m - n)^2] \]

The optimal choice of \( c \) in this case can be shown (using first and second-order conditions) to be:

\[ c = \frac{\lambda(m-n)+t}{1-\lambda} \]

So, this is exactly the case outlined in the main paper (and Appendix A), where the firm is forced to keep wage offers time-consistent. Clearly, none of the other constraints will bind then, when constraint \( i \) binds. In this case, firm profits are:

\[ \Pi_i = \frac{(m-n-t)^2}{2(m-n)(1+\lambda)} \]

Now, suppose that constraint \( iv \) binds, (\( W_F = m \)). In this case, \( \hat{\rho} = \frac{c-(1-\lambda)(m-n)}{\lambda(m-n)} \), and the firm profit function reduces to:  

\[ \Pi_{iv} = \frac{(c-t)(m-n-c)}{\lambda(m-n)} \]

Here, it can be shown,
through first and second-order conditions, that the profit-maximizing choice of
apprenticeship fee is $c = \frac{m-n+t}{2}$, and that in this case firm profits are: $\Pi_{iv} = \frac{(m-n-t)^2}{4\lambda(m-n)}$ (*).

Next, suppose that constraint $ii$) binds ($c = (1-\lambda)(W_F-n)$). In this case, $\hat{\rho} = 0$, and $\Pi = \frac{1}{2}(1-\lambda)(m+W_F-2n) - t$. Since $\frac{\partial \Pi}{\partial W_F} = \frac{1}{2}(1-\lambda) > 0$, to maximize profits when constraint $ii$) binds, the firm would increase $W_F$ until constraint $iv$) binds, so that $W_F = m$. However, substituting the two binding constraints, we get that firm profits are: $\Pi_{ii,iv} = (1-\lambda)(m-n)-t$. This is obviously not the same solution as the maximum profit when $iv$) binds alone. Therefore, we expect that firm profits here should be less than the optimal case where $iv$) binds above (*). That is we would like to show that $\Pi_{iv} > \Pi_{ii,iv}$.

This amounts to showing that $\frac{(m-n-t)^2}{4\lambda(m-n)} > (1-\lambda)(m-n) - t$. But this inequality is equivalent to $[(2\lambda-1)(m-n)-t]^2 > 0$, which is always true. Therefore $\Pi_{iv} > \Pi_{ii,iv}$.

Next, suppose that constraint $iii$) binds ($c = m-n$). First, consider the apprenticeship decision in this case. The expected returns to apprenticeship are:

$-(m-n) + \rho m + (1-\rho)[\lambda n + (1-\lambda)W_F]$, while the returns from not apprenticing are $n$.

Therefore, to induce apprenticeship we need that the returns from apprenticeship are at least as large as $n$. This can be shown to be equivalent to $W_F \geq \frac{m-\lambda n}{1-\lambda}$. But, again since $\frac{m-\lambda n}{1-\lambda} > m$, this lies outside our constraint set (and the equality version of this is just the critical point that we ruled out earlier). Therefore, constraint $iii$) does not bind at the optimal solution.

So, the cases that have not yet been ruled out are when constraint $i$) binds, and when constraint $iv$) binds. To show the profits on the region where constraint $i$) binds are never larger than the profits when constraint $iv$) binds, consider the following. For $\Pi_i > \Pi_{iv}$, $\frac{(m-n-t)^2}{2(m-n)(1+\lambda)} > \frac{(m-n-t)^2}{4\lambda(m-n)}$. But, this can be shown to be equivalent to $\lambda > 1$. 

56
However, by the definition of $\lambda$ as a proportion, $0 < \lambda < 1$. Therefore, $\Pi_i < \Pi_{iv}$, and so the profit maximizing solution occurs when constraint $iv$) binds, that is when, $W_F = m$, 
\[ c = \frac{m-n+t}{2}, \hat{\rho} = \frac{c-(1-\lambda)(m-n)}{\lambda(m-n)}, \text{ and } \Pi_{iv} = \frac{(m-n-t)^2}{4\lambda(m-n)}. \]
So, in this case, firms do pay their former apprentices their marginal products in terms of a wage. Firms still make more profits here than in the case of constraint $i$) binding (Appendix A), as they charge a very high apprenticeship fee, $c_{iv} = \frac{m-n+t}{2} > \frac{t+\lambda(m-n)}{1+\lambda} = c_{monop}$ in Appendix A. The difference in profits between this case and the case of Appendix A
\[ \left( \frac{(m-n-t)}{2\lambda(m-n)} - \frac{(m-n-t)^2}{2(1+\lambda)(m-n)} \right) > 0 \]
can be seen as the foregone profits that firms lose as a result of being unable to engage in binding contracts.

**Case of Free Entry:**

We can also consider the equilibrium that would result when there is free entry into the provision of apprenticeships. As in the main case in the text, here the result will be driven by a zero-profit condition. This case can be solved by setting equation (30) above equal to zero subject to the same constraints as in the previous section, namely: $i)$ $W_F \geq n$, $ii)$ $\hat{\rho} \geq 0$, $iii)$ $\hat{\rho} \leq 1$, $iv)$ $W_F \leq m$, and $v)$ $c \leq m - n$. It can be shown that in this context, neither constraint $iii)$ nor $v)$ will bind. Provided that $\hat{\rho} > 0$, equation (30) will give:
\[ [(c - t)\rho + (1 - \lambda)(m - W_F)\rho - \frac{1}{2}\rho^2] e^{(1-\lambda)(m-W_F)} = 0 \]
This will result in the following condition, after some simplification:
\[ (m - n - c)[2(c - t)(m - \lambda n - (1 - \lambda)W_F) + (1 - \lambda)(m - W_F)(m - n - c)] = 0 \]

The first case where profits will be zero is the case $c = m - n$. This case was outlined above for constraint $iii)$ binding, and the wage required to induce apprenticeship
lies outside our constraints in this case. The case that remains where profits will be zero occurs when the second factor is 0, that is when:

\[
2(c - t)(m - \lambda n - (1 - \lambda)W_F) + (1 - \lambda)(m - W_F)(m - n - c) = 0 \quad (42)
\]

This equation is a function of the two choice variables for the firm, \(c\), and \(W_F\). It defines a hyperbola in the \((c, W_F)\) space. Using extensive functional analysis\(^{42}\) and ruling out the inconsistent cases, the single segment of the hyperbola that defines the solution set within the constraints is a concave, upward-sloping curve between points \(A(\frac{2t-(1-\lambda)(m-n)}{1+\lambda}, n)\) and \(B(t, m)\). This is the solution provided that \(\hat{\rho} > 0\) always, which can be shown to be true if \((1 - \lambda)(m - n) \leq t\). Now, consider when constraint \(ii\) binds, that is when \(\hat{\rho} = 0\). It can be shown that this will occur for a non-trivial set provided that \((1 - \lambda)(m - n) > t\). In this case, the zero-profit solutions consist of the hyperbola defined by equation (42) between points \(A(\frac{2t-(1-\lambda)(m-n)}{1+\lambda}, n)\) and \(C(2t - (1 - \lambda)(m - n), \frac{2t}{1-\lambda} + 2n - m)\) combined with the contiguous line \(CB\), with \(B\) defined above.\(^{43}\) Therefore, in either case \((1 - \lambda)(m - n) \leq t\) or \((1 - \lambda)(m - n) > t\), the zero-profit solution will consist of an upward-sloping continuous curve joining points \(A\) and \(B\), defined above. Note that point \(A\) corresponds to the solution from the time-consistent equilibrium outlined in the main paper and Appendix A. Since \(n > 0\) and \(m > 0\), we know that this curve from \(A\) to \(B\) lies within the first two quadrants of the \((c, W_F)\) space (thankfully—otherwise workers would be paying firms to work). However, \(\frac{2t-(1-\lambda)(m-n)}{1+\lambda}\) can be positive or negative depending on parameters. That is, in some cases \(c < 0\), and the

\(^{42}\)The full details are available from the author upon request, but this work is too tedious to take up space here.

\(^{43}\)The equation of this line is given by setting \(\hat{\rho} = 0\) in equation (30).
firm will pay the apprentices in order to apprentice within the firm. The upper bound on $c$ is the cost of training, $t$, and $n \leq W_F \leq m$ on this segment. It should also be noted that 
\[
\frac{\partial c}{\partial W_F} > 0, \text{ that is if the wage } W_F \text{ is raised, then the apprenticeship fee must also be raised}
\]
to maintain zero profits.

In summary, the case of free entry allows for a continuum of equilibria, with wages ranging anywhere within their bounds, from $n$ to $m$. Therefore, the free-entry equilibria with full contracting commitments will all but surely be consistent with the data. Consider the more interesting case of profit-maximizing firms. Here, firms will pay their former apprentices their marginal product, $m$, in equilibrium. Therefore, this model predicts that workers who have apprenticed within the current firm should be both more productive than apprentices from elsewhere, and paid more than apprentices from elsewhere. The latter point differs from the main model of Appendix A, and therefore these different predictions can be compared in the data.
9 Appendix D - Production Function Estimation

For simplicity, we will rewrite the production function in equation (19) as follows:

\[ y_{ft} = \beta_{0}^{**} + (X_{V})_{ft}(\beta_{V}) + \beta_{k}k_{ft} + \omega_{ft} + \varepsilon_{ft} \]  \hspace{1cm} (43)

where \( \beta_{0}^{**} = \lambda_{0} + \beta_{0}^{*}\beta_{H} \), \( X_{V} \) is the row vector containing \( \{ \log L, \overline{S_{W}}, \overline{X_{W}}, \overline{X_{W}^{2}}, \overline{T_{W}}, \overline{App_{W}}, \overline{N_{A}}, \overline{S_{A}}, \overline{T_{A}} \} \), and \( \beta_{V} \) is the column vector containing \( \beta_{H}, \beta_{H}\beta_{1}, \beta_{H}\beta_{2}, \beta_{H}\beta_{3}, \beta_{H}\beta_{4}, \beta_{H}\beta_{5}, \beta_{H}\beta_{6}, \beta_{H}\beta_{7}, \beta_{H}\beta_{8} \). In basic least squares production function estimation, the residual in the estimation consists of both the firm productivity, \( \omega \), as well as the conventional noise component, \( \varepsilon \). The basic idea (although far more formally elaborated) of both Olley and Pakes (1996) and Levinsohn and Petrin (2003) is to try to control more carefully for firm productivity, as it could well be correlated with the inputs, particularly the labor inputs, \( X_{V} \), which would result in biased parameter estimates. Olley and Pakes develop a proxy measure for productivity based on a firm’s use of investment, and Levinsohn and Petrin develop a proxy based on the firm’s use of intermediate inputs. It is the Levinsohn and Petrin procedure followed here. Levinsohn and Petrin consider the following intermediate input demand function:

\[ m_{ft} = f_{t}(\omega_{ft}, k_{ft}) \]  \hspace{1cm} (44)

In order to write the function in this way, input prices are assumed to be common across firms within a given time period, and so the intermediate input function is indexed by \( t \). In order for intermediate inputs to be a valid proxy for productivity, the intermediate input use must be monotonically increasing in \( \omega \), as in the case of investment. Levinsohn and Petrin provide the conditions under which this is the case. Then, equation (44) can be
inverted to get:

$$\omega_{ft} = f_t^{-1}(m_{ft}, k_{ft})$$  \hspace{1cm} (45)

We can then substitute this expression into equation (43) to get:

$$y_{ft} = \beta^*_0 + (X_V)_{ft}(\beta_V) + \beta_k k_{ft} + f_t^{-1}(m_{ft}, k_{ft}) + \varepsilon_{ft}$$  \hspace{1cm} (46)

The $f_t^{-1}$ is allowed to be a nonparametric function of $m$ and $k$. As a result, initially, the $\beta_k$ coefficient cannot be separately identified from the nonparametric function. Therefore, define this combined function as $\phi_t(m_{ft}, k_{ft}) = \beta_k k_{ft} + f_t^{-1}(m_{ft}, k_{ft})$. Then, equation (46) can be rewritten as:

$$y_{ft} = \beta^*_0 + (X_V)_{ft}(\beta_V) + \phi_t(m_{ft}, k_{ft}) + \varepsilon_{ft}$$  \hspace{1cm} (47)

The $\phi$ function is nonparametric function, and so equation (47) is a partially linear model, which can be estimated using any semiparametric estimation technique. The method that I use is that of Robinson (1988), which is also described in Levinsohn and Petrin (2003). This estimation is labelled as the ‘first stage’ of the Levinsohn and Petrin estimation procedure, and gives a consistent estimate of the parameters $\beta_V$. The ‘second stage’ of the Levinsohn and Petrin procedure involves the identification of the capital coefficient, and the assumptions of this second stage are given in Levinsohn and Petrin (2003). While I do estimate the capital coefficient according to this procedure, the capital coefficient is of no interest for the current study, and so I do not describe the rest of the procedure here. None of the assumptions (limited as they are) of the second stage procedure are required for the identification of $\beta_V$, and so those assumptions are not described here either.
10 Appendix E - Net Benefits of Apprenticeship

This appendix measures the net present discounted value of apprenticeship, according to the model, and the coefficient estimates of Table 4. It should be noted that these estimated benefits are very rough calculations, as they are calculated at the averages of the variables, and using the point estimates. I do not bother to calculate confidence bounds on these estimates, as once the standard errors on each of the coefficient estimates are taken into account, the bounds would be very large indeed. That being said, these estimated net benefits can provide some "back-of-the-envelope" measures of the net benefits of apprenticeship, and at least ensure that they are positive for those who undertake apprenticeship.

The net benefits of apprenticeship can be calculated from the terms in equation (2). In order to calculate the net benefits in the population as a whole, we use the GLSS data, and the estimated coefficients from Table 4. Here, we do not know whether a former apprentice is working in the same firm as the firm in which an apprenticeship was completed or not. Therefore, a more general measure of the ex-ante net expected wages if a worker completes an apprenticeship is:

\[ NPV_{app} + \rho NPV_{A-S} + (1 - \rho) NPV_{A-EE} \] (48)

The first term, \( NPV_{app} \), is the net monetary returns during the apprenticeship, and will include the apprentice wages minus apprenticeship fees (this was labelled \(-c\) in the simple two period model). While in the model, the value \( c \) included both pecuniary and non-pecuniary transfers between the apprentice and the firm owner, here we will focus on the pecuniary transfers (i.e. including both cash and in-kind allowances, but abstracting
from how ‘hard’ the apprentice works in apprenticeship versus alternative employment, for example). The variable $\rho$ is the usual probability of obtaining capital to start self-employment. Here, $NPV_{A-SE}$ is the net present discounted value of the post-apprenticeship wage returns in self-employment, while $NPV_{A-EE}$ is the net present discounted value of the post-apprenticeship wage returns in employee wage work. On the other hand, the ex-ante net expected wages if a worker does not complete an apprenticeship are:

$$\rho NPV_{NA-SE} + (1 - \rho) NPV_{NA-EE}$$

(49)

where the $NPV_{NA-SE}$ is the net present discounted value of the self-employment wage for non-apprenticed workers, and the $NPV_{NA-EE}$ is the net present discounted value of the employee wage for non-apprenticed workers. Therefore, the ex-ante returns to apprenticeship are positive if $Net\ Benefit = (48) - (49) > 0$. We will now discuss how expressions (48) and (49) can be estimated using the data, and the results of the paper.

First, consider the net monetary returns during the apprenticeship period, $NPV_{app}$. Let the discount factor be $\beta$ and the apprenticeship wages in a single year be $W_{app}$. Also, we will assume for simplicity that the apprenticeship fees are paid at the beginning of the apprenticeship period (the data in the GLSS does not distinguish when the apprenticeship fees are paid). Then, the net monetary return during apprenticeship, using an apprenticeship period of 3 years, which is the average (technically 2.92) in the data, is given by:

$$NPV_{app} = -fee + W_{app} + \beta W_{app} + \beta^2 W_{app}$$

(50)

Now, consider the net present discounted value of the post-apprenticeship earnings
in self-employment, \( NPV_{A-SE} \). If a worker ceases to work in \( T \) periods from now, then \( NPV_{A-SE} \) will be calculated over the periods from \( t = 3 \) until \( t = T \) (as the apprentice is engaging in the apprenticeship in periods \( t = 0, 1, 2 \)). Therefore, we get:

\[
NPV_{A-SE} = \sum_{t=3}^{T} \beta^t W_{A-SE} = \frac{\beta^3(1 - \beta^{T-2}) W_{A-SE}}{1 - \beta}
\]

(51)

where \( W_{A-SE} \) is the single-year earnings for former apprentices in self-employment. Similarly, we get that \( NPV_{A-EE} = \frac{\beta^3(1 - \beta^{T-2}) W_{A-EE}}{1 - \beta} \), where \( W_{A-EE} \) is the single-year wage for former apprentices in employee work. On the other hand, workers that do not engage in an apprenticeship can begin earning their returns immediately, and therefore

\[
NPV_{NA-SE} = \sum_{t=0}^{T} \beta^t W_{NA-SE} = \frac{(1 - \beta^{T+1}) W_{NA-SE}}{1 - \beta}, \quad \text{and}
\]

\[
NPV_{NA-EE} = \sum_{t=0}^{T} \beta^t W_{NA-EE} = \frac{(1 - \beta^{T+1}) W_{NA-EE}}{1 - \beta}, \quad \text{where } W_{NA-SE} \text{ and } W_{NA-EE} \text{ are the single-year earnings to non-apprenticed workers in self-employment and employee work, respectively.}
\]

Since in this exercise we are interested in the ex-ante returns to apprenticeship, each of the earnings measures, \( W_{A-SE}, W_{A-EE}, W_{NA-SE}, \text{ and } W_{NA-EE} \), is calculated at the mean values of all individuals in the Ghanaian labor force, that is using the data in the first panel of Table 1b, and using the relevant regression coefficients of Table 4 (i.e. column (6) for the self-employment earnings, and column (9) for the employee earnings, including the correction terms). The apprenticeship wage, \( W_{app} \), is merely taken as the average earnings among all those who are currently engaged as apprentices in the GLSS survey. The apprenticeship fee, \( fee \), is taken as the average fee paid by those who are currently engaged in apprenticeship. The parameters upon which we have discretion are the probability of receiving capital, \( \rho \), the discount factor, \( \beta \), and the length of one’s remaining working life at
the date of the apprenticeship decision, $T$. In Ghana, in 1999, the life expectancy at birth for men was 57 years. Since the life expectancy upon reaching one’s early twenties would be higher than this, we will arbitrarily set the end of one’s working life at 60 years.\textsuperscript{44} Since the average age of apprentices is 23 years old, and the apprenticeship is 3 years in length, on average people are making the apprenticeship decision at age 21, with 39 years of working life remaining after the current year ($T = 39$). For simplicity, we will take $\rho = 0.5$. Therefore, initially, we will measure the net benefits of apprenticeship for those who have a 50 percent chance ex-ante of receiving the capital to start their own firm. Higher values of $\rho$ would naturally increase the net benefit, while lower values would decrease it. For the discount factor, we will take $\beta$ to be 0.9 or 0.95 (corresponding to a discount rate of 0.05 or 0.10). For an average labor force member, with $\rho = 0.5$, the present discounted value of the net benefit of apprenticeship is 85 000 cedis (for $\beta = 0.90$) and 270 000 cedis (for $\beta = 0.95$), that is approximately 2 or 7 months earnings for an average labor force member. Now, we can also measure the ex-ante net benefit of apprenticeship, using the characteristics of those who actually take up an apprenticeship. Since the fraction of current apprenticed workers who become self-employed is 0.658, we can take that value for $\rho$, along with the characteristics of the sub-set of the population that takes up apprenticeship to calculate the ex-ante net benefit for this group. For ex-post apprenticed workers, then, the ex-ante net benefit of apprenticeship is 110 000 cedis (for $\beta = 0.90$) and 350 000 cedis (for $\beta = 0.95$), or approximately 3 or 8 months earnings, respectively. Again, given the assumptions and use of averages, etc., these numbers should be taken as indicating order of magnitude rather than precise predictions, but they do indicate positive

\textsuperscript{44}Retirement is not an issue before this age in Ghana.
discounted net present value returns from apprenticeship, for those who apprentice.
References


Frazer, G., 2003. Linking Firms and Workers: Heterogeneous Labor and Returns to Education, mimeo, University of Toronto.


### Table 1
Summary Statistics - GLSS Survey

#### Table 1a - All Individuals over the Age of 15

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Schooling</td>
<td>5.63</td>
<td>5.27</td>
</tr>
<tr>
<td>Annual Earnings (1991 cedis)</td>
<td>487938</td>
<td>2818455</td>
</tr>
<tr>
<td>Age</td>
<td>35.7</td>
<td>16.96</td>
</tr>
<tr>
<td>Currently an apprentice</td>
<td>0.050</td>
<td>0.217</td>
</tr>
<tr>
<td>Completed an apprenticeship</td>
<td>0.117</td>
<td>0.321</td>
</tr>
<tr>
<td>Male</td>
<td>0.465</td>
<td>0.499</td>
</tr>
<tr>
<td>Hours worked per week</td>
<td>38.6</td>
<td>19.87</td>
</tr>
</tbody>
</table>

N=14991

#### Table 1b - All Self-employed and Wage Workers in the Labor Force*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Schooling</td>
<td>7.49</td>
<td>5.72</td>
</tr>
<tr>
<td>Annual Earnings (1991 cedis)</td>
<td>493347</td>
<td>2848986</td>
</tr>
<tr>
<td>Age</td>
<td>38.0</td>
<td>12.50</td>
</tr>
<tr>
<td>Completed an apprenticeship</td>
<td>0.203</td>
<td>0.402</td>
</tr>
<tr>
<td>Male</td>
<td>0.428</td>
<td>0.495</td>
</tr>
<tr>
<td>Hours worked per week</td>
<td>45.8</td>
<td>20.64</td>
</tr>
</tbody>
</table>

N=4151

#### Table 1c - All Self-employed and Wage Workers in the Manufacturing Sector*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Schooling</td>
<td>6.06</td>
<td>5.35</td>
</tr>
<tr>
<td>Annual Earnings (1991 cedis)</td>
<td>405960</td>
<td>645909</td>
</tr>
<tr>
<td>Age</td>
<td>38.4</td>
<td>12.58</td>
</tr>
<tr>
<td>Completed an apprenticeship</td>
<td>0.372</td>
<td>0.484</td>
</tr>
<tr>
<td>Male</td>
<td>0.522</td>
<td>0.500</td>
</tr>
<tr>
<td>Hours worked per week</td>
<td>46.0</td>
<td>20.09</td>
</tr>
</tbody>
</table>

N=559

---

**Note:**
- The labor force definition does not include farmers or apprentices, and includes those of age 15 or older.
- Apprenticeships include all non-tailoring apprenticeships. See discussion in the text.
- The manufacturing sector considered here excludes textiles and garments.
Table 2
Summary Statistics for Manufacturing Workers from the GMES and GLSS

<table>
<thead>
<tr>
<th></th>
<th>GMES Summary Statistics</th>
<th>GLSS Summary Statistics for Manufacturing Workers (not Self-Employed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>11.20</td>
<td>3.52</td>
</tr>
<tr>
<td>Annual Earnings (1991 cedis)</td>
<td>420761</td>
<td>378806</td>
</tr>
<tr>
<td>Age</td>
<td>36.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Male</td>
<td>0.853</td>
<td>0.354</td>
</tr>
<tr>
<td>Hours worked per week</td>
<td>43.3</td>
<td>7.5</td>
</tr>
<tr>
<td>Completed an apprenticeship</td>
<td>0.251</td>
<td>0.434</td>
</tr>
<tr>
<td>Within this firm</td>
<td>0.081</td>
<td>0.272</td>
</tr>
<tr>
<td>At another firm</td>
<td>0.171</td>
<td>0.376</td>
</tr>
<tr>
<td></td>
<td>Years of Schooling</td>
<td>8.95</td>
</tr>
<tr>
<td></td>
<td>Annual Earnings (1991 cedis)</td>
<td>406820</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>36.6</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td>Hours worked per week</td>
<td>49.5</td>
</tr>
<tr>
<td></td>
<td>Completed an apprenticeship</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>N=2122</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N=109</td>
<td></td>
</tr>
<tr>
<td>Activity</td>
<td>Fraction</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Continued working for the firm</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>Worked for another firm in the industry</td>
<td>0.152</td>
<td></td>
</tr>
<tr>
<td>Worked for a firm in another industry</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>Started their own business</td>
<td>0.285</td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>Don't know</td>
<td>0.178</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All Workers</td>
<td>Self-Employed</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>(1) OLS</td>
<td>(2) OLS</td>
</tr>
<tr>
<td>Education</td>
<td>0.075 (.005)</td>
<td>0.066 (.005)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.046 (.005)</td>
<td>0.044 (.005)</td>
</tr>
<tr>
<td>Experience²/100</td>
<td>-0.062 (.008)</td>
<td>-0.060 (.008)</td>
</tr>
<tr>
<td>Was An Apprentice</td>
<td>0.169 (.055)</td>
<td>0.069 (.054)</td>
</tr>
<tr>
<td>Male</td>
<td>0.266 (.044)</td>
<td>0.247 (.047)</td>
</tr>
<tr>
<td>l</td>
<td>-0.243 (.104)</td>
<td>0.111 (.048)</td>
</tr>
<tr>
<td>R²</td>
<td>0.137 0.148 0.141</td>
<td>0.083 0.074 0.077</td>
</tr>
<tr>
<td>N</td>
<td>4099 4099 3608</td>
<td>2791 2409 2409</td>
</tr>
<tr>
<td>Test for normality - F-stat</td>
<td>1.720</td>
<td>0.950</td>
</tr>
<tr>
<td>prob.</td>
<td>0.163</td>
<td>0.415</td>
</tr>
<tr>
<td>Overid-test-X²</td>
<td>14.184</td>
<td>13.182</td>
</tr>
<tr>
<td>prob.</td>
<td>0.077</td>
<td>0.106</td>
</tr>
<tr>
<td>F-stat on excluded variables</td>
<td>1.270</td>
<td>1.920</td>
</tr>
<tr>
<td>prob.</td>
<td>0.273</td>
<td>0.078</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.

The variables in the selection equation which are excluded from the wage equation are father's education and occupation.

Further details on the procedure are in the text.
Table 5
Returns to Apprenticeship - Within the Manufacturing Sector (GMES Data)
Dependent Variable: log(Wage)

<table>
<thead>
<tr>
<th></th>
<th>All Non-Apprentice Workers in GMES Data</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td></td>
<td>0.063</td>
<td>0.063</td>
<td>0.062</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.010)</td>
<td>(.009)</td>
<td>(.010)</td>
<td>(.009)</td>
</tr>
<tr>
<td>Experience</td>
<td></td>
<td>0.008</td>
<td>0.007</td>
<td>0.008</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.014)</td>
<td>(.014)</td>
<td>(.014)</td>
<td>(.014)</td>
</tr>
<tr>
<td>Experience^2/100</td>
<td></td>
<td>0.009</td>
<td>0.011</td>
<td>0.009</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.029)</td>
<td>(.028)</td>
<td>(.029)</td>
<td>(.028)</td>
</tr>
<tr>
<td>Did an Apprenticeship</td>
<td></td>
<td>-0.159</td>
<td>-0.148</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.074)</td>
<td>(.073)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apprenticed in this Firm</td>
<td></td>
<td>-0.177</td>
<td>-0.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.084)</td>
<td>(.085)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apprenticed Elsewhere</td>
<td></td>
<td>-0.151</td>
<td>-0.138</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.079)</td>
<td>(.079)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>-0.113</td>
<td></td>
<td>-0.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.136)</td>
<td></td>
<td>(.136)</td>
<td></td>
</tr>
<tr>
<td>R^2</td>
<td></td>
<td>0.117</td>
<td>0.120</td>
<td>0.117</td>
<td>0.120</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>2129</td>
<td>2129</td>
<td>2129</td>
<td>2129</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses.
### Table 6
Production Function - Ghanaian Manufacturing Sector
Dependent Variable - log(Value-Added)  \( N=457 \)

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) LP</th>
<th>(3) OLS</th>
<th>(4) LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(employment) ( (b_{1t}) )</td>
<td>0.835</td>
<td>0.272</td>
<td>0.830</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.066)</td>
<td>(0.074)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Share of apprentices ( (b_{1t} * b_6) )</td>
<td>-1.166</td>
<td>-0.178</td>
<td>-1.135</td>
<td>-0.102</td>
</tr>
<tr>
<td></td>
<td>(0.753)</td>
<td>(0.588)</td>
<td>(0.756)</td>
<td>(0.764)</td>
</tr>
<tr>
<td>Worker Education ( (b_{1t} * b_1) )</td>
<td>0.023</td>
<td>0.071</td>
<td>0.023</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.022)</td>
<td>(0.027)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Worker Experience ( (b_{1t} * b_2) )</td>
<td>-0.017</td>
<td>-0.016</td>
<td>-0.015</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.031)</td>
<td>(0.037)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Worker Experience^2/100 ( (b_{1t} * b_3) )</td>
<td>0.010</td>
<td>0.028</td>
<td>0.007</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.078)</td>
<td>(0.089)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>Worker Did an Apprenticeship ( (b_{1t} * b_5) )</td>
<td>-0.431</td>
<td>-0.325</td>
<td>-0.431</td>
<td>-0.325</td>
</tr>
<tr>
<td></td>
<td>(0.267)</td>
<td>(0.212)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Worker Apprenticed in this Firm: -0.290, -0.055
Worker Apprenticed Elsewhere: -0.537, -0.521
Apprentice Tenure \( (b_{1t} * b_8) \): -0.208, -0.098, -0.209, -0.100
Apprentice Education \( (b_{1t} * b_7) \): 0.070, 0.032, 0.069, 0.031
log(Capital): 0.157, 0.286, 0.159, 0.292

Note: Standard errors are in parentheses. The standard errors in the Levinsohn and Petrin (LP) procedure are bootstrapped \( b=100 \). Controls for industry sector were included in all of the estimation procedures.
### Appendix Table A1 - First-Stage Probit Estimates of Self-Employment Status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>-0.086</td>
<td>0.008</td>
</tr>
<tr>
<td>Experience</td>
<td>-0.011</td>
<td>0.006</td>
</tr>
<tr>
<td>Experience$^2/100$</td>
<td>0.010</td>
<td>0.009</td>
</tr>
<tr>
<td>Male</td>
<td>-1.105</td>
<td>0.084</td>
</tr>
<tr>
<td>Did an Apprenticeship</td>
<td>0.346</td>
<td>0.068</td>
</tr>
<tr>
<td>Father's Schooling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>-0.180</td>
<td>0.116</td>
</tr>
<tr>
<td>Middle</td>
<td>-0.047</td>
<td>0.084</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.195</td>
<td>0.134</td>
</tr>
<tr>
<td>Father's Occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farming</td>
<td>-0.115</td>
<td>0.086</td>
</tr>
<tr>
<td>Clerical/Professional</td>
<td>-0.179</td>
<td>0.096</td>
</tr>
<tr>
<td>Construction</td>
<td>-0.177</td>
<td>0.163</td>
</tr>
<tr>
<td>N</td>
<td>3610</td>
<td></td>
</tr>
<tr>
<td>F-stat on excluded variables</td>
<td>2.72</td>
<td></td>
</tr>
<tr>
<td>p-value on F-stat</td>
<td>0.014</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Standard Errors are in Parentheses