Measuring Firm-level Inefficiencies in the Ghanaian Manufacturing Sector*

Andrea Szabó

Economics Department, University of Houston

E-mail: aszabo2@uh.edu

First version: June 2010

August 17, 2016

Abstract

This paper measures firm level inefficiencies in input use among manufacturing firms in Ghana by explicitly estimating their production function. I find that the fraction of undercapitalized firms is 46%, but overall firms use 77% more capital and 40% less labor than would be optimal. Underutilization of labor is especially prevalent among firms with a unionized workforce. Firms with formal loans and firms with more human capital are closer to their efficient capital stock. The findings suggest large potential gains in value added from adjusting input use in the optimal direction.

*A previous version of this paper circulated under the title "Input misallocation in developing countries: Structural estimates from Ghana." I would like to thank Hunt Allcott, Patrick Bajari, Tom Holmes, Amil Petrin, Chris Timmins, Gergely Ujhelyi, two anonymous referees, as well as participants at NEUDC 2010 and the 2014 Workshop on Manufacturing and Economic Development at Harvard for comments and suggestions.

1 Introduction

Factor markets in developing countries exhibit well-documented inefficiencies and the use of inputs (such as labor and capital) is far from the neoclassical ideal. As a result, output is not produced in the most efficient way possible and resource misallocation is recognized to be a major obstacle to growth. Quantifying these inefficiencies and understanding which types of firms are most likely to operate inefficiently is a crucial first step in the evaluation of a wide range of development policies from microfinance programs to labor market or banking reforms. However, only a handful of existing micro-level studies are able to provide direct evidence on the magnitude of inefficiencies. This paper uses data from a panel of manufacturing firms in Ghana, 1991-2002, which contains information on detailed firm characteristics as well as capital, labor, output, firm-specific prices, and interest rates from a variety of financing sources. Using direct production function estimates, I measure firm level inefficiencies in input use, investigate which types of firms are most likely to underutilize specific inputs, and quantify the effects of inefficiency on aggregate value added.

The estimation uses the Wooldridge (2009) modification of the Levinsohn and Petrin (2003) procedure which addresses identification concerns related to earlier methods. This method is appropriate for typical developing country data sets, where investment is zero for many observations.[1] I also allow for various extensions to the basic estimation procedure, including firm exit, dynamic labor choice, and firm-specific input prices.

I compute two measures of inefficiency in input use based on the parameter estimates of the production function. The first measure is similar in spirit to Fernandes and Pakes (2008) and quantifies the underutilization of an input with the ratio of optimal to observed use of that input. Specifically, I compute the cost minimizing input levels for a firm producing the observed output and divide it by the input use observed in the data.[2] When this ratio

[1] In such cases, much information would be lost in dropping these observations, as would be required by the Olley and Pakes (1996) technique.

[2] In their counterfactual, Fernandes and Pakes (2008) compute the firm's optimal use of each input one at a time, holding other inputs fixed at their observed level. By contrast my counterfactual is the fully cost minimizing solution where all inputs are optimized simultaneously.

is above 1, it shows the extent to which a firm would need to increase its input usage in order to operate efficiently (minimize costs). The second measure of inefficiency is due to Petrin and Sivadasan (2013) and computes the difference between an input's marginal value product and its price at the firm level. This shows the gain in value added that could be achieved by moving a firm's use of a particular input by 1 unit in the optimal direction, holding everything else constant. Summing the absolute value of these gaps across firms yields a useful measure of lost value added in the economy. For example, in the case of capital, this sum shows the gain in aggregate value added that can be achieved by moving all firms' capital use by one unit in the optimal direction.

These two measures of inefficiency differ in terms the underlying counterfactual and thus might be relevant for the evaluation of different policies. For underutilization, the counterfactual thought experiment is cost minimization: firms attempt to produce the current output efficiently, adjusting all inputs as necessary. For the gap measure, the thought experiment is a much more limited adjustment of one unit of a specific input in the optimal direction, holding all other inputs fixed.[3] It should be noted that both of these counterfactuals are partial equilibrium in nature in that they hold constant all prices faced by firms in the economy at a predetermined level.

Overall, manufacturing firms in this sample are found to use 40% less labor and 77.4% more capital than would be efficient (minimize costs). These patterns echo recent findings from India (Fernandes and Pakes, 2008) and Chile (Petrin and Sivadasan, 2013) where the underutilization of labor is found to be especially large. Based on the estimated gap measures, adjusting the labor force by 1 worker in the optimal direction at every manufacturing firm in Ghana would yield a 0.16% increase in GDP. Adjusting capital usage in the optimal direction by the value of 1 worker's annual wage (about 0.3 million Cedis) at every firm holding everything else constant would increase GDP by 0.53%.

[3]For a given input, the optimal adjustment might be in different directions under these two counterfactuals. For example, a firm using too little capital for given labor may need to reduce its capital stock further to minimize costs if labor can be adjusted as well.

Looking at the firm-level correlates of inefficiency, three main findings emerge. First, firms without formal loans tend to underutilize capital while firms with formal loans are more likely to overutilize it. At the same time, firms with formal loans are closer to their efficient capital stock than firms with no formal loans. The gap measures therefore imply that remedying inefficiencies among firms without formal loans would result in larger gains in the economy. Second, firms with more human capital (measured with the level of education and experience of their workforce) are closer to using capital efficiently and have lower underutilization ratios for this input. This may indicate that these firms substitute skilled labor for the missing capital. Third, firms with a unionized workforce are further from the efficient use of the labor input, which may reflect constraints to the free adjustment of labor for these firms.

A large micro literature studies factor misallocation in developing countries using a variety of indirect methods to infer marginal value products.[4] For example, Banerjee and Munshi (2004) measure the propensity to invest in the garment industry in two communities in Southern India, and argue that the difference likely reflects different access to capital. Banerjee and Duflo (2014) study the natural experiment of an Indian banking reform to measure the returns to firms newly eligible for loans. They find that the gap between the marginal value product of capital and the market interest rate is at least 70 percentage points. De Mel, McKenzie and Woodruff (2008) and McKenzie and Woodruff (2008) conduct randomized field experiments in, respectively, Sri Lanka and Mexico, giving firms money and in-kind grants of about 100 USD and estimate the real return on this capital shock. From this experiment, they estimate an average real return on capital of 55-63 percent per year, which is substantially higher than the market interest rate.

My paper complements this literature by using a direct approach based on production function estimates. The structural estimates of the marginal value of capital in the current paper are quantitatively similar to these earlier estimates, which offers a validation of some

[4] Banerjee and Duflo (2005) provides a good summary of this literature.

previously used methods. The paper also contributes by comprehensively documenting the heterogeneity of inefficiency across firms using the rich set of covariates in the Ghanaian manufacturing dataset. Furthermore, the direct approach allows me to go beyond previous analyses by considering counterfactual scenarios.

Soderbom and Teal (2004) also study inefficiency and misallocation in Ghanaian manufacturing by estimating a production function. Unlike the approach used here, they use an estimation method that treats firm productivity as constant over time. Because I include later rounds of the survey in the analysis, I work with a substantially longer panel data, making the assumption of constant productivity unappealing in this case. Frazer (2005) also uses the shorter panel and the control function approach but focuses on firm exit and not the question of misallocation studied here.

Related studies in the macro literature include Basu and Fernald (2002), Caselli and Feyrer (2007), Restuccia and Rogerson (2008), Hsieh and Klenow (2009) and Vollrath (2014). The most important difference relative to this literature concerns the definition of "inefficiency" and, as a consequence, the type of counterfactual scenarios being considered. In particular, following Fernandes and Pakes (2008) and Petrin and Sivadasan (2013) the present paper considers the impact of relatively small adjustments in the optimal direction, rather than counterfactuals that involve, e.g., large-scale changes in the technology used by firms in the country. Some of these methodological differences are discussed further in Section 3 below.

The remainder of the paper is organized as follows. Section 2 describes the data used in the empirical analysis and Section 3 discusses the measurement of inefficiency. Section 4 describes the estimation of the production function. Section 5 and 6 present the results on inefficiency, and Section 7 discusses various robustness checks. Finally, Section 8 concludes.

2 Data

2.1 Ghanaian dataset

The main data source for this study is the Ghanaian Manufacturing Survey, 1991-2002, conducted by the World Bank, the Centre for the Study of African Economies at Oxford University, the Ghana Statistical Office, and the University of Ghana.[5] This data is particularly suited to analyze the question of inefficiency and misallocation for a number of reasons. First, of all African datasets used in previous research, this is the longest panel, containing 12 years of data collected in seven rounds. Second, the data is extremely detailed, containing an extensive list of questions about general firm characteristics and the labor market and financial market activities of the firms. Information collected includes formal and informal financing as well as measures of workers' human capital. Third, given the nature of the Ghanaian economy, firms in the data tend to operate a single plant and produce one major product. The survey contains explicit questions about the number of plants and products of each firm. In the data, 86% of firms have a single plant. Most of the multi-plant firms are concentrated in two industries (Bakery/Food and Metal). 38% of firms have a single product, and among multi-product firms, the average revenue share of their dominant product is 53.2%. This makes it unlikely that measures of inefficiency will be biased by the unobserved allocation of inputs within a firm across plants or products (De Loecker et al., 2016). Finally, the data allows circumventing a problem often encountered in the production function literature when only revenue data is available. Obtaining consistent estimates of the production function coefficients from data on revenues is only possible under certain assumptions on the correlation of firms' technology, input use, and prices. To alleviate this issue, studies typically deflate firm revenues using price indices corresponding to groups of firms (e.g., an industry). The survey used here contains *firm-specific* price indices (these are

[5] The newest rounds of the data were published only recently. Teal (2011) describes the construction of the dataset. The questionnaire and the data is available from http://www.csae.ox.ac.uk/datasets/Ghana-rped/Ghmain.html. Studies using earlier rounds of the dataset include Jones (2001), Söderbom and Teal (2004), Schündeln (2005), and Frazer (2005, 2006).

computed by the survey team using information collected on quantities and prices of each product produced by a firm). Using these as revenue deflators allows the consistent estimation of production coefficients under weaker assumptions than is typical in the literature (see Section 4.2).

In the first round of the survey, a sample of 200 firms was selected, designed to be representative based on size and industry structure according to the 1987 National Industrial Census.[6] Approximately half of these firms appear in all subsequent waves of the survey. In each wave, exiting firms were replaced by similar firms to keep the sample representative and the number of firms constant across waves. Over 12 years, a total of 312 firms were interviewed. The estimation below will account for firm exit. When I discuss the implications for the entire Ghanaian manufacturing sector (Section 6.4), I use the firm-size distribution of all manufacturing firms from the 2003 Census (the first census after 1987) to construct the appropriate weights.

The data used in my analysis is restricted by the availability of the information necessary to estimate the production function. The final sample consists of 1602 firm-year observations. Table 12 in the Appendix presents the sectoral distribution of the sample.

2.2 Output, capital, input and labor data

Real output is obtained as firm revenue deflated with firm-specific price deflators provided by the survey team (these price deflators were constructed using information collected on each firm's products and their prices). Capital is measured as the replacement value of the stock of plants and equipment.[7] To measure intermediate inputs, I use the total cost of raw material inputs per year. Real values are constructed using firm-specific material price indices provided in the survey (see Teal (2011) section 4.3). Employment at the firm includes

[6] The National Industrial Census was collected only three times by the Ghana Statistical Service in 1962, 1987 and 2003. It contains basic information, such as firm location and the number of employees. It does not contain the information necessary to estimate production functions.

[7] The capital variable is calculated as described in Teal (2011) but assuming a 6 percent depreciation rate. Results using alternative depreciation rates are discussed in section 7.

all salaried employees. The summary statistics for these variables are in Table 1. The values of all monetary variables in the paper are deflated to 1991 Ghanaian Cedis.

Table 1: Summary statistics for gross output production function estimation

	Mean	Median	Std. dev	10%	90%	N
Output	486.75	24.54	4265.88	2.23	607.8	1602
Capital	329.92	6.40	1370.11	0.19	560.79	1602
Material inputs	203.3	11.18	1601.92	0.96	243.87	1602
Employment	71.48	21	154.66	4	176.1	1602

Notes: All monetary values are in Million 1991 Ghanaian Cedis.

2.3 Interest rate and wage data

In developing countries such as Ghana, borrowing can come from many sources, including informal sources such as family and friends, or from overdraft facilities. Different lending sources operate with a wide variety of interest rates. An advantage of the survey used here is that it provides information on the various financing sources used by the firm. In the data, I can observe the loan amount with the interest rate provided by a formal financial institution in a given year. I also have data on the loan amount from various informal sources and the expected repayment (either in 1991 Cedis, or in-kind where the monetary value is given in the survey). I calculate the interest rate for loans coming from family using the loan amount and the expected repayment. Summary statistics of the interest rate data are in Table 2.[8] Table 13 in the Appendix presents the average of the highest observed interest rates in a given year, as well as the risk-free interest rate on deposits from the World Bank for comparison. As expected, interest rates on formal loans generally follow the trend observed in the deposit interest rate. The wedge between these two measures is 9 percentage points on average. Yearly averages computed including the informal interest rates are lower, reflecting the fact that interest rates on loans from family and friends have a median of zero.[9]

[8] Variable codes for the interest rate data are listed in Table 14 in the Appendix.

[9] It is not surprising to observe such low interest rates on loans coming from friends and family. In fact, when asked why they chose to borrow from informal sources, 29% of respondents in the survey cited the low

Firms report their yearly wage bill, which I divide by the number of employees to get the price of labor. I have non-zero wage data for 1423 firms. In some cases, workers receive (in-kind or cash) allowances or bonuses in addition to wages. As a robustness check, I compute some of the results below with the available earnings data which includes these allowances. Note that there is very little difference in the averages of these two variables. The summary statistics are in Table 2.

Table 2: Summary statistics, loans and wages

	Mean	Median	Std. dev	10%	90%	N
Formal loan amount	696.49	61.66	2726.27	1.47	948.10	422
Formal interest rate	34.65	36.00	10.92	20.00	45.00	422
Informal loan amount	16.28	0.21	143.56	0.01	13.06	186
Informal interest rate	5.56	0.00	18.85	0.00	23.46	186
Portfolio interest rate	26.55	32.98	18.12	0.00	45.00	583
Highest observed formal interest rate	35.83	37.50	10.51	22.00	45.00	422
Highest observed interest rate	27.83	35.00	18.30	0.00	45.00	583
Wage	0.27	0.16	0.72	0.03	0.57	1423
Earnings	0.30	0.26	0.23	0.06	0.58	1495

Notes: Monetary amounts are in Million 1991 Ghanaian Cedis. Interest rates are in percentage. Portfolio interest rate is the firm's average interest rate weighted by the loan amounts. Highest formal interest rate is the highest interest rate on loans from banks and overdraft facilities. Highest interest rate also includes interest rates from informal sources (friends and family). Wage is the firm's total reported wage bill divided by the number of employees. Earnings include the wage and other (in kind or cash) allowances, such as food or housing allowance and bonuses. Source: Ghanaian Manufacturing Survey, 1991-2002, and author's calcualations.

2.4 Other firm characteristics

To comprehensively describe the heterogeneity in inefficiency across firms in the sample, I exploit the rich set of covariates in the dataset. In particular, in addition to basic firm characteristics such as size, industry and location, the survey provides information on the characteristics of each firm's workforce (average years of education, worker age, unionization, and the share of management workers in the firm's workforce), the firm's ownership structure, and several measures related to the firm's export and import activities. The list of variables

interest rates (49% cited easier formalities, and 11% that no collateral was required).

used below is in Table 3, along with summary statistics.

Table 3: Summary statistics of firm characteristics

	Mean	Std. dev	10%	90%	N
Workforce					
Management workers as a share of all workers	2.71	5.23	0.00	8.33	1597
Firm average years of education	9.83	2.62	7.33	12.62	1557
Firm average of worker age	31.84	8.66	21.00	43.67	1557
Some or all employees unionized	0.35	0.48	-	-	1587
Loans					
Dummy for firms with formal loans	0.26	0.44	-	-	1602
Dummy for firms with informal loans	0.12	0.32	-	-	1602
Ownership					
Private Ghanaian	0.77	0.42	-	-	1602
Stata ownded	0.05	0.22	-	-	1602
Foreign ownership	0.18	0.38	-	-	1602
Trade					
Percentage of output exported within Africa	2.02	9.64	0.00	0.00	1434
Percentage of output exported outside Africa	7.78	23.95	0.00	20.00	1434
Percentage of raw materials imported	24.43	36.16	0.00	95.00	1514
Location					
Accra	0.53	0.50	-	-	1602
Cape Coast	0.03	0.18	-	-	1602
Kumasi	0.36	0.48	-	-	1602
Takoradi	0.08	0.27	-	-	1602

Notes: State owned firms are firms with some fraction of state ownership. Foreign firms are private firms with some fraction of foreign ownership. Source: Ghanaian Manufacturing Survey, 1991-2002.

3 Measuring inefficiency

Inefficiency of input use is a relative measure: it depends on how we define the counterfactual, unconstrained world. Several recent papers define inefficiency in somewhat different ways.

In approaches such as Hsieh and Klenow (2009), one source of inefficiency is technology (in the production function context, the different elasticities of inputs). They assume that observed US industry level elasticities approximate the technology of an unconstrained economy. Consequently, they compute the counterfactual "optimal" input allocation which an economy can achieve by first using the corresponding US technology, and then using the optimal input combinations (equating the price and the marginal value product of each input). One advantage of this approach is that it does not require estimating a production function for the country under study - instead, one can simply use the corresponding factor shares from the US.

In the "micro" approaches of Fernandes and Pakes (2008) and Petrin and Sivadasan (2013), inefficiency is measured in a different way.[10] In both cases, the goal is to determine the potential gain in value added that can be achieved by changing firms' inputs under their current technology. In Fernandes and Pakes, the counterfactual allocation is the cost minimizing solution. Specifically, firms are assumed to adjust one of their inputs optimally, holding output and all other inputs fixed. In Petrin and Sivadasan (2013), the counterfactual is where firms adjust one of their inputs in the profit-maximizing direction by one unit. In both cases technology and factor prices are obtained from the data. In the first case, the counterfactual assumes that all firms can obtain inputs at the same prices (taken to be the average of those observed in the data). In the second case, the adjustments take firm-level input price differences as given. One advantage of these micro approaches is that they lend themselves naturally to the study of the correlates of inefficiency at the firm level.

In this paper, I follow the micro approach, estimating the parameters of the production technology and the corresponding firm-level productivity terms. I use these production function estimates to compute measures of inefficiency and use the wide variety of firm characteristics available in the Ghanaian dataset to study the correlates of firm-level inefficiency. I extend the previous approaches by considering the fully cost minimizing solution as one of the counterfactual benchmarks.

It should be noted that, like all the above approaches, I do not analyze a fully dynamic general equilibrium model with endogenous prices and adjustment costs. In line with these previous studies, my measures of inefficiency are based on limited adjustments in firms' behavior, either holding output constant or considering one-unit adjustments in the optimal

[10] See Shenoy (2014) for a related approach.

direction. Incorporating the dynamic general equilibrium effects of these adjustments in measures of inefficiency is left for future research. See Matsuyama (2007) for a theoretical discussion of some of these issues.

4 Production function estimation

After describing the basic framework and the estimation procedure, I describe the three extensions which I use to account for the specific developing country environment studied here.

4.1 Basic framework

To estimate the marginal value of inputs and the optimal input allocations I estimate the firms' production function. I assume that in a given industry, firm i faces a Cobb-Douglas production function given by

$$q_{it} = \beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it} + \beta_a a_{it} + \varepsilon_{it}, \tag{1}$$

where q_{it} is output in period t, l_{it} is the number of employees, k_{it} is the real capital stock, m_{it} is the quantity of intermediate inputs (materials), and a_{it} is the age of the firm, all in logs. Including age is useful as a proxy for learned productivity, and it will also enter the proxy function for productivity (see equation (3) below). The term ε_{it} is a productivity shock that satisfies

$$\varepsilon_{it} = \omega_{it} + \eta_{it}. \tag{2}$$

Here ω_{it} is the "transmitted component," which is known by the firm but not by the econometrician. The firm sees ω_{it} before choosing its input combination, i.e., ω_{it} is a state variable for the firm. The term η_{it} is an unpredictable (both to the firm and to the econometrician) productivity shock assumed to be uncorrelated with input choices and distributed independently across firms. To estimate (1), ω_{it} is assumed to follow an exogenous first order Markov process so that

$$p(\omega_{it+1}|\{\omega_{i\tau}\}_{\tau=0}^{t}, I_{it}) = p(\omega_{it+1}|\omega_{it}),$$

where I_{it} is the firm's entire information set at time t. This assumption includes firm-level fixed effects as a special case when ω_{it} is fixed over time (i.e., $\omega_{it} = \omega_i$). Like in Olley and Pakes (1996) or in Levinsohn and Petrin (2003), I assume that $p(\omega_{it+1}|\omega_{it})$ is stochastically increasing in ω_{it}, that is, if a firm has a higher value of ω_{it} today, then it has a better distribution of ω_{it+1} tomorrow.

To proxy the transmitted component of the productivity shock I invert the firm's material demand and use

$$\omega_{it} = g(k_{it}, m_{it}, a_{it}), \tag{3}$$

where ω_{it} depends on the firm's state variable k_{it} and the proxy variable m_{it}, the intermediate inputs. The choice of intermediate inputs as the proxy variable follows Levinsohn and Petrin (2003). This is particularly important since every year between 46 and 80 percent of the firms do not report investments above the startup capital. Therefore much information would be lost in dropping these cases, as would be required by Olley and Pakes (1996), who use investment as a proxy.[11] Equation (3) relies on the assumption that the firm's demand for materials is monotonic in the productivity term ω (in which case the demand function can be inverted to obtain (3)). There are reasons to believe that this monotonicity may not hold in a developing country context unless one conditions on other variables as well (in particular, the price of materials), and I deal with this issue in Section 4.5 below.

Firms are assumed to solve a standard dynamic programming problem with the state variables k, a, and ω, choosing their level of investment. Investment I_{it} determines the

[11] The zero firm level investment does not mean that we do not have information on the firm's investment activity. A zero reported investment means that the firm did not invest in a particular year. This investment data can be used in the construction of the capital stock, but not as a proxy for transmitted productivity.

evolution of the capital stock according to

$$k_{it} = (1 - \delta)k_{it-1} + I_{it},$$

where δ is the depreciation rate. In the standard formulation, labor l is taken to be a non-dynamic input chosen freely in every period. Alternatively, the firm's labor choice can also be modeled as dynamic (see Fernandes and Pakes (2008) and Petrin and Sivadasan (2013)). As I explain in section 4.4, I allow for both of these possibilities in the estimation. Finally, I also allow for the possibility that firms may decide to exit after observing their realized productivity (see section 4.3).

4.2 Estimation procedure

Traditionally, equation (1) is estimated in two steps. The first stage involves estimating the inverse intermediate input demand function as well as the coefficient on labor. The second stage identifies the capital and age coefficients. The method proposed by Wooldridge (2009) combines these two stages into a single set of moment conditions and estimates the parameters in one step using GMM. This takes into account the simultaneity problem as described in Levinsohn and Petrin (2003) and Olley and Pakes (1996), and deals with the identification problem described by Ackerberg, Caves and Frazer (2015). The method also yields more efficient parameter estimates than the two-step procedures.

Another advantage of the Wooldridge (2009) method that is particularly important in the present context is that it allows separating the predictable (transmitted) component ω of the error term from the shock η in equation (2). This is important because studying firm-level inefficiency requires estimating the marginal product of inputs. Arguably this should depend only on the transmitted component of the productivity shock, which is a state variable observed by the firm, and not on the unpredictable shock η. Similarly, if the output data contains measurement error, this will be captured by η. This should not be included in the marginal product, and therefore an estimation method that can separate the two error terms is needed.

Following Wooldridge (2009), the production function parameters are estimated from the system[12]

$$q_{it} = \beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it} + \beta_a a_{it} + g(k_{it}, m_{it}, a_{it}) + \eta_{it} \text{ for } t = 1, ..., T \tag{4}$$

$$q_{it} = \beta_L l_{it} + \beta_K k_{it} + \beta_M m_{it} + \beta_a a_{it} + f[g(k_{it-1}, m_{it-1}, a_{it-1})] + u_{it} \text{ for } t = 2, ..., T \tag{5}$$

where $u_{it} = \omega_{it} - E(\omega_{it}|\omega_{it-1}) + \eta_{it}$. In my implementation, f is a second degree polynomial and g is a general third degree polynomial. The GMM estimation and the choice of instruments follows Wooldridge (2009). After parametrization of g and f, the residual function is defined for each $t > 1$ and can be written as:

$$\begin{pmatrix} r_{it1}(\theta) \\ r_{it2}(\theta) \end{pmatrix} = \begin{pmatrix} q_{it} - \alpha_0 - \beta_L l_{it} - \beta_K k_{it} - \beta_M m_{it} - \beta_a a_{it} - \mathbf{c}_{it}\boldsymbol{\lambda} \\ q_{it} - \varphi_0 - \beta_L l_{it} - \beta_K k_{it} - \beta_M m_{it} - \beta_a a_{it} - \rho_1 \mathbf{c}_{it-1}\boldsymbol{\lambda} - \rho_2(\mathbf{c}_{it-1}\boldsymbol{\lambda})^2 \end{pmatrix}$$

where $\mathbf{c}_{it}$ is a vector of the terms in the polynomial function g, and all greek letters denote parameters. This yields the following moment conditions for identification:

$$E[\mathbf{Z}'_{it}\mathbf{r}_{it}(\theta)] = 0 \text{ for } t = 2, ..., T,$$

where $\mathbf{Z}_{it}$ is a matrix of instruments given by

$$\mathbf{Z}_{it} = \begin{pmatrix} (1, l_{it}, \mathbf{c}_{it}, k_{it-1}, l_{it-1}, a_{it-1,}\mathbf{c}_{it-1}, \mathbf{h}_{it-1}) & 0 \\ 0 & (1, k_{it-1}, a_{it-1,} l_{it-1}, \mathbf{c}_{it-1}, \mathbf{h}_{it-1} \end{pmatrix}$$

where $\mathbf{h}_{it-1}$ contains the terms of a second degree polynomial of $\mathbf{c}_{it-1}$.[13]

[12] These equations correspond to equations (2.10) and (2.11) in Wooldridge (2009).

[13] For example, if $g(x_1, x_2) = \lambda_1 x_1 + \lambda_2 x_1 x_2$ and $f(x_1, x_2) = \phi_1 g(x_1, x_2) + \phi_2 g(x_1, x_2)^2$, then $\mathbf{c} = [x_1, x_1 x_2]$ and $\mathbf{h} = [x_1, x_1 x_2, (x_1)^2, (x_1 x_2)^2, (x_1)^2 x_2, x_1(x_2)^2]$.

I estimate separate production functions for the following four industries (the lowest level of aggregation allowed by the data): Furniture/Wood, Textile/Garment, Metal/Machinery, and Bakery/Food/Alcohol. For each of these, I also include dummy variables for ownership status (foreign, state, private Ghanaian) to account for the potentially different technologies available to firms in these groups within an industry.[14]

A common issue in the production function literature is that, due to the available data, the production function in (1) has to be estimated using data on revenues rather than the physical quantity of output (Olley and Pakes (1996) refer to this as a "sales generating function"). This has the potential to result in inconsistent coefficient estimates if firm-specific output prices are correlated with technology or input use (e.g., due to demand shifts or markups under imperfect competition). This can be alleviated if industry-specific price indices are available to deflate the revenue data (Petrin and Sivadasan (2013) refer to this as a "gross output production function"). In this case, the estimates are valid as long as the deviation of firms' prices from the industry average is uncorrelated with technology or input use. As described in Section 2, the data used here allows the identification of production function parameters under weaker assumptions because it contains *firm-specific* price indices. Using these to deflate firm revenue yields consistent production function coefficients as long as technology and input use is uncorrelated with changes in a firm's output price within an industry. Following Petrin and Sivadasan (2013), I will refer to the estimates below as the parameters of a gross output production function.

4.3 Accounting for selection

Firm exit may create a selection bias if firms exit based on unobserved productivity. Since the present dataset contains exit, it is important to correct for this. To do this, I follow

[14]There are potentially other variables that could affect a firm's technology choice (e.g., unionization) and with a larger dataset one could estimate separate production functions on finer cuts of the data. I have experimented with disaggregating the technology further but this resulted in many groups with too few firms to indentify the production function parameters.

Olley and Pakes (1996) who specify the following exit rule for firms:

$$\chi_{it} = \begin{cases} 1 \text{ (continue) if } \omega_{it} \geq \overline{\omega}_t(k_{it}, a_{it}, m_{it}) \\ 0 \text{ (exit) if } \omega_{it} < \overline{\omega}_t(k_{it}, a_{it}, m_{it}) \end{cases}$$

The endogeneity problem arises because we cannot observe the productivity cutoff $\overline{\omega}_t(k_{it}, a_{it}, m_{it})$. We can control for $\overline{\omega}_t(k_{it}, a_{it})$ using data on observed exit conditional on the information available at $t-1$:[15]

$$P_{it} \equiv \Pr(\chi_{it} = 1 | k_{it-1}, m_{it-1}, a_{it-1}) = \Pr(\omega_{it} \geq \overline{\omega}_t(k_{it}, a_{it}, m_{it}) | k_{it-1}, m_{it-1}, a_{it-1}) \quad (6)$$

I estimate equation (6) using a flexible functional form, modelling the probability of surviving in t as a function of $k_{it-1}, m_{it-1}, a_{it-1}$ using a probit model with a 4th order polynomial.[16] As long as the density of ω_t conditional on ω_{t-1} is positive around $\overline{\omega}_t(k_{it}, a_{it}, m_{it})$, equation (6) can be inverted to obtain $\overline{\omega}_t$ as a function of ω_{it-1} and $\widehat{P}_{it}$, $\overline{\omega}_t(k_{it}, a_{it}, m_{it}) = h(\omega_{it-1}, \widehat{P}_{it})$.

Now the evolution of ω_{it} depends both on ω_{it-1} and $\overline{\omega}_t = h(\omega_{it-1}, \widehat{P}_{it})$, so that $f(\omega_{it-1})$ becomes $f(\omega_{it-1}, h(\omega_{it-1}, \widehat{P}_{it}))$ or, using (3),

$$f[g(k_{it-1}, m_{it-1}, a_{it-1}), \widehat{P}_{it}].$$

In the estimation, I change the moment conditions to include $\widehat{P}_{it}$ in the function f.[17]

[15] As discussed by Olley and Pakes (1996, p1276), this is similar to the techniques used in single-index models like Ichimura (1993) and Ahn and Powell (1993).

[16] I experimented with different polynomials and found that, consistent with the findings reported in Olley and Pakes (1996), the exact choice of which terms to include matters little for the results. The results below include the full set of interactions up to order 3 of the variables (k, a, m), as well as the 4th order terms a^2m^2, a^2k^2 and m^2k^2. When accounting for labor market frictions and different input prices (see below), further terms are added as allowed by the size of the dataset.

[17] The fact that $\hat{P}$ is itself estimated creates a further source of error in the production function estimation. The literature offers little guidance of how to control for this, especially in the context of the 1-step Wooldridge procedure used here. One option is to estimate (6) as part of the same GMM system. In my dataset because I estimate production functions separately by industry there are not enough observations to estimate corresponding industry-specific exit rules, but investigating this issue further would be a useful area for future research. I thank an anonymous referee for this point.

4.4 Accounting for labor market frictions

In the standard formulation, labor l is taken to be a non-dynamic input chosen freely in every period. This assumption may not hold if hiring and firing is associated with high fixed costs. In this case, labor becomes a dynamic variable, which is chosen by the firm conditional on expected productivity next period.

Hiring and firing costs can arise in the presence of powerful labor unions or other government regulations. Petrin and Sivadasan (2013) discuss this in the Chilean context where the government adopted extensive regulations to promote job security. Similarly, Fernandes and Pakes (2008) allow for dynamic labor choices in the case of India, where firms had to obtain permission before firing employees. Fernandes and Pakes (2008) show that allowing labor to become dynamic significantly alters their estimates of the production function coefficients.

In the 1990's, labor unions in Ghana were generally fragmented, with different organizations representing workers in different sectors of the economy. This period was characterized by privatization and a reduction in jobs in state-owned enterprises that historically had higher levels of unionization. These trends were accompanied by a decline in economy-wide union membership. The union currently representing manufacturing workers, the Ghana Federation of Labour was formed in 1997 and had a growing membership in subsequent years. In my dataset on average about 35% of employees are unionized over the period 1991-2002, indicating that there may be some constraints to the free adjustment of labor inputs.

To control for any constraints to hiring and firing, I present estimates that allow labor to be a dynamic variable. In this case, I include labor in the set of state variables, so that (3) becomes

$$\omega_{it} = g(k_{it}, m_{it}, a_{it}, l_{it}),$$

and the moment conditions change accordingly.

4.5 Accounting for different intermediate input prices

The above estimation procedure assumes that firms face the same intermediate input (material) prices. For the present application, it is worth relaxing this assumption. Small firms in Ghana use a variety of financing sources to purchase materials, including loans from banks, loans from family, and their own financial assets. Firms using different financing sources effectively face different input prices: for example, if they finance the purchase from bank loans, the corresponding interest rate will increase the price of materials.

The notion that loans are used to purchase intermediate inputs may be unfamiliar, as a typical Western firm would use loans mainly for purchasing investment goods. It would deal with liquidity problems using trade credit or other short term business credits, such as overdrafts. By contrast, among small firms in Ghana, investment is not common. At the same time, they accumulate substantial debt which suggests that loans are used to deal with liquidity problems, including the purchase of materials. In the data, the mean value of material purchases is on average 12 times higher than the mean value of investment.

Since firms that can get lower interest rates are effectively facing lower material prices, they can purchase more materials for given productivity. This may violate the assumption that input demand is monotonic in productivity, which is needed to write down equation (3). In this case, monotonicity may only hold conditional on the material price, and I therefore include a measure of material prices based on the source of financing in the estimation.[18]

To calculate the interest rate on firms' portfolio, I take a weighted average of the formal and informal interest rates, using the relative loan amounts as weights. Denote as D_{it} a firm's total loans (from formal or informal sources) and r_{it}^{p} the average interest rate on its portfolio. Since in the survey material use is reported in monetary amounts, I write the unit

[18] I know of only one other attempt to deal with the heterogeneity of input prices across firms in the estimation of production functions. De Loecker et al. (2016) deal with unobserved input prices by proxying for them with an index of output quality. In the Ghanaian context, the variation in firms' sources of financing is likely to be a more important determinant of input price differences.

price of materials as

$$p_{it}^{m} = \begin{cases} 1 \; if \; D_{it} \leq 0 \; or \; (D_{it} > 0 \; and \; I_{it} \geq D_{it}) \\ 1 + \frac{r_{it}^{p}}{100} \frac{D_{it} - I_{it}}{M_{it}} \; if \; D_{it} > 0 \; and \; I_{it} < D_{it} \\ 1 + \frac{r_{it}^{p}}{100} \; if \; D_{it} > 0 \; and \; I_{it} = 0 \end{cases} \tag{7}$$

where I_{it} is the firm's investment in capital. If the firm does not borrow, or if investment is greater than the loan amount, then the firm is assumed to pay the market price for the materials. This assumes that the firm uses the loan first to purchase investment goods and only the remaining part of the loan is used for purchasing materials. Similarly, if the firm makes an investment, then only the remaining part of the loan will count toward an increase in the material price. If the firm does not invest, then the firm spends the entire loan on purchasing materials. Table 4 shows the summary statistics of the material price variable.

Table 4: Summary statistics of material prices

	Mean	Median	Std. dev	10%	90%	N
Input price	1.012	1.000	0.047	1.000	1.000	1602
Input price conditional on Debt $>$ 0	1.032	1.000	0.073	1.000	1.105	583
Price conditional on Debt $>$ 0 and Investment $=$ 0	1.082	1.055	0.098	1.000	1.238	228

Notes: Material prices are computed based on (7). Prices are increased with the interest rate if materials are purchased using a loan. Prices are deflated to 1991 Ghanaian Cedis.

Using the firm-specific material prices, equation (3) becomes

$$\omega_{it} = g(k_{it}, m_{it}, a_{it}, p_{it}^{m}).$$

Below I present estimation results using different combinations of these extensions and use all of them together in the preferred specification.

4.6 Estimation results

Table 5 reports the estimates of the gross output production function parameters using the procedures described above. The estimation is done separately for each industry. In each case, column (1) treats firms' labor choice as static. Columns (2)-(4) treat labor as a dynamic variable and present estimates with or without adjusting for the correlation of productivity and exit decisions, and with or without conditioning on material price differences. At the bottom of the table, I present the returns to scale measure implied by these estimates, as well as a J-test for the joint validity of the instruments. We can never reject the validity of the instruments at the 10 percent level.[19]

[19] The model relies on the assumption that the capital and material demand rules (and labor in a dynamic labor input case) are monotonic in productivity, thus (3) can be inverted. For the Olley and Pakes (1996) two step estimation procedure, this assumption can be tested as described in Ackerberg et al. (2006), section 2.4.1. This test relies specifically on the two step nature of the estimation. I am not aware of a corresponding test for the 1-step Wooldridge estimation used here.

Table 5: Gross output production function parameter estimates

	Food / Bakery / Alcohol				Furniture / Wood			
	Static labor	Dynamic labor			Static labor	Dynamic labor		
	No adjustment for exit	No adjustment for exit	Adjustment for exit	Adjustment for exit with prices	No adjustment for exit	No adjustment for exit	Adjustment for exit	Adjustment for exit with prices
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Capital	0.050***	0.144**	0.137**	0.128***	0.068***	0.075**	0.051*	0.058**
	(0.015)	(0.060)	(0.058)	(0.016)	(0.019)	(0.032)	(0.029)	(0.027)
Material	0.773***	0.508***	0.580***	0.808***	0.762***	0.821***	0.824***	0.825***
	(0.027)	(0.091)	(0.046)	(0.026)	(0.038)	(0.083)	(0.054)	(0.040)
Worker	0.306***	0.409***	0.324***	0.106**	0.235***	0.156**	0.173***	0.165***
	(0.057)	(0.091)	(0.061)	(0.053)	(0.054)	(0.070)	(0.067)	(0.055)
Age	0.025	0.412*	0.478***	0.063	0.214***	0.203*	0.263**	0.283***
	(0.036)	(0.223)	(0.153)	(0.039)	(0.045)	(0.111)	(0.103)	(0.061)
Returns to scale	1.129	1.061	1.041	1.042	1.065	1.052	1.048	1.048
p-value	0.000	0.3683	0.5074	0.0992	0.005	0.1425	0.0449	0.061
Hansen J statistic	25.547	33.023	41.493	42.185	18.156	31.056	23.314	23.303
p-value	0.143	0.104	0.147	0.108	0.445	0.121	0.384	0.385
N	381	381	381	381	461	461	461	461
	Garment / Textiles				Machines / Metal			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Capital	0.108***	0.151**	0.064***	0.088***	0.062	0.094***	0.079***	0.149***
	(0.024)	(0.070)	(0.018)	(0.019)	(0.085)	(0.023)	(0.030)	(0.046)
Material	0.573***	0.634***	0.747***	0.700***	0.661***	0.786***	0.832***	0.805***
	(0.046)	(0.090)	(0.048)	(0.029)	(0.100)	(0.043)	(0.042)	(0.042)
Worker	0.314***	0.297***	0.185***	0.250***	0.552***	0.117*	0.071	0.075
	(0.046)	(0.081)	(0.031)	(0.043)	(0.149)	(0.071)	(0.078)	(0.083)
Age	0.177***	-0.415	0.079*	0.066	-0.503	0.005	0.101*	0.035
	(0.044)	(0.276)	(0.047)	(0.049)	(0.365)	(0.045)	(0.057)	(0.039)
Returns to scale	0.995	1.082	0.996	1.038	1.275	0.997	0.982	1.029
p-value	0.8977	0.4331	0.9251	0.2413	0.0488	0.9339	0.6659	0.4563
Hansen J statistic	24.333	12.826	33.127	36.221	16.951	30.035	37.180	25.834
p-value	0.184	0.802	0.193	0.167	0.151	0.184	0.115	0.309
N	369	369	369	369	391	391	391	391

Notes: Gross output (revenues deflated with firm-specific price deflators) production function parameter estimates obtained using the Wooldridge extension of the Levinsohn-Petrin procedure. The estimation controls for three ownership dummies. Robust standard errors clustered at the firm level in parentheses. Columns (2)-(4) treat labor as a state variable. Returns to scale is the sum of the capital, material, and labor coefficients; the corresponding p-value is for the null that the sum of these is equal to 1. Hansen's J statistic is a test of the overidentifying restrictions. * significant at 10 percent, ** significant at 5 percent, *** significant at 1 percent.

Column (1) corresponds to the standard Levinsohn and Petrin (2003) approach. Coefficient estimates tend to be highly significant. As expected, materials have the highest and capital the lowest share in each industry. Starting from column (2), the estimation accounts for any barriers to the free adjustment of the labor input. In most cases moving from column (1) to (2) reduces the coefficient on labor. This likely reflects the fact that in column (1) increases in output associated with increases in labor are fully attributed to labor, while in column (2) changes in labor can also result in increased productivity. Column (3) adjusts for firm exit. In the literature, increases in the capital coefficient are typically interpreted as reflecting the ability of larger firms (with more capital) to stay in the market even when they face low productivity (Fernandes and Pakes, 2008). In the present context, decreases in the capital coefficient are possible, e.g., if smaller firms are more likely to receive loans from family and friends and these have a much lower interest rate than those available from the formal sector (Szabó and Ujhelyi, 2015). In this case, it is the smaller firms that may have access to financing that allows them to stay in the market. Finally, column (4) also controls for different intermediate input prices across firms. This can raise the capital coefficient if larger firms face lower material prices. Holding everything else fixed, firms with a lower material price need lower productivity to produce the same amount of output. Once the lower material prices are accounted for, the role of capital in the production process can increase, as seems to be the case particularly for the machinery sector (which is highly capital intensive).[20]

My preferred specification is column (4) which combines all the adjustments to the standard production function estimates that seem relevant in the Ghanaian context (see sections 4.3-4.5). The following calculations in the main text are based on the parameter estimates from column (4) for each industry. Section 7 describes various robustness checks with results

[20] In De Loecker et. al (2016) correcting for input price variation using an index of output quality has different effects on the production function parameters depending on the sector. The authors argue that without controlling for input prices, the productivity measure includes both unobserved input and output price variation. Since these may or may not offset each other, there is no a priori expectation on how the parameters should change.

reported in the Online Appendix.

5 Inefficient input use

Using the production function estimates, I quantify inefficiency of input use and document its heterogeneity across firms. This section studies inefficiency relative to the cost-minimizing counterfactual, while the following section considers more limited counterfactual adjustments.[21]

5.1 Optimal input combinations

Using the production function parameter estimates, we can solve the firm's cost minimization problem to derive the conditional factor demands for the observed output levels and factor prices:

$$L_{it}^* = \left(\frac{Y_{it}}{\exp(\eta_{it})}\right)^{\frac{1}{\beta_L+\beta_K+\beta_M}} \left(\frac{\beta_L(r_{it}-\delta)}{\beta_K w_{it}}\right)^{\frac{\beta_K}{\beta_L+\beta_K+\beta_M}} \left(\frac{\beta_L p_{it}^m}{\beta_M w_{it}}\right)^{\frac{\beta_M}{\beta_L+\beta_K+\beta_M}} \exp\left(\frac{-\omega_{it}-\beta_a a_{it}}{\beta_L+\beta_K+\beta_M}\right) \quad (8)$$

$$K_{it}^* = \left(\frac{Y_{it}}{\exp(\eta_{it})}\right)^{\frac{1}{\beta_L+\beta_K+\beta_M}} \left(\frac{\beta_K w_{it}}{\beta_L(r_{it}-\delta)}\right)^{\frac{\beta_L}{\beta_L+\beta_K+\beta_M}} \left(\frac{\beta_K p_{it}^m}{\beta_M (r_{it}-\delta)}\right)^{\frac{\beta_M}{\beta_L+\beta_K+\beta_M}} \exp\left(\frac{-\omega_{it}-\beta_a a_{it}}{\beta_L+\beta_K+\beta_M}\right) \quad (9)$$

$$M_{it}^* = \left(\frac{Y_{it}}{\exp(\eta_{it})}\right)^{\frac{1}{\beta_L+\beta_K+\beta_M}} \left(\frac{\beta_M w_{it}}{\beta_L p_{it}^m}\right)^{\frac{\beta_L}{\beta_L+\beta_K+\beta_M}} \left(\frac{\beta_M (r_{it}-\delta)}{\beta_K p_{it}^m}\right)^{\frac{\beta_K}{\beta_L+\beta_K+\beta_M}} \exp\left(\frac{-\omega_{it}-\beta_a a_{it}}{\beta_L+\beta_K+\beta_M}\right) \quad (10)$$

Note that firms should condition their input decision on the predictable part of productivity (ω), which is a state variable in the firms' problem.[22] The estimation procedure used here allows separating this predictable term from the unpredictable error η in computing (8)-(10).

When computing these quantities, I first replace r_{it}, w_{it} and p_{it}^m with the corresponding

[21] All results use all available observations in the data.
[22] A recent paper by Midrigan and Xu (2014) highlights the importance of allowing firms to condition on persistent productivity in measuring inefficiency.

average factor prices observed in the dataset. Specifically, for capital, I use the average interest rate on formal loans. For each year, I take the average interest rate on formal loans for firms with nonzero formal loans (using the highest rate for firms with multiple formal loans). I then assign this rate to all firms observed in a given year in the dataset. For labor, I take the average annual wages, and for materials (which are measured in monetary amounts) I use a price of 1. Using the average price for all firms follows Fernandes and Pakes (2008) and has several advantages. First, it provides a measure of inefficiency relative to the neoclassical benchmark, where all firms face the same price and there are no frictions. Second, since firms without loans have no observed interest rate in the data, using the average rate for all firms avoids biases that would result from treating firms with and without loans asymmetrically. In the Appendix, I repeat some of the analysis using the firm-specific prices observed in the data instead of average prices, and find similar results.

To quantify the extent of inefficiency, I consider the ratio of the optimal and observed level of each input, which I will refer to as "underutilization" following Fernandes and Pakes (2008). For example, the underutilization of capital by firm i in year t is given by $K_{it}^{*}/K_{it}^{observed}$. When this ratio is above 1, it shows how much more capital would be required to efficiently produce the current level of output in the counterfactual, cost-minimizing scenario.[23]

Note that this counterfactual thought experiment holds prices fixed. Large changes in the amount of an input available in the economy are likely to affect both input and output prices. Such effects are not reflected in the above computation of firms' counterfactual decisions. Thus, the underutilization measures should not be interpreted as estimates of the impact of large scale policy experiments. They are intended to provide an estimate of the effects of small changes.

[23] In Fernandes and Pakes (2008) the counterfactual inputs are optimized one at a time, holding all other inputs fixed. For example, K^{*} is the optimal level of K holding constant the observed level of other inputs. By contrast, my counterfactual is the fully cost minimizing solution given by equations (8)-(10), where all inputs are set optimally to produce the current level of output.

Table 6: Underutilization of inputs

	All firms	Firms with formal loans	Firms with formal or informal loans	Firms with no loans
Capital				
Fraction of underutilizing firms	0.458	0.211	0.334	0.528
Average underutilization ratio	3.511	1.805	2.816	3.909
Median underutilization ratio	0.784	0.255	0.404	1.163
Total underutilization ratio - all firms	0.226	0.228	0.224	0.232
Total underutilization ratio - underutilizing firms	2.032	1.918	1.931	2.239
Labor				
Fraction of underutilizing firms	0.202	0.332	0.271	0.162
Average underutilization ratio	0.743	1.169	0.990	0.602
Median underutilization ratio	0.425	0.655	0.567	0.361
Total underutilization ratio - all firms	1.403	1.693	1.657	0.909
Total underutilization ratio - underutilizing firms	2.788	3.002	2.989	2.157
Material				
Fraction of underutilizing firms	0.692	0.751	0.729	0.671
Average underutilization ratio	1.269	1.257	1.259	1.275
Median underutilization ratio	1.185	1.183	1.179	1.188
Total underutilization ratio - all firms	0.810	0.778	0.781	0.913
Total underutilization ratio - underutilizing firms	1.205	1.201	1.202	1.216
N	1602	422	583	1019

Notes: Optimal input use is based on equations (8) - (10), computed using the gross output production function parameter estimates from Table 5, column (4). Average underutilization is the ratio of the optimal and observed input quantity averaged across all firms in a group. Median underutilization is the median of these ratios within a group. Total underutilization is total optimal input quantity divided by total observed input quantity in a group. All input prices are set equal across firms, equal to the yearly average observed prices in the data. The price of labor is based on wages. All values are in 1991 Ghanaian Cedis.

5.2 Summary statistics of input underutilization

Table 6 shows summary statistics of firm-level underutilization ratios for each input as defined in section 5.1. Ratios above 1 indicate that the current input use is below what would be optimal while ratios below 1 indicate higher-than-optimal input use. I present results for all firms as well as firms grouped by whether they have formal or informal loans.

In this sample, 45.8% of manufacturing firms are undercapitalized. However, overall firms use 40% less labor and 77.4% more capital than would be efficient based on cost minimization. These patterns echo recent findings from India (Fernandes and Pakes, 2008) and Chile (Petrin and Sivadasan, 2013) where the underutilization of labor is found the be especially large.

Labor is the input underutilized by the lowest fraction of firms: 20.2% among all firms in the data. However, total labor underutilization ratios (total counterfactual labor divided by total actual labor) are large, at around 2.8 across all groups. This is surprising since the official unemployment rate in Ghana was around 10% throughout the 1990s, indicating a substantial excess supply of labor.[24] The underutilization figures indicate that large inefficiencies exist in the allocation of the existing labor input.

Among all firms, the average undercapitalization ratio is 3.5. However, total undercapitalization (total counterfactual capital divided by total actual capital) is 0.23, indicating a surplus of capital in the economy. This suggests that the allocation of the existing capital stock may be more problematic than the low level of available capital in the economy.

This could be so for at least three reasons. First, since physical capital can be used as a collateral on loans while labor cannot, firms using external financing may have added incentives to substitute capital for labor. Consistent with this, firms with formal loans operate with the lowest undercapitalization ratios. In this group, only 21% have less-than-optimal capital, while over 52% of firms with no loans are undercapitalized. Second, firms

[24] Including the self-employed who would be willing to accept a salaried job, the excess supply of labor is likely to be much higher.

may substitute capital for labor in response to frictions on the labor market. This could happen if there is a shortage of workers with the right skills in the economy, or if there are constraints to the flexible adjustment of labor (e.g., hiring or firing costs due to unionization). For example, Soderbom and Teal (2004) estimate larger Ghanaian firms face substantially higher relative labor costs and argue that this is an important source of the dispersion of input choices. Third, if capital is "lumpy," small firms purchasing capital may necessarily overutilize capital in the short run. For example, capital may come in the form of indivisible machinery.[25] While testing a specific mechanism is beyond the scope of this paper, when I look at the firm-level correlates of underutilization ratios below, I present suggestive evidence consistent with the first and second and inconsistent with the third channel.

5.3 Underutilization and firm characteristics

This section presents correlations between firm-level underutilization and various firm characteristics. The literature has explored several channels through which specific frictions lead to inefficiencies and the Ghanian dataset offers an opportunity to look at corresponding associations between firm characteristics and the underutilization measures computed above. While I cannot isolate causal channels, exploring these correlations may motivate future research to take a closer look at specific mechanisms.

For this analysis, I run regressions of the form

$$Y_{it}^{j} = \boldsymbol{\beta}^{j}\mathbf{X}_{it} + u_{it}^{j}, \; j = 1, 2, 3 \tag{11}$$

where the dependent variable Y^j is one of the three underutilization measures. The right hand side variables $\mathbf{X}_{it}$ include firm-level characteristics which may be correlated with the underutilization measures (all three equations include the same set of right-hand side variables). This includes industry dummies (the same industries used in the production function esti-

[25] See Bigsten et al. (2005) for an analysis of African firms' investment behavior in the presence of adjustment costs and irreversibility.

mation, see Table 5); variables indicating the ownership structure (Ghanaian private, state or foreign owned); indicator variables for four geographic areas: Accra, Kumasi, Takoradi and Cape Coast; and indicator variables for whether the firm has formal and/or informal loans. I also include three variables to describe the human capital of the firm: share of management workers among all workers, education and age of the workers. The latter two variables are computed from individual worker interviews at the firms (answers are averaged out by firm-year-occupation, weighted by the number of employees in each occupation for a given firm-year). See Teal (2011) for details. To take into account different access to raw materials, I include the percentage of raw materials imported at the firm level. Different export opportunities are captured through two variables: percentage of output exported within and outside Africa. Finally, a unionization variable indicates whether some of the firm employees belong to a worker union.

The coefficient estimates are in Table 7. It is important to emphasize that the estimated coefficients in these regressions represent correlations rather than causal effects.

Firms with formal loans have significantly lower underutilization ratios for capital and higher underutilization ratios for labor. By contrast, firms with informal loans have lower underutilization ratios for labor. This is consistent with the explanation for capital overutilization discussed in the previous section where firms may be induced to substitute labor with capital in order to satisfy collateral requirements to obtain formal external financing.[26]

Another possible explanation for capital overutilization discussed above is that small firms may be forced to hold too much "lumpy" capital in the short run. Results in the Online Appendix appear inconsistent with this explanation, showing that firms in the third and fourth employment quartile in fact have lower underutilization ratios than smaller firms. Thus, it is larger firms that are disproportionately likely to overutilize capital. More generally, these patterns may reflect the fact that the availability of credit is connected to observed firm characteristics rather than future profitability, as has been observed by Bigsten et al.

[26]See Gelos and Werner (2002) for an analysis of how the effects of increased access to formal financing can be mediated by banks' collateral requirements.

Table 7: Underutilization measures and firm characteristics

	Capital	Material	Labor
State owned	-1.045	0.001	0.050
	(1.145)	(0.116)	(0.610)
Private Ghanaian	0.826	-0.099	-0.327
	(0.539)	(0.092)	(0.220)
Dummy for firms with formal loans	-1.191***	-0.093**	0.391***
	(0.435)	(0.041)	(0.147)
Dummy for firms with informal loans	0.487	0.036	-0.170**
	(0.999)	(0.057)	(0.079)
Management workers as a share of all workers	-0.011	0.003	-0.002
	(0.076)	(0.004)	(0.012)
Workers' average years of education	-0.276*	0.031***	-0.014
	(0.148)	(0.008)	(0.014)
Workers' average age	-0.157**	0.004	-0.004
	(0.061)	(0.003)	(0.006)
Unionization	-0.138	-0.039	0.380***
	(0.667)	(0.083)	(0.138)
Percentage of output exported within Africa	-0.022**	-0.000	0.005**
	(0.010)	(0.001)	(0.002)
Percentage of output exported outside Africa	0.136	-0.003	0.005*
	(0.138)	(0.002)	(0.003)
Percentage of raw materials imported	0.006	-0.000	-0.000
	(0.009)	(0.001)	(0.002)

Notes: N = 1299. Each column corresponds to a separate regression with the dependent variable listed in the column heading. Regressions include year, location and industry indicators. Underutilization measures are computed as described in Section 5.1 using the gross output production function parameter estimates from Table 5 column (4) for each industry. All input prices are set equal across firms, equal to the yearly average observed prices in the data. The price of labor is based on wages. All values are in 1991 Ghanaian Cedis. Standard errors clustered by firm in parentheses. * significant at 10 percent, ** significant at 5 percent, *** significant at 1 percent.

(2003), Bloom et. al. (2010), and others. In particular, Bigsten et al. (2003) show that in six African countries firms in the manufacturing sector are more likely to get a bank loan if they are larger, holding profitability constant.

Turning to measures of firm level human capital, Table 7 shows that firms with more educated or more experienced (older) employees have significantly lower underutilization ratios for capital. This could indicate that these firms are able to substitue skilled labor for missing capital in the production process. Firms where at least part of the labor force belongs to a union have a higher underutilization ratio for labor. This is consistent with these firms having a lower flexibility to adjust labor and / or higher than average wages.[27]

It may be argued that some of the correlations described above may reflect different technology choices by some firms. Production functions in this paper allow for some differences in technology based on industry and ownership status (see section 4.2), but other factors, such as the availability of different sources of financing or the presence of unions may also affect technology under some circumstances. At the same time, while industry and ownership tends to be time-invariant, other covariates included in the regressions in Table 7 change more frequently. In this sense, the latter may be less important determinants of firms' technology choices.

6 Potential gains from small adjustments

In this section I use the production function estimates to compute the marginal value of inputs and the gaps relative to input prices. Following Petrin and Sivadasan (2013), this allows me to quantify the gain in value added from small (one unit) adjustments, and study which firms are likely to yield the highest gains.

[27] See Hirsch (2004) and Lee and Mas (2012) for recent work on the many channels through which unions can affect firms' operations.

6.1 Input gaps

After estimating the parameters of the production function, the marginal product of capital multiplied by the firm's output price yields the marginal value product of capital,

$$MVPK_{it} = p_{it}^{output} \frac{\partial Q_{it}}{\partial K_{it}},$$

where Q_{it} is the quantity produced and p_{it}^{output} is the price of the firm's output. Similarly, the value of the marginal product of labor is given by $MVPL_{it}$ and the value of the marginal product of the intermediate input is $MVPM_{it}$. As above, firms are assumed to condition their input decision on the predictable part of productivity (ω), which is a state variable in the firms' problem. The value of the marginal product of capital is therefore given by:[28]

$$MVPK_{it} = \frac{\beta_K p_{it}^{output} Q_{it}}{\exp(\eta_{it}) K_{it}}.$$

To calculate these marginal value products, I use the estimated coefficients β and shocks η, and, from the data, the nominal value of firm revenue $p_{it}^{output} Q_{it}$ (real revenue multiplied by the firm-level price deflator) and the value of the capital stock K.

An alternative view of the error terms estimated from the production functions is that the full error represents productivity that should be taken into account in the calculation of marginal products (see Petrin and Sivadasan, 2013). In the Robustness section below I therefore present an alternative computation of marginal value products that conditions on the full error term. I find that this does not affect the main findings.

In the absence of constraints, the standard first order conditions for profit maximization should hold. E.g., for capital, this is $MVPK_{it} = r_{it} - \delta$, where r_{it} is the firm specific interest rate. In the presence of constraints, the gap between the value of the marginal product and the marginal input price provides a measure of input inefficiency. The value of this absolute

[28] Given the Cobb-Douglas technology, the marginal product of capital conditional on ω is given by $\frac{\partial Q}{\partial K} \frac{1}{\exp(\eta)} = \frac{\beta_K Q}{\exp(\eta) K} = \frac{\beta_K}{K} \exp(\omega) K^{\beta_K} L^{\beta_L} M^{\beta_M}$.

gap in real terms (deflated with the consumer price index) is[29]

$$GapK_{it} = |MVPK_{it}/CPI_t - r_{it} - \delta| \tag{12}$$

Constructing the marginal product measures similarly for the other inputs, we have

$$GapL_{it} = \frac{|MVPL_{it} - w_{it}|}{CPI_t}$$

$$GapM_{it} = \frac{|MVPM_{it} - p^m_{it}|}{CPI_t}.$$

Inefficiencies can arise from many different sources. Some of these induce lower demand for inputs than would be optimal (for example a tax on capital), so the difference between the value of the input and its marginal value product is positive. On the other hand, some sources of inefficiency (e.g., firing costs) lead to higher input usage than would be optimal, resulting in a negative difference. The absolute value of the gap captures both types of allocative inefficiencies. A wider gap corresponds to higher levels of inefficiency.

At the firm level, these gap measures will allow me to study the correlates of inefficiency. Aggregating them across firm, they permit making statements about the impact of improving input allocation on total value added in the economy (Petrin and Sivadasan, 2013). Specifically, write total value added as $\sum_i \sum_j (p_i^{output} Q_i - w_{ij} X_{ij})$, where X_{ij} denotes an input with corresponding price w_{ij}. We then get that the gap for an input at a given firm is exactly the derivative of aggregate value added with respect to that input ($p_i^{output} \frac{\partial Q_i}{\partial X_{ij}} - w_{ij}$). Thus, adjusting a given input by one unit in the optimal direction in every firm, holding all else constant, leads to an increase in average value added equal to the average absolute gap.

Note that in this framework, all prices are taken to be the actual prices observed in the data. In this sense, this measure of input inefficiency is conditional on any distortions that exist in those prices. This allows the evaluation of small-scale adjustments in input use,

[29] Since r_{it} is a percentage, it does not need to be deflated separately.

which may be more realistic for some policies. An alternative view of inefficiency is that it should include all deviations from the neoclassical ideal in which an input's price would be the same across all firms. If one is willing to take a stance on what the prices in this ideal world would be, it is possible to compute the gaps using these counterfactual prices. This can be used as a measure of lost value added due to both price distortions and input inefficiency conditional on those prices. Below, I will emphasize the gap measures obtained using actual firm-specific prices but also present results using the same (average) prices for all firms in the Appendix.

Firm-specific prices are constructed as follows:

Price of capital. I use the highest interest rate reported by the firm or the risk-free deposit rate, whichever is highest. The latter is equivalent to assuming that firms which were able to raise some capital at unusually low interest rates cannot raise unlimited capital at those rates. Thus, the marginal cost of capital for these firms must be the opportunity cost of their own resources: the deposit rate.[30] Among the 422 firm-year observations with positive formal loans, 14% have a lower interest rate than the deposit interest rate reported in the WDI. These low interest rates are associated with small loan amounts: the median loan amount among these firms is 20.5 million Cedis (by comparison, the median loan amount for loans with an interest rate above the deposit rate is 108 million Cedis). Informal loans, most of which have an interest rate of 0, have a median loan amount of only 0.21 million Cedis. These figures lend some support to the assumption that loans with interest rates below the deposit rate are used to finance inframarginal units of capital. For firms with no loans, the price of capital is equal to the opportunity cost of using their own resources, which I approximate using the risk-free deposit rate.

The above assumes that the interest rate on an additional unit of capital is equal to the opportunity cost of the marginal capital currently held by the firm. In reality, the cost of an

[30] There are 79 such firm-year observations in the dataset. Since the gap measure for capital is the difference between the marginal value product and the interest rate, using the higher deposit-free rate for these firms yields more conservative estimates of the gap.

extra unit of capital may differ from the cost of the last unit. In computing the gaps below, a potential concern arises if this difference is correlated with the firm's capital stock. For example, if firms that have 0 loans would face unusually high interest rates if they decided to borrow, then the capital gap computed with the deposit rate will be overestimated for these firms. To address this, I compare gap estimates with firm-specific prices to those obtained using the average formal interest rates in the data, and find similar patterns.

Price of labor. I construct the price of labor in two alternative ways, using wages or using total earnings (wage plus in-kind or cash allowances). In some cases, especially at smaller firms, allowances might be large. Teal (2011) provides an earning measure that includes food allowance, housing allowance, clothing allowance, transportation allowance, Christmas bonus (13th month), production or merit bonuses, and other allowances.[31] Earnings are about 26% higher on average than the reported wage in the sample. Small firms (1st quartile, less than 8 employees) report about 70% higher earnings than wages, while in the second quartile (8 - 13 employees) earnings are only 24% higher. At larger firms (3rd and 4th quartile) additional allowances increase the wage by 12%.

Price of materials. For materials, I use the same prices that were constructed for the production function estimation (see section 4.5).

6.2 Returns to capital

Table 8 shows the estimated marginal value product of capital. As before, these values are presented for three groups of firms: firms that have nonzero formal loans, firms with nonzero loans from either formal or informal sources, and all firms. Among all firms, the median (mean) value of the marginal product of capital is 37.4% (136.3%) and, as expected, it is lower for firms that have access to loans. Firms that have access to either formal or informal loans operate with a median (mean) $MVPK$ of 18.4% (97.0%). For firms that have access to formal loans, this value is 12.6% (58.9%). By comparison, the average risk free deposit

[31] Table 1 shows the summary statistics of this measure.

interest rate during the study period was 26.7%.[32]

The recent literature in development economics finds similarly high returns to capital.[33] Using bank data and an indirect method of computing the value of the marginal product of capital, Banerjee and Duflo (2014) find 74-100% per year for India. Udry and Anagol (2006) estimate returns to capital between 50-250% among small agricultural farmers in Ghana. De Mel, McKenzie and Woodruff (2008) implemented a randomized experiment in Sri Lanka, where they gave cash and in-kind grants to small retail firms, generating an exogenous capital shock. The estimated return to capital was at least 4.6-5.3% *per month*. Fafchamps et al. (2014) study a similar experiment in Ghana and find a return to capital of 15% per month. These studies use different methods to compute the marginal value of capital and none of them uses direct estimates of the production function. Nevertheless, the results reported here confirm that the marginal value of capital is often substantially higher than the average interest rate in the economy.

Table 8: Marginal value product of capital

	Mean	Median	Std.	10%	90%	N
Firms with formal loans	0.589	0.126	1.349	0.028	1.527	422
Firms with formal or informal loans	0.970	0.184	1.874	0.030	2.635	583
All firms	1.363	0.374	2.171	0.035	4.372	1602

Notes: The table reports summary statistics for the estimated marginal value product of capital (MVPK) for three groups of firms. MVPK is computed as described in Section 6.1, using estimates from Table 5 column (4) for each industry. Estimated MVPK values were winsorized by 2.5 percent on both tails before computing the summary statistics. All values are in 1991 Ghanaian Cedis.

6.3 Summary statistics of input gaps

Tables 9 and Tables 10 show the gap estimates capital, labor and materials, respectively. As described in Section 6.1, gaps provide a measure of inefficiency based on the absolute difference between the value of the marginal product of inputs and the prices the firms pay for them. The average gap across firms is equal to the increase in average value added

[32] Risk free deposit interest rate, averaged over 1991-2002. Source: World Development Indicators.

[33] A good summary of the recent results is in McKenzie and Woodruff (2008).

that could be achieved if the usage of a given input was adjusted by 1 unit in the optimal direction at every firm, taking input prices as given. A larger gap indicates more inefficient input allocation in this sense.

Table 9 reports the capital gap using the firm-level price of capital observed in the data (see Section 6.1). Among firms with formal loans the average gap is 0.649 (64.9 percentage points). Among all firms the average gap is 1.259. This shows that firms without formal loans are especially far from their efficient capital stock. Recall that these firms also have especially high underutilization ratios for capital (Table 6). Firms with either formal or informal loans are located in between, with an average gap of 0.970. Table 16 in the Appendix shows the corresponding estimates assuming that the price of capital is the same for all firms in a given year, and equal to the average observed interest rate on formal loans reported in the data for that year. The distribution of the estimates tends to be similar to those in Table 9, indicating that the patterns observed above are not sensitive to how the price of capital is measured (while the average values are larger, recall that estimates using firm-specific prices were winsorized). Below, I present some simple back-of-the-envelope calculations to gauge the implications of these numbers for the potential gain in aggregate value added from small adjustments.

Table 9: Gap for capital

Parameter estimates	Mean	Median	Std.	10%	90%	N
Firms with formal loans	0.649	0.356	1.212	0.131	1.144	422
Firms with formal or informal loans	0.959	0.373	1.730	0.142	2.252	583
All firms	1.259	0.359	2.043	0.105	4.022	1602

Notes: The gap measure is computed as described in Section 6.1. The marginal product is computed using the gross output production function parameter estimates from Table 5 column (4) for each industry. Input prices are based on the firm-level prices observed in the data. Estimated gap values were winsorized by 2.5 percent on both tails before computing the summary statistics. All values are in 1991 Ghanaian Cedis.

For labor, I present estimates both using wages and using earnings as the input price.[34] With the wages observed in the data, the average absolute gap for labor is 0.19 million Cedis

[34] See Section 4.5 for the construction of material prices.

and the median is 0.11 million Cedis. Using earnings instead of wages yields similar values (0.21 million and 0.13 million, respectively). By comparison, a worker's yearly wage at the average firm in the sample is 0.3 million Cedis.

Table 10: Gap for labor and materials

Parameter estimates	Mean	Median	Std.	10%	90%	N
Labor - Wage	191959	107703	235439	17952	467999	1423
Labor - Earnings	206769	134853	228436	20855	459167	1495
Material	1.308	0.865	1.381	0.147	3.256	1594

Notes: The gap measure is computed as described in Section 6.1. The marginal product is computed using the gross output production function parameter estimates from Table 5 column (4) for each industry. Input prices are the firm-level prices observed in the data. For labor, these are either wages or total earnings. Estimated gap values were winsorized by 2.5 percent on both tails before computing the summary statistics. All values are in 1991 Ghanaian Cedis.

6.4 Implications for aggregate value added

Summing the gap measures for a specific input across firms yields a measure of the total gain in value added that can be achieved if the utilization of that input is moved by 1 unit in the optimal direction at every firm. I now perform some simple back-of-the-envelope calculations to gauge what the estimated gap measures would imply for the Ghanaian economy.

First, I provide an estimate of the economy-wide gap for each input. To do this, I use the 2003 National Industrial Census to obtain the distribution of firm size (number of employees) in Ghana for the universe of 26,088 manufacturing firms.[35] For the 9 categories given in the census (from 1-4 employees to 500+ employees), I compute the average gaps for the corresponding firms in my sample. To estimate the economy-wide gap, I calculate the weighted sum of these gaps, using the number of firms from the census in each category as weights. For labor, I obtain a total gap of 4015 million Cedis ($\approx$ 8.1 million USD). According to these estimates, this is the gain in value added that could be achieved by adjusting the

[35] There were three rounds of this national census in Ghana in the past 50 years, in 1962, 1987, and 2003. I take the 2003 firm size distribution because that is closest to my period of study. See Ghana Statistical Service (2006).

labor force by 1 worker in the optimal direction at every manufacturing firm in Ghana. The size of this gain is approximately 0.16% of the country's GDP.[36]

For capital and materials, I calculate the total gain that can be achieved by adjusting input usage by the equivalent of 1 worker's annual wage (0.3 million Cedis $\approx$ 604.5 USD) at every firm. For capital, I get 13,763 million Cedis, or 0.53% of GDP. For materials, the corresponding figure is 8121 million Cedis, or 0.32% of GDP.

6.5 The correlates of firm-level input gaps

As seen above, adjusting inputs by one unit in the optimal direction at every firm can yield substantial gains in aggregate value added. Which firms are mainly responsible for these gains? To answer this question, I look at the firm-level correlates of input gaps. I estimate regressions similar to (11), where the dependent variables Y^j are the three gap measures (computed with the observed prices) and the firm-level productivity term. Although the focus of this paper is not the determinants of total factor productivity in the Ghanaian manufacturing sector, it is interesting to study the relationship between inefficiency of input use and productivity. I therefore also discuss some notable features of manufacturing productivity based on the findings.[37] The right hand side variables are the same in all four equations and identical to those in (11).

The estimates are shown in Table 11. In terms of ownership structure, domestic private firms have lower productivity than foreign firms. They also exhibit lower gaps for labor while state owned firms have higher gaps (although these differences are not statistically significant at conventional levels). This difference between state owned and private firms could reflect the former facing higher adjustment costs for this input as in Cooper et al. (2015).

[36] By comparison, Petrin and Sivadasan (2013) find an increase equal to 0.5% of GDP from adjusting the blue collar worker input in Chile.

[37] In a recent paper, Asker et al (2014) argue that observed misallocation can be explained entirely by time-series volatility in productivity. That is, firms hold suboptimal levels of capital, not because they are "constrained" but rather because they are hedging against a bad productivity shock. Although I do not have enough year-industry observations to study this question in a meaningful way, I analyze the firm-level correlates of both productivity and misallocation below.

Turning to the sources of financing, I find significantly lower capital gaps and higher labor gaps among firms that have formal loans. By contrast, firms with informal loans have lower gaps for labor. These patterns echo those seen above for the underutilization measures and may reflect the increased incentives to purchase capital goods when these can serve as collateral for formal loans. I do not find significantly different capital gaps for firms that have informal loans relative to those that have no loans. This could be due to the fact that informal loan amounts are too small to matter for efficiency: the average formal loan amount is 42 times larger than the average informal loan. Interestingly, we do not see any difference in estimated productivity among firms with different types of loans, holding everything else constant.

In terms of our measures of firms' human capital, firms with a more educated workforce and firms with older employees achieve lower gaps for capital. This could be consistent with higher human capital leading to more efficient operations or with more efficient firms being able to attract a higher quality workforce. Firms with a unionized workforce have significantly larger gaps for labor.

Together with the findings on underutilization patterns described in Section 5, three main findings emerge. First, firms with formal loans are less likely to underutilize capital and more likely to overutilize it. However, these firms are closer to their efficient capital stock than firms with no formal loans, implying that remedying inefficiencies among firms without formal loans would result in larger gains in the economy. Second, firms with more human capital are closer to using capital efficiently and have lower underutilization ratios for this input. This may indicate that these firms successfully substitute skilled labor for the missing capital input. Third, firms with a unionized workforce are further from their optimal use of labor, which may reflect constraints to the free adjustment of the labor input faced by these firms.[38]

[38] Alternatively, firms that require less flexibility in their labor input may find it less costly to have a union.

Table 11: Gap measures, productivity, and firm characteristics

	Gap for Capital	Gap for Material	Gap for Labor	Productivity
State owned	-0.140	0.287	43.765	0.110
	(0.542)	(0.449)	(82.162)	(0.123)
Private Ghanaian	0.235	-0.519	-61.003	-0.139**
	(0.185)	(0.337)	(41.143)	(0.059)
Dummy for firms with formal loans	-0.310**	-0.172	64.834**	-0.021
	(0.133)	(0.132)	(25.414)	(0.033)
Dummy for firms with informal loans	0.046	-0.025	-40.806*	-0.005
	(0.284)	(0.127)	(20.885)	(0.028)
Management workers as a share of all workers	-0.008	-0.019*	-1.946	-0.006**
	(0.023)	(0.010)	(1.874)	(0.003)
Workers' average years of education	-0.120**	0.087***	-2.247	0.014**
	(0.050)	(0.026)	(3.788)	(0.007)
Workers' average age	-0.042***	0.017	0.075	0.001
	(0.014)	(0.011)	(1.471)	(0.002)
Unionization	-0.113	0.064	76.929**	0.005
	(0.211)	(0.282)	(36.694)	(0.052)
Percentage of output exported within Africa	-0.004	-0.005***	0.497	-0.001**
	(0.003)	(0.002)	(0.409)	(0.000)
Percentage of output exported outside Africa	0.013	-0.008*	1.477	-0.000
	(0.020)	(0.005)	(0.944)	(0.001)
Percentage of raw materials imported	0.004	-0.002	0.519	-0.000
	(0.003)	(0.004)	(0.644)	(0.001)

Notes: Each column corresponds to a separate regression with the dependent variable listed in the column heading. Regressions include location, industry, and year indicators. Gaps are computed as described in Section 6.1 using the gross output production function parameter estimates from Table 5 column (4). Input prices are the firm-level prices observed in the data. Productivity refers to the transmitted component of the productivity term (ω_{it}). All values are in 1991 Ghanaian Cedis. N = 1175. Estimated gap values were winsorized by 2.5 percent on both tails. Standard errors clustered by firm in parentheses. * significant at 10 percent, ** significant at 5 percent, *** significant at 1 percent.

7 Robustness Checks

This section examines the robustness of the findings to using different depreciation rates, an alternative definition of the productivity residual, and to using earnings instead of wages as the price of labor. For each of these, I compute the corresponding gap measures and repeat the regressions in Table 11. I find that the patterns of the estimates is similar to those seen above.

Depreciation rate. The results above followed Bigsten et al. (2005) in assuming a 6% deprecation rate. There are other figures used in the literature: Söderbom and Teal (2004) and Frazer (2005) use 2%, which is also the figure originally used by the survey team in their construction of the capital stock variable. Some of the rates used for other countries are larger (see Schündeln, 2013). In the above exercise, the choice of the depreciation rate matters at several different stages. First, it affects the computed values of firm's capital stock. Second, the capital stock variable in turn affects the production function parameter estimates. Third, the optimal input combinations and gap values computed in sections 5.1 and 6.1 are affected by both the capital stock and the production function parameters, as well as directly by the chosen depreciation rate.

To check the sensitivity of the findings to alternative deprecation rates, the Online Appendix presents results using alternatively a 2% and a 10% depreciation rate. Remarkably the correlations between gap measures and firm characteristics always follow the same patterns as those discussed in the main text.

Alternative productivity residual. Marginal products computed in Section 6.1 use only the transmitted component of the productivity term, which was assumed to be observed by the firms. An alternative view is that input gaps are affected similarly by both error terms $(\omega + \eta)$. Using the production function coefficients estimated above, I recalculate marginal value products and gap measures conditioning on the entire error term.[39] The resulting

[39] In this case the marginal value product for capital only depends on the data and the estimated β_K. This provides a straightforward way to construct confidence intervals for $MVPK$ and I present these in the Online Appendix.

regressions are in the Online Appendix. The main findings are broadly similar to those seen above.

Alternative wage measure. I also repeat the regressions using the alternative wage measure discussed in Section 6.1 (total earnings). This includes in-kind and cash allowances as well as wages. The findings are again similar to those reported above.

The Online Appendix contains further robustness checks using alternative production function specifications.

8 Conclusion

This paper directly estimates the production function of manufacturing firms in Ghana in order to evaluate the allocation of inputs in the economy. I use a long panel dataset that contains detailed information on firm-specific prices, including various wage measures and interest rates on loans from various sources. The estimation method incorporates recent developments in the production function literature, and explicitly accounts for the input price differentials created by the fact that firms use different financing sources to purchase materials. Using the estimates, I compute two measures of inefficiency: (1) the ratio of the observed and cost-minimizing input use given the observed output, and (2) the gap between the firm level input price and the marginal value product of inputs. I use these measures to quantify the effect of inefficient input use on aggregate value added, and study the firm-level correlates of inefficiency and productivity.

I find that the fraction of undercapitalized firms is 46%, but overall firms use 77% more capital and 40% less labor than would be optimal. Regressing measures of inefficiency on firm characteristics yields three robust patterns. First, firms with formal loans are less likely to underutilize capital and more likely to overutilize it. However, these firms are closer to their efficient capital stock than firms with no formal loans, implying that remedying inefficiencies among firms without formal loans would result in larger gains in the economy. Second, firms with more human capital are closer to using capital efficiently and have lower underutilization ratios for this input. This may indicate that these firms successfully substitute skilled labor for the missing capital input. Third, firms with a unionized workforce are further from their optimal use of labor, which may reflect constraints to the free adjustment of the labor input faced by these firms.

My results speak to a large micro literature that uses a variety of indirect methods to infer the marginal value of capital. Similarly to these prevous studies, the structural estimates obtained in this paper indicate large potential gains in value added from adjusting input use in the optimal direction. Further research may usefully refine these estimates by modeling relevant sources of heterogeneity in the firms' optimization problem, such differences between cash and in-kind investments or entrepreneurs' ability (De Mel et al., 2008, Fafchamps et al., 2014).

References

[1] Ackerberg, D., C. L. Benkard, S. Berry, and A. Pakes (2006): "Econometric Tools for Analyzing Market Outcomes", in: J.J. Heckman and E.E. Leamer (eds.): *Handbook of Econometrics*, Vol. 6A, Elsevier: Oxford, UK.

[2] Ackerberg, D., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica* 83(6), 2411–2451,

[3] Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic Inputs and (Mis)Allocation", *Journal of Political Economy* 22(5), 1013-1063.

[4] Banerjee, V. A. and K. Munshi (2004): "How Efficiently Is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur," *The Review of Economic Studies* 71(1), 19-42.

[5] Banerjee, V. A. and E. Duflo (2005): "Growth Theory through the Lens of Development Economics", in: S. Durlauf and P. Aghion (eds.): *Handbook of Economic Growth*, Elsevier: Oxford, UK.

[6] Banerjee, V. A. and E. Duflo (2014): "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," *Review of Economic Studies* 81(2), 572-607.

[7] Basu, S., and J. Fernald (2002): "Aggregate Productivity and Aggregate Technology", *European Economic Review* 46, 963-991.

[8] Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Patillo, M. Söderbom, F. Teal, and A. Zeufack (2003): "Credit Constraints in Manufacturing Enterprises in Africa", *Journal of African Economies* 12(1), 104-125.

[9] Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, R. Oostendorp, C. Patillo, M. Söderbom, and F. Teal (2005): "Adjustment Costs and Irreversibility as Determinants of Investment: Evidence from African Manufacturing", *B.E. Journal of Economic Analysis & Policy* 4(1), article 12.

[10] Bloom, N., A. Mahajan, D. McKenzie, and J. Robert (2010): "Why Do Firms in Developing Countries Have Low Productivity?," *American Economic Review: Papers & Proceedings* 100, 619-623.

[11] Caselli, F. and J. Feyrer (2007): "The Marginal Product of Capital", *Quarterly Journal of Economics* 122 (2), 535-568.

[12] Cooper, R., G. Gong, and P. Yan (2015): "Dynamic labor demand in China: Public and private objectives," *The RAND Journal of Economics* 46(3), 577–610.

[13] De Loecker, J., P.K. Goldberg, A.K. Khandelwal, and N. Pavcnik (2016): "Prices, markups and trade reform," *Econometrica* 84(2), 445–510.

[14] De Mel, S., D. McKenzie and C. Woodruff (2008): "Returns to Capital in Microenterprises: Evidence from a Field Experiment", *The Quarterly Journal of Economics* 73(4), 1329-1372.

[15] Fafchamps, M., D. McKenzie, S. Quinn, and C. Woodruff (2014): "Microenterprise Growth and the Flypaper Effect: Evidence from a Randomized Experiment in Ghana," *Journal of Development Economics* 106(C), 211-26.

[16] Fernandes, A. M. and A. Pakes (2008): "Factor Utilization in Indian Manufacturing: A look at the World Bank Investment Climate Surveys Data", NBER Working Paper, 14178.

[17] Frazer, Garth (2005): "Which Firms Die? A Look at Exit from Manufacturing in Ghana", *Economic Development and Cultural Change* 53(3), 585-617.

[18] Frazer, Garth (2006): "Learning the Master's Trade: Apprenticeship and Human Capital in Ghana", *Journal of Development Economics* 81(2), 259-298.

[19] Gelos, R.G., and A.M. Werner (2002): "Financial liberalization, credit constraints, and collateral: Investment in the Mexican manufacturing sector," *Journal of Development Economics* 67, 1–27.

[20] Ghana Statistical Service (2006): *2003 National Industrial Census Report, Background and Results*, Accra, Ghana.

[21] Hirsch, B.T. (2004): "Reconsidering union wage effects: Surveying new evidence on an old topic," *Journal of Labor Research* 25 (2), 233–266.

[22] Hsieh, C. T., and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and in India", *Quarterly Journal of Economics* 124(4), 1403-1448.

[23] Jones, Patricia (2001): "Are educated workers really more productive?", *Journal of Development Economics* 64, 57–79.

[24] Levinsohn J., A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies* 70(2), 317-342.

[25] Lee, D., and A. Mas (2012): "Long-run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961-1999, *Quarterly Journal of Economics* 127(1), 333-378.

[26] Matsuyama, K. (2007): "Aggregate Implications of Credit Market Imperfections," in: *NBER Macroeconomics Annual Volume 22*.

[27] McKenzie D. and C. Woodruff (2008): "Experimental Evidence on Returns to Capital and Access to Finance in Mexico", *The World Bank Economic Review* 22(3), 457-482.

[28] Midrigan, V. and D. Yi Xu (2014): "Finance and Misallocation: Evidence from Plant-Level Data", *American Economic Review* 104(2), 422-58.

[29] Olley, S. G. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica* 64(6), 1263-1297.

[30] Petrin, A. and J. Sivadasan (2013): "Estimating Lost Output from Allocative Inefficiency, with an Application to Chile and Firing Costs," *Review of Economics and Statistics* 95(1), 286-301.

[31] Restuccia, D. and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics* 11, 707-720.

[32] Schündeln, M. (2005): "Modeling Firm Dynamics to Identify the Cost of Financing Constraints in Ghanaian Manufacturing," working paper.

[33] Schündeln, M. (2013): "Appreciating depreciation: physical capital depreciation in a developing country," *Empirical Economics* 44, 1277-1290.

[34] Shenoy, A. (2014): "Market failures and misallocation: The costs of factor and financial market failures in rural Thailand," working paper.

[35] Söderbom, M., and F. Teal (2004): "Size and efficiency in African manufacturing firms: evidence from firm-level panel data," *Journal of Development Economics* 73, 369-394.

[36] Szabó, A., and G. Ujhelyi (2015): "Family Financing and Productivity Among Small Manufacturing Firms in Ghana," working paper.

[37] Teal, F. (2011): "Ghana Manufacturing Enterprise Survey, Rounds I - VII (12 years: 1991-2002), Explanatory Notes on Dataset," University of Oxford, Centre for the Study of African Economies.

[38] Udry, C. and S. Anagol (2006): "The Return to Capital in Ghana," *American Economic Review* 96(2), 388-393.

[39] Vollrath, D. (2014): "The Efficiency of Human Capital Allocations in Developing Countries," *Journal of Development Economics*, 108, 106-118.

[40] Wooldridge, J. M. (2009): "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters* 104(3), 112-114.

9 Appendix

Table 12: Sectoral distribution of the sample

Sector	N	Percentage
Alcohol	14	0.87
Bakery	159	9.93
Food	208	12.98
Furniture	320	19.98
Garment	329	20.54
Machines	63	3.93
Metal	328	20.47
Textile	40	2.5
Wood	141	8.8
Total	1602	100

Source: Ghanaian Manufacturing Survey, 1991-2002.

Table 13: Interest rate by year

	Formal loans	Formal and informal loans	WDI deposit interest rate
1992	0.29	0.21	0.16
1993	0.30	0.20	0.24
1994	0.35	0.24	0.23
1995	0.37	0.31	0.29
1996	0.38	0.32	0.34
1997	0.38	0.27	0.36
1998	0.34	0.28	0.32
1999	0.39	0.33	0.24
2000	0.38	0.35	0.29
2001	0.38	0.35	0.31
2002	0.35	0.28	0.16

Notes: The first two columns contain the average annual interest rates from the dataset. The third column shows the Ghanaian deposit interest rate reported in the World Bank's World Development Indicators http://data.worldbank.org/data-catalog/world-development-indicators

Table 14: Source of interest rate data

Wave	Source of loan	Question number
Wave 1	Overdraft facilities	[R37Q02AA- R37Q02AE]
		[R37Q04A- R37Q04E]
	Formal borrowing	[R39Q06AA- R39Q06AE]
		[R39Q07A- R39Q07E]
	Informal borrowing	R44Q06
		R44Q09
		R44Q07
		R44Q10
Wave 2	Overdraft facilities	Z32Q03
		Z32Q04
	Formal borrowing	[Z34Q09A - Z34Q09D]
		[Z34Q11A - Z34Q11D]
	Informal borrowing	[Z38Q05A- Z38Q05F]
		[Z38Q08A- Z38Q08F]
		[Z38Q06A- Z38Q06F]
		[Z38Q09A- Z38Q09F]
Wave 3	Overdraft facilities	L39Q03
		L39Q04
	Formal borrowing	
	Informal borrowing	[L43Q03A- L43Q03G]
		[L43Q04A- L43Q04G]
		L44Q05
		L44Q05P
Wave 4	Overdraft facilities	S9CQ3
		S9CQ4
	Formal borrowing	[S9CQ8A- S9CQ8E]
		[S9CQ10AA- S9CQ10AE]
	Informal borrowing	[S9DQ3A- S9DQ3H]
		[S9DQ4A- S9DQ4H]
		S9DQ5A
		S9DQ5B
Wave 5	Overdraft facilities	F7DQ5A, F7DQ5B
		F7DQ6
	Formal borrowing	[F7DQ9AA- F7DQ9AE , F7DQ9BA- F7DQ9BE]
		[F7DQ12A- F7DQ12E]
	Informal borrowing	[F7FQ3A- F7FQ3H]
		[F7FQ4A- F7FQ4H]
		F7FQ5A
		F7FQ5B
Wave 6	Overdraft facilities	N7DQ5A , N7DQ5B
		N7DQ6A , N7DQ6AB
	Formal borrowing	[N7DQ9AA- N7DQ9AE]
		[N7DQ9BA- N7DQ9BE], [N7DQ13A- N7DQ13E]
		[N7DQ14AA- N7DQ14AE]
		[N7DQ14BA- N7DQ14BE], [N7DQ18A- N7DQ18E]
	Informal borrowing	[N7FQ3A- N7FQ3J]
		[N7FQ4A- N7FQ4J]
		N7FQ5A
		N7FQ5B
Wave 7	Overdraft facilities	P7DQ5A , P7DQ6A
		P7DQ5B , P7DQ6B
		P7DQ5C, P7DQ6C
	Formal borrowing	[P7DQ9AA- P7DQ9AE], [P7DQ9BA- P7DQ9BE]
		[P7DQ9CA- P7DQ9CE], [P7DQ13A- P7DQ13E]
		[P7DQ14AA- P7DQ14AE], [P7DQ14BA- P7DQ14BE]
		[P7DQ14CA- P7DQ14CE], [P7DQ18A- P7DQ18E]
	Informal borrowing	[P7FQ3A- P7FQ3J], [P7FQ4A- P7FQ4J]
		[P7FQ5A- P7FQ5J]
		P7FQ6A
		P7FQ6B

Notes: Question numbers refer to the original survey questions in the Ghanaian Manufacturing Survey. The dataset and documentation can be downloaded from http://www.csae.ox.ac.uk/datasets/ghana-rped/Ghmain.html.

Table 15: Underutilization measures and firm characteristics, firm-specific prices

	Capital	Material	Labor
State owned	-0.825	0.011	-0.139
	(0.740)	(0.108)	(0.276)
Private Ghanaian	0.729*	-0.084	-0.143
	(0.387)	(0.082)	(0.130)
Dummy for firms with formal loans	-0.885***	-0.080**	0.248***
	(0.285)	(0.038)	(0.078)
Dummy for firms with informal loans	0.252	0.043	-0.108*
	(0.666)	(0.045)	(0.057)
Management workers as a share of all workers	-0.025	0.003	0.000
	(0.042)	(0.004)	(0.006)
Workers' average years of education	-0.323***	0.031***	-0.012
	(0.109)	(0.008)	(0.013)
Workers' average age	-0.106***	0.004	-0.002
	(0.030)	(0.003)	(0.004)
Unionization	-0.197	-0.019	0.323***
	(0.477)	(0.074)	(0.105)
Percentage of output exported within Africa	-0.013**	-0.001	0.003***
	(0.005)	(0.001)	(0.001)
Percentage of output exported outside Africa	0.021	-0.003	0.004*
	(0.039)	(0.002)	(0.002)
Percentage of raw materials imported	0.006	-0.000	0.001
	(0.007)	(0.001)	(0.002)

Notes: N = 1175. Each column corresponds to a separate regression with the dependent variable listed in the column heading. Regressions include year, location and industry indicators. Underutilization measures are computed as described in Section 5.1 using the gross output production function parameter estimates from Table 5 column (4) for each industry. Input prices are set equal to the firm-specific prices observed in the data. The price of labor is based on wages. Estimated gap values were winsorized by 2.5 percent on both tails. All values are in 1991 Ghanaian Cedis. Standard errors clustered by firm in parentheses. * significant at 10 percent, ** significant at 5 percent, *** significant at 1 percent.

Table 16: Gap for capital, average prices

Parameter estimates	Mean	Median	Std.	10%	90%	N
Firms with formal loans	15.983	0.360	188.966	0.136	1.104	422
Firms with formal or informal loans	12.152	0.371	160.852	0.151	2.179	583
All firms	5.566	0.387	97.183	0.131	3.927	1602

Notes: The gap measure is computed as described in Section 6.1. The marginal product is computed using the gross output production function parameter estimates from Table 5 column (4) for each industry. The price of capital is set equal for all firms, based on the average interest rate reported in the data. All values are in 1991 Ghanaian Cedis.

Table 17: Gap for labor and materials, average prices

Parameter estimates	Mean	Median	Std.	10%	90%	N
Labor - Wage	307302	141614	1493018	31962	397764	1602
Labor - Earnings	316884	160533	1492505	35127	381116	1602
Material	1.565	0.872	4.332	0.151	3.286	1594

Notes: The gap measure is computed as described in Section 6.1. The marginal product is computed using the gross output production function parameter estimates from Table 5 column (4) for each industry. Input prices are the average prices observed in the data. For labor, these are either wages or total earnings. All values are in 1991 Ghanaian Cedis.