World Development Vol. 39, No. 8, pp. 1333–1346, 2011

0305-750X/$ - see front matter
ELSEVIER
www.elsevier.com/locate/worlddev
doi:10.1016/j.worlddev.2010.12.002

Wages Equal Productivity. Fact or Fiction? Evidence from Sub Saharan Africa

JOHANNES VAN BIESEBROECK *
K.U.Leuven, Belgium
CEPR, United Kingdom

Summary. — If labor markets operate with only minor frictions, productivity premiums associated with worker characteristics should equal the corresponding wage premiums. We evaluate this for labor market experience, schooling, job tenure, and training using matched employer–employee data from the manufacturing sector of three sub-Saharan countries. Equality holds remarkably well in Zimbabwe (the most developed country in the sample), but not at all in Tanzania (the least developed), while results are intermediate in Kenya. Where equality fails, the pattern is for more general human capital characteristics (such as experience) to receive a wage return that exceeds productivity, while the reverse applies to more firm-specific characteristics (such as tenure). Localized labor markets and imperfect substitutability of worker-types provide a partial explanation.

Key words — sub-Saharan Africa, human capital, wage, labor market

1. INTRODUCTION

In the textbook economics world, markets are the most efficient institution to allocate scarce resources. Demand and supply are equalized, profit opportunities are arbitraged away, and factor prices are determined by the marginal productivity of the market clearing production factor. In the real world, there are frictions, unobservable characteristics, adjustment costs, erroneous expectations, and maybe discrimination; all of which can distort the market equilibrium away from efficient allocation. This should not necessarily worry us, researchers, as the theory is only intended to be a stylized version of reality. However, a systematic gap between costs (wages) and benefits (productivity) can provide information about crucial omissions from the theory.

A well-functioning labor market should perform at least two tasks: allocating workers to firms and determining market clearing wage levels. A large literature studies the importance of frictions and search costs in the allocation process (Mortensen & Pissarides, 1999). In the conclusion to his chapter on labor markets in the *Handbook of Development Economics*, Rosenzweig (1988) stresses that solid empirical evidence on market distortions, which play a prominent role in many theories, is rare. Hsieh and Klenow (2009) find recently that workers in the manufacturing sectors of China and India are not allocated to the most productive firms, leaving room for productivity-enhancing reallocation of workers.

In this paper, we do not study the allocation process directly, but estimate the differences in productivity and wage premiums associated with human capital characteristics. A second aspect of the labor market is to determine wage rates. If frictions and informational asymmetries are of lesser importance, we would expect arbitrage to equalize the remuneration of a characteristic to its productivity contribution. The lack of suitable data has hampered extensive study of this prediction. Employee surveys do not contain information on firm-level output and inputs, necessary to assess productivity. Data sets of firms or plants generally lack information on all but a few basic characteristics of their workforce.

Matched employer–employee data sets contain the necessary information, but these are not widely available. [1] The observed employees are used to estimate average values of worker characteristics by firm. Hellerstein, Neumark, and Troske (1999) pioneered the approach, jointly estimating a plant level wage equation with a production function. Using US administrative record information, they test for equality of the wage and productivity premiums associated with a number of characteristics. They only find a discrepancy for the gender dummy: women are estimated to be only 16% less productive than their male coworkers, but are paid 45% less. [2]

The bulk of the evidence for developed countries points toward equal wage and productivity returns for most worker characteristics. Most recently, using 1990 US data, Hellerstein and Neumark (2007) cannot reject equality of the two premiums for black or married workers, and for different occupation categories. They do confirm that the gender wage gap systematically exceeds the productivity gap and that relative wages for older workers (55+ years) exceed their relative productivity. A gap is also present for education (“some college”), but it is a lot smaller. The productivity premium exceeds the wage premium by 15%, while the discrepancies for gender and age exceed 40% of the respective premiums.

Similar work for France in Crepon, Deniau, and Pérez-Duarte (2003), for Israel in Hellerstein and Neumark (1999), and for Norway in Haegeland and Klette (1999) finds no gender discrimination. Only a few characteristics in those studies are associated with a wage premium that differs significantly from the productivity premium: older workers are overpaid in France, engineers are underpaid in Israel, Norwegian workers with 8–15 years of experience earn too little, and those with more than 15 years of experience too much. Dearden, Reed, and Van Reenen (2006) focus on the effects of training using industry-level data in the UK. They separately estimate wage and production equations and find that a productivity effect of training that substantially exceeds the wage effect, but no formal test is reported.

* Seminar participants at the University of Illinois, Kellogg School of Management, K.U.Leuven, the NBER Productivity meetings, and four anonymous referees provided useful suggestions. Funding by the Connaught Foundation is gratefully acknowledged. Final revision accepted: December 6, 2010.

The evidence for developing countries is even more limited. Jones (2001) estimates a firm-level production function jointly with an individual-level wage equation for Ghana.[3] She finds that women are 42–62% less productive and are paid 12–15% less. No formal test is reported, but standard errors are large. Her focus is on the reward for an extra year of schooling, which is estimated to equal the productivity gain associated with education, both are 7% per year. When discrete levels of education attainment are used, the results are ambiguous. The differences in point estimates are large but estimated imprecisely and none of the formal tests finds a statistically significant difference.[4] Bigsten *et al.* (2000) gauge the link between wages and productivity indirectly. First, they estimate the returns to education in five sub-Saharan countries using a standard wage equation. Then, they separately estimate the production function, including lagged levels of education as a proxy for human capital. They find that the implied rate of return to human capital is very low, in particular only a fraction of the return to physical capital.

In this paper, I provide information on the wage-productivity gap for three sub-Saharan countries: Tanzania, Kenya, and Zimbabwe. All three countries are relatively poor, but during the sample period the GDP per capita for Zimbabwe exceeded that for Tanzania by a factor of six, while Kenya was intermediately developed. The initial focus is on two human capital characteristics, experience and schooling, but the analysis is subsequently extended to include job tenure and training programs. I extend the Hellerstein *et al.* (1999) methodology of comparing wage and productivity effects to worker characteristics that are measured as continuous variables. Consistent aggregation to the firm level requires a Taylor expansion approximation as in Frazer (2001).[5]

The main findings are the following. In Tanzania, the poorest country of the three, the wage premiums deviate substantially and significantly from the corresponding productivity premiums. The gaps are much smaller and all are insignificant in relatively more developed Zimbabwe. Results for Kenya, an intermediate country in level of development, are intermediate: equal remuneration can be rejected for experience, but not for schooling. Several robustness checks are provided for this pattern. Allowing for imperfect substitution between male and female workers attenuates the discrepancy, especially in Kenya.

When we add tenure and training to the analysis, the way in which equality fails also proves interesting. Even though formal education (building general skills) is rewarded well everywhere, productivity returns to schooling are modest or nonexistent in Tanzania and Kenya. In contrast, the productivity advantages associated with formal training programs, which are likely to generate more firm-specific human capital are large in all countries, ranging from 26% in Zimbabwe to 61% in Tanzania.[6] None of these productivity gains accrue to the worker in the form of higher salary in Tanzania or in Kenya. Such a mismatch between productivity and salary premiums provides low incentives for workers to enroll in training. Of course, it strengthens the incentives for firms to offer training to their workers and can contribute to the financing of training programs. The pattern for experience (general skills) and tenure (firm-specific skills) is broadly similar.

Finally, when we conduct the analysis separately in the two most important industrial cities or regions of each country, we always find that the equality of wages and productivity is least likely to be rejected in the most developed city. The reduced samples lead to less precise estimates, but in each country the point estimates are closer together in the richer city. It suggests that the differences at the national level are not purely driven by localized labor markets.

Equality of relative wages and productivity is an important maintained assumption in several literatures. When productivity growth is calculated by subtracting labor growth from output growth, different categories of workers are weighted by their wage shares (Jorgenson & Griliches, 1967). If the equality between wages and productivity fails to hold systematically in developing countries, productivity growth measures will be biased. When assessing the importance of firm-specific human capital, proxied by tenure, it is assumed that wage increases reflect productivity advances (Topel, 1991). For indirect evidence on the wage-productivity link one can refer to Brown (1989), who finds that wage increases within a plant occur predominantly when on-the-job training is taking place. Our approach provides direct evidence on the link.

The remainder of the paper is organized as follows: The measurement framework to compare wage and productivity premiums is introduced in Section 2 and the data in Section 3. Benchmark results are presented in Section 4, followed by a number of robustness checks and extensions to multiple human capital characteristics and local labor markets. The implications of the results are discussed with some general conclusions in the final section.

2. EMPIRICAL MODEL

The methodology owes a great deal to Hellerstein *et al.* (1999). If firms minimize costs and labor markets operate as a spot market, the wage premium of a worker should equal its productivity premium. Barring imperfect information, any difference will be arbitraged away. As an example, assume that the average productivity of male workers exceeds the productivity of female workers by φ_M per cent. With perfect substitutability between workers, which is relaxed later, the production function can be written as[7]

$$Q = Af(K, L_F + (1+\varphi_M)L_M). \tag{1}$$

The first order conditions for cost minimizing input choices of the firm entail that in an efficient labor market the relative wage for both types of workers should equal their relative productivity:

$$\frac{W_M}{W_F} = \frac{MP_M}{MP_F} \equiv 1+\varphi_M,$$
$$\lambda_M \equiv \frac{W_M - W_F}{W_F} = \frac{MP_M - MP_F}{MP_F} \equiv \varphi_M.$$

Jointly estimating the wage (λ_M) and productivity (φ_M) premiums associated with each characteristic makes it possible to test for the equality in Eqn. (1) for several characteristics individually or jointly.[8]

To implement this approach, we need to specify a production function, introduce several human capital characteristics, and aggregate the wages and effective productivity contributions of all employees to the firm level. For now, we assume a Cobb–Douglas functional form for the production function, which is relaxed in a robustness check. It can be written in logarithms as

$$\ln Q = \ln A + \alpha_K \ln K + \alpha_L \ln \tilde{L} + \epsilon_q. \tag{2}$$

The labor aggregate $\tilde{L}$ is the sum of the effective labor over all employees, where each employee's contribution is multiplied by its respective productivity level. Following the human capital theory of Mincer (1974), this is defined as

$$\tilde{L} = \sum_{i=1}^{L} \exp(\varphi_M M_i + \varphi_S S_i + \varphi_X X_i),$$

where φ_M, φ_S, and φ_X are the productivity premiums associated with males, and each year of schooling or experience.[9] They measure how the effective labor contribution of each individual varies with schooling in percentage terms, e.g., $\varphi_S \equiv \partial \ln \tilde{L}_l / \partial S_i$.

In order to use the labor aggregate without observing the characteristics of all workers, we follow Frazer (2001) and take the first order approximation of $\tilde{L}$, which amounts to

$$\ln \tilde{L} = \ln L + \varphi_M \frac{\sum_i M_i}{L} + \varphi_S \frac{\sum_i S_i}{L} + \varphi_X \frac{\sum_i X_i}{L} \equiv \ln L + \varphi_M \overline{M} + \varphi_S \overline{S} + \varphi_X \overline{X}. \quad (3)$$

The total number of workers is adjusted by the different productivity premiums multiplied by the average value over all employees for each characteristic. Substituting Eqn. (3) in the production function allows estimation of the different productivity premiums from just the average proportion of male workers in each firm, average levels of schooling and experience, and the usual output and input variables.

The same first order approximation can be used to aggregate individual-level wages that vary by worker characteristics to the firm level. If we define the firm's aggregate wage bill as the sum over the wages of all employees i as

$$\overline{W}L \equiv \sum_{i=1}^{L} W_i = \sum_{i=1}^{L} W_0 \cdot \exp(\lambda_M M_i + \lambda_S S_i + \lambda_X X_i)$$

with the wage premiums defined as percentage increases over the baseline wage of a female worker with 0 year of schooling and education. It can then be approximated at the firm level as

$$\ln \overline{W} \simeq \ln W_0 + \lambda_M \overline{M} + \lambda_S \overline{S} + \lambda_X \overline{X} \quad (4)$$

In the empirical application, a random error ε_w is added to Eqn. (4) and the baseline wage can be made dependent on average hours worked, total employment, and other firm characteristics. Eqns. (2) and (4) are estimated jointly with Zellner's seemingly unrelated regression (SUR) estimator, allowing for correlation between the two error terms. Using nonlinear least squares allows direct estimation of the coefficients.

The approach suggested here differs from the original method from Hellerstein *et al.* (1999) that only considers discrete characteristics. Assume, as an example, that male and female workers (M or F) can achieve high or low schooling (S or U). The effective labor force aggregate in Eqn. (1) now has to distinguish between four categories of workers

$$\tilde{L} = L_{UF} + (1 + \varphi_{UM}) L_{UM} + (1 + \varphi_{SF}) L_{SF} + (1 + \varphi_{SM}) L_{SM}. \quad (5)$$

Estimation now requires data on employment shares of all four worker categories (L_{ij}/L). If more characteristics are introduced the data requirements would become prohibitive quickly. Lacking the necessary data to estimate employment shares on narrowly defined cells, one can still proceed if one is willing to assume that the education premium is the same for male and female workers and that the relative share of high and low schooling workers is also the same for male and female workers. In that case, Eqn. (5) simplifies to

$$\tilde{L} = L\left(1 + \varphi_M \frac{L_M}{L}\right)\left(1 + \varphi_S \frac{L_S}{L}\right). \quad (6)$$

and similar calculations for the wage bill. Estimation now remains feasible if the number of employees sampled remains constant and more characteristics are introduced.

In our approach, we make the similar assumption that the percentage gain in effective human capital (or salary) for a year of schooling is independent of any other worker characteristic. Most of the vast literature estimating Mincerian wage equations relies on the same assumption, but we relax it for male and female workers in a robustness check. However, we do not need to assume that the relative proportion of workers with different values of a characteristic is independent of other characteristics.[10]

Several alternatives to (2) and (3) are estimated in the robustness checks. First, generalizing the production function to a translog specification does not require any change in the key definition of the labor aggregate. Second, we can allow decreasing returns to schooling and experience by incorporating quadratic terms in S_i and X_i in the specification of each worker's productivity (and also in the wage equation). As the quadratic terms would drop out of the first order approximation, we now use a second order approximation that involves the within-firm variances and covariances for all characteristics.[11] These terms are missing in the firm-level production function estimated by Jones (2001). Third, augmenting the model with additional human capital characteristics is straightforward. A continuous measure of the years of job tenure can be added in the same way as experience ($\varphi_T T_i$) and a dummy variable capturing whether a worker has followed a training program will appear in the same way as the fraction of male workers ($\varphi_R R_i$).

3. DATA

We estimate the model on an unbalanced panel of manufacturing firms from Tanzania, Kenya, and Zimbabwe. The data comes from firm surveys coordinated by the Regional Program of Enterprise Development (RPED) at the World Bank in three consecutive years between 1992 and 1995. Sampling of firms was stratified by size, giving each employee in the manufacturing sector an equal probability of having its employer selected. This yields a sample of 110–191 firms per year, accounting for 12–31% of manufacturing value added in each country. Large firms are strongly overrepresented; Van Biesebroeck (2005) compares the sample with the universe of manufacturing firms.

Each year, at most 10 employees per firm were interviewed as well. While firms can be linked over time as a panel, this was not possible for the workers. In most cases, employees were selected randomly from a few broad occupation categories in proportion to the total employment in each category, which was collected first from the employer. Because of missing data and some firms employing fewer than 10 workers, we observe an average of just over six employees per firm.[12] Appendix provides more information on the exact sampling procedure used in each country.

Because of the many small and medium sized firms in the sample, the fraction of all employees that is sampled is not as low as it might appear. The median fraction of employees interviewed in the first year is 26% for firms in Tanzania and 20% for Kenya, as median employment is only 14 and 25 workers, respectively. The coverage is lower in Zimbabwe, where only a quarter of firms in the sample have fewer than 27 employees, and the median sampling rate of employees is only 6%.

The following firm-level variables are used: value added is total sales minus raw materials and intermediate inputs; labor input is the total number of full-time employees; capital is the replacement value of the plant and equipment at the end of the year; salary is the annual wage bill per employee, including all monetary employment expenditures. Nominal variables are deflated with GDP deflators. At the worker level, we observe the fraction of male workers in each firm exactly. In addition, we construct a weighted average by firm, using the occupation category weights of the following variables: schooling, labor force experience, and tenure (the number of years with the current employer), all measured continuously in years; a training dummy capturing completion of a formal training program, excluding on-the-job training, and the number of hours worked per week. Summary statistics on the countries, firms, and workers are in Table 1; additional information on variable construction is in Appendix A.

Information on productivity is only available at the firm level and worker characteristics and wages have to be aggregated to carry out the joint estimation. Identification of the different premiums comes from variation across firms in the composition of the workforce and average salaries or output. Employee-level wage regressions in the working paper version, Van Biesebroeck (2003), which are discussed in Appendix, confirm that the aggregation does not obscure how an individual's characteristics are rewarded.

The level of development of the three sample countries differed substantially. GDP per capita (in PPP in 1992) ranged from $395 in Tanzania, less than half the $1089 in Kenya, to a level almost six times as high in Zimbabwe ($2459). The same development ranking is apparent from the U.N. human development index, labor productivity in industry, relative manufacturing-agriculture wages and employment, and a host of infrastructure statistics. It also holds for the firms in the sample. The median firm in Tanzania achieved only 38% of the labor productivity of the median firm in Kenya, while labor productivity in Zimbabwe was 42% higher than in Kenya, with similar differences for total factor productivity. These comparisons confirm that Zimbabwe is by far the most developed country of the three, while Tanzania is lagging far behind.

4. RESULTS

(a) Benchmark estimates

Estimation results by country for Eqns. (2) and (4) jointly are in Table 1. In this and all following specifications, hours worked and year, industry, and location fixed effects are included as controls in both equations and robust-White standard errors are reported.[13] The results are robust to the inclusion of additional controls, as illustrated in robustness checks below. Inputs and outputs in the production function are normalized by employment, consistent with the definition of the dependent variable in the wage equation. Log employment is included in both equations, allowing the base wage to vary with firm size in the wage equation and variable returns to scale in the production function.

The input coefficients in the production function are estimated precisely and the point estimates are plausible. The relative importance of capital and labor is similar across

Table 1. *Summary statistics*

	Tanzania	Kenya	Zimbabwe
Country characteristics			
Population	27.1 m	25.0 m	10.3 m
% employed in industry	4.9%	7.3%	8.6%
GDP/capita (PPP)	$395	$1,089	$2,459
VA/employee in industry (USD)	$983	$1,705	$7,049
Firm characteristics			
Number of firms	171	191	110
Sample total as share of			
Manufacturing GDP	0.31	0.17	0.26
Manufacturing labor force	0.15	0.12	0.31
Relative level of development			
Median *LP* in sample[a]	38	100	142
Median *TFP* in sample[a]	54	100	143
Monthly wage (USD)	55.9	117.0	203.3
Variables used in the analysis[b]			
Value added (log)	11.2 (2.5)	10.1 (2.6)	9.9 (2.2)
Capital stock (log)	11.7 (3.2)	10.3 (3.1)	9.2 (2.6)
Employment	112.6 (320.0)	100.8 (271.7)	252.1 (534.4)
Salary (log)	6.8 (1.3)	5.4 (0.9)	4.3 (0.7)
Workers interviewed per firm	6.0 (3.2)	6.3 (3.4)	5.7 (2.6)
Worker characteristics			
Workers in the sample	1,018	1,206	619
Variables used in the analysis[b]			
Male (%)	0.84 (0.26)	0.90 (0.19)	0.86 (0.23)
Experience (years)	16.4 (7.0)	16.0 (7.3)	18.5 (8.0)
Schooling (years)	11.9 (3.3)	10.7 (2.4)	10.9 (2.8)
Tenure (years)	7.5 (5.0)	7.6 (5.1)	10.1 (6.1)
Received training (%)	0.07 (0.17)	0.08 (0.19)	0.16 (0.27)
Hours worked (log)	3.8 (0.2)	3.8 (0.2)	3.8 (0.1)

Sources: World Bank (2000) and own calculations for the sample statistics.
Notes: Aggregate statistics refer to 1992. Firm and worker level statistics refer to the first year of the survey, 1992 for Tanzania and 1991 for the other two countries. The proportion of male workers is exact; the other worker characteristics are a weighted average for the firm constructed from the employee information and total employment by occupation category.
[a] Relative to Kenya, see Van Biesebroeck (2005).
[b] Means and standard deviations calculated across firms.

countries and returns to scale are moderately increasing, although never significantly so. After normalizing by employment, the fit for the production function is comparable to that of the wage equation, but standard errors on the human capital characteristics are higher.

When comparing coefficients on worker characteristics in both equations, a word of caution is appropriate. The wage equation coefficients are merely representations of the different premiums. Under the assumption of a competitive spot labor market, they capture the productivity effect of the different human capital characteristics. As a result, they should equal the coefficients from the production function. If equality can be rejected, (some of) the underlying assumptions are rejected. Under alternative assumptions of labor market operations, for example monopsony power for firms or matching of heterogeneous workers and firms, it is not obvious how to compare the coefficients from both equations. Wage coefficients will now also capture other effects than the productivity of human capital.

The coefficients on the gender dummy tend to be estimated imprecisely. The negative coefficient for males in the wage equation is the result of labor sorting, not within-firm gender differences, as discussed in Appendix. Only for Kenya can we reject equality of the wage and productivity premiums for male workers, and even there only at a 10% significance level. As the coefficients on the gender dummy are unstable throughout, we do not focus on them.

The returns to experience and schooling in the wage equation are precisely estimated and correspond well to results at the individual level (Van Biesebroeck, 2003). Salaries rise most rapidly with experience in Tanzania, twice as fast as in Zimbabwe. Comparing across the three countries, wage premiums for education seem to vary inversely with the experience premium. In contrast, the contributions of experience and education in the production function both rise with the level of development. In Tanzania, experience contributes negatively to productivity, perhaps an age effect, and education contributes nothing. In Kenya, there is no significant effect of experience on production, while schooling contributes positively, although not in proportion to the wage premium. In Zimbabwe, the return to the worker—in the form of higher salary—and the return to the firm—in the form of higher output—for experience and especially for schooling are estimated very closely.

The point estimates for the gap between wage and productivity premiums for experience and schooling are largest in Tanzania, at respectively 4.3% and 4.2%, and still sizeable in Kenya, at 2.2% and 4.1%. In Zimbabwe, the gaps are only 0.5% and 0.1% and equality of the returns cannot be rejected. The p-values for the statistical tests, reported at the bottom of Table 2, are 0.58 and 0.99. In the other two countries, equality of the returns to experience can firmly be rejected. The p-values for equality of the schooling coefficients are a lot lower than in Zimbabwe, but do not justify a formal rejection of equality in either country.

The joint test for equality of the returns to schooling and experience confirms the pattern. In Tanzania, by far the least developed economy, the p-value of the Wald test is only 1%. In Kenya, the p-value still tends toward rejection at 9%, largely due to the wage premium for experience. In Zimbabwe, none of the differences between the estimated coefficients is even remotely significant, and the same is true for the joint test. The discrepancies between the wage and productivity premium are most striking for labor market experience. In the two least developed countries, workers receive substantial pay increases over their careers, which are not backed up by any discernible productivity effect. Moreover, these pay increases are very uniform as they are estimated very precisely.

The different results for the different countries cannot simply be attributed to less precisely estimated coefficients for Zimbabwe. Its standard errors are only higher in the wage equation, but for the test statistic that combines information from both equations it has the lowest standard error for experience and is comparable to the other two countries for schooling.

Table 2. *Joint estimation of wage and production equations*

	Tanzania		Kenya		Zimbabwe	
	Wage	Output	Wage	Output	Wage	Output
Labor		0.819		0.888		0.935
		(.077)***		(.056)***		(.068)***
Capital		0.221		0.314		0.237
		(.036)***		(.033)***		(.041)***
Male	0.155	0.474	0.086	0.657	−0.095	0.186
	(.122)	(.387)	(.106)	(.270)**	(.224)	(.300)
Experience	0.016	−0.027	0.013	−0.009	0.008	0.013
	(.005)***	(.015)*	(.004)***	(.010)	(.007)	(.009)
Schooling	0.044	0.002	0.059	0.018	0.073	0.072
	(.010)***	(.031)	(.010)***	(.026)	(.021)***	(.029)**
Observations	316	316	544	544	210	210
R^2	0.24	0.23	0.44	0.41	0.44	0.53
Test for equality of coefficients (p-value)						
Experience ($\lambda_X - \varphi_X$)	0.00		0.04		0.58	
Schooling ($\lambda_S - \varphi_S$)	0.16		0.13		0.99	
Joint test	0.01		0.09		0.82	

Notes: Joint estimation (SUR) of the wage equation and production function at the firm level. The sample for Tanzania covers 3 years, 1992–94, for Kenya also three years, 1991–94, and for Zimbabwe 2 years, 1991–92. Controls in both equations include employment, hours worked and year, industry, and location fixed effects. Robust-White standard errors are in parentheses.
* Significance at 10% level.
** Significance at 5% level.
*** Significance at 1% level.

(b) Robustness checks

The general pattern of the results—that the failure of wage and productivity premiums to be equalized is negatively related to the level of development of the country—survives a host of robustness checks. We performed the analysis with discrete definitions for the human capital characteristics, diminishing returns to schooling and experience, with a translog production function, with additional controls in both equations, taking sampling error explicitly into account, controlling for measurement error in the capital stock and the wage variable, and limiting the sample to the first year for each country. We briefly discuss these results and report the *p*-values on the joint test for equality of the wage and productivity coefficients for experience and schooling in Table 3. [14]

The first robustness check is to define human capital characteristics discretely, as in the existing literature. The limited number of workers interviewed per firm limit us from defining experience and schooling categories more finely than high or low. If we assume the relative proportion of workers according to one characteristic is independent of other characteristics, the human capital aggregate can be derived as in Eqn. (6) without requiring an approximation in the aggregation. The results are largely unchanged. The rejection of equality for Tanzania and Kenya is even more pronounced: the *p*-values for the test of joint equality of the schooling and experience premiums in line (a) of Table 3 are lower than in Table 2.

In the next variation, the assumption of a constant marginal impact of schooling and experience is relaxed. In order for the quadratic terms to survive the aggregation procedure to the firm level, it requires a second order approximation that incorporates within-firm variance and covariance terms. The results point to decreasing returns in most cases, as expected, but the fit of the regressions improves only marginally. The probability of rejecting the equality test again varies inversely with the level of development, but the *p*-values for both Tanzania and Kenya are larger than in the benchmark case. This is more the result of less precisely estimated coefficients than because of smaller gaps in the point estimates. For Tanzania, the mismatch between the premiums is still much more pronounced for experience. For Kenya, much of the discrepancy for the returns to experience is eliminated, but the effect of schooling in both equations is now almost entirely unrelated.

In the results in line (c) of Table 3, we use a translog instead of Cobb–Douglas production function. The labor aggregate ($\ln \tilde{L}$) is unchanged, but quadratic and interaction terms with capital are included. Only in Kenya are the second order terms jointly significant. It is no surprise then that the results go through relatively unchanged. As in Hellerstein and Neumark (2007) or Fox and Smeets (forthcoming), coefficients on worker characteristics in the production function are also barely affected when the estimation method is changed from least squares to the Olley–Pakes semi-parametric estimator.

In the next set of regressions, we included additional controls in both equations to control for heterogeneity across firms: state and foreign ownership dummies to capture some firm heterogeneity, the fraction of unionized workers and family members, which could influence firms' remuneration practices, and (log) capital in the wage equation. Even though these additional controls tend to be jointly significant in most equations, the main findings go through unchanged. The most notable difference is that schooling now attracts a higher return in the labor market than its productivity effect in Kenya and Zimbabwe, but the gaps still decrease with the level of development, respectively 4.4%, 3.0%, and 1.1% per year. [15]

Because not all workers are observed, we implement two approaches to control for the fact that the averages of the employee characteristics have to be estimated for each firm. Generalizing the approach in Hellerstein and Neumark (1999) to continuous variables, we create new samples of workers by sampling with replacement from the universe of workers implied by the observed workers. Alternatively, we create a sample with randomly generated employee-characteristics and construct weights for each firm using Bayes' law based on the observed sample of workers. Details are in Van Biesebroeck (2003). The average *p*-values for estimates on 100 such samples, reported in lines (e) and (f) of Table 3, suggest that the results are robust to sampling error.

The next two estimations are aimed at addressing potential measurement error. The capital coefficient is estimated the lowest in Tanzania, as would be the case if measurement errors are largest there. As the capital stock is likely to be correlated with human capital levels, the bias would also affect the coefficients of interest. However, running the regressions for the three countries as a system, enforcing uniform capital, and labor coefficients, had little impact on the estimated skill premiums or on the *p*-values for the joint tests. Ignoring non-wage compensation, as we have done so far, is likely to be more important in the least developed countries, such as Tanzania. Approximately 60% of the firms report whether they make

Table 3. *Robustness checks: p-values on the joint test for experience and schooling*

	Tanzania	Kenya	Zimbabwe
(a) Schooling and experience defined discretely	0.00	0.06	0.83
(b) Diminishing return to schooling and experience	0.05	0.17	0.61
(c) Translog production function	0.01	0.04	0.84
(d) Additional controls in both equations	0.01	0.08	0.81
(e) Controlling for sampling error I	0.00	0.06	0.60
(f) Controlling for sampling error II	0.00	0.06	0.13
(g) Identical input elasticities in all three countries	0.00	0.10	0.81
(h) Measurement error in wages: payment in-kind	0.01	0.08	0.79
(i) Single year of data	0.02	0.16	0.86

Notes: Estimation is always as in Table 2:
(a) Schooling and experience are defined as dummy variables: high or low; (b) squared terms for schooling and experience are included and a second order approximation is used; (c) second order terms–$\log^2(L)$, $\log^2(K)$, $\log(L) * \log(K)$–are added to the production function; (d) additional controls–foreign and state ownership dummies, the fraction of the workforce that is unionized or a family member, and capital-added to both equations; (e) average *p*-values for estimates on 100 samples generated by subsampling from the universe of workers implied by the observed sample of workers, extends the approach in Hellerstein *et al.* (1999) to continuous variables; (f) average *p*-values for estimates on 100 samples with randomly generated characteristics for each firm and weights constructed using Bayes' law based on the observed sample; (g) identical capital and labor coefficients in the production function of the three countries; (h) adding payments in-kind to the wage variable (sample size is reduced); (i) limit the sample to only the first year of data for each country.

payments in-kind. In some cases it is zero and in almost all instances it amounts to less than 10% of total compensation. Adding these to the dependent variable of the wage equation increases the premiums associated with schooling and experience slightly in Tanzania, with little change for Kenya and Zimbabwe. This exacerbates the excess wage premium for Tanzania, relative to the productivity premium.

Finally, limiting the sample to only the first year of data for each country has limited impact on the results. As expected, most standard errors are increased, but given that a similar number of observations are now used for each country, it is reassuring to find that the standard errors are now extremely similar. The *p*-values for the tests all rise, but Tanzania (0.02) and Zimbabwe (0.86) are still at opposite extremes.

Unfortunately, only three countries could be included in the analysis. A partial analysis was possible with data from Cameroon (almost as developed as Zimbabwe) and Burundi (even less developed than Tanzania). The sample sizes are smaller, some variables (e.g., capital) are measured less accurately, and workers cannot be weighted by the importance of their occupation category in Burundi. Results for these countries fall in between the extremes of Tanzania and Zimbabwe. Reassuringly, the failure of the equalities to hold is more pronounced for Burundi than Cameroon: the *p*-values for the joint test for schooling and experience were, respectively, 0.03 and 0.35.

(c) Two alternative specifications

Following the many robustness checks, we now turn to two alternative specifications that change the estimation strategy and the assumptions in the model more radically.

Following the estimation procedure in Van Biesebroeck (2007), it is possible to estimate an individual-level wage equation jointly with a plan-level production function. This is likely to increase precision and make the equality tests more powerful. It does require some additional assumptions on the variance–covariance structure of the errors. In addition to the correlation between the error terms in the two equations, we allow for a random firm effect in the wage equation, as several observations now share a single employer. The transformation required to implement the feasible generalized least squares estimator also requires a correction for the fact that our panel is unbalanced (Wooldridge, 2000, p. 450) which relies on exit from the sample being a random event. Estimation is in two steps and the results are reported in Table 4.

The estimated returns to experience in the wage equation rise by approximately 0.5% in all three countries, while the returns to experience in the production function barely change. They fall by 0.1–0.2% in Tanzania and Kenya and increase by a comparable amount in Zimbabwe, but changes are not statistically significant. As the standard errors shrink by a factor of three in the wage equation and are almost cut in half in the production function, the test of equal returns to experience is now rejected a lot more resoundingly in both Tanzania and Kenya.

In contrast, the point estimates on the returns to schooling in both equations are estimated closer together for Tanzania and Kenya. A year of schooling still attracts a higher wage premium than its productivity effect would warrant, but the gaps shrinks from 4.2% per year of education to 2.2% in Tanzania, and from 4.1% to 1.8% in Kenya. The *p*-values for the equality tests show that we cannot reject equal returns anymore. They have not risen more, in Tanzania there was not even a chance, because the coefficients are now estimated a lot more precisely, especially in the wage equation.

The difference between the salary reward for experience and its productivity effect is now estimated so significantly in those two countries that the joint test for equality for experience and schooling is also rejected strongly. It is worth noting that the standard errors for Zimbabwe are on average smaller than for Tanzania and scarcely higher than for Kenya. The different conclusions are clearly not the result of imprecise estimates.

An important maintained assumption thus far is that all types of workers are perfect substitutes. We now relax this for male and female workers as they are often employed in dif-

Table 4. *Joint estimation of the wage equation at the individual level and the production function at the firm level*

	Tanzania		Kenya		Zimbabwe	
	Wage	Output	Wage	Output	Wage	Output
Labor		0.841		0.783		0.872
		(.035)***		(.023)***		(.028)***
Capital		0.213		0.337		0.266
		(.017)***		(.015)***		(.017)***
Male	0.105	0.610	−0.035	0.224	0.042	0.153
	(.047)**	(.338)*	(.035)	(.118)**	(0.068)	(0.143)
Experience	0.023	−0.029	0.018	−0.008	0.012	0.014
	(.002)***	(.007)***	(.001)***	(.005)	(.002)***	(.005)***
Schooling	0.061	0.039	0.053	0.038	0.045	0.056
	(.004)***	(.016)**	(.004)***	(.015)***	(.007)***	(.014)***
Observations	1345	316	3209	544	1157	210
R^2	0.77	0.23	0.44	0.42	0.26	0.58
Test for equality of coefficients (p-value)						
Experience ($\lambda_X - \varphi_X$)	0.00		0.00		0.61	
Schooling ($\lambda_S - \varphi_S$)	0.16		0.32		0.52	
Joint test	0.00		0.00		0.79	

Notes: Joint estimation (SUR) of the wage equation at the individual level and the production function at the firm level. The sample for Tanzania covers 3 years, 1992–94, for Kenya also three years, 1991–94, and for Zimbabwe 2 years, 1991–92. Controls in both equations include hours worked and year, industry, and location fixed effects. The two-step estimation procedure uses a feasible generalized least squares transformation, drawing on Wooldridge (2000) and explained in some detail in Van Biesebroeck (2007), and includes a random firm effect in the wage equation.

* Significance at 10% level.

** Significance at 5% level.

*** Significance at 1% level.

ferent occupation categories. At the same time, we allowed wage and productivity premiums for experience and schooling also to vary for male and female workers.

We tried modeling the labor aggregate using the C.E.S. functional form for the two groups of workers, but the elasticity of substitution converged to unity for Kenya and Zimbabwe. For consistency, we kept the unitary elasticity assumption from the Cobb–Douglas functional from for all three input factors—capital and two worker categories—in each country and estimated three input coefficients. In the results in Table 5 it is further enforced that for Tanzania the ratio of the input elasticity of female to male workers equals 4.7, the average for the other two countries. Without this restriction, the input elasticity of female workers would be estimated negative. [16]

The first order approximation for the male and female labor aggregates is calculated separately, leading to the following production function:

$$\ln Q = \ln A + \alpha_K \ln K + \alpha_{LM}(\ln L_M + \varphi_{SM}\overline{S_M} + \varphi_{XM}\overline{X_M}) + \alpha_{LF}(\ln L_F + \varphi_{SF}\overline{S_F} + \varphi_{XF}\overline{X_F}) + \epsilon_q.$$

The wage equation now similarly includes separate measures of male and female average levels of experience and schooling, both multiplied with the respective employment shares for each gender.

The schooling coefficients for both genders always bracket the average premiums obtained in Table 2 that lump male and female workers together. While this makes intuitive sense, nothing ensured it came out this way. The average education level by firm is not even contained within the separate averages for male and female workers, which is made possible by a varying proportion of male workers across firms. The relative effect of one extra year of schooling on productivity is estimated to be higher for female workers, which is also intuitive as their average schooling level is a lot lower and diminishing returns should be less important.

Comparing schooling premiums across equations reveals that the gaps are a lot lower for male than for female workers. In order, they stand at 3.2%, 0.1%, and 0.5%. Especially for Kenya this is a large reduction from the 4.1% difference in Table 2 and even for Tanzania the difference is not statistically significant anymore. In contrast, the gaps for female workers are a lot larger than in Table 2 for all three countries. Especially for Tanzania, where a year of schooling boosts salaries less for female than for male workers, equality can be strongly rejected. In Kenya, the salary premium associated with education for female workers is a lot higher than for males, but not quite as high as the productivity benefit. The latter is estimated very imprecisely, making all differences not statistically significant.

For experience, the gaps are also larger for female than for male workers for all three countries, but even for male workers the differences in the point estimates are not always small. The gap shrinks from 4.3% for Tanzania in Table 2 to 3.2% and from 2.2% to 0.9% for Kenya. While clearly narrowing, compared to wage premiums of 1.8% and 1.4% these are not negligible. For Tanzania we can still reject equality, for Kenya not, partly because standard errors also rise slightly. For Zimbabwe, the difference disappears entirely, dropping from 0.5% to 0.1% per year of experience.

The difference in salary reward and productivity effect of a year of experience for female workers is very far apart for all countries. For Kenya, the gap is not statistically significant

Table 5. *Joint estimation of wage and production equations with imperfect substitution between male and female workers*

	Tanzania		Kenya		Zimbabwe	
	Wage	Output	Wage	Output	Wage	Output
Male workers		0.639		0.782		0.684
		(.062)***		(.084)***		(.099)***
Female workers		0.136		0.078		0.232
		–		(.081)		(.082)***
Capital		0.232		0.306		0.249
		(.036)***		(.033)***		(.042)***
Male	−0.321		0.619		0.525	
	(.325)		(.442)		(.791)	
Experience (male)	0.018	−0.014	0.014	0.005	0.008	0.009
	(.005)***	(.015)	(.004)***	(0.011)	(.007)	(.011)
Experience (female)	0.020	−0.058	0.006	0.118	−0.014	0.069
	(.013)	(.099)	(.011)	(.157)	(.014)	(.032)**
Schooling (male)	0.067	0.035	0.047	0.046	0.055	0.060
	(.013)***	(.031)	(.011)***	(.022)**	(.022)***	(.027)**
Schooling (female)	0.025	0.328	0.108	0.368	0.138	0.079
	(.018)	(.124)***	(.033)***	(.352)	(.054)***	(.057)
Observations	316	316	544	544	210	210
R^2	0.29	0.23	0.42	0.41	0.40	0.53
Test for equality of coefficients (p-value)						
Experience (male and female)		0.08	0.60		0.04	
Schooling (male and female)		0.02	0.76		0.71	
Experience and schooling (only male)		0.06	0.72		0.99	
Experience and schooling (only female)		0.04	0.69		0.04	

Notes: Estimation as in Table 2, but separating male and female workers in two labor aggregates; see Section 4(c) for further details. The input elasticity of female workers in the production function for Tanzania is fixed at 1/4.7 of the male worker coefficient to avoid negative estimates. Robust-White standard errors are in parentheses.
* Significance at 10% level.
** Significance at 5% level.
*** Significance at 1% level.

as the standard error is again very high. This was already the case for schooling and not surprising as female workers only make up 10% of the workforce at Kenyan firms. Even for Zimbabwe the gap is almost as high as for Tanzania and both are significantly different from zero at the 5% level.

The joint tests at the bottom of Table 5 summarize the findings. For Tanzania, equality can always be rejected and most forcefully for schooling and returns for female workers. For Kenya, no difference is significantly different anymore. For the characteristics for male workers this is mostly due to a narrowing of the gaps, for female worker characteristics, the differences are huge, but estimated extremely imprecisely. For Tanzania, gaps are negligible for characteristics of male workers, but large and significantly different from zero for female worker characteristics, especially experience.

(d) Extensions

Finally, we take a look at labor markets defined by city rather than country, and we introduce two additional more firm-specific human capital characteristics.

One feature of labor markets in developing countries that might help explain the failure of arbitrage to eliminate the gap in returns is geographic segregation of economic activities. Reardon (1997) surveys evidence suggesting that localized labor markets are important in Africa. If workers rarely migrate between cities and poor transportation infrastructure limits daily commuting, firms operating in different areas should not be pooled. If regions differ in the relative supply of or demand for human capital characteristics, the estimated premiums are only averages and might not represent the trade-off for any particular firm. Location dummies, included in all previous regressions, might not suffice.

The sample size allows estimation by city, but only in the one or two largest cities should we expect to find any significant coefficients. In Tanzania, the main center of manufacturing activity is Arusha near the border with Kenya. Nairobi is one of the most important manufacturing centers of East Africa. In Zimbabwe, manufacturing activity is less concentrated than in the other countries. Still, 42% of the firms in the sample are located in the capital, Harare. To give some sense of the patterns we illustrate the results graphically in Figure 1, ordering cities by increased level of development within each country. The main manufacturing center for each country is always listed second and the runner-up first.

In all three countries, but especially in Tanzania and Kenya, the gap between the wage and productivity effect of experience narrows substantially from the first to the second listed city. In Dar Es Salaam (Tanzania) and Bulowayo (Zimbabwe), we can reject equality for the experience premiums. While the gaps are the largest of all in Mombasa (Kenya), there are two few observations to obtain precision and we cannot reject equality. In Kenya, almost 65% of all firms in the sample are located in Nairobi, and the gap for experience is narrower than anywhere else.

The pattern is broadly similar for the schooling premiums, but the high p-values for the equality tests do not allow any unambiguous rejections. At least in terms of point estimates, we see narrower gaps in the main manufacturing center for the country.

Comparing the results in Figure 1 with those in Table 2, two things have changed. On the one hand, limiting the sample to firms in a single city facilitates arbitrage and the distorting effect of isolated labor markets should disappear. On the other hand, the number of observation and estimation precision declines, which makes it harder to statistically reject any test. In

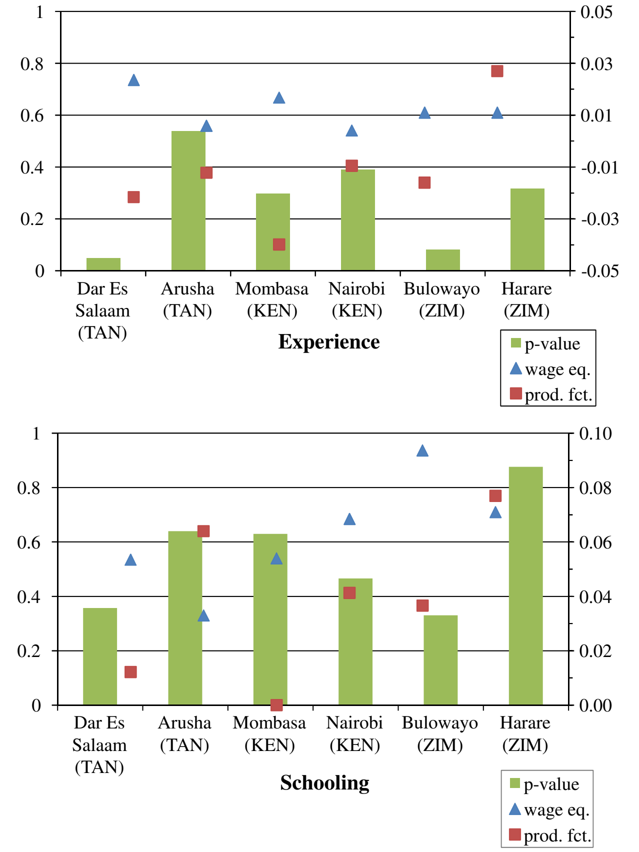

Figure 1. *Results by city. Note: Similar regressions as in Table 2, but separately by city. City-order within each country is by increasing level of development. Coefficient estimates on the years of experience and schooling in both regressions are indicated on the right axis. The p-value of the test for a zero difference in the wage and production function coefficient is on the left axis.*

addition, a manufacturing center such as Nairobi is a lot richer and much more developed than the rest of the country. The U.N. estimates that Nairobi alone generates 45% of Kenyan GDP. The fact that equality of returns cannot be rejected anymore for Nairobi, while it could for Kenya as a whole could be due to localized labor markets or to a more developed labor market in Nairobi.

In the final set of results, depicted in Figure 2, we include two additional human capital characteristics in both equations: tenure and training. Given that tenure at the current employer is correlated positively with general labor market experience, the estimates on experience also change somewhat from Table 2, but the general picture is still that equality of wage and productivity premiums can be rejected in Tanzania and Kenya, but not in Zimbabwe. In comparison, for tenure, equality cannot be rejected in any country. The gaps in premiums are reduced, especially for Tanzania, and the *p*-values for the formal test grow commensurately. [17]

To facilitate comparison with the discrete training variable, the firm-level measure is the fraction of workers that received training; we also changed the definition of schooling. The measure is now the fraction of workers that finished at least high school. The *p*-value for equality of the wage and productivity premiums is still lowest in Kenya, but is now a lot higher for Tanzania.

Training is associated with large productivity effects in all three countries, estimates range from 26% (Zimbabwe) to

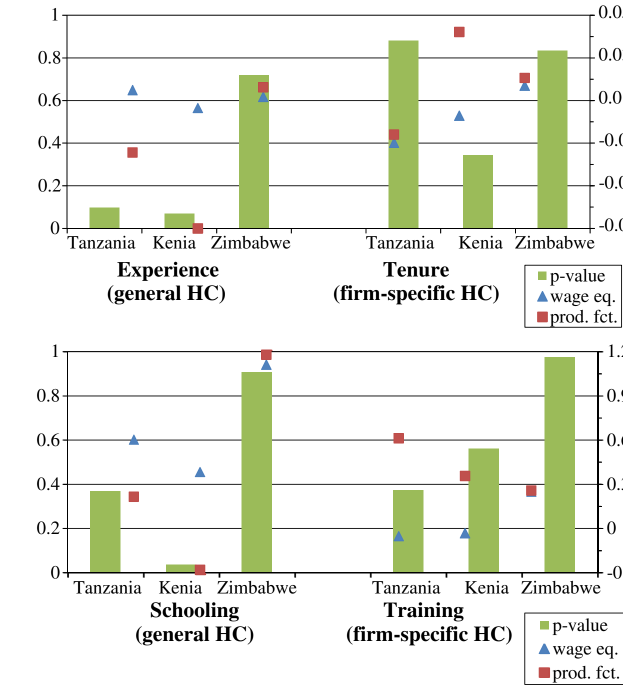

Figure 2. *Results with four human capital characteristics. Note: Similar regressions as in Table 2, with two additional human capital characteristics: years of tenure at current employer and fraction of workers that received formal training. For comparability with the coefficients on training, schooling is now measured as fraction of workers that finished at least high school.*

61% (Tanzania), but the standard errors are high. The effect combines human capital accumulation and selection, as firms can offer training selectively or disproportionately choose to hire or retain workers that received training. These trained workers do not receive any salary boost in Tanzania or Kenya, even though the point estimate for their productivity contribution is higher than in Zimbabwe. Even though the point estimates are far apart, especially for Tanzania, equality of the returns to training can never be rejected.

Part of the problem is the low incidence of training, 7–8% in the two poorest countries, which leads to imprecise estimates. The strong productivity effects warrant some attention though. In the vast quantity of research summarized in the latest *Handbook of Development Economics* and *Handbook of Labor Economics* there is only a short discussion of on-the-job training in the chapter by Gibbons and Waldman (1999). This is in sharp contrast to the many chapters evaluating the effects of formal education. Training seems a topic worthy of further research, especially in a developing country context.

A final pattern that is worth pointing out is the reversal of the wage and productivity estimates in the two countries where equality is least supported. For both general human capital characteristics on the left, experience and schooling, wage premiums exceed the productivity premiums. For both firm-specific measures on the right, tenure and training, the productivity premiums exceed the wage premiums. The pattern leading toward rejection of the equalities is in both cases and both countries for the general human capital characteristics to achieve a higher return in the labor market than in production and the reverse for firm-specific characteristics. Such a compensation pattern will help reduce worker turnover, especially of those valuable employees that received training, which is borne out by a positive correlation between training and tenure at the individual level.

Separate tests grouping the more firm-specific aspects of human capital—tenure and training—and the more general attributes—experience and schooling—clearly identifies the general characteristics as the cause of rejection in the lesser developed countries. The *p*-values are 65% and 55% for a joint test on the first set of variables in Tanzania and Kenya and 17% and 4% for the second set. As a result, we would expect workers to underinvest in firm-specific skills. In Zimbabwe all premiums are estimated remarkably close to one another and all *p*-values are extremely high.

5. DISCUSSION AND CONCLUSIONS

The results indicate that the failure of wage and productivity premiums to equalize is particularly pronounced in the poorest economy and sometimes appears in the country with intermediate level of development too. In Zimbabwe, which the World Bank (2000) attributes a manufacturing labor productivity four times higher than in Kenya and seven times the level of Tanzania, equality can almost never be rejected. This general pattern is robust to many specification checks.

What can we make of it? It is beyond this paper to look at the causality of this relationship. Badly functioning labor markets could impede economic development, but it is equally likely that low levels of development make it harder for labor markets to function efficiently. One should also be careful not to generalize broadly from the experience of just three countries.

The alternative specifications provided some additional evidence that can shed light on the failure of premiums to equalize in the benchmark model. Using individual-level information on wages allowed the incorporation of firm-specific random effect and a more flexible control of unobserved heterogeneity. This reduced the gaps in wage and productivity premiums associated with schooling. Some of the pattern at the firm level could be due to the matching of more educated workers to better firms and rationalize the high remuneration of schooling.

Allowing imperfect substitution between male and female workers and different returns to schooling and experience, also eliminated some of the discrepancies, most notably for schooling of male workers. In general, the wage premiums deviated a lot more from productivity premiums for female workers.

As labor markets do not operate as spot markets under perfect information, it is possible that human capital characteristics influence worker-firm matching rates and the relative power in the bargaining over wages. Imagine, for example, that workers are randomly matched with firms and bargain over the surplus of the match. Firms will make wage offers that lie between the worker's outside alternative, which is potentially very low and the worker's productivity level. If a worker's bargaining position improves with schooling or experience, we would expect the wage return to be more elastic to these human capital characteristics than the productivity return.

Recall that all effects are estimated relative to the benchmark worker, a young, uneducated woman. The first order condition for cost minimization (in logarithms) for the benchmark worker is $\lambda_0 = \ln \alpha_L + \ln(Q/L)$, which characterizes the relationship between the constant term in the wage equation and the labor input coefficient in the production function. For Tanzania, we could reject at a 1% significance level that this equation held at the sample mean. It indicates that the benchmark worker is paid below her productivity level. Excess wage returns for schooling and experience merely bring salaries closer to the productivity level and are affordable for firms.

It does offer a way to rationalize the observed wage-productivity gaps while maintaining cost minimizing behavior. It does leave the explanation for the negative relationship between low bargaining power for unskilled workers and development level for further study. It is not implausible that in a poor agriculture-dominated economy the outside options for unskilled manufacturing workers are especially bad.

The estimation results by city suggest an additional explanation for the failure of premiums to equalize at the country level. If the relevant labor market is very localized and arbitrage is only relevant at a sub-national level, estimates that pool firms from different regions only recover weighted averages of the true (local) premiums. The loss of precision when the model is estimated separately by city makes it nearly impossible to reject equality. Especially for schooling no unambiguous conclusions are possible, even though the point estimates on the gaps seem even wider at the city level than at the national level.

At least for experience in Tanzania, we can still reject equality in the less developed city Dar Es Salaam, while much of the difference vanishes for Arusha. The difference in development level seems at least as important as the local labor market explanation. A similar conclusion for Kenya is not possible as there are two few observations in the poorer city, Mombasa, but the point estimates at least point in the same direction.

Finally, a crucial aspect of remuneration practices is the trade-off between paying workers for general experience *versus* firm-specific tenure. This mirrors a similar trade-off between general pre-employment education and training programs for employees. General skills (experience and schooling) are rewarded relatively higher than firm-specific skills (tenure and training) in Kenya, even though the latter are associated with larger productivity gains. In Tanzania, the same holds for the experience–tenure combination. In Zimbabwe, all wage premiums match the productivity gains that are associated with them, and the returns to firm-specific investments are higher than in the other countries. A richer model of human capital accumulation and remuneration is needed to understand these relationships better.

A final caveat is in order. We have focused on observable differences between workers, motivated in part by the observation in Abowd, Kramarz, and Margolis (1999) that 90% of inter-industry wage differentials are explained by person effects—including unobservables. A large literature uses search and matching models of the labor market, that leave a lot of scope for unobservables, including on the firm side, in explaining the association of firms and employees (Mortensen & Pissarides 1999). If matching is positive assortative, the low remuneration of highly skilled workers in our data can reflect an equilibrium situation where good workers are employed at firms with high levels of unobservable assets. Even though the empirical literature on assortative matching using linked employer–employee data tends to find a negative correlation between unobserved worker quality and unobserved firm quality, this reverse pattern might be due to estimation problems, as argued by Andrews, Gill, Schank, and Upward (2008). Unfortunately, it is beyond this paper and our data to investigating this explanation explicitly.

NOTES

1. A conference symposium in the *Monthly Labor Review* (July 1998) and the book by Haltiwanger, Lane, Spletzer, Theeuwes, and Troske (1999) provide overviews.

2. Attaining "some college" is also estimated to bring higher productivity than wage benefits in their benchmark results (significantly different at an 11% significance level), but this finding was not confirmed in their robustness checks.

3. No details are given on the assumptions on the variance–covariance matrix when the individual and firm data is combined. Van Biesebroeck (2007) outlines one possible set of assumptions and finds results for discretely measured characteristics in line with those in this paper.

4. Many differences are large in absolute value—five of the eight estimated differentials exceed 20%—but the direction of the difference varies by schooling level.

5. The consistent aggregation method is especially important when diminishing returns are incorporated, see Van Biesebroeck (2003). When squared terms on experience or schooling are included in the individual-level Mincer wage equation, one needs to include the average within-firm variance of experience or schooling in the firm-level wage regression.

6. Note that the productivity effect of training could be due to selection or to human capital building.

7. Firm and time subscripts are omitted.

8. Concerns over potential bias introduced by unobserved worker ability in the wage equation or unobserved productivity in the production function should be alleviated by joint estimation as the bias works in the same direction in both equations. The unobservables are to a large extent two sides of the same coin, which Frazer (2001) exploits to control for unobserved ability in the wage equation with the productivity residual. We are only interested in the relative magnitudes of the coefficients in the two equations.

9. In this specification the gender premium is introduced just as the education or experience premiums, but given that the M_i characteristic is a dummy variable, we could define the premium as in (1), by defining (1), by defining $\varphi'_M \equiv \exp(\varphi_M)$.

10. In the working paper version (Van Biesebroeck, 2003), we present results using both continuous and discrete characteristics. Only in the latter case do we rely on the constant proportions assumption.

11. Full details on the derivation are in Appendix B of Van Biesebroeck (2003).

12. The three occupation categories combine workers designed as (1) production, maintenance, supervisory, instructors; (2) clerical, administration, management, sales; (3) apprentices and trainees. If, for example, the first category employed 60% of all workers, six of the ten workers interviewed would be selected from this category.

13. Results in Van Biesebroeck (2007) illustrate that additionally allowing for arbitrary correlation of errors within a firm over time hardly increases the standard errors.

14. Full coefficient estimates for all robustness checks are available upon request; several are reported in the working paper version, Van Biesebroeck (2003), although those results do not weigh workers properly to construct the firm-level human capital characteristics.

15. The firm-level controls have the expected signs, which are invariably the same in both equations: negative for state and positive for foreign ownership. Tellingly, family members receive higher salaries in Tanzania and Kenya, even though firms that employ a high fraction of family have lower productivity.

16. The effective input share for female workers using the C.E.S. specification for the production function was estimated positive, and the ratio with the male share was approximately 1 to 7.

17. The failure to reject equality for Kenya, even though the point estimates are far apart, is the result of high standard errors. This is partly caused by the multicollinearity between tenure and experience.

REFERENCES

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica, 67*(2), 251–333.
Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias?. *Journal of the Royal Statistical Society: Series A, 171*(3), 673–697.
Appleton, S., Hoddinott, J., & Mackinnon, J. (1996). Education and health in sub-Saharan Africa. *Journal of International Development, 8*(3), 307–339.
Bigsten, A. *et al.* (2000). Rates of return on physical and human capital in Africa's manufacturing sector. *Economic Development and Cultural Change, 48*(4), 801–827.
Brown, J. N. (1989). Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms. *American Economic Review, 79*(5), 971–991.
Crepon, B., Deniau, N., & Pérez-Duarte, S. (2003). Productivité et salaire des travailleurs agés. *Revue Francaise d'Economie, 18*(1), 157–185.
Dearden, L., Reed, H., & Van Reenen, J. (2006). The impact of training on productivity and wages: Evidence from British panel data. *Oxford Bulletin of Economics and Statistics, 68*(4), 397–421.
Fox, J. & Smeets V. (forthcoming). Does input quality drive measured differences in firm productivity? *International Economic Review*.
Frazer, G. (2001). *Linking firms and workers: Heterogeneous labor and returns to education*. Mimeo: Yale University.
Gibbons, R., & Waldman, M. (1999). *Careers in organizations: Theory and evidence*. In O. C. Ashenfelter, & D. Card (Eds.). *Handbook of labor economics* (Vol. 3B, pp. 2373–2437). Amsterdam: North Holland.
Haegeland, T., & Klette, T. (1999). Do higher wages reflect higher productivity? Education, gender and experience premiums in a matched plant-worker data set. In J. Haltiwanger, J. Lane, J. Spletzer, J. Theeuwes, & K. R. Troske (Eds.), *The creation and analysis of employer–employee matched data* (pp. 231–259). Amsterdam: North Holland.
Haltiwanger, J., Lane, J., Spletzer, J., Theeuwes, J., & Troske, K. R. (1999). *The creation and analysis of employer–employee matched data*. Amsterdam: North Holland.
Hellerstein, J. K., & Neumark, D. (1999). Sex, wages, and productivity: An empirical analysis of Israeli firm-level data. *International Economic Review, 40*(1), 95–123.
Hellerstein, J. K., & Neumark, D. (2007). Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer–employee data set. In E. R. Berndt, & C. M. Hulten (Eds.), *Hard to measure goods and services: Essays in honor of Zvi Griliches*. Chicago: University of Chicago Press.
Hellerstein, J. K., Neumark, D., & Troske, K. R. (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labor Economics, 17*(3), 409–446.
Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics, 124*(4), 1403–1448.
Jones, P. (2001). Are educated workers really more productive?. *Journal of Development Economics, 64*(1), 57–79.
Jorgenson, D. W., & Griliches, Z. (1967). The explanation of productivity change. *Review of Economic Studies, 34*(99), 349–383.
Mincer, J. (1974). *Schooling, experience, and earnings*. New York: Colombia University Press.
Mortensen, D. T., & Pissarides, C. A. (1999). New developments in models of search in the labor market. In O. C. Ashenfelter, & D. Card (Eds.). *Handbook of labor economics* (Vol. 3B, pp. 2567–2627). Amsterdam: North Holland.
Reardon, T. (1997). Using evidence of household income diversification to inform study of the rural nonfarm labor market in Africa. *World Development, 25*(5), 735–747.
Rosenzweig, M. R. (1988). *Labor markets in low-income countries*. In H. Chenery, & T. N. Srinivasan (Eds.). *Handbook of development economics* (Vol. 1). Amsterdam: North Holland.
Topel, R. H. (1991). Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of Political Economy, 99*(1), 145–176.
Van Biesebroeck, J. (2003). Wages equal productivity. Fact of fiction?" NBER Working Paper No. 10174. Boston.
Van Biesebroeck, J. (2005). Firm size matters: Growth and productivity growth in African manufacturing. *Economic Development and Cultural Change, 53*(3), 545–584.
Van Biesebroeck, J. (2007). Wage and productivity premiums in sub-Saharan Africa. In S. Bender, J. Lane, K. Shaw, F. Andersson, & T. von Wachter (Eds.), *The analysis of firms and employees: Quantitative and qualitative approaches* (pp. 345–371). Chicago: University of Chicago Press.
Wooldridge, J. M. (2000). *Econometric Analysis of Cross Section and Panel Data*. Boston: South-Western.
World Bank (2000). *African Development Indicators*. Computer file.

APPENDIX A. DATA

(a) *Countries*

The three countries included in the sample—Tanzania, Kenya, and Zimbabwe—are middle-sized former British colonies in East Africa that differed substantially in level of development. Their GDP per capita (in PPP in 1992) ranged from \$395 in Tanzania, less than half the \$1089 of Kenya, to a level almost six times as high in Zimbabwe (GDP per capita of \$2459). The differences are smaller comparing the U.N. human development index, which also takes education and life expectancy into account, but the ordering is the same. In the ranking for 1992, Tanzania occupies the 148th place with 0.306, putting it in the low development category. Kenya and Zimbabwe rank closely at places 125 and 121, with scores of 0.434 and 0.474, near the bottom of the medium development group.

The relative development levels of the countries are mirrored in the share of workers employed in industry (manufacturing employment was not available for Tanzania). Only 4.7% of employment in Tanzania is in industry, while it is almost twice as high in Zimbabwe (8.6%) and intermediate in Kenya (7.3%). In Tanzania, the transition from agriculture to other sectors had only just begun; agriculture comprised almost half the workforce at the end of the 1990s. In Kenya, the transformation was in full swing; the employment share of agriculture declined from 42% in 1975 to 27.5% by the sample period. Zimbabwe, in contrast, has seen a stable 18.5% of its workforce employed in agriculture for the last 25 years.

The difference in labor productivity in industry is even more stark. While industry workers in Kenya produce twice as much as Tanzanian workers, Zimbabwe's output per worker outstrips Tanzania by a factor of one to seven and Kenya one to four. World Bank (2000) statistics further indicate that manufacturing workers in Tanzania earn on average 3.5 times more than agricultural workers, while the ratio stands at 5.7 in Kenya and even at 9.9 in Zimbabwe.

Infrastructure statistics, from the World Bank Development Report, are in line with the relative levels of development. Zimbabwe has 22 km of paved highways per 1000 km^2 of land, while the corresponding numbers for Kenya and Tanzania are 15 km and 4 km. The ranking is preserved in kilometers of railroad by area, at respectively 8, 5, and 4 km, or airports per million inhabitants, 1.4 in Zimbabwe, 0.6 in Kenya, and 0.3 in Tanzania. In fact, almost any conceivable statistic that one expects to be correlated with development leads to the same ranking: access to clean water, telephone penetration, school enrollment, infant mortality, etc. Only life expectancy at birth gives a reverse ranking, due to the staggering HIV infection rate, affecting one third of the adult population in Zimbabwe and almost one sixth in Kenya.

(b) *Firms*

A sample of manufacturing firms, surveyed in three consecutive years between 1992 and 1995, provides the micro data used in the analysis. The data was collected by three different research teams coordinated by the Regional Program of Enterprise Development at the World Bank. The firm-level data for the three countries is available online at the site of the Centre for the Study of African Economics at the University of Oxford. That site also lists data appendix with information on the survey, the sampling frame, and the variable construction: http://www.csae.ox.ac.uk/datasets/main.html.

Firms were sampled to give the firm of each manufacturing worker equal probability of being included in the sample—an implicit stratification by employment size. Approximately 200 firms were surveyed each year in each country, covering four broadly defined manufacturing sectors: food processing, textile and clothing, wood and furniture, and metal and equipment. The resulting sample is an unbalanced panel of firms with, on average, 110–191 observations per year in each country.

A large fraction of the manufacturing sector is covered by this sample. The value added produced by the sample firms amounts to 31% of official manufacturing GDP in Tanzania, 17% in Kenya, and 26% in Zimbabwe. The share of manufacturing workers that are employed by firms included in the sample is substantially lower in the first two countries, as large firms tend to have higher labor productivity (Van Biesebroeck, 2005). The stratified sampling yielded significantly larger than average firms. The absence of reliable firm censuses in these countries makes it impossible to compare precisely with the universe, but the average firm size in these countries is surely smaller than in the US manufacturing sector, where firms employed on average 61 workers in 1993 (Van Biesebroeck, 2005).

The differences in level of development between the countries are equally apparent when we compare the firms in the sample. The median firm in Tanzania achieved only 38% of the labor productivity of the median firm in Kenya, while labor productivity in Zimbabwe was 42% higher than in Kenya (calculations are detailed in Van Biesebroeck (2005)). Total factor productivity numbers show similar differences: the median firm in Kenya is twice as productive as in Tanzania, but achieves only two thirds of the productivity level of the median firm in Zimbabwe. The salary differences between the countries match the labor productivity differences rather well. Workers in Tanzania earn 27.4% of the average salary in Zimbabwe, while the labor productivity at their employers stands at 26.8%. Salaries in Kenya, on average \$120, are slightly below what one would predict from the relative labor productivity, which would imply a salary of approximately \$140. These comparisons confirm that Zimbabwe is by far the most developed country of the three, while Tanzania is lagging far behind.

(c) *Workers*

Each year, a maximum of ten employees per firm were interviewed as well. In Zimbabwe, workers were only interviewed in the first and second year, while 3 years of employee data are available for Kenya and Tanzania. In the first year of interviews, the principle instructions regarding the sampling frame for the interviewers in Kenya and Zimbabwe are as follows:

> You will be interviewing 10 workers from each firm. The sample selection should work as follows:
> Based on the information from Question 2 of Section A in the Labor Market questionnaire previously, calculate the share of each of the following groups of workers in the firm's total labor force:
> a. Production, maintenance, supervisors, masters.
> b. Clerical, admin/mgmt., sales.
> c. Apprentices.
> Use these calculated shares to allocate the 10 workers to be sampled across groups, with one worker for each 10% of the firm's labor force.

From this we can straightforwardly construct appropriately weighted average worker characteristics, e.g., experience, schooling, that are representative for the firm.

The instructions in the questionnaire for Tanzania, however, suggest a different sampling frame. They are as follows:

> We suggest you interview 20 workers from each firm (or as many as are available).
> Please observe the following guidelines in selecting workers to interview:
> Interview at least one worker from each category in which the firm shows employees in Part A, Question 3 of the Labor Market questionnaire, page 18. These categories are:

01 Management	07 Other production workers
02 Administrative/clerical	08 Supervisors/foremen
03 Commercial/sales	09 Support Staff
04 Equipment/maintenance	10 Trainees
05 Technicians	11 Craftsman
06 Skilled production workers	12 Apprentices

These latter instructions are the same as for Ghana, the first country in which the RPED survey was administered, one year before other countries were visited. From the resulting dataset, however, there are several indications that the first-listed set of instructions was followed in all three countries. First, only for 4 of the 171 firms in Tanzania are there more than 11 workers observed in the first year; in subsequent years there are never more than 10 workers interviewed. Second, the sampling of workers in the less popular occupation categories is not higher for Tanzania than in the other two countries, while this should be the case if the latter instructions were followed. Third, a lot of workers indicate an occupation category for which the firm does not list any employees. In light of these patterns, I used the same three broad occupation categories to construct weighted averages for all characteristics also for the first year of Tanzania.

In the second and third year of interviewing, all countries used the same set of instructions. Ten workers now need to be selected from seven categories, which resulted in one or two workers being chosen in most categories in almost all cases. All questionnaires for the worker surveys with instructions for the interviewers are available online at: http://www.econ.kuleuven.be/public/N07057/Africa/.

In the first year, the sample firms employed 19,383 to 58,108 workers and 619 to 1,206 of them were interviewed. The number of workers interviewed in each firm averages around six and varies between one and 17; only for a few Tanzanian firms more than 10 workers are included. Workers are selected randomly from three occupation categories in proportion to the overall importance of each category in total employment at the firm.

Workers in Zimbabwe work on average in larger firms, are slightly older, stay longer with the same firm and are more likely to receive, or choose to enroll in, formal training once they are employed. The sample of workers in Kenya is the most dominated by males. In Tanzania, workers receive the lowest salaries, but paradoxically they have the highest years of schooling. Firm averages of the worker characteristics, as summarized in Table 1, are constructed using occupation category weights. Compared to the unweighted averages, standard deviations are much reduced, schooling and training levels are lower, but differences are smaller for tenure or experience.

Information on productivity is only available at the firm level and individual wages have to be aggregated to carry out the comparison with productivity. Identification of the wage and productivity premiums associated with worker characteristics comes from variation across firms in the composition of the workforce and average salaries or output. Employee wage regressions in the working paper version, Van Biesebroeck (2003), confirm that the aggregation does not obscure how an individual's characteristics are rewarded. A survey of the returns to education estimated from Mincer wage regressions in sub-Saharan Africa is in Appleton, Hoddinott, and Mackinnon (1996).

Individual wage regressions capture both the variation within and between firms. For example, the higher salary for male workers can be the result of men getting on average higher salaries than women within a given firm or men being disproportionately employed in firms that pay higher salaries, a between effect, even without differential pay by gender. When we separately identify the magnitude of both effects, it turns out that in almost all cases they work in the same direction. Only two variables warrant caution: the gender dummy in Tanzania and Zimbabwe and tenure in Zimbabwe.

The average male worker receives a higher salary in all three countries. In Tanzania and Zimbabwe, this is solely the result of higher wages for men within firms. The pay differential is reduced by sorting of men toward lower-paying employers. Comparing average earnings across firms will show a negative wage premium for men, because firms that employ a high proportion of men pay lower salaries on average, even though men employed in those firms still earn more than their female coworkers.

A positive coefficient on tenure can be the result of firms raising salaries for employees with high tenure or, alternatively, workers could choose to stay for a longer time with employers that offer high pay in general. Both interpretations are plausible, but only the second one is backed up by the data in Zimbabwe. The between-firm effect, which is the only part we pick up in the firm-level regressions, dominates the total.