WIDER Working Paper 2024/5

Sheepskin effects and heterogenous wage-setting behaviour

Evidence from Mozambique

David A. Jaeger[1] and Sam Jones[2]

January 2024

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: Using a unique panel survey of final-year undergraduates at six of the largest universities in Mozambique, we study the wage premium associated with completing an undergraduate degree. Conditional on a very rich set of controls, including pre-degree earnings, objective measures of ability, and academic performance, we find heterogeneity in 'sheepskin effects' across different kinds of firms. We propose a simple model of wage-setting in which productivity is only partially observable in some firms and fully observable in others. In this setting, education serves both to enhance productivity and as a productivity signal. Consistent with the theory, positions where productivity is likely to be less observable offer larger sheepskin effects.

Key words: returns to education, wage-setting, productivity, Mozambique

JEL classification: I26, J31

Acknowledgements: This study falls under the project 'Inclusive growth in Mozambique – scaling-up research and capacity', implemented in collaboration between UNU-WIDER, University of Copenhagen, University Eduardo Mondlane, and the Mozambican Ministry of Economics and Finance and financed through specific programme contributions by the governments of Denmark, Finland, and Norway. The authors thank seminar participants at the Universities of Essex, Glasgow, and Sussex as well as the 2022 Scottish Economic Society annual conference for comments.

[1] University of St Andrews, St Andrews, Scotland; CESifo, IZA, and CEPR; [2] UNU-WIDER, Maputo, Mozambique, jones@wider.unu.edu

This study has been prepared within the UNU-WIDER project Inclusive growth in Mozambique – scaling up research and capacity implemented in collaboration between UNU-WIDER, University of Copenhagen, University Eduardo Mondlane, and the Mozambican Ministry of Economy and Finance. The project is financed through specific programme contributions by the governments of Finland, Norway, and Switzerland.

ISSN 1798-7237 ISBN 978-92-9267-463-2

https://doi.org/10.35188/UNU-WIDER/2024/463-2

Typescript prepared by Mary Boss.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland and Sweden, as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

The relationship between schooling and earnings is one of the most examined in economics. The well-documented positive association between schooling and earnings is likely due to the causal role of education in enhancing productivity (Becker 1962; Mincer 1974) but also may reflect a signal of innate ability in a world where true productivity is unobservable (Spence 1973; Layard and Psacharopoulos 1974). A large number of empirical attempts to discriminate between these hypotheses have focused on diploma or 'sheepskin' effects, which seek to measure the change in earnings purely associated with receipt of a certificate of completion of some major educational level (e.g., a high school diploma) *conditional* on time spent in school. Positive diploma effects, which would be consistent with a signalling role for education, have been documented in both developed countries (e.g., Jaeger and Page 1996; Gibson 2000; Denny and Harmon 2001; Bauer et al. 2005) and developing country contexts (e.g., Mora 2003; Schady 2003; Aslam et al. 2012; Crespo and Reis 2009; Olfindo 2018).

Interpreting the results in many studies of sheepskin effects as causal is problematic. Most rely on cross-sectional variation in degree attainment, with identification relying on non-linearities or discontinuities in earnings observed at the end of specific educational levels. As Mehta and Villarreal (2008) note, positive effects on earnings at degree completion could be driven by quite different mechanisms. Completing a certificate or degree could signal a positive ability or other productive attributes like persistence or could be driven by actual above-average increases in productivity obtained in the final year of schooling, implying that skills acquisition is non-linear in time spent in school. The same cross-sectional evidence may be interpreted in favour of either main hypothesis. While a small number of studies use longitudinal data to investigate the relationship between education to earnings, these tend to follow adults who take a break from employment to upgrade their skills, raising concerns about self-selection (see Park 2011).

In this paper we exploit a unique high frequency data set in a developing country (Mozambique), where we observe individuals both during and after they complete their Bachelor's degree (*liceniatura*). For many individuals, we observe their earnings immediately both pre- and post-graduation as students often finish their education and obtain work but formally graduate at a later date. We argue that because the specific timing of graduation is largely random, estimates of the earnings premium associated with degree completion can be interpreted as reflecting a causal sheepskin effect. In addition, our data set contains extremely rich information, including measures of ability and personality traits (from psychometric tests) as well as lagged earnings and future earnings expectations, thereby allowing us to control for multiple factors that may otherwise be confounded with degree completion and its timing.

Our main result is that holding a Bachelor's degree diploma yields an immediate earnings premium of around 7 per cent. This is substantially lower than previous estimates for developing countries, which are typically four to five times larger. Our results also are robust to the inclusion of proxies for innate ability; exclusion of various sub-groups, such as individuals in self-employment or those who register changes in job type coinciding with graduation; and choice of estimator, including those that incorporate individual fixed effects. The same findings are supported by (dynamic) event study estimates. These reveal no pre-trend to earnings before graduation, but, contrary to theories that posit sheepskin effects may decline over time (e.g., see Psacharopoulos 1979), we observe large significant positive returns that emerge six months after degree completion. At the same time, there is evidence that final grades *do* matter—e.g., a one standard deviation increase in the grade maps to an expected 3 per cent increase in earnings.

We also explore heterogeneity in sheepskin effects, focusing on variation across different types of jobs. A main motivation here is that conventional textbook models of screening make the binary assumption that ability cannot be observed by firms. An important strand of literature suggests that individual productivity may be only imperfectly observable, however, allowing for employer learning over time (Farber and Gibbons 1996; Mansour 2012). Incorporating these insights, we posit a model in which productivity is differentially observable across different kinds of employers, plausibly associated with differences in the nature of tasks undertaken. The implications of the model are not only that sheepskin effects should be larger where individual productivity is difficult to observe (e.g., in the public sector) but also that, for workers with the same level of education, lower-ability individuals will generally sort into jobs that offer fixed education-based earnings premia as opposed to productivity-based earnings premia.

Our results are consistent with the model's predictions. Compared to a reference category of public sector workers, we find individuals working in the private sector score significantly higher on ability tests (administered before labour market entry), including an almost one standard deviation difference on a short-form Raven's test. Wages in the public sector also show significantly lower dispersion, being the only type of employer associated with a significant sheepskin effect in the data. This suggests that, perhaps regardless of the respective contributions of education or innate ability, time spent in education can operate as a signalling device for productivity. But the salience of this signalling property will vary in accordance with the difficulty of observing individual productivity in different tasks.

2 Data and context

We examine the labour market for university graduates in Mozambique. As with many other low-income countries, Mozambique's education system has witnessed significant expansion

over recent decades. Reflecting serious historical human capital deficits, most of this has occurred at primary and lower secondary levels. According to Barro and Lee (2013), average levels of educational attainment among adults aged 15–64 increased from 0.07 to 2.03 years between 1900 and 2010.[1] This is not to say access to higher education has been stagnant; rather, it has grown but from a very low base. In 2010, approximately 10,000 workers had a university degree, or about 0.03 per cent of the population.

Mozambique remains predominantly an agricultural economy (Jones and Tarp 2016). In urban areas, in which around two-fifths of the population reside, informal activities largely of a commercial nature are common, leaving a comparatively small formal or modern employment sector. Estimates from the most recent published household budget survey (in 2014–15) suggest that a small minority of all workers receive a regular wage. And among these, the public sector has played a key role. From our own survey, described below, we find that among recent graduates from university (hereafter, graduates) who find employment, around one in three do so in the public sector, often as teachers.

Data on Mozambique's labour market are limited. Although graduates continue to be relatively scarce, it is not clear if they are absorbed either quickly or easily into the modern sector, at least in higher-skilled occupations. To shed light on this, Jones and Santos (2022) implemented an education-to-employment transition (tracer) survey, focusing on the experiences of graduates as they enter the labour market. The baseline survey of final year undergraduates at six of the country's largest universities was conducted in 2017, jointly covering approximately three-quarters of the population of university students.[2] In total, we surveyed 2,174 final-year undergraduates at baseline, classified into seven study fields.[3] Of those surveyed, 2,100 agreed to be contacted again by telephone—we refer to these as our baseline sample. We did so on a quarterly basis over an 18-month period, starting in 2018 after their studies were to have finished. In each follow-up survey round we collected information on their employment status, including current labour market earnings.

[1] Data available from: www.barrolee.com, last accessed 14 August 2022.

[2] The survey encompassed a grand total of 27 faculties across 106 different courses at the following universities: Universidade Eduardo Mondlane (UEM), Universidade Pedagógica (UP), Universidade Politécnica (AP), and Universidade São Tomás de Moçambique (USTM) in their Maputo campuses; and Universidade Católica de Moçambique (UCM) and Universidade Zambeze (UZ) in their Beira campuses. See Jones et al. (2018, 2019) for details.

[3] The main study fields were: Education; Languages and Humanities; Social Sciences, Management, and Law; Natural Sciences; Engineering, Industry, and Construction; Agriculture; and Health and Welfare. All participants were aged 18 and older and gave written informed consent to participate in the baseline survey and follow-up telephone rounds. They were also free to desist at any time, and the data collected were anonymized. In the absence of a full institutional review board at both UNU-WIDER and local institutions at the time of the surveys, approval for undertaking the research was received from each participating university prior to fieldwork and upon discussion of the research and survey design and procedures.

Table 1 and Figure 1 summarize key characteristics of the sample, focusing on the follow-up survey rounds. Importantly, there was low attrition over time. In the final survey round, around two years after the first survey, we were able to contact nearly 90 per cent of the baseline sample. While there is little evidence that attrition is systematic (for elaboration see Jones and Santos (2022)), we nonetheless adjust the original sample weights constructed from the baseline survey to ensure that the sampling strata (i.e. gender and field of study) represent a fixed share of observations in each round.[4]

Table 1: Descriptive statistics, by follow-up survey round

							Graduated		Not graduated	
	N	Female	Maputo	Working	Formal	Grad.	Working	Wage	Working	Wage
1	2,051	0.44	0.78	0.48	0.14	0.03	0.64	20,628	0.47	9,751
2	2,004	0.44	0.78	0.58	0.17	0.25	0.68	12,494	0.55	10,367
3	1,971	0.44	0.78	0.60	0.19	0.35	0.67	12,605	0.57	9,933
4	1,921	0.44	0.76	0.52	0.20	0.50	0.58	14,118	0.47	11,174
5	1,880	0.44	0.76	0.66	0.27	0.55	0.71	14,737	0.59	11,993
6	1,849	0.44	0.74	0.70	0.33	0.68	0.72	17,119	0.63	13,029

Note: row numbers give the survey round; the first labelled column ('N') indicates the observation count; all other columns, except 'wage', are proportions; 'Maputo' is the share residing in the capital city (Maputo region); 'Formal' gives the share in a formal job, defined as having a written contract and registered in the social security system (INSS); 'Graduated' refers to individuals reporting to have graduated (certified) from the university course they were attending in 2017; 'Working' and 'Wage' are conditional on graduation status; wages are average monthly earnings (after tax), reported by participants (workers) in the national currency (USD1 $\approx$ MZN60) in constant 2019 prices.
Source: authors' own estimates.

Figure 1: Evolution of employment status by graduation status and survey round

Note: figures summarize information from Table 1; employment rates are conditional on graduation status.
Source: authors' own estimates.

[4] See Chen et al. (2015) for a similar procedure. The adjusted weights are used henceforth (unless indicated otherwise). Quantitatively, the adjustments are small (less than $\pm 10\%$ or $\pm 3\%$ on average), and none of our results are significantly affected by this procedure (full details on request).

Two other points merit note. First, final graduation from their studies—by which we mean conclusion of all formal requirements, including any thesis component, thereby allowing the individual to receive a completion certificate—did not occur on a uniform basis. In the first survey round, less than 5 per cent of the sample had graduated, but this steadily rose to over two-thirds in the final round. Qualitative investigation suggests this is normally driven by incomplete coursework (e.g., a final thesis essay), which can be delayed due to difficulties in finding a responsive supervisor as well as tardiness (lack of urgency) on the part of students. This variation in the timing of graduation provides a valuable opportunity to investigate wage-setting mechanisms (including sheepskin effects), especially because many individuals start paid employment *before* having graduated. This is evident from Figure 1, which shows that the likelihood of being in work is only marginally higher among graduates versus non-graduates across all rounds. Put differently, around half of all observations of individuals in paid employment refer to participants who have not (yet) graduated.

Second, not all university leavers find good jobs, if any work at all. This latter group—covering those continuing to study and those formally unemployed—represents around half of the observed sample in round one and declines to 30 per cent in the final round. While the reasons for being out of work vary, they do not appear to be driven by distortionary wage-setting *institutions*. Raw average after-tax monthly earnings reported by participants, which encompass individuals working part-time, are at least double the minimum wage. Furthermore, employment conditions are often precarious—e.g., even by the last round, one-third of individuals reporting to be in work stated they did not have a written contract, and half were not registered in the (mandatory) national social security scheme. This situation is consistent with the generally lax enforcement of labour regulations in the private sector, as well as the general weakness of existing trade unions (see Dibben et al. 2015).

3 Empirical strategy

The Mozambique tracer survey provides rich longitudinal employment information about university leavers over the period following the conclusion of their final year of studies. Delays in concluding the thesis component of their degree, often despite already having found employment, suggests the following naïve model to estimate the presence of sheepskin effects, defined as the wage premium associated purely with graduation:

$$y_{ijt} = \mu_t + \gamma \text{Grad}_{it} + X'_{ijt}\lambda + \varepsilon_{ijt} \tag{1}$$

where y is log monthly earnings, 'Grad' is a dummy variable taking a value of one if all courses including the thesis are complete; X is a vector of fixed and time-varying covariates, associated with differences in earnings; and i, j, and t index individuals, occupations, and time (in rounds),

respectively. On the assumption that the timing of graduation is strictly exogenous, coefficient γ in equation (1) should give an unbiased estimate of the (average) sheepskin effect.

An immediate objection to the naïve model is that the timing of graduation is likely to be influenced by other factors, including innate ability or (perceived) employment prospects, which also may be associated with earnings. To address this, we extend the model in four directions. First, in addition to the core set of controls (X), which cover the characteristics of the individual and her position of employment, we add a set of variables measured at baseline to capture differences in ability. These include two objective measures—one of academic performance, based on short tests of verbal and numerical reasoning, and the other of general intelligence, measured by a short-form Raven's test. Both of these metrics are scored using Item Response Theory (IRT), taking a mean of zero and standard deviation of one. To these, we add the individuals' self-reported academic level (taking a value of one if they rate themselves as excellent) and their self-reported proficiency in English (taking a value of one if they rate themselves as good or very good).

Second, we control for earnings obtained in the preceding period. On one hand, this simply excludes observations of individuals who simultaneously graduate and find employment, effectively focusing attention on the immediate changes in earnings associated with graduation among the already employed.[5] On the other hand, to the extent that differences in productivity associated with human capital are indeed observed and rewarded by firms, such differences should be reflected in prior earnings. Thus, the estimate for γ conditional on prior earnings will represent the rent from holding a diploma that is unrelated to productivity differences observed by the firm. This interpretation assumes that any changes to human capital (productivity) in the period immediately preceding graduation is negligible. We believe this is reasonable on two accounts, however: (1) the final graduation hurdle typically only involves organizing an (oral) thesis examination, by which time all substantive work has been complete; and (2) by including lagged earnings, the individual must already be working, implying limited time for studies.

Third, individuals may by stimulated to finalize their university education on the basis of an expected increment to earnings or employment opportunity (e.g., a minimum condition for a new post). If we are willing to accept such expectations play an independent role in behaviour, such as by motivating a more intense job search, bargaining, or proactive organization of the thesis defence, then failing to control for these expectations may represent a form of omitted variable bias. Consequently, we add information on log earnings expectations anchored to a fixed future date, collected in each round of the survey including the baseline as a further control.

[5] Earnings are set to missing for individuals not in paid employment.

Fourth, in addition to receiving a formal certificate, the act of graduation may allow (prospective) employers to acquire additional information about individual ability, such as through a final transcript that covers all courses including the thesis. In order to distinguish between the news effect of this information from the core diploma (sheepskin) effect, we interact graduation status with the individual's final grade. Taken together, this implies the following model:

$$y_{ijt} = \mu_t + (\gamma + \delta \text{Final}_{it}) \times \text{Grad}_{it} + \alpha y_{ijt-1} + \varphi \text{E}_{it-1}(y_{ijT}) + A_i'\beta + X_{ijt}'\lambda + \varepsilon_{ijt} \qquad (2)$$

where we standardize the final grade ('Final') to take a mean of zero and standard deviation of one, and individual (baseline) ability measures are represented by the vector A. This represents our most complete specification, which addresses a range of potentially confounding effects associated with the specific timing of graduation. As discussed further below, we take this specification as our point of departure for further investigation of heterogeneity in sheepskin effects.

4 Results

We begin with the first-order question of whether there is any evidence for the presence of sheepskin effects in log monthly earnings among university leavers. Table 2 sets out the results, distinguishing focusing on the sample employed by a third party, which excludes the self-employed and workers in family businesses where capital and labour income are typically difficult to separate. Unless otherwise indicated, we exclude this latter group henceforth.[6]

Columns (1) to (4) sequentially build up to our naïve specification. In addition to time fixed effects, we add controls for fixed individual characteristics; employment type (e.g., employer and sector dummies); and employment conditions, which capture various aspects of the quality or precariousness of the position, such as having a permanent position, holding a written contract, and being registered in the social security system. Columns (5) to (8) continue to build up to the full specification, based on the four extensions described above. Namely, we add the set of proxies for individual ability; we control for actual earnings in the previous period; we control for expected earnings, as reported in the previous period; and, last, we add the interaction of graduation status with the (standardized and centred) overall final grade. Standard errors are clustered at the individual level.

Focusing on panel (a), a moderate sheepskin effect is present. Based on the full specification, this is equal to around a 7 per cent earnings premium in the short run, or 16 per cent in the longer run. This latter estimate takes into account the shift in interpretation of the coefficient on graduation status when lagged earnings are included in the model (compare column 5 to

[6] Results for the full sample are available on request and do not change in any meaningful way.

6). Importantly, the derived long-term estimate from column (8) is neither statistically nor economically significantly different from the corresponding estimate in the ability-augmented naïve specification (column 5). This suggests that, despite the large increase in the model goodness-of-fit when prior earnings is included, this term is not strongly associated with the timing of graduation (*ceteris paribus*). And our results do not change if we restrict attention to individuals working in both current and prior rounds (panel b).

Table 2: Regression analysis of reported log earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a) *Employee sample*:								
Graduated	0.244***	0.198***	0.199***	0.174***	0.173***	0.071***	0.061***	0.061***
	(0.035)	(0.034)	(0.032)	(0.028)	(0.028)	(0.017)	(0.017)	(0.017)
Prior period earnings (log.)						0.659***	0.612***	0.610***
						(0.025)	(0.027)	(0.027)
Expected future earnings (log.)							0.111***	0.110***
							(0.018)	(0.018)
Graduated × final grade								0.032***
								(0.010)
Obs.	4,674	4,674	4,674	4,674	4,674	3,163	3,078	3,078
R^2(adj.)	0.08	0.20	0.25	0.37	0.38	0.67	0.67	0.67
Oster's delta	1.00	0.42	0.60	0.76	0.73	0.69	0.57	0.57
(b) *Employee sample, with prior earnings*:								
Graduated	0.211***	0.156***	0.175***	0.154***	0.156***	0.071***	0.061***	0.061***
	(0.038)	(0.037)	(0.035)	(0.031)	(0.031)	(0.017)	(0.017)	(0.017)
Prior period earnings (log.)						0.659***	0.612***	0.610***
						(0.025)	(0.027)	(0.027)
Expected future earnings (log.)							0.111***	0.110***
							(0.018)	(0.018)
Graduated × final grade								0.032***
								(0.010)
Obs.	3,163	3,163	3,163	3,163	3,163	3,163	3,078	3,078
R^2(adj.)	0.07	0.20	0.25	0.37	0.37	0.67	0.67	0.67
Oster's delta	0.34	0.34	0.60	0.75	0.76	0.69	0.57	0.57
Controls:								
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual chars.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment type	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment conditions	No	No	No	Yes	Yes	Yes	Yes	Yes
Ability proxies	No	No	No	No	Yes	Yes	Yes	Yes

Note: significance: *** 1%, ** 5%, * 10%. The table summarizes estimates of the determinants of reported log earnings across all follow-up survey rounds; moving left to right in the columns builds up to the full specification, as per equation (2), with terms entering each specification noted in the footer; panel (a) only includes individuals reporting to be employed by a third party; panel (b) restricts the sample to all individuals not self-employed who report earnings in the previous period; observations are dropped when missing (e.g., no prior earnings); standard errors (in parentheses) are clustered by unique individual.

Source: authors' own estimates.

As a first step to assessing the robustness of our results, we calculate Oster's delta (Oster 2019). For each specification, this gives the estimated proportional degree of selection on unobservable factors relative to observed factors (included in a given model) required to drive the estimated coefficient on the binary graduation variable to zero. These are reported in the footer of Table 2 and suggest a good degree of robustness, particularly in light of the very rich set of controls we employ. For instance, the final column of panel (a) suggests that a selection of unobservables

would need to be at least 0.6 as large as a selection on observables to yield an estimate for γ not different from zero.

This is supported by an analysis of the propensity to graduate, which we explore in Table 3. Here we consider the pairwise association between baseline covariates (in rows) and having formally graduated.[7] The first column looks at graduation in a cross-section—that is, the dependent variable takes a value of one if the individual graduates by the end of the survey period and zero otherwise. Next, we consider each round individually but treat graduation as an absorbing state such that, in each consecutive round, we exclude individuals who have already graduated. The final column pools the data, also maintaining the previous exclusion rule (and controls for round effects). Each cell reports the difference in graduation rates associated with a unit increase in the row variable, most of which are dummy variables. Two main points emerge. First, reflecting differences in academic rules and departmental norms, there are systematic differences in graduation (timing) according to the individual's university and study field. For instance, students of STEM courses tend to graduate earlier, and there is some clustering of graduation in certain fields (e.g., education) at the end of the calendar year (round 4), likely associated with the timing of graduation ceremonies. There is also a moderate positive association between self-assessed academic excellence and graduation, but, critically, we observe no systematic unconditional relationship between the outcome and objective measures of ability (Raven's and numerical/literacy scores). Overall, this suggests that after controlling for the very rich set of controls that we have at our disposal, the possibility for substantive confounding of the relationship between graduation and earnings is likely to be minimal.

[7] All estimates here are based on simple linear regressions with robust standard errors. In the final column, we control for period effects and also cluster standard errors at the individual level.

Table 3: Correlates of propensity to have graduated

		Round						
	Ever	1	2	3	4	5	6	Pool
Female	0.01	-0.01	-0.03	0.01	0.00	0.02	0.06	0.00
Age	0.00	0.00	0.00	-0.00	0.01	0.00	-0.01	0.00
Private university	-0.02	0.02†	-0.03	0.04	0.07	0.04	-0.16†	0.00
Education study field	0.07†	-0.01	0.04	-0.01	0.11†	0.02	0.03	0.03
STEM study field	0.03	0.05†	0.16†	0.02	-0.08†	0.00	-0.08	0.03†
Raven's score	0.00	-0.00	-0.00	0.02	-0.02	-0.01	0.01	0.00
Ability score	0.01	0.00	0.00	0.02	0.04†	-0.01	-0.03	0.01
Academic level (self)	0.15†	-0.00	0.09†	0.08†	0.05	0.08†	0.13†	0.06†
English proficiency	-0.01	0.01	0.05	-0.02	-0.03	0.03	-0.03	0.00
Obs.	2,069	2,051	1,945	1,477	1,265	989	886	8,613

Note: significance: † 1%. Cells report pairwise associations based on linear regressions of graduation status against the row variables (separately); all explanatory variables are taken from the baseline survey (time invariant); column 'Ever' is a cross-section where the dependent variable takes a value of one if the individual graduated by the last survey round; estimates by survey round (1–6) exclude individuals who have graduated in prior rounds; final column pools the round-specific analysis.

Source: authors' own estimates.

Table 4 continues to explore robustness, focusing on possible confounding associated with job dynamics, such as movements between employment positions. To do so, we narrow the analysis to sub-samples in which the characteristics of employment remain stable across consecutive periods. For reference purposes, column (1) replicates the full specification as before. Column (2) excludes all individuals who self-report to be in a different employment position to the previous survey round. Column (3) excludes individuals reporting changes in the sector of employment. Column (4) excludes individuals reporting changes in the type of employer. Column (5) excludes individuals reporting changes in hours worked (part- versus full-time). Column (6) excludes individuals reporting changes in overall employment quality, measured by the sum of a vector of eight dummy variables. And column (7) combines all these changes into an aggregate indicator, excluding all individuals with changes in any of the aforementioned dimensions. Evidence for a moderate sheepskin effect is confirmed across all these specifications, in both the employees and full samples.

Table 4: Robustness analysis of reported log earnings, controlling for job dynamics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Graduated	0.061***	0.051***	0.057***	0.058***	0.045***	0.045*	0.044**
	(0.017)	(0.014)	(0.016)	(0.016)	(0.017)	(0.023)	(0.022)
Graduated × final grade	0.032***	0.028***	0.024**	0.026***	0.024**	0.032**	0.029*
	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.015)	(0.015)
Prior period earnings (log.)	0.610***	0.686***	0.664***	0.660***	0.666***	0.635***	0.649***
	(0.027)	(0.028)	(0.028)	(0.028)	(0.031)	(0.051)	(0.055)
Expected future earnings (log.)	0.110***	0.095***	0.095***	0.105***	0.097***	0.111***	0.128***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)	(0.028)	(0.027)
Obs.	3,078	2,786	2,767	2,763	2,449	1,270	1,139
R^2(adj.)	0.67	0.71	0.69	0.69	0.70	0.68	0.68
Sample exclusions	–	Position	Sector	Employer	Hours	Quality	Any

Note: significance: *** 1%, ** 5%, * 10%. The table reports estimates for the full model (as per column 8 of Table 2); column (1) replicates the reference estimates; columns (2)–(7) exclude individuals who report changes in the characteristics of their employment across rounds according to the dimension indicated in the footer (e.g., 'Position' excludes individuals self-reporting they are not in the same position as in the previous round).

Source: authors' own estimates.

As a fourth exercise, we consider temporal dynamics, namely variation over time in earnings differences associated with graduation status. Table 5 presents this analysis for the employee sample. In columns (2) to (4), we respectively interact graduation status with an indicator for being in a new or different position (as per the dummy variable used in Table 4, column 8), experience in the current job (in rounds), and time since graduation (also in rounds). Neither of these first two interaction terms appear significant or alter the magnitude of the estimated short-run sheepskin effect (but the statistical significance does vary). However, column (4) suggests that the earnings increment associated with graduation is not always immediate but becomes more consistent over time. We confirm this in columns (5) and (6), where we interact graduation status with a set of dummy variables, indicating the number of periods before and after the time of graduation, as per an event study design. Here we constrain so-called estimates for the treatment (ever-graduating) and control (never-graduating) groups to be equivalent in the round immediately preceding graduation, and we now include a dummy variable to capture any unobserved differences between these two groups.

The event study estimates support the general finding that, conditional on prior earnings and other controls, the formal act of graduation is associated with an increment to earnings. Here, the weighted average effect across all post-graduation periods is around 11 per cent, and the effect magnitude does not seem to diminish over time. Although this latter finding is counter to the theoretical expectation that sheepskin effects might decline, this expectation depends on the extent of employer learning. If differences in productivity remain unobserved, then sheepskin effects would be constant. Also, bureaucratic delays in the physical emission of certificates or updating of salary schedules may account for some of this variation across rounds.

Table 5: Robustness analysis of reported log earnings, controlling for temporal dynamics (employees sample)

	(1)	(2)	(3)	(4)	(5)	(6)
Graduated	0.061***	0.041*	0.048	0.032		
	(0.017)	(0.022)	(0.034)	(0.021)		
Grad. × new job		0.033				0.040
		(0.027)				(0.028)
Grad. × experience			0.005			-0.004
			(0.011)			(0.013)
Grad. × time				0.015**		
				(0.007)		
Graduate in t+4					-0.015	-0.014
					(0.059)	(0.059)
Graduate in t+3					-0.006	-0.003
					(0.048)	(0.048)
Graduate in t+2					0.058	0.060
					(0.039)	(0.039)
Graduate at t					0.088**	0.072
					(0.035)	(0.049)
Graduated at t-1					0.051	0.037
					(0.033)	(0.048)
Graduated at t-2					0.122***	0.111**
					(0.036)	(0.055)
Graduated at t-3					0.188***	0.178***
					(0.040)	(0.061)
Graduated at t-4					0.105**	0.093
					(0.046)	(0.068)
Graduated × final grade	0.032***	0.031***	0.032***	0.033***	0.032***	0.031***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Expected future earnings (log.)	0.110***	0.111***	0.110***	0.110***	0.111***	0.111***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Prior period earnings (log.)	0.610***	0.610***	0.610***	0.607***	0.606***	0.606***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Obs.	3,078	3,078	3,078	3,078	3,078	3,078
R^2(adj.)	0.67	0.67	0.67	0.67	0.67	0.67

Note: significance: *** 1%, ** 5%, * 10%. The table reports estimates for the full model (as per column 8 of Table 2) with the employee sample, extending the specification to account for temporal dynamics; column (1) is the reference estimates; columns (2)–(4) interact graduation status (respectively) with entry into a new job, time in current job, and time since graduation; columns (5) and (6) report event study estimates.

Source: authors' own estimates.

Our preferred specification, given by equation (2), controls for both lagged earnings outcomes and lagged future expectations. It implicitly takes into account the contribution of fixed individual characteristics to earnings levels, as well as prior (time-varying) factors and private information regarding expected future changes to earnings. In the context of panel data, Imbens and Wooldridge (2009) explicitly recommend this kind of dynamic specification: 'It is difficult to see how making treated and control units comparable on lagged outcomes will make the causal interpretation of their difference less credible, as suggested by the DiD [difference-in-difference] assumptions'. Furthermore, the specification captures the plausible state de-

pendence of individual earnings, especially over relatively short timeframes (see Meghir and Pistaferri 2004; Brodaty 2018).

Even so, it is helpful to validate whether our core results are sensitive to alternative functional form assumptions. As a first step, we move to a static specification, in which we replace the lagged outcome with individual fixed effects. These results are shown in column (2) of Table 6, and column (1) reports results from our complete baseline specification. Next, we allow for *both* individual fixed effects and dynamics by adopting a simple one-step robust GMM estimator—see column (3). Last, we allow for the possibility that individual earnings contain a unit root, in which case the coefficient on the lagged outcome in equation (2) is unity. Thus, column (4) re-runs our baseline model, switching from log earnings levels to differences.[8] Overall, despite some moderate differences in the point estimates (and standard errors), we continue to find a positive and significant sheepskin effect, which also clearly is of the same order of magnitude as our earlier results.

Table 6: Robustness analysis of reported log earnings, alternative models

	(1)	(2)	(3)	(4)
Graduated	0.063***	0.043*	0.070*	0.046***
	(0.015)	(0.025)	(0.040)	(0.015)
Graduated × final grade	0.025***	0.035	0.005	0.015*
	(0.009)	(0.022)	(0.036)	(0.009)
Obs.	2,778	3,487	2,653	2,778
Outcome	Level	Level	Level	1st diff.
Lagged outcome	Yes	No	Yes	No
Individual FEs	No	Yes	Yes	No
Estimator	OLS	OLS	GMM	OLS

Note: significance: *** 1%, ** 5%, * 10%. The table summarizes estimates for alternative estimators of the full model, without sample weights; column (1) is the full specification (as per Table 2); column (2) is a static specification with unit FEs only; column (3) applies a one-step GMM estimator allowing for dynamics and unit FEs; column (4) returns to the baseline specification but assumes the coefficient on the lagged outcome is unity; round fixed effects are included throughout.

Source: authors' own estimates.

5 Variation in wage-setting mechanisms

Graduates in Mozambique primarily sort into the private and public sectors.[9] Conditions in the public sector differ from those in the private sector. Most positions are of a permanent nature and based on a formal contract. Perhaps even more critically, public administration has

[8] To ensure consistency with the GMM estimator results, all estimates in Table 6 do not incorporate sample weights and use only round-specific time fixed effects.

[9] Less than 10 per cent work in a family business or are self-employed. Signalling is not an issue in these sectors, however, since workers and their family members presumably know their own productivity.

a fixed wage scale that automatically rewards university degree holders. As set out in *Decreto* number 54/2009, individuals with a *licenciatura* (undergraduate degree) receive a 60 per cent supplement to their base wage, while individuals with a technical education (*técnicos médios*) receive a 40 per cent supplement. In a very typical case, for example, where a teacher moves from having a *técnico médio* to a *licenciado* qualification, they would expect to receive an approximate 14 per cent earnings increase.[10] The wage-setting mechanisms in the private and public sectors are likely to be quite different, with implications for which type of worker sorts into which sector.

5.1 Conceptual framework

Under conventional models of screening, education plays no direct role in augmenting labour productivity, with certificates of completed education providing a signal of higher innate ability that firms reward with higher wages. A fundamental assumption of these models is that firms neither observe individual ability nor (marginal) labour productivity, perhaps due to prohibitive monitoring costs or some other feature of the production technology. As such, firms offer a fixed wage schedule, such as:

$$w_i = w_L + \gamma E_i \tag{3}$$

where we make the simplifying assumption that $E_i = 1$ if the individual obtains a degree and $E_i = 0$ otherwise. Assuming the cost of acquiring this education is lower for higher-ability individuals: $C_i = f(C^*, A_i)$, where A is innate ability and $\partial C_i / \partial A_i < 0$. Then a separating equilibrium will emerge, under which individuals with ability above a certain threshold invest in education to obtain the wage premium and the remainder opt for no education (e.g., Spence 1973; Bedard 2001; Chevalier et al. 2004).

To see this relationship more concretely, define $C_i = C^* - \delta A_i$, where C^* is fixed costs of education (e.g., fees) and δA_i is net income earned while attending education (e.g, obtained from tutoring, which may be negative). Assume individuals live for two periods without discounting. In the first period, individuals choose either to work or to engage in education, and in the second period, everyone works. Where wages follow equation (3) in all periods (with certainty), it follows that individual i will be better off investing in education if:

$$\begin{aligned} & w_L + \gamma - (C^* - \delta A_i) > 2w_L \\ \implies & A_i > (1/\delta)(w_L + C^* - \gamma) \end{aligned} \tag{4}$$

[10] Calculated as 1.6/1.4-1.

Under this model, it is worth noting that, conditional on education, there is no relation between wages and ability. Also, only individuals holding an education diploma receive the higher wage, so fractional educational investments (dropping out) are irrational.

Our model provides a rationale for 'sheepskin effects', defined as an increment in wages associated with completing some level of education regardless of ability or skills (Jaeger and Page 1996). The assumption that individual productivity is never visible to employers is extreme, however, as is the assumption that education never augments productivity. To see what happens if we relax these key assumptions for some firms, consider the simple case where the (marginal) product of labour in some specific occupation is defined as :

$$p_i = w_L + \beta \max(A_i - A^*, 0) E_i,$$

which allows education to increase productivity but only for those with ability levels above A^*.[11] Retaining previous assumptions, including education as a binary choice, if firms in this segment offer wages directly equal to each individual's marginal product, then individuals faced only with this wage schedule will invest in education if both $A_i > A^*$ and:

$$\begin{aligned} & w_L + \beta(A_i - A^*) - (C^* - \delta A_i) > 2w_L \\ \implies & A_i > (w_L + C^* + \beta A^*)/(\delta + \beta) \end{aligned} \tag{5}$$

The distinctive feature here is that knowledge gained through education increases productivity, which is observed by the firm.[12] And, although for this type of firm it is irrelevant whether or not the individual holds a diploma, it remains beneficial for higher-ability individuals to pursue education because it augments their productivity. Furthermore, contrary to the previous case, it now holds that wages increase with ability conditional on education.

Combining these two models of firm wage-setting behaviour allows for a richer choice set. Specifically, assuming just three types of firms (or occupations) across which workers freely move, individuals will obtain no education and choose the low-wage occupation if: $A_i < (1/\delta)(w_L + C^* - \gamma) \equiv A_L$. They will obtain education and join the first type of firm, which screens workers, if: $A_L < A_i < \gamma/\beta + A^* \equiv A_H$. The latter threshold represents the switching point at which educated workers prefer the second type of firm, which pays according to

[11] This may be justified from the perspective that the production technology is highly complex (skills-intensive) or that relevant knowledge gained through education is lumpy.

[12] To assure the second constraint (in the equation) binds more tightly, we might further impose: $A^* \leq (w_L + C^*)/\delta$.

productivity. This implies observed wages in the second period will be:

$$w_i = \begin{cases} w_L & \text{if } A_i < A_L \\ w_L + \gamma & \text{if } A_L < A_i < A_H \\ w_L + \beta(A_i - A^*) & \text{otherwise, i.e. } A_i > \gamma/\beta + A^* \end{cases} \tag{6}$$

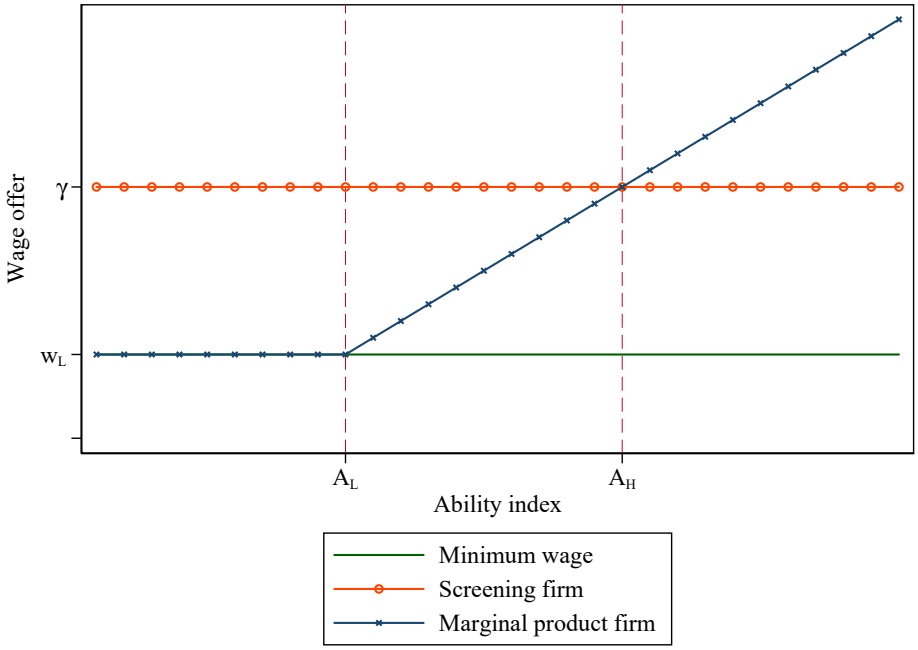
For all three regimes to exist, it must be the case that $A^* > A_L - \gamma/\beta$. In turn, the existence of this intermediate (screening) regime provides a formal definition of the sheepskin effect, namely the rent earned from completing education in excess of productivity in a counterfactual occupation, given by the difference in earnings between the two higher-wage occupations: $\gamma - \beta(A_i - A^*)$.

Figure 2 illustrates the options facing individuals in a hypothetical case, showing both wage offers (panel a) and net lifetime earnings (panel b) for individuals of differing abilities.[13] To the left of the first vertical dashed line (at $A = A_L$), earnings are always higher without education; in the middle regime, $A_L < A < A_H$, individuals maximize net earnings (i.e. wages minus education costs) by obtaining education and then joining the screening firm; and to the right of the second vertical dashed line (at $A = A_H$), net earnings are highest at the firm paying according to marginal productivity, as augmented through education.

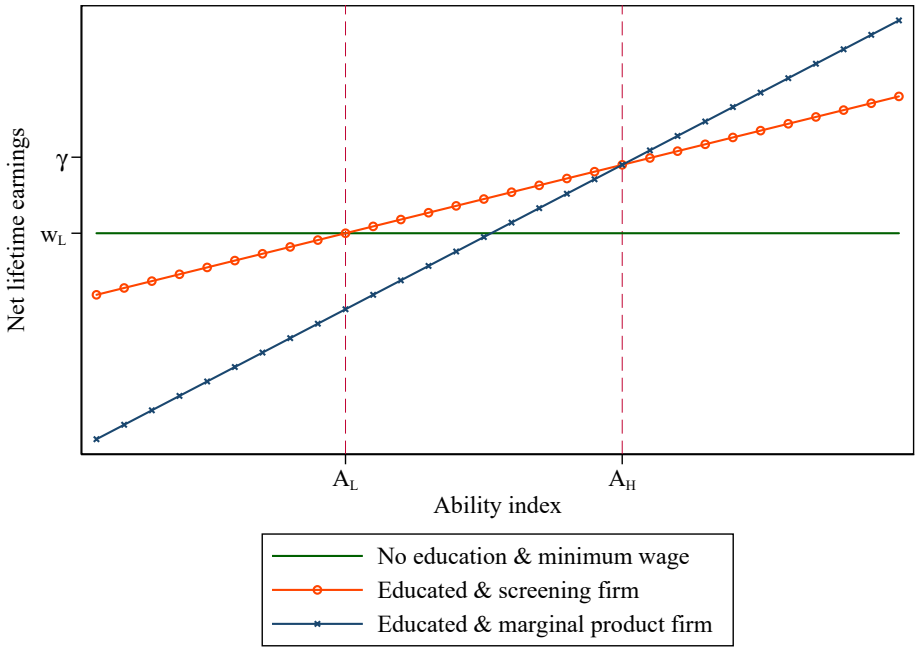
[13] Note that for individuals who attend education, net lifetime earnings depend on the cost of acquiring education, which in turn vary with ability (captured by parameter δ).

Figure 2: Earnings under different education and firm choices, by ability

(a) Wage offers by firm type



(b) Net lifetime earnings



Note: figure plots simulated estimates for alternative wage offers (panel a) and net lifetime earnings (panel b) for individuals with different fixed ability endowments (x-axis) and (in panel b) under alternative education-employment choices.
Source: authors' own estimates.

Generalizations of this set-up, allowing for partial observability of ability and employer learning are feasible. For instance, consider the general firm-specific wage function:

$$w_{ijt} = w_L + (1-\theta_j^{t_i})\beta(A_i - A^*)E_i + \theta_j^{t_i}\gamma_j E_i \tag{7}$$

where j indexes firms, t is employee tenure, and $0 \leq \theta \leq 1$ is the degree to which the contribution of ability to output is not by the firm. Continuing to assume two periods,[14] individuals will obtain education if there is some firm for which net lifetime earnings exceeds the no-education alternative:

$$\begin{aligned} &\exists j : (1-\theta_j)\beta(A_i - A^*) + \theta_j\gamma_j > w_L + (C^* - \delta A_i) \\ \implies &A_i > \left[w_L + C^* + (1-\theta_j)\beta A^* - \theta_j\gamma_j\right] / \left[(1-\theta_j)\beta + \delta\right] \end{aligned} \tag{8}$$

And individuals with the same level of education ($E = 1$) will choose firm m over firm n when:

$$\begin{aligned} &(1-\theta_m)\beta(A_i - A^*) + \theta_m\gamma_m > (1-\theta_n)\beta(A_i - A^*) + \theta_n\gamma_n \\ &\implies A_i > A^* + \left[\theta_n\gamma_n - \theta_m\gamma_m\right] / \left[\beta(\theta_n - \theta_m)\right] \end{aligned} \tag{9}$$

This confirms that the highest-ability individuals will generally prefer firms where their productivity is more easily observed. But, there also is a plausible intermediate region where educated individuals select into firms that offer a fixed education wage premium. Put differently, those firms that cannot perfectly observe ability must set $\gamma_j\theta_j$ to attract sufficient workers—i.e. higher fixed education wage premia raise the ability threshold above which individuals prefer payment according to productivity.

The testable empirical implications of variation in wage-setting behaviour associated with differences in the observability of individual productivity θ_j are of prime interest. Three of these stand out, to which we return later:

1. *Heterogeneous sheepskin effects*: variation in θ_j, as well firm-specific preferences to attract higher-ability workers, is expected to map directly to differences in the magnitude of fixed education-based wage premia offered by firms. Generally, as per equation (9), we expect to find larger sheepskin effects in firms where individual labour productivity is more difficult to observe. And, in the limit, occupations in which individual productivity can be easily verified will offer no such wage premia.

[14] Educated individuals accumulate a maximum tenure of one period ($t = 1$).

2. *Sorting across firm types by ability*: also as a direct consequence of equation (9), it follows that among workers with the same formal level of education, those with relatively lower ability will tend to sort into firms offering fixed education wage premia.

3. *Differential within-firm wage dispersion*: a corollary of the previous point is that the dispersion of wages across workers within firms (occupations) will vary across different types of firms, based on the degree to which individual productivity is observable. This can be seen from the conditional variance of equation (7):

$$\text{Var}(w_{ijt} \mid E_i = 1, \delta_j, t_i = t) = \beta^2(1-\delta_j^t)^2\text{Var}(A_i)$$

which indicates that, for $0 \leq \delta_j < 1$, wage dispersion (within-firms) also will end to increase with tenure.

5.2 Sheepskin effects in different wage regimes

Following the theoretical intuition elaborated above, a key distinction concerns the extent to which individual productivity can be observed and rewarded by employers, represented by parameter θ in equation (9). Where this is difficult *and* firms wish to attract employees of a sufficient quality (e.g., university graduates), then we would expect these firms to offer a wage premium based on the individuals' certified level of education.

The immediate empirical challenge is that θ is not observed. Nonetheless, we posit there are likely to be systematic differences between alternative types of employers (e.g., public versus private) and possibly across sectors (e.g., services versus industry). The theoretical prediction is that material differences in θ will be associated with both ability-sorting and differential wage dispersion. This constitutes our first point of entry. Concretely, columns (1a) and (1b) of Table 7 present the relationship between the two objective measures of ability—as measured prior to labour market entry—and the type of employer. To remove contamination from other factors, we use a least squares regression to control for other relevant individual characteristics, including time, gender, field of study, and university. Also, as ability is fixed, we focus on the first employment position in which the individual is observed (the same results hold across all observations). The constant term or reference category refers to individuals employed in the public sector, and coefficients on the remaining variables indicate whether there are systematic differences in the dependent variables (ability measures) across other employers.

The results clearly point to ability sorting, with the public sector (civil servants) attracting university graduates with comparatively lower scores on *both* the Raven's and academic proficiency tests. By way of example, individuals working in the private sector score close to 0.4 standard deviations higher on the Raven's test than those working in the public sector. Column

(1c) extends the same analysis to wage dispersion, where the dependent variable is the absolute value of the residual estimated from a regression of log earnings on the core set of individual and employment characteristics (as per Table 2 column 5). This supports the same conclusion, and the public sector again stands out as showing the lowest conditional wage dispersion among all employer types. As might be expected, wage dispersion is highest in the self-employment category.

Table 7: Ability sorting and wage dispersion across firms (in first employment position, all workers)

	(1a) Raven's	(1b) Academ.	(1c) Dispers.	(2a) Raven's	(2b) Academ.	(2c) Dispers.
Constant	-0.075	0.005	0.274*	-0.180	-0.055	0.386**
	(0.154)	(0.140)	(0.159)	(0.163)	(0.149)	(0.170)
Civil servant	-0.390***	-0.153**	-0.203***	-0.404***	-0.157**	-0.198**
	(0.070)	(0.070)	(0.074)	(0.074)	(0.074)	(0.077)
Self-employed	-0.122*	-0.105	0.387***	-0.179**	-0.150**	0.451***
	(0.071)	(0.071)	(0.094)	(0.075)	(0.075)	(0.098)
Secondary sector				0.311**	0.292**	-0.205
				(0.138)	(0.132)	(0.161)
Private services				0.219**	0.140	-0.268**
				(0.097)	(0.096)	(0.113)
Ed./health services				0.180*	0.110	-0.162
				(0.102)	(0.102)	(0.120)
Obs.	1,528	1,528	1,528	1,528	1,528	1,528

Note: significance: *** 1%, ** 5%, * 10%. The table summarizes estimates of the relationship between ability and employment outcomes; columns (1a) and (2a) regress alternative ability metrics (as measured at baseline) against employer types, where the constant (reference category) is workers in the private sector, including NGOs; in column (1c) the dependent variable is the squared residual from our main earnings regression; columns (2a)–(2c) replicate the same analysis adding sector-specific controls; full sample included.

Source: authors' own estimates.

Columns (2a)–(2c) repeat the analysis, adding the reported occupational sector to the set of regressors. The reference category is public sector workers employed only in primary occupations, such as agriculture or extractive industries. These controls add little to the analysis, indicating that there are no additional systematic differences in ability or wage dispersion between sectors, and it is the *type* of employer that matters most.

Together, these preliminary insights recommend two approaches to test for heterogeneity in sheepskin effects. The first is whether graduation earnings premia vary directly with ability or wage dispersion across firms. The hypothesis is that the presence of positive sheepskin effects tend to attract individuals of lower average ability, whereas higher-ability individuals tend to be found in positions without sheepskin effects. We implement this test by extending the full model (Table 2 column 8) to include interaction terms constructed from standardized employer-by-sector means of the three correlates of θ considered previously. These results are reported in columns (1) to (3) of Table 8. Consistent with our framework, in which the ease of observing individual productivity varies across firms, there is strong evidence that sheepskin effects are found predominantly in occupations attracting lower-ability graduates and those

with lower earnings dispersion. For instance, from column (3) we see that a one standard deviation reduction in mean earnings dispersion is associated with an approximate doubling of the graduation premium (from 0.053 to 0.095). In this sense, sheepskin effects represent a mechanism to attract university graduates, presumed to be of higher ability than non-graduates but a substitute for payment-by-ability.

The second test of heterogeneity concerns variation across employer types. These results are shown in column (4) of Table 8. In line with previous results, we note a significantly larger positive sheepskin effect in the public sector of almost 10 per cent and a smaller and statistically weaker relationship among all other workers captured by the non-interacted term ('graduated'). Last, column (5) combines all these interaction terms to capture heterogeneity, leaving the reference category as private sector employees. Here, the broad pattern of previous results is maintained, but due to substantial multicollinearity we cannot determine which specific dimension of heterogeneity is most important.

Table 8: Analysis of differences in wage setting across firms

	(1)	(2)	(3)	(4)	(5)
Graduated	0.059***	0.060***	0.046**	0.037*	0.036
	(0.017)	(0.017)	(0.019)	(0.022)	(0.028)
Graduated × Raven's score	-0.018				-0.019
	(0.012)				(0.049)
Graduated × academic score		-0.010			0.028
		(0.013)			(0.039)
Graduated × earnings dispersion			-0.038*		-0.016
			(0.021)		(0.031)
Graduated × civil servant				0.060**	0.045
				(0.027)	(0.064)
Graduated × final grade	0.032***	0.031***	0.032***	0.032***	0.032***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Prior period earnings (log.)	0.611***	0.611***	0.611***	0.611***	0.611***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Expected future earnings (log.)	0.110***	0.110***	0.109***	0.109***	0.109***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Obs.	3,078	3,078	3,078	3,078	3,078
R^2(adj.)	0.67	0.67	0.67	0.67	0.67

Note: significance: *** 1%, ** 5%, * 10%. The table reports estimates for the full model (as per column 8 of Table 2), focusing on the full sample and adding interaction terms that proxy for possible dimensions of heterogeneity in sheepskin effects.

Source: authors' own estimates.

6 Conclusion

This study added to the literature on diploma effects in private returns to education by exploiting high frequency longitudinal data on Mozambican university students. Concretely, following a baseline survey, we tracked individuals as they finished their studies and entered the labour market. In many cases, due to delays in defending their written thesis, most university leavers obtained paid work before formally graduating. As such, in contrast to typical cross-section analyses, we observed earnings for the same individual both immediately before and after receiving their graduation diploma. Conditional on other factors, including lagged earnings and future expectations, this provided a rare opportunity to study the existence and magnitude of any increment to earnings that can be attributed to receiving such a certificate.

Our main finding was that, even after controlling for multiple measures of ability and prior earnings, there is a small but statistically significant short-run sheepskin effect of about 7 per cent in Mozambique. While this effect is substantially smaller than those estimated previously in other developing countries, we found variation across different kinds of employers. Consistent with a stylized model where productivity is differentially observable across different types of firms, the evidence points to larger sheepskin effects in the public sector, as well as ability sorting across occupations.

In sum, the evidence from Mozambique supports a theoretical conjecture that differences in the observability of individual productivity map to differences in wage-setting mechanisms. Fixed premia associated with completing education, such as a university degree, represent a simple device to attract higher-ability individuals within the overall population of workers. But since such premia are held fixed, the highest-ability individuals prefer to find employment where individual productivity can be rewarded. So, within the pool of university graduates, lower-ability individuals appear to gravitate towards employers offering fixed earnings premia, which are predominantly in the public sector. It is these same employers that also show the lowest dispersion of earnings. Moreover, as expected, there is no evidence of sheepskin effects among the self-employed.

References

Aslam, M., Bari, F., and Kingdon, G. (2012). 'Returns to Schooling, Ability and Cognitive Skills in Pakistan'. *Education Economics*, 20(2): 139–73. https://doi.org/10.1080/09645292.2010.488470

Barro, R. J., and Lee, J. W. (2013). 'A New Data Set of Educational Attainment in the World, 1950–2010'. *Journal of Development Economics*, 104(-): 184–98. https://doi.org/10.1016/j.jdeveco.2012.10.001

Bauer, T. K., Dross, P. J., and Haisken-DeNew, J. P. (2005). 'Sheepskin Effects in Japan'. *International Journal of Manpower*, 26(4): 320–35. https://doi.org/10.1108/01437720510609528

Becker, G. S. (1962). 'Investment in Human Capital: A Theoretical Analysis'. *Journal of Political Economy*, 70(5): 9–49. https://doi.org/10.1086/258724

Bedard, K. (2001). 'Human Capital Versus Signaling Models: University Access and High School Dropouts'. *Journal of Political Economy*, 109(4): 749–75. https://doi.org/10.1086/322089

Brodaty, T. (2018). 'Is the Ladder Sticky? Measuring Semi-Parametrically State Dependence in Earnings Mobility'. *Applied Economics*, 50(2): 143–56. https://doi.org/10.1080/00036846.2017.1316478

Chen, Q., Gelman, A., Tracy, M., Norris, F. H., and Galea, S. (2015). 'Incorporating the Sampling Design in Weighting Adjustments for Panel Attrition'. *Statistics in Medicine*, 34(28): 3637–47. https://doi.org/10.1002/sim.6618

Chevalier, A., Harmon, C., Walker, I., and Zhu, Y. (2004). 'Does Education Raise Productivity, or Just Reflect It?' *The Economic Journal*, 114(499): F499–F517. https://doi.org/10.1111/j.1468-0297.2004.00256.x

Crespo, A., and Reis, M. C. (2009). 'Sheepskin Effects and the Relationship Between Earnings and Education: Analyzing Their Evolution Over Time in Brazil'. *Revista Brasileira de Economia*, 63(3): 209–31. https://doi.org/10.1590/S0034-71402009000300002

Denny, K. J., and Harmon, C. P. (2001). 'Testing for Sheepskin Effects in Earnings Equations: Evidence for Five Countries'. *Applied Economics Letters*, 8(9): 635–7. https://doi.org/10.1080/13504850010028625

Dibben, P., Wood, G., and Williams, C. C. (2015). 'Pressures Towards and Against Formalization: Regulation and Informal Employment in Mozambique'. *International Labour Review*, 154(3): 373–92. https://doi.org/10.1111/j.1564-913X.2014.00014.x

Farber, H. S., and Gibbons, R. (1996). 'Learning and Wage Dynamics'. *The Quarterly Journal of Economics*, 111(4): 1007–47. https://doi.org/10.2307/2946706

Gibson, J. (2000). 'Sheepskin Effects and the Returns to Education in New Zealand: Do They Differ by Ethnic Groups?' *New Zealand Economic Papers*, 34(2): 201–20. https://doi.org/10.1080/00779950009544323

Imbens, G. W., and Wooldridge, J. M. (2009). 'Recent Developments in the Econometrics of Program Evaluation'. *Journal of Economic Literature*, 47(1): 5–86. https://doi.org/10.1257/jel.47.1.5

Jaeger, D. A., and Page, M. E. (1996). 'Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education'. *The Review of Economics and Statistics*, 78(4): 733–40. https://doi.org/10.2307/2109960

Jones, S., Mambo, F., Mazive, E., Paris, Y., Santos, R., and Xirinda, G. (2018). *Baseline Survey on University Students' School-to-Employment Transition.* Helsinki: UNU-WIDER, University of Copenhagen, University of Eduardo Mondlane. Available at: www.wider.unu.edu/publication/baseline-survey-school-work-transitions-university-graduates-mozambique (accessed in January 2024)

Jones, S., and Santos, R. (2022). 'Can Information Correct Optimistic Wage Expectations? Evidence from Mozambican Job-Seekers'. *Journal of Development Economics*, 159(-): 102987. https://doi.org/10.1016/j.jdeveco.2022.102987

Jones, S., Santos, R., and Xirinda, G. (2019). *Final Report of the Survey on the Education-*

Employment Transitions of University Finalists. Helsinki: UNU-WIDER, University of Copenhagen, University of Eduardo Mondlane. Available at: https://www.wider.unu.edu/publication/survey-education-employment-transitions-university-finalists (accessed in January 2024)

Jones, S., and Tarp, F. (2016). 'Mozambique: Growth Experience Through an Employment Lens'. In F. Tarp and H. Bhorat (eds), *Africa's Lions: Growth Traps and Opportunities for Six African Economies* (pp. 145–80). Washington, DC: Brookings Institution Press. https://doi.org/10.35188/UNU-WIDER/2015/998-5

Layard, R., and Psacharopoulos, G. (1974). 'The Screening Hypothesis and the Returns to Education'. *Journal of Political Economy*, 82(5): 985–98. https://doi.org/10.1086/260251

Mansour, H. (2012). 'Does Employer Learning Vary by Occupation?' *Journal of Labor Economics*, 30(2): 415–44. https://doi.org/10.1086/663590

Meghir, C., and Pistaferri, L. (2004). 'Income Variance Dynamics and Heterogeneity'. *Econometrica*, 72(1): 1–32. https://doi.org/10.1111/j.1468-0262.2004.00476.x

Mehta, A., and Villarreal, H. J. (2008). 'Why Do Diplomas Pay? An Expanded Mincerian Framework Applied to Mexico'. *Applied Economics*, 40(24): 3127–44. https://doi.org/10.1080/00036840600994120

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: NBER Press.

Mora, J. J. (2003). 'Sheepskin Effects and Screening in Colombia'. *Colombian Economic Journal*, 1(1): 95–108.

Olfindo, R. (2018). 'Diploma as Signal? Estimating Sheepskin Effects in the Philippines'. *International Journal of Educational Development*, 60(-): 113–9. https://doi.org/10.1016/j.ijedudev.2017.11.001

Oster, E. (2019). 'Unobservable Selection and Coefficient Stability: Theory and Evidence'. *Journal of Business & Economic Statistics*, 37(2): 187–204. https://doi.org/10.1080/07350015.2016.1227711

Park, S. (2011). 'Returning to School for Higher Returns'. *Economicsste of Education Review*, 30(6): 1215–28. https://doi.org/10.1016/j.econedurev.2011.07.006

Psacharopoulos, G. (1979). 'On the Weak Versus the Strong Version of the Screening Hypothesis'. *Economics Letters*, 4(2): 181–5. https://doi.org/10.1016/0165-1765(79)90232-5

Schady, N. R. (2003). 'Convexity and Sheepskin Effects in the Human Capital Earnings Function: Recent Evidence for Filipino Men'. *Oxford Bulletin of Economics and Statistics*, 65(2): 171–96. https://doi.org/10.1111/1468-0084.00042

Spence, M. (1973). 'Job Market Signaling'. *The Quarterly Journal of Economics*, 87(3): 355–74. https://doi.org/10.2307/1882010