Safety valve or sinkhole? Vocational schooling in South Africa

Todd Pugatch

Correspondence:
todd.pugatch@oregonstate.edu
Department of Economics, Oregon
State University and IZA, 303 Ballard
Extension Hall, Corvallis OR 97331,
USA

Abstract

As an alternative to traditional academic schooling, vocational schooling in South Africa may serve as a safety valve for students encountering difficulty in the transition from school to work. Yet if ineffective, vocational schooling could also be a sinkhole, offering little chance for success on the labor market. After defining the terms "safety valve" and "sinkhole" in a model of human capital investment with multiple schooling types, I test for evidence of these characteristics using a panel of urban youth in South Africa. I find support for the safety valve role of vocational schooling, with a 1 percentage point decrease in vocational enrollment in response to grade failure, compared to a decline of 40 percentage points for academic enrollment. In contrast, I fail to find evidence that vocational schooling is a sinkhole, with wage and employment returns at least as large as those for academic schooling. The results suggest that vocational schooling plays an important role in easing difficult school to work transitions for South African youth.

JEL classification: I25; J24; J31; O12.

Keywords: Human capital investment; Vocational schooling; Youth unemployment; South Africa

1 Introduction

Youth unemployment far exceeds adult unemployment in most countries. Developing countries are particularly concerned about youth unemployment because their younger populations make the problem more severe. Low productivity levels and mismatch between youth skills and employer demands are often cited as major contributors to youth unemployment. Vocational education may be an effective means to provide youth with skills that employers seek. Countries with high youth unemployment face the largest gaps between the supply and demand of qualified young workers, and therefore might gain the most from effective vocational education.

South Africa is a prime example of a country in which vocational education could play this important role. Youth unemployment is exceptionally high, exceeding 50% in recent years,[1] with surprisingly few opportunities in the informal sector. Students face wide variation in school quality and high rates of grade failure and repetition. Improving the difficult school to work transition for youth, particularly for those from vulnerable circumstances, is a key social and political issue. Vocational schooling, in which about 400,000 students enroll at the secondary and post-secondary levels annually each (Bunting and Cloete 2010; UNESCO Institute for Statistics 2011), offers the promise of an

alternative educational path for students struggling in the general education system. The South African government intends to expand vocational secondary programs by over 6 times current enrollment in the coming years (Oxford Business Group 2012). Yet little is known about what types of students enroll in vocational programs in South Africa, why they choose to enroll, or how they fare on the labor market.

This paper analyzes the role of vocational schooling in the school to work transition of South African youth. Specifically, I examine whether vocational schooling serves as a "safety valve" for students struggling in traditional academic schooling, experiencing adverse household shocks, or forming households. I also examine whether vocational schooling benefits students after they leave, or whether (consistent with concerns expressed by many South African students) it is a "sinkhole" offering little chance for success on the labor market. After defining the terms safety valve and sinkhole in a model of human capital investment with multiple schooling types, I test for evidence of these characteristics in a panel of urban youth in South Africa.

I find support for the safety valve role of vocational schooling, with a 1 percentage point decrease in vocational enrollment in response to grade failure, compared to a decline of 40 percentage points for academic enrollment. These findings account for persistent unobserved heterogeneity and remain after extending the model to account for multi-dimensional ability and splitting the sample among various subsamples of interest. In contrast, I fail to find evidence that vocational schooling is a sinkhole, with wage and employment returns at least as large as those for academic schooling. Labor market returns to both types of schooling are precisely estimated, so that these findings are not due to large standard errors (the sample is underpowered to detect such differences when distinguishing between secondary and post-secondary schooling of each type, however). Although not causal, I use the method of Altonji et al. (2005) to present evidence that these returns estimates are not driven by selection on unobserved characteristics. The results suggest that vocational schooling plays an important role in easing difficult school to work transitions for South African youth. Students who struggle to advance in academic programs may find it optimal to enter vocational programs, where they can earn comparable returns.

Vocational schooling has been the subject of considerable previous economic research, much of it focused on estimating comparative returns between vocational and general/academic programs. For the developing world, Bennell (1996) provides an early review, questioning the then-conventional wisdom that academic returns exceeded vocational. Since then, a number of researchers have continued to estimate comparative returns for individual developing countries. A partial list includes Egypt (El-Hamidi 2006; Krafft 2013), Indonesia (Newhouse and Suryadarma 2011), Taiwan (Rodgers et al. 2006), Tanzania (Hawley 2003, Hawley 2008), and Thailand (Kahyarara and Teal 2008), with varying results regarding the direction and magnitude of comparative returns. Two recent studies (Oosterbeek and Webbink 2007; Malamud and Pop-Eleches), rely on policy changes in the Netherlands and Romania, respectively, to isolate plausibly exogenous changes in vocational schooling[2]. Both studies find little difference in labor market outcomes between students with vocational and academic schooling as a result of the policy change, and argue that non-random selection into these tracks drives the varying findings of other studies in the literature. Hanushek et al. (2011), using data from OECD countries, find initial gains to vocational schooling that are overtaken by academic schooling later

in working careers, providing another plausible explanation for contradictory findings regarding returns. Because my sample is composed entirely of young adults, the Hanushek et al. (2011) results call for caution in interpreting my results on labor market returns, as differences between academic and vocational returns may emerge later in working careers.

The notion of vocational schooling as safety valve or sinkhole has been discussed in previous literature. Arum and Shavit (1995) consider whether vocational schooling in the United States "is simply the crude mechanism of social exclusion" (i.e., a sinkhole), but conclude that "vocational education is a safety net that reduces the risk of falling to the bottom of the labor queue" (p. 187). For South Africa, Needham and Papier (2011) refer to a stigma attached to vocational schooling, and the common perception that it is a sinkhole: "Some school [general education] students saw FET college [Further Education and Training, or vocational secondary] education as a second choice education that would result in low-paying jobs with no career prospects" (p. 36). They also reference vocational schooling's safety valve role, citing qualitative evidence of poor previous schooling outcomes among vocational students: "[S]tudents viewed vocational education, particularly FET colleges, as second chance programmes that you went to if you could not make it at [traditional] school" (p. 37)[3]. This paper contributes to the literature by formalizing the safety valve and sinkhole concepts in a human capital investment model, documenting the patterns of differential selection into vocational and academic schooling in South Africa, and estimating the wage and employment returns to each type of schooling. These returns estimates, while observational in nature, use high-quality panel data that include complete schooling histories, measures of ability, and actual (rather than potential) work experience, allowing for bias reductions relative to similar estimates from a census or labor force survey.

In the next section, I define the terms "safety valve" and "sinkhole" in a simple model of human capital investment. Section 3 describes the data and presents descriptive statistics. Section 4 analyzes the role of vocational schooling in youth enrollment choices and labor market outcomes, and Section 5 concludes.

2 Theory

2.1 Safety valves and sinkholes

Consider a generalization of the standard Becker (1967)/Mincer (1974) human capital model in which agents may choose between two types of schooling, academic (a) and vocational (v). Students may invest in one type of schooling only, both, or neither[4]. Each schooling type has associated return and cost functions, $R_j(\cdot)$ and $C_j(\cdot)$, for schooling types $j = \{a, v\}$. The cost functions are general and include both the monetary and psychic costs of schooling. These return and cost functions depend on the type-specific stock of schooling S_j, the agent's ability A, and exogenous shifters Z. As is standard, assume that R_j is increasing and weakly concave in S_j and C_j is increasing and convex in S_j [5]. Also assume that R_j is increasing in A and C_j is decreasing in A: higher-ability individuals earn higher returns and face lower costs at any level of schooling. The signs of R_j and C_j are ambiguous with respect to Z.

I assume that schooling type v carries a fixed cost:

$$C_v(S_v, A_e, Z) = c_v(S_v, A_e, Z | S_v > 0) + \phi \cdot \mathbb{I}(S_v > 0) \tag{1}$$

where $\mathbb{I}(\cdot)$ is the indicator function and ϕ is the fixed cost. The fixed cost ϕ stands in for many features that distinguish vocational from academic schooling in the South African context, such as the cost of acquiring information about vocational programs, greater travel distance, an application fee, or non-pecuniary costs such as a stigma for attending a type v school.

The agent's lifetime, discounted income Y is the sum of the schooling return functions: $Y = R_a(S_a, A, Z) + R_v(S_v, A, Z)$. The agent observes his or her own ability imperfectly, and must base decisions on its expectation A_e. The agent's optimization problem is then[6]:

$$\max_{S_a, S_v} V = \{[R_a(S_a, A_e, Z) + R_v(S_v, A_e, Z)]$$
$$- [C_a(S_a, A_e, Z) + C_v(S_v, A_e, Z)]\} \tag{2}$$

The first order conditions are:

$$\frac{\partial R_j}{\partial S_j}(S_j, A_e, Z) \geq \frac{\partial C_j}{\partial S_j}(S_j, A_e, Z) \text{ for } j = a, v \tag{3}$$

where the weak inequality allows for discrete increments of S_j and corner solutions. This is the standard result that the agent invests in each type of schooling S_j to equate (as closely as possible) its marginal benefits and costs. The fixed cost to schooling type v adds the following participation constraint:

$$Y - c_v(S_v, A_e, Z) \geq \phi \tag{4}$$

The first order conditions in (3) and the participation constraint (4) implicitly define the optimal levels of schooling $S_j^* = S_j(A_e, Z)$. If there is heterogeneity in the population in the fixed cost ϕ, the first result is that there exists a threshold ϕ^* above which the agent will choose not to acquire any type v schooling.

Now consider comparative statics with respect to ability. Suppose the agent revises his expected ability, such as after failing a grade:

$$\frac{\partial V}{\partial A_e} = \sum_{j=\{a,v\}} \left[\frac{\partial R_j}{\partial A_e}(S_j, A_e, Z) - \frac{\partial C_j}{\partial A_e}(S_j, A_e, Z) \right] \tag{5}$$

The first term inside brackets on the right-hand side of (5) is the (perceived) marginal return resulting from the revision to expected ability, while the second term is the change in schooling costs. The first term is positive and the second negative, making the sign of the derivative unambiguously positive. This positive sign is intuitive: perceived returns will fall when one revises expected ability downward after failing a grade, or rise with gains in expected ability. Costs will rise in response to a failure, or decrease with a gain in expected ability.

To see how the change in expected ability affects schooling choices, first define the schooling type j-specific change in net returns as:

$$\Delta_j^A \equiv \frac{\partial R_j}{\partial A_e}(S_j, A_e, Z) - \frac{\partial C_j}{\partial A_e}(S_j, A_e, Z) \tag{6}$$

Although Δ_j^A is positive for both types of schooling, it need not be equivalent across schooling types, and therefore the agent may alter schooling investments differentially between types in order to satisfy the first-order conditions in (3). I refer to schooling type v as a *safety valve* if $|\Delta_a^A| > |\Delta_v^A|$: if this inequality holds, then schooling of type v is less responsive to revisions in expected ability. This implies that students experiencing a grade failure would be more likely to drop out of type a schooling than type v, or switch from

a to *v* than from *v* to *a*. Because of discrete increments of S_j and the possibility of corner solutions, we may observe no change in either type of schooling; those whose fixed cost of type *v* schooling (ϕ) is sufficiently high will continue to abstain from investing in type *v* schooling. However, we are more likely to observe no change in the safety valve schooling type *v* than in *a*.

The safety valve result applies equally to external shocks *Z*: schooling type *v* would also exhibit the safety valve property if $|\Delta_a^Z| > |\Delta_v^Z|$, so that schooling type *v* would be less responsive to shocks than type *a*. For negative shocks, such as the death or job loss of a household member, the safety valve schooling option would offer a more forgiving course of study, making the student less likely to drop out. For positive shocks, such as a gain in household income, investment in the safety valve schooling type should also be less responsive, because students whose household financial position is more secure are better insured against failure in the less forgiving, non-safety valve schooling option.

The presence of a safety valve type of schooling is a boon for students experiencing failure or negative shocks, as they can shift their schooling choices in favor of the relatively less affected type (*v* in the running example). They do so until the first-order conditions in (3) are satisfied, and the marginal benefits and marginal costs of each type of schooling chosen return to equilibrium. Nonetheless, the presence of the safety valve may impose a significant cost on students through inferior labor market returns, $\beta_j \equiv \frac{\partial R_j}{\partial S_j}$, earned at the new optimal schooling choices.

If these inferior returns are systematic, i.e., if $\beta_v < \beta_a$ for all $S_v = S_a$, then it would be evidence that vocational schooling is a *sinkhole*, consistent with the concerns expressed by its graduates in Section 1. That is, if schooling type *v* generates lower labor market returns than type *a* at comparable schooling levels, it is a sinkhole, because students must pay for the lower costs (pecuniary or non-pecuniary) of this type of schooling in terms of labor market performance[7]. In other words, a sinkhole schooling type draws in students with low costs, but fails to offer competitive returns on the labor market[8].

2.2 Empirical implementation

In this paper, I investigate the hypothesis that vocational schooling acts as a safety valve for South African youth, as well as the hypothesis that it is a sinkhole with respect to its labor market returns. To test the safety valve hypothesis, I seek evidence that enrollment in vocational schooling responds less than academic schooling to revisions of expected ability and external shocks. To do so, I run variations of the following regression:

$$D_{it}^j = \alpha_0^j + \alpha_f^j F_{it} + \alpha_z^j Z_{it} + \gamma^j X_{it} + \epsilon_{it}^j \tag{7}$$

where $j = \{a, v\}$ indexes academic schooling *a* and vocational schooling *v*; *i* indexes individuals; and *t* indexes time periods. The outcome variable D^j is an indicator for enrollment in schooling type *j*; *F* is a vector of grade failure history; *Z* is a vector of external shocks; *X* is a vector of additional controls, which may include fixed effects for individual, age, schooling level, and calendar year, as well as other control variables; and ϵ is an error term. In this regression, grade failure history variables *F* proxy for revisions to expected ability, making the coefficient α_f^j analogous to Δ_j^A in the model of Section 2.1. Individual fixed effects will capture a student's prior expected ability, while age and completed schooling terms will capture average ability expectations associated with those variables, allowing the failure terms to isolate changes in expected ability. Comparing the coefficients α_f^v and

α_f^a therefore provides a test of the safety valve hypothesis: if vocational schooling is a safety valve, I should find that $\alpha_f^a < \alpha_f^v < 0$.

Similarly, the safety valve hypothesis suggests that vocational enrollment should be less responsive to external shocks Z, i.e., $|\alpha_z^a| > |\alpha_z^v|$. Due to data limitations that I will explain in the following sections, I will run variations of the safety valve regression (7) on subsets of the failure and shock variables F and Z.

To test whether vocational schooling is a sinkhole, I run a variant of the standard Mincer regression:

$$Y_{it} = \beta_0 + \beta_a S_{it}^a + \beta_v S_{it}^v + \gamma X_{it} + \epsilon_{it} \tag{8}$$

where Y is a labor market outcome (employment or log wages); S^a and S^v are the stock of academic and vocational schooling, respectively; and all other notation is as before. This model assumes that the type-specific schooling return functions $R_j(\cdot)$ are linear in S_j, leading to the extension of the standard Mincer regression of (log) wages on years of each type of schooling shown above. I address the limitations of Mincer regressions in the empirical setting of this paper in Section 4.

3 Data

Vocational schooling is available to South African students beginning in Grade 10, when schooling is no longer compulsory. Students opting for a vocational program may enroll in technical schools in Grades 10-12 (also known as FET colleges, for Further Education and Training) or National Technical Certificate (NTC) programs at equivalent grade levels. Such programs tend to focus on preparing students for skilled trades and service professions, such as metal and electrical work, plumbing, and hospitality. At the post-secondary level, students with sufficiently high scores on the high school exit ("matric") exam may qualify for admission to public universities. Such universities include traditional academically oriented institutions and universities of technology (formerly known as Technikons). Although academic universities and universities of technology have some overlap in courses of study offered, the universities of technology focus more on trade and professional skills than their academic counterparts, offering accounting, teacher training, design, nursing, health services and other subjects.

I define "vocational" and "academic" schooling as mutually exclusive categories within a year of enrollment, although students may switch between them across years. Vocational schooling includes enrollment in a National Technical Certificate program (such as an FET college) or university of technology; enrollment in all other programs is academic schooling. Private higher education institutions are also available at the post-secondary level. Although such institutions may offer both academic or vocational courses of study, I consider them academically oriented institutions because I am unable to distinguish what course of study a particular student chooses. Because employment while enrolled in school is rare in South Africa, I treat school enrollment and labor market participation as mutually exclusive activities[9].

Vocational secondary and post-secondary programs have distinct requirements and attract different students. Nonetheless, they have a shared focus on training students in a narrow set of occupation-specific skills (e.g., welding and plumbing in secondary, nursing and tourism management in post-secondary), whereas academic programs tend to

focus on more general skills. In the analysis presented in Sections 4.1.1 and 4.3, I will probe robustness of the key results of the paper to separating each type of schooling into secondary and post-secondary levels.

I use data from the Cape Area Panel Study (CAPS), a longitudinal study of youth in metropolitan Cape Town, South Africa (Lam et al. 2008). CAPS sampled about 4,800 youth aged 14-22 in Wave 1 (2002) and currently includes five publicly available waves with the most recent conducted in 2009. In Wave 1, retrospective life histories were collected for each year stretching back to birth, and include information on school enrollment and advancement, job search, and employment. I update this retrospective life history data with information from Waves 2-5 to construct the panel used in this paper. I make several sample restrictions. I drop white youth, as they tend to enjoy living standards comparable to those in developed countries, while my focus here is on youth who tend to face high uncertainty in their schooling and employment outcomes. I keep only those youth observed until at least age 18. I also drop those with inconsistent schooling histories (details described in Appendix A), leaving 2,927 individuals in the sample. In the analysis, I use only those person-years with at least 9 years of completed schooling, at which point enrollment in vocational schooling becomes a viable option[10]. The resulting dataset includes 23,997 person-year observations.

Table 1 shows the panel balance at selected ages, conditional on vocational enrollment history. The sample size drops sharply at later ages due to both the young ages of entry into the panel (i.e., right-censoring) and attrition in later waves. There is evidence of selective censoring and attrition by vocational enrollment status, with vocational enrollees leaving the sample at lower rates than those who never enter a vocational program, arguing for caution when interpreting my results. However, the direction of the potential bias resulting from selective attrition is unclear: if vocational enrollees were both less successful on the labor market than others (i.e., the sinkhole hypothesis) and more likely to leave the sample as a result, then we would expect to see the opposite pattern of attrition than the one observed.

Table 2 presents summary statistics for the sample. Panel A presents characteristics by individual, while Panel B describes the person-year observations used in the panel estimation I conduct in Section 4. Coloured[11] youth make up almost two thirds of the (weighted) sample and black/Africans the remainder, as is characteristic of metropolitan Cape Town. Only 9% of the sample ever enrolls in a vocational program, with a nearly even split between secondary and post-secondary enrollees. Such small proportions are

Table 1 Panel attrition, by vocational enrollment history

	N	18	20	22	24	26	28
			Proportion censored at age				
Full sample	2,927	0.00	0.08	0.33	0.57	0.77	0.92
vocational enrollment							
Never	2,663	0.00	0.08	0.34	0.58	0.78	0.92
Any	264	0.00	0.04	0.20	0.51	0.69	0.92
Secondary	118	0.00	0.02	0.20	0.50	0.67	0.89
Post-secondary	149	0.00	0.05	0.19	0.52	0.72	0.94

Cells show number of observations (unweighted) or percent of sample (weighted) with missing enrollment information by age. Sample drawn from CAPS life history data: black and coloured young adults observed until at least age 18 with valid schooling histories. Other sample restrictions described in Appendix A.

Table 2 Summary statistics

Variable	N	Mean	S.D.	Variable	N	Mean	S.D.
Panel A: individuals				**Panel B: person-years**			
Female	2,927	0.55	0.50	Enrollment			
Black	2,927	0.35	0.48	Academic	23,997	0.37	0.48
Coloured	2,927	0.65	0.48	Vocational	23,997	0.02	0.14
Age, Wave 1	2,927	18.1	2.4	Wage	3,840	27,183	21,090
Literacy and numeracy z-score	2,927	-0.10	0.83	Employed	14,454	0.25	0.44
Ever enrolled in:				Formal job acquisition	1,297	0.28	0.45
Any vocational	2,927	0.09	0.29	Failed last grade			
Vocational secondary	2,927	0.04	0.20	Any program	23,997	0.30	0.46
Vocational post-secondary	2,927	0.05	0.23	Academic	23,997	0.29	0.46
Completed schooling (years), conditional on enrollment:				Vocational	23,997	0.01	0.10
All schooling	2,927	11.4	1.6	Previous grades failed			
Academic	2,927	11.3	1.6	Any program	23,997	0.63	0.86
Any vocational	264	0.8	0.8	Academic	23,997	0.62	0.85
Vocational secondary	118	0.2	0.5	Vocational	23,997	0.00	0.06
Vocational post-secondary	149	0.8	0.9	Not employed last period	23,997	0.86	0.34
Vocational enrollees ever enrolled in:				Previous periods employed	23,997	0.42	0.94
Academic secondary	118	0.97	0.17	Household shocks			
Academic post-secondary	149	0.39	0.49	All	2,015	0.16	0.37
				Health	2,015	0.09	0.28
				Financial	2,015	0.08	0.28
				Pension-eligible grandmother	23,997	0.13	0.33
				Pension-eligible grandfather	23,997	0.04	0.20
				Pregnancy	14,053	0.02	0.15
				Marriage	14,053	0.01	0.11

Sample drawn from CAPS life history data: black and coloured young adults observed until at least age 18 with valid schooling histories. Other sample restrictions described in Appendix A. Vocational secondary includes technical colleges and National Training Certificate (NTC) programs. Literacy and numeracy z-score is age-adjusted based on standardized test administered to all respondents in Wave 1. Vocational post-secondary includes universities of technology. Wage is annualized full-time equivalent based on 160 hours/month, denominated in constant South African rand (base year 2002). Wage refers to "typical" earnings on current jobs held among those not enrolled in school, regardless of current employment status. Employment restricted to periods when not enrolled in school. Formal method of job acquisition refers to responding to ad, sending CV, visiting firm, or using employment agency. Household shocks include death; serious illness or injury; job loss; major financial loss (including business failure or bankruptcy); abandonment or divorce; theft, fire or property damage; or other shock. Health shocks include death, serious illness, or injury. Financial shocks include job loss or major financial loss. "Pension-eligible grand(mother/father)" refers to co-residence with grandmother at least 60 years old or grandfather at least 65 years old, respectively. Marriage refers to event, not marital status. Marriage refers to event, not marital status. Marriage and pregnancy statistics for females only. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. Survey weights applied.

consistent with the presence of a fixed cost to vocational schooling, as in the model of Section 2.1[12]. Those who enroll in vocational programs complete 0.8 years of study, on average.

Among vocational secondary students, 97% have also enrolled in an academic secondary program (grades 10-12) at some point in their careers, while 39% of vocational post-secondary students have also been enrolled in an academic post-secondary program. This stands in contrast to many vocational systems, such as those in Europe, in which schooling tracks are rigidly defined and switching is difficult or impossible. By contrast, South African vocational students often arrive in their programs through experimentation and switching. Such behavior also generates useful identifying variation for the Mincer regressions based on (8) that I will use to test the sinkhole hypothesis.

Turning to the person-year characteristics in Panel B, we see that the enrollment rate in academic programs is far higher than in vocational (37% to 2%). Among the employed, mean wages (annualized full-time equivalent) are R27,183, or about USD2,567. Grade failure is quite prevalent, with 30% of person-year observations having failed their last grade enrolled. Household shocks are also common, with a frequency of 16%.

Table 3 presents enrollment rates by selected characteristics. Gender and racial differences in enrollment are modest. Although we saw in Table 2 that the overall enrollment

Table 3 Vocational enrollment, by selected characteristics

| | | Ever enrolled in vocational: | | |
	N	Any	Secondary	Post-secondary
Male	1,227	0.10	0.04	0.06
Female	1,700	0.09	0.04	0.05
Race				
Black	1,522	0.09	0.04	0.05
Coloured	1,405	0.10	0.04	0.06
Household income				
Poorest two quintiles	1,564	0.06	0.04	0.03
Richest three quintiles	1,363	0.12	0.04	0.07
Schooling				
9-11 years	1,602	0.03	0.03	0.00
12 years	840	0.13	0.08	0.06
> 12 years	485	0.21	0.01	0.20
Ability				
Low	1,809	0.06	0.04	0.02
High	1,118	0.13	0.04	0.10
Re-enrollment history				
Never	1,528	0.05	0.02	0.04
At least once	1,399	0.14	0.07	0.08
At least once	1,097	0.12	0.09	0.04
Before completing secondary				

Sample drawn from CAPS life history data: black and coloured young adults observed until at least age 18 with valid schooling histories. Other sample restrictions described in Appendix A. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. Sample sizes are raw counts, with survey-weighted means reported in remaining columns. Vocational secondary includes technical colleges and National Training Certificate (NTC) programs. Vocational post-secondary includes universities of technology. "High ability" refers to scoring above full sample age-adjusted median score on literacy and numeracy evaluation administered in Wave 1. Household income quintiles based on Wave 1 data.

rate for vocational programs was relatively low, in Table 3 we see that enrollees in vocational secondary programs come from across the income distribution, while vocational post-secondary students tend to be from wealthier backgrounds. Vocational students are also highly educated: 20% of those with more than 12 years of schooling have enrolled in a vocational program. The bulk of this group are enrollees in a post-secondary vocational program (i.e., a university of technology); only 1% of those with more than 12 years of schooling ever enrolled in a vocational secondary program. Vocational enrollment follows a similar pattern with respect to the ability distribution: the highly able choose to enroll in vocational post-secondary more than the less able, where high ability refers to scoring above the full sample, age-adjusted median on the literacy and numeracy evaluation administered in CAPS Wave 1. Vocational enrollment also exhibits an interesting pattern with respect to student re-enrollment histories. Vocational enrollment is low (just 5%) among those who never re-enrolled, i.e., those who never interrupt their studies with a period of dropout. But vocational enrollment rises to 14% among those who re-enroll at least once in their schooling career, with this activity split equally between secondary and post-secondary vocational programs. Enrollment in vocational secondary programs is 9% among those who re-enroll before completing secondary, compared to an overall vocational secondary enrollment rate of 4%. The disproportionate prevalence of re-enrollees among vocational students suggests that these programs attract many "second-chance" students who left the traditional academic system. This descriptive evidence is consistent with the safety valve role of vocational schooling, but remains to be verified in the formal analysis.

4 Results

4.1 Testing the safety valve hypothesis

4.1.1 *Do struggling students select into vocational schooling?*

If vocational schooling is a safety valve, it would drive away a smaller share of students who struggle in school. Table 4 shows results of regressions (linear probability models) of enrollment in academic or vocational schooling on measures of grade failure and unemployment. All regressions include individual fixed effects and a full set of completed schooling, age and calendar year dummies, so that identification is achieved through variation in school or labor market outcomes that is uncorrelated with time-invariant unobserved heterogeneity, including ability or motivation to the extent these are permanent characteristics; factors common to youth at particular ages or levels of schooling; or macroeconomic conditions[13]. The identification strategy is thus similar to the "schooling as a lottery" idea of Lam et al. (2011), in which school performance depends on innate characteristics and idiosyncratic stochastic shocks[14].

Panel A of Table 4 presents results for the full sample. In column (1), I find that students who failed their last grade are 40 percentage points less likely to enroll in an academic program. Each previous failure (prior to the result of the last grade) results in an additional 17 percentage point decline in the probability of enrollment in an academic program. Both coefficients are statistically significant at the 1% level. The analogous regression for vocational enrollment in column (3) shows that students who failed their last grade are only 1 percentage point less likely to enroll in vocational schooling, with significance at 10%. The differential response of academic and vocational enrollment to grade failure is large in both absolute and relative terms (mean academic and vocational enrollment are

Table 4 Enrollment response to grade failure and previous employment, by program type

	(1)	(2)	diff.	(3)	(4)	diff.
Panel A: full sample	Academic		(3)-(1)	Vocational		(4)-(2)
Failed last grade enrolled	-0.40	-0.39	***	-0.01	-0.01	***
	(0.01)***	(0.01)***		(0.005)*	(0.005)*	
Previous grades failed	-0.17	-0.17	***	0.004	0.004	***
	(0.01)***	(0.01)***		(0.004)	(0.004)	
Not employed last period		0.07			0.02	***
		(0.01)***			(0.004)***	
Observations	23,997	23,997		23,997	23,997	
R-squared	0.70	0.70		0.26	0.26	
No. individuals	2,927	2,927		2,927	2,927	
Includes work experience		x			x	
Panel B: by level	Secondary		diff.	Post-secondary		diff.
	Academic	Vocational	(2)-(1)	Academic	Vocational	(4)-(3)
Failed last grade enrolled	-0.44	-0.02	***	-0.27	-0.16	
	(0.01)***	(0.005)***		(0.06)***	(0.05)***	
Previous grades failed	-0.22	0.01	***	-0.41	-0.12	***
	(0.01)***	(0.01)		(0.09)***	(0.06)**	
Observations	17,023	17,023		6,974	6,974	
R-squared	0.77	0.32		0.60	0.49	
No. individuals	2,927	2,927		1,325	1,325	
Panel C: by previous period choice	Dropout only		diff.	Re-enrollment only		diff.
	Academic	Vocational	(2)-(1)	Academic	Vocational	(4)-(3)
Failed last grade enrolled	-0.38	-0.01	***	-0.38	-0.06	***
	(0.02)***	(0.01)		(0.03)***	(0.02)***	
Previous grades failed	-0.14	0.00	***	-0.39	-0.03	***
	(0.02)***	(0.01)		(0.03)***	(0.01)**	

Table 4 Enrollment response to grade failure and previous employment, by program type (*Continued*)

Observations	11,914	11,914		12,083	12,083	
R-squared	0.65	0.35		0.69	0.38	
No. individuals	2,912	2,912		2,702	2,702	
Panel D: by program of last enrollment	Last enrollment: academic		diff.	Last enrollment: vocational		diff.
	Academic	Vocational	(2)-(1)	Academic	Vocational	(4)-(3)
Failed last grade enrolled	-0.42	-0.01	***	0.03	-0.43	***
	(0.01)***	(0.003)**		(0.07)	(0.13)***	
Previous grades failed	-0.18	0.003	***	0.003	-0.25	*
	(0.01)***	(0.004)		(0.03)	(0.17)	
Observations	23,120	23,120		877	877	
R-squared	0.71	0.22		0.65	0.75	
No. individuals	2,927	2,927		245	245	

Robust standard errors in parentheses, clustered by individual. *significant at 10%; **significant at 5%; ***significant at 1%. Sample is person-years from estimation sample (sample selection criteria described in Appendix A), with completed education 9 years or more, from CAPS life history panel. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. All regressions are survey-weighted linear probability models that include individual fixed effects and age, completed schooling, and calendar year dummies. Work experience variables include dummy for non-employment in last period of non-enrollment and cumulative work experience. Vocational programs include technical colleges, NTC and universities of technology. Secondary and post-secondary refer to person-years with completed schooling corresponding to indicated level. Dropout refers to non-enrollment following a period of enrollment. Re-enrollment refers to enrollment after at least one year of non-enrollment.

37% and 2%, respectively, as reported in Table 2, Panel B)[15]. The difference between coefficients across models is also statistically significant for both recent and previous grade failure[16]. This is strong evidence in favor of the safety valve role for vocational schooling: struggling students are deterred from enrollment in academic programs but are not similarly deterred from vocational programs. Columns (2) and (4) add work history to the regressions to check if differential employment opportunities are driving the results, but there is almost no change in the failure coefficients. I find that non-employment in the previous period out of school increases the likelihood of enrollment in both academic and vocational programs, but by a greater amount (7 percentage points to 2) for academic programs. Thus schooling of both types serve a safety valve function for youth who were unsuccessful in the labor market.

Panel B splits the sample between secondary and post-secondary levels, to explore whether the safety valve role is driven by one level only. I omit work experience variables from these and all subsequent regressions in the table for simplicity, although their inclusion leaves the grade failure coefficients almost unchanged. For secondary schooling (columns 1-2), the safety valve results are repeated with similar magnitudes as for the full sample, with a 44 percentage point decline in academic enrollment in response to grade failure compared to a 2 percentage point decline for vocational. Moreover, each previous grade failed leads to a 22 percentage point decrease in academic secondary enrollment, versus no significant change for vocational. For post-secondary schooling, academic enrollment is also more responsive to grade failure than vocational enrollment, although here the discrepancy is statistically significant only for previous grades failed. The contrasting results probably arise because academic and vocational post-secondary programs are more similar to each other in entrance requirements and rigor than are secondary programs. Nonetheless, the safety role of vocational schooling exists to some extent at both levels.

In Panel C, I examine enrollment choices separately by last enrollment status. Columns (1)-(2) condition on enrollment in the previous period, so that non-enrollment represents dropping out of the program. Columns (3)-(4) condition on non-enrollment in the previous period, so that the outcome is re-enrollment following dropout. Again I find strong support for the safety valve hypothesis. Grade failure leads to dropout from an academic program, with similar magnitudes to those from the full sample. However, grade failure does not result in dropout from vocational programs. Analysis of re-enrollment yields qualitatively similar results: struggling students are much less likely to re-enroll in academic programs, with the coefficient on previous failure even larger than for the full sample. Grade failure is now a significant negative predictor of re-enrollment in vocational schooling, but the coefficients are less than one-sixth the magnitude of those for academic schooling.

I analyze enrollment transitions within and between programs in Panel D by conditioning on the type of program in which a student was last enrolled. In columns (1)-(2), students whose last enrollment was in an academic program respond to grade failure in much the same manner as the full sample, further strengthening the safety valve hypothesis. In column (4), however, students whose last enrollment was in a vocational program are deterred from enrollment in vocational programs in response to grade failure to a similar extent as the analogous results for academic programs in column (1). This result reveals a limitation of the safety valve role for vocational schooling: it offers a welcoming

environment for students who struggle in the academic system, but students who struggle within the vocational system tend not to continue.

One concern with the results presented in Table 4 is that they treat all grade failure as equivalent regardless of schooling type. Instead, failure in an academic program may send a different signal than failure in a vocational program. I address this possibility, as well as other potential shortcomings in the baseline specification, in the robustness checks presented in Section 4.3.

4.1.2   Do students experiencing shocks or forming households select into vocational schooling?

Another way in which vocational schooling could serve as a safety valve is by deterring fewer students who experience negative household shocks, such as a household member's job loss or death. However, in regressions (linear probability models) of academic or vocational enrollment on indicators of contemporaneous and lagged negative household shocks, such shocks do not lead to significant changes in either academic or vocational enrollment. This is true for shocks of all types and separately for health and financial shocks. For positive shocks like an increase in household income, youth might respond by choosing vocational schooling less frequently because its safety valve function is less necessary. Following the literature that finds that public pension receipt in South Africa influences decisions by other household members (Case 2004; Duflo 2003; Edmonds 2006; Souza 2010), I used a regression discontinuity design that specifies enrollment as a function of a grandmother or grandfather's pension eligibility and a polynomial in grandparent ages. I found no statistically significant association between household pension eligibility and enrollment in either academic or vocational programs. Results of both these exercises are omitted for brevity but available upon request.

Despite this lack of clear evidence on the safety valve role of vocational schooling in response to household shocks, the results of Table 4 are highly robust to the inclusion of household shocks as controls. Table A1 presents results, in which the baseline specification of column (1) is augmented with inclusion of the shocks considered above in columns (2)-(8). Coefficients remain stable, with significant differences between academic and vocational enrollment throughout, suggesting that it is revisions to expected ability resulting from grade failure, and not shocks, that drive the safety valve result.

In addition to negative shocks, household formation can lead to interruptions in human capital investment, making it more difficult for students, particularly women, who choose to marry or have children to succeed in the labor market. Vocational schooling could also serve as a safety valve by being particularly welcoming to female students who are expanding their families. I explore this possibility in Table 5 by limiting the sample to females and regressing academic or vocational enrollment on pregnancy and marriage events (i.e., getting married rather than marital status) and their lags, along with individual fixed effects, and dummies for age, completed schooling, and calendar year.

In column (1) of Table 5, I find that females who become pregnant are 9 percentage points less likely to enroll in academic programs in the year of the pregnancy, and 12 percentage points less likely the year following pregnancy, with coefficients significant at the 1% level. By contrast, in column (2) I find corresponding declines in vocational enrollment of just 1 percentage point each, with statistical significance only for lagged pregnancy. The

Table 5 Female enrollment and household formation, by program type

| Program type | Academic | Vocational | Dropout only | | Re-enrollment only | |
| | | | Academic | Vocational | Academic | Vocational |
	(1)	(2)	(3)	(4)	(5)	(6)
Pregnancy	-0.09	-0.01	-0.18	-0.02	-0.01	-0.01
	(0.03)***	(0.01)	(0.05)***	(0.01)	(0.02)	(0.01)
Pregnancy ($t-1$)	-0.12	-0.01	-0.31	-0.02	-0.01	-0.02
	(0.03)***	(0.004)***	(0.06)***	(0.01)*	(0.02)	(0.01)**
Marriage	-0.04	0.004	-0.09	0.01	-0.02	0.001
	(0.02)	(0.01)	(0.16)	(0.08)	(0.01)	(0.004)
Marriage ($t-1$)	0.06	0.02	0.16	0.02	0.02	0.02
	(0.03)**	(0.02)	(0.16)	(0.02)	(0.03)	(0.02)
Observations	14,053	14,053	7,071	7,071	6,982	6,982
R-squared	0.65	0.24	0.60	0.33	0.64	0.38
No. individuals	1,700	1,700	1,692	1,692	1,567	1,567
p-value of difference						
Pregnancy		0.00		0.00		0.96
Pregnancy ($t-1$)		0.00		0.00		0.52
Marriage		0.16		0.52		0.15
Marriage ($t-1$)		0.13		0.35		0.96

Robust standard errors in parentheses, clustered by individual. *significant at 10%; **significant at 5%; ***significant at 1%. Sample is person-years from estimation sample described in Appendix A, females only, with completed education 9 years or more, from CAPS life history panel. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. All regressions are survey-weighted linear probability models that include individual fixed effects and age, completed schooling, and calendar year dummies. Vocational programs include technical colleges, NTC and universities of technology. Dropout refers to non-enrollment following a period of enrollment. Re-enrollment refers to enrollment after at least one year of non-enrollment.

coefficients on pregnancy differ significantly from each other between the models for academic and vocational enrollment. Thus I find that pregnancy is associated with schooling interruptions, but much less so in vocational than academic programs. The only marriage coefficient that is significant is lagged marriage in column (1), which is associated with a 6 percentage point increase in academic enrollment. These patterns are broadly consistent with a safety valve role for vocational schooling among females who are forming households, although recently married women are also comfortable in the traditional academic system. Decomposing the results by dropout and re-enrollment choices in columns (3)-(6), I find that the results are driven by dropout behavior. This is an important result, as vocational schooling appears to allow pregnant girls fewer interruptions in their studies than academic schooling. Enrollment, marriage and fertility are often decided simultaneously, of course, so I hesitate to interpret the coefficient estimates causally. However, the results do answer whether enrollment and household formation decisions differ by type of schooling, thereby offering further insight into the safety valve hypothesis[17].

4.2 Testing the sinkhole hypothesis

Although I have found evidence supportive of the safety valve hypothesis for vocational schooling, a natural follow-up question is whether vocational schooling facilitates good subsequent outcomes for its students. Even if vocational schooling is a safety valve for those currently enrolled, is it a sinkhole when it comes to labor market returns?

Table 6 presents results of regressions estimating the wage and employment returns to schooling, differentiated by type of program (academic or vocational) and level (secondary or post-secondary). These are variations on standard Mincer regressions, and

Table 6 Labor market outcomes and schooling type

	(Log) wages			Employment		
	(1)	(2)	(3)	(4)	(5)	(6)
Schooling	0.09			0.05		
	(0.01)***			(0.004)***		
Academic		0.09			0.05	
		(0.01)***			(0.004)***	
Vocational		0.10			0.10	
		(0.04)**			(0.03)***	
Academic secondary			0.08			0.04
			(0.01)***			(0.004)***
Academic post-secondary			0.13			0.11
			(0.03)***			(0.01)***
Vocational secondary			0.06			0.02
			(0.10)			(0.05)
Vocational post-secondary			0.16			0.19
			(0.08)**			(0.05)***
Observations	3,840	3,840	3,840	14,454	14,454	14,454
R-squared	0.27	0.27	0.27	0.21	0.21	0.22
No. individuals	1,676	1,676	1,676	2,811	2,811	2,811
H_0: academic=vocational		0.69			0.09	
H_0: academic=vocational (secondary)			0.80			0.79
H_0: academic=vocational (post-secondary)			0.74			0.10
U/O: schooling	3.8			-4.5		
U/O: academic		3.9			-4.6	
U/O: vocational		1.8			-2.9	
U/O: academic secondary			2.7			13.5
U/O: academic post-secondary			-6.2			-1.5
U/O: vocational secondary			0.3			2.1
U/O: vocational post-secondary			-1.3			-2.0

Robust standard errors in parentheses, clustered by individual. *significant at 10%; **significant at 5%; ***significant at 1%. Unit of observation is person-year. Sample restricted to person-years with at least 9 years completed schooling and not currently enrolled. Wages are log of monthly full-time equivalent wages based on 160 hours/month. Wages refer to self-reported earnings on current or most recent jobs held among those not enrolled in school, regardless of current employment status, FTE based on 160 hours worked/month. Employment restricted to periods when not enrolled in school. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. All regressions include race, gender, and calendar year dummies; and controls for age-adjusted LNE *z*-score, experience, experience squared, age, and age squared. Vocational schooling includes NTC and other technical secondary, and universities of technology. Reported test statistics are *p*-values of indicated null hypothesis. U/O ratio: ratio of selection on unobservables/observables required to explain away effect of schooling variable on outcome, obtained as $\frac{\beta_C}{\beta_{NC} - \beta_C}$, where β_C is schooling coefficient from model with included controls and β_{NC} is schooling coefficient from model without controls. Controls as indicated and described above. See Appendix B for details. Survey weights applied.

therefore come with the usual caveats against a causal interpretation of coefficients. Indeed, in the preceding sections I have documented the dimensions of self-selection into academic and vocational schooling. However, to the extent that the previous results show that vocational students are adversely selected relative to their peers in academic programs, it would be especially notable to find high Mincerian returns to vocational schooling.

Columns (1)-(3) of Table 6 present regressions of log wages on years of schooling. When scaled by 100, coefficients represent the approximate percentage return to an additional year of schooling. Because the sample includes only those person-years with at least 9 years schooling and not currently enrolled, the estimated returns refer to schooling beyond the compulsory level, when vocational schooling becomes an option. Data

on wages were collected only in Waves 1-5, resulting in far smaller sample sizes than the safety valve regressions that use life history data. Regressions also include age-adjusted literacy and numeracy *z*-score; work experience, age and their squares; and race, gender, and calendar year dummies. The inclusion of an ability proxy and actual (rather than potential) work experience allow for bias reductions relative to results from a census or labor force survey, which typically lack such measures. Column (1) is the canonical Mincer regression of log wages on overall years of schooling and these controls. I find that each additional year of schooling correlates with wage gains of 9 percent. Column (2) distinguishes between academic and vocational schooling, with vocational returns of 10 percent versus 9 percent for academic; the *p*-value on the difference is 0.69.

Column (3) further distinguishes each type of schooling by level (primary/secondary and post-secondary for academic, secondary and post-secondary for vocational). Within each type of schooling, returns are larger in post-secondary than secondary, consistent with earlier evidence on the convexity of returns to schooling in South Africa (Case and Yogo 1999). The point estimate for vocational secondary returns is both lower than its academic counterpart (6% to 8%) and imprecisely estimated. The return to vocational post-secondary exceeds that for academic post-secondary, 16% to 13%, and is statistically significant at 5%. In neither case are the within-level academic and vocational coefficients significantly different from each other. This suggests that the estimated return for vocational secondary is hampered by lack of statistical power, rather than being a precisely estimated zero. The results show no clear advantage for academic schooling compared to vocational, contrary to the sinkhole hypothesis. The absence of evidence in favor of the sinkhole hypothesis, though not conclusive, is striking because the potentially negative selection into vocational schooling found in the preceding sections of the paper should lead to downward bias in estimates of vocational returns.

The obvious concern with these results is that students may select into levels and types of schooling based on unobserved attributes, leading to inconsistent estimates of schooling returns. To assess the importance of this concern, I examine whether estimates of returns vary substantially when excluding the control variables mentioned above. If these controls help to mitigate bias due to omitted ability or other unobserved characteristics, then finding similar returns to schooling regardless of their inclusion suggests that any remaining omitted variable bias is likely to be small. On the other hand, if estimated returns attenuate substantially upon including the controls, then one has more cause to worry that selection on unobservables drives the results.

To formalize this idea, I follow Altonji et al. (2005) in calculating the ratio of selection on unobservables to selection on observables required to explain away the entire estimated returns for each schooling measure[18]. A ratio greater than one in absolute value indicates that the covariance between unobserved characteristics and schooling must exceed the covariance between included control variables and schooling to explain away the entire return. A negative ratio indicates that the schooling coefficient estimated with controls *increases* relative to the coefficient estimated without controls, a surprising result if we expect the controls to proxy for (upward) ability bias. Altonji et al. (2005) interpret ratios greater than three as large, and ratios greater than one as indicative of a non-zero treatment effect even in the presence of selection on unobservables. I report this ratio for each coefficient as "U/O" in Table 6. Details of the procedure are in Appendix B.

Given the high-quality data of the panel, I expect the included controls to account for much of the selection into schooling, and the results in Table 6 support this claim. In column (2), I find U/O ratios of 3.9 and 1.8 for academic and vocational schooling, respectively, indicating that selection on unobservables is unlikely to be so large as to explain away these coefficients. In column (3), only the ratio for the vocational secondary coefficient is less than one in absolute value, which is not surprising given that this coefficient is not statistically different from zero. The negative U/O ratios for academic and vocational post-secondary reflect increases in these coefficients relative to their counterparts estimated without control variables.

Columns (4)-(6) show regressions analogous to the first three columns, but with employment as the outcome[19]. As these are linear probability models, coefficients (scaled by 100) may be interpreted as percentage point changes in the likelihood of employment for an additional year of schooling. In column (4), I find the employment return to schooling is 5 percentage points. In column (5), I find that the employment return to vocational schooling is 10 percentage points versus 5 percentage points for academic schooling, with the *p*-value of the difference equal to 0.09. When differentiating schooling by level, I find larger returns to post-secondary than secondary schooling within each type of program. As with wages, for employment the point estimate for post-secondary vocational exceeds that for academic while the opposite is true for secondary, although the coefficients are not significantly different from each other. All of the U/O ratios in columns (4)-(6) exceed one in absolute value, suggesting that selection on unobservables does not drive the results. As with wage returns, the sinkhole hypothesis of systematically lower labor market returns for vocational schooling does not find support in the data[20]. Due to lack of precision, I am unable to conclude that wage and employment returns to vocational secondary schooling are strictly positive, however.

In results not shown but available upon request, I have also allowed for non-linearity in returns at particular levels of academic or vocational schooling by including dummies for each completed grade of academic and vocational schooling in the Mincer wage and employment regressions. Although grade level-specific point estimates of academic returns exceed those of vocational for both (log) wages and employment at some levels of schooling, the differences are never statistically significant. These results rule out the presence of differential "sheepskin effects" between academic and vocational schooling, providing further evidence against the sinkhole hypothesis.

Another way to gauge the importance of selection on unobservables is to restrict the analysis of Table 6 to a subsample for whom the sinkhole hypothesis is likely to hold: the negatively selected subsample of those who previously failed a grade. These results (not shown but available upon request) are quite similar to those for the full sample, with no significant differences between academic and vocational schooling coefficients with either log wages or employment as the outcome.

An important caveat for all the labor market results, however, is that the sample is quite young (with maximum age 29), so that the returns we observe correspond to early career outcomes only. Academic schooling may provide general skills whose returns rise faster than those of vocational skills over a worker's lifetime, as shown for OECD countries in Hanushek et al. (2011)[21].

4.3 Robustness checks

One might object that the model of Section 2 treats ability as uni-dimensional when in fact students possess expectations about both their academic- and vocational-specific abilities. If this is the case, then the safety valve regressions in Table 4 are misspecified. A straightforward extension of the model would allow for such multi-dimensional ability (i.e., allowing academic- and vocational-specific expected abilities A_e^a and A_e^v, respectively) and modify the safety valve definition so that vocational enrollment is less responsive to own-ability revisions than is academic enrollment[22].

Table A2 presents regressions testing this extension of the model. In column (1), I find that academic enrollment declines by 40 percentage points in response to a recent academic failure, and an additional 18 percentage points for each previous academic failure. Additionally, academic enrollment falls 9 percentage points in response to a recent vocational failure, suggesting that students who fare poorly in either type of program tend to leave academic schooling. Coefficients barely change when work experience controls are added in column (2). By contrast, in column (3) I find that although vocational enrollment does decline in response to a recent vocational failure, the magnitude (17 percentage points) is less than half that of a recent own-program failure in column (1). Previous vocational failures have no statistically significant effect on vocational enrollment. Moreover, each previous academic failure leads to a 1 percentage point *increase* in vocational enrollment. Results are again robust to inclusion of work experience in column (4). Differences across models are significant only for academic failure coefficients.

Next, I explore heterogeneity in the enrollment response of academic and vocational schooling to grade failure in Table A3. Specifically, I regress enrollment on previous schooling and labor market outcomes, analogous to columns (1) and (3) of Table 4, Panel A, separately for students who attain secondary/post-secondary levels, males and females, blacks and coloureds, and the bottom two/top three quintiles of household income. Panel A of Table A3 shows the academic enrollment response, while Panel B shows the vocational enrollment response.

In Panel A, columns (1)-(2), I split the sample by whether the student's final attainment is secondary or post-secondary, to explore whether students who achieve at different levels respond to failure differently (note that this differs from Table 4, Panel B, which splits the sample by secondary and post-secondary person-years; here I split the sample by individuals). I find that students who failed their last grade are less likely to enroll in academic programs regardless of whether their eventual attainment is the secondary or post-secondary level. The magnitude of this response is considerably higher for those who stop at the secondary level: a 49 percentage point decline for secondary, compared to 16 percentage points for post-secondary. The discrepancy also appears in the coefficients for previous grade failure. Looking at the same columns in Panel B, we see that students who have failed are not similarly deterred from enrolling in vocational programs, with only a 2 percentage point drop in vocational enrollment among those who stop at secondary schooling and no significant decline for students who go on to post-secondary. The safety valve role of vocational programs exists for students who attain both major levels of schooling.

In columns (3)-(8) of Table A3, I find similar patterns in the enrollment response to failure across genders, races, and economic background: grade failure leads to sizable

and statistically significant declines in academic enrollment, but little to no effect on vocational enrollment. Within each subsample, differences between coefficients in the academic and vocational enrollment equations are significant at the 1% level. Vocational schooling is a safety valve for both males and females, blacks and coloureds, and youth from the poorest and least poor households.

Given this robustness in the safety valve role of vocational schooling for enrollment choices, it is also sensible to explore heterogeneity in labor market outcomes among these subsamples. Table A4 presents Mincer regressions for wages and employment, respectively, analogous to columns (2) and (4) of Table 6, separately for black and coloured youth. In columns (1)-(2) of Table A4, I find no significant differences between wage returns for vocational and academic schooling for both black and coloured youth. In columns (3)-(4), I find that vocational schooling generates larger employment returns for youth of both races. For coloureds the difference is striking: an additional year of vocational schooling is associated with an 8 percentage point increase in the probability of employment, more than one and a half times as large as an additional year of academic schooling (with the difference significant at 10%). For black youth, vocational schooling does not appear to be a sinkhole—vocational wage and employment returns are not significantly different from academic—but they do not reap the employment *gains* accruing to coloureds. Selection on unobservables is unlikely to explain any of these results: the ratio of selection on unobservables to observables required to explain away the schooling coefficients is often negative—indicating that returns increase upon including observable characteristics—or otherwise greater than one.

Given the importance of social networks in securing employment for South African youth (Magruder 2010), one potential mechanism explaining this result is the greater success enjoyed by coloured youth in using vocational schooling to acquire jobs through formal means rather than informal networks[23]. If youth need credible evidence of skills to access jobs through formal channels rather than relying on social networks, vocational schooling might provide such a credential. In columns (5)-(6), I explore this possibility by regressing an indicator for job acquisition through formal channels (defined as responding to a job advertisement, sending a CV, visiting a firm, or using an employment agency, as opposed to using informal social networks) on years of schooling of each type among the sample of employed youth. Sample sizes drop substantially in these regressions because information on search method is available only in the concurrent waves of the panel, rather than in retrospective life histories, and because I restrict the sample to employed youth in order to avoid conflating the employment returns found in column (3)-(4) with the search method. The results reveal that the marginal increase in success rates from formal search associated with academic schooling is similar for black and coloured youth (6 versus 4 percentage points). However, each year of vocational schooling is associated with a 22 percentage point increase in successful formal search for coloured youth, compared to a 6 percentage point (though not significant) decrease for blacks. Paired with the results from columns (3)-(4), this result suggests that employers view black youth with vocational schooling less favorably than otherwise similar coloureds when evaluating job candidates. Vocational schools catering to black youth may do well to place extra emphasis on job placement to rectify this disparity.

5 Conclusion

This paper analyzes the role of vocational schooling in the school to work transition of South African youth. I find evidence that vocational schooling acts as a safety valve for students who struggle academically: grade failure leads to modest decreases in vocational enrollment, compared to substantial declines in academic enrollment. Vocational schooling also serves as a safety valve for pregnant girls, in contrast to academic schooling, which exhibits large enrollment declines following pregnancy. Household shocks have no statistically significant effects on enrollment in either type of schooling, however. The safety valve hypothesis is robust to allowing for multi-dimensional ability, with own-program failure leading to smaller enrollment declines for vocational than academic schooling. The results are present in both secondary and post-secondary schooling (although somewhat muted for the latter) and persist across a number of subsamples of interest.

Given this evidence for the safety valve role for vocational schooling, one might worry that an associated cost is lower labor market returns compared to academic schooling, i.e., that vocational schooling is a sinkhole. Estimates of wage and employment returns fail to provide evidence for the sinkhole hypothesis, however. Returns estimates for vocational secondary are hampered by lack of statistical power, but are nonetheless positive and not statistically different from those for academic secondary. Although I have shown that these estimated returns are unlikely to be explained away by selection on unobserved characteristics, they come with the usual caveats against the causal interpretation of Mincer regressions, and the additional caveat that the relative returns may shift in favor of academic schooling as workers age (Hanushek et al. 2011). Even so, the absence of evidence favoring academic schooling is notable in this context, considering the low esteem South Africans hold for vocational schooling and the evidence in this paper that its students are adversely selected.

If vocational schooling has the salutary effects that I find, particularly for struggling students, why don't more students enroll? At the post-secondary level, limited open seats and relatively strict admissions requirements at universities of technology are likely barriers to expanded enrollment. At the secondary level, qualitative evidence suggests that stigma and information costs limit student enrollment in vocational programs, drawing students who feel they have no other choice (Needham and Papier 2011). If this is the case, then efforts by school administrators to promote vocational programs as a practical alternative to the traditional system should yield benefits.

South African policymakers frequently express consternation about slow student progression through schools and high rates of youth unemployment, with economists finding ample evidence to support such concerns. The evidence presented here suggests that vocational schooling, a small and often overlooked educational institution, plays an important role for vulnerable students both while they remain in school and when they enter the labor market.

Endnotes

[1]Based on data for 15-24 year olds in 2005, compiled in Banerjee et al. (2008). Unemployment follows the "narrow" definition of the International Labor Organization (ILO), which classifies working age individuals as being in the labor force if during a week of reference they were employed or wanted to work and were available to start working within a week but also had actively looked for work during the past four weeks.

[2]Several recent studies (Attanasio et al. 2011; Card et al. 2011; Hicks et al. 2011; Maitra and Mani 2012; Macours et al. 2013; Blattman et al. 2013), report results from randomized control trials of vocational training programs in developing countries (Colombia, Dominican Republic, India, and Nicaragua, respectively), with mixed results for employment and wage returns. Although the nature of the evidence in these studies is compelling, the programs under study are short-term job training programs implemented outside the public education system, making them not strictly comparable to the South African vocational programs that I consider in this paper.

[3]A recent South African newspaper article put it more bluntly with the headline, "Vocational education not an 'option for losers' " (Nkosi 2012).

[4]Without loss of generality, I assume that each agent has acquired sufficient prior schooling to make both types of schooling an option. In South Africa, students who have completed 9 years of schooling may choose to continue in the academic system or pursue vocational study.

[5]Although there is evidence that returns to schooling in South Africa are convex (Case and Yogo 1999), the concavity assumption is made for convenience here to ensure an interior solution, conditional on the agent choosing some positive amount of investment in a schooling type.

[6]I frame the problem as static for simplicity of exposition. The model could be made dynamic in a straightforward manner, adapting the dynamic discrete choice framework of (Pugatch 2012).

[7]If academic and vocational schooling and the occupations associated with them require distinct skills, then finding $\beta_v < \beta_a$ need not indicate that vocational students are poorly served by their choice. However, vocational students frequently began their careers at comparable levels of academic schooling, with 97% of vocational secondary students and 39% of vocational post-secondary students previously enrolled in academic programs in each level. This suggests that students are experimenting with different types of programs and should be concerned about their comparative returns.

[8]Note that the presence of the fixed cost makes it possible to observe an equilibrium in which the marginal returns to schooling type v exceed those of type a, but many students still optimally choose not to invest in type v schooling.

[9]In my sample, only 3% report ever working and attending school simultaneously. This figure rises to just 5% among those who ever enrolled in a vocational program.

[10]I also exclude person-years corresponding to Wave 5 interviews suspected of fraudulent data collection. See Appendix A for details.

[11]"Coloured" refers to people of mixed racial heritage and descendants of Cape Malay slaves; the term is in common use among South Africans of all races. Coloured people occupied an intermediate position under apartheid in which they faced major restrictions on civil, political and economic activity, but not as severe as the majority black population.

[12]Schools of all types are free to set their own admissions criteria and fees, but may not discriminate on the basis of race, test scores, or the ability to pay (Fiske and Ladd 2004). Reported school fees are considerably higher for vocational secondary programs than academic, which may partly explain low vocational secondary enrollment. At the post-secondary level, there are no appreciable differences in reported fees between academic and vocational programs. The regression of (log) school fees on enrollment in various programs yields: $\ln fee = \underset{(184.5)}{5.78} + \underset{(8.1)}{1.81} vocsec + \underset{(43.5)}{2.91} acadpostsec + \underset{(32.0)}{2.86} vocpostsec$, where *vocsec*, *acadpostsec*, and *vocpostsec* are dummies for enrollment in vocational secondary, academic post-secondary, and vocational post-secondary programs, respectively; enrollment in academic upper secondary (grades 10-12) is the omitted category. *t*-statistics from standard errors clustered by individual in parentheses.

[13]I use linear probability models because their coefficients may be interpreted as marginal effects of changes in individual covariates, which is my focus in this paper.

Results from conditional (fixed effects) logit estimation of models analogous to those in this section yielded similar findings.

[14]Direct measures of shocks are excluded from these regressions because data limitations would reduce the sample size and hamper the identification of individual fixed effects. I examine the direct effect of shocks on enrollment and robustness of grade failure measures to their inclusion in Section 4.1.2.

[15]Although the response to grade failure relative to mean enrollment in a schooling program is informative, the safety valve hypothesis concerns the absolute response only. Moreover, inclusion of individual fixed effects serves to control for an individual's mean enrollment rate in each program.

[16]To conduct these tests, I correct the variance-covariance matrix of coefficients from paired academic/vocational enrollment regressions to account for correlation between error terms. The procedure takes advantage of the equivalence between seemingly unrelated regression (SUR) and equation-by-equation OLS when the equations have identical explanatory variables (Greene 2012).

[17]Table A1, column (9) verifies that the safety valve results for grade failure are not driven by household formation for females.

[18]Intuitively, the motivation for the procedure is the assumption that "selection on the unobservables is the same as selection on the observables" (Altonji et al. 2005). Formally, in the Mincer regressions (8) that I run in this paper, the assumption is $cov(S_v, \epsilon) = cov(S_v, X)$, and analogously for S_a. Altonji et al. (2005) argue that this assumption is no more restrictive than the standard OLS assumptions $cov(S_v, \epsilon) = 0$ and $cov(X, \epsilon) = 0$.

[19]The sample for all regressions in Table 6 is restricted to youth not enrolled in school. I do not distinguish between full-time or part-time employment, as information on hours worked is available only in the concurrent waves of the panel, not the life history. Where data on working hours are available, 79% of the employed sample work full-time (defined as an average of at least 35 hours per week).

[20]I also find evidence against the sinkhole hypothesis using pooled, semi-annual samples of the South Africa Labour Force Survey (LFS): vocational schooling provides larger wage and employment returns than academic schooling, with coefficients significantly different from each other at the 1% level. I restrict the sample to black and coloured youth from the Western Cape ages 14-26 observed between September 2002-September 2006 for maximum comparability with CAPS. Unlike CAPS, however, the LFS lacks data on vocational post-secondary schooling, work experience and ability. Results not shown but available upon request.

[21]Auxiliary analysis using a sample of workers aged 16-55 from the South African Labour Force Survey suggest that this is the case in South Africa as well. Younger workers with vocational schooling earn an experience premium relative to observationally equivalent workers with academic schooling. However, the returns to experience for workers with academic schooling overtakes that of vocationally-trained workers at age 44 for wages, 39 for employment. Results not shown but available upon request.

[22]This extension would make the framework similar to a Roy (1951) model in which the agent has expectations over academic- and vocational-specific ability endowments. The agent would choose the schooling track that offers higher expected return, with revisions in expected returns due to failure altering choices for those at the margin.

[23]Note that examining job acquisition through formal versus informal channels is not equivalent to the question of whether the jobs themselves are in the formal or informal sector. South Africa is notable for its low rates of informality (Banerjee et al. 2008) and lack of small firms (Magruder 2012). Unfortunately the data are not well suited to explore this question because information that could be used to determine formality of job sector is collected only in CAPS Wave 2.

[24]The Cape Area Panel Study Waves 1-2-3 were collected between 2002 and 2005 by the University of Cape Town and the University of Michigan, with funding provided by

the US National Institute for Child Health and Human Development and the Andrew W. Mellon Foundation. Wave 4 was collected in 2006 by the University of Cape Town, University of Michigan and Princeton University. Major funding for Wave 4 was provided by the National Institute on Aging through a grant to Princeton University, in addition to funding provided by NICHD through the University of Michigan. Wave 5 was collected in 2009 by the University of Cape Town. Major funding for Wave 5 was provided by the Health Economics & HIV/AIDS Research Division (HEARD) at the University of KwaZulu-Natal, with additional funding from the Andrew W. Mellon Foundation (through the CSSR at UCT), the European Union (through the Microcon research partnership on the microfoundations of violent conflict, via the CSSR) and the NICHD (through the University of Michigan).

[25]A total of 289 young adult interviews in Wave 5 are suspected of occurring under fraudulent conditions based on an investigation conducted by CAPS. These observations are flagged in the data. Individuals with Wave 5 interviews suspected of fraud account for 5% of my estimation sample.

[26]This derivation of omitted variable bias follows (Wooldridge 2001) closely.

Appendix
A Data definitions

The data come from retrospective life histories collected in Wave 1 of the Cape Area Panel Study (CAPS), augmented with events recorded in Waves 2-5[24]. The Wave 1 retrospective life histories record events by youth's age, where age refers to the age at which the event occurred in the case of living arrangements and marriage, and to age at the beginning of the calendar year in the case of enrollment and progression through school, labor force participation, and pregnancy. I follow this convention in mapping Wave 2-5 responses to youth's age.

I make several sample restrictions. I exclude whites. I keep only those observed until at least age 18. Those who report advancing two or more grades in a year, or without continuous information on enrollment, are dropped from the sample. I drop those whose educational histories, by the definitions below, place them with less than 9 or more than 16 years of completed schooling. I exclude person-years corresponding to Wave 5 for which fraudulent fieldwork was suspected[25]. Excluding whites and those not observed until age 18 accounts for 56% of the observations dropped from the sample.

Schooling level covers grades 1-16, with National Technical Certificate (NTC) 1, 2, and 3 mapped to grades 10, 11, and 12, respectively. Students enrolled in university or university of technology programs that include grade 12 are considered enrolled in grade 12. I define "vocational" and "academic" schooling as a set of mutually exclusive categories. Vocational schooling includes enrollment in National Training Certificate, university of technology or other technical program that does not require matric (secondary completion); enrollment in all other programs, including traditional secondary schools, universities, and diploma or certificate programs that do require matric, is academic schooling.

Reporting successful completion of the grade level or reporting enrollment in a higher grade level in a subsequent year is considered passing the level for grades 1-12. Beginning at grade 13, reporting successful completion of the grade level or "no grade/continuing" are considered passing the level, up to a maximum of 16 years completed schooling.

Table A1 Enrollment response to grade failure by program type: robustness to other shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: academic									
Failed last	-0.40	-0.48	-0.42	-0.48	-0.42	-0.48	-0.42	-0.40	-0.40
Grade enrolled	(0.01)***	(0.03)***	(0.04)***	(0.03)***	(0.04)***	(0.03)***	(0.04)***	(0.01)***	(0.02)***
Previous grades	-0.17	-0.22	-0.21	-0.22	-0.21	-0.22	-0.21	-0.17	-0.17
Failed	(0.01)***	(0.03)***	(0.04)***	(0.03)***	(0.04)***	(0.03)***	(0.04)***	(0.01)***	(0.01)***
R-squared	0.70	0.81	0.67	0.81	0.67	0.81	0.67	0.70	0.70
Panel B: vocational									
Failed last	-0.01	-0.02	0.00	-0.02	0.00	-0.02	0.00	-0.01	-0.01
Grade enrolled	(0.005)*	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.005)*	(0.01)
Previous grades	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Failed	0.00	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)
R-squared	0.26	0.55	0.52	0.55	0.52	0.55	0.52	0.26	0.24
Observations	23,997	6,805	6,614	6,805	6,614	6,805	6,614	23,997	14,053
No. individuals	2,927	2,900	2,855	2,900	2,855	2,900	2,855	2,927	1,700
p-value of difference									
Failed last	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grade enrolled									
Previous grades	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Failed									

Robust standard errors in parentheses, clustered by individual. * significant at 10%; ** significant at 5%; *** significant at 1%. Sample is person-years from estimation sample described in Appendix A, with completed education 9 years or more, from CAPS life history panel. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. Vocational programs include technical colleges, NTC and universities of technology. Re-enrollment refers to enrollment after at least one year of non-enrollment. Sample sizes drop in regressions including shocks because these only measured in Waves 1, 3 and 4. All regressions are survey-weighted linear probability models that include individual fixed effects and age, completed schooling, and calendar year dummies. Key to included shocks in each regression, by column:

(1) none.

(2) any household shock, including death; serious illness or injury; job loss; major financial loss (including business failure or bankruptcy); abandonment or divorce; theft, fire or property damage; or other shock.

(3) any household shock ($t-1$).

(4) death, serious illness or injury.

(5) death, serious illness or injury ($t-1$).

(6) job loss; major financial loss; theft, fire or property damage.

(7) job loss; major financial loss; theft, fire or property damage ($t-1$).

(8) pension eligibility (regression discontinuity specification).

(9) pregnancy, marriage and their first lags (females only).

Table A2 Enrollment response to program-specific grade failure and previous employment, by program type

Program type	Academic		Vocational		Differences	
	(1)	(2)	(3)	(4)	(3)-(1)	(4)-(2)
Failed last academic grade enrolled	-0.40	-0.40	0.000	0.000	***	***
	(0.01)***	(0.01)***	(0.004)	(0.004)		
Previous academic grades failed	-0.18	-0.18	0.01	0.01	***	***
	(0.01)***	(0.01)***	(0.004)**	(0.004)*		
Failed last vocational grade enrolled	-0.09	-0.08	-0.17	-0.17		
	(0.04)**	(0.04)**	(0.04)***	(0.04)***		
Previous vocational grades failed	-0.02	-0.01	-0.03	-0.03		
	(0.03)	(0.03)	(0.08)	(0.08)		
Not employed last period		0.06		0.02		***
		(0.01)***		(0.004)***		
Observations	23,997	23,997	23,997	23,997		
R-squared	0.70	0.70	0.27	0.27		
No. individuals	2,927	2,927	2,927	2,927		
Includes work experience		x		x		

Robust standard errors in parentheses, clustered by individual. *significant at 10%; **significant at 5%; ***significant at 1%. Sample is person-years from estimation sample described in Appendix A, with completed education 9 years or more, from CAPS life history panel. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. All regressions are survey-weighted linear probability models that include individual fixed effects and age, completed schooling, and calendar year dummies. Work experience variables include dummy for non-employment in last period of non-enrollment and cumulative work experience. Vocational programs include technical colleges, NTC and universities of technology.

I make this distinction because "no grade/continuing" is the modal response for those enrolled in the post-secondary education sector, indicating that most youth are continuing in their programs of higher education, whereas "passing" reports at these levels drop considerably. Unfortunately, I am unable to determine whether students are making satisfactory progress towards post-secondary degree completion. All other results while enrolled are considered failure. I define "dropout" as non-enrollment following a year of enrollment, and "re-enrollment" as enrollment following a year of non-enrollment. Grades failed represent the accumulation of periods of enrollment in which the agent did not pass the grade, and therefore may include events such as withdrawal, illness or residential moves rather than outright academic failure.

Employment, wages, and work experience are conditional on non-enrollment at a given age, where reports of enrollment supersede reports of labor market participation. Only 3% of individuals in the sample ever worked and enrolled in school simultaneously. Wages are full-time annual equivalent based on 160 working hours per month (those reporting monthly hours above 160 are considered full-time and do not receive an adjustment). Wages are available only at the time of the interview, rather than as retrospective histories, resulting in diminished sample sizes relative to employment histories. Wages refer to self-reported earnings on current or most recent jobs held among those not enrolled in school, regardless of current employment status. Wages are denoted in real South African rand per year (base year 2002), unless otherwise noted. The South African rand traded at 10.3 per US dollar in August 2002 when CAPS Wave 1 began.

Other covariates are largely self-explanatory. Ability quantiles refers to full sample (i.e., before all restrictions imposed) rank of age-adjusted score on the literacy and numeracy evaluation (LNE) administered to all CAPS respondents in Wave 1. Household income

Table A3 Enrollment response to grade failure and previous employment, subsamples by program type

Subsample	Secondary	Post- secondary	Male	Female	Black	Coloured	Poorest two quintiles	Richest three quintiles
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: academic								
Failed last	-0.49	-0.16	-0.39	-0.41	-0.36	-0.41	-0.44	-0.35
Grade enrolled	(0.016)***	(0.019)***	(0.020)***	(0.016)***	(0.015)***	(0.019)***	(0.016)***	(0.019)***
Previous grades	-0.24	-0.03	-0.18	-0.17	-0.15	-0.18	-0.19	-0.15
Failed	(0.014)***	(0.012)**	(0.016)***	(0.013)***	(0.013)***	(0.015)***	(0.014)***	(0.014)***
R-squared	0.73	0.69	0.69	0.70	0.68	0.71	0.70	0.70
Panel B: vocational								
Failed last	-0.02	0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
Grade enrolled	(0.005)***	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.006)*	(0.01)
Previous grades	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00
Failed	(0.00)	(0.01)	(0.007)*	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
R-squared	0.23	0.26	0.28	0.24	0.24	0.27	0.25	0.27
Observations	11,846	12,151	9,944	14,053	11,625	12,372	12,185	11,812
No. individuals	1,602	1,325	1,227	1,700	1,522	1,405	1,564	1,363
p-value of difference								
Failed last	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grade enrolled								
Previous grades	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Failed								

Robust standard errors in parentheses, clustered by individual. *significant at 10%; **significant at 5%; ***significant at 1%. Sample is person-years from estimation sample described in Appendix A, with completed education 9 years or more, from CAPS life history panel. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. Regressions for secondary include only individuals with final attainment of 9-11 years schooling, while regressions for post-secondary include only individuals with final attainment of at least 12 years schooling. Poorest and richest quintiles based on household per capita income in Wave 1. All regressions are survey-weighted linear probability models that include individual fixed effects and age, completed schooling, and calendar year dummies. Vocational programs include technical colleges, NTC and universities of technology. Re-enrollment refers to enrollment after at least one year of non-enrollment.

Table A4 Schooling type and labor market outcomes, by race

Outcome	(log) wages		Employment		Formal search	
Race	Black	Coloured	Black	Coloured	Black	Coloured
	(1)	(2)	(3)	(4)	(5)	(6)
Academic	0.08	0.09	0.04	0.05	0.06	0.04
	(0.02)***	(0.01)***	(0.01)***	(0.01)***	(0.03)**	(0.02)**
Vocational	0.17	0.10	0.06	0.13	-0.06	0.22
	(0.01)*	(0.04)**	(0.03)*	(0.04)***	(0.18)	(0.10)**
Observations	1,388	2,452	6,576	7,878	449	848
R-squared	0.18	0.22	0.16	0.25	0.07	0.06
No. individuals	678	998	1,437	1,374	356	676
H_0: academic=vocational	0.34	0.79	0.53	0.07	0.51	0.07
U/O: academic	8.0	2.9	5.2	-3.3	3.6	1.6
U/O: vocational	-2.8	1.2	-2.2	-2.4	-1.7	56.9

Robust standard errors in parentheses, clustered by individual. * significant at 10%; ** significant at 5%; *** significant at 1%. Unit of observation is person-year. Sample restricted to person-years with at least 9 years completed schooling and not currently enrolled. Excludes person-years corresponding to Wave 5 for which fraudulent fieldwork suspected. Wages are log of monthly full-time equivalent wages based on 160 hours/month. Wages refer to self-reported earnings on current or most recent jobs held among those not enrolled in school, regardless of current employment status, FTE based on 160 hours worked/month. Employment restricted to periods when not enrolled in school. Dependent variable in formal search regressions is indicator for job acquisition through formal method. Formal method of job acquisition refers to responding to ad, sending CV, visiting firm, or using employment agency. All regressions include race, gender, and calendar year dummies; and controls for age-adjusted LNE z-score, experience, experience squared, age, and age squared. Vocational schooling includes NTC and other technical secondary, and universities of technology. Reported test statistics are p-values of indicated null hypothesis. U/O ratio: ratio of selection on unobservables/observables required to explain away effect of schooling variable on outcome, obtained as $\frac{\beta_C}{\beta_{NC}-\beta_C}$, where β_C is schooling coefficient from model with included controls and β_{NC} is schooling coefficient from model without controls. Controls as indicated and described above. See Appendix B for details. Survey weights applied.

quintiles are derived from the distribution of household per capita income reported in Wave 1 of CAPS. The pension-eligible grandfather/grandmother indicator is set to one if there is at least one co-resident grandmother/grandfather on the household roster who is of pension-eligible age (60 for females, 65 for males) in the relevant young adult person-year. Grandparent ages are determined by extrapolating from ages reported in the most recent panel wave, or Wave 1 for the retrospective histories. Co-residence is reported in each year of the life histories. The variable definition also makes note of changes in the household roster reported in Waves 2-5.

B The sinkhole hypothesis and selection on unobservables

To investigate the sinkhole hypothesis that vocational schooling generates lower labor market returns than vocational schooling, I regress a labor market outcome y on S_a and S_v, the stocks of academic and vocational schooling, respectively. The parameters of interest are β_a and β_v, the coefficients on academic and vocational schooling. The concern with such regressions is that individuals will select into academic and vocational schooling non-randomly, leading to inconsistent estimates of β_a and β_v.

To formalize this concern, suppose there is an index u that captures this selection effect, so that the true model (with constant term and covariates as in (8) omitted for ease of exposition) is[26]:

$$y = \beta_a S_a + \beta_v S_v + \gamma u + \epsilon \tag{9}$$

Omitting u from the regression means that the composite error term $\gamma u + \epsilon$ will be correlated with the regressors of interest. Focusing on the consequence for estimating β_a,

consider the linear projection of u on S_a and S_v:

$$u = \delta_a S_a + \delta_v S_v + r \tag{10}$$

Substituting (10) into (9) yields:

$$y = (\beta_a + \gamma \delta_a) S_a + (\beta_v + \gamma \delta_v) S_v + \gamma r + \epsilon \tag{11}$$

Written in this form, the familiar omitted variables bias formula for $\hat{\beta}_a$ follows:

$$\text{plim } \hat{\beta}_{a,NC} = \beta_a + \gamma \delta_a \tag{12}$$

where the subscript NC denotes "no controls".

Now suppose there is a set of controls X that is related to the selection term u in the following way:

$$u = X \beta_X + \tilde{u}$$

The availability of these controls allows me to rewrite the model as:

$$y = \beta_a S_a + \beta_v S_v + \gamma (X \beta_X + \tilde{u}) + \epsilon \tag{13}$$

Analogously to (10), write the linear projection of the omitted variable $\tilde{u}$ on the observables as:

$$\tilde{u} = \theta_a S_a + \theta_v S_v + X \theta_X + v \tag{14}$$

Substituting (14) into (13) yields the counterparts to (11) and (12):

$$y = (\beta_a + \gamma \theta_a) S_a + (\beta_v + \gamma \theta_v) S_v + X(\beta_X + \gamma \theta_X)$$
$$+ \gamma v + \epsilon \tag{15}$$

$$\text{plim } \hat{\beta}_{a,C} = \beta_a + \gamma \theta_a \tag{16}$$

where C refers to the model with controls.

Following (Bellows and Miguel 2009) application of (Altonji et al. 2005) methodology, taking the ratio of (16) over the difference in the probability limits (12)-(16) under the null hypothesis $\beta_a = 0$ yields:

$$\text{plim } \frac{\hat{\beta}_{a,C}}{\hat{\beta}_{a,NC} - \hat{\beta}_{a,C}} = \frac{\theta_a}{\delta_a - \theta_a}$$

The numerator on the right-hand side is the partial correlation between academic schooling S_a and $\tilde{u}$, the remaining selection term after including the controls X. The denominator is this term subtracted from the partial correlation between academic schooling S_a and u, the selection term that includes the controls X. If the denominator is small relative to the numerator and the controls X are representative of all possible controls (Altonji et al. (2005) formalize this argument), then the selection term remaining after controls are included ($\tilde{u}$) is not likely to be an important source of inconsistency for β_a.

The interpretation of this ratio becomes somewhat clearer if I further assume that $\delta_v = \theta_v = \theta_X = 0$, in which case we have:

$$\delta_a = \frac{cov(S_a, u)}{var(S_a)}$$

$$\theta_a = \frac{cov(S_a, \tilde{u})}{var(S_a)}$$

$$\text{plim}\left(\hat{\beta}_{a,NC} - \hat{\beta}_{a,C}\right) = \gamma \frac{cov(S_a, u - \tilde{u})}{var(S_a)}$$

$$= \gamma \frac{cov(S_a, X\beta_X)}{var(S_a)}$$

Again setting $\beta_a = 0$ under the null hypothesis yields:

$$\text{plim} \frac{\hat{\beta}_{a,C}}{\hat{\beta}_{a,NC} - \hat{\beta}_{a,C}} = \frac{cov(S_a, \tilde{u})}{cov(S_a, X\beta_X)} \tag{17}$$

The ratio on the right-hand side represents how strong the relationship between the residual unobservable $\tilde{u}$ and S_a must be relative to that of the controls X to explain a non-zero estimate of $\hat{\beta}_a$ when $\beta_a = 0$. A large ratio suggests that omitted variable bias is unlikely to explain away the entire effect of academic schooling on y. (Altonji et al. 2005) interpret ratios greater than 3 as large, and ratios greater than 1 as suggestive of non-zero treatment effects in the presence of possible selection on unobservables. A similar argument can be made for the corresponding expression for the vocational schooling coefficient β_v.

The assumptions used to arrive at (17) are admittedly restrictive. In particular, the assumptions that $\delta_v = \theta_v = 0$ are potentially problematic, as they require that vocational schooling S_v is unrelated to the unobservables u and $\tilde{u}$ that determine selection into academic schooling. (The assumption $\theta_X = 0$ is not problematic, as it follows from the definition of a linear projection.) Nonetheless, these assumptions are required to arrive at the familiar omitted variables bias formula when there are more than two regressors, and are therefore often invoked (Wooldridge 2001).

References
Altonji JG, Todd EE, Christopher RT (2005) Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.
    J Pol Econ 113(1): 151–184
Arum R, Shavit Y (1995) Secondary vocational education and the transition from school to work. Sociol Educ 68(3):
    187–204
Attanasio O, Adriana K, Costas M (2011) Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a Randomized Trial. Am Econ J Appl Econ 3(3): 188–220
Banerjee A, Sebastian G, James AL, Zoe McL, Ingrid W (2008) A symposium on fostering growth in South Africa: Why has unemployment risen in the New South Africa? Econ Transit 16(4): 715–40
Becker GS (1967) Human Capital a Theoretical and Empirical Analysis, With Special Reference to Education. Columbia University Press

Bellows J, Miguel E (2009) War and local collective action in Sierra Leone.
    J Public Econ 93(11-12): 1144–1157
Bennell Paul (1996) General versus vocational secondary education in developing countries: A review of the rates of
    return evidence. J Dev Stud 33(2): 230–247
Blattman C, Nathan F, Sebastian M (2013) Generating skilled self-employment in developing countries: Experimental
    evidence from Uganda. SSRN Scholarly Paper ID 2268 552,. Social Science Research Network, Rochester, NY
Bunting I, Cloete N (2010) Institutional Types in Higher Education in South Africa
Card D, Pablo I, Ferdinando R, Rosas S, Yuri S (2011) The labor market impacts of youth training in the Dominican
    Republic. J Labor Econ 29(2): 267–300
Case A (2004) Does money protect health status? Evidence from South African Pensions. Perspect Econ Aging: 287–305
Case A, Yogo M (1999) Does school quality matter? Returns to education and the characteristics of schools in South
    Africa. National Bureau of Economic Research Working Paper Series. No. 7399
Duflo E (2003) Grandmothers and Granddaughters: Old-age pensions and intrahousehold allocation in South Africa.
    World Bank Econ Rev 17(1): 1–25
Edmonds EV (2006) Child labor and schooling responses to anticipated income in South Africa. J Dev Econ 81(2): 386–414
El-Hamidi F (2006) General or vocational schooling? Evidence on school choice, returns, and 'Sheepskin' effects from
    Egypt 1998. J Policy Reform 9(2): 157–176
Fiske EB, Ladd HF (2004) Elusive equity: Education reform in post apartheid South Africa. Brookings Institution Press
Greene WH (2012) Econometric analysis, Boston, Prentice Hall
Hanushek EA, Ludger W, Lei Z (2011) General education, vocational education, and labor-market outcomes over the
    life-cycle. National Bureau of Economic Research Working Paper Series. No 17504
Hawley JD (2003) Comparing the payoff to vocational and academic credentials in Thailand over time. Int J Educ Dev
    23(6): 607–625
Hawley JD (2008) Vocational-technical schooling and occupational matching in Thailand: differences between men and
    women. Asia Pac J Educ 28: 189–205
Hicks JH, Michael K, Isaac M, Edward M (2011) Vocational education voucher delivery and labor market returns: A
    randomized evaluation among Kenyan youth. Technical Report, Report for Spanish Impact Evaluation Fund (Phase II)
Kahyarara G, Teal F (2008) The returns to vocational training and academic education: Evidence from Tanzania. World Dev
    36(11): 2223–2242
Krafft C (2013) Is school the best route to skills? Returns to vocational school and vocational skills in Egypt, Technical
    Report 2013-09 Minnesota Population Center Working Paper Series
Lam D, Cally A, Nicola B, Anne C, Murray L, Alicia M, Jeremy S, Meredith S (2008) The cape area panel study: A very short
    introduction to the integrated waves 1-2-3-4 data
Lam D, Cally A, Murray L (2011) Schooling as a lottery: Racial differences in school advancement in urban South Africa. J
    Dev Econ 95(2): 121–136
Macours K, Patrick P, Renos V (2013) Demand versus returns? Pro-poor targeting of business grants and vocational skills
    training
Magruder JR (2010) Intergenerational networks, unemployment, and persistent inequality in South Africa. Am Econ J:
    Appl Econ 2(1): 62–85
Magruder JR (2012) High unemployment yet few small firms: The role of centralized bargaining in South Africa. Am Econ
    J: Appl Econ 4(3): 138–166
Maitra P, Mani S (2012) Learning and earning: Evidence from a randomized evaluation in India. Monash University
    Department of Economics Discission Paper 44/12
Malamud O, Pop-Eleches C General education versus vocational training: Evidence from an economy in transition. Rev
    Econ Stat 92(1): 43–60
Mincer J (1974) Schooling, experience, and earnings. 1st edn. National Bureau of Economic Research
Needham S, Papier J (2011) Practical matters: What young people think about vocational education in South Africa.
    Technical Report, City & Guilds Centre for Skills Development, London
Newhouse D, Suryadarma D (2011) The value of vocational education High school type and labor market outcomes in
    Indonesia. World Bank Econ Rev 25(2): 296–322
Nkosi Bongani (2012) Vocational educational not an 'option for losers. The M&G Online
Oosterbeek H, Webbink D (2007) Wage effects of an extra year of basic vocational education. Econ Educ Rev 26(4):
    408–419
Oxford Business Group (2012) South Africa: Raising higher education.
    http://www.oxfordbusinessgroup.com/economic_updates/south-africa-raising-higher-education
Pugatch T (2012) Bumpy rides: School to work transitions in South Africa
Roy AD (1951) Some thoughts on the distribution of earnings. Oxford Econ Papers 3(2): 135–146. ArticleType:
    research-article / Full publication date: Jun., 1951, Copyright ©1951, Oxford University Press
Souza P (2010) Moral hazard in the family. Unpublished paper, Yale University
UNESCO Institute for Statistics (2011)
van der Meulen R, Zveglich YJE, Wherry L (2006) Gender differences in vocational school training and earnings premiums
    in Taiwan. Feminist Econ 12(4): 527–560
Wooldridge JM (2001) Econometric analysis of cross section and panel data, 1st ed. The MIT Press