Southern Africa Labour and Development Research Unit

Estimating the short run effects of South Africa's Employment Tax Incentive on youth employment probabilities using a difference-in-differences approach

by

Vimal Ranchhod and Arden Finn

Working Paper Series
Number 134

About the Author(s) and Acknowledgments

Vimal Ranchhod is a Chief Research Officer in the Southern Africa Labour and Development Research Unit at the University of Cape Town.

Arden Finn is a Research Officer in the Southern Africa Labour and Development Research Unit at the University of Cape Town.

Recommended citation

Ranchhod, V., Finn, A. (2014).Estimating the short run effects of South Africa's Employment Tax Incentive on youth employment probabilities using a difference-in-differences approach. A Southern Africa Labour and Development Research Unit Working Paper Number 134. Cape Town: SALDRU, University of Cape Town

ISBN: 978-1-920517-75-5

Orders may be directed to:
The Administrative Officer, SALDRU, University of Cape Town, Private Bag, Rondebosch, 7701,
Tel: (021) 650 5696, Fax: (021) 650 5697, Email: brenda.adams@uct.ac.za

Estimating the short run effects of South Africa's Employment Tax Incentive on youth employment probabilities using a difference-in-differences approach[1]

Vimal Ranchhod[2]

Arden Finn[3]

Abstract:

What effect did the introduction of the Employment Tax Incentive (ETI) have on youth employment probabilities in South Africa in the short run? The ETI came into effect on the 1st of January 2014. Its purpose is to stimulate youth employment levels and ease the challenges that many youth experience in finding their first jobs. Under the ETI, firms that employ youth are eligible to claim a deduction from their taxes due, for the portion of their wage bill that is paid to certain groups of youth employees. We utilize nationally representative Quarterly Labour Force Survey (QLFS) data for the period from January 2011 to June 2014, and implement a difference-in-differences methodology at the individual level to identify the effects of the ETI on youth employment probabilities.

Our primary finding is that the ETI did not have any statistically significant and positive effects on youth employment probabilities. The point estimate from our preferred regression is -0.005 and the 95% confidence interval is from -0.017 to 0.006. We thus obtain a fairly precisely estimated 'zero effect'. We also find no evidence that the ETI has resulted in an increase in the level of churning in the labour market for youth. What our results imply is that any decrease in tax revenues that arise from the ETI are effectively accruing to firms which, collectively, would have employed most of these youth even in the absence of the ETI. We conclude with a discussion of some of the policy implications of our findings.

JEL codes: H25, H32, J38

[1] The authors wish to thank Ihsaan Bassier for excellent research assistance, and participants in the SALDRU seminar series for very useful comments and suggestions. All opinions, errors and omissions expressed in this paper remain the sole responsibility of the authors.

[2] Southern Africa Labour and Development Research Unit, University of Cape Town. Vimal.ranchhod@uct.ac.za. Vimal Ranchhod is a Chief Research Officer in the Southern Africa Labour and Development Research Unit. He acknowledges support from the National Research Foundation's Human and Social Dynamics in Development Grand Challenge.

[3] Southern Africa Labour and Development Research Unit, University of Cape Town. aj.finn@uct.ac.za. Arden Finn is a Research Officer in the Southern Africa Labour and Development Research Unit. He acknowledges support from the National Research Foundation's Human and Social Dynamics in Development Grand Challenge.

Section 1: Introduction

Stubbornly high levels of unemployment in general, and extremely high levels of youth unemployment in particular, have been constant features of South African society since the transition to democracy in 1994. The youth (defined in this paper as adults aged between 18 and 29) experience an unemployment rate that is more than double that of the overall population unemployment rate, and about two-thirds of unemployed youth have never had a job (National Treasury, 2011). Policy discussions aimed at boosting youth employment have been taking place since the mid-2000s, with the objective of raising employment levels and mitigating the "wasting" effect of being unemployed for an extended period of time. These discussions led to the ETI, which was initiated on the 1st of January 2014. The ETI takes the form of a direct intervention in the labour market to reduce the cost to firms of employing young workers. This is one of the most ambitious youth employment initiatives in the country's history, with a proposed cost to government of R5 billion over 3 years intended to create 178 000 new jobs for youth over that period.

In this paper we ask the question "Did the youth employment tax incentive have an impact on youth employment rates in South Africa in the short run?" We utilize data from the nationally representative Quarterly Labour Force Survey (QLFS) from 12 quarters before and 2 quarters after the implementation of the ETI in order to identify the short run effects of the intervention. We employ a number of difference-in-difference estimators with different control groups to identify what effect the programme had on the probability of youth employment. We find no evidence of a positive effect on youth employment probabilities in the short run, and no evidence that the implementation of the ETI has led to increased levels of churning amongst youth in the labour market.

The remainder of this paper is structured as follows. Section 2 presents a background to the ETI in South Africa and summarizes both local and international evidence on youth employment incentives. In Section 3 we detail how the ETI works and how, in theory, the introduction of a tax rebate might increase the probability of employment amongst eligible workers. Section 4 describes the datasets and variables that we use in our analysis, while Section 5 outlines the methods that we use in order to identify the effects of the incentive. Section 6 presents the main results of our analysis, Section 7 presents some robustness tests, and Section 8 provides some concluding discussion.

Section 2: Background to the ETI in South Africa and international evidence

In order to contextualize the aims of this paper, we outline what prior research has been conducted in South Africa to understand and address youth unemployment. We then compare this to some related international evidence.

Economic theory has long established that education and work experience are important ways for work-seekers to signal their potential productivity. Levinsohn (2007) frames the problem in a paper from which much of the motivation for South Africa's youth employment incentive is drawn. When firms hire workers, they predict a productivity level based on training and experience. For youth in South Africa, education is discounted by firms because of its apparently low quality and weak applicability to skills-based jobs (Levinsohn, 2007). By virtue of being young, experience is also low or absent. Moreover, for firms who can ascertain a young worker's productivity potential, there are still substantial costs in having to train inexperienced workers.

The result of this means that firms may hire older workers who can signal productivity through experience. Related to this, the 'scarring effect' suggests that youth without prior work experience struggle to find employment as they age because (a) they have not been able to gain experience, and (b) firms see unemployment as a signal of low productivity. Firms who have to hire youth use observable characteristics as signals of productivity, and this generates the potential for statistical discrimination.

Based on the causes of youth unemployment identified above, most interventions have taken the form of one of two strategies. The first branch aims to address the lack of relevant skills or reliable education by raising the quality of workers through a series of supply-side interventions. The second branch offsets the cost (thereby lowering the risk) of employing inexperienced/young workers through demand-side initiatives.

Most previous initiatives in South Africa have been premised on supply-side changes. A study undertaken by the National Treasury (2011) emphasised the need to address the problem of youth unemployment and reviewed the measures that had already been implemented in the country. The Second Chance programmes (where drop-out learners are helped to complete schooling), the Further Education and Training (FET) schools, and the Sector Education and Training Authority programmes (SETAs) are some examples of these initiatives. While these programmes are supposed to establish a transferrable skills base for youth, alignment with firms' needs remains a problem. This is highlighted by Schoer and Rankin (2011), who find that only 44% of firms that provided training were SETA approved, indicating that the majority require non-transferrable skills. There have also

been non-training based supply-side programmes, and these have been implemented through entrepreneurship programmes, and job search assistance and sanctions schemes. [4]

One of the criticisms of these programmes is that the pathways through which they are designed to achieve results are strictly via changes in the supply of labour: plausibly, labour demand may be a bigger driver of employment since the labour supply is already high, and so these programmes would only have a limited effect on employment. As such, there have been two major initiatives from the demand-side; the Expanded Public Works Programme (EPWP) and Learnership Agreements.[5] The EPWP provided 1.6 million jobs in its first phase, but is widely recognised as a short-term and unsustainable solution (Meth, 2011 and Betcherman et al, 2004). A Learnership is a policy whereby firms are subsidised for providing approved training to employees, integrated with a job. The majority of those enrolled in this programme were previously unemployed. Despite this, the overall success of the programme is yet to be verified due to a lack of studies indicating how many of the beneficiaries would have been employed without the programme.

Learnerships may shed some light on how firms would react to the ETI, as they share many characteristics with wage subsidy programmes. The firm study by Schoer and Rankin (2011) found that just 19% of firms offer Learnerships. A report by the National Treasury (2011) unpacks this further by showing that 73% of Learnership uptake is through large firms, suggesting that small firms, potentially a source of substantial employment, may be discouraged by high administration costs. The programme also mainly benefits medium-skilled workers who earn relatively higher salaries, which excludes the majority of the low-skilled unemployment base. Burns, Edwards and Pauw (2010) further suggest that the subsidy that is given to firms is too low to cover the related training costs.

Youth employment trends gathered from labour market data make it clear that these initiatives have not solved the aggregate unemployment problem. For this reason, the ETI was proposed as an incentive for firms to hire youth, compensating firms for both the risks that firms take in hiring youth with uncertain productivity levels, as well as for the training that firms may have to provide these relatively inexperienced workers.

In addition to addressing the aforementioned problems associated with youth unemployment, the experience gained through this policy would also make future employment more likely. Schoer and Rankin (2011) find that 61% of firms consider referrals as the best mechanism for job matching, indicating how important these networks are for finding employment in the labour market.

[4] The sanction takes the form of a deduction from the UIF grant if a job offer is turned down.
[5] As defined in Section 17 of the Skills Development Act.

A number of criticisms have been levelled against the ETI. The first is that a 'deadweight loss' is incurred through companies who would have hired young workers in the absence of the programme, as has been documented in other countries with similar schemes (Betcherman et al, 2004).[6] Moreover, it is possible that companies with market power will capture these subsidies as economic rents, rather than pass them on to consumers. Destructive churning is also a potential problem if companies release workers after the subsidy expires, and employ new workers who qualify for the subsidy as replacements. The fear of young hires replacing established employees entails a similar response, and this point has been forcibly made by the Congress of South African Trade Unions (COSATU) (COSATU, 2013).

In an as yet unpublished paper, Levinsohn et al (2014) attempt to measure the effectiveness of a wage subsidy in South Africa by conducting a controlled experiment whereby a voucher (to be presented to a prospective employer) was given to unemployed individuals in a treatment group. Compared to the control group of those who did not receive the voucher, employment was higher by 25% (7.4 percentage points higher than the 31% probability of employment in the comparison group), and this persisted for 18 months after the expiry of the voucher. This result, however, has limited relevance in light of the ETI, since the current initiative entails a tax credit to firms rather than a voucher given to potential workers. In addition, it must be noted that the effect recorded by the RCT includes the increase in relative value of selected youth against non-selected youth. This implies that the findings from an RCT need to be interpreted cautiously when considering the effects of an intervention such as the ETI, which is national in scope and will affect all youth equally.

Schoer and Rankin (2011) investigate the demand-side directly by asking firms how they would react to the implementation of a youth employment incentive. Of the respondents in their study, 38% said that they would hire new labour, while 62% indicated that they would favour youth in new job openings. This lends some support to the hypothesis that older workers will be replaced in favour of younger workers who are eligible for the incentive. At the same time, 77% of firms confirm that it is unlikely that any older workers will be replaced as a result of the subsidy, if only because of high retrenchment costs and loss of experience. Finally, Burns, Edwards and Pauw (2010) create a micro-simulation for the effects of the subsidy, and find a 4.7% increase in employment, for what they term a "medium-high" assumed wage elasticity of demand of 0.7. A pro-poor bias for the initiative is suggested, since most new jobs would be for low-skilled workers in poor households. These findings

[6] In technical economic terms, this would not be a 'deadweight loss', at least not in any direct sense, as there is no obvious decrease in efficiency that would arise from such a flat transfer. There is, nonetheless, an opportunity cost to the fiscus, due to the reduction in resources available for other state activities.

are qualified in the paper with the warning that these gains are not long-term and do not address deeper structural issues in the economy.

2.1 Evidence from youth employment initiatives in other countries

The evidence in South Africa is limited in that no national cases of youth employment initiatives have previously been implemented. International comparisons can be instructive in this regard. Betcherman et al (2007) analyse the Youth Employment Inventory, a database of worldwide youth employment interventions released by the World Bank. 39% of interventions were skills-based, with subsidies constituting approximately 12% overall. Fewer than half of all programmes were cost-effective and successful, an indication that interventions need to be considered very carefully. Three general conclusions emerge: finance is important in determining success, interventions work better in countries with flexible labour markets, and context - not type of programme - determines the success of an initiative.

More specifically, wage subsidies tend to be most successful when they are woven into 'comprehensive packages' – that is, programmes that include other facets such as training. Wage subsidies have been successful in Europe especially, with increases in employment of 12 and 13% in Poland and the Czech Republic respectively. In line with the RCT conducted by Levinsohn et al (2014), average earnings conditional on employment decrease as more low-income jobs are created. These results are all found to be short-term.

In an analysis of global youth employment initiatives, Smith (2006) argues against demand-side subsidies. Demand side programmes were implemented in Australia, where administrative costs in establishing eligibility severely limited outcomes; in the Hungarian subsidy of the 1990s, where no impact was found; and in Poland's 'Intervention Works' programme in which a negative effect was found due to biased state administration. This is in contrast to a number of successful international supply-side subsidies: the USA's Earned Income Tax Credit scheme increased employment by 6%, and programmes in Canada and the UK recorded positive effects as well. Smith (2006) argues that any benefits of a demand-side subsidy are found through the supply-side, while the former incurs more administrative costs for firms. However, these findings are all in OECD countries (the impact may be very different in a developing economy) and for targeted subsidies. The implication is that a low-income, non-targeted youth employment incentive in South Africa could potentially avoid many of the administrative costs that reduce the impact of these initiatives. At the very least, two issues should be kept in mind when considering the lessons from the international literature: the potential for deadweight loss from jobs that would have been created anyway, and whether the intervention is aimed at the demand-side or the supply-side of the labour market.

Finally, an analysis of a number of Latin American and the Caribbean (LAC) countries by Puerto (2007) is useful since the regions comprise developing economies that are similar in many respects to South Africa. Three waves of schemes have been adopted in the region. In the 1970s, supply-driven vocational training programmes comparable to South Africa's SETAs dominated. By the 1990s, a demand-side had approach become more prevalent, starting with Chile's Jovenes programme and replicated in Venezuela, Argentina, Paraguay, Peru, Columbia, Panama and the Dominican Republic. Including costs, an average return of 4% in *sustainable* employment was created compared to control groups. In more recent times, the 'Extra 21' programme has been the showcase example, and has been adopted in various forms in eighteen LAC countries. Generally, programmes that are demand-driven and involve the private sector have been more successful (Puerto, 2007). Nonetheless, Burns, Edwards and Pauw (2006) observe that in Argentina's Proempleo Experiment, the take-up of firm vouchers was low while the increase in employment was high, which lends support to the findings of Smith (2006) and Levinsohn et al (2014) that the gains are supply driven. Potentially of relevance to South Africa, the successes in the LAC countries are obtained despite the region recording the world's third highest employment rigidity index, an indication that the presence of highly-regulated markets is not an insurmountable obstacle to the implementation of programmes.

Evidence from South Africa and international evaluations of employment subsidies indicates that some interventions do generate positive returns. Concerns about labour substitution are not supported by available evidence, and flexibility in the market is seen as a positive feature rather than a barrier. With this evidence in mind we now turn to a description of the ETI, along with its theoretical implications.

Section 3: Outline of the programme and its theoretical implications

The ETI was presented for public comment in a 2011 National Treasury discussion document (National Treasury, 2011). The central stated aim was to spend R5 billion over 3 years to create 178 000 new jobs for the youth (at a cost of approximately R28 000 per job). The National Planning Commission cited youth unemployment as a key policy issue for the country, and the ETI was signed into law in 2013, with effect from the 1st of January 2014. The budget for the first year of implementation was R1 billion (National Treasury, 2014).

The main eligibility criterion for an employer is registration for employees' income tax (PAYE). There is no limit to the number of qualifying workers that an employee can hire. The employer claims the incentive on a monthly basis, based on the number of employees supported by the incentive and the relevant monthly salaries of those employees. Penalties are levied against employers who claim the

incentive for workers that are paid less than the relevant minimum wage (or under R2 000 per month for cases where a minimum wage is not applicable). The programme also builds in a disincentive to displace existing employees by levying a penalty on the employer of R30 000 for each worker that is found to have been displaced in order to take advantage of the ETI.

Employees qualify for the ETI if they are aged between 18 and 29 years and hold a South African ID.[7] The employee cannot have been hired for the current job before the 1st of October 2013.

Employers claim the incentive every month by filling out the ETI field on the EMP201 form that is submitted to the South African Revenue Service (SARS). The incentive is activated by lowering the amount of PAYE tax that is payable to SARS each month. The benefits of the programme extend to a maximum of 24 months, and the programme in its current form is effective until the 31st of December 2016.

The Employment Tax Incentive Bill (Republic of South Africa, 2013) sets out the amounts that employers can claim from the incentive each month, and this information is presented in the Table 1 below.

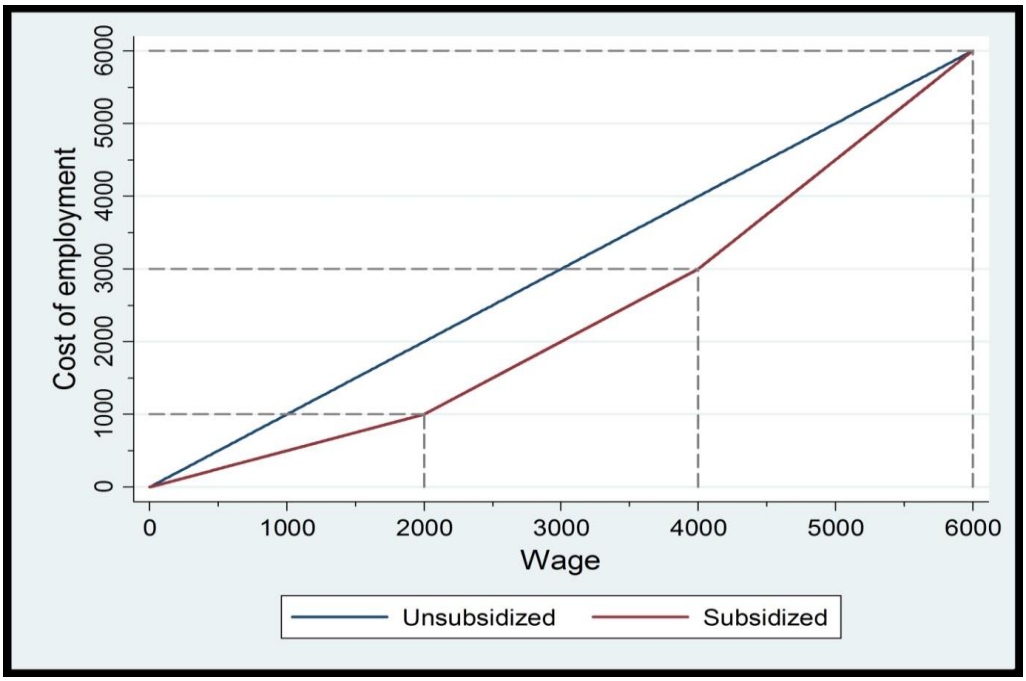
Table 1: Details of the ETI per qualifying employee

Monthly remuneration	**Year 1 ETI per month for qualifying employee**	**Year 2 ETI per month for qualifying employee**
R0 - R2 000	50% of monthly pay	25% of monthly pay
R2 001 - R4 000	R1 000	R500
R4 001 - R6 000	R1 000 - (0.5*(monthly pay - R4 000))	R500 - (0.25*(monthly pay - R4 000))

Figure 1 presents a graphical contrast of the cost to company versus the wage of a subsidised worker over the R0 to R6 000 range. As in the table above, the wage is subsidised by 50% up to R2000, beyond which the subsidy is a flat rate of R1 000, up to a wage of R4000. In the final segment the cost to company and the wage received by the subsidised employee converge until they meet at R6000.

[7] The age restriction is not applicable for employees working in special economic zones (SEZs). The SEZs are geographically designated areas that are earmarked for specific economic activities aimed primarily at boosting labour-intensive, export-led industries. They grew out of the Industrial Development Zone (IDZ) programme which aimed to increase South Africa's value-added exports, and are a key part of government's 2014/2015 – 2016/2017 Industrial Policy Action Plan.

Figure 1: Cost to company of an eligible employee as a function of wage per month



Administrative data on the take-up of the ETI during the first six months of the programme is not yet publicly available. Nonetheless, we have two pieces of relevant information about take-up rates obtained from official speeches of government representatives. First, the erstwhile Minister of Finance noted in his February 2014 budget speech that 56 000 beneficiaries were recorded in the first month of the year (Gordhan, 2014). Second, in his June 2014 State of the Nation Address, President Zuma stated that in the first five months after its introduction, there were 133 000 beneficiaries of the ETI, most of whom were employed in the wholesale and retail, manufacturing, and finance sectors (Zuma, 2014).

3.1 Theoretical implications of the ETI

The primary contribution of this paper is empirical and not theoretical. Nonetheless, it is useful to briefly consider the behaviour of profit maximizing firms from a simplified theoretical perspective for two reasons. First, having a clear sense of how firms might react to the ETI will help to shape our *a priori* expectations of the likely impact of the ETI, which in turn will provide us with some frame of reference within which we can interpret our empirical results. Second, theoretical expectations have a role to play in informing our choice of empirical strategies, their credibility and their possible limitations.

The overall objective of the ETI is to address the youth unemployment problem in South Africa, and its stated objective is to create 178 000 new jobs[8] over a three year period. Given the substantial unemployment levels amongst youth, the binding constraint on aggregate youth employment levels is probably not labour supply amongst youth, but aggregate labour demand for youth labour. Aggregate demand for youth labour at a point in time, in turn, is defined as the sum total of the demand for youth labour from each firm at that same point in time. Being explicit about the time horizons is useful, as a purely static analysis would not take cognisance of the fact that the number and type of firms is itself changing over time. Thus, we can describe the ETI as a policy intervention with the objective of either increasing the number of firms that employ youth labour, or of increasing the amount of youth labour that existing firms would like to employ, or both. In the short run, it seems reasonable to ignore the firm entry and exit margin. Thus, any short run effects of the ETI are likely to be obtained primarily through increases in the demand for youth labour by (at least some) existing firms.

How does the introduction of the ETI affect the economic environment of an individual profit maximizing firm? The design of the ETI is such that informal sector firms, firms in the formal sector that are not registered for PAYE, and formal sector firms that pay sufficiently low wages such that they are not liable for any PAYE; will all be ineligible to benefit from the ETI. As such, the ETI does not affect these firms in any direct way. For the ETI to have a direct impact, this impact must thus be derived from eligible firms; namely formal sector firms that are registered for PAYE and have at least one sufficiently high earning employee against whose PAYE obligations the ETI can be set off against. Amongst these eligible firms, the introduction of the ETI has two implications. First, it makes a subset of the labour force less costly than it used to be in absolute terms. Second, it makes that same subset of the labour force less costly than it used to be in relative terms, in comparison to all other subsets of the labour force.

The first effect, which can be attributed to the decrease in the absolute price of a factor of production, leads to an increase in output under most conditions. This is called the 'output effect', and can increase the usage of all factors of production. In the short run, however, we assume that both the capital stock and the technology of production is fixed, hence the increase in output will have to be generated by an increase in the usage of various forms of labour, including youth labour.

[8] In the public discussion paper, National Treasury (2011) estimates that the ETI would subsidise 423 000 new jobs, of which 178 000 would be "net" new jobs, that is, jobs that would not have been created without the implementation of the ETI. Furthermore, they assume that three quarters of these net new jobs are "sustainable jobs" – jobs that last 6 months or more – meaning that 133 000 net sustainable jobs should be created by the programme.

Note that the employment of other categories of labour are potentially also increasing due to the output effect.

The second effect, which can be attributed to the decrease in the relative price of labour, leads to a substitution towards the relatively cheaper factor of production, and is called the 'substitution effect'. In this example, the substitution effect implies that for any given level of output, firms will employ relatively more youth than they would have in the absence of the ETI. Moreover, since we are assuming that the capital stock and technology of production are fixed, this implies that they will employ relatively fewer workers from at least one other category of workers, for a given level of output.

The total effect within an eligible firm is the combination of the substitution effect and output effect. For youth labour, since the substitution effect and the output effect are both positive, or at least not negative, we expect to observe an increase in the total demand from eligible firms. By aggregating across all firms, we thus expect that an increase in the aggregate demand for youth labour will occur, as a result of the ETI.

What is relevant for the ETI to have a substantial impact is not just the existence of an increase in aggregate demand for youth labour, but also the magnitude of such an increase. The 'wage elasticity of labour demand for youth labour' is a measure of the sensitivity of the aggregate demand for youth labour to changes in the youth wage rate, and it follows from our preceding discussion that it will be affected by several considerations. These include the number of eligible firms, the size of these firms, the technologies of production in these firms, the substitutability of youth labour for other groups of labour within eligible firms, the sensitivity of these firms' output to the cost of youth labour, and the amount of youth labour that these firms optimally utilize per unit of output.

Ultimately, the wage elasticity of labour demand for youth labour could be very high or very low, and can differ substantially by industry, due to the various possible parameter values of each of the aforementioned factors that affect this elasticity. The possible dispersion in the wage elasticity of labour demand for youth labour could, in turn, result in the ETI having a very large or a very small impact on the youth unemployment problem. This provides the theoretical motivation for our empirical analyses.

Finally, although it is not the focus of this paper, it is worth noting that the ETI could also affect the demand for other groups of labour. As with youth labour, the output effect should be weakly positive for all of the other groups of labour as well. The sign of the substitution effect for the other groups of labour, however, is uncertain, depending on whether a particular group is a complement

or a substitute to youth labour in production. The total effect of the ETI on the aggregate demand for a particular group of non-youth labour is thus uncertain. Without more information about the number of firms, their size, scale of production and production technologies, we cannot make any definitive statements about the theoretical effect of the ETI on aggregate demand for other groups of labour.

Section 4: Data and variables

In this paper we make use of 14 waves of the Quarterly Labour Force Surveys (QLFS) conducted by Statistics South Africa (StatsSA). These are nationally representative individual-level surveys which are conducted four times per annum. The waves that we make use of begin with QLFS 2011:1 and include all subsequent waves, up to and including QLFS 2014:2. The time period that this spans is from January 2011 until June 2014, thus capturing trends in the data from three years prior to the introduction of the ETI as well as data from the first six months in which the intervention was implemented.

The QLFS surveys are treated as repeated cross-sections in our study.[9] In each wave, enumerators attempt to contact approximately 30 000 dwellings and interview all adult resident members of these dwellings. The sampling frame is obtained from the Census Master sample and sampling follows a stratified clustered design, as per various StatsSA documents that are released with each cross-section.[10] In order to obtain unbiased population estimates from each wave, while accounting for the complex survey design as well as selective non-response rates, StatsSA releases a weight variable that is calibrated on the population mid-year estimates obtained from the relevant Census. We apply these weights to all computations that are intended to reflect population-level statistics.

Table 2 below indicates the aggregate sample size of our dataset, as well as the cell sizes of various groups of respondents within each wave. We restrict our sample to working aged adults aged 18 to 64.[11] Our overall sample across all waves is 667 610 observations, and each of the fourteen waves has more than 45 000 observations. Within age groups, we have 246 941 youth aged 18 to 29, 79 289 'almost youth' aged 30 to 34, 199 801 'prime aged' respondents aged 35 to 49, and 141 579

[9] Technically, they are not repeated cross-sections as they have a rotating panel design at the dwelling level with a 25% rotation rate. (QLFS Metadata documents, various waves, StatsSA). Unfortunately, we cannot make use of the panel component of this data as StatsSA has not yet released the QLFS panel for waves from 2014.

[10] Each wave comes with a set of related documentation including a user guide, metadata document and questionnaire. The datasets, together with the documentation, are available for free public download from *www.datafirst.uct.ac.za.*

[11] This is slightly different to the StatsSA age range, in that they include respondents aged 15 to 17 as well. We chose not to include this subset of the sample as they are not eligible for any tax incentives under the ETI.

'older adult' respondents aged 50 to 64. Even the smallest group, 'almost youth', contains more than 5 000 observations in each wave. The major advantage of having such a large dataset is that, even when we consider subsets of the sample within a particular wave, we still have a substantial amount of statistical power. This implies that our results are likely to be relatively precisely estimated.

4.1 Questionnaire and variables

Each wave of data contains questionnaires with multiple modules. First, a household roster is populated with information including demographic characteristics such as the respondent's age, gender, race and educational attainment. Thereafter, a series of sections explores whether a respondent is employed, unemployed or not economically active. Within each of these categories of labour market statuses, additional questions are asked. Of those who are employed, questions are

Table 2: Sample sizes by wave and age group

	Youth	**Non-Youth**				**Total**
Age group	***18-29***	***30-34***	***35-49***	***50-64***	***30-64***	***18-64***
Wave	**Total: Youth**	**Almost Youth**	**Prime Aged**	**Older Adults**	**Total: Non-Youth**	**Youth & Non-Youth**
1	17,101	5,227	13,888	9,496	28,611	45,712
2	17,020	5,112	13,714	9,613	28,439	45,459
3	17,707	5,342	14,226	9,705	29,273	46,980
4	17,903	5,482	14,262	9,886	29,630	47,533
5	17,486	5,558	14,343	9,961	29,862	47,348
6	17,451	5,635	14,329	10,064	30,028	47,479
7	17,631	5,750	14,511	10,297	30,558	48,189
8	17,619	5,757	14,486	10,241	30,484	48,103
9	17,412	5,776	14,345	10,256	30,377	47,789
10	17,829	5,973	14,608	10,436	31,017	48,846
11	18,099	5,944	14,429	10,439	30,812	48,911
12	18,172	6,012	14,495	10,491	30,998	49,170
13	18,088	5,913	14,370	10,437	30,720	48,808
14	17,423	5,808	13,795	10,257	29,860	47,283
All Waves	**246,941**	**79,289**	**199,801**	**141,579**	**420,669**	**667,610**

asked about the conditions of employment, the characteristics of the employer, the worker's job description, and when the respondent started their current job. Of those who are not employed, a set of questions ask about when they were last employed, as well as the job description and the characteristics of the employer in their most recent job.

There are also derived variables which are released by StatsSA for each respondent which are not obtained in direct response to any question in the questionnaire. Some of these variables are geographic, such as the province of the household or whether the household is located in an urban or non-urban area. Other variables, such as a worker's occupation, industry and sector of employment, are derived from responses to multiple questions in the questionnaire.

Our analysis makes use of several variables. We have two main variables of interest; *employed* is an indicator variable that takes on a value of one if a respondent is classified as employed in the StatsSA derived *status* variable, and *employed_formal* is an indicator variable that takes on a value of one if a respondent is classified as *employed* and the sector within which the employment occurs is the formal sector.[12] StatsSA classifies a respondent as employed if, in the week preceding the interview, he/she had any type of paid employment including a regular job, casual job or piece work; or was self-employed; or helped in a family business for no pay.

Since our analysis is focussed on identifying the effects of a new policy, our primary source of variation is temporal. This is captured by the wave that the data is obtained from. Waves 1 to 12 represent the period prior to the introduction of the incentive, while wave 13 and wave 14 were conducted in the period in which the incentive was in effect. To utilize this information, we construct several temporal variables. First, we generate a trend variable which is equal to the wave number. Second, we generate a '*pre-*' variable and a '*post-*' variable, which identifies whether a row of data is obtained from waves 1 to 12, or waves 13 to 14 respectively. Third, in our main regressions, we separate the *post-* variable into *post1* and *post2*, which denotes wave 13 and wave 14 respectively. Fourth, to account for seasonality in the labour market, we generate a set of *quarter_i* dummy variables, which represents the quarter within a calendar year that the relevant survey was conducted.

Of the demographic variables, the most important variable is *Age*, as we use it to construct age groups which capture whether a respondent was a potential beneficiary of the incentive or not. The age groups are defined above in our discussion of Table 2. In addition, we use race, gender and the highest educational attainment. The race variable captures the race of the respondent into one of four categories; namely African, Coloured, Indian and White. The education variable is converted into a dummy variable called *tertiary* that takes on a value of one if the respondent has attained

[12] The sector of employment is derived by StatsSA using a combination of variables including whether the employing organization is registered for tax purposes and the number of employees employed by the employing organization.

some form of a post-secondary school level qualification. For some of our analysis, we refer to 'skilled' and 'unskilled' youth, and the *skilled* variable is identical to the *tertiary* variable. Thus, for example, 'unskilled youth' are respondents aged 20 to 29 who have at most a matric level education.

Finally, we make use of a number of other variables. Of the household-level variables we use the *province* variable and also generate an *urban* dummy variable. Amongst those who are currently employed, we utilize the *industry* and *occupation* variables, which are derived by StatsSA. In the latter part of our paper, we consider whether there has been a change in the levels of churning in the labour market amongst youth. To do this, we construct four dummy variables. Amongst the employed, we generate *recentjob_3* and *recentjob_6*, which identify whether a respondent began his/her job within the last three or last six months respectively.[13] Amongst the people who are not employed, we make use of a variable that captures when a person was last employed. From this, we construct *recentloss_3* and *recentloss_6*, which identifies whether a respondent who was not employed at the time of the survey lost their most recent job within the past three or past six months respectively.[14]

In Table 3 we contextualize the data by summarizing some of the variables that are used in our study. The mean age amongst youth is 23.12 years, while amongst the non-youth it is 44.8 years. About 49% of youth are male, while amongst the non-youth this is slightly lower at approximately 44%. About 84% of the youth are African, while this is substantially lower amongst the non-youth, at close to 75%. Correspondingly, the proportion of the sample made up of the other race groups is higher amongst the non-youth than the youth. As expected, the youth are slightly less likely to have acquired a tertiary qualification, at about 7%. This reflects the fact that many youth, especially those in their early twenties or younger, are still studying or that they may still enrol for a tertiary qualification at some future date. The proportion of youth residing in urban areas is slightly lower than that of non-youth, at 61% and 66% respectively. The largest provinces, as defined by where our

[13] In order to construct these variables we used data on the month and year in which the current job began. This is imprecisely measured for two reasons. First, we do not know the exact date when the interview was conducted, so we assumed that every respondent was interviewed at the very end of the relevant quarter. Second, we do not know the exact date that the job began, so we assumed that it began on the first day of the month in which it began. This must lead to a downward bias in the number of jobs that are classified as 'recent jobs'. Nonetheless, we have no reason to believe that this bias would not be consistent across the different waves, and what we are testing for is whether there is an increase in the fraction of these recent jobs in the *post-* period relative to the *pre-* period.

[14] Note the *recentjob_3* and *recentloss_3* are nested within *recentjob_6* and *recentloss_6* respectively. By this we mean that if a respondent found their current employment within the previous three months then it implies that they also found their current employment within the previous six months, and similarly for having recently lost a job within the previous three or six months (conditional on not being employed).

sample is obtained from, are Gauteng and KwaZulu-Natal. These are followed by the Western Cape, Eastern Cape, Limpopo and Mpumalanga.

To summarize, the youth in our sample do look somewhat different to the non-youth, but in ways that are not surprising. The youth are mostly unskilled, most likely to be African, almost balanced in terms of gender, and are just over 23 years old on average. Approximately three out of five of these youth live in urban areas, and more than a third of them live in either KwaZulu-Natal or Gauteng.

Table 3: Means of covariates used in regressions

	Youth			**Non-Youth**			**Full Sample**
Variable	***Pre-***	***Post-***	**Diff.**	***Pre-***	***Post-***	**Diff.**	**(All waves)**
Age (years)	23.12	23.20	0.08	44.81	44.86	0.05	36.79
Male	0.49	0.49	0.00	0.44	0.44	0.00	0.46
African (%)	83.68	84.17	0.49	74.40	74.73	0.33	77.89
Coloured (%)	10.34	10.29	-0.05	13.05	13.25	0.20	12.07
Indian (%)	1.65	1.81	0.16	2.66	2.62	-0.04	2.29
White (%)	4.33	3.73	-0.60	9.89	9.41	-0.48	7.76
Tertiary	0.07	0.07	0.01	0.13	0.13	0.00	0.11
Urban	0.61	0.61	0.00	0.66	0.67	0.00	0.64
Province (%)							
Western Cape	10.80	10.92	0.12	13.36	14.37	1.01	12.51
Eastern Cape	10.52	10.96	0.44	11.30	11.38	0.08	11.04
Northern Cape	4.77	4.82	0.05	5.79	5.62	-0.17	5.40
Free State	8.77	8.07	-0.70	8.91	8.84	-0.07	8.82
KwaZulu-Natal	18.47	19.03	0.56	16.00	15.67	-0.33	16.91
North West	7.50	7.30	-0.20	8.12	7.76	-0.36	7.85
Gauteng	16.06	15.40	-0.66	17.59	17.21	-0.38	16.95
Mpumalanga	10.70	10.83	0.13	8.86	8.94	0.08	9.55
Limpopo	12.42	12.69	0.27	10.08	10.22	0.14	10.97

Notes:
1. These are sample means and are unweighted.
2. The "Diff." column is the difference between the mean of the relevant variable in the Post- period relative to the Pre- period.
3. The Pre- and Post- periods are represented by waves 1-12, and waves 13-14 respectively.

That the youth and non-youth do indeed look different to each other in some respects is not a matter of great concern for us. What we would ideally like to see is that the underlying samples are stable within each group, across the *Pre-* and *Post-* periods. The differences in the group level means are presented in the relevant *Diff* column in Table 3. While there are a number of small changes, such as the observation that the percentage of youth that are African increased by about half a percentage point in the *post-* period relative to the *pre-* period, the overall observation is that the

samples within each group and time period are extremely stable. This implies that the sampling framework was consistent across time, and that there was no substantial selective migration over time. This makes us more confident in the validity of our empirical findings, as either a change in the underlying samples, or substantial selective migration, would potentially have confounded our subsequent analyses.

Section 5: Methods

The methods that we employ in this paper are reasonably common in the empirical labour economics literature. First, we present a number of summary statistics from our sample. These take the form of the means and conditional means of the employment to population ratio[15], for various age groups within each wave. We then present the distribution of industries in which youth are employed, as well as the proportion of the employees within each industry that are youth. We perform these calculations for both the *pre-* and *post-* periods. We next repeat the analysis just described, but do so for occupational category instead of industry.

We then estimate two sets of regression results. The first is a set of *before-after* regressions that focuses only on youth. We fit a model of the form:

$$y = \beta_0 + \beta_1 trend + \beta_2 post1 + \beta_3 post2 + \lambda X + \varepsilon$$

Our regressions are OLS regressions run at the individual level and are linear probability models. We have two dependent variables of interest, namely *employed* and *employed_formal*. We chose to investigate employment in the formal sector in isolation, as the design of the ETI is such that it does not have any direct bearing on firms in the informal sector. For each of these dependent variables, we fit the regression model to the data for two different estimation samples; all of the youth in our data, and youth that reside in urban areas only. Our rationale for focussing only on urban areas is again informed by the design of the ETI, as we expect that a disproportionate number of eligible firms will be based in urban areas.[16] By focussing on the formal sector and/or on urban youth we thus hope to have a better targeted regression model, in terms of our ability to identify the effects of the ETI.

[15] This is also referred to as the 'labour absorption rate'.

[16] In waves 1 to 12 of our dataset, 63% of rural youth who are employed are employed in the formal sector, while the corresponding statistic for urban youth is just below 81%.

We further include *age*, *age-squared*, *male*, *urban* (where relevant), *tertiary*, as well as indicator variables for *province* categories and *race* categories as control variables. The regressions are weighted and the standard errors are clustered at the PSU level.

Our coefficients of interest in these regressions are β_2 and β_3, which correspond to the variables *post1* and *post2*. We chose to separate the 'treatment period' by including two dummy variables separately as this allows for a more flexible specification, and there is some chance that the ETI would not have had much of an impact in wave 13 as it had just been introduced at that time. The parameters β_2 and β_3 measure the magnitude of any trend break in the dependent variable (in the relevant estimation sample), in each of the two waves after the introduction of the ETI. In the absence of any contemporaneous confounding factors, these parameters measure the effects of the ETI in wave 13 and wave 14 respectively.

Our second set of results is obtained from a regression model that uses the *differences-in-differences* (DD) approach popularized by Card and Kruger (1994). While the *before-after* estimator is extremely intuitive, its primary limitation is that any changes in the general economic environment that affect youth employment probabilities, which are also coincident with the introduction of the ETI, will also be captured by the relevant coefficients. Some examples of external events that would affect the general economic climate include an oil price shock, an economy wide expansion or recession, or a financial crisis. If any of these were to have occurred beginning in 2014, they would also affect our estimates of β_2 and β_3. If we were to then interpret these estimates as the effects of the ETI on youth employment probabilities, we would thus be making an attribution error in terms of our estimate of the impact of the ETI. The DD estimator is a slightly more complicated estimator, but is also more robust to such threats to the identification of the effects of the ETI.

To implement the DD estimator, we fit a model of the form:

$$y = \beta_0 + \beta_1 trend + \beta_2 youth + \beta_3 post1 + \beta_4 post2 + \alpha_1(post1 \times youth) + \alpha_2(post2 \times youth) + \lambda X + \varepsilon$$

We use the same dependent variables as with the *before-after* regression models described above, and also regress our dependent variables first on the full dataset as well as on the restricted sub-sample that resides in urban areas only. Our estimation samples now also include all of the *non-youth* in our data. In addition to all of the covariates from the *before-after* regression models, we also include *youth, youth_post1* and *youth_post2*. In conforming to the terminology used in the DD

literature; our 'treated group' variable is *youth*, the 'treatment period' is captured by *post1* and *post2*, and the 'treated group in the treatment period' variables are the interaction dummies *youth_post1* and *youth_post2*.

Our primary coefficients of interest are α_1 and α_2, which correspond to the *youth_post1* and *youth_post2* variables. The *trend* variable captures general trends in employment rates over the fourteen waves of data, while the *youth* variable captures mean differences in the employment rate of youth relative to the non-youth in the estimation sample. The variables *post1* and *post2* allow for a break in the general trend in employment in the estimation sample in wave 13 and wave 14, respectively. What the coefficient on the *youth_post1* thus represents is mean changes in the probability of employment amongst youth relative to the non-youth, in the first of the post-ETI periods, in addition to any pre-existing differences in employment probabilities that were observed between youth and non-youth prior to the introduction of the ETI, and also in addition to any general changes in the economic climate that occurred in wave 13 that affected the employment probabilities of both youth and non-youth equally.[17] In the absence of any additional confounding factors, this represents the average treatment effect of the ETI in wave 13.[18]

There is one nuanced issue that might arise with respect to our methods; the average treatment effect amongst all youth might not be a very interesting question. Skilled youth have much higher employment probabilities and probably earn substantially more than unskilled youth do.[19] In addition, the way that the ETI is designed is that it is unlikely to significantly stimulate demand for skilled youth due to its R6 000 per month cut-off. By combining skilled and unskilled youth into a single category, we might not observe some potentially important heterogeneity in the effects of the ETI. For example, it is possible that the true effect of the ETI on employment probabilities is zero amongst skilled youth, and positive and statistically significant amongst unskilled youth. In this case, if we were to combine these two groups, we would estimate an average treatment effect that is smaller than the true effect amongst unskilled youth and the estimate may not be statistically significantly different from zero. In response to this concern, we repeat all of our analyses described above, but further restrict the estimation sample used for the regressions to the set of unskilled respondents in the dataset.

[17] The interpretation of the coefficient corresponding to the *youth_post2* variable is analogous.
[18] We discuss additional possible confounding factors further in the section on robustness tests.
[19] The labour absorption rate amongst unskilled and skilled youth in 2013 was 26.7% and 60.9% respectively. Moreover, about 92.5% of the youth in our sample are categorized as *unskilled*. Taken together, this implies that about 96% the youth in our sample who are not employed are unskilled. (At present, we do not have the relevant wage data required to identify the fraction of these employed youths in the two groups that earn below the various relevant income thresholds to qualify for the ETI).

Section 6: Results

6.1: Labour absorption rates

We start by presenting and discussing a series of summary statistics. In Table 4 below, we present the weighted labour absorption rates in each wave. The overall mean, across all the waves and all the age groups is 0.465, which means that slightly below half of all the adults aged 18 to 64 were employed during this time period. One interesting observation is how the labour absorption rate varies across the life cycle. Amongst the youth aged 18 to 29, the labour absorption rate is only 0.292. The absorption rate increases further with age until it peaks at just below two thirds amongst the prime aged adults, and then drops sharply amongst the older adults to 0.47. The absorption rate is thus about 18 percentage points higher for the older adults than the youth.

Table 4: Mean labour absorption rates by wave and age group

	Youth	Non-Youth				Total:
Wave	Total: All Youth	Almost Youth (30-34)	Prime Aged (35-49)	Older (50-64)	Total: All Non-Youth	Youth & Non-Youth
1	0.288	0.566	0.626	0.456	0.565	0.460
2	0.287	0.562	0.618	0.465	0.563	0.458
3	0.291	0.568	0.622	0.469	0.567	0.462
4	0.296	0.574	0.624	0.474	0.571	0.466
5	0.294	0.569	0.621	0.464	0.566	0.463
6	0.287	0.571	0.627	0.463	0.568	0.462
7	0.294	0.571	0.631	0.468	0.572	0.467
8	0.289	0.561	0.633	0.463	0.569	0.464
9	0.288	0.561	0.630	0.463	0.568	0.463
10	0.290	0.563	0.628	0.471	0.569	0.465
11	0.295	0.575	0.637	0.478	0.579	0.472
12	0.304	0.575	0.631	0.480	0.576	0.474
13	0.292	0.561	0.632	0.483	0.574	0.469
14	0.293	0.557	0.628	0.484	0.572	0.468
All Waves	0.292	0.567	0.628	0.470	0.570	0.465

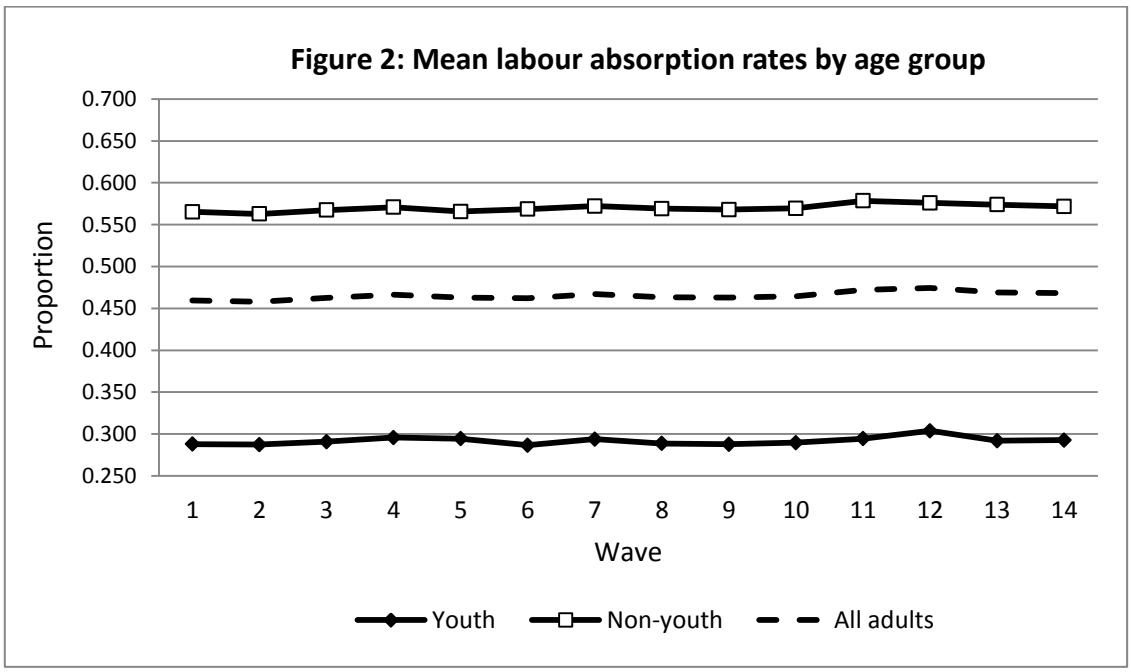
Notes:

1. All means are weighted using the sampling weights.

2. The sample includes any respondent of a particular group, including those who are not economically active.

Of particular interest in this study is whether there is a change in the trend line of youth employment relative to non-youth employment in either wave 13 or wave 14. The trend lines are depicted graphically in Figure 2 below. When we consider the aggregate absorption rate, there does

seem to be a small upward trend between wave 1 and wave 12. Thus, between January 2011 and December 2013, the proportion employed increased from 46% to a peak of 47.4%, although most of the upward movement occurred during 2013. This may be small in terms of percentage points, but it is substantial in terms of the number of new jobs that were created. The trends amongst the youth as well as amongst the non-youth are both similar to that of the aggregate, albeit at different levels of absorption. In 2014 however, the absorption rate amongst the non-youth seems to have stabilized or decreased slightly, from a peak of 47.4% in wave 12 to 46.9% and then 46.8% in wave 13 and wave 14 respectively. The pattern is also observed amongst the youth, although the decline appears to be more pronounced. The absorption rate amongst the youth decreases from a peak of 30.4% in wave 12, prior to the introduction of the ETI, to 29.2% and 29.3% in wave 13 and wave 14 respectively. Taken together, this suggests that the ETI did not result in a significant number of new jobs for its intended recipients.



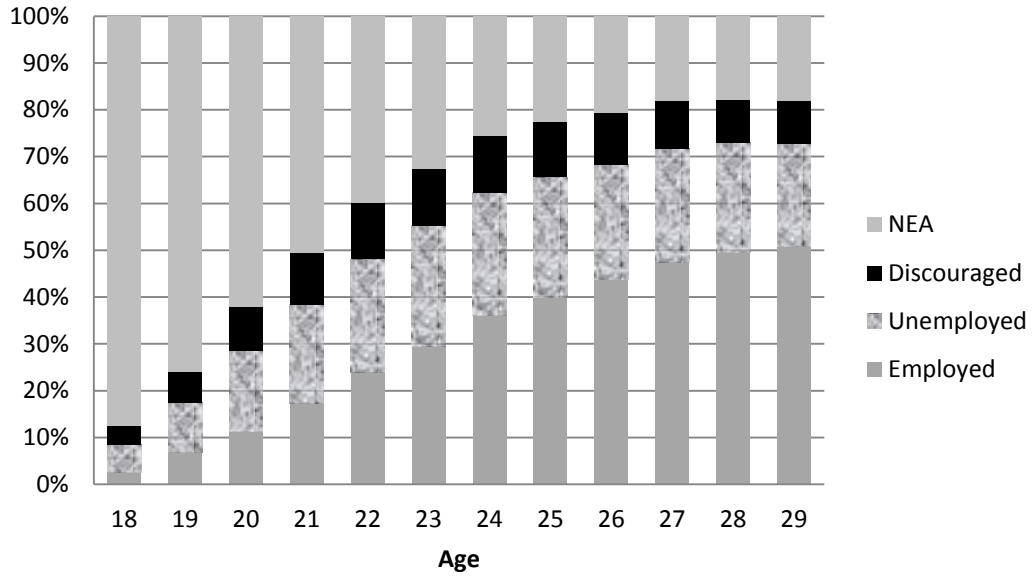
Section 6.2: Labour market status across the life cycle amongst youth

We next present the differences in labour market outcomes that youth of different ages experience. The mean percentage in each labour market state is presented graphically, for each age, using the data from waves 1 to 12. Amongst teenagers, the levels of labour force participation are very low, as evidenced by the light grey category of “Not Economically Active”. This group decreases rapidly with age, such that by age 21 about half of the youth are in the labour force, and this rises further to almost 80% by age 26.

The percentage employed, however, does not rise as rapidly as the entry into the labour force, as evidenced by the dark grey category in the bar chart. At ages 21, 25 and 29, the mean percentage employed are approximately 17%, 40% and 51% respectively. This difference between the rate of entry into the labour market and the rate at which youth find jobs results in a rapid increase in unemployment in the early and mid-twenties, as shown by the middle two groups in each bar. Graphically, this reflects the essential problem that led to the introduction of the ETI; namely that youth in South Africa experience exceptional difficulty in finding their first jobs.

In their late twenties, the rate of labour force entry amongst youth stabilizes at just over 80%, while the rate at which youth find employment continues to increase, albeit rather slowly. Thus, by age 29, 62% of youth who are in the labour market are employed, which is the highest employment rate for any age within the youth group, although the corresponding unemployment rate is still exceptionally high at 38%.

Fig 3. Distribution of labour market states prior to ETI: by age



Section 6.3: An analysis of the industries that employ youth

In the first two columns of Table 5 we document the industries in which employed youth are employed, as observed in the QLFS. In the period from 2011 to 2013, the industry that was by far the most likely to employ youth was the wholesale and retail sector, at 27.6%. This was followed by the community, social and personal services; financial; manufacturing; and construction sectors at

16.4%, 14.7%, 13% and 8.9% respectively. Together, these sectors accounted for just over 80% of employment amongst youth in our sample.[20]

In 2014, we observe some changes in the industries in which youth are employed. There is a 1.72 percentage point decrease in the percentage of employed youth that are employed in the manufacturing sector, and a 1.34 percentage point increase in the fraction of employed youth that are employed in the community, social and personal services industry. With the exception of these two changes, all of the other compositional changes are below 1 percentage point in absolute value. The main industries that employ youth, as well as their ranking, remain unchanged and account for about 81.1% of all employment amongst youth. Thus, the overall distribution of employment by industry appears to be quite stable.

Table 5: Distribution of industry conditional on youth employment, and proportion of employed within industry that are youth

	Distribution of Industry amongst employed youth(%)			**Prop. of employed within an industry that are youth**		
Industry	**Pre-**	**Post-**	**Diff.**	**Pre-**	**Post-**	**Diff.**
Agriculture, hunting etc	5.32	5.07	-0.25	0.263	0.258	-0.005
Mining and quarrying	2.18	2.3	0.12	0.200	0.191	-0.009
Manufacturing	13.04	11.32	-1.72	0.243	0.223	-0.021
Electricity and utilities	0.67	0.65	-0.02	0.214	0.184	-0.030
Construction	8.89	9.18	0.29	0.273	0.269	-0.004
Wholesale and retail	27.63	27.71	0.08	0.301	0.304	0.003
Transport and related	5.82	6.21	0.39	0.231	0.236	0.005
Financial industry	14.72	15.11	0.39	0.264	0.260	-0.004
Community, social work etc.	16.4	17.74	1.34	0.176	0.178	0.002
Private households	5.33	4.69	-0.64	0.149	0.130	-0.019
Total	**100**	**100**		**0.237**	**0.232**	**-0.005**

Notes:

1. All means and proportions are weighted.
2. The "Diff." column is the difference between the mean of the relevant variable in the Post- period relative to the Pre- period.
3. The Pre- and Post- periods are represented by waves 1-12, and waves 13-14 respectively.

In terms of the impact of the ETI, we are more interested in whether any sector has shifted its employment disproportionately in favour of youth relative to non-youth employees. In columns 4 and 5 of Table 5, we present the mean proportion of the employed that are youth, within each

[20] Note that the sampling frame excludes workers' hostels (QLFS metadata documents, various waves, StatsSA). This will affect our estimates of the employment levels in the mining sector substantially.

industry. Overall, 23.7% of those employed in the *Pre-* period are youth, and this decreases to 23.2% in the *Post-* period. The most youth intensive employing industries are wholesale and retail, construction, agriculture, finance and manufacturing, where the proportion of the employed that are youth range from 30.1% to 24.3%. In most industries, the change in the proportion of employees that are youth are very small in magnitude, and they are often negative in sign. Manufacturing is the only large sector that does show a substantial change, although it indicates a decrease in the proportion of employees that are youth, of 2.1 percentage points. What we conclude from this component of our analysis is that the ETI did not lead to any substantial shift towards youth employment in any of the individual industrial categories that we have.

Section 6.4: An analysis of the occupations in which youth are employed

In Table 6 we document the occupational categories in which youth are employed, as well as the proportion of employees within each occupational category that are youth. In the period from 2011 to 2013, the most frequently observed occupational category amongst employed youth was elementary occupations, at almost 25%. This was followed by the service workers and sales, clerks, crafts and trades, and technical and associate professionals; at 18.1%, 14.1%, 12.7% and 9.6% respectively. Together, these sectors accounted for almost 80% of the occupational categories in which employed youth are employed. In 2014, we observe relatively small changes in the occupations in which youth are employed. None of the occupations show a change that is greater than one percentage point in absolute value, and the overall distribution of occupations seems very stable across the *Pre-* and *Post-* periods.

When we look at the fraction of workers within each occupational category that are youth, we find that clerks, and service works and sales are the most youth intensive occupations, with 31.6% and 29.4% of the employees in these categories being youth respectively. These are followed by elementary occupations, and crafts and trades, in which 27% and 25.1% of employees are youth. As with our analysis of the industries that employ youth, when we compare the *Pre-* and *Post-* proportions of employees that are youth within occupational categories, we observe that the changes are generally small in magnitude and often negative in sign. The one exception is the crafts and related trades occupation, in which the proportion of employees that are youth increased by 2.3 percentage points. Overall, this component of the analysis implies that the ETI did not lead to any substantial shift towards youth employment in most of the occupational categories captured in the QLFS.

Table 6: Distribution of occupations conditional on youth employment, and proportion of employed within occupations that are youth.

	Distribution of Occupations amongst employed youth (%)			Prop. of employed within occupations that are youth		
Occupation	**Pre-**	**Post-**	**Diff.**	**Pre-**	**Post-**	**Diff.**
Legislators, senior officials etc.	4.1	4.53	0.43	0.117	0.120	0.003
Professionals	5.35	5.21	-0.14	0.214	0.202	-0.012
Technical and assoc. professionals	9.58	9.44	-0.14	0.203	0.208	0.005
Clerks	14.14	14.55	0.41	0.316	0.311	-0.004
Service workers	18.09	17.89	-0.20	0.294	0.274	-0.019
Skilled agricultural workers	0.28	0.2	-0.08	0.137	0.111	-0.026
Craft and related trades	12.67	13.62	0.95	0.251	0.274	0.023
Plant and machine operators	7.32	6.63	-0.69	0.204	0.184	-0.021
Elementary occupations	24.79	24.65	-0.14	0.270	0.262	-0.008
Domestic workers	3.66	3.29	-0.37	0.128	0.114	-0.014
Total	**100**	**100**		**0.237**	**0.232**	**-0.005**

Notes:

1. All means and proportions are weighted.
2. The "Diff." column is the difference between the mean of the relevant variable in the Post- period relative to the Pre- period.
3. The Pre- and Post- periods are represented by waves 1-12, and waves 13-14 respectively.

6.5 Before-After regression results

In this section, we present and discuss the results from our before-after regression models. In column 1 of Table 7, the estimation sample includes the entire set of youth in our sample from ages 18 to 29 inclusive, from both urban and non-urban areas. The dependent variable is employed, and we have suppressed the coefficients on the control variables for brevity.[21] As discussed in our methods section, our coefficients of interest correspond to the *post1* and *post2* variables. They measure the difference in the employment probability of youth in the first and second quarters of 2014 relative to what one would expect to observe based on all of the covariates and the pre-existing trend.

The first thing to note is that the coefficients in column 1 are negative in sign, when we would have anticipated that they would be positive if the ETI was having a positive impact on youth employment probabilities. The second pertinent observation is that they are extremely small in magnitude. For

[21] The full regression results are available from the authors upon request.

example, the coefficient on *post1* of 0.00162 indicates that that out of every 10 000 youth, we estimate that there are 16.2 fewer employed youth than we would have expected, in the first quarter of 2014. Similarly, the coefficient on *post2* of 0.000142 indicates that that out of every 10 000 youth, we estimate that there are 1.42 fewer employed youth than we would have expected, in the second quarter of 2014. The third relevant piece of information is that the corresponding standard errors are very small. This derives from the large sample sizes that we have, and indicates that our coefficients are quite precisely estimated. Thus, the 95% confidence intervals for the coefficients on *post1* and *post2* are (-0.0127, 0.0095) and (-0.0123, 0.0120) respectively. For all intents and purposes, this is consistent with a 'zero effect' of the ETI on the employment probability of youth.

Table 7: Before-After regressions on youth employment probability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Employed	Employed	Employed formal	Employed formal	Employed	Employed	Employed formal	Employed formal
Skill level	All skills	All skills	All skills	All skills	Unskilled	Unskilled	Unskilled	Unskilled
Area-type	All areas	Urban	All areas	Urban	All areas	Urban	All areas	Urban
Variables								
t	0.000177	-0.000887	0.000507	-0.000248	0.000119	-0.00111	0.000344	-0.000595
	(0.000557)	(0.000737)	(0.000507)	(0.000687)	(0.000569)	(0.000761)	(0.000507)	(0.000692)
post1	-0.00162	0.00189	-0.00368	-0.000784	0.00121	0.00723	-0.000530	0.00525
	(0.00568)	(0.00760)	(0.00518)	(0.00706)	(0.00589)	(0.00805)	(0.00535)	(0.00742)
post2	-0.000142	9.60e-05	-0.00145	0.000556	0.000356	0.000735	-0.000489	0.00201
	(0.00620)	(0.00820)	(0.00550)	(0.00736)	(0.00638)	(0.00860)	(0.00557)	(0.00759)
Constant	-1.713***	-2.249***	-1.284***	-1.864***	-1.807***	-2.344***	-1.418***	-2.010***
	(0.0803)	(0.109)	(0.0750)	(0.104)	(0.0815)	(0.112)	(0.0751)	(0.106)
Observations	246,941	149,603	246,941	149,603	230,366	136,660	230,366	136,660
R-squared	0.198	0.196	0.190	0.184	0.167	0.164	0.142	0.135

Notes:

1. Standard errors in parentheses.

2. Asterisks denote statistical significance as follows:

*** p<0.01, ** p<0.05, * p<0.1

3. Suppressed coefficients on age, age squared, race dummies, male, tertiary (Col 1-4), urban (Col 1,3,5,7), province dummies and quarter dummies.

4. Regressions incorporate survey weights.

5. Standard errors are clustered at the PSU level.

In the remaining columns in Table 7, the substantive findings are the same as those obtained from column 1. Our overall findings remain unchanged when we consider only formal sector employment, or focus only on youth in urban areas, or unskilled youth only, or combinations of these groups and

dependent variables; the coefficients are small in magnitude, are not statistically significantly different from zero, and have fairly small standard errors.

6.6 Difference-in-differences regression results

In this section, we present and discuss the results from our main regression models. In column 1 of Table 8 we present our most preferred regression results.[22] The dependent variable is *employed*, the estimation sample includes the entire working aged population in our sample from ages of 18 to 64 inclusive, and we have not restricted the sample by geography or by skill level. As with our presentation and discussion of our 'before-after' regression models, we have suppressed the coefficients on the control variables.[23]

We notice that the *trend* variable has a very small coefficient and is not statistically significantly different from zero. This implies that the trend in employment probabilities was flat over the period of observation. The fact that the coefficient estimates on *post1* and *post2* are also small and not statistically significant indicates that there was no break in the trend for the non-youth in either the first or second quarters of 2014. The *youth* coefficient of 0.0517 is surprising in that it is positive and significant, but it needs to be interpreted in conjunction with the *age* and *age squared* coefficients, as there is a mechanical mapping from age to the *youth* variable. Once one accounts for the effects of the *age* and *age squared* variables, the results accord with our expectations based on the summary statistics discussed above.[24]

As discussed in the methods section, the primary coefficients of interest to us are those that correspond to *youth_post1* and *youth_post2*. The first of these, at -0.0056, indicates that, conditional on all the other covariates, the youth in the first quarter of 2014 were just over one half of a percentage point less likely to be employed than the comparison group of non-youth in that quarter. Similarly, the coefficient of -0.0053 indicates that, conditional on all of the other covariates, the youth in the second quarter of 2014 were 0.53 of a percentage point less likely to be employed than the comparison group of non-youth in that quarter.

[22] These are our more preferred results for two reasons. First, this specification allows for the largest sample size, which in turn gives us the most statistical power with the available data. Second, the ETI applies to youth in both urban and rural areas, and to both skilled and unskilled youth. Moreover, it is not clear how good a proxy the Stats SA definition of the formal sector is in terms of its ability to identify the eligibility of a firm.

[23] As before, these are available from the authors on request.

[24] This point is true for all of the regression results presented in this table.

Note that neither of the two point estimates is statistically significantly different from zero, and the corresponding confidence intervals are (-0.016, 0.005) and (-0.017, 0.006) for the first and second quarters of 2014 respectively. These are fairly precisely estimated coefficients, which just barely include the value of zero. What this implies is that, even if the true population parameter is positive, it is likely to be extremely small in magnitude.

Table 8: Difference-in-Differences regressions on youth employment probability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Employed	Employed	Employed formal	Employed formal	Employed	Employed	Employed formal	Employed formal
Skill level	All skills	All skills	All skills	All skills	Unskilled	Unskilled	Unskilled	Unskilled
Area-type	All Areas	Urban	All Areas	Urban	All Areas	Urban	All Areas	Urban
Variable								
t	0.000493	-0.000560	0.000650*	1.80e-05	0.000691	-0.000491	0.000735*	-4.97e-05
	(0.000406)	(0.000488)	(0.000393)	(0.000501)	(0.000433)	(0.000530)	(0.000409)	(0.000528)
post1	0.00136	0.00352	0.00411	0.00561	0.00173	0.00345	0.00512	0.00698
	(0.00414)	(0.00508)	(0.00437)	(0.00552)	(0.00461)	(0.00582)	(0.00483)	(0.00629)
post2	0.000857	0.00149	-0.000622	-0.000296	0.00193	0.00223	0.00175	0.00227
	(0.00465)	(0.00564)	(0.00459)	(0.00569)	(0.00501)	(0.00620)	(0.00492)	(0.00624)
youth	0.0517***	0.0660***	0.0345***	0.0497***	0.0570***	0.0710***	0.0422***	0.0577***
	(0.00512)	(0.00636)	(0.00477)	(0.00612)	(0.00553)	(0.00703)	(0.00513)	(0.00677)
youth_post1	-0.00557	-0.00589	-0.00911*	-0.00900	-0.00456	-0.00257	-0.00800	-0.00569
	(0.00547)	(0.00720)	(0.00540)	(0.00723)	(0.00590)	(0.00795)	(0.00583)	(0.00801)
youth_post2	-0.00529	-0.00700	-0.00322	-0.00398	-0.00758	-0.00951	-0.00629	-0.00735
	(0.00591)	(0.00762)	(0.00555)	(0.00728)	(0.00625)	(0.00820)	(0.00582)	(0.00778)
age	0.0870***	0.0943***	0.0621***	0.0717***	0.0873***	0.0956***	0.0615***	0.0718***
	(0.000794)	(0.00102)	(0.000766)	(0.000991)	(0.000833)	(0.00109)	(0.000791)	(0.00104)
Age squared	-0.00102***	-0.00111***	-0.000744***	-0.000860***	-0.00102***	-0.00112***	-0.000737***	-0.000863***
	(9.21e-06)	(1.18e-05)	(8.90e-06)	(1.14e-05)	(9.60e-06)	(1.25e-05)	(9.09e-06)	(1.19e-05)
Constant	-1.378***	-1.396***	-1.010***	-1.094***	-1.384***	-1.424***	-0.993***	-1.101***
	(0.0184)	(0.0225)	(0.0181)	(0.0220)	(0.0197)	(0.0244)	(0.0191)	(0.0235)
Observations	667,610	429,286	667,610	429,286	596,984	371,004	596,984	371,004
R-squared	0.217	0.199	0.222	0.201	0.175	0.158	0.147	0.129

Notes:

1. Standard errors in parentheses

2. Asterisks denote statistical significance as follows:

*** p<0.01, ** p<0.05, * p<0.1

3. Suppressed coefficients on race dummies, male, tertiary (Col 1-4), urban (Col 1,3,5,7), province dummies and quarter dummies.

4. Regressions incorporate survey weights.

5. Standard errors are clustered at the PSU level.

When we consider the other regression models that we have fit to the data, we notice a remarkably consistent set of results. First, all of the relevant coefficients are negative in sign. Second, they are all small in magnitude, at less than one percentage point in absolute value. Third, with one exception, they are not statistically significantly different from zero.[25] Fourth, all of the relevant standard errors are small in magnitude, which implies correspondingly narrow confidence intervals. Taken together, we can be reasonably confident that the ETI has not had any substantial positive effect on aggregate youth employment probabilities in the short run.

Section 7: Robustness tests

We might have some remaining concerns related to the analysis and subsequent results presented above. These relate to the choice of estimator, the sensitivity of results to the choice of the dependent variable and the statistical power that we have given our sample size. In this section, we discuss the major concerns and address them to the extent possible.

7.1 Robustness checks on the validity of using a difference-in-differences estimator

There are three main assumptions required for a DD estimator to provide an unbiased estimate of the average treatment effect of the ETI. First, we require that the trend of the dependent variable was the same for both the youth and the non-youth in the *pre-* period. To test this, we estimated our DD regressions from Table 8 above, but included an interaction term *youth*trend*. This variable allows for the trends to differ between the treatment and control groups in the *pre-* period. The results of these regressions are included in Appendix Table A1. None of the coefficients on the *youth*trend* interaction term are statistically significant, and they are all extremely small in magnitude, with values below 0.001 in absolute value. We are thus empirically satisfied that our first requirement is met.

The second requirement is that there are no contemporary shocks in the economic environment that affect the treatment and control group differently. This requirement is impossible to test, but we can point out that the time trend for both youth and non-youth employment rates are quite stable across the *Pre-* and *Post-* periods. This is shown graphically in Figure 2, and is observable for the control groups by investigating the coefficients on *Post1* and *Post2* in the regressions in Table 8 above, all of which are small and statistically insignificant. Similarly, the results from the *before-after* estimates indicated that there was no trend break in the employment probabilities amongst youth either.

[25] The one exception is the coefficient corresponding to *youth_post1* in column 3. In this case, it is negative and statistically significantly different from zero.

The third requirement is that the ETI did not affect the employment probabilities of the control group. As mentioned in the theory section, this condition could be violated if there are significant substitution or complementary effects towards the groups of non-youth. We are not too concerned about this for two reasons. First, as discussed in the preceding paragraph, neither the youth nor the non-youth show any evidence of a trend break in their employment probabilities in the *Post-* period. This critique does not apply to the *before-after* estimator, and the *before-after* estimator and the *difference-in-differences* estimator both suggest the same conclusions.

Nonetheless, we explored the possibility that such substitution or complementary effects may be relevant for our estimates by repeating our main *DD* regression analyses but restricting the estimation sample to include only subsets of the non-youth as the control group. The subsets that we used were the *Almost youth* (aged 30-34), *Prime aged* (aged 35-49) and *Older adult* (aged 50-64). Our thinking was that some age groups, such as the *Almost youth*, may be more affected by the ETI while others, such as the *Prime adults*, would be relatively less so. The relevant coefficient estimates from these regressions are presented in Appendix Table A2. While the coefficients do show considerably greater variation, the general pattern is maintained in that the coefficients are not large, they are generally not statistically significant, and in all of the cases where they are significantly different from zero, they are negative in sign.

7.2 Sensitivity of the results to our choice of dependent variable

Our results up to this point indicate that the ETI has not had any significant positive effect on aggregate youth employment probabilities. Nonetheless, it may still be the case that the ETI is having an effect on the youth labour market. One possibility is that the 'net rate' at which youth are finding jobs has indeed increased, but that the time period has been too short for this to have an observable impact on the aggregate youth employment probabilities.[26] Alternatively, there may be an increase in the rate at which youth find employment but this is offset by a corresponding increase in the rate at which youth lose employment. This would be rational from a firm's perspective as only newly hired youth are eligible for the ETI. A different version of this behaviour would be that firms could hire youth who were previously employed in a different firm. This could result in a stable aggregate labour absorption rate, but a higher level of churning in the youth labour market.

[26] By 'net rate' we mean the rate at which youth who are not employed find employment, less the rate at which youth who are employed lose jobs.

To explore these possible scenarios, we regress the variables that capture whether an employed youth was recently employed[27] on the trend variable and the quarter dummies, as well as the *post1* and *post2* indicator variables. In essence, we are testing whether there is evidence of a positive trend break in the job finding rate amongst youth.[28] We perform the analogous analysis for a recent job loss on the subset of youth who are not employed.

Table 9: Regressions to test for changes in outflows or inflows into employment amongst youth

	(1)	(2)	(3)	(4)
Variables	recentjob_3	recentjob_6	recentloss_3	recentloss_6
t	0.00129***	0.00234***	0.000540**	0.000757**
	(0.000462)	(0.000793)	(0.000223)	(0.000323)
post1	0.00766	0.00368	-0.000681	-0.000317
	(0.00771)	(0.0103)	(0.00310)	(0.00388)
post2	-0.00337	0.00364	0.00306	0.00247
	(0.00603)	(0.0105)	(0.00302)	(0.00404)
2.qtr	-0.0243***	0.0180***	-0.00763***	-0.00287
	(0.00434)	(0.00546)	(0.00170)	(0.00210)
3.qtr	-0.0215***	-0.00786	-0.00660***	-0.00732***
	(0.00446)	(0.00633)	(0.00169)	(0.00218)
4.qtr	-0.0286***	-0.0176***	-0.00792***	-0.0100***
	(0.00452)	(0.00579)	(0.00178)	(0.00214)
Constant	0.0815***	0.165***	0.0390***	0.0631***
	(0.00407)	(0.00634)	(0.00200)	(0.00269)
Observations	66,764	66,764	180,177	180,177
R-squared	0.003	0.002	0.000	0.000

Notes:

1. Robust standard errors in parentheses

2. *** p<0.01, ** p<0.05, * p<0.1

The results of our regressions are presented in Table 9. In column 1, with the dependent variable being a recent job within the previous three months, we are interested in whether the coefficients on *post1* or *post2* are positive and statistically significant. The coefficient on *post1* is 0.00766, which is positive, but small and not statistically significant. The coefficient on *post2* is -0.00337, which is negative, but not statistically significant. In the second column, we are only interested in the

[27] As discussed in the data section, these variables are *recentjob_3* and *recentjob_6* for the job finding rate, and *recentloss_3* and *recentloss_6* for the job loss rate.

[28] The use of repeated cross sections is not ideal for this purpose. Longitudinal data, if available, are much better suited to the study of transitions into and out of employment. Unfortunately, at present there is no longitudinal data that covers the period in which the ETI came into effect.

coefficient on *post2*, as it is only in the second quarter of 2014 that a job found within the previous six months could be attributed to the introduction of the ETI. This point estimate is positive but small at 0.00364, and it is also not statistically significant. We thus find no evidence that the job finding rate had increased as a response to the introduction of the ETI.

With regard to job loss within the previous three or six months, we also find no evidence of a trend break in either the first or second quarters of 2014. The coefficients on the *post1* and *post2* variables are all small in magnitude and are not statistically significant. In summation, we find no evidence that the ETI has led to any increases in the rate at which youth find jobs. We also find no evidence to suggest that the ETI had any impact on the rate at which youth lose jobs. Thus, the churning hypothesis has no empirical support, to the extent that this is captured by the proxies that we used from the QLFS.

Section 7.3 Sample sizes and statistical power

A final concern with our results relates to the sample size and statistical power that we have, in terms of our ability to identify any effects of the ETI. *A priori*, this is a valid concern as, despite our large sample sizes, the plausible parameter values that the ETI could realistically generate are probably quite small.

It is useful to work through a simple numerical example that uses reasonable assumptions to gain a sense of the types of coefficients that we could expect to find. Based on the weights in the QLFS, which are themselves calibrated on the mid-year population estimates from the 2011 Census, we would expect that there are just below 12 million youth aged 18 to 29 in the country. The ETI has the stated objective of creating 178 000 new jobs over three years. If we pro-rate this 178 000 for a six month period, we would expect just below 30 000 new jobs for youth by the end of wave 14. Using a base population of almost 12 million, this would imply a coefficient of about 0.0025 in our main regressions, which is very small in magnitude.

The above example illustrates two things. First, it provides us with a sense of perspective with regard to the overall objective of the ETI, in relation to the magnitude of the overall youth unemployment level. In all of the 2013 QLFS data combined, about 60% of youth aged 18 to 29 were in the labour force, of whom about 51% are unemployed. This corresponds to approximately 3.66 million unemployed youth. Even if the longer term objective of 178 000 new jobs was achieved, and everything else remained the same, this would reduce the number of unemployed youth to 3.48

million. Given all of our assumptions, this would result in a new unemployment rate amongst youth of approximately 48.36%. Thus, even if the ETI does achieve its stated objective, there will remain an extremely high youth unemployment rate.

The purpose of our study is to measure the effects of the ETI on aggregate youth employment probabilities, and not to explicitly test whether the ETI is on course to achieve its stated objective. It is possible that the true effect of the ETI is sufficiently small in magnitude, such that we cannot reject that the ETI is on track to achieve its objective, and simultaneously, that we cannot reject that the true population parameter is different to zero. This arises due to a combination of sample sizes, survey design, statistical power and most importantly, that a coefficient of 0.0025 and a coefficient of 0 are very close together, given that our dependent variable is binary.

The second point that the illustrative example makes clear is that some nuance is required in how we interpret our results. For example, a coefficient of 0.01 could be interpreted as economically substantial or insubstantial, depending on whether one benchmarks on the absolute number of new jobs that this coefficient entails and the ensuing improvement in welfare, or on the number of youth that are unemployed, which will remain large even if 120 000 new jobs became available for youth.

We thus encounter a challenging empirical problem. We want to test whether the ETI is having a statistically significant and positive effect on youth employment probabilities, but all we have shown is that there is insufficient statistical evidence to claim that this is the case. The point is subtle but very important, because we have not shown any evidence that the ETI has not had a positive and significant effect on youth employment probabilities either. What we can do is test whether the coefficients from our regression models are greater than or equal to some plausible expectations, but then it becomes crucial to have plausible expectations. The key word here is ‘plausible’, and one needs a clear sense of what determines whether an expected effect is plausible or not. One criterion would be to use the official objectives of the ETI, but we have no information about how these objectives were reached and why they should be accepted as plausible. Theoretically, what we would like to know are two key parameters; the wage elasticity of aggregate labour demand for youth labour, and the rate at which adjustments in the youth labour market occur. If we knew these two parameters, we could calculate what effects the ETI would likely have had within a six month period, and we could then test whether these were in fact achieved. Unfortunately, we also have no good empirical measures of either of these parameters.

Given all of these limitations, we decided to perform multiple one-tailed hypothesis tests on the coefficients on *youth_post1* and *youth_post2* from our *DD* regressions, over a range of values. To be precise, we tested whether the true population parameter corresponding to the relevant coefficient was greater than or equal to 0.01, 0.005, 0.0025 and 0.001. The p-value from each of these hypothesis tests are presented in Appendix Table A3.

As with our discussion of our main regression results, we focus primarily on the hypothesis tests that correspond to our most preferred regression, that is the regression with *employed* as the dependent variable, youth of all skills as the treatment group, non-youth of all skills and all ages as the control group, and including both urban and non-urban areas. The p-values for the tests for this regression are presented in column 1. For the coefficient on *post2*, the p-value that the true parameter is greater than or equal to 0.01 (or 1%) is 0.005, so we reject the null hypothesis at the 1% level of significance. Similarly, we reject the null hypothesis that the true population parameter is greater than or equal to 0.005 at the 5% level of significance. When testing for whether the true population parameter is greater than or equal to 0.0025, which is approximately the value that we would expect if the ETI was on course to achieve its overall objective assuming a *pro rata* objective at month 6, we reject the null hypothesis at the 10% level of significance. What this implies is that, while we can never be absolutely certain about the precise magnitude of the effects of the ETI on youth employment probabilities in the short run, we can be confident that the upper bound of these effects are, at best, small in magnitude.[29]

Section 8: Conclusion

The Employment Tax Incentive was introduced in January 2014 to address a large and persistent unemployment problem amongst youth, particularly unskilled youth. In this paper, we make use of fourteen waves of nationally representative QLFS data to investigate the short run effects of the ETI on the aggregate employment probabilities amongst youth.

In the first six months since the introduction of the ETI, we find no evidence that the ETI had any substantial, positive and statistically significant effect on aggregate youth employment probabilities. Our preferred estimate of the effects of the ETI has a 95% confidence interval of (-0.017, 0.006). We can thus be fairly confident that, at best, the effects of the ETI are small in magnitude. We also find

[29] The hypothesis tests from the other regressions indicate qualitatively the same result. There is some variation in terms of the value of the threshold that we would reject, but the overall point remains that the upper bound of the effects of the ETI are fairly small in probability space.

no evidence that the rate at which youth find or lose employment has changed since the ETI was introduced.

There are several reasons why the ETI may not be having an impact on youth employment probabilities in the short run, and these explanations are not mutually exclusive. First, it may be that the ETI has not had any effects yet, but that with time it may become more effective. This could arise if firms take some time to understand the rules of the ETI, or if they require more time to adjust their methods of production.[30] Whether the ETI does or does not generate substantial additional youth employment with enough time is a hypothesis can be tested with future waves of data.

Second, the way in which the ETI has been implemented may itself limit its effects. The ETI is a tax incentive and not a wage subsidy, and as such should not have any effect on informal sector firms. In addition, relatively small firms that are not registered for PAYE are also excluded from benefiting from the incentive. In all likelihood, the ETI thus effectively serves as an incentive that is targeted at medium sized and larger formal sector firms, which would further limit its potential impact.[31]

Third, the value of the incentive may be too low to substantially affect firms' hiring decisions. Related to the potential issue of the value of the incentive is a question about the magnitude of the wage elasticity of labour demand for young and relatively unskilled workers. For the ETI to have a substantial impact, the labour demand curve for eligible workers, amongst firms that are effectively able to take up the incentive, needs to be relatively elastic. There is very little evidence to suggest that this is the case, and there is very little evidence to suggest that this is not the case. A qualitative study amongst firms to investigate their awareness and responsiveness to the ETI, as well as their reasons for optimizing the way that they do, could be very useful here. Such an analysis would also have to seriously consider the multiple reasons, and their relative importance, for the persistently high youth unemployment rates in South Africa.

[30] Note that for these effects to occur with time we would nonetheless require a shift in behaviour from firms relative to their responses during the first six month of 2014. The issue is not about statistical power. Our sample size is so large that we would probably identify even the 56 000 jobs that were subsidized in January 2014 alone (Gordhan, 2014), if they had been new hires in new positions rather than subsidies accruing to new hires in existing positions. This argument becomes even stronger when we consider instead the 133 000 subsidized jobs mentioned by President Zuma in his State of the Nation Address (Zuma, 2014).

[31] This expectation about which firms are likely to benefit from, and thus respond to, the ETI is a conjecture. It would be interesting to obtain data from SARS that would give the distribution and value of ETI related tax relief amongst firms of various sizes.

Fourth, within the firms that are eligible to benefit from the ETI, it may be the case that the decision makers who decide on whether to employ a person or not are personally unaffected by the ETI. For example, in a large franchised supermarket chain, it may be that the employment decisions are made by local branch managers but that the ETI benefits accrue to a centralized headquarters. Unless the incentive is passed on to the branch managers, the ETI will have no effect on employment in such a firm. Similarly, if the incentive accrues to labour brokers, who do not pass the incentive on to the ultimate employers, then the employment related effects will be also be muted.[32]

From a policy perspective, there are several observations that flow from our study that are worth considering. First, high levels of youth unemployment have potential implications for aggregate welfare, for economic growth and development, for human capital formation and possible implications for social and political stability; and the problem remains as large as it did prior to the introduction of the ETI. Moreover, the scale of the program is actually quite modest relative to the number of unemployed youth. Thus, even if the ETI were to achieve its stated objective of creating 178 000 net new jobs over a three year period, the youth unemployment levels would remain exceptionally high.

Second, while the state might care about the aggregate levels of youth employment, a profit maximizing firm will optimize its employment decisions by considering the net marginal cost of an employee, including any incentive attached to the marginal employee, relative to the marginal benefit of employing this marginal employee. The aggregate level of tax relief received due to the ETI is thus not relevant for a profit maximizing firm's decision on whether or not to hire an additional employee. If we conceptualize the state as the principal and the firm as the agent, this lack of convergence in objectives means that it is unlikely that an optimal outcome will be reached, in terms of youth employment levels, from the perspective of the state.

For example, consider an eligible firm that planned on hiring 50 youth workers at R4000 per worker per month, without the ETI. With the introduction of the ETI, the net cost to this firm of these 50 employees will drop from R200 000 to R150 000 per month, since each worker will qualify for tax relief of R1000 per month. The firm will thus receive aggregate tax relief of R50 000, even though there will be no impact on youth employment levels. This firm will then decide on whether the 51st youth employee, i.e. the marginal employee, will add enough value to the firm to warrant their net wage. The relevant question for the firm is whether the benefit of this marginal employee is greater

[32] There is some evidence that labour brokers have accounted for a significant position of the take-up (van Rensburg, 2014).

than the net marginal cost of R3000 per month.[33] The ETI thus only has an impact on youth employment levels by changing the firm's decisions with regard to marginal employees, while the R50 000 of tax relief per month that the firm receives for its original planned employment has no bearing on its cost-benefit calculation at the margin.

Third, the lack of effectiveness of the ETI has implications for the efficacy of policy from a public finance perspective. As illustrated in the example just discussed, if there is substantial take up of the incentive for employment that would have arisen even in the absence of the ETI, this represents a pure transfer from taxpayers to a subset of firms who are not doing anything differently. These transfers have opportunity costs, and it is difficult to believe that this is desirable from a policy perspective.

Fourth, it is worth pointing out that for a modest increase in the number of employment positions available for youth, the effective aggregate subsidy per new post from the government's perspective could still be exceptionally large. This is because all new appointments can benefit from the ETI, but only new appointments in new positions will have an impact on the labour absorption rate.[34] Moreover, this possibility is made even more likely because the youth labour market is characterized by high levels of turnover.[35]

For example, by applying the sampling weights to all of our data from 2011 to 2013, we estimate that there were about 1.737 million youth who found new employment in the formal sector in firms with 5 or more employees in the three years prior to the introduction of the ETI. The ETI expects to subsidize 423 000 jobs, of which 178 000 are expected to be in new positions, over a period of three years. This implies that the expectation is that the ETI would subsidize 245 000 jobs over the three years, that would have arisen even in the absence of the ETI. It is unclear how this number of 245 000 was estimated, but it is much smaller than our estimate of the number of new hires in the three years prior to the introduction of the ETI. Of course, not all new hires are eligible for the subsidy, and not all employees will take up the incentive, and the turnover rates might change; but it is nonetheless extremely difficult to reconcile these two numbers. One has to consider the possibility

[33] R4000 wages less R1000 tax relief per month.

[34] In the previous example, suppose that the firm raised its employment from the original planned 50 up to 55, because of the ETI. Under these assumptions, the aggregate tax relief would be R55 000 per month, for 5 additional jobs, which pay a total of R20 000 per month to these additional employees.

[35] In the four waves of QLFS data from 2013, the average proportion of recent appointments in formal sector appointments was 0.064 amongst youth aged 18 to 29, in each wave. Amongst *almost youth* (30-34), *prime aged* (35-49) and *older adult* (50-64) respondents, the corresponding proportions were 0.038, 0.025 and 0.014 respectively.

that the 245 000 new hires in existing posts is a gross underestimate of the number of subsidized jobs into pre-existing posts or posts that would have been created even in the absence of the ETI. Since subsidized new hires in existing posts represents a pure cost to the ETI with no benefit in terms of aggregate employment, one can clearly see how the ratio of aggregate subsidy claimed, to aggregate wages paid in new positions, might realistically exceed 100%.[36]

Our overall conclusions are somewhat disappointing. The ETI does not seem to be increasing youth employment levels substantially, and there is some chance that it might never have the impact that was desired. At the same time, the labour market issue of youth unemployment, and the public finance implications of the ETI, are both extremely important policy topics in contemporary South Africa. Further research is required to understand all of the many layers of complexity of both the ETI as well as the youth unemployment problem.

[36] Note that for certain parameter values, this ratio could easily exceed 100% by an order of magnitude. This would arise if the ETI only generates a `small' number of new hires in new positions, while the take up rates are high for new hires in existing positions. We cannot make precise predictions without two sources of additional data; administrative data on take up rates and the average value of the tax relief claimed per employee, and more waves of the QLFS. Unfortunately, we can only get more waves of the QLFS as time passes, and by then the effective costs will already have been incurred.

References

Betcherman, G., Olivas, K., & Dar, A. (2004) “Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries”, World Bank, Social Protection Discussion Paper Series, Paper No. 402, Washington.

Betcherman, G., Godfrey, M., Puerto, S., Rother, A. & Stavreska, A. (2007) ‘A review of interventions to support young workers: findings of the Youth Employment Inventory”, *SP Discussion Paper No. 0715*, World Bank.

Burns. J, Edwards. L and Pauw. K (2010) “Wage Subsidies to Combat Unemployment and Poverty: Assessing South Africa’s Options “, *Southern Africa Labour and Development Research Unit Working Paper Number 45*, Cape Town: SALDRU, University of Cape Town.

Card, D. & Krueger, A. (1994) “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania”, *American Economic Reviewew* 84(4), pp. 772-793.

Congress of South African Trade Unions (COSATU). (2012) “The Youth Wage Subsidy in South Africa: Response of the Congress of South African Trade Unions (To the National Treasury and Democratic Alliance)”. COSATU.

Gordhan, P. (2014). “Budget Speech, February 2014”. Pretoria: National Treasury.

Levinsohn, J. (2007) “Two policies to alleviate unemployment in South Africa”, *Centre for International Development at Harvard University Working Paper 166*.

Levinsohn, J., Rankin, N., Roberts, G. & Schoer, V. (2014) “Wage subsidies and youth employment in South Africa: Evidence from a randomised control trial”, *Stellenbosch Economic Working Papers: 02/14*.

Meth, C. (2011) “Employer of Last Resort? South Africa’s Expanded Public Works Programme (EPWP)”, *A Southern Africa Labour and Development Research Unit Working Paper Number 58*. Cape Town: SALDRU, University of Cape Town

National Treasury (2011). “Confronting Youth unemployment: Policy options for South Africa”, Pretoria: National Treasury.

National Treasury (2014). “Budget Review 2014”. Pretoria: National Treasury.

Puerto, O. (2007) “Interventions to support young workers in Latin America and the Caribbean: Regional Report for the Youth Employment Inventory”, World Bank.

Schoer, V. & Rankin, N. (2011) “Youth employment, recruitment and a youth targeted wage subsidy: Findings from a South African firm level survey”, *Human Development Unit African Region Paper No. 72601*, World Bank.

Smith, C. (2006) “International experience with worker-side and employer-side wage and employment subsidies, and job search assistance programmes: Implications for South Africa”, HSRC.

van Rensburg, D. (2014). “Brokers profit from youth tax incentive”. City Press. 31 August 2014.

Zuma, JG. (2014). “State of the Nation Address, June 2014”. Pretoria: The Presidency.

Section 9: Appendix

Table A1: Difference-in-Differences regressions on youth employment probability - with youth trend differential

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	employed	employed	employed formal	employed formal	employed	employed	employed formal	employed formal
Skill level	All skills	All skills	All skills	All skills	Unskilled	Unskilled	Unskilled	Unskilled
Area-type	all areas	urban	all areas	urban	all areas	urban	all areas	urban
Variables								
trend	0.000740	-0.000402	0.000834*	0.000185	0.00108**	-0.000197	0.00102*	0.000211
	(0.000501)	(0.000589)	(0.000497)	(0.000618)	(0.000548)	(0.000659)	(0.000531)	(0.000675)
trend*youth	-0.000655	-0.000444	-0.000489	-0.000472	-0.000982	-0.000777	-0.000711	-0.000692
	(0.000671)	(0.000873)	(0.000636)	(0.000841)	(0.000705)	(0.000930)	(0.000658)	(0.000887)
post1	-0.000232	0.00251	0.00292	0.00454	-0.000764	0.00157	0.00331	0.00530
	(0.00430)	(0.00525)	(0.00455)	(0.00573)	(0.00479)	(0.00603)	(0.00505)	(0.00659)
post2	-0.000979	0.000321	-0.00199	-0.00154	-0.000954	6.00e-05	-0.000344	0.000335
	(0.00498)	(0.00598)	(0.00491)	(0.00607)	(0.00544)	(0.00669)	(0.00531)	(0.00674)
youth	0.0560***	0.0689***	0.0377***	0.0528***	0.0634***	0.0761***	0.0469***	0.0622***
	(0.00662)	(0.00856)	(0.00634)	(0.00835)	(0.00713)	(0.00942)	(0.00674)	(0.00910)
youth_post1	-0.00134	-0.00303	-0.00596	-0.00596	0.00178	0.00244	-0.00341	-0.00123
	(0.00612)	(0.00805)	(0.00583)	(0.00777)	(0.00654)	(0.00877)	(0.00632)	(0.00864)
youth_post2	-0.000408	-0.00370	0.000425	-0.000469	-0.000255	-0.00372	-0.000992	-0.00220
	(0.00716)	(0.00913)	(0.00656)	(0.00853)	(0.00765)	(0.00996)	(0.00691)	(0.00918)
age	0.0870***	0.0943***	0.0621***	0.0717***	0.0873***	0.0956***	0.0615***	0.0718***
	(0.000794)	(0.00102)	(0.000766)	(0.000991)	(0.000833)	(0.00109)	(0.000791)	(0.00104)
Age squared	-0.00102***	-0.00111***	-0.000744***	-0.000860***	-0.00102***	-0.00112***	-0.000737***	-0.000863***
	(9.21e-06)	(1.18e-05)	(8.90e-06)	(1.14e-05)	(9.60e-06)	(1.25e-05)	(9.09e-06)	(1.19e-05)
Constant	-1.379***	-1.397***	-1.012***	-1.095***	-1.387***	-1.426***	-0.994***	-1.102***
	(0.0185)	(0.0227)	(0.0182)	(0.0222)	(0.0198)	(0.0246)	(0.0192)	(0.0237)
Observations	667,610	429,286	667,610	429,286	596,984	371,004	596,984	371,004
R-squared	0.217	0.199	0.222	0.201	0.175	0.158	0.147	0.129

Notes:

1. Standard errors in parentheses

2. Asterisks denote statistical significance as follows:

*** p<0.01, ** p<0.05, * p<0.1

3. Suppressed coefficients on race dummies, male, tertiary (Col 1-4), urban (Col 1,3,5,7), province dummies and quarter dummies.

4. Regressions incorporate survey weights.

5. Standard errors are clustered at the PSU level.

Table A2: Coefficient estimates from main Difference-in-Differences regressions on youth employment probability - various control groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	employed	employed	employed formal	employed formal	employed	employed	employed formal	employed formal
Skill level	All skills	All skills	All skills	All skills	Unskilled	Unskilled	Unskilled	Unskilled
Area-type	All areas	urban	All areas	urban	All areas	urban	All areas	urban
Control group: Almost youth								
Variables								
youth_post1	0.00655	0.00348	-0.00202	0.000953	0.00608	0.00496	-0.00382	0.000876
	(0.00868)	(0.0108)	(0.00846)	(0.0105)	(0.00961)	(0.0122)	(0.00937)	(0.0120)
youth_post2	0.00831	0.00451	0.0135	0.0173	0.00445	0.00115	0.00722	0.0117
	(0.00920)	(0.0112)	(0.00901)	(0.0113)	(0.00997)	(0.0123)	(0.00976)	(0.0125)
Control group: Prime aged								
Variables								
youth_post1	-0.00547	-0.00683	-0.0105	-0.0129	-0.00459	-0.00349	-0.00953	-0.00999
	(0.00647)	(0.00825)	(0.00674)	(0.00877)	(0.00709)	(0.00927)	(0.00739)	(0.00988)
youth_post2	-0.00265	-0.00482	-0.00364	-0.00704	-0.00322	-0.00497	-0.00518	-0.00868
	(0.00699)	(0.00880)	(0.00667)	(0.00856)	(0.00753)	(0.00964)	(0.00716)	(0.00936)
Control group: Older adult								
Variables								
youth_post1	-0.0153**	-0.0120	-0.0122*	-0.00987	-0.0124	-0.00681	-0.00826	-0.00300
	(0.00756)	(0.00989)	(0.00715)	(0.00963)	(0.00809)	(0.0109)	(0.00748)	(0.0104)
youth_post2	-0.0212**	-0.0210*	-0.0156**	-0.0152	-0.0247***	-0.0262**	-0.0189**	-0.0199*
	(0.00828)	(0.0109)	(0.00762)	(0.0101)	(0.00882)	(0.0119)	(0.00799)	(0.0109)

Notes:

1. Standard errors in parentheses

2. Asterisks denote statistical significance as follows:

*** p<0.01, ** p<0.05, * p<0.1

3. Regression models are identical to those in Table 8, but all other coefficients are suppressed.

4. Regressions incorporate survey weights.

5. Standard errors are clustered at the PSU level.

6. Control groups: Almost youth aged 30-34, Prime aged 35-49, Older adult aged 50-64.

Table A3: One tailed tests for whether true population parameter is greater than some threshold.

	All youth and all non-youth				Unskilled youth and unskilled non-youth			
Depvar	Employed	Employed	Employed Formal	Employed Formal	Employed	Employed	Employed Formal	Employed Formal
Area-type	All areas	Urban	All areas	Urban	All areas	Urban	All areas	Urban
Coefficients and Standard Errors								
youth_post1	-0.00557	-0.00589	-0.00911*	-0.00900	-0.00456	-0.00257	-0.00800	-0.00569
	(0.00547)	(0.00720)	(0.00540)	(0.00723)	(0.00590)	(0.00795)	(0.00583)	(0.00801)
youth_post2	-0.00529	-0.00700	-0.00322	-0.00398	-0.00758	-0.00951	-0.00629	-0.00735
	(0.00591)	(0.00762)	(0.00555)	(0.00728)	(0.00625)	(0.00820)	(0.00582)	(0.00778)
Coef. on youth_post1	P-value for corresponding hypothesis test							
Ho: B>=0.01	0.002	0.014	0.000	0.004	0.007	0.057	0.001	0.025
Ho: B>=0.005	0.027	0.065	0.005	0.026	0.053	0.170	0.013	0.091
Ho: B>=0.0025	0.070	0.122	0.016	0.056	0.116	0.262	0.036	0.153
Ho: B>=0.001	0.115	0.169	0.031	0.083	0.173	0.327	0.061	0.202
Coef. on youth_post2	P-value for corresponding hypothesis test							
Ho: B>=0.01	0.005	0.013	0.009	0.027	0.002	0.009	0.003	0.013
Ho: B>=0.005	0.041	0.058	0.069	0.109	0.022	0.038	0.026	0.056
Ho: B>=0.0025	0.094	0.106	0.151	0.187	0.053	0.072	0.065	0.103
Ho: B>=0.001	0.144	0.147	0.224	0.247	0.085	0.100	0.105	0.141

Notes:

1. These coefficients and hypothesis tests correspond to the regressions in Table 8

southern africa labour and development research unit

The Southern Africa Labour and Development Research Unit (SALDRU) conducts research directed at improving the well-being of South Africa's poor. It was established in 1975. Over the next two decades the unit's research played a central role in documenting the human costs of apartheid. Key projects from this period included the Farm Labour Conference (1976), the Economics of Health Care Conference (1978), and the Second Carnegie Enquiry into Poverty and Development in South Africa (1983-86). At the urging of the African National Congress, from 1992-1994 SALDRU and the World Bank coordinated the Project for Statistics on Living Standards and Development (PSLSD). This project provide baseline data for the implementation of post-apartheid socio-economic policies through South Africa's first non-racial national sample survey.

In the post-apartheid period, SALDRU has continued to gather data and conduct research directed at informing and assessing anti-poverty policy. In line with its historical contribution, SALDRU's researchers continue to conduct research detailing changing patterns of well-being in South Africa and assessing the impact of government policy on the poor. Current research work falls into the following research themes: post-apartheid poverty; employment and migration dynamics; family support structures in an era of rapid social change; public works and public infrastructure programmes, financial strategies of the poor; common property resources and the poor. Key survey projects include the Langeberg Integrated Family Survey (1999), the Khayelitsha/Mitchell's Plain Survey (2000), the ongoing Cape Area Panel Study (2001-) and the Financial Diaries Project.

www.saldru.uct.ac.za
Level 3, School of Economics Building, Middle Campus, University of Cape Town
Private Bag, Rondebosch 7701, Cape Town, South Africa
Tel: +27 (0)21 650 5696
Fax: +27 (0) 21 650 5797
Web: www.saldru.uct.ac.za