The Employment effects of the South African Youth Wage Subsidy

Amina Ebrahim,[1] Murray Leibbrandt[2], and Vimal Ranchhod[3]

October 2018

DRAFT: Please do not cite

[1] PhD Canditate, SALDRU, University of Cape Town, South Africa, corresponding author: amina.ebrahim@gmail.com
[2] Director, SALDRU, University of Cape Town, South Africa
[3] Deputy Director, SALDRU, University of Cape Town, South Africa

Abstract: South Africa's Employment Tax Incentive, launched in 2014, aimed to address high youth unemployment rates by reducing the cost of hiring young workers. We make use of anonymized tax data from the 2012–2015 tax years to examine the effect of the subsidy on youth employment. We match firms claiming the subsidy with similar firms not claiming the subsidy and observe their hiring behaviour before and after the implementation of the policy. We find a statistically significant impact on youth employment on average. We see a positive and statistically significant effect on youth and non-youth employment in firms with fewer than 200 employees. We confirm that the increase in employment is not simply employment growth within the firms claiming the incentive.

Keywords: wage subsidy, South Africa, difference-in-differences, youth, unemployment.

JEL classification: H25, H32, J23, J38

1 Introduction

Reducing unemployment has emerged as a policy priority on the agenda of the ANC-led government. The government's New Growth Path strategy aims to create 5 million jobs by 2020 (Department of Economic Development, 2011) and is emphasised in the National Development Plan 2030. In 2011, President Jacob Zuma announced the governments' intention to spend 9 billion rand on job creation (Zuma, 2011). The same year, the National Treasury proposed the Youth Employment Subsidy as one of multiple policies to combat youth unemployment (National Treasury, 2011).

The South African government launched the Employment Tax Incentive (ETI) in January 2014, a direct intervention in the labour market aimed at stimulating demand for youth labour. The policy gives firms a tax credit for hiring individuals between the ages of 18 and 29 years.

More than three years after the ETI began, the impact on the labour market, for both subsidised and unsubsidised workers, is unclear. Levinsohn et al. (2014) conducted a randomized control trial before the launch of the ETI in which participants were given a wage subsidy voucher to present to employers when seeking employment. They found that those with a wage subsidy voucher were more likely to be employed, and that the employment effect persisted even after the subsidy had ended. Ranchhod and Finn (2014, 2015), on the other hand, found no change in the employment probabilities for youth when they analysed the ETI in the first year it was implemented.

The main contribution of the present work is our analysis of the ETI at the firm level. We do this by examining the growth in the employment of youth. In particular, we quantify the jobs created for young workers. Different from previous work on the ETI, this research uses administrative tax data giving us detailed information on the take-up of the subsidy.

We use annual job-level and firm-level tax data from 2011 to 2015 from the South African Revenue Service (SARS). We use propensity score matching to construct a group of ETI and non-ETI firms that are statistically similar in terms of pre-policy firm characteristics. We then use a difference-in-differences estimation and find a positive statistically significant effect on youth employment, non-youth employment and total employment at the firm. ETI-firms have 3.902 additional youth compared to non-ETI firms as a result of the policy. These positive effects are limited to firms we were able to match and with fewer than 1200 employees.

To evaluate whether the ETI has affected firms differently depending on their firm size, we separately match and estimate difference-in-differences for each firm size subgroup. The results are positive and significant in firms with fewer than 200 employees. Firms with more than 200 employees have positive but no statistically significant results for the employment of youth, non-youth and total employment.

The estimated impact of the ETI on firm employment is larger than estimates for similar programmes in other countries. Crichton and Maré (2013) estimates the impact of a wage subsidy programme in New Zealand and show that firms with less than 50 employees increase employment by 1.09 additional persons. The largest impact is 1.57 for firms with more than 250 employees. Kaiser and Kuhn (2016) estimate an impact of 0.458 additional workers for the highly skilled targeted group that were subsidised in Danish firms.

This paper is organized in the following manner: In Section 2, we describe the policy and in Section 3 we provide some context within the literature. We lay out the evaluation process from the defining the outcomes we assess and our identification strategy in Section 4. Section 5 describes the data and study sample of firms, Section 6 details the matching process including estimating propensity scores, assessing our matching algorithm and evaluating the matched treatment and control groups. Section 7 presents the results of conditional difference-in-differences methods. Section 8 reviews the cost of the policy and finally, in Section 9, we draw some conclusions.

2 The Employment Tax Incentive policy

Several studies conducted after the ETI was proposed in 2010 were of the opinion that it had the potential to change the employment prospects for youth (Burns et al., 2010, Levinsohn and Pugatch, 2014, Levinsohn et al., 2014, Mtembu and Govender, 2015, National Treasury, 2011).

Burns et al. (2010) argued that the wage subsidy might be successful in creating jobs in South Africa if it was associated with skills training, especially in industries that are sensitive to labour costs, and should have a focus on youth. Levinsohn and Pugatch (2014) suggested that the wage subsidy could decrease the share of youth experiencing long-term unemployment.

Levinsohn et al. (2014) conducted a randomized control trial before the policy was enacted to examine how a wage subsidy might affect youth unemployment in the South African context. Participants in the trial were given a voucher for a wage subsidy that the firm could claim monthly for up to six months. The authors found that participants who were given a wage subsidy voucher were more likely to be in waged employment both one year and two years after they were given the voucher.

Mtembu and Govender (2015) examined the perception of the wage subsidy among unemployed youth and employers in Kwazulu-Natal and found that both parties were in support of the policy, hoping that it would decrease youth unemployment and ease the wage burden.

The ETI policy, which was enacted in December 2013, was implemented on 1 January 2014 and retroactively applied to new hires from 1 October 2013. The ETI was to run for a period of three years, ending on 31 December 2016. It aimed to subsidize 423,000 youth jobs and create 178,000 new youth jobs over the policy period at a cost to the government of ZAR5 billion. We outline the details of the policy below.

2.1 Policy details

The ETI is subject to a set of criteria for the firms wishing to claim the subsidy as well as for the individuals for whom firms can claim the subsidy. The firm-level criteria are:

- Firms in the public sector are ineligible.
- Firms need to be registered for Pay-As-You-Earn (PAYE), income tax being deducted by the firm from an employee's income.
- Claims can be made only if firms do not owe the South African Revenue Service (SARS) any money.
- Claims can be made for individuals between the ages of 18 and 29 years hired after 1 October 2013 and earning less than ZAR6,000 per month but more than the minimum wage.
- No older worker must be displaced to make way for an eligible worker.

The policy does not require any training for the employed youth and is available to all industries. No requirements are placed on length of unemployment for eligible youth.

To reduce the displacement of older workers that might result from the policy, there is a ZAR30,000 penalty per employee displaced. Penalties are also imposed on firms that claim the ETI for workers who are paid less than the minimum wage.

Firms can claim the subsidy for a 24-month period for an eligible employee. However, the amount of the subsidy is greater for the first 12 months of the employment contract than the second 12 months. The amount claimed per employee is based on the employee's salary, as shown in Table 1.

Table 1. ETI monthly subsidy per employee

	Monthly subsidy	
Chapter 6. Monthly pay (ZAR)	*First 12 months*	Next 12 months
0–2,000[4]	50% of monthly pay	25% of monthly pay
2,001–4,000	ZAR1,000	ZAR500
4,001–6,000	ZAR1,000 - 0.5 x (monthly pay - 4,000)	ZAR1,000 - 0.25 x (monthly pay - 4,000)
Source: Authors' computation based on Employment Tax Incentive Act (2013).		

2.2 Policy evaluation

As the policy period is still under way, only a handful of studies have been published since its implementation. De Jongh et al. (2016) studied perceptions of the ETI among 13 local businesses in the Vaal triangle in Gauteng. The authors found that firms were in support of the policy, but that 8 of the 10 businesses claiming subsidies had not created any new jobs. They found that firms were more concerned with the skill level of young employees than they were interested in the tax incentive.

Odendaal (2016) conducted a comparative analysis of the ETI and similar policies enacted in other countries. The author suggested that the policy, as it was enacted, was unlikely to reach its goal of combatting youth unemployment, citing lack of firm awareness of the policy, the short duration of the ETI, and the absence of compulsory skills training among other reasons for his conclusion.

Ranchhod and Finn (2014) examined the policy six months after its inception. They measured the effects of the ETI using nationally representative survey data. Using an individual-level difference-in-differences (DID) approach, the authors found no change in youth employment probabilities in the first six months after the policy was implemented. Extending their analysis to the first year of the policy Ranchhod and Finn (2015), they did not find a statistically significant change in the probability of youth employment.

The limited and conflicting results from the quantitative and qualitative literature leaves fertile ground for further evaluation of the policy. Access to administrative tax data for individuals and firms provides us with an opportunity to examine the policy at the firm level. In the next section, we look to the international literature for the analysis of similar policies elsewhere.

[4] Minimum wages occur only in some sectors. The subsidy applies to part-time employment covering those earning between ZAR0 and ZAR2,000.

3 Previous literature

We examine the international literature on wage subsidies with a focus on studies that use administrative data. Previous studies on the ETI have not made use of administrative data and thus we look to the international literature for guidance on the methodology.

Betcherman et al. (2010) implement a DID method to examine the effects of two successive employment subsidy policies in Turkey. The authors use monthly administrative panel data for the period 2002–2005. The employment subsidy policies were expanded in a progressive manner across neighbouring provinces, a fact that the authors use to identify appropriate treatment and control groups for estimation. The two policies varied in their incentives, which included a subsidy on social security contributions, an income tax subsidy, an energy consumption subsidy, and a five-year land subsidy.

The authors find that the employment subsidy schemes led to significant net increases in registered jobs in provinces where the policy was implemented, despite deadweight loss (considerable in the case of the first policy). Furthermore, they find that the employment subsidy policies led to an increase in the number of registered firms; in other words, informal or unregistered firms were incentivized to register to benefit from the policy.

Crichton and Maré (2013) use a propensity score matching approach to analyse a wage subsidy policy in New Zealand. The authors use tax administrative data covering a 10-year period from 2000 to 2010. The wage subsidy was targeted at disadvantaged jobseekers, lasted for up to one year, and represented approximately 50 per cent of the weekly wage. Firms employing a subsidized worker were matched to a subset of firms that had a similar likelihood of hiring a subsidized worker but that had not yet hired one. They restricted their sample to firms continuously hiring in the three months prior to hiring a subsidized worker to ensure that firms with similar employment trends were matched. The probability of hiring a subsidized worker is modelled as a function of past employment trends, workforce composition, industry, and region. The authors run separate logistic regressions in firm size categories. Each treated firm is matched to a minimum of five control firms. The authors find that firms increase the hiring of subsidized workers and see an increase in their total employment relative to the matched comparison firm. The authors cannot, however, establish whether the growth in total employment is due to the subsidy, as firms are increasing their employment at the same time.

Rotger and Arendt (2010) use a DID matching estimator to estimate the employment effects of a wage subsidy on small private firms in Denmark. The wage subsidy amounted to approximately 50 per cent of the monthly wage and was available for up to one year.

The authors use monthly administrative data including individual, firm, and firm-level data for 10 months in 2006. The authors use a logit model to estimate the propensity score for treatment before conducting the DID estimation. They find little evidence of deadweight loss and substitution effects. Their results show that the wage subsidy has a net employment effect of 0.26 employees who would not have been hired without the subsidy.

Kaiser and Kuhn (2016) measure the effects of another Danish wage subsidy programme, which aimed to increase the employment of highly skilled workers. The subsidy lasted between 6 and 12 months and subsidized up to half of eligible employees' wages. The authors examine the performance of the firms that hired subsidized workers, using a sample of 316 firms. They match treatment and potential control groups on observed characteristics in the year before the wage subsidy programme was introduced. They estimate a dynamic fixed-effects regression and find a positive significant effect on the number of employees per firm in the year of the programme.

Bruhn (2016) examines the effect of a wage subsidy on firms in the manufacturing industry in Mexico. Subsidies were granted to firms that retained workers instead of letting them go during an economic crisis. The policy lasted eight months. Monthly administrative data for the period 2004–2013 is used in a propensity score matching and DID regression. A positive but not statistically significant effect is found. The effect ranges from a 5 per cent to a 13 per cent increase in employment. After the policy, the author finds that employment at firms in eligible industries recovered from crises more quickly than in ineligible industries.

Hujer et al. (2002) estimate the effects of wage subsidies on labour demand in West Germany. Subsidies were targeted at individuals with poor labour market prospects, including the long-term unemployed and those over 50 years old. The subsidy ranged between 30 and 80 per cent of the monthly wage, depending on the programme, and lasted around 24 months. The authors make use of annual firm survey data to calculate the effect of the subsidy using a conditional DID approach. They measure the change in labour demand by examining the change in actual employment at the firm. No significant effect of the subsidy on employment is found. They suggest that this is due to displacement or substitution effects.

Kangasharju (2007) examines the effect of a wage subsidy on employment in subsidized firms in Finland. The subsidy was available to the long-term unemployed and unemployed youth under 25 years old and was equivalent to approximately one-third of their average monthly wages. The subsidy had a maximum length of 10 months. The author uses a DID approach preceded by regression and matching methods using annual tax administrative data for the period 1995 to 2002. The author measures the change in employment by the change in payroll, and concludes that there is roughly a 9 per cent increase in employment at the subsidized firm based on the change in payroll.

In summary, wage subsidy programmes are mostly targeted at the unemployed with low labour market prospects. This often includes youth, the long-term unemployed, and in some cases those over 50. Some of the literature reports no statistically significant effects of the wage subsidy on employment. In many cases, the observed effects are modest, ranging from 0.17 to 1.09 additional jobs (Crichton and Maré, 2013, Kaiser and Kuhn, 2016, Rotger and Arendt, 2010). Kangasharju (2007) and Betcherman et al. (2010) et al. (2010), on the other hand, see a 5 and 12 per cent increase in employment in their respective studies. Many of these studies suggest that the long-term effects of the policies are either modest or short-lived. Only a few estimate deadweight loss and substitution effects, as these are often harder to measure. Betcherman et al. (2010), additionally, look at expenditure on the subsidy programmes to estimate the cost of job creation. They warn that the cost of subsidized employment is especially high in cases where there is large deadweight loss.

In terms of methodology used, all studies have access to data before and after the policy, which is well suited to a DID estimation. The studies face the same challenge of creating a convincing counterfactual to firms that take up the policy. Most match firms on a set of observables to deal with the selection issue. Betcherman et al. (2010) and Bruhn (2016) exploit implementation and eligibility criteria to create suitable control groups.

4 Evaluation process

4.1 Outcome Measure

We start out evaluation process of the ETI by choosing suitable outcome measures. This is to clarify how we think the policy works and what is defined as a policy success. The ETI was aimed at addressing the youth unemployment problem through stimulating the aggregate demand for youth labour. We consider the aggregate demand for youth labour as the sum of the demand for youth labour in each firm. We use the number of youth employed at a firm as proxy for labour demand for youth.

We want to test whether there is an increase in youth employment at the firm due to the implementation of the ETI. Wage subsidies decrease the relative cost of employing youth, thereby, in theory, increasing the demand for youth workers. However, if youth workers are substitutes for non-youth workers the decrease in the relative cost of employing youth will result in substitution (of employment) away from non-youth workers.

At the same time, a decrease in the wage bill lowers production costs, resulting in lower prices and an increase in demand for produced goods - termed the output effect. The effect on non-youth therefore depends on whether the substitution effect or the output effect is larger. We are thus also interested and want to test whether there is any change in non-youth employment as a result of the ETI.

Many wage subsidy policies see an increase in total employment, greater than the change in employment for the subsidised group (Crichton and Maré, 2013, Kaiser and Kuhn, 2016, Rotger and Arendt, 2010). This is viewed as a positive externality; the increase of unsubsidised employment can amplify the effect of the policy through enabling firms to grow. We examine total employment, log employment and payroll as additional outcomes of the policy.

Overtime, wage subsidies can increase the number of formal firms through incentivising informal firms to register with the government, as was the case in the Turkey (Betcherman et al., 2010). Existing firms could also increase the number of youth employed in established firms. These outcomes are not mutually exclusive and can occur simultaneously. We restrict our sample to established firms thus measuring the change in the number of youth employed in established firms not allowing for any firm entry or exit.

We outline, in the subsequent sections, the evaluation problem and our solution to assessing the above-mentioned outcomes.

4.2 The Microeconometric Evaluation problem

The difficulty in program evaluation is finding a credible counterfactual. The impact of the wage subsidy policy on the outcome of the ETI claiming firm requires us to think about how the ETI firm would have behaved in the absence of the policy. We can think of two outcomes, Y^T, where the firm claims the ETI and Y^C, where the firm does not claim the ETI. We borrow language from the evaluation literature and name Y^T the "treated firm" and Y^C the "control firm". Let ETI be an indicator for whether a firm claimed ETI or did not claim the ETI.

The treatment effect is the difference between the treated and control firm:

$$\Delta = Y^T - Y^C \qquad (1)$$

However, we cannot observe the ETI firm not claiming the subsidy, Y^C, if the firm has claimed the subsidy, a challenge referred to as the "Fundamental Problem of Causal Inference" (Holland, 1986). The observed outcome for each firm is:

$$Y = ETI.Y^T + (1 - ETI)Y^C \tag{2}$$

This means we cannot observe Y^T and Y^C at the same time. When $ETI = 1$, that is the firm is treated, Y is Y^T and $Y^C$is the counterfactual. When $ETI = 0$, that is the firm is not treated, Y is Y^C and $Y^T$is the counterfactual. The average treatment effect on the treated is thus:

$$E(\Delta|ETI = 1) = E(Y^T|ETI = 1) - E(Y^C|ETI = 1) \tag{3}$$

This equation asks whether there is difference in outcome between the firm who claimed the ETI compared to the hypothetical situation where the same firm did not claim the ETI. However, we cannot simultaneously observe the hypothetical firm not claiming the ETI represented by $E(Y^C|ETI = 1)$.

In cases where a randomised experiment is conducted, the control group is constructed such that it is a valid counterfactual because the following condition holds true:

$$E(Y^C|ETI = 1) = E(Y^C|ETI = 0) \tag{4}$$

That is, the control firm is a comparable to the hypothetical firm.

Firms claiming the ETI were not randomly assigned thus the condition is not true:

$$E(Y^C|ETI = 1) \neq E(Y^C|ETI = 0) \tag{5}$$

Using the mean outcome of the non-claiming firms to approximate the outcomes for the hypothetical firm that did not claim the ETI is not recommended as ETI and non-ETI firms may differ without the policy. Using the non-ETI firms as a control group will thus lead to selection bias on observable and unobservable characteristics (Hujer et al., 2002). In the next subsection we deal with the selection bias.

4.3 Selection on Observables and Unobservables

One way to deal with selection on observables is through matching. The idea behind matching is to search for a similar and comparable non-ETI firm from a large group of non-ETI firms. Similarity is based on firm pre-treatment covariates and matching can be conducted by covariates (Rubin, 1978) or by propensity scores (Rosenbaum and Rubin, 1983).

The difference in outcomes computed between the treated and matched control groups can then be attributed to the policy (Rosenbaum and Rubin, 1983).

Matches are made based on the identifying assumption that, conditional on all relevant pre-treatment covariates (X), the potential outcomes (Y^T, Y^C) are independent of participation (conditional independence assumption or CIA).

$$Y^T, Y^C \coprod ETI|X \qquad (6)$$

If the CIA holds, that is the treatment and control groups are balanced, then, the control group is a valid counterfactual for the treatment group. This means we can use the observed outcome of the non-ETI firm to estimate the counterfactual outcome of the hypothetical ETI firm if it did not claim the ETI. This is described in the equation below:

$$E(Y^C|X, ETI = 1) = E(Y^C|X, ETI = 0) = E(Y^C|X) \qquad (7)$$

The role of matching is to ensure balance between the treatment and control groups based on pre-treatment covariates. This confirms independence between the potential outcome and treatment and gives an unbiased estimator.

The most intuitive way to match firms is to match directly using firm characteristics. For example, small ETI firms are matched with other small non-ETI firms. However, the number of covariates that determine selection is large and it becomes impossible to match directly; also known as the curse of dimensionality. There are two ways to proceed, through mapping covariates into a metric measuring the closeness of two observations (Rubin, 1978) or matching by propensity score (Rosenbaum and Rubin, 1983). Both methods make the distribution of covariates in the treatment group and the distribution of covariates in the control group the same. Rosenbaum and Rubin (1983) recommend the use of a balancing score, that is a function of the covariates. The propensity score is an example of a balancing score summarising information of covariates (X) into a single index function that gives us the probability of claiming the ETI based on the observed characteristics of the firm.

$$Y^T, Y^C \coprod ETI|\, P(X) \qquad (8)$$

The CIA then extends to propensity scores, that is, conditional on propensity scores $P(X)$, the potential outcomes Y^T and Y^C are independent of treatment as described in equation (8).

If we go back to equation (7) and amend the conditioning from X to $P(X)$ we get the following equation:

$$E(Y^C|P(X), ETI = 1) = E(Y^C|P(X), ETI = 0) = E\big(Y^C\big|P(X)\big) \tag{9}$$

The curse of dimensionality is avoided with the use of $P(X)$. This means our counterfactual can be represented as follows:

$$E(Y^C|ETI = 1) = E_{P(X)}[E(Y^C|P(X), ETI = 0)|ETI = 1] \tag{10}$$

We update equation (3), the average treatment effect on the treated (ATT), with our new counterfactual:

$$E(\Delta|ETI = 1) = E(Y^T|ETI = 1) - E_{P(X)}[E(Y^C|P(X), ETI = 0)|ETI = 1] \tag{11}$$

This means that we can match ETI firms with non-ETI firms where they have the same propensity score. The propensity score can be estimated using a standard probability model such as a logit or probit model and the quality of matching depends on how well the propensity score is estimated and the ability to find comparison firms. An examination of the firms not matched will give us confidence whether any incomplete matching will bias our results.

4.4 Conditional Difference-in-Differences Approach

Thus far, we have argued that matching allows us to deal with selection on observables, what remains is for us to deal with the selection on unobservables. Changes in the economic environment, such as an economic recession or expansion, at the same time as the introduction of the ETI, will affect our assessment of the policy. To account for this, we use a difference-in-differences (DID) approach. The DID measures the difference between the treatment and control group before the policy and subtracts from it the difference between the treatment and control after policy implementation. This is represented in equation (12).

$$DiD = E(Y_1^T - Y_0^T|ETI_1 = 1) - E(Y_1^C - Y_0^C|ETI_1 = 0) \tag{12}$$

The DID approach requires a parallel time trend assumption and allows for time invariant selection bias.

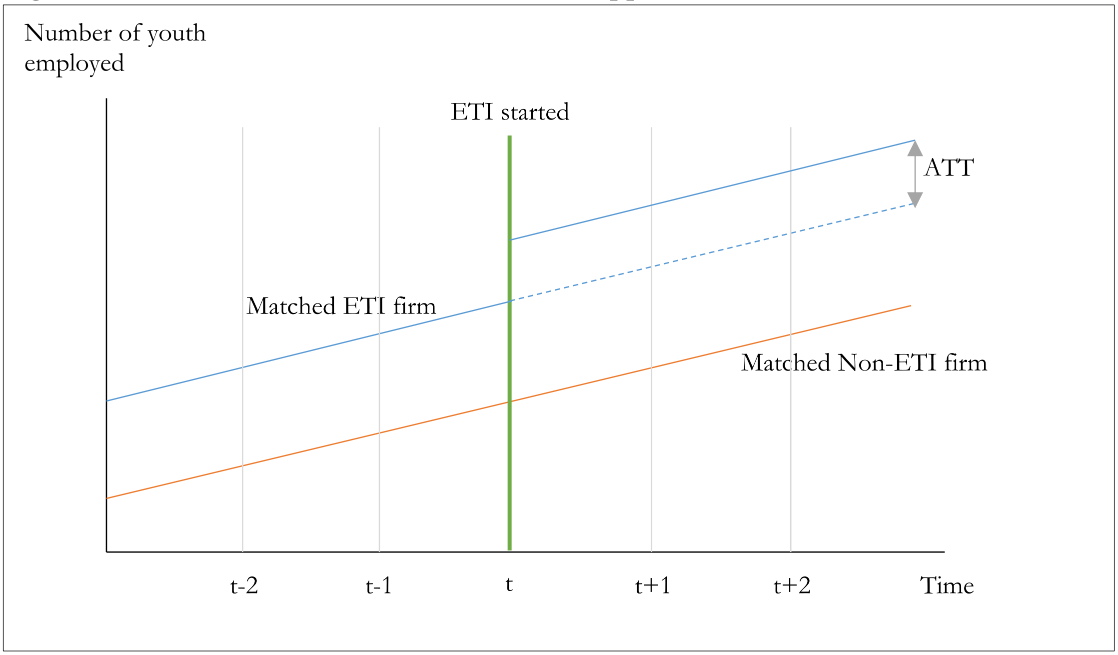
Heckman et al. (1998) introduced the conditional DID (cDID) estimator extending the usual DID through conditioning the outcomes on the propensity score. The identifying assumption is thus:

$$E(Y_1^T - Y_0^T | P(X), ETI_1 = 1) = E(Y_1^C - Y_0^C | P(X), ETI_1 = 0) \quad (13)$$

Equation (13) demonstrates that conditional on propensity score marching, we can estimate the average treatment effect on the treated by examining the differences between the matched treated-control firms before and after the policy.

This cDID approach ensures that our estimation is the true treatment effect having accounted for selection on observables and unobservables. This idea is illustrated in the Figure 1 below.

Figure 1. Conditional Difference-in-differences approach



ETI and non-ETI firms are matched at time $t-1$. We then compare their differences in employment of youth at $t-1$ and subtract from the difference between the two groups at $t+1$. This gives us the average treatement on the treated as indicated in Figure 1.

The estimation equation for the ATT is thus:

$$\Delta_{t+1} = \left[Y_{t+1}^{T'} - Y_{t-1}^{T'}\right] - \left[Y_{t+1}^{C'} - Y_{t-1}^{C'}\right] \quad (14)$$

where:

Y	Refers to the outcome measured
T'	Refers to the matched ETI firm
C'	Refers to the matched non-ETI firm
$t-1$	Refers to the pre-policy period
$t+1$	Refers to the period after policy implementation at time t.

The ATT estimates the change in employment beyond the number that would have been employed in the absence of the subsidy. This means that the interpretation of the ATT estimate is the number of jobs created due to the wage subsidy policy.

As a scenario, Firm A and Firm B are both expanding their firms and hiring more workers. Both firms are in industries that have a need for semi or low-skilled youth. In the absence of the subsidy these firms will continue to hire and expand. The subsidy is introduced in 2014 but only Firm A claims the subsidy. Firm B does not claim the subsidy as they are unaware of the policy. Firms A and B are matched on their pre-policy characteristics, including industrial classification, firm size, and employment growth rate. During the policy years, it is assumed that a recession takes place that affects both firms' hiring patterns. Firm A sees an increase in youth employment beyond the recession compared with the period before. A before–after comparison will be limited, as it does not account for any change in the economic environment. A DID approach, on the other hand, takes the economic environment into account and reflects the true effect of claiming the ETI. This means that a cDID approach will distinguish any increase in youth employment at Firm A and allow us to attribute it to the policy.

This is similar to the approaches of Bruhn (2016), Crichton and Maré (2013), Hujer et al. (2002), Kaiser and Kuhn (2016), Kangasharju (2007) and Rotger and Arendt (2010) for the evaluation of wage subsidy programmes, using administrative data, in other countries.

5 Data

The South African Revenue Service (SARS) made anonymized tax data available to measure the effects of the ETI through a joint SARS/National Treasury/UNU-WIDER initiative. These data include, but are not limited to, Company Income Tax (CIT), Employee Tax Certificate (or IRP5), and EMP501 data. The data are reported by tax year. The tax year for individuals in South Africa runs from 1 March to 28/29 February. This means that for the 2014 tax year we see the ETI being claimed only for two months, whereas in the 2015 tax year, we see the ETI being claimed for the full year, starting 1 March 2014 and ending 28 February 2015. At the time of writing, IRP5 data is available only until 2015 and CIT data only until 2014. Companies have 12 months from their financial year end in which to complete their tax returns; thus, some company data may be incomplete

5.1 Description of IRP5 data

The IRP5 data is job-level administrative tax data. For each job with an annual income greater than ZAR2,000 there is an entry if the firm is registered for PAYE. Employees who earn less than ZAR2,000 in a tax year are not issued with an IRP5 form.

If an individual has multiple jobs in a year (within the same firm or in different firms), that individual will be seen multiple times in the data. Government or public entities also issue IRP5 forms to their employees; consequently, the data includes employees in both the formal private sector and the public sector. An individual working in a firm that is not registered for PAYE will not be included in the IRP5 data. We use IRP5 data from 2012 to 2015.

The individual identifier is an anonymized South African national identity (ID) number. Where no South African ID number existed but an anonymized passport number was present, we assume that the individual is a non-resident or non-citizen of South Africa.

The data also includes some basic individual-level information—dates of the start and end of employment, an annual amount for the ETI claimed—and firm-level information, such as income source and amount earned.

The PAYE reference number (or payroll reference) serves as the firm identifier in the IRP5 data. Larger firms may have several payrolls and therefore several PAYE reference numbers but only one company income tax reference number. Within the CIT data, public firms have no company income tax reference number.

For the analysis, we organize the data by the tax reference number within the tax year. This allows us to create a panel of firms. The analysis is thus restricted to data on formal private sector and public sector firms.

5.2 Description of the CIT-IRP5 panel

The Company Income Tax-IRP5 (CIT-IRP5) panel was created through a joint initiative between SARS, the National Treasury, and UNU-WIDER. The dataset matches employer–employee variables from the IRP5 and CIT datasets as well as value added tax data from firms and customs data from firms that trade. The panel includes tax information from 2008 to 2014 and makes use of the company income tax reference number as the unique identifier for the firm. Pieterse et al. (2016) provide a more detailed description of the CIT-IRP5 panel. They discuss how the panel was constructed and any biases it might contain and compare the panel with other data sources.

To complement our analysis, we merge firm-level variables from the CIT-IRP5 panel with the IRP5 panel we create. Firm-level variables from the CIT-IRP5 panel include firm sales, age of firm, firm assets, firm debt, and firm labour broker status. The CIT and IRP5 data are extractions of data from the revenue collection by SARS.

5.3 Study sample

Firms in the final empirical equation are a subset of firms in the data. First, we need to select the group of control firms. We remove firms in the public sector as they are ineligible for the subsidy. There are 323,005 firms in the unbalanced panel from 2011 to 2015, of which 905 public sector firms are dropped.

Second, we supplement our panel data with firm characteristics from the Company Income Tax panel (CIT-IRP5). This gives us more information on firms that enables us to find more suitable matches between firms. The downside of this, however, is that not all firms in the IRP5 have corresponding CIT data. This can be a result of the lag in submitting a firm tax return to SARS. Approximately 82% of firms in the panel have CIT information and there is currently no company tax information on firms in 2015 and 2016. We use the firm characteristics to match in 2013 and drop 39,814 firms with no CIT firm characteristics in 2013.

Third, we require firms with 'pre' and 'post' intervention data to conduct the DID. For example, a firm that registered for the first time in the 2014 tax year will not be included in our analysis as we do not have sufficient pre-intervention information with which we can match to a firm not claiming the ETI. We further restrict our sample to firms that were operating in 2011 and 2012 to include lagged employment information in our estimation. Thus, no firm that registered after 2011 is included in our model giving us a balanced panel of 130,808 firms representing about 41% of the total firms in the unbalanced panel.

Table 2. Firm size of all firms versus balanced panel, 2015

	All firms	**Balanced panel**
Micro: 0 – 5	115,658	47,537
Small: 6 - 10	45,634	26,093
Medium: 11 – 50	62,329	42,264
Large: 51 - 200	14,537	11,231
Very Large: 200+	4,326	3,683
Total	242,484	130,808

This leaves us with a balanced panel of 130,235 firms. From the 31,524 firms that claim the ETI in 2015, we used 20,055 firms claiming the ETI in 2015 and the remainder of firms are used to form our control group of firms. We later examine a control group of firms that employed at least 1 eligible youth. In this set up treatment is defined in terms of claiming the subsidy in 2015, it is then possible that a firm in the control group in 2015 goes on to claim the subsidy in 2016. In section 7.6 we restrict the control group to firms who do not go on to claim in 2016.

Lastly, the take-up rate for firms with more than 1,200 employees is approximately 72% in 2015. We are unable to find credible matches for these firms as there are too few non-ETI firms to match with. There are 573 firms with more than 1,200 employees in 2013 of which 400 claim the ETI in 2015. This represents only about 0.4% of firms in our balanced panel and 2% of ETI firms in 2015. This may seem like very few firms we are dropping but the value of ETI claimed and number of youth they employ is large. Table 3 compares the very large firms to firms in our sample panel. In terms of the value of ETI claims, our sample of ETI firms claim 2.5 billion rand in incentives and 400 very large firms claim 52.6% of this. Similarly, of the 1.4 million youth employed in ETI firms, more than half of these youth are employed in these very large firms.

Table 3. Dropped firms, very large ETI-firms vs sample panel of ETI-firms

	>1200 employee ETI firms	**Sample ETI firms**	**Percentage**
Value of claims in 2015 (billions of Rands)	1.3	2.4	54%
Number of youth in 2015	974,444	1,400,036	70%
Number of firms	400	20,455	2%
Source: Authors' own estimates based on IRP5 data			

This leaves us with a balanced panel of 130,235 firms. From the 31,524 firms that claim the ETI in 2015, we used 20,055 firms claiming the ETI in 2015 and the remainder The way in which we construct our study sample affects the external validity of our results. Our estimation will therefore be limited to formal sector, private firms; operating from 2011 onwards; with company tax information in 2013 and with less than 1200 employees representing 46% of the value of ETI claimed at firms.

<u>6 Estimating the propensity score, matching and achieving balance</u>

<u>6.1 Propensity scores</u>

We estimate the probability of take-up of the ETI based firm characteristics in 2013. Caliendo and Kopeinig (2008) advise the use of a logit or probit model for the binary treatment case as the two models have similar results and decision is thus not critical. The logit model is:

$$y_i^* = \beta' x_i + u_i \quad (15)$$

where y_i is equal to 1 if firm i claims the ETI in 2015 and 0 if firm i does not claim the ETI. The drivers of participation in the ETI are given by the vector x_i.

Caliendo and Kopeinig (2008) provides some guidance regarding the variables to include in the propensity score model. Variables that affect the choice to claim the ETI and influence the number of employees should be included in the model. Variables should be unaffected by the ETI and thus we estimate the propensity score model based on the 2013 (pre-ETI) firm characteristics. Where the probability of take-up cannot be explained by any of the covariates then matching cannot be used to as these firms never claim the ETI. This is the case for firm in the public sector who are ineligible for the ETI and the common support condition required for matching fails. We remove all public-sector firms before we establish the propensity scores.

In the previous chapter we provide a rigorous discussion of the factors affecting the take-up of the ETI. We use these variables in our logit model and present them in Table 4 below. We include firm sales, assets and average wages as they are standard determinants of labour demand (Kangasharju, 2007). We also control for other youth employment programmes such as the Learnership Tax Incentive, a programme giving firms a tax incentive for employing and training workers. To account for the regional differences in economic environment we include firm location. To account for the differences in take-up rates among industries we include a firm sector variable.

Using the same tax data, Bhorat et al. (2016) find that firm age, trade status and firm size affect wages at the firm level. Also using the same tax data, Matthee et al. (2017) find that firm exporter status affects the labour demand at firms. Based on this evidence we include these firm-level characteristics when estimating the propensity score.

Lastly, to ensure that we are not assessing firms that are already growing and expanding we include the employment growth rates in previous years.

The following variables were found not to be statistically significant in predicting take-up of the ETI: Lagged employment growth rate, lagged youth employment growth rate firm import status relative to no trade status, firm location in the Eastern Cape relative to the Western Cape and Wholesale and Retail and the Education sector relative the Agricultural sector.

Table 4. Logit model estimation for propensity scores

	Estimate	Std. Errors
Firm assets	-0.000****	(0.00000000001)
Firm sales turnover	-0.0000000001***	(0.00000000002)
Firm age	0.00526***	(0.000978)
Firm learnership claim (dummy)	0.821***	(0.0641)
Firm size (continuous)	0.00926***	(0.000179)
Mean employee age (in years)	-0.0644***	(0.00204)
Mean employee wage (in rands)	-1.18e-05***	(1.26e-06)
Employment growth rate (%)	0.158***	(0.0376)
Lagged employment growth rate (%)	0.0506	(0.0309)
Youth employment growth rate (%)	0.0540**	(0.0247)
Lagged youth employment growth rate (%)	0.00740	(0.0233)
Firm trade status (reference No trade)		
- Import only	0.0486	(0.0422)
- Export only	0.194***	(0.0474)
- Import and export	0.269***	(0.0313)
Firm province (reference Western Cape)		
- Eastern Cape	0.0748*	(0.0420)
- Northern Cape	0.177**	(0.0695)
- KwaZulu-Natal	-0.148***	(0.0544)
- Gauteng	-0.0402	(0.0316)
- North West	-0.217***	(0.0572)
- Mpumalanga	-0.251***	(0.0257)
- Free State	-0.214***	(0.0499)
- Limpopo	-0.207***	(0.0625)
Firm industry (reference Agriculture)		
- Mining	-0.445***	(0.107)
- Manufacturing	-0.147***	(0.0388)
- Electricity supply	-0.234*	(0.138)
- Water management	-0.476***	(0.125)
- Construction	-0.290***	(0.0507)
- Wholesale & retail	0.0448	(0.0403)
- Transportation & storage	-0.123*	(0.0630)
- Accommodation & service activities	0.354***	(0.0532)
- Information & communication	-1.383***	(0.241)
- Financial & insurance services	-0.363***	(0.0461)
- Real Estate Activities	-0.621***	(0.107)
- Professional, scientific, & technical	-0.267***	(0.0490)
- Administrative & support service activities	-0.174*	(0.0923)
- Education	0.0389	(0.0866)
- Health & social work activities	-0.508***	(0.0844)
- Arts & entertainment	-0.327***	(0.115)
- Other service activities	-0.365***	(0.0946)
Intercept	0.944***	(0.0802)

Observations	72,823	
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Source: Authors' own estimates based on IRP5 and CIT-IRP5 data		

From the 20,055 ETI firms in the study sample we are able to estimate the propensity score for 16,025 ETI firms.

6.2 Matching

We argued that matching allows us to compare ETI firms with non-ETI firms if done correctly. The mechanics of this are as follows: each ETI firm is matched to one or more comparable non-ETI firm based on their propensity score to minimise the differences between the firms. There are various ways in which to use propensity scores to conduct matching described in detail by Caliendo and Kopeinig (2008). Matching methods trade bias against precision. Bias relates to the differences between the treated and control groups while precision relates to the size of the control group. Matching itself is not identifying the treatment effect in our model and we have a large control group of firms.

Nearest neighbour matching uses the 'most similar' non-ETI firm to be compared with an ETI firm. The downside of this matching method arises if the nearest neighbour is far away. The distance of the nearest neighbour can be imposed by using a caliper, but this may reduce the number of matches made. Another way of improving the matching method may be to conduct matching within groups. An example of this can be seen in the study by Hujer et al. (2002) where matches are made within firm size and industry groups. This method requires enough treatment and control units in each of these sub-categories within which to match.

In our case, where some ETI firms have many close neighbours and other ETI firms have few neighbours that may be far away, kernel matching is favoured over nearest neighbour matching. Nearest neighbour matching could result in poor matches when compared to kernel matching (Bryson et al., 2002).

Kernel matching is a non-parametric method using all firms in the control group to build a weighted composite according to the distance between propensity scores. Observations closer in absolute propensity score distance, $|P(X_i) - P(X_j)|$, are assigned a greater weight. Kernel matching produces smaller standard errors as more information is used to the build the weighted counterfactual.

Caliendo and Kopeinig (2008) assert that the choice of kernel function is not as important as the choice of bandwidth. The choice of bandwidth is a trade-off between a variance and bias (Silverman, 1986, Pagan and Ullah, 1999). A higher bandwidth gives us a lower variance but could result in a biased estimate. The overall take-up rate for the ETI is around 13% so there is a sufficient number of firms that can be used beyond a limited number of nearest neighbours. We compare the more popular Epanechnikov and Gaussian kernel functions and find no difference in our results and choose to use the Gaussian kernel function (See appendix Table A for a comparison of the results). The Gaussian kernel function can be written as:

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

One way of choosing a bandwidth is proposed by Silverman (1986). The optimal h or bandwidth that minimises the mean square error of the kernel estimate is given as

$$h = \frac{0.9A}{n^{\frac{1}{5}}} \ where\ A = \min\left(\sqrt{Var(X)}, \frac{IQR(X)}{1.349}\right) \tag{16}$$

IQR is the interquartile range of the sample. We calculate h accordingly and use it in our kernel matching process. To check the sensitivity of our results to the choice of bandwidth we estimate the cDID for our outcomes of varying h (See appendix Figure A).

The results are sensitive to bandwidth below 0.4 the estimator but this remains constant beyond this point. Our choice is thus to use the Gaussian kernel function with a bandwidth of 0.45 as per our calculation.

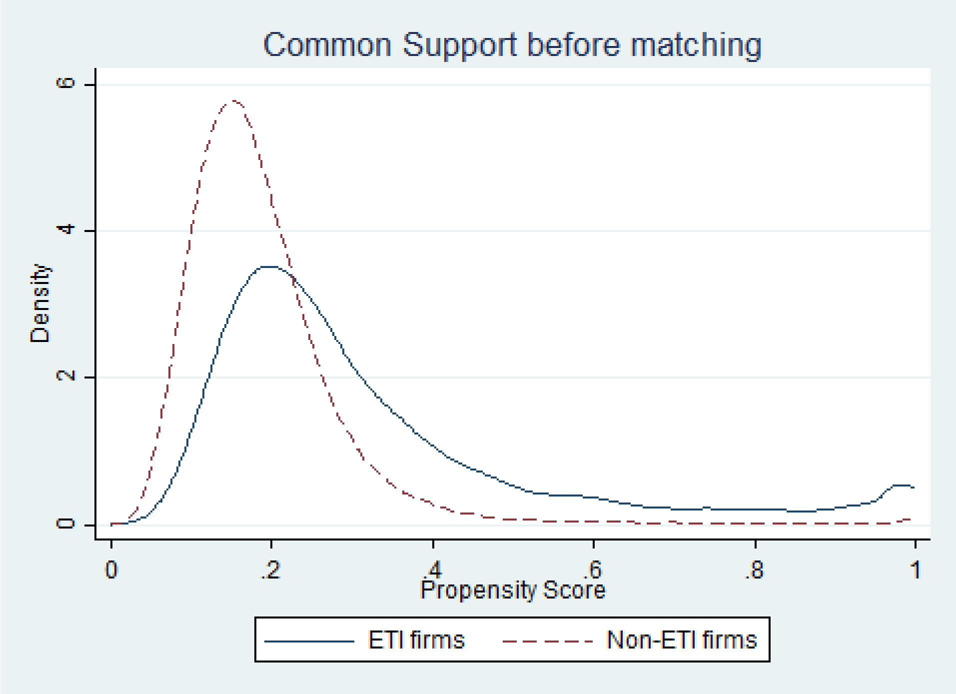
Our estimations are conducted using Stata 14.0 along with the "diff" and "psmatch2" user-written module by Villa (2016) and Leuven and Sianesi (2003) . The "psmatch2" command is used to the conduct the propensity score matching while the "diff" command is used to calculate the difference-in-differences from the matched groups.

6.3 Common Support

We check whether the matching process has reduced the difference between the two groups and renders them comparable. This will give us confidence in estimating the ATT.

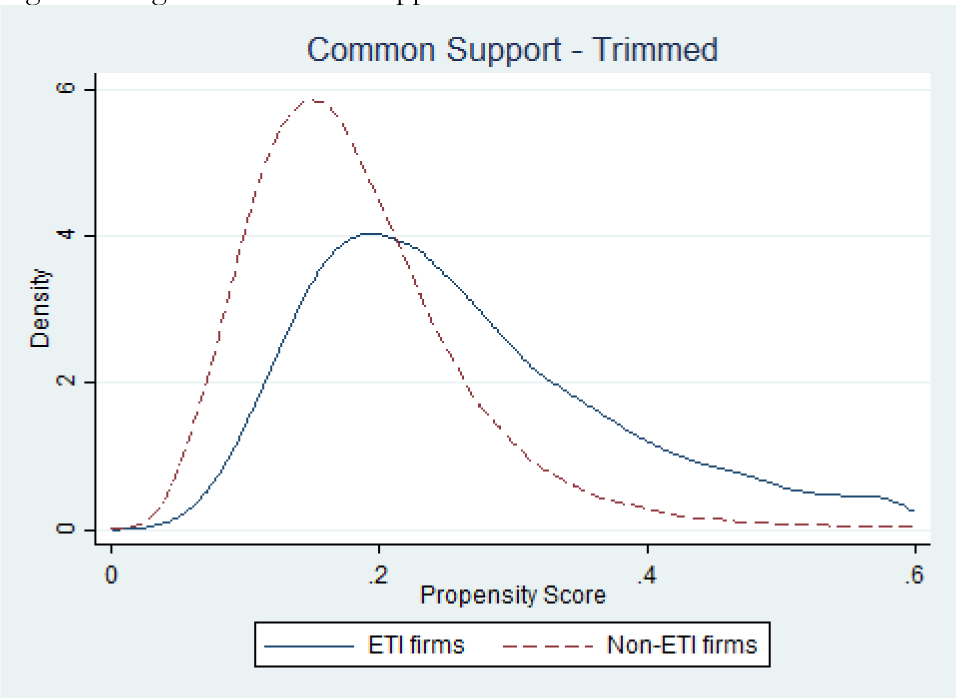
We start with the region of common support as the ATT is only defined in this region. Figure 2 below is a visual analysis containing the density distribution of propensity scores for the treatment and control groups of firms.

Figure 2. Region of common support, before matching



The densities are very low above the propensity score of 0.6 for both ETI and non-ETI firms.[5] Smith and Todd (2005) suggest trimming the common support region where the density is low. We drop observations with propensity score of 0.6 thus reducing our panel further. Below is the common support graph after trimming.

Figure 3. Region of common support trimmed



[5] Results for the cDID estimation before trimming can be found in the Appendix, Table C.

The trimming process drop 1892 ETI firms representing about 8% of ETI firms in our study sample. This leaves us with 14,133 ETI firms with which estimate the cDID.

Figure 4. Region of common support – Before and after matching

[Insert figure]

6.4 Firms not matched and dropped

Our matching process is incomplete if we do not find matches for all the treated firms in 2015. However, firms with a propensity score equal to zero cannot be matched. This has the potential to create a bias if the unmatched firms are systematically different from the matched firms. Table 5 shows the firm size distribution, the firm sector association, the value of the ETI claims and the number of youth employed in the firms that were not matched firms. In terms of the firm size, no large differences appear between the matched and unmatched groups. In terms of the industrial classification, firms in Manufacturing, Construction and Wholesale and Retail have a higher percentage of firms matched than unmatched. Firms in the education sector have a lower percentage of matched compared to unmatched firms. This appears to be linked to the take-up high take-up rates in Manufacturing, Construction and Wholesale and Retail and the low take-up rate in the Education sector.

In terms of the relationship to the ETI, the firms not matched account for 18% of the youth employed in ETI firms and 19% of the value of claims from firms in our study sample.

Table 5. Firm size distribution, unmatched versus matched ETI firms

Firm size	**Unmatched**	**%**	**Trim**	**%**	**Matched**	**%**	**Study sample of ETI firms**	**%**
0-5	1,039	8.8	-	-	782	42.9	1,821	8.3
6-10	989	13.0	-	-	1,644	62.4	2,633	12.0
11-50	2,329	44.4	1	0.0	7,328	75.9	9,658	44.0
51-200	1,141	24.1	324	5.6	4,305	74.6	5,770	26.3
201+	364	9.8	1,574	75.4	149	7.1	2,087	9.5
Total	5,862	100	1,899	8.6	14,208	64.7	21,969	100
Industry								
Agriculture	749	13.0	279	12.6	1,179	8.3	2,207	10.1
Mining	43	0.7	49	20.9	142	1.0	234	1.1
Manufacturing	1,037	17.9	503	9.0	4,022	28.3	5,562	25.4
Electricity	9	0.2	6	7.0	71	0.5	86	0.4
Water supply	24	0.4	7	6.1	83	0.6	114	0.5
Construction	232	4.0	60	5.1	893	6.3	1,185	5.4
Wholesale & retail	735	12.7	268	6.8	2,939	20.7	3,942	18.0
Transport	109	1.9	100	15.2	451	3.2	660	3.0
Accommodation & Food service	300	5.2	137	10.3	891	6.3	1,328	6.1
Information & Communication	12	0.2	5	11.1	28	0.2	45	0.2
Financial & Insurance services	608	10.5	241	10.6	1,422	10.0	2,271	10.4
Real Estate Activities	75	1.3	5	2.6	114	0.8	194	0.9
Professional, Scientific & Technical	630	10.9	115	6.2	1,123	7.9	1,868	8.5
Administrative & Support service	74	1.3	48	15.5	187	1.3	309	1.4
Education	360	6.2	21	3.6	203	1.4	584	2.7
Healthcare	267	4.6	26	5.3	199	1.4	492	2.3
Art & entertainment	94	1.6	17	8.0	102	0.7	213	1.0
Other service activities	424	7.3	12	2.0	159	1.1	595	2.7
Total	5,782	100	1,899	8.7	14,208	100	21,889	100
Value of claims in 2015 (millions of Rands)	231	19.5	475	39.2	481	40.4	1,190	100
Number of youth in 2015 ('000)	124	18.2	267	39.2	267	39.2	682	100
Source: Authors' own estimates based on IRP5 data								

No group of firms is over or under represented in the matched group of firms and therefore no systematic bias is suspected. This gives us confidence in proceeding with the cDID estimation. Overall, there are 20,055 ETI firms in our study sample, of which we find matches for 14,133 ETI firms representing 70% of study sample of ETI firms.

6.5 Achieving balance

We examine the covariate means before and after matching. Before matching we expect the means are very different and after matching we want to ensure that these differences have been reduced. Table 6 displays the before-matching means (columns 2 and 3) for ETI and non-ETI firms and the after-matching means (columns 4 and 5) for ETI and non-ETI firms. Columns 2 and 3 indicate that firms in the treatment group have greater sales turnover and assets, are slightly older, have a lower mean employee age and wage, and have higher employment rates. Column 4 and 5 indicates that many of these differences remain with the exception of the employment growth rates. The differences in employment growth rates appear to have been reduced by the matching process.

Table 6. Comparison between ETI firms and matched control firms

	Before Matching		**After Matching**	
	Non-ETI	**ETI**	**Non-ETI**	**ETI**
Number of firms	133,459	21,969	56,009	14,208
Firm sales turnover (in millions, R)	25.1	77.4	29.4	53.0
Firm assets (in millions, R)	32.6	64.2	33.6	46.7
Firm age (in years)	13.9	14.9	14.0	14.4
Firm learnership claim	0.00	0.04	0.01	0.03
Number of employees	15.9	79.6	19.5	45.5
Mean employee age (in years)	40.8	36.0	37.4	35.6
Mean employee wage (Rands)	11,135	9,295	11,139	9,657
Employment growth rate	0.053	0.106	0.051	0.090
Lagged employment growth rate (%)	0.065	0.116	0.080	0.106
Youth employment growth rate	-0.007	0.064	-0.032	0.044
Lagged youth employment growth rate (%)	0.041	0.094	-0.032	0.100
Firm trade status (reference No trade)				
- Import only	0.044	0.051	0.056	0.057
- Export only	0.030	0.041	0.041	0.045
- Import and export	0.069	0.134	0.107	0.141
Firm province (reference Western Cape)				
- Eastern Cape	0.06	0.07	0.057	-0.016
- Northern Cape	0.02	0.02	0.056	-0.009
- KwaZulu-Natal	0.03	0.04	0.003	0.007
- Gauteng	0.13	0.16	0.073	0.009

- North West	0.03	0.03	-0.002	0.003
- Mpumalanga	0.45	0.37	-0.148	0.022
- Free State	0.04	0.04	0.012	-0.006
- Limpopo	0.03	0.03	0.004	-0.008
Firm industry (reference Agriculture)				
- Mining	0.008	0.011	0.010	0.010
- Manufacturing	0.221	0.254	0.277	0.283
- Electricity supply	0.005	0.004	0.006	0.005
- Water management	0.008	0.005	0.009	0.006
- Construction	0.062	0.054	0.075	0.063
- Wholesale & retail	0.143	0.180	0.172	0.207
- Transportation & storage	0.030	0.030	0.034	0.032
- Accommodation & service activities	0.030	0.061	0.035	0.063
- Information & communication	0.003	0.002	0.004	0.002
- Financial & insurance services	0.141	0.104	0.130	0.100
- Real Estate Activities	0.028	0.009	0.018	0.008
- Professional, scientific, & technical	0.107	0.085	0.094	0.079
- Administrative & support service activities	0.009	0.014	0.011	0.013
- Education	0.023	0.027	0.014	0.014
- Health & social work activities	0.070	0.022	0.024	0.014
- Arts & entertainment	0.009	0.010	0.008	0.007
- Other service activities	0.039	0.027	0.014	0.011
Source: Authors' own estimates based on IRP5 and CIT-IRP5 panel data.				

Kernel matching uses all the firms in the control group to be matched with each treated firm. This means that differences will remain between the treatment and control groups but firms with a closer propensity score will be more heavily weighted that a firm with a very different propensity score. With the matched treatment and control groups of firms we proceed with the cDID estimation.

7 Results

7.1 Conditional difference-in-differences estimation

Our matching process produces 14,133 matched ETI firms which we used in our estimation. Table 7 displays the results from the conditional DID estimation. The coefficients represent the effect of claiming the ETI for youth, non-youth and total employment compared with matched firms that did not claim.

Table 7. Results of the cDID estimators from matched firms, 2015

Number of employees	**Number of firms**	**Coefficient**	**Standard Error**	**R^2**
Youth	14,133	3.902***	0.204	0.09
Non-youth	14,133	4.780***	0.380	0.06
All	14,133	8.704***	0.594	0.07
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Source: Authors' own estimates based on IRP5 data				

For firms that claimed in the 2015, there is a positive statistically significant difference in youth employment at ETI-firms. This means that ETI-firms have 3.902 additional youth employed in 2015 when compared with non-ETI firms. At the same time, we see positive and significant increases for non-youth employees and total employees at the firm.

The estimated effect on the number non-youth employment are larger than for youth employment. This implies that firms hiring additional youth workers would on average hire 4.780 additional non-youth workers. There are two possible ways in with non-youth can be affected by the policy. One, non-youth employment can decrease if youth are substitutes for non-youth workers and two, we could also see an increase in the demand for non-youth labour should the firm use the savings from the tax incentive to employer additional older workers who they know are more productive and have more experience. The larger, positive, significant result implies that the ETI enables firms to grow both youth and non-youth employment.

The positive significant results for the total employment that is greater than the result of the target group is similar to results found in many other studies [citations needed]

7.2 cDID estimation within firm size subgroups in 2015

The assumption in the model above is that of a "constant effect" at the firm level, that the ETI affects our sample of firms in the same way (Holland, 1986). We check this assumption by estimating the average treatment effect on firm size subgroups as we have already seen that the subsidy take-up rates differ for these subgroups. If the estimated effects vary then the constant effect of the policy is violated.

Table 8 contains the results of the cDID estimation within firm size subgroups.

Table 8. Results of the cDID estimators, matching within firm size subgroups, 2015

		Number of employees		
Firm size category	**Number ETI firms**	**Youth**	**Non-youth**	**Total**
Micro firms 0–5 employees	817	1.766*** (0.067)	1.363*** (0.124)	2.996*** (0.193)
Small firms 6-10 employees	1643	2.226*** (0.183)	1.920*** (0.277)	4.154*** (0.446)
Medium firms 11–50 employees	7327	2.511*** (0.125)	2.624*** (0.183)	5.145*** (0.268)
Large firms 51–200 employees	4619	7.428*** (1.034)	9.746*** (1.541)	17.230*** (2.345)
Very Large firms 201+ employees	2038	19.71 (56.7)	24.570 (134.2)	44.59 (182.4)
Reduced sample firms 201 – 1200 employees	1707	14.07** (6.787)	23.64** (10.489)	38.07** (15.122)

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Source: Authors' own estimates based on IRP5 data.

Column 2 shows the positive significant effects on youth employment at firms with fewer than 200 employees. Very large firms, however, do not see a significant increase in the hiring of youth employees, although the coefficient is positive. Since we are concerned with the quality of matches for firms with more than 1200 employees we examine a reduced group of very large firms. If we consider firms the employ between 201 and 1200 people we find a positive and significant result for youth but not for non-youth employment. This implies that the insignificant result we find in the group of very large firms may be driven by the no change in youth employment at firms with more than 1200 employees. Firms with more than 1200 employees may be substituting the employment of unsubsidised workers with subsidised youth workers creating a deadweight loss but we cannot be certain of this as we are unable to find any credible counterfactual firms. A deadweight loss occurs if subsidised workers would have been hired even without the subsidy. In large firms, wage subsidies might be redundant if the subsidy is not being used to create new employment simply absorbed by the firm.

The estimates within firm size subgroups allow us to compare with results from other studies. Crichton and Maré (2013) find and increase of 1.1 subsidised workers in firms with less than 50 employees in New Zealand. Rotger and Arendt (2010) find an increase of 0.26 additional workers in small subsidised firms. Our estimates are, however, larger.

The third column in Table 8 show the additional employment expansion of non-youth workers. In the case of Micro firms and larger firms the coefficients for non-youth workers is larger than for youth employees. Crichton and Maré (2013) see a similar trend where the increase in unsubsidised workers is greater than the increase in subsidised worker in firms with more than 50 employees.

The literature often does not test the outcomes for the non-subsidised group, measuring the impact on the subsidised group and total employment at the firm. However, when tested, it is often the case, where an increase in employment of the subsidised group is found, there is an increase in total employment greater than the increase of the subsidised group (Kaiser and Kuhn, 2016, Hujer et al., 2002).

A visual representation of the results from Table 8 gives a sense of the magnitude of the estimated impact on youth employment. The figures below give some more nuance to the assertion that the ETI affects small and big firms in different ways. Based on pre-policy youth employment trends, it appears as though micro firms and small firms are not likely to employ many youths. The ETI changes this and a trend break can be seen. For medium sized and large firms the pre-policy period predicts and steady increase in the employment of youth. However, non-ETI firms see a decline in the employment of youth in the post-policy period while ETI firms continue to increase. This implies that the ETI enables firms to grow where they might be declined their hiring of youth. Another explanation is that youth that would have been employed in non-ETI firms are now being employed in ETI-firms, so no real job creation has occurred. No break in employment trend can be seen in the very large ETI claiming firms.

Figure 5. Youth employment, Micro firms, 2011-2015

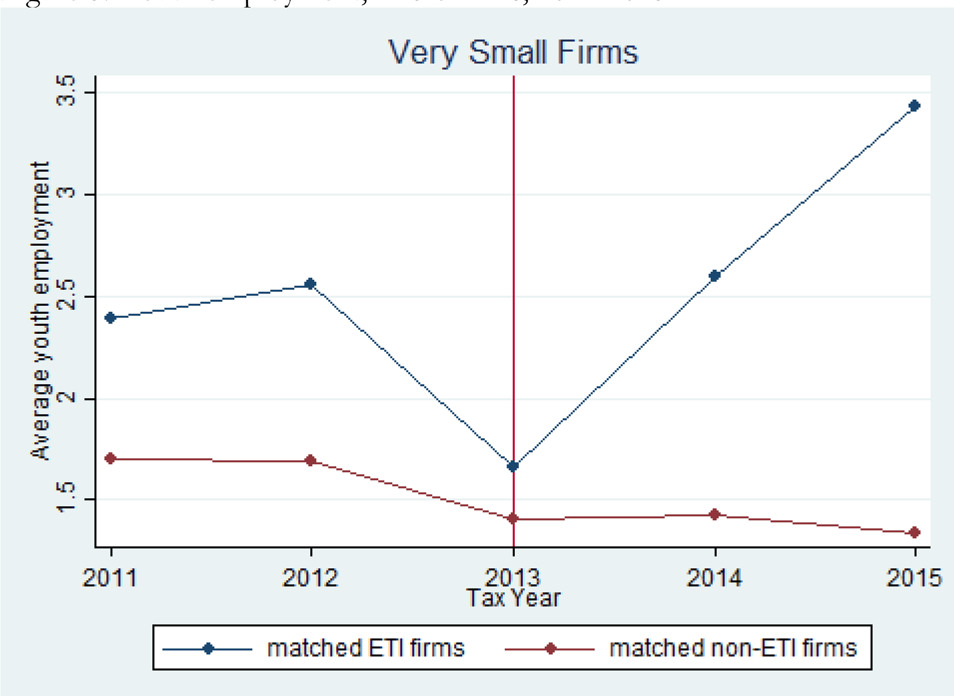


Figure 6. Youth employment, Small firms, 2011-2015

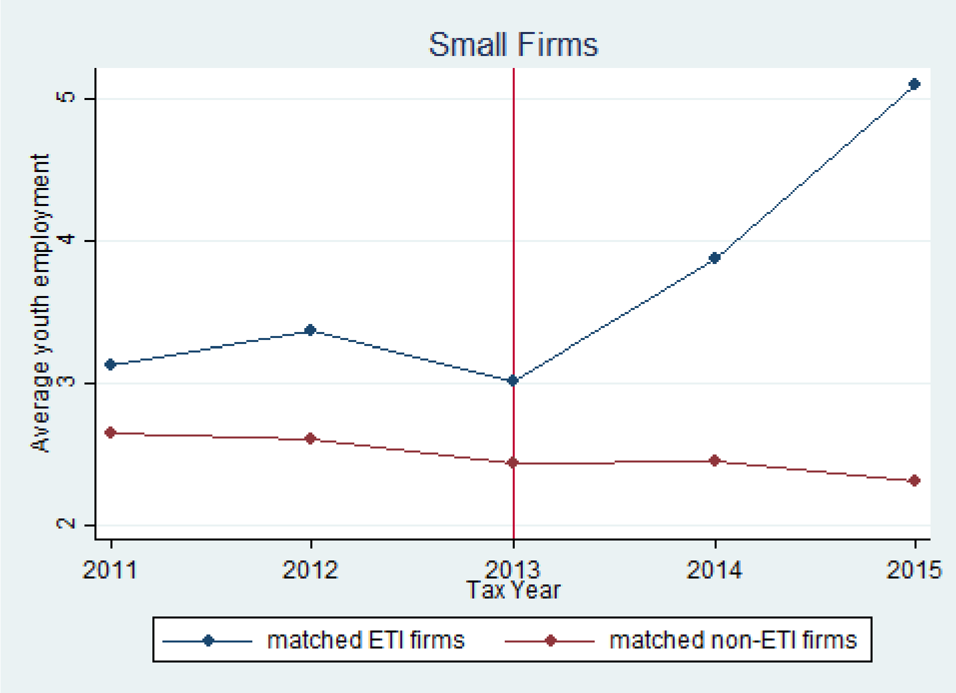


Figure 7. Youth employment, Medium firms, 2011-2015

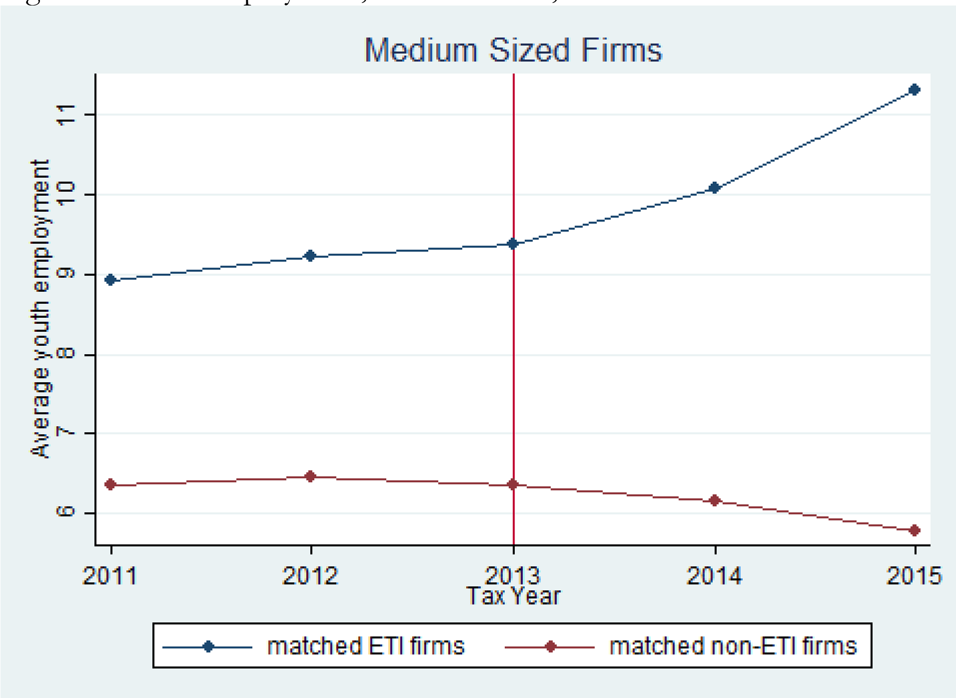


Figure 8. Youth employment, Large firms, 2011-2015

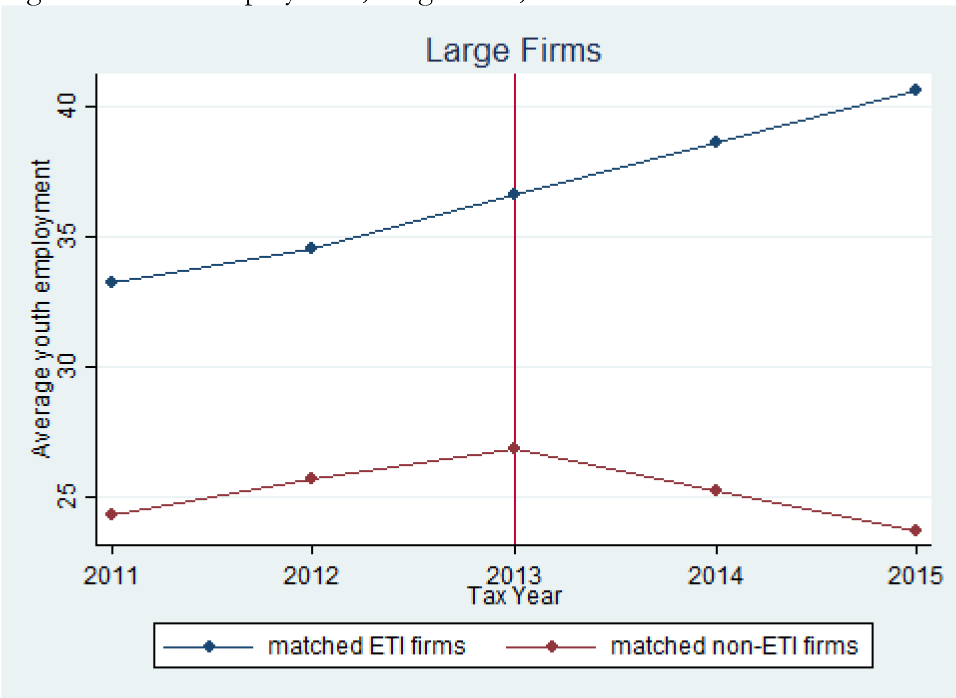


Figure 9. Youth employment, Very large firms, 2011-2015

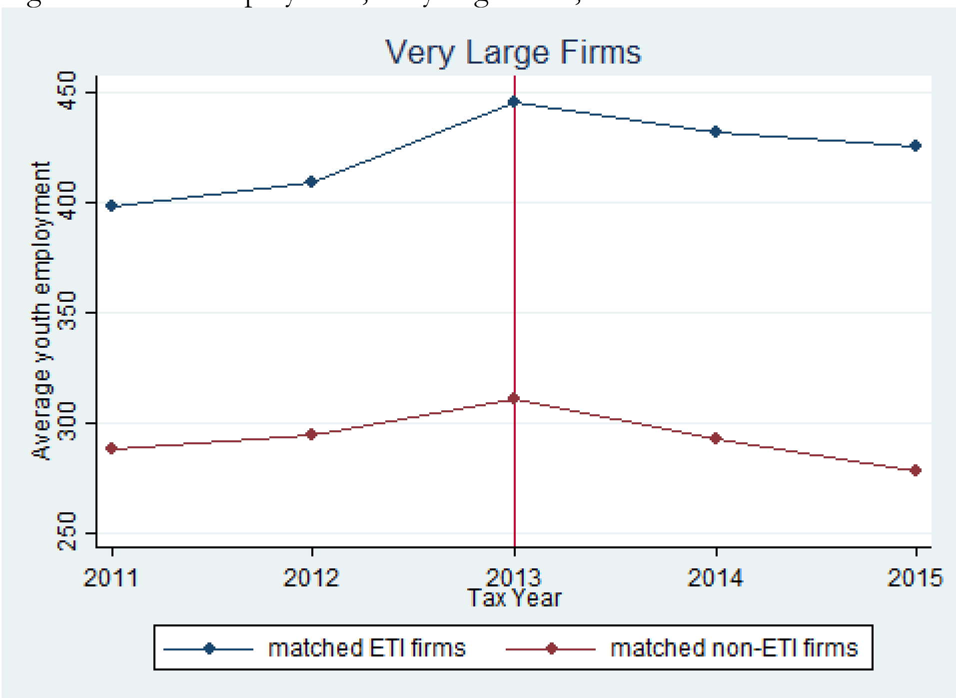
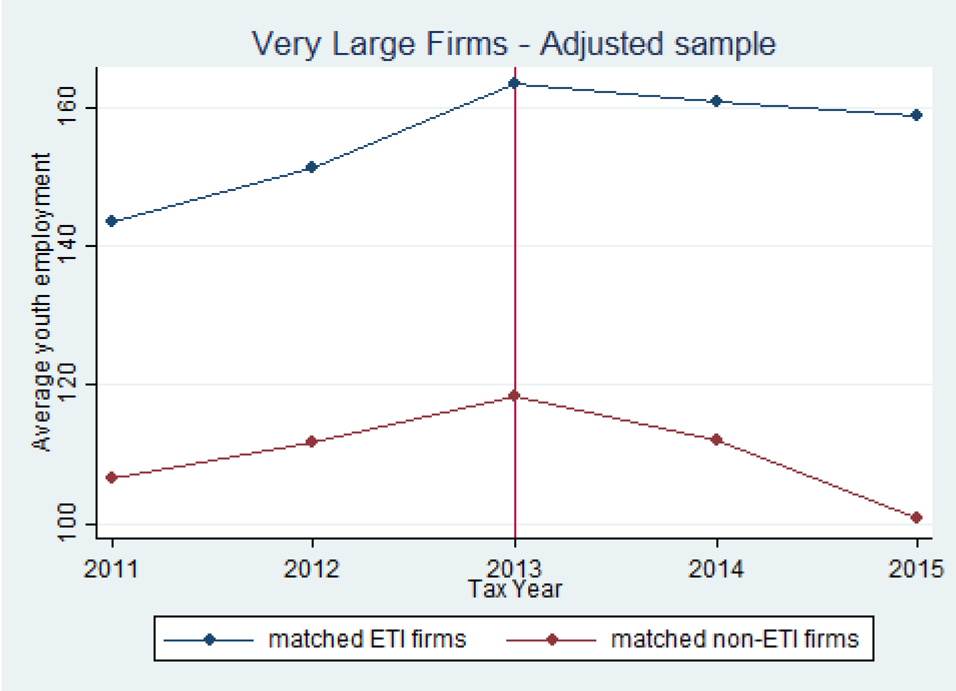


Figure 10. Youth employment, Very large firms – Adjusted Sample, 2011-2015



In summary, the constant effect assumption is violated. This means that for the sample of firms we examine the ETI does not affect firms in the same way.

Birch (1979) argues that job creation happens at small firms and not at large corporations when he analyses employment generation in the USA. In a later work, Birch and Medoff (1994) claim that it is not all small firms that create jobs but a subset of small firms growing rapidly, which the authors term 'gazelles'. Henrekson and Johansson (2010) conduct a meta-analysis of the literature and find that the Birch and Medoff (1994) assertion can also be found in 20 studies in developed countries. Henrekson and Johansson (2010) find that Gazelles create all or a large share of net new jobs and they are, on average, younger and smaller firms. The findings in 29Table 8 support the argument that small firms are creating jobs and large firms are not, where the ETI, an employment incentive, was available to all firms irrespective of size.

7.3 Other dependent variables

We consider three further outcomes to check the sensitivity of the results to our choice of dependent variable. Bruhn (2016) estimates the log of total employment for a wage subsidy programme in manufacturing firms in Mexico. The author uses a similar matching and DID approach and finds positive but not statistically significant effect ranging from 5.7 to 13.2 percent in magnitude. Kangasharju (2007) use payroll as a proxy for total employment at the firm and finds positive statistically significant effects of a wage subsidy on payroll.

We adapt these approaches to estimate the effect of the ETI on the log of youth employment and total payroll at the firm. In addition, we also estimate the impact on youth ratio at the firm, that is, the ratio of youth to total employment. We use kernel matching and remove firms with more than 1200 employees. The results presented below are in line with the results found above.

Table 9. Results of the cDID estimators, log employment and log payroll, 2015

	Number of firms	Estimate	Standard Errors
Log youth employment	14,208	0.134***	0.0122
Log payroll	14,208	0.137***	0.0139
Youth ratio	14,208	0.057***	0.0020

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Source: Authors' own estimates based on IRP5 data.

The estimated effect using log of youth employment is 0.134. This means there is a 14.3% ($e^{0.111} - 1 = 0.143$) increase in youth employment at ETI firms versus non-ETI firms. The estimated effect using log of payroll 0.137. This means there is a 14.7% ($e^{0.137} - 1 = 0.147$) increase in payroll at ETI firms relative to non-ETI firms. And lastly, youth ratio is positive and statistically significant at the 1% level. This means that ETI firms have increase their concentration of youth when compared to non-ETI firms.

7.4 Displacement effects

A big concern raised by COSATU is the displacement of older workers (COSATU, 2013). The policy makes youth relatively cheaper to hire than non-youth, which may induce displacement. Displacement effects have been found in other programmes (Crépon et al., 2013)

If we consider non-youth workers as a substitute for youth workers, there may be a decline in the employment of non-youth workers. This is not the result we see in Table 7 or Table 8. Instead we see an increase in the employment of non-youth. Older workers will have more years of work experience and perhaps earn a higher wage that ETI employees. Firms maybe employing older workers to compensate for the lower productivity in younger workers.

We examine two different sets of ineligible workers that are more likely to be substituted eligible workers. We examine the employment of workers between the ages of 30 and 40 years old and workers between the age of 18 and 29 years old earning between R6,000 and R7,000. We do not consider the prospect of including workers younger than 18 years as we assume a large share of 16-18 years are in school. While there is evidence that students leave school to benefit from the employment subsidy (Webb, 2016), we do not believe that this is the case in South Africa as the policy information was disseminated amongst firms and was less likely to have reached the youth.

While the ETI aims to create jobs for youth in general, the policy is aimed at a very specific group of youth; those between 18 and 29 years old and earning less than R6,000. Among youth, 18 to 29-year-olds, the policy is making a subset of them cheaper to hire. If youth earning just below R6,000 are substitutes for youth earning just above R6,000 then the policy will likely displace ineligible youth.

Table 10. Results of the cDID estimators, displacement effects, 2015

	Number of firms	**Estimate**	**Standard Errors**
Number of 30-40 year-old employees	14,208	2.538***	0.227
Number of youth, not eligible	14,208	0.274***	0.023
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Source: Authors' own estimates based on IRP5 data.			

The results indicate a positive statistically significant increase in the number of 30-40 year olds employed. In terms of magnitude this is less than the additional number of youth employed in Table 7. For the ineligible youth the effect is positive and significant, but the magnitude is much lower than for youth in general. This indicates that no displacement of workers was induced by the policy.

7.5 Firm Growth

In the matching process we take into account firm employment growth to ensure that our estimation does not count intended firm employment growth as a result of the policy. To examine this point more explicitly we look at positive growth firms and negative growth firms separately. We define employment growth as the difference in employment between 2012 and 2013. Positive employment growth refers to firms with a positive growth in 2013, negative employment growth refers to firm with zero or negative employment growth 2013.

Of the 155,428 firms in 2013, 48% have an employment growth of 0 or less. This means that in the period from 2012 to 2013 these firms were shrinking either through laying off existing workers or laying off more workers than they had hired in 2013. On the other side, 52% of these firms report a positive employment growth rate, either through employing additional workers or through hiring more workers than they are laying off. This is slight different for ETI firms. Approximately 38% of firms claiming the ETI in 2015 report an employment growth rate of zero. The results for the two groups are presented in Table 11.

Table 11. cDID estimation, by employment growth

	Number of ETI firms matched	Youth	Standard Errors
Positive-growth firms	11,494	3.169***	0.604
Negative-growth firms	5,645	1.548***	0.570
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Source: Authors' own estimates based on IRP5 data.			

Positive growth ETI firms have a positive significant increase in youth employment. The effect is also positive and significant for negative growth firms but the effect less pronounced. This helps us understand that it is not only positive growth firms creating jobs but the ETI is allowing even the negative firms to create jobs.

Figure 11. Positive growth firm, youth employment trend, 2015

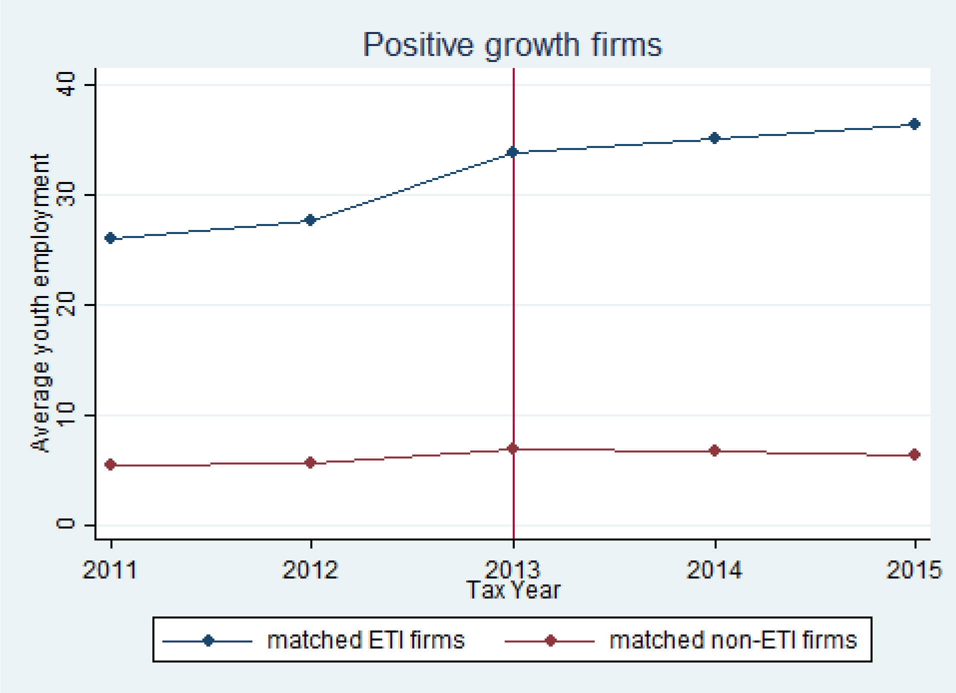
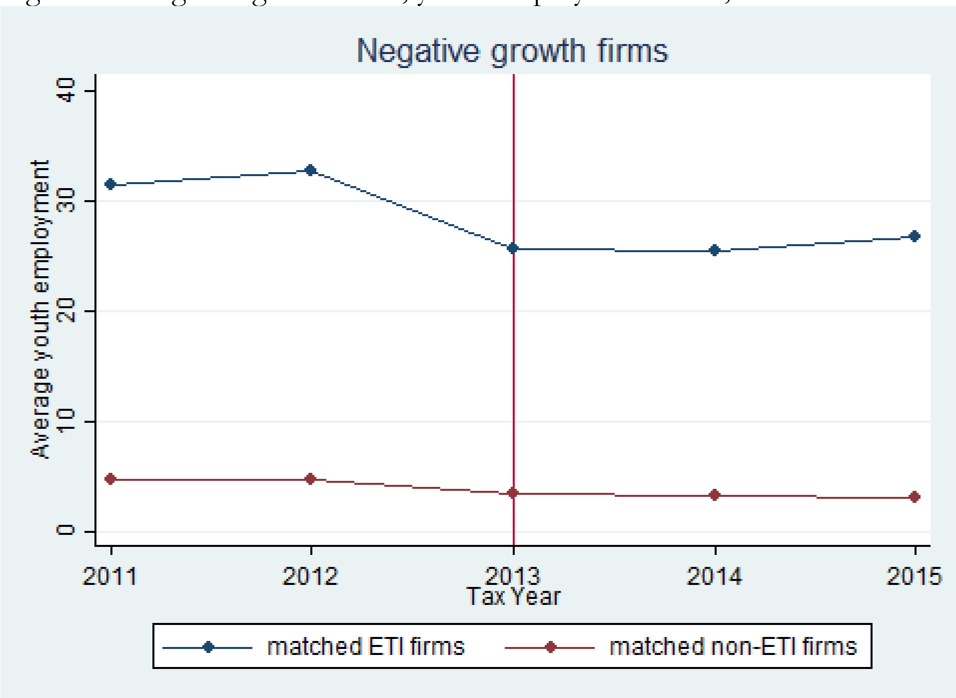


Figure 12. Negative growth firm, youth employment trend, 2015



7.6 Alternate control groups

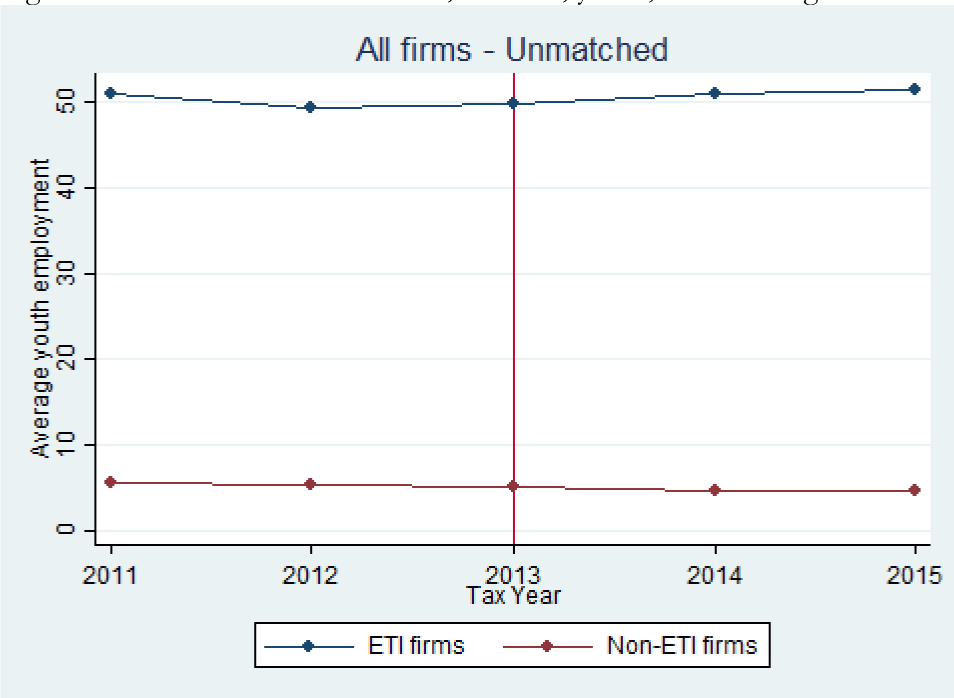
Can try to restrict the control group to firms with at least one eligible youth like the article.

Can restrict control group to firms that do not claim in 2016

7.7 Discussion

The results from the cDID is nested in the distribution of all firms claiming the ETI. If we look at all firms claiming the ETI versus firms not claiming the ETI then there appears to be no change in youth employment.

Figure 13. Difference-in-differences, all firms, youth, no matching

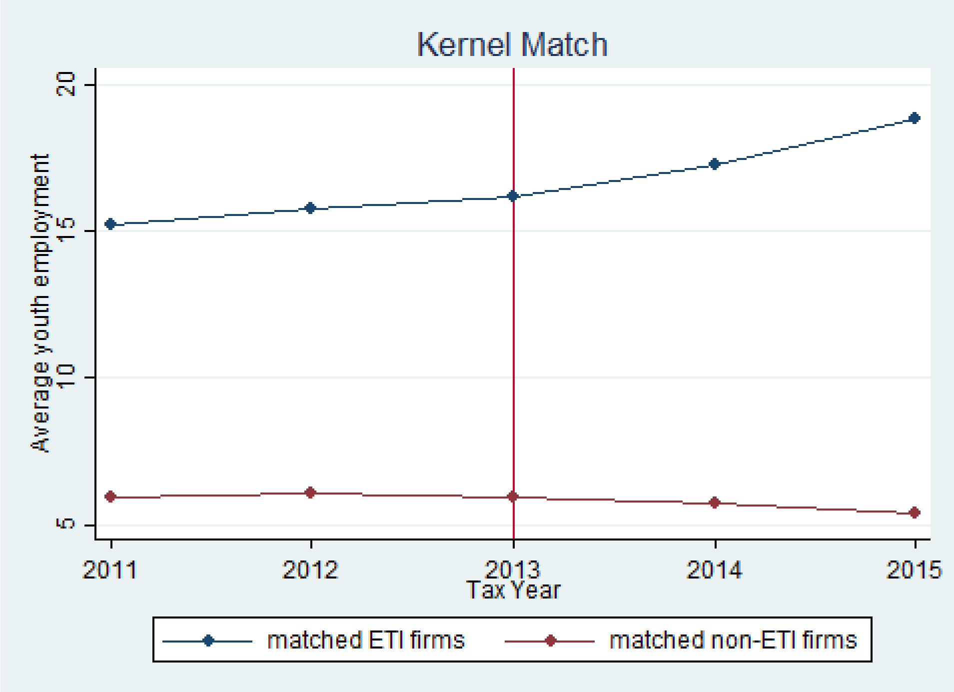


This is perhaps why Ranchhod and Finn (2015) find no change in the probability of youth being employed when estimating the impact of the ETI.

However, we know from our descriptive statistics that small firms are reacting differently to the ETI than very big firms. If we take away the very big firms a slight trend break emerges but this is at the cost of removing firms that claim about half of the ETI and represent the employment of half of the youth in the ETI firms in our study sample. We suspect that these firms are not creating any additional jobs for youth, but we cannot credibly match these firms to firms not claiming the ETI. This is because there are very few large firms not claiming the ETI. The cDID estimation within large firms shows a positive but insignificant result for the employment of additional youth. This gives some validation to our suspicion that large firms are not necessarily creating jobs but the credibility of the matching of large firms cannot be ignored.

This leaves us with a group of firms that are large in number but do not represent the bulk of the ETI claims and youth employment. Acknowledging this, we proceeded with our estimation of the change in youth employment at these firms.

Figure 14. cDID, matched firms, youth employment



This describes a trend break for the firms we were credibly able to match and visually describes Table 7.

The question remains as to why the difference in firm size subgroup matters. It matters in two ways, first, it matters in the sense of policy efficiency. The policy can be better targeted to firms that are creating jobs or likely to create. The second is that of the larger public finance question of who is benefitting from this policy besides young workers. Relating this discussion to the question of the highest 20 ETI claiming firms we see that these are firms with more than 1200 employees.

8 Cost of the policy

To put our findings into context we take a brief look at the cost of the policy vis-à-vis the number of jobs created for youth. We use the estimates from Table 8 to calculate the number of jobs created by multiplying the cDID estimate by the number of ETI firms in the category. This gives us 81,646 new jobs created in the 16,107 firms we matched. We sum up the amount of ETI claimed by these firms. We divide the amount of ETI claimed by the number of jobs created to get the cost per job created.

We calculate a total of 81,646 jobs created in the 2015 tax year at a cost of R955.6 million. This indicates a cost of R11,704 per job created.

Table 12. Estimated costs and number of jobs created in ETI-claiming firms, by tax year

	2015
Estimated number of jobs created	81,646
Estimated cost (in millions Rands)	955.6
Cost per job created (in Rands)	11,704
Source: Authors' own estimates based on IRP5 and CIT-IRP5 data.	

The National Treasury, in its public discussion paper, predicted that 423,000 jobs would be subsidized and 178,000 jobs would be created over three years National Treasury (2011). A deadweight loss is created when a subsidy is claimed for a youth employee a firm already intended to hire in the absence of the subsidy. This means that 245,000 jobs would be subsidized unnecessarily, an expected deadweight loss of approximately 57 per cent. If we consider that for each year of the policy 59,333 jobs were going to be created ($178{,}000 \div 3 = 59{,}333$) then just focusing on the firm we were able to match, the number of jobs created is above the target.

The other side of the policy relates to the cost. In the 2015 tax year, the total amount claimed for the ETI was approximately R2.9 billion rand. Based on our analysis we cannot tell whether the difference, R1.9 billion claimed by firms, was used to create jobs or merely absorbed by firms as economic rent.

9 Conclusion

There is a small body of literature evaluating the effect of wage subsidy policies on labour demand using administrative data. We add to this literature with our examination of the ETI as a policy intervention on labour demand in a developing country. We use administrative tax data with the population of firms that were eligible for the subsidy to estimate the effects of the ETI on employment at firms in South Africa. Using a conditional difference-in-differences approach we examine the change on the youth and non-youth labour markets across firms matched in terms of firm characteristics.

At the aggregate level of youth employment, we see a positive significant change in the number of youth employed the 2015 tax year. The estimated effect is 3.902 additional jobs were created in ETI firms.

When we break down firms into different size categories, we see that firms employing up to 1200 workers have a positive significant effect on youth employment. We are unable to credible match firms with more than 1200 employees as the take-up rate for these very large firms is very high, leaving few very large non-ETI firms with which to match.

In both the aggregate level and within firm subgroups we see that the increase in total employment is greater than the estimation of the increase in youth employment. This is in line with other international studies implying that wage subsidies can enable firm to grow their total employment.

We consider other outcomes of the policy, log of youth employment, log payroll and youth ratio, all with positive significant results indicating that at the firms we are matching there is an increase in youth employment.

We find no evidence of displacement of older works as a result of the policy. We consider that displacement could take place among middle aged workers (30 to 40-year olds) but again find no evidence of this. We also see no displacement of youth workers earning just about the ETI eligibility income criteria of R6,000.

Lastly, we consider the positive and negative growth firms. Estimation of the ETI on both groups are positive and significant indicating that our estimation is not being driven by firm employment growth.

These findings suggest that the youth wage subsidy may be an effective tool for increasing the demand for youth at small to medium size firms. The positive, statistically significant results for non-youth indicate that firms are not laying off workers in order to employ youth and benefit from the subsidy. This means that the subsidy has enabled firms to increase overall employment. The number of jobs created relative to the cost of policy in 2015 leaves questions about the efficiency of the policy. While we are unable to credibly estimate the youth employment impact at very large firms, it is clear from Figure 9 that there is no trend break in youth employment and thus these large firms are not likely to be creating any new jobs but claim in excess of 30% of the subsidy in 2015.

10 References

BETCHERMAN, G., DAYSAL, N. M. & PAGÉS, C. 2010. Do employment subsidies work? Evidence from regionally targeted subsidies in Turkey. *Labour Economics,* 17**,** 710-722.

BHORAT, H., CASSIM, A., KANBUR, R., STANWIX, B. & YU, D. 2016. Minimum Wages and Youth: The Case of South Africa. *Journal of African Economies,* 25**,** i61-i102.

BIRCH, D. L. 1979. *The job generation process: MIT program on neighborhood and regional change*.

BIRCH, D. L. & MEDOFF, J. 1994. Gazelles. *Labor markets, employment policy and job creation,* 159167.

BRUHN, M. 2016. Can wage subsidies boost employment in the wake of an economic crisis? evidence from Mexico. *World Bank Policy Research Working Paper* [Online].

BRYSON, A., DORSETT, R. & PURDON, S. 2002. The use of propensity score matching in the evaluation of active labour market policies.

BURNS, J., EDWARDS, L. & PAUW, K. 2010. Wage subsidies to combat unemployment and poverty: Assessing South Africa's options. *Southern Africa Labour and Development Research Unit Working Paper* [Online].

CALIENDO, M. & KOPEINIG, S. 2008. Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys,* 22**,** 31-72.

CRÉPON, B., DUFLO, E., GURGAND, M., RATHELOT, R. & ZAMORA, P. 2013. Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *The quarterly journal of economics,* 128**,** 531-580.

CRICHTON, S. & MARÉ, D. C. 2013. The Impact of Wage Subsidies on Jobseekers' Outcomes and Firm Employment. *Motu Working Paper* [Online].

DE JONGH, J. J., MEYER, N. & MEYER, D. F. 2016. Perceptions of local businesses on the Employment Tax Incentive Act: the case of the Vaal Triangle region. *Journal of Contemporary Management,* 13**,** 409-432.

DEPARTMENT OF ECONOMIC DEVELOPMENT. 2011. The New Growth Path: Framework.

HECKMAN, J., ICHIMURA, H., SMITH, J. & TODD, P. 1998. Characterizing selection bias using experimental data.

HENREKSON, M. & JOHANSSON, D. 2010. Gazelles as job creators: a survey and interpretation of the evidence. *Small Business Economics,* 35**,** 227-244.

HOLLAND, P. W. 1986. Statistics and Causal Inference. *Journal of the American Statistical Association,* 81**,** 945-960.

HUJER, R., CALIENDO, M. & RADIC, D. 2002. Estimating the effects of wage subsidies on the labour demand in West-Germany using the IAB establishment panel. *Working Paper 47*. Germany: Wirtschaftswissenschaften.

KAISER, U. & KUHN, J. M. 2016. Worker-level and firm-level effects of a wage subsidy program for highly educated labor: Evidence from Denmark. *Research Policy,* 45**,** 1939-1943.

KANGASHARJU, A. 2007. Do Wage Subsidies Increase Employment in Subsidized Firms? *Economica,* 74**,** 51-67.

LEUVEN, E. & SIANESI, B. 2003. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Available: http://ideas.repec.org/c/boc/bocode/s432001.html. .

LEVINSOHN, J. & PUGATCH, T. 2014. Prospective analysis of a wage subsidy for Cape Town youth. *Journal of Development Economics,* 108**,** 169-183.

LEVINSOHN, J., RANKIN, N., ROBERTS, G. & SCHÖER, V. 2014. Wage subsidies and youth employment in South Africa: Evidence from a randomised control trial. *Stellenbosch economic working papers* [Online].

MATTHEE, M., RANKIN, N. & BEZUIDENHOUT, C. 2017. Labour demand and the distribution of wages in South African manufacturing exporters. Helsinki, Finland: UNU-WIDER.

MTEMBU, V. N. & GOVENDER, L. N. 2015. Perceptions of employers and unemployed youth on the proposed youth employment wage subsidy incentive in South Africa : a KwaZulu-Natal study : original research. *SA Journal of Human Resource Management,* 13**,** 1-9.

NATIONAL TREASURY 2011. Confronting youth unemployment: policy options for South Africa. *In:* TREASURY, N. (ed.). Pretoria: National Treasury.

ODENDAAL, P. J. L. 2016. *A comparative analysis of the Employment Tax Incentive Act, no. 26 of 2013.* Masters in Commerce Masters Thesis, University of Cape Town.

PAGAN, A. & ULLAH, A. 1999. *Nonparametric econometrics*, Cambridge university press.

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. *Southern Africa Labour and Development Research Unit Working Paper* [Online].

RANCHHOD, V. & FINN, A. 2015. Estimating the Effects of South Africa's Youth Employment Tax Incentive–An Update. *Southern Africa Labour and Development Research Unit Working Paper* [Online].

ROSENBAUM, P. R. & RUBIN, D. B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika***,** 41-55.

ROTGER, G. P. & ARENDT, J. N. 2010. The Effect of a Wage Subsidy on Employment in the Subsidised Firm. *AKF Working Paper* [Online].

RUBIN, D. B. 1978. Bias reduction using Mahalanobis metric matching. *ETS Research Report Series,* 1978.

SILVERMAN, B. W. 1986. *Density estimation for statistics and data analysis*, CRC press.

SMITH, J. & TODD, P. 2005. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics,* 125**,** 305-353.

VILLA, J. M. 2016. diff: Simplifying the estimation of difference-in-differences treatment effects. *Stata Journal,* 16**,** 52-71.

ZUMA, J. G. 2011. State of the Nation Address, February 2011. *In:* PRESIDENCY, T. (ed.). Pretoria.

11 Appendix

Table A. Gaussian versus Epanechnikov kernel matching

	Number of ETI firms matched	Youth	Non-youth	Total
Gaussian	14,133	3.902*** (0.230)	4.780*** (0.398)	8.704*** (0.594)
Epanechnikov	14,128	4.154*** (0.246)	4.878*** (0.443)	9.060*** (0.652)

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Source: Authors' own estimates based on IRP5 and CIT-IRP5 data.

Table B. Steps taken in matching process

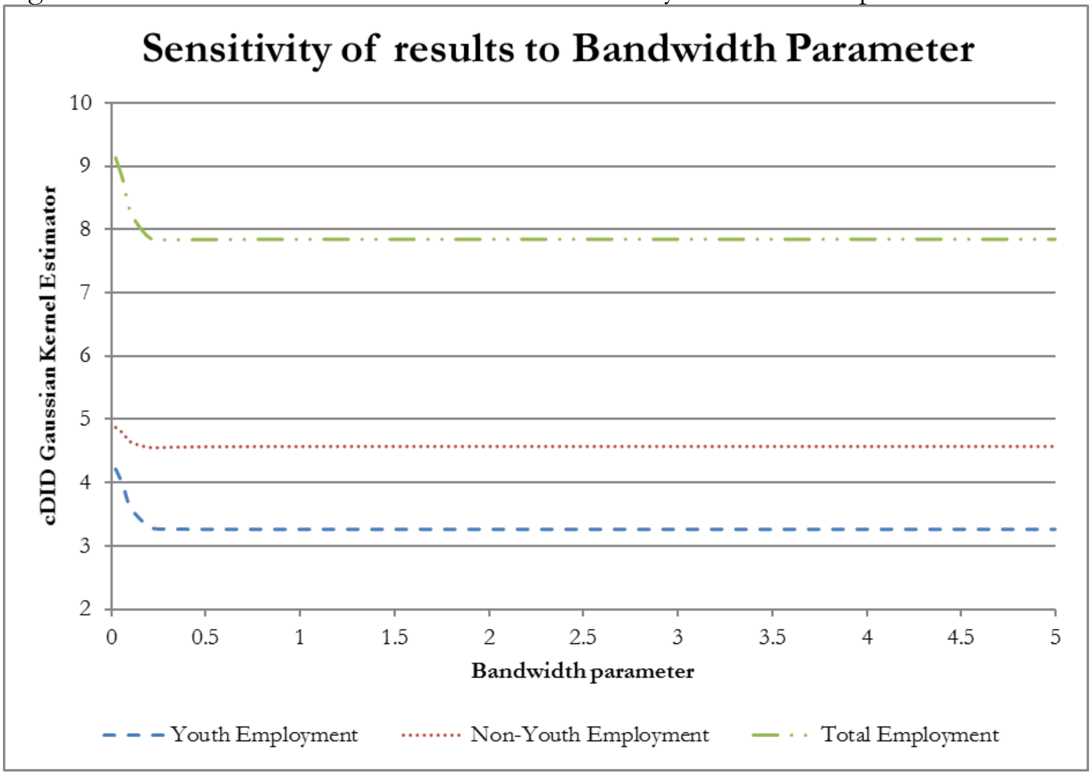
Step 1:	Drop public sector firms Drop firms with no CIT data Keep firms existing since 2011 (balanced panel)	
Step 2:	Drop firms with very high take-up rate (unable to match)	
Step 3:	Estimate propensity scores	21,969
Step 4:	Drop firms with propensity score == 0 (cannot match them)	16,107
Step 5:	Drop firm with very low propensity score (close to 0) (not sufficient common support)	14,133
Step 6:	Use propensity scores to match treatment firms to control firms	
Step 7:		

Table C. cDID estimation over the full common support region, 2015

Number of employees	Number of firms	Coefficient	Standard Error	R^2
Youth	16,104	2.711***	0.519	0.07
Non-youth	16,104	5.629***	0.820	0.07
All	16,104	8.332***	1.262	0.08

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Source: Authors' own estimates based on IRP5 data

Figure A. cDID Gaussian Kernel Estimator sensitivity to bandwidth parameters



band