

The Effect of Language on Labour Market Success for Immigrants and Citizens in South Africa

Samaa Zien¹, Colette Muller² and Claire Vermaak³

February 2019

¹ UKZN (SAEF), Durban, South Africa. Samaasiddig50@gmail.com.

² UKZN (SAEF), Durban, South Africa. Mullerc2@ukzn.ac.za.

³ UKZN (SAEF), Durban, South Africa. Vermaakc@ukzn.ac.za.

Abstract

A sharp increase in the number of immigrants in South Africa has raised questions about the impact of immigration on the labour market. Although immigrants are usually regarded as disadvantaged in their adoptive countries, the use of English as the primary business language in South Africa could affect the success of immigrants in the country's labour market. This study uses Census 2011 data to compare the effect of speaking English on labour market outcomes for both immigrants and local South Africans. In contrast to research in other countries, the study finds that immigrants are more successful than locals in South Africa. Part of immigrants' success can be attributed to the advantage gained from speaking English as a second language, whereas speaking English as a second language is less valued by the labour market for locals. Individuals who speak English are more active, more likely to obtain employment and earn higher wages relative to non-English speakers. The study finally shows that, although female immigrants benefit more from the ability to speak English, male immigrants are more successful overall in South Africa's labour market.

Keywords: immigrants, language, labour force participation, employment, earnings, probit, interval regression.

JEL classification: J21 J31 F22 Z13

Acknowledgements: First and foremost, I would like to express my deepest gratitude to the Mighty **Allah** (God) for without His graces and blessings, this study would not have been possible. Second, I would like to express my sincere appreciation and thanks to my supervisors **Dr. Claire Vermaak** and **Dr. Colette Muller** for allowing me the opportunity to work with them and learn from them. Their mentorship and support were key factors for the completion of the study. Furthermore, I would like to thank **UNU-WIDER** for their financial support and for the opportunity to share the study on a wider platform and get feedback from experts in the field. Third, I acknowledge the patronage and loving inspiration from my family and friends.

1 Introduction

Recent estimates of immigration by the United Nations indicate that immigration has increased worldwide (United Nations, 2018). The scale of immigration to South Africa (SA) has also increased substantially in the post-apartheid period. This was to be expected as the immigration policies were stricter prior to democracy, especially from neighbouring countries (Facchini *et al*, 2013; Zuberi and Sibanda, 2004). The magnitude of these inflows has not gone unnoticed in popular discourse around the impact of immigration on the South African labour market; however, there is only a limited literature that formally considers the labour market effects of immigration in South Africa to date.

International research that considers the performance of immigrants relative to citizens in their host country's labour market typically reveals that immigrants' position is unfavourable (Bell, 1997; Chiswick, 1978; Clark and Lindley, 2009; Euwals *et al*, 2010; Evans, 1987; Hayfron, 2001; Lindemann and Kogan, 2013; Paolo and Raymond, 2012). Upon entry to a foreign country, immigrants usually face an initial disadvantage compared to native-born individuals (Bell, 1997; Chiswick, 1978; Euwals *et al*, 2010) that is partially a result of the immobility of knowledge and skills across countries. A limited knowledge of customs, culture, language, employment opportunities, as well as firm-specific required skills in the host country, can negatively affect immigrants' assimilation into a new labour market. The extent to which immigrants' performance in the host country's economy improves over time is often attributed to their ability to assimilate into the new labour market (Chiswick and Miller 2002; Dustmann and Fabbri, 2003; Evans, 1987; Shields and Price, 2002). One of most important human capital factors that influences assimilation is that of language: immigrants who speak the titular language of their host country are likely to assimilate and perform better relative to those who do not (Dustmann and Fabbri, 2003; Lindemann and Kogan, 2013; Paolo and Raymond, 2012).

In South Africa, there are eleven official languages, and despite English being the first language of only a minority of citizens, it is the *de facto* language of business in the country. Research shows that individuals who speak English proficiently perform better in the South African labour market than those who do not (Cornwell and Inder, 2008; McKenzie and Muller, 2017; Posel and Casale, 2011). However, the extant immigrant literature, which, in contrast to much of the international literature, reveals that immigrants are more likely to participate and be employed relative to natives (Peters and Sundaram, 2015; Zuberi and Sibanda, 2004), fails to acknowledge or consider the possible effect that speaking English may have on the labour market performance of immigrants

as compared to citizens. Given that immigrants to South Africa are much more likely than locals to speak English as either a first or a second language, an interrogation of whether and how speaking English affects the labour market performance of immigrants relative to citizens is warranted.

In this paper we address this knowledge gap by interrogating whether and how speaking English affects labour market outcomes for immigrants relative to citizens and in particular, whether immigrants who speak English are advantaged in the South African labour market. Using data from the 2011 Census, we distinguish between citizens and immigrants, and account for whether immigrants and citizens speak English as a first or second language, or not at all.

2 Data and Methodology

The primary data source for this study is the 10 percent sample⁴ of the 2011 South African Census, which is the most recent Census conducted by Statistics South Africa (STATS SA). Prior to the advent of democracy in 1994, the reliability of the information collected for the country as a whole was disputable. However, since the abolition of apartheid, STATS SA has collected reliable Census data in 1996, 2001 and most recently in 2011. For the purposes of exploring changes in immigration over time, the 10 percent samples of the 1996 and 2001 Census data are also utilised.

The Census captures demographic information about individuals as well as information about their households, and a large number of immigrants can be identified, even when working age and population group restrictions are imposed on the dataset. While the identification of a reasonable sample of immigrants is critical for this study, one concern is that the immigrants who respond to surveys and whose information is contained in the data used could be more likely to speak English and/or be highly educated relative to immigrants who do not respond (for instance, illegal immigrants). This concern may be partially mitigated by the *Post Enumeration Survey*, conducted directly after the Census by STATS SA, which aims to address any over and/or undercount of individuals (Vermaak and Muller 2019).

In addition to enabling an identification of immigrants, the Census collects data on the key explanatory variable “language”. In the language acquisition section of the questionnaire respondents were asked to indicate: “What are the two languages that the individual speaks most

⁴ Due to dataset size and confidentiality concerns, the complete Census data are not released to researchers. The 10 percent sample therefore is the publicly accessible form of the Census data. The dataset contains weights that are used to make this sample representative of the population of South Africa.

often within the household?” In response to this question, it is possible to report speaking only one language, in which case we term the individual as monolingual.⁵ Based on the individuals’ responses we categorise language into five groups: monolingual English speakers, monolingual non-English speakers; bilingual first-language English speakers; bilingual second-language English speakers and bilingual non-English speakers.

Much of the literature on language emphasises the importance of language proficiency, rather than just speaking ability, so the limited questioning of the Census around language (as well as limited questioning more generally) does restrict the analysis to some extent. Census data has been used in other studies of immigration in South Africa, however, and given the paucity of immigration information contained in other national household survey data sets, the Census provides the best data option for this study.

We formally investigate the effect of language on labour market success for black African immigrants and citizens of working age (15 to 65 years)⁶ in South Africa by estimating the following equation:

$$Y_i = \beta_0 + \beta_1 \text{immig}_i + \boldsymbol{\beta}' \text{LangCat}_i + \boldsymbol{\delta}' \text{immig}_i * \text{LangCat}_i + \boldsymbol{\alpha}' \mathbf{X}_i + \varepsilon_i$$

Outcome variables, namely labour force participation (defined using the strict definition), employment and earnings.

The two main variables of interest of the study are immigrant status ($\text{immig}_i=1$ if the individual is an immigrant and 0 if s/he is not) and languages spoken (LangCat_i). The five language categories are included in the regressions using four dummy variables, with monolingual non-English speakers as the reference category.

The language variables are also interacted with the immigrant status dummy, resulting in four interaction terms that provide the basis for the immigrant-citizen language comparison. Therefore, β_1 represents the difference in labour market success between immigrants and citizens who are monolingual non-English speakers (the language reference category). The vector $\boldsymbol{\beta}'$ represents the effect of language (relative to monolingual non-English speakers) on the labour market outcomes

⁵ Note that the survey does not permit individuals to report speaking more than two languages. It is therefore not possible to determine whether individuals who are classified as bilingual are actually multilingual (speaking 3 or more languages). Furthermore, among those classified as bilingual non-English speakers, it is not possible to determine whether English is spoken as perhaps a third or fourth language.

⁶ The black African population group is South Africa’s largest population group, and immigrants to South Africa are also largely of African descent. To ameliorate against sample size concerns (particularly in terms of the immigrant-citizen comparison), the sample is restricted to Africans only. Furthermore, as the focus is on labour market outcomes, the analysis focuses on the working age population.

for citizens. The relevant element of $\beta' + \delta'$ represents the total effect of the relevant language combination on immigrants' labour market success, which also captures the comparison between immigrants and citizens in the same language category.

As presented in the equation, the models also include a standard set of explanatory variables typically included in models explaining labour market participation, employment and earnings. These are broadly categorised as human capital variables (age and education controls), marital status, selected household composition variables and, for the employed, selected job characteristics. The different language combinations and education levels among immigrants and citizens help to account for differences in labour market outcomes and performance while the characteristics of the household are included to control for household conditions - penalties or incentives - that may affect the performance of immigrants and citizens (Dustmann and Fabbri, 2003; Posel and Casale, 2011). Most of the explanatory variables comprise dummy variables and are binary in nature, however, some are continuous as shown in Table 2.1 below.

Table 2.1 Types of the Explanatory Variables

<i>Dummy variables</i>	<i>Continuous variables</i>
Immigrant status	Age
Language	Household size
Education (five categories ⁷)	Number of other employed members residing in the household
Marital status (four categories)	Number of children (aged < 7 years) residing in the household
Sector of employment (three categories, for the employed only)	Duration of residency
Industry of employment (seven categories, for the employed only)	

The same set of explanatory variables are used when the dependent variable is labour force participation or employment, while sector and industry related job characteristics are included only when estimating the earnings equation. The different sectors and industries have different propensities to employ based on many factors, including language, therefore their inclusion in the model is of essence (Paolo and Raymond, 2012).

All three labour market outcomes are estimated separately by gender to account for different employment propensities, men and women's choices and for differences in the nature of the types of jobs men and women occupy. Several unfolding model specifications are used, which

⁷ Education is included as categories rather than as continuous years of schooling to allow for varying returns to different levels of education, as is standard in the South African labour literature (for example, Posel and Casale, 2011).

sequentially include immigrant status and language categories, followed by the language interaction variables, and that finally include a full set of control variables. A probit model is used to estimate labour force participation and employment, as both of the dependent variables are binary in nature (1 if the individual participates or is employed, and 0 if the individual does not participate or is not employed). To estimate the earnings equation, an interval regression model is used as the income data collected in the Census are categorical.

A key concern in terms of the proposed methods of estimation presented above relates to the possible endogeneity of immigrant status. For example, immigrants may enter South Africa with confirmed employment, and in addition are unlikely to comprise a random sample of individuals from their home countries. As noted by Vermaak and Muller (2019:4), “if immigrants are, on average, positively (negatively) selected into the South African labour market then any immigrant disadvantage will be understated (overstated)”. Not only is the potential selectivity of immigrants a concern, but self-selection into employment may also confound the results, particularly when estimating the earnings equation. Unobservable factors that are correlated with employment (such as ability and/or motivation) could also be correlated with immigrant status, thereby biasing the estimated coefficients not just for immigrant status but also for the language interaction terms. Language acquisition may follow from engagement with the labour market, in addition to preceding it, which may bias the estimated language effects through bidirectional causality.

Addressing issues related to selection and endogeneity biases are typically challenging, even more so in cross-sectional datasets of limited breadth such as the Census. Attempts to control for the endogeneity of immigrant status by Vermaak and Muller (2019), who, working with the 2011 Census data, used a combination of individual-level and macroeconomic variables to instrument for immigrant status, yielded mixed results.⁸ Furthermore, their study highlighted that it was not possible to also control for selectivity into employment in addition to the endogeneity of immigrant status in the earnings estimation. In this study, although no formal attempts are made to control for endogeneity or selection in the estimations, the likely bias and the implications for the results are highlighted.

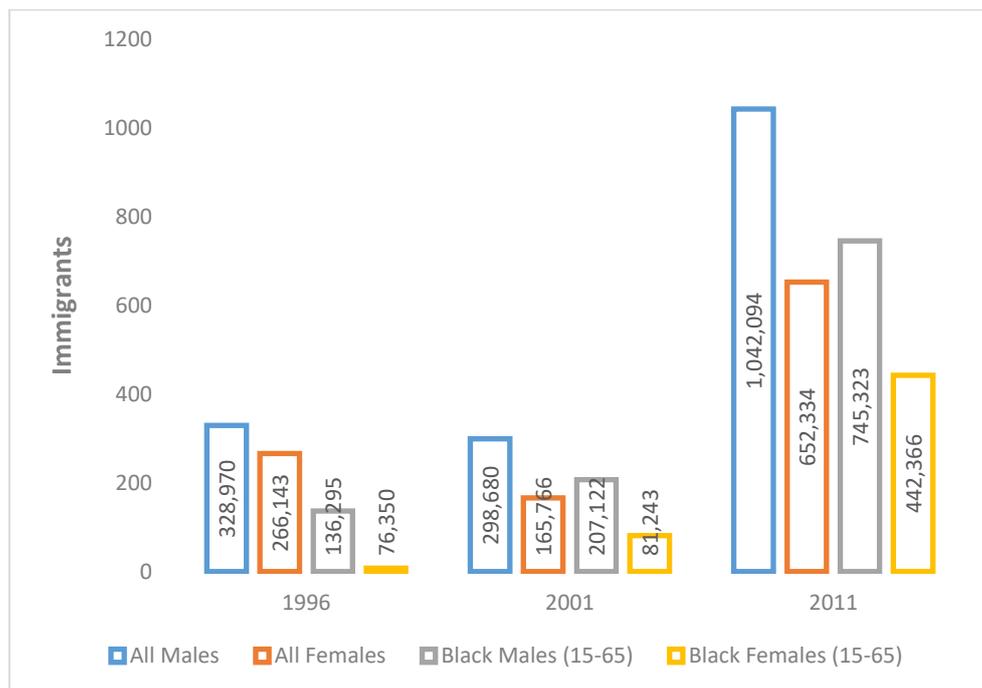
⁸ In particular, Vermaak and Muller (2019) distinguished between two immigrant groups in their analysis of the labour market outcomes for immigrants versus citizens, namely naturalised immigrants (citizens born abroad), and foreigners. Their findings suggested that there may be differences in the direction of the selection effect for the different immigrant groups, with foreigners being negatively selected and naturalised citizens being positively selected.

3 Descriptive Analysis

Individuals often immigrate to countries that are economically more advanced, enjoy high employment rates and are more stable than their home country, such as developed and industrialised countries (Trebilcock, 2003; Zuberi and Sibanda, 2004). This renders South Africa a very interesting case. Being a developing country and yet the largest economy in Southern Africa, individuals from all over the world and especially from neighbouring African countries immigrate to South Africa (Facchini *et al*, 2013).

One of the longstanding features of South African history is international immigration, especially from neighbouring countries. Figure 3.1 uses the 1996, 2001 and 2011 Census data to show the rise in immigration that South Africa has experienced since the abolition of apartheid. The figure further shows the total number of immigrants (all race and age groups) and the restricted sample of immigrants (Black immigrants aged 15-65), by gender.

Figure 3.1 The Immigration Trend



Note: The results are weighted.

Source: Authors' illustration based on Census 1996, 2001 and 2011.

Despite a slight decline in 2001, from 1996 to 2011 the total number of immigrants in South Africa almost tripled, from just fewer than 600,000 in 1996 to nearly 1.7 million in 2011. The number of black immigrants of working age (who comprise roughly of 82 percent of immigrants in SA) also

increased over the period: by almost 36 percent from 1996 to 2001 and more than 300 percent from 2001 to 2011. This increase may either be driven by economic factors such as higher earnings and employment opportunities (see Borjas, 1994; Facchini *et al*, 2013; Marquez, 1997; Trebilcock, 2003; Zuberi and Sibanda, 2004), or non-economic factors such as racial and cultural implications and security (Hainmueller and Hiscox, 2007; Mayda, 2006).

In addition, although the black immigrant population comprises mostly males, immigration to South Africa has become increasingly feminized, with black female immigration increasing nearly five-fold to more than 440 000 in 2011. The rise in female immigration to South Africa is in line with international immigration trends, where an increase in women’s immigration has been recognized. Likely reasons for this may include a growing demand for immigrant women’s labour in destination countries (healthcare, domestic and manufacturing sectors), and the fact that women have become independent immigrants and/or primary economic providers.

Descriptive statistics comparing the sample of working age black immigrants to citizens by gender are shown in Table 3.1. More detailed descriptive statistics showing the characteristics of the respective subsamples of labour force participants and the employed, are provided in the Appendices.

Table 3.1 Sample Characteristics

	Male citizen	Male immigrant	Female citizen	Female immigrant
English first, no second language	0.011 (0.000)	0.027*** (0.001)	0.012 (0.000)	0.028*** (0.001)
English first, other second language	0.014 (0.000)	0.094*** (0.001)	0.015 (0.000)	0.096*** (0.002)
Other first, English second language	0.325 (0.000)	0.338*** (0.002)	0.315 (0.000)	0.360*** (0.003)
Other first, other second language	0.208 (0.000)	0.302*** (0.002)	0.193 (0.000)	0.263*** (0.002)
Other first, no second language	0.441 (0.001)	0.240*** (0.002)	0.466 (0.000)	0.253*** (0.002)
Age	32.432 (0.013)	31.183*** (0.037)	33.610 (0.013)	29.810*** (0.046)
No education	0.062 (0.000)	0.080*** (0.001)	0.077 (0.000)	0.073*** (0.001)
Primary education	0.177 (0.000)	0.198*** (0.002)	0.161 (0.000)	0.168*** (0.002)
Incomplete secondary education	0.436 (0.001)	0.432* (0.002)	0.425 (0.000)	0.466*** (0.003)
Matric	0.262 (0.000)	0.203*** (0.002)	0.262 (0.000)	0.207*** (0.002)
Tertiary education	0.063 (0.000)	0.086*** (0.001)	0.075 (0.000)	0.087*** (0.001)
Married	0.354 (0.000)	0.537*** (0.002)	0.360 (0.000)	0.643*** (0.003)

Table 3.1 Continued

	Male citizen	Male immigrant	Female citizen	Female immigrant
Widower/widow	0.010 (0.000)	0.004*** (0.000)	0.045 (0.000)	0.023*** (0.001)
Divorced	0.013 (0.000)	0.008*** (0.000)	0.022 (0.000)	0.018*** (0.001)
Never married	0.623 (0.000)	0.451*** (0.002)	0.573 (0.000)	0.316*** (0.002)
Household size	3.095 (0.002)	2.095*** (0.006)	3.186 (0.002)	2.176*** (0.007)
No. of young children (age<7) in the household	0.627 (0.001)	0.317*** (0.003)	0.928 (0.001)	0.591*** (0.004)
No. of other employed members in the household	0.558 (0.001)	0.612*** (0.004)	0.589 (0.001)	0.740*** (0.004)
Narrow labour force participation	0.569 (0.001)	0.848*** (0.001)	0.484 (0.000)	0.656*** (0.002)
Broad labour force participation	0.664 (0.000)	0.888*** (0.001)	0.607 (0.000)	0.745*** (0.002)
Employment (among official labour force participants)	0.682 (0.001)	0.846*** (0.002)	0.585 (0.001)	0.656*** (0.003)
Annual income (among the employed)	68,661 (270.712)	48,425*** (587.441)	59,419 (266.759)	37,700*** (910.220)
Sample size	983,732	60,179	1,112,027	36,816
Population size	11,720,997	734,872	12,915,809	437,031

Notes: Sample is restricted to black Africans aged (15-65). The table results are weighted, robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1 for significance test “Male immigrant subject to Male citizen” and “Female immigrant subject to Female citizen”. Annual income is measured in 2011 prices.

Source: Authors’ calculations based on Census 2011.

The descriptive statistics for the language categories reveal that despite English being the dominant language of business in South Africa, it is not the dominant spoken language for African citizens. Only 2.5 percent of monolingual male citizens and less than 3 percent of monolingual female citizens report speaking English as a first language. Substantially more citizens report speaking English as a second language (more than 30 percent of men and of women), but the majority of citizens in South Africa are monolingual non-English speakers. More than 44 percent of male citizens and more than 46 percent of female citizens are reported not to speak English as a first language, and have no reported second language (see also Posel and Casale (2011) who indicate that most South African citizens speak an official language other than English). In contrast, both male and female immigrants are more likely than citizens to report speaking English as a main language - more than 12 percent of male immigrants and more than 12 percent of female immigrants speak English as their first language. Immigrants are also significantly more likely than citizens to report speaking English as a second language: nearly 34 percent of male immigrants and 36 percent of female immigrants are second-language English speakers. The proportions of male

and female immigrants who are monolingual but do not speak English are also much lower than for citizens.

In terms of other characteristics, immigrants tend to be younger than citizens on average. With the exception of tertiary education, which a significantly higher proportion of male and female immigrants report having completed relative to citizens, the proportions of immigrants within the lower educational attainment categories are typically higher than for citizens. More than 70 percent of immigrants have not achieved a matric level of education or higher, as compared to less than 70 percent of citizens. A significantly higher percentage of citizens (more than a quarter of men and of women) than immigrants (only about one-fifth of immigrant men and women) report having completed matric or its equivalent. In addition, immigrants are more likely than citizens to be married, while the majority of citizens report having never married. Immigrants typically live in smaller households, with fewer young children than citizens, but where the number of other household members who are employed is higher, on average, than for citizens.

The final rows of Table 3.1 reveal some of the characteristics of immigrants and citizens in terms of labour market outcomes. Both the narrow⁹ and broad¹⁰ definitions of labour force participation reveal that immigrants are significantly more likely than citizens to participate in South Africa's labour market and that, within the immigrant and citizen subsamples, men are more likely to participate than women. The broad definition of labour market participation, for example, shows that nearly 85% of male immigrants, and almost three quarters of female immigrants participate, as compared to two thirds of male citizens and about sixty percent of female citizens. The problem of discouraged workers among South Africans is also evident in these statistics, as the broad rates of labour market participation are much higher relative to the strict participation rates for citizens than for immigrants.

Not only are immigrants more likely to participate than citizens, but immigrants, particularly immigrant men, are also more likely to be employed. However, once employed, the average annual income for immigrants is much lower than for citizens, with female immigrants receiving the lowest income (less than R40 000 per annum), on average.

⁹ Includes individuals who are employed or who are unemployed and have been actively seeking work.

¹⁰ Includes individuals who are employed or who are unemployed and have been either actively seeking work or are discouraged seekers.

4 Regression Analysis and Results

To further interrogate labour market outcomes among immigrants and citizens, Table 4.1 shows labour force participation probabilities (marginal effects) by gender.¹¹ Three models are estimated for each gender. Model **I** includes controls for immigrant status and English acquisition levels, with no other interaction terms or control variables included. The marginal effects shown in Model **I** therefore partially capture the effect of other variables that influence labour force participation, but which are not yet accounted for in the model. Model **II** includes the immigrant-language interaction terms in addition to the controls included in Model **I**, and Model **III** further includes individual and household characteristics. The marginal effects are expected to change in each model as the additional factors that influence labour force participation are accounted for.

Across all specifications, both female and male immigrants are significantly more likely to participate in the labour force compared to female and male citizens, in the base category. In Model **I**, all of the language categories have positive and significant effects. This suggests that, irrespective of immigrant status, women and men with any language combination are more likely to participate in the labour force compared to monolingual non-English speakers. For example, bilingual women who speak English as a first language are, on average, 19.1 percent more likely to participate than monolingual women who do not speak English, after controlling for immigrant status. In addition, the effects of language on participation are larger for women than for men. Part of these participation advantages, however, are likely the result of not having controlled for other correlates associated with labour market participation in this specification.

The second model specification (**II**) includes the interaction terms for immigrants and language acquisition levels. Here, the effect of the language dummy variables is interpreted exclusively for citizens, while the language effect for immigrants is given by $\beta' + \delta'$. To ease the viewing and interpretation of the results, the relevant coefficients for immigrants have been summed and converted into marginal effects, and are displayed in Table 4.1 as 'Language effects for immigrants only'.

¹¹ The labour market estimations are presented here using the strict definition of participation. The results obtained using the broad definition of participation are qualitatively similar, although the magnitudes of the marginal effects differ.

Table 4.1 Probit models for Labour Force Participation

	Females			Males		
	I	II	III	I	II	III
Immigrant	0.142*** (0.003)	0.192*** (0.005)	0.170*** (0.006)	0.260*** (0.002)	0.307*** (0.003)	0.248*** (0.004)
<i>Language effects for all</i>						
English first, no second language	0.101*** (0.004)			0.073*** (0.004)		
English first, other second language	0.191*** (0.003)			0.152*** (0.003)		
Other first, English second language	0.145*** (0.001)			0.131*** (0.001)		
Other first, other second language	0.121*** (0.001)			0.142*** (0.001)		
<i>Language effects for citizens only</i>						
English first, no second language		0.101*** (0.004)	0.050*** (0.005)		0.075*** (0.004)	0.029*** (0.005)
English first, other second language		0.204*** (0.004)	0.089*** (0.004)		0.177*** (0.003)	0.068*** (0.004)
Other first, English second language		0.146*** (0.001)	0.077*** (0.001)		0.133*** (0.001)	0.069*** (0.001)
Other first, other second language		0.123*** (0.001)	0.084*** (0.001)		0.144*** (0.001)	0.090*** (0.001)
<i>Language effects for immigrants only</i>						
English first, no second language		0.065*** (0.017)	-0.070*** (0.019)		-0.001 (0.015)	-0.115*** (0.018)
English first, other second language		0.077*** (0.010)	-0.071*** (0.011)		-0.011 (0.009)	-0.145*** (0.010)
Other first, English second language		0.091*** (0.007)	0.003 (0.008)		0.036*** (0.006)	-0.051*** (0.007)
Other first, other second language		0.040*** (0.007)	0.028*** (0.008)		0.057*** (0.006)	0.039*** (0.008)
Immigrants' duration of residence			0.000 (0.000)			0.000** (0.000)

Table 4.1 Continued

	Females			Males		
	I	II	III	I	II	III
Primary education			0.093*** (0.002)			0.089*** (0.002)
Incomplete secondary education			0.168*** (0.002)			0.132*** (0.002)
Matric			0.266*** (0.002)			0.212*** (0.002)
Tertiary education			0.378*** (0.002)			0.241*** (0.002)
Age			0.082*** (0.000)			0.086*** (0.000)
Age squared			-0.001*** (0.000)			-0.001*** (0.000)
Married			-0.036*** (0.001)			0.157*** (0.001)
Widower/widow			-0.011*** (0.003)			0.061*** (0.005)
Divorced			0.075*** (0.004)			0.039*** (0.005)
Household size			-0.032*** (0.000)			-0.045*** (0.000)
No. of young children in the household			-0.020*** (0.001)			-0.018*** (0.001)
No. of other employed in the household			0.103*** (0.001)			0.112*** (0.001)
Sample	1156307	1156307	1151294	1051571	1051571	1046476
Pseudo R ²	0.0165	0.0167	0.1569	0.0290	0.0293	0.2163

Notes: Sample is restricted to black Africans aged 15-65. The results are weighted, robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1 for significance levels.

Source: Authors' calculations based on Census 2011.

Reference categories: citizens, another first and no second language, no schooling and never married.

For citizens, all language groups are more likely to participate than monolingual non-English speakers. Relative to the reference group, bilingual citizens who speak English as a first language have the highest participation probability. English as a second language has the second largest effect for women, and the third largest effect for men. Among immigrants, speaking English as a second language is most beneficial for women. The language effects are mostly smaller and less significant for immigrant men than immigrant women. For both men and women, the participation benefit of speaking English is lower for immigrants than for citizens, especially among men. However, immigrants retain an average participation advantage.

Model **III** includes the full complement of control variables. It shows that immigrants, irrespective of gender, retain a significantly higher probability of participating in the labour force compared to citizens, even after controlling for their observable characteristics. This difference may be explained by immigrants having less access than citizens to sources of financial support outside of the labour market, such as government grants and extended family resources. In addition, it may reflect a reverse causality in that participating in the labour market is the reason that immigrants come to South Africa. Male immigrants have a larger participation advantage over citizens than female immigrants. The language effects for citizens are smaller in magnitude compared to Model **II**, suggesting that citizens' language skills are positively correlated with other characteristics that enhance participation. For immigrants, this correlation and decline in the marginal effects is much larger than for citizens, such that the language effects are mostly negative or insignificant.

Comparing English acquisition levels only, the results reveal that each acquisition level has a larger positive impact for citizens' participation probability than immigrants' participation probability. Among immigrants, irrespective of gender, the impact of English on the participation probability is smaller (or more negative) for individuals who speak English as a first language relative to those who speak it as a second language. The differences in the language combination effects will be discussed in more detail in the later models.

In terms of the other correlates included in Model **III**, the results for the duration of residency, which is used as a proxy for assimilation in the host country's labour market, shows that, on average, the longer immigrants reside in South Africa the higher is their probability of participating in the labour force. However, these effects are very small, and are only significant for immigrant males.

In terms of education, as both females and males achieve higher education levels, *ceteris paribus*, their probability of participating in the labour force increases. This occurs particularly for tertiary education: females with tertiary education, for instance, are almost 40 percent more likely to

participate in the labour force than those with no schooling, *ceteris paribus*. The participation probabilities are lower for males than for females at all education levels. Table 4.1 also displays a non-linear relationship, for both females and males, between labour force participation and age: as age increases the probability of participating in the labour force increases, but at a decreasing rate, suggestive of an inverted U-shaped relationship.

Females who are married or widowed are less likely to be active in the labour force in comparison to females who have never married, while females who are divorced and men in all marital status categories are more likely to participate in the labour force in comparison to those who never married. Moreover, when looking at household characteristics, having more children within the household specifically, and larger households in general, affect both females' and males' labour force participation negatively. However, having more employed members in the household raises both females' and males' labour force participation, perhaps through providing motivation and communicating knowledge of job opportunities.

The findings thus far show that black African immigrants in South Africa aged 15-65 have a higher probability of participating in the labour force in comparison to citizens, which is similar to other South African studies' findings (Peters and Sudaram, 2015; Zuberi and Sibanda, 2004). Furthermore, the ability to speak English as either a first or second language – without accounting for citizenship - positively affects individual's labour market participation.

To address specifically the role played by the English language in influencing the probability of finding employment, the estimating sample was restricted to black African labour force participants of working age. Restricting the sample to labour market participants only may result in sample selection bias, however. Therefore, the results shown in Table 4.2, which presents the employment probability for females and males on the basis of language and other characteristics, should be interpreted as conditional on someone having entered the labour market.

Similar to other South African studies, Table 4.2 shows that both male and female immigrants are more likely to be employed in comparison to citizens (Peters and Sundaram, 2015; Zuberi and Sibanda, 2004). Model I also shows the positive impact of speaking English on employment for both immigrants and citizens irrespective of gender.

Table 4.2 Probit models for Employment

	Females			Males		
	I	II	III	I	II	III
Immigrant	0.049*** (0.003)	0.051*** (0.006)	0.087*** (0.007)	0.152*** (0.002)	0.167*** (0.003)	0.131*** (0.004)
<i>Language effects for all</i>						
English first, no second language	0.111*** (0.005)			0.072*** (0.005)		
English first, other second language	0.201*** (0.004)			0.144*** (0.003)		
Other first, English second language	0.097*** (0.001)			0.068*** (0.001)		
Other first, other second language	0.055*** (0.002)			0.069*** (0.001)		
<i>Language effects for citizens only</i>						
English first, no second language		0.104*** (0.006)	0.071*** (0.006)		0.067*** (0.005)	0.032*** (0.006)
English first, other second language		0.204*** (0.004)	0.124*** (0.005)		0.148*** (0.004)	0.074*** (0.005)
Other first, English second language		0.096*** (0.002)	0.056*** (0.002)		0.067*** (0.001)	0.032*** (0.001)
Other first, other second language		0.058*** (0.002)	0.028*** (0.002)		0.073*** (0.002)	0.037*** (0.002)
<i>Language effects for immigrants only</i>						
English first, no second language		0.180*** (0.017)	0.088*** (0.022)		0.097*** (0.013)	0.033** (0.016)
English first, other second language		0.187*** (0.010)	0.092*** (0.013)		0.118*** (0.008)	0.049*** (0.010)
Other first, English second language		0.121*** (0.008)	0.079*** (0.009)		0.071*** (0.006)	0.027*** (0.006)
Other first, other second language		0.015 (0.009)	0.023** (0.010)		0.011* (0.006)	0.009 (0.006)
Immigrants' duration of residence			0.000 (0.000)			0.000*** (0.000)

Table 4.2 Continued

	Females			Males		
	I	II	III	I	II	III
Primary education			-0.002 (0.004)			-0.016*** (0.003)
Incomplete secondary education			0.001 (0.003)			-0.011*** (0.003)
Matric			0.100*** (0.003)			0.057*** (0.003)
Tertiary education			0.283*** (0.003)			0.158*** (0.003)
Age			0.037*** (0.000)			0.028*** (0.000)
Age squared			-0.000*** (0.000)			-0.000*** (0.000)
Married			-0.058*** (0.002)			0.180*** (0.001)
Widower/widow			0.015*** (0.004)			0.101*** (0.005)
Divorced			0.063*** (0.004)			0.031*** (0.005)
Household size			-0.055*** (0.001)			-0.065*** (0.000)
No. of young children in the household			-0.037*** (0.001)			-0.007*** (0.001)
No. of other employed in the household			0.150*** (0.001)			0.136*** (0.001)
Sample	562659	562659	559751	609405	609405	605823
Pseudo R ²	0.0082	0.0083	0.1455	0.0145	0.0148	0.1526

Notes: Sample is restricted to the black African aged 15-65, who participate in the labour market (using the narrow definition). The table results are weighted, with standard errors in parentheses, ***p<0.01, ** p<0.05,

* p<0.1 for significance level.

Source: Authors' calculations based on Census 2011.

Reference categories: citizens, other first language and no second language, no schooling and never married.

For citizens, the disaggregated language effects in Model **II** differ only marginally from the aggregate effects for the full sample in Model **I**. Bilinguals with English as a first language are the most likely group of citizens to be employed. Bilingual citizens with English as a second language benefit almost as much as monolingual English speakers, and the language effects are mostly larger for female than male citizens. Among immigrants, Model **II** shows that first language English speakers – monolingual or bilingual – are the most likely to be employed relative to the other language groups. However, similarly to citizens, the language effects are larger for female than male immigrants. In Model **III**, when controlling for other characteristics, the effect of language on the probability of employment declines for all immigrants and citizens, but remains positive and significant.

When comparing the language effects between citizens and immigrants in the full specification, the table shows that most English language categories are more influential among female immigrants than among female citizens. Comparing immigrants to citizens within a language group, female immigrants are at the greatest employment advantage when they speak English as a second language. These female immigrants have a 2.3 percentage point greater marginal effect of language than female citizens (7.9 percent compared to 5.6 percent), in addition to the base immigrant effect of 8.7 percent. For males, the base immigrant effect is larger than for females, but the language benefits are smaller.

Male immigrants are at the greatest employment advantage relative to citizens when they are monolingual English speakers or second language English speakers. The higher employment probability that immigrants gain by speaking English, the dominant language of business of the country, is similar to what other studies have found (Dustmann and Fabbri, 2003; Lindemann and Kogan, 2013; Paolo and Raymond, 2012). Speaking English in South Africa enables immigrants to communicate with citizens, be exposed to various employment opportunities, and understand and engage with employers' requirements. The likelihood of missing these opportunities and benefits increases when immigrants are not able to speak English.

The importance of speaking English for gaining employment stems from the fact that most job sectors (formal sector, and private households) and industries (community, social and personal services, and wholesale and retail) where immigrants may find employment require a minimum level of English proficiency (Lindemann and Kogan, 2013). In addition, the larger language effects for women than for men, regardless of immigrant status, speak to the types of jobs that women typically access which are often in the service sector.

As mentioned earlier in Section 2, Table 4.2 uses the same set of control variables as in the previous table. The remainder of the table further shows that higher levels of education improve the individual's chances of being employed at the matric and/or tertiary education level. Age, marital status, immigrants' duration of residency and the selected household characteristics follow similar patterns as in Table 4.1. However, Table 4.2 shows that having more employed members in the household has a larger effect on the individual's probability to be employed than to participate. This may be the result of the greater exposure to employment opportunities and greater networks (Evans, 1987).

Thus far, immigrants have been shown to be in an advantageous position in the South African labour market relative to citizens. The findings reveal that immigrants and citizens who speak English – conditional on other factors – have significantly higher probabilities of working in the South African labour market. Immigrants are more likely to be economically active and obtain employment than citizens in the same English language group. This finding may partly reflect that immigrants have a higher propensity to accept any job opportunity offered to them independent of pay considerations. The composition of the English language group (i.e. English as a first/second language) affects the participation and employment probabilities differently by gender. Both female and male immigrants have the greatest participation advantage over citizens when they are second language English speakers, but for employment, male immigrants benefit most if they are monolingual English speakers while females benefit most as second language English speakers relative to citizens. In general, language plays a stronger role in providing access to the labour market for women than for men, particularly for employment. This suggests that there are differences in the importance of language to the types of work that women and men access.

With the aim of investigating the last labour market channel, Table 4.3 estimates the level of income for employed individuals using an interval regression model. This table includes a fourth model specification, in which employment characteristics (sector and industry) are added to the previous control variables, as they are expected to affect the level of income earned by individuals. As was the case previously, these results should be interpreted conditional on participation and finding employment. In contrast to the previous two tables, immigrants perform worse than citizens in the first model specification. The average immigrant income penalty is very large for females, but much smaller for males, which confirms the low average annual earnings that immigrants enjoy relative to citizens as displayed in Table 3.1. This suggests that part of the explanation for immigrants having a greater propensity to be employed than citizens, as shown in the previous table, is that immigrants may be more willing to accept low-paying jobs than citizens.

Appendix **B** presents descriptive statistics including the sector in which female and male citizens and immigrants work, illustrating the different quality of jobs that these groups access. Citizens of both genders are much more likely than immigrants to work in the formal sector, with a higher proportion of men than women having formal sector jobs. Immigrants are more likely than citizens to work in the informal sector, which is especially low paying for women. Formal sector jobs are less precarious than other sectors, and more likely to be associated with non-wage benefits in addition to higher wages, indicating that citizens on average work in better quality jobs than immigrants.

However, Table **4.3** shows that after disaggregating the language effects and controlling for personal and job characteristics, immigrants earn more than citizens in the base language group (the immigrant status effect), *ceteris paribus*. This implies that, within a given job category and education level, immigrants in the language reference group outperform citizens. This finding of an income premium to immigrants is in contrast to many international studies (Bell, 1997; Chiswick, 1978).

In Model **I**, speaking any language combination is positively associated with income from employment, for females and males, compared to monolingual non-English and after controlling for immigrant status. The largest premium accrues to bilinguals with English as a first language. The role of language for citizens remains similar when disaggregating the language effects by citizenship status. In the second model, for male immigrants, the language effects are smaller than for male citizens, but remain positive and significant. For female immigrants, all English language groups receive an income premium, which is largest for monolinguals. In Models **III** and **IV**, the language effects decline for all groups as other characteristics that explain income (and are correlated with language) are included. For immigrants, the largest income premium is earned by second language English speakers, with females benefitting more from language than males in both citizenship categories.

Table 4.3 Interval Regressions for Annual Earnings

	Females				Males			
	I	II	III	IV	I	II	III	IV
Immigrant	-0.506*** (0.029)	-0.419*** (0.065)	0.287*** (0.067)	0.283*** (0.067)	-0.038*** (0.013)	0.095*** (0.028)	0.435*** (0.028)	0.451*** (0.029)
<i>Language effects for all</i>								
English first, no second language	0.822*** (0.060)				0.642*** (0.048)			
English first, other second language	1.364*** (0.039)				1.028*** (0.031)			
Other first, English second language	0.916*** (0.015)				0.616*** (0.012)			
Other first, other second language	0.289*** (0.018)				0.270*** (0.013)			
<i>Language effects for citizens only</i>								
English first, no second language		0.744*** (0.064)	0.299*** (0.061)	0.309*** (0.061)		0.637*** (0.056)	0.287*** (0.052)	0.295*** (0.053)
English first, other second language		1.416*** (0.043)	0.653*** (0.041)	0.656*** (0.041)		1.087*** (0.040)	0.477*** (0.037)	0.473*** (0.038)
Other first, English second language		0.914*** (0.015)	0.463*** (0.015)	0.464*** (0.015)		0.624*** (0.013)	0.323*** (0.012)	0.326*** (0.012)
Other first, other second language		0.302*** (0.018)	0.243*** (0.017)	0.240*** (0.017)		0.290*** (0.014)	0.196*** (0.013)	0.194*** (0.013)
<i>Language effects for immigrants only</i>								
English first, no second language		1.362*** (0.163)	0.259 (0.162)	0.278* (0.162)		0.568*** (0.088)	-0.168* (0.087)	-0.151* (0.088)
English first, other second language		1.088*** (0.107)	0.081 (0.106)	0.077 (0.107)		0.814*** (0.049)	-0.045 (0.049)	0.005 (0.050)
Other first, English second language		0.906*** (0.076)	0.379*** (0.077)	0.373*** (0.078)		0.501*** (0.033)	0.075** (0.033)	0.105*** (0.034)
Other first, other second language		0.014 (0.087)	0.097 (0.088)	0.074 (0.089)		0.068** (0.034)	0.030 (0.034)	0.034 (0.035)
Immigrants' duration of residence			0.000 (0.000)	0.000 (0.000)			0.000* (0.000)	0.000* (0.000)

Table 4.3 Continued

	Females				Males			
	I	II	III	IV	I	II	III	IV
Primary education			0.198*** (0.035)	0.204*** (0.035)			0.085*** (0.024)	0.080*** (0.024)
Incomplete secondary education			0.760*** (0.033)	0.777*** (0.033)			0.551*** (0.023)	0.551*** (0.023)
Matric			1.778*** (0.033)	1.814*** (0.034)			1.187*** (0.023)	1.187*** (0.024)
Tertiary education			3.305*** (0.033)	3.357*** (0.034)			2.337*** (0.025)	2.340*** (0.026)
Age			0.273*** (0.004)	0.273*** (0.004)			0.209*** (0.004)	0.211*** (0.004)
Age squared			-0.003*** (0.000)	-0.003*** (0.000)			-0.002*** (0.000)	-0.002*** (0.000)
Married			-0.080*** (0.013)	-0.073*** (0.013)			0.692*** (0.011)	0.688*** (0.011)
Widower/widow			0.096*** (0.026)	0.107*** (0.027)			0.419*** (0.045)	0.420*** (0.046)
Divorced			0.203*** (0.028)	0.205*** (0.028)			0.076* (0.042)	0.076* (0.042)
Household size			-0.081*** (0.005)	-0.079*** (0.005)			-0.135*** (0.005)	-0.138*** (0.005)
No. of young children in the household			-0.251*** (0.008)	-0.248*** (0.008)			-0.142*** (0.007)	-0.139*** (0.007)
No. of other employed in the household			-0.125*** (0.010)	-0.126*** (0.010)			-0.063*** (0.008)	-0.064*** (0.008)
Formal sector				-0.195*** (0.016)				0.179*** (0.014)
Informal sector				-0.284*** (0.022)				-0.017 (0.018)
Agriculture, forestry and fishing				0.269*** (0.030)				0.216*** (0.020)
Mining				0.225*** (0.049)				0.637*** (0.021)
Manufacturing				0.034 (0.026)				0.148*** (0.017)

Table 4.3 Continued

	Females				Males			
	I	II	III	IV	I	II	III	IV
Infrastructure				-0.175*** (0.031)				-0.030* (0.017)
Wholesale and retail				0.021 (0.016)				-0.007 (0.014)
Financial, insurance, real estate and business				0.088*** (0.020)				0.053*** (0.016)
Sample	317254	317254	315415	309313	407594	407594	404947	394284
Wald χ^2	4636.61	4679.77	52569.48	52860.04	3370.43	3443.24	44220.84	46323.30

Notes: Sample is restricted to employed black Africans aged 15-65. The table results are weighted, with standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations based on Census 2011.

Reference categories: citizens, other first language and no second language, no schooling, never married, private household sector and community, social and personal services industry.

Analysing the effect of the various English language groups, Table 4.3 shows that the key language group – irrespective of gender – that is beneficial for immigrants’ earnings is English being a second language. The findings in the third and fourth specifications for both genders and irrespective of citizenship show not only the importance of English in influencing the individual’s level of earnings, but also indicate that English is more influential for earnings than for labour market participation and employment, as also found by Posel and Casale (2011).

Although, in the initial specification, immigrants in the base language group earn less than citizens, the later models show that this can be attributed to immigrants having characteristics that are less highly rewarded in the labour market, and working in poorer quality jobs, than citizens. There are several plausible reasons for why immigrants may work in low quality jobs. First, they may have less access than citizens to sources of financial support outside the labour market, such as government grants and assistance from extended family members. They may therefore have lower reservation wages than citizens and be willing to accept lower wage offers. Second, employers may prefer to hire citizens for high quality jobs, due to employment equity requirements that prioritise South African candidates or due to immigrant difficulties in securing work permits, but may be willing to employ immigrants in low paying jobs. Third, most of the employed in South Africa find work through social networks (Posel *et al*, 2014). The networks that immigrants access may contain information about poorer-quality jobs than the networks that citizens access.

After controlling for individual characteristics and job type, immigrants in the language reference group earn a premium, which has the effect of raising the incomes of all immigrants relative to citizens. For male immigrants, this premium is partially offset because they are less rewarded for language than citizens. Female immigrants who are monolingual or second language English speakers receive the same language benefits as citizens¹², and thus retain their full immigrant income premium. Female immigrants who are first language English speakers (monolingual) or second language English speakers therefore earn more than female citizens in the analogous English category. Among the English language categories, male immigrants earn more than male citizens.

For both gender regressions, the results show that age, marital status and immigrants’ duration of residence follow similar patterns as in the previous tables. Higher levels of education are associated with higher incomes, across all education levels. Moreover, education impacts the level of earnings for females more than males.

¹² The difference between the language effects is not statistically significant for these groups.

Like in the previous tables, income is lower in larger households with more children within the household. However, having more employed members within the household also negatively affects the individual's level of earnings. On average, females who are employed within the formal and informal sectors earn less relative to females who are employed within private households. This may be in part because of the existence of a minimum wage for domestic work, but the private household sector also includes jobs at any occupational level that are located in private households. Moreover, the table indicates that males who work within the formal sector tend to have higher earnings, and those in the informal sector have lower earnings, compared to others who work within private households.

Females who are employed within all industries, except for infrastructure, earn more in comparison to females who are employed within the community, social and personal services industry, which is the modal industry for women (see Appendix **B**). The pattern of the relationship between industry and the level of earnings is similar between genders, although the magnitudes of the industry wage differentials differ. The findings for the control variables in the earnings model are typically in line with the results of such models for South Africa.

5 Discussion of Immigrant-Citizen Comparisons

The tables presented thus far have illustrated differences in the role of language for immigrants and for citizens. However, the results showed the immigrant effect and the language effect separately. The results for the overall immigrant-citizen comparison, comprising both the language effect and the immigrant status effect, are summarised in Table **5.1**. The table shows the marginal effect of immigrant status on each labour market outcome, for individuals in each language group, using the full model specification. For labour force participation and employment, the values show the difference in the predicted probability of the outcome between immigrants and citizens in the given language group, while for income, the values show the proportional difference in predicted income between immigrants and citizens in the given language group.

Immigrants significantly outperform citizens in terms of access to the labour market within every language group after controlling for all the variables presented in Section 4, as illustrated in Table **5.1**. Among the English language categories, speaking English as a second language is most beneficial in raising immigrants' participation and employment compared to citizens. For example, bilingual female immigrants who speak English as their second language are 9.2 percent more likely to be employed than female citizens in the same language category, *ceteris paribus*.

Table 5.1 Difference in Predicted Outcome between Immigrants and Citizens

	Females			Males		
	LFP	EMP	INC	LFP	EMP	INC
English first, no second language	0.045***	0.086***	0.286	0.115***	0.119***	0.004
English first, other second language	0.011	0.042***	-0.256***	0.060***	0.090***	-0.017
Other first, English second language	0.083***	0.092***	0.211***	0.126***	0.115***	0.257***
Other first, other second language	0.097***	0.071***	0.124*	0.166***	0.098***	0.337***
Other first, no second language	0.145***	0.077***	0.327***	0.215***	0.125***	0.569***
Sample	1151294	559751	309313	1046476	605823	394284

Notes: Sample is restricted to employed black Africans aged 15-65. The values displayed are the marginal effects of immigrant status within each language group calculated using the Stata “margins” command. The final model specification was used for each outcome. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors’ calculations based on Census 2011.

Among the English language categories, only immigrants who have English as a second language earn significantly more than similar citizens, by more than 20 percent. English monolingual immigrants – irrespective of gender - are more likely to participate and be employed relative to citizens, however, their earnings are not affected by being in that language group.

Immigrants also outperform citizens in the non-English language categories, across all labour market outcomes. Immigrants, especially those from neighbouring countries, may speak official South African languages other than English¹³, hence they may access employment opportunities in those languages. In addition, those who speak two non-English languages may speak English as a third language, but this information is not collected in the Census data.

As presented in Section 4, the role of language in influencing labour market success was shown to be larger for women than for men. This may be the result of the different sectors and industries that women occupy, see Appendix B. They mostly work within the “wholesale and retail” and service industries, such as domestic work, child care and sales persons, where these type of jobs require good language skills. On the other hand, the type of jobs that men occupy require relatively less language skills and more physical skills, such as logistics work, inventory monitoring and loading and offloading of goods, for example.

Although the language benefits were higher for female immigrants than male immigrants, Table 5.1 shows that once incorporating the overall advantage that immigrants experience in South Africa’s labour market, male immigrants tend to be more successful in the labour market compared than male citizens and female immigrants more than female citizens. Male immigrants were also relatively more successful than female immigrants.

¹³ For example, Zimbabwe shares six official languages with South Africa.

The study further shows that among the English language groups, English as a second language is most influential for immigrants' labour market participation, employment and earnings.¹⁴ Monolingual English speakers followed by bilinguals who speak English as a second language are also more successful than citizens in the labour market. The study further shows that immigrants outperform citizens even when both groups do not speak English either as a first or a second language. This may reflect that immigrants have greater proficiency than citizens do in a third language, although it is not possible to explore such multilingualism using Census, which collects data on two spoken languages. However, this effect may also reflect other limitations of this study. If immigrants are a positively selected sample from their home countries, they may outperform citizens at any given set of observable characteristics, although Vermaak and Muller (2019) find limited evidence of positive selection. Rather, foreigners may be pushed to migrate to South Africa, which suggest that their labour market performance may be under-estimated. Immigrants may also self-select into employment at a greater rate than citizens do, by being more willing to accept low paying jobs, and therefore have better participation and employment outcomes than citizens across all language groups. This study shows how language helps to explain why some immigrants perform better than others, but it cannot completely account for South Africa's unusual phenomenon of immigrants being more successful in the labour market than citizens.

The Census is the only data source with the required information and sufficiently large sample to analyse the role of language and immigrant status in South Africa's labour market. However, this paper acknowledges a number of limitations that arise from the analysis of these data. The proficiency aspects of language are not analysed in the study as a result of how the language data were collected, where individuals were only asked to specify the two languages they speak most often. The results are potentially affected by endogeneity regarding the language variable, in addition to the immigrants' selectivity issue. Furthermore, sample selection bias may occur as the sample analysis is narrowed to labour force participants and further to employed individuals. However, due to the limited number of variables available in the Census data it is impossible to identify appropriate instruments to instrument for the endogeneity of immigrant status or find exclusion restrictions that separate the outcome equations from the selection equations in order to account for selection bias. As it is not possible to address these potential issues. Therefore, the

¹⁴ Speaking English as a first language is uncommon in the African race group, as shown in Appendices **A** and **B**. For immigrants, being first language English speakers is strongly associated with high levels of education. Therefore, the independent labour market benefit of English as a first language for immigrants may be smaller than for second language speakers, after isolating the effect of education.

employment and income results should be interpreted as conditional on labour market participation and employment respectively.

6 Conclusion

The main objective of this study was to analyse the effect of language on labour market success in a comparative framework between immigrants and citizens in South Africa, using the 10 percent sample of the 2011 Census. The existing literature for South Africa explores either labour market outcomes between immigrants and citizens, or the role of language proficiency in the labour market, but no study has combined the two areas of analysis. In addition, this study includes the female population, whereas in previous national studies the main focus was on the working age male population. Therefore, the paper highlights gender comparisons between immigrants and citizens regarding their labour force participation, employment and earnings.

In conclusion the study reveals South Africa to be, in part, an attractive place for immigrants regarding labour market outcomes. It also highlights the importance of language in influencing labour market outcomes, and particularly the role of English. In a country like South Africa, where English is effectively the language of business, immigrants who speak English as either a first or second language are advantaged in the labour market, both compared to immigrants without English and compared to citizens in the same language group. Even after accounting for this differential language effect, immigrants retain a general advantage over citizens in the South African labour market. Therefore, as language in this study justifies partially the success of immigrants in South Africa, it cannot fully explain their extraordinary labour market performance in the country.

This paper has shown that one of the reasons why immigrants are more successful than citizens in the labour market may be how their ability to speak English is valued. African immigrants are more likely than citizens to speak English as a first language, and immigrants in this language group are more likely to participate and be employed than citizens. However, they do not earn more after controlling for their individual and job characteristics. The largest benefit of speaking English is as a second language: immigrants in this category are more successful than citizens across all three of the labour market outcomes, even though second language English is equally common for immigrants and citizens. Speaking English as a second language may be more highly valued for immigrants than for citizens, perhaps due to unmeasured proficiency. This confirms other South African studies, without an immigration focus, that have emphasised the importance of speaking English proficiently in the South African labour market (Cornwell and Inder, 2008; Deumert *et al*,

2005; McKenzie and Muller, 2015; Posel and Casale, 2011). The study suggests possible implications for the South African language policies. The move towards bilingualism in South Africa, through supporting African mother tongue education has been highly encouraged as it influences English acquisition skills. However, alongside the policies that support and develop African languages, the development of the English language ability, fluency and literacy among citizens needs to be maintained and strengthened. This is highly recommended in order to ensure that future labour market entrants are not disadvantaged in the South African labour market.

References

- Bell, B.D. (1997). 'The performance of immigrants in the United Kingdom: Evidence from the GHS'. *The Economic Journal*, 107(3): 33-44.
- Borjas, G.J. (1994). 'The Economics of Immigration'. *Journal of Economic Literature*, 32(4): 1667-1717.
- Chiswick, B.R. (1978). 'The Effect of Americanization on the Earnings of Foreign-born Men'. *Journal of Political Economy*, 86(5): 897-921.
- Chiswick, B.R. and P.W. Miller (2002). 'Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle'. *Journal of Population Economics*, 15(1): 31-57.
- Clark, K. and J. Lindley (2009). 'Immigrant Assimilation Pre and Post Labour Market Entry: Evidence from the UK Labour Force Survey'. *Journal of Population Economics*, 22 (1): 175-98.
- Cornwell, K. and B. Inder (2008). 'Language and Labour Markets in South Africa'. *Journal of African Economies*, 17(3): 490-525.
- Deumert, A., B. Inder and P. Maitra (2005). 'Language, Informal Networks and Social Protection'. *Global Social Policy*, 5(3): 303-28.
- Dustmann, C. and F. Frabbi (2003). 'Language Proficiency and Labour Market Performance of Immigrants in the UK'. *The Economic Journal*, 113(489): 695-717.
- Euwals, R., J. Dagevos, M. Gijsberts and H. Roodenburg (2010). 'Citizenship and Labor Market Position: Turkish Immigrants in Germany and the Netherlands'. *The International Migration Review*, 44(3): 513-38.

- Evans, M.D.R. (1987). 'Language Skill, Language Usage and Opportunity: Immigrants in the Australian Labour Market'. *Sociology*, 21(2): 253-74.
- Facchini, G., A.M. Mayda and M. Mendola (2013). 'What Drives Individual Attitudes towards Immigration in South Africa?'. *Review of International Economics*, 21(2): 326–41.
- Hainmueller, J. and M.J. Hiscox (2007). 'Educated Preferences: Explaining Individual Attitudes toward Immigration in Europe'. *International Organization*, 61: 399-442.
- Hayfron, J.E. (2001). 'Language training, language proficiency and earnings of immigrants in Norway'. *Applied Economics*, 33(15): 1971-79.
- Lindemann, K. and I. Kogan (2013). 'The Role of Language Resources in Labour Market Entry: Comparing Estonia and Ukraine'. *Journal of Ethnic and Migration Studies*, 39(1): 105-123.
- Marquez, S.A. (1997). 'Race, Class and Gender: Reformulating the Push and Pull Factors Explanation of Hispanic Immigration'. *Race, Gender & Class journal*, 4(2): 45-55.
- Mayda, A.M. (2006). 'Who is Against Immigration? A Cross-country Investigation of Individual Attitudes toward Immigrants'. *Review of Economics and Statistics*, 88(3): 510–30.
- McKenzie, T. and C. Muller (2017). 'English language proficiency and labour market participation in South Africa'. *Interdisciplinary Journal of Economics and Business Law*, 6(2): 49-76.
- Paolo, A.D. and J.L. Raymond (2012). 'Language Knowledge and Earnings in Catalonia'. *Journal of Applied Economics*, 15(1): 89-118.
- Peters, A.C. and A. Sundaram (2015). 'Country of Origin and Employment Prospects among Immigrants: An Analysis of South–South and North–South migrants to South Africa'. *Applied Economics Letters*, 22(17): 1415– 18.
- Posel, D. and D. Casale (2011). 'Language Proficiency and Language Policy in South Africa: Findings from New Data'. *International Journal of Educational Development*, 31(5): 449–57.
- Posel, D., D. Casale and C. Vermaak, (2014). 'Job Search and the Measurement of Unemployment in South Africa'. *South African Journal of Economics*, 82(1): 66-80.
- Shields, M.A and S.W. Price (2002). 'The English Language Fluency and Occupational Success of Ethnic Minority Immigrant Men Living in English Metropolitan Areas'. *Journal of Population Economics*, 15(1): 137-60.

Trebilcock, M.J. (2003). 'The Law and Economics of Immigration Policy'. *American Law and Economics Review*, 5(2): 271-317.

United Nations, *International migration and development*, [website] (2018), https://digitallibrary.un.org/record/1640258/files/A_73_286-EN.pdf, (Accessed 5 February 2019).

Vermaak, C. and C. Muller (2019). 'Do immigrants have better labour market outcomes than South Africans?'. *Development Southern Africa*, 1-21.

Zuberi, T. and A. Sibanda (2004). 'How Do Migrants Fare in a Post-Apartheid South African Labour Market?'. *International Migration Review*, 38(4): 1462-91.

Appendix A - Characteristics of Labour Force Participants

Table A1: The Mean of the Explanatory Variables for Labour Force Participants

	Male citizen	Male immigrant	Female citizen	Female immigrant
English first, no second language	0.011 (0.000)	0.026*** (0.001)	0.012 (0.000)	0.028*** (0.001)
English first, other second language	0.017 (0.000)	0.091*** (0.001)	0.019 (0.000)	0.099*** (0.002)
Other first, English second language	0.359 (0.001)	0.340*** (0.002)	0.362 (0.001)	0.380*** (0.003)
Other first, other second language	0.234 (0.001)	0.308*** (0.002)	0.213 (0.001)	0.258*** (0.003)
Other first, no second language	0.379 (0.001)	0.235*** (0.002)	0.394 (0.001)	0.234*** (0.003)
Age	34.729 (0.015)	31.527*** (0.039)	34.880 (0.015)	30.449*** (0.054)
No education	0.048 (0.000)	0.076*** (0.001)	0.048 (0.000)	0.061*** (0.002)
Primary education	0.157 (0.000)	0.199*** (0.002)	0.130 (0.000)	0.158 (0.002)
Incomplete secondary education	0.387 (0.001)	0.436*** (0.002)	0.373 (0.001)	0.474*** (0.003)
Matric	0.319 (0.001)	0.202*** (0.002)	0.325 (0.001)	0.210*** (0.003)
Tertiary education	0.089 (0.000)	0.087*** (0.001)	0.124 (0.000)	0.097*** (0.002)
Married	0.458 (0.001)	0.562*** (0.002)	0.406 (0.001)	0.622*** (0.003)
Widower/widow	0.009 (0.000)	0.004*** (0.000)	0.039 (0.000)	0.025*** (0.001)
Divorced	0.014 (0.000)	0.008*** (0.000)	0.028 (0.000)	0.022*** (0.001)
Never married	0.519 (0.001)	0.426*** (0.002)	0.526 (0.001)	0.330*** (0.003)
Household size	2.801 (0.002)	2.030*** (0.006)	2.986 (0.002)	2.089*** (0.008)
No. of young children (age<7) in the household	0.547 (0.001)	0.307*** (0.003)	0.826 (0.001)	0.518*** (0.005)
No. of other employed members in the household	0.610 (0.001)	0.653*** (0.005)	0.689 (0.001)	0.763*** (0.006)
Employment status official	0.682 (0.001)	0.846*** (0.002)	0.585 (0.001)	0.656*** (0.003)
Sample size	553,392	508,35	534,347	23,985
Population size	6,674,832	623,384	6,255,088	286,625

Notes: Sample is restricted to labour force participants who are black Africans aged 15-65. The table results are weighted, robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1 for significance test “Male immigrant subject to Male citizen” and “Female immigrant subject to Female citizen”.

Source: Authors’ calculations based on Census 2011.

Appendix B - Characteristics of the Employed

Table B1: The Mean of the Explanatory Variables for the Employed

	Male citizen	Male immigrant	Female citizen	Female immigrant
English first, no second language	0.011 (0.000)	0.027*** (0.001)	0.013 (0.000)	0.032*** (0.001)
English first, other second language	0.020 (0.000)	0.098*** (0.001)	0.024 (0.000)	0.118*** (0.003)
Other first, English second language	0.372 (0.001)	0.348*** (0.002)	0.390 (0.001)	0.406*** (0.004)
Other first, other second language	0.246 (0.001)	0.303*** (0.002)	0.216 (0.001)	0.237* (0.004)
Other first, no second language	0.351 (0.001)	0.224*** (0.002)	0.357 (0.001)	0.206*** (0.003)
Age	36.514 (0.018)	31.888*** (0.044)	37.492 (0.020)	31.495*** (0.070)
No education	0.050 (0.000)	0.072*** (0.001)	0.049 (0.000)	0.057*** (0.002)
Primary education	0.153 (0.001)	0.196*** (0.002)	0.129 (0.001)	0.144*** (0.003)
Incomplete secondary education	0.353 (0.001)	0.430*** (0.003)	0.314 (0.001)	0.455*** (0.004)
Matric	0.324 (0.001)	0.204*** (0.002)	0.318 (0.001)	0.217*** (0.003)
Tertiary education	0.120 (0.001)	0.098*** (0.001)	0.189 (0.001)	0.127*** (0.003)
Married	0.554 (0.001)	0.584*** (0.002)	0.434 (0.001)	0.570*** (0.004)
Widower/widow	0.010 (0.000)	0.004*** (0.000)	0.050 (0.000)	0.033*** (0.001)
Divorced	0.015 (0.000)	0.008*** (0.000)	0.037 (0.000)	0.029*** (0.001)
Never married	0.421 (0.001)	0.404*** (0.002)	0.479 (0.001)	0.368*** (0.004)
Household size	2.577 (0.003)	1.985*** (0.007)	2.806 (0.003)	2.035*** (0.011)
No. of young children (age<7) in the household	0.507 (0.001)	0.298*** (0.003)	0.684 (0.002)	0.439*** (0.006)
No. of other employed members in the household	0.655 (0.002)	0.698*** (0.006)	0.771 (0.002)	0.822*** (0.008)
Formal sector	0.744 (0.001)	0.593*** (0.002)	0.720 (0.001)	0.550*** (0.004)
Informal sector	0.143 (0.001)	0.221*** (0.002)	0.128 (0.001)	0.180*** (0.003)
Private households	0.113 (0.001)	0.186*** (0.002)	0.152 (0.001)	0.271*** (0.004)
Agriculture, forestry and fishing	0.062 (0.000)	0.078*** (0.001)	0.044 (0.000)	0.074*** (0.002)
Mining	0.050 (0.000)	0.053** (0.001)	0.017 (0.000)	0.013*** (0.001)
Manufacturing	0.111 (0.001)	0.095*** (0.001)	0.065 (0.000)	0.057*** (0.002)
Infrastructure	0.127 (0.001)	0.166*** (0.002)	0.048 (0.000)	0.036*** (0.002)

Table B1: Continued

	Male citizen	Male immigrant	Female citizen	Female immigrant
Wholesale and retail	0.237 (0.001)	0.251*** (0.002)	0.204 (0.001)	0.213** (0.003)
Financial intermediation, insurance, real estate and business	0.145 (0.001)	0.136*** (0.002)	0.118 (0.001)	0.106*** (0.003)
Community, social and personal services	0.267 (0.001)	0.221*** (0.002)	0.504 (0.001)	0.500 (0.004)
Annual income (among the employed)	68,661 (270.712)	48,425*** (587.441)	59,419 (266.759)	37,700*** (910.220)
Sample size	353,736	40,086	294,406	14,480
Population size	4,286,613	493,843	3,454,914	174,465

Notes: Sample is restricted to employed individuals who are black Africans aged 15-65. The table results are weighted, robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ for significance test “Male immigrant subject to Male citizen” and “Female immigrant subject to Female citizen”.

Source: Authors’ calculations based on Census 2011.