

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Globalization and Poverty

Volume Author/Editor: Ann Harrison, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-31794-3

Volume URL: <http://www.nber.org/books/harr06-1>

Conference Date: September 10-12, 2004

Publication Date: March 2007

Title: Globalization and the Returns to Speaking English in South Africa

Author: James Levinsohn

URL: <http://www.nber.org/chapters/c10714>

---

# Globalization and the Returns to Speaking English in South Africa

James Levinsohn

---

## 15.1 Introduction

The literature on globalization and wages is, by the standards of economics, huge. It is a literature that compensates for its volume by offering precious little in the way of convincing results. This is not (usually) the fault of the researchers. Rather, it is just very difficult to identify the role of international trade and/or investment on wages relative to the multitude of other factors that influence wages (and which frequently occur simultaneously with globalization). This has led researchers to debate, for example, whether trade explains a growing wage gap between high-wage and low-wage earners or whether the real determinant of increasing wage disparity is coincident skill-biased technical change. Yet others (correctly) claim that even this dichotomy is a false one since international trade and investment and skill-biased technical change are themselves codetermined.

With this cacophony as background, this paper steps back and experiments with a very different approach to investigating the impact of globalization on wages. Noting the special circumstances around South Africa's emergence from the apartheid era (and the relatively closed economy that accompanied the apartheid era), this paper asks whether the return to speaking English (measured in a narrow way) increased as the South

James Levinsohn is the J. Ira and Nicki Harris Family Professor of Public Policy, a professor of economics and associate dean of the Gerald R. Ford School of Public Policy, University of Michigan, and a research associate of the National Bureau of Economic Research (NBER).

I would like to thank without implicating Raquel Fernandez, Ann Harrison, Mark Rosenzweig, and Duncan Thomas as well as participants at the NBER conference on globalization and poverty. Thanks to Nzinga Broussard for research assistance.

African economy embarked upon its integration with the rest of the industrialized world.

There is a certain logic to trying to measure the impact of globalization on wages in this manner. Following the advent of democracy in South Africa in 1994, there were several huge changes in the economy, many of which might be expected to change wages. One, but only one, of these changes was South Africa's reintegration with the global economy. Others included legislated changes in the labor market (with an emphasis on affirmative action) and the outbreak of the HIV/AIDS pandemic. Decomposing the changes in South African wages into those fractions due to increased disease, the dismantling of apartheid and ensuring affirmative action, changes in technology during the 1990s, and increased integration into the global economy is a Herculean (or outright impossible) task. Measuring changes in the return to speaking English is a simple task, and I argue below that it is one that at least stands a chance of shedding light on the impact of globalization on wages in South Africa.

The underlying idea is that as South Africa reintegrated with the rest of the world, the return to speaking an international language of commerce might plausibly increase. In South Africa, English is that language. (The other widely spoken languages, such as Zulu and Afrikaans, are not used much in international commerce.) It is less obvious why some of the other changes concurrent with the fall of apartheid should change the return to speaking English. It is, for example, unclear why AIDS should have much of an impact on the returns to speaking English (although it almost surely impacts wages).<sup>1</sup> Nor is it clear why the sort of skill-biased technological change that occurred worldwide in the 1990s ought to impact returns to speaking English. Skill-biased technical change probably changes the returns to different levels of education, but, conditional on education, it is hard to see why this sort of technical change would elevate the returns to speaking English. It is easier to suspect that affirmative action might impact the return to speaking English. This is a confounding influence that is explicitly discussed when presenting econometric specification and when interpreting results.

When substantial parts of the world did not openly trade with or invest in South Africa, there was still a return to speaking English. South Africa, after all, was not Albania. There remained some international trade, the mining industry produced traded goods, and there was some, albeit minimal, international investment in South Africa. Each of these areas might support a return to speaking English. Furthermore, speaking English was probably coincident with other factors that impacted wages given South

1. One can of course concoct stories, some of them plausible, but few involve as direct a link between global integration and wages as that associated with the returns to speaking English.

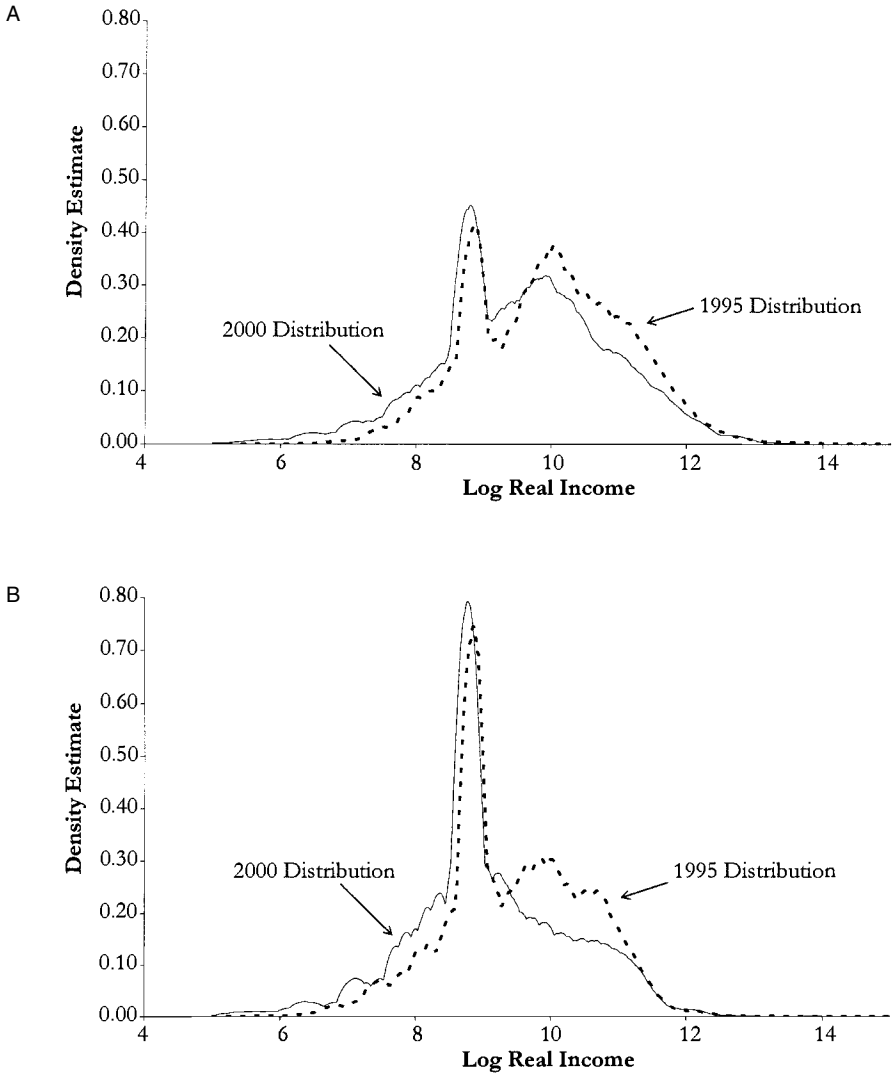
Africa's history (see the Boer War). For these reasons, this paper focuses on whether the return to speaking English *changed*. Of course, if there is no return to speaking English in the first place, searching for changes in that return is not especially informative.

This approach to investigating the impact of globalization on wages is intended as a complement to the way economists usually address this question. My aims are pretty modest. This approach will not offer the definitive word on the impact of globalization on wages in postapartheid South Africa. Put another way, it is hard to imagine evidence on language as being dispositive. Nonetheless, when the cultural situation is appropriate, this approach might usefully add to the trade and wages debate (a.k.a. "cacophony"). Furthermore, this approach uses the sort of survey data that have for the most part been ignored in the trade and wages literature.

That is what this paper is about. It is also important to note what this paper is not about. This paper does not directly speak to the issue of poverty. While it is indirectly related, since wages and incomes are of course related to poverty, this paper does not focus on poverty. The evidence from other sources suggests that the overall distribution of real individual incomes have shifted to the left in the period covered by this paper. This is shown in figure 15.1. That figure shows an almost 40 percent decline in real incomes throughout much of the income distribution from 1995 to 2000. Further examination reveals that Blacks have become relatively worse off since the fall of Apartheid. See Leibbrandt, Levinsohn, and McCrary (2005) for details. Finally, yet other evidence points to increased poverty for Blacks over the period covered by the language data in this paper. See Hoogeveen and Ozler (2004) for further information. To the extent that the return to English has increased, and English is spoken mostly by non-Blacks, the evidence on language and the evidence on incomes and poverty are consistent with one another. It would be wrong, though, to infer any causality.

This paper is not the first to examine the economic implications of speaking English. One paper even does so in the context of considering globalization. Munshi and Rosenzweig (2003) use Indian data to show that lower-caste families are increasingly sending their female children to English schools and that this has encouraging implications for occupational outcomes. Most of the literature on the returns to speaking English uses U.S. data and focuses on the role of language on immigrant earnings. See, for example, Bleakley and Chin (2004a, 2004b) and the literature cited therein. A paper in this vein using U.K. data is Shields and Price (2002).

The paper proceeds in section 15.2 by first describing some of the changes in openness in South Africa since the fall of apartheid. Section 15.3 introduces the data that are used and provides some descriptive statistics. Section 15.4 estimates changes in the return to speaking English, while section 15.5 concludes.



**Fig. 15.1** Distribution of real incomes in South Africa, 1995 and 2000: *A*, men; *B*, women

*Notes:* Figure gives weighted kernel density estimates of log real total income (2000 rand) for men (panel A) and women (panel B) in 1995 and 2000. All four density estimates use an Epanechnikov kernel and a bandwidth selector three-quarters the size of the Silverman (1986) rule of thumb (cf. Silverman's equation [3.31]). Sample sizes for 1995 and 2000 are 21,882 and 16,893 for men, respectively, and 18,868 and 17,776 for women, respectively.

## 15.2 Background

In 1993, the first year of my data, South Africa was preparing for its first nationally representative election in decades. It was clear to all that a new government would be taking power in 1994. There was, though, considerable uncertainty regarding just what economic policies would be pursued by President Nelson Mandela. There were competing pressures to assure the international financial community of continued stability on the one hand, and to dramatically improve the lot of those who had for decades been excluded under the policies of the previous governments (and who were principally responsible for electing the new government) on the other hand.

South Africa quickly implemented a policy of macroeconomic stabilization to reassure the international financial community. Called GEAR for Growth, Employment, and Redistribution, the policy seemed to contribute to stabilization of key macroindicators such as inflation, real interest rates, and the budget deficit. It is less obvious that the policy enhanced growth, employment, and redistribution, but this of course depends on the counterfactual. Each component of the GEAR moniker might have been that much worse in the absence of the policy.

Encouraged by the sober fiscal policies of GEAR, companies from around the world that had hesitated before investing substantially in South Africa began to get off the sidelines. Foreign direct investment skyrocketed. Table 15.1 presents data for foreign direct investment (FDI) in millions of rand (R). The data in table 15.1 show that annual FDI inflows went from only R33 million to over R1.3 billion as soon as the new government was ensconced and proceeded to increase to over R6 billion by 2000. The

**Table 15.1** Foreign direct investment (FDI) in South Africa

Year	FDI
1993	33
1994	1,348
1995	4,502
1996	3,515
1997	17,587
1998	3,104
1999	9,184
2000	6,083
2001 <sup>a</sup>	53,000

*Source:* OECD Global Forum on International Investment.

*Note:* All figures are nominal millions of rand.

<sup>a</sup>Data are only for the first half of 2001.

huge inflow for the first half of 2001 is not typical and represents the one-off purchase of De Beers by the London-listed Anglo American Corporation. Even excluding that transaction, 2001 showed continued healthy increases in FDI inflows. According to the South African Reserve Bank, FDI was split pretty evenly between mining, manufacturing, and the financial sector.

South Africa also joined the World Trade Organization (WTO) on January 1, 1995. Tariffs, never that high anyway, fell into the single-digit range. The largest barrier to trade during the Apartheid era, though, was never tariffs. Rather, it was the willingness of the rest of the world to trade with South Africa. Under the new government, South Africa entered into regional free trading agreements with the European Union and with the Southern African Development Community. Trade, as a percentage of GDP, increased substantially. Table 15.2 presents these figures. From 1991 to 1993, a period during which it became pretty clear that apartheid was going to be replaced with a representative democracy, the ratio of trade to GDP was pretty flat. It was with the new government in 1994 that trade as a fraction of GDP started to really increase. By 2000, the last year of my survey data, the trade-GDP ratio had risen almost 50 percent from .424 to .611. The ratio continued to rise and was .704 in 2002. By almost any standard, these are meteoric increases. Considering that they were mirrored by the even greater increases in FDI, there is little doubt that the South African economy “globalized.” South Africa clearly became more integrated with the global economy after 1993. I turn now to the question of whether the return to speaking English increased over the course of this period.

**Table 15.2** Trade/GDP in South Africa

Year	Trade/GDP
1991	.440
1992	.423
1993	.424
1994	.452
1995	.48
1996	.514
1997	.520
1998	.550
1999	.541
2000	.611
2001	.656
2002	.704

*Source:* South African Reserve Bank web site (<http://www.resbank.co.za>).

*Notes:* Data are for imports and exports of goods and services and annual GDP.

### 15.3 Data

This study uses data drawn from three South African household surveys—one from 1993 and two from 2000. The 1993 data are from the Living Standards Measurement Study (LSMS), a household survey conducted by the World Bank. This survey included about 44,000 individuals comprising just over 8,800 households.<sup>2</sup> The version of the data often used by researchers contains about 300 variables.<sup>3</sup> Information on language and income are key variables for the study at hand. The data on language are not ideal due to the way that the survey instrument was worded. In particular, language is a household-level variable, and the head of the household was asked to identify the “main language spoken at home.” The fact that language is a household-level variable is not of particular concern, since the language spoken at home typically does not vary within the household. The fact that there is no information on whether a person could speak English instead of whether it was the main language spoken at home is a cause for concern, and the results presented below must be considered in light of this. One would of course like to know whether respondents could speak English and how well, not whether it was spoken at home.

This is an example of one-sided measurement error. Some of those who are reported as not speaking English (as measured by the language spoken at home) can in fact speak English quite fluently. On the other hand, few or none of those who stated that English was their first language were in fact unable to speak English. This is because the answer to the language question was asked at the outset and determined the language in which the survey was administered. For example, if someone who spoke only Zulu stated that English was their language, that individual would then have to complete a multihour survey in English. It would not be hard to detect the misstatement of language in this instance.

The 1993 data on income are pretty good. The measure I use in this paper is an individual’s total monthly income and is a constructed variable composed mostly of wage income. It is common in developing countries to highlight the importance of accounting for self-production of food to properly compute income, but this is not an issue in South Africa. Own production is negligible.

For the 2000 data, I combine two surveys, the September 2000 Labour Force Survey (LFS) and the 2000 Income and Expenditure Survey

2. A cleaned and ready-to-use version of the data set, along with a primer for analyzing household survey data in Stata and the survey instruments, is available at the South Africa Distance Learning Project web site, <http://saproject.psc.isr.umich.edu/>.

3. The original data set includes over 2,000 variables, although many of these are essentially individual-level variables that are easily aggregated. Researchers who have used these data include Case and Deaton (1998), Thomas (1996), and Duflo (2000).



(IES).<sup>4</sup> Although the surveys are not explicitly linked, it turns out that the same households were included in both. The merged surveys result in a data set with about 101,000 individuals comprising about 26,000 households. The language question is in the 2000 LFS, while income comes from the 2000 IES. The wording of the language question is the same as in 1993: it again asks about the language spoken most often at home. The 2000 individual income data are used to compute total individual income in a manner most comparable with the 1993 definition. This amounts to subtracting various grants and pensions from individual income.

Table 15.3 presents frequency counts of language by race for each year of the sample. The sample is taken only among those reporting positive income and between the ages of twenty and sixty. The lower bound is intended to exclude students, and results are robust to a lower bound of twenty-five years instead of the somewhat arbitrary twenty. The advantage of using age twenty instead of twenty-five is that the sample size increases substantially. The upper bound is intended to exclude those receiving old-age pensions, since those clearly do not depend on language spoken. Also, old-age pensions are not going to be impacted by globalization as wage income might be. Some women begin to collect these pensions at age sixty, hence the upper bound. There are three key messages from table 15.3. First, very few Blacks list English as their primary language.<sup>5</sup> This is especially true in 1993, and it suggests that the return to speaking English within Blacks is going to be identified off of precious few observations. Second, Coloreds and Whites have a substantial numbers of English speakers. For each, Afrikaans is the majority language, and for each there are substantial shifts in the fraction of the population group reporting English as their primary language. That fraction declines for Coloreds and increases for Whites. Third, English is essentially colinear with Indian, so that it will not be possible to separately identify the impact of English from the impact of being Indian on wages.

Whereas table 15.3 indicated the racial composition of English speakers, table 15.4 illustrates in which sectors of the economy these English speakers work. Tabulating only individuals between the ages of twenty and sixty, table 15.4 shows what fraction of workers in each of eleven sectors list English as their first language. That fraction is highest in business services (composed mostly of the financial sector) at 35.75 percent in 1993 and 32.2 percent in 2000. Other sectors with large fractions of English speakers (or,

4. There was also an IES in 1995 and a linkable household survey (the October Household Survey, or OHS). The 1995 IES and OHS are an attractive data source since questions are asked in the same way and one can be comfortable that income is measured consistently across the 1995 and 2000 surveys. Alas, the 1995 survey “forgot” to include the standard question on language.

5. I use the term “Blacks” since this seems to be preferred by most South Africans to the term “Africans,” which is used in the survey instrument. For data purposes, the two terms are interchangeable. I use the term “Colored” to refer to the well-defined racial group used in South African government statistics.

**Table 15.3** Primary language among wage earners twenty to sixty years old

Language	Black	Colored	Indian	White	Total
<i>A. 1993</i>					
English	6	160	232	202	600
Afrikaans	11	411	0	522	944
Xhosa	421	0	0	0	421
Zulu	624	0	0	0	624
Tswana	493	0	0	0	493
N. Sotho	285	0	0	0	285
S. Sotho	91	0	0	0	91
Venda	79	0	0	0	79
Tsonga	209	0	0	0	209
Swazi	173	0	0	0	173
Ndebele	69	0	0	0	69
Other	7	0	3	6	16
Total	2,468	571	235	730	4,004
<i>B. 2000</i>					
English	118	534	626	783	2,075
Afrikaana	230	3,032	6	1,540	4,838
Ndebele	435	0	0	0	435
Xhosa	3,789	17	0	0	3,806
Zulu	5,487	3	0	0	5,490
N. Sotho	2,238	1	0	0	2,239
S. Sotho	2,681	6	0	1	2,688
Tswana	2,601	17	0	0	2,618
Swazi	871	4	0	0	876
Venda	577	0	0	0	577
Tsonga	1,109	0	0	0	1,109
Other	72	2	22	19	120
Missing	14	0	0	0	14
Total	20,222	3,616	654	2,343	26,885

*Note:* 2000 row totals do not sum properly due to the exclusion of nonresponses to the race question.

more accurately, English “listeners”) include manufacturing, electricity, wholesale and retail trade, and community services (which includes doctors, teachers, and lawyers). In all sectors, the fraction listing English declined from 1993 to 2000, usually modestly. There are in principle two ways that the economy might adapt to an increased demand for English: the supply could increase, or the return could increase (or both). Table 15.4 suggests that the supply did not increase. I turn next to examining whether the return to speaking English increased.

#### 15.4 The Return to Speaking English

The question at hand is whether the return to speaking English (as imperfectly measured) increased as South Africa opened up to the interna-

**Table 15.4** Share of industry employment by language

Sector	1993		2000	
	Other	English	Other	English
Agriculture	98.09	1.91	98.36	1.64
Mining	95.45	4.55	97.25	2.75
Manufactures	75.26	24.74	82.70	17.30
Electric	83.30	16.70	83.77	16.23
Construction	83.25	16.75	90.54	9.46
Wholesale and retail	81.35	18.65	85.61	14.39
Transport	82.78	17.22	84.47	15.53
Business services	64.25	35.75	67.76	32.24
Community services	81.67	18.33	84.56	15.44
Private households	99.11	0.89	97.58	2.42
Other	79.42	20.58	82.46	17.54
Total	84.87	15.13	86.95	13.05

*Notes:* Cell entries give the share of employment in a given industry that lists English as the first language. The 1993 and 2000 data sets had different industry categories, and the categories in this table reflect a concordance to the 2000 industry definitions. In particular, the 1993 categories of wholesale and retail trade and restaurant and hotel services were combined. Also, the 1993 categories of education, medical, and legal services were combined to form “community services.” Industry names are from the Statistics South Africa Labour Force Survey 2000 report, page vii.

tional economy from 1993 to 2000. The return to speaking English is not directly observable and so needs to be inferred from econometric evidence. The approach adopted here is to estimate Mincer-like wage regressions and include as an explanatory variable whether the wage-earner listed English as his or her primary language. While this is simple in principle, several issues arise in practice.

First, it is necessary even in the cross section to include as explanatory variables key determinants of wages.<sup>6</sup> Omission of an explanatory variable that itself might be correlated with speaking English will bias the estimate on the return to speaking English. Second, the many changes in South Africa from 1993 to 2000 probably impacted many of the determinants of wages. It is widely believed, for example, that the return to education and the wage differentials apparently due to race changed over this period. Holding them constant and only allowing the return to English to change will yield biased estimates of the true change in the return to speaking English. (On the other hand, such an approach pretty much guarantees find-

6. A difference-in-differences approach is not advised because of concurrent changes in many other variables that impact wages. That is, although one could measure the difference in wages between those who list English as their first language and those who do not, and one could then examine the difference over time in this difference, the result would be hard to interpret. This is because many other variables changed over this period, and some of those changes are not orthogonal to an observed return to speaking English.

ing a pretty big change in the return to speaking English.) Third, the fact that about 40 percent of English speakers are Indian and there is virtually no language variation within this population group poses a challenge. The most flexible approach to estimating the returns to speaking English examines the change in that return within population group, yet this approach is going to be noninformative for Indians.

The simplest specification regresses log individual income ( $y_i$ ) on indicator variables for each value of  $j$  years of education (ED), experience (EX), experience squared, an indicator variable for whether the worker is male ( $M$ ), indicator variables for population group (CO for Colored, IN for Indian, and WH for White, with Blacks as the excluded group), and an indicator for whether English is the language spoken at home (ENG). Experience is defined as age minus 20. Hence,

$$(1) \ln y_i = \beta_0 + \sum_{j=2}^{j=13} \beta_{1,j} \text{ED}_j + \beta_2 \text{EX} + \beta_3 \text{EX}^2 + \beta_4 M + \beta_5 \text{CO} + \beta_6 \text{IN} \\ + \beta_7 \text{WH} + \beta_8 \text{ENG} + \varepsilon_i.$$

Equation (1) is estimated separately for each year of the sample using ordinary least squares (OLS) with the appropriate sample weights. Estimating the regression separately for each year is necessary to capture the changes in returns to education between 1993 and 2000 as well as changes in the return to being male and/or of a particular population group. Use of indicator variables for each level of education permits returns to vary nonlinearly with years. The coefficient on English,  $\beta_8$ , is interpreted as the percentage wage differential attributable to speaking English conditional on the other included regressors. The results from this specification applied to the 1993 and 2000 data are presented in table 15.5.

The results from 1993 are discussed first to fix ideas. The first twelve rows show the usual returns to education. For example, someone with twelve years of education, all else being equal, earns about 168 percent more than those with one year or less of education conditional on the other covariates. The wage premium for being a member of a race other than Black ranges from 32 percent for Colored to 98 percent for White. Males earn 46 percent more than similar females. The coefficient of interest for this study, though, is that on "English." Conditional on education, experience, gender, and race, people who list English as their primary language earn about 18 percent more than those who list another language. This differential is quite precisely estimated.

Equation (1) is estimated using the 2000 data, and the results are shown in the second column of table 15.5. While there are several interesting comparisons between 1993 and 2000 to be made (the changing pattern of the return to education, for instance), the focus here is on the impact of speaking English. The "English premium" jumps from .183 in 1993 to .252 in

**Table 15.5**                      **The returns to speaking English among those twenty to sixty years old**

	1993	2000
ED2	-.093 (.071)	.092 (.035)
ED3	.219 (.067)	.233 (.038)
ED4	.299 (.065)	.224 (.036)
ED5	.449 (.060)	.295 (.034)
ED6	.472 (.052)	.380 (.031)
ED7	.733 (.051)	.549 (.030)
ED8	.622 (.060)	.646 (.032)
ED9	.948 (.050)	.820 (.030)
ED10	.993 (.063)	.908 (.031)
ED11	1.274 (.047)	1.240 (.027)
ED12	1.688 (.054)	1.796 (.031)
ED13	1.788 (.075)	2.126 (.036)
EX	.062 (.004)	.099 (.002)
EX2	-.001 (.000)	-.001 (.000)
Colored	.326 (.037)	.360 (.020)
Indian	.394 (.071)	.421 (.041)
White	.984 (.037)	.921 (.021)
English	.183 (.043)	.252 (.024)
Male	.463 (.024)	.501 (.012)
Constant	5.041 (.050)	7.183 (.030)
$R^2$	.58	0.46
No. of observations	3,979	26,616

*Note:* Standard errors in parentheses.

**Table 15.6** The returns to speaking English: Within-group results

	1993	2000
All	0.183 (0.043)	0.252 (.024)
Black	0.592 (0.290)	0.380 (.084)
Colored	0.521 (0.0765)	0.410 (.042)
White	-0.017 (.050)	0.145 (.036)

*Note:* Standard errors in parentheses.

2000. The 2000 premium is precisely estimated, and the change between the two years is significantly different from zero. Allowing the entire pattern of returns to years of schooling to vary from 1993 to 2000, and allowing for differing returns to race, gender, and experience, it is still the case that the return to speaking English increased substantially. This change in the English premium, as well as its level, is of an economically large magnitude. By 2000, English speakers were earning about 25 percent more after conditioning on other observables, and the premium had increased by 7 percentage points since 1993.

The specification reported in table 15.5 imposes that the returns to education, experience, and gender are identical across racial groups. A convincing body of research suggests that this is too strong an assumption. I proceed by looking for the English premium within each of the racial groups. Doing so allows the returns on all the other observables to vary by racial group. This flexibility is clearly a good thing, for it lets the data speak more freely. The flexibility, though, will carry a price. Of the sample that listed English as the language spoken at home, 30 to 40 percent are Indian, and virtually all Indians list English as the primary language. There is, then, no within-group language variation for Indians. Hence, it is not feasible to estimate a return to speaking English for Indians since that return is not identifiably different from the return to simply being Indian.<sup>7</sup>

Table 15.6 reports results from the within-group regressions for Blacks, Coloreds, and Whites. In the interest of parsimony, only the coefficient on speaking English is reported.<sup>8</sup> This approach is pretty flexible. It allows

7. It is possible to estimate a return to speaking English among Indians, but the effect is identified off of three individuals who listed "Other" in 1993 and about six Afrikaans-speaking Indians in 2000. The English premium, when separate regressions are run for Indians, is never significantly different from zero.

8. Were it the case that the returns on the other observables (not reported in this table) did not vary significantly across racial groups, it would be efficient to pool groups. Alas, coefficients vary across groups, and one can readily reject the hypothesis that the returns to observables other than the English premium are the same across groups.

the returns on all observables to vary both over time and across racial groups.

As is usually the case with a more flexible specification, the messages are more mixed than those reported in table 15.5. The first row of table 15.6 reports the English premium from table 15.5 for comparison's sake. The next three rows report the English premium for the other racial groups (except Indian, for reasons discussed above). For Blacks, the English premium stayed constant from 1993 to 2000. It was huge (about 60 percent) but did not increase over time, although the precision of the estimate did increase. One should recall, though, that this premium is being identified off of very few individuals: 6 out of 2,468 in 1993 and 118 out of 20,222 in 2000. For Coloreds, the return to speaking English fell about 11 percentage points. The decline, while not large, is statistically significantly different from zero. The largest change from 1993 to 2000 in the English premium impacted Whites. The return went from being basically nonexistent in 1993 to a precisely estimated 14.5 percent. Another way to interpret this result is that the "penalty" to speaking Afrikaans among Whites skyrocketed.

The general pattern reported in table 15.6 is robust to many alternative specifications. For example, the inclusion of indicator variables for the province in which a household lives, using a single variable for years of education instead of the more flexible set of indicator variables, the exclusion of the variable for male, using age twenty-five as a lower age bound, using sixty-five as an upper age bound, and interacting the return to English with education all yield basically the same message when it comes to the English premium. Namely, that premium became much larger for Whites and fell slightly for Coloreds and for Blacks.<sup>9</sup> These results, though, exclude a large number of those with positive income for whom English is the first language—Indians.

### 15.5 Concluding Remarks

Did globalization really cause the return to speaking English to increase in South Africa? The evidence in this paper is, in some cases, corroborating, but hardly conclusive. The strongest and most robust result is that the return to speaking English increased for Whites over the period during which South Africa reintegrated with the world economy. This result is strong because it results from the flexible within-group estimates, and it is robust because it arises in all the investigated specifications. When Indians are included and a (necessarily) less flexible estimation strategy is adopted, I again find that the return to English increased and that the increase is precisely estimated. These are the results in table 15.5. If one thinks of these

9. For some specification, the premium rises slightly for Blacks.

results as indicating an average effect of speaking English, that effect is positive. There is less evidence, though, that the return to speaking English increased among Blacks and Coloreds.

One explanation for the lack of an increase in the return to speaking English among Blacks is the following. In 1993, there were few Blacks who spoke English, and they earned a premium for their language skills. With the advent of affirmative action, the premium for speaking English fell as more Blacks were promoted into higher-paying jobs. In this case, it was no longer just the few English-speaking Blacks earning the relatively higher wages. This scenario illustrates one of the difficulties of disentangling the impact of globalization (which might actually increase the return to speaking English) with the impact of affirmative action for Blacks (which was concurrent with globalization and which might actually decrease the extra return to speaking English).

There was no affirmative action for Whites, and among this group the return to speaking English clearly rose. Put another way, the penalty for speaking Afrikaans rose for Whites. This is consistent with capturing an impact of globalization. Afrikaans is much less useful than English in international commerce. Those Whites whose first language was English benefited conditional on education, gender, and experience. This is, as noted above, corroborating but not conclusive evidence.

The finding that the return to speaking English did not increase for Coloreds muddies the waters. Coloreds did not benefit from affirmative action as did Blacks under the new government. Still, the return to speaking English did not rise; in fact, it fell. If globalization is what moves the return to speaking English, one should have found an increase to speaking English among Coloreds, and this was not the case.

The evidence in the end is mixed. On the whole, the return to speaking English increased, but within racial groups the pattern is not consistent.<sup>10</sup> The approach adopted in this paper is perhaps a novel way to revisit the wages and globalization issue. It is an approach that is especially well suited to developing countries, many of which have a rich variety of languages spoken, as they integrate with the global economy. In other contexts (India, for example), or with better data (industry of employment data, for example), the approach adopted here may prove more conclusive. Or not. Even if language is an accurate way to isolate an impact of globalization on wages, it may simply be that globalization has differing impacts on differing segments of a population. This appears to be the case in South Africa.

10. It should be noted that precious few of the English speakers are among the very poor. In 1993, virtually none are, while in 2000 only a handful are. Hence, this approach does not speak to the role of globalization on the incomes of the very poor.



## References

- Bleakley, H., and A. Chin. 2004a. Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics* 86 (2): 481–96.
- . 2004b. What holds back the second generation? The intergenerational transmission of language human capital among immigrants. Working Paper no. 104. University of California, San Diego, Center for Comparative Immigration Studies, October.
- Case, A., and A. Deaton. 1998. Large cash transfers to the elderly in South Africa. *Economic Journal* 108:1330–61.
- Duflo, E. 2000. Grandmothers and granddaughters: Old age pensions and intra-household allocation in South Africa. Massachusetts Institute of Technology, Department of Economics. Working Paper.
- Hoogeveen, J., and B. Ozler. 2004. Not separate, not equal: Poverty and inequality in postapartheid South Africa. Washington, DC: World Bank. Unpublished Manuscript.
- Leibbrandt, M., J. Levinsohn, and J. McCrary. 2005. Incomes in South Africa since the fall of apartheid. NBER Working Paper no. 11384. Cambridge, MA: National Bureau of Economic Research.
- Munshi, K., and M. Rosenzweig. 2003. Traditional institutions meet the modern world: Caste, gender and schooling choice in a globalizing economy. Working Paper 03-23. Massachusetts Institute of Technology, Department of Economics.
- Shields, M., 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* 15:137–60.
- Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. Monographs on Statistics and Applied Probability no. 26. London: Chapman and Hall.
- Thomas, D. 1996. Education across the generations in South Africa. *American Economic Review* 86 (2): 330–34.

## Comment      Raquel Fernández

This paper by James Levinsohn seeks to add to the large literature on globalization, trade, and wages by examining the change in the returns to speaking English in South Africa. This strategy has the potential to avoid the usual problem that plagues this literature—that of distinguishing the effects of globalization from other concurrent changes—if, as the author argues, the change in the returns to speaking English is independent of other sources of changes in the economy.

The period examined is from 1993 to 2000. This was a time of important changes in South Africa. The end of apartheid was followed by the first nationally representative election, and the country joined the WTO and var-

ious regional free trade agreements. Trade and foreign direct investment as a percentage of GDP increased dramatically over this period as South Africa integrated with rest of the world.

Using household data sets from 1993 and 2000, Levinsohn finds that the returns to speaking English (where the latter is really “the main language spoken at home”) changes over this time period. He finds that, constraining the returns to speaking English to be the same across population groups (Whites, Blacks, Coloreds, and Indians), and running a standard Mincer wage regression that controls for age, education, experience, gender, and population group, the returns to speaking English increase by 7 percentage points.

Acknowledging that the returns to education or to any of the other explanatory variables are likely to differ across groups, Levinsohn runs the same wage regression for each racial group separately. In this case, however, the results look disturbingly different. The return to speaking English dropped by 21 percentage points for Blacks and by 11 points for Coloreds. It increased only for Whites, by 13 percentage points.

I found the results difficult to interpret. Levinsohn argues that an explanation for the decrease for Blacks may be due to affirmative action’s ending the premium for speaking English as more Blacks were promoted to high-wage jobs. Maybe. It is very hard to make sense of a result that, in any case, relies on only six observations of Blacks who declare English as the primary language spoken in the household in 1993. This hypothesis does not, in any case, help explain the decrease in the return to speaking English for Coloreds, as this group probably did not benefit from affirmative action. Furthermore, although the author interprets the increase in the return to English for Whites (or the decrease in the return to speaking Afrikaans) as resulting from globalization, another possibility is again the end of apartheid. Presumably this would hurt primarily those Whites who were most associated with the apartheid regime—those Whites of Dutch heritage—rather than all Whites. Hence, it is not at all clear what is driving these results.

Another possible explanation for the results is the traditional one of skill-biased technological change. The author’s argument for discarding this possibility is that skill-biased technological change should show up in the returns to education. A plausible counterargument, however, is that the quality of education is higher precisely in those schools attended by students who primarily speak English at home. This would also explain why the returns only went up for Whites if only this group attends high-quality schools. Is this so? It is hard to answer this question without knowing more about who attends which type of school in South Africa.

In general this paper would have benefited from presenting the socio-economic and political-economic history and present of South Africa in greater detail. As it stands, it is very hard to evaluate the validity of alter-

native hypotheses. For example, to what extent were Afrikaans-speaking Whites at a greater disadvantage than other Whites in the new South Africa? What was and what happened to the position of individuals of mixed race?

I also had some questions about the data. Neither the racial proportions in the data sets nor the languages spoken seem very comparable across years. Whites, for example, constitute a bit over 18 percent of population in 1993 and under 9 percent in 2000. Is this due to Whites leaving the country? In general, the percentage of Blacks increased significantly relative to all other races. Similarly, how are we to interpret the large increase in the proportion of Coloreds that declare Afrikaans to be their primary language? Since it is doubtful that the primary language spoken at home changed during this seven-year period, it renders the results for this group questionable at best.

In conclusion, I think that the author has explored an interesting hypothesis that could fruitfully be applied to other countries. It would be very interesting to examine whether increased integration changed the returns to speaking English in India or Argentina, for example. For the present study, the analysis unfortunately suffers from an important data limitation (in particular the inability to identify whether individuals can speak English) and, more easily transcended, from not placing individuals and their race, education, and language abilities in a socioeconomic context.