Making the 'Next Billion' Demand: the Effect of Local Content through Google.co.za in Setswana

Bastiaan Quast

January 2016

Abstract

Internet connectivity provides enormous improvements in quality of life as well as opportunities for the newly connected. Attempts to connect the "next billion" in Africa have not met expectations, even in places where infrastructure is in place. This paper shows that an exogenous increase in local content creates an enormous incease in demand among native speakers, even when demand as a whole is falling as a result of the economic malaise.

The introduction of the Setswana language in the South African Google Search website was a side of effect of this translation work being done for the Botswanan Google Search website (where Setswana is the official language, together with English) and provides and exogenous event which catalysed a huge increase in the number of internet-connected native speakers, as well as actual usage of the Setswana language online.

1 Introduction

Tswana language translation was done for Botswana (www.google.co.bw), so result are free of endogeneity.

2 Methods

- Difference in Differences
- no logit because DiD
- should try Imbens etc.

3 Data

- NIDS

4 Results

- Significant only for Setswana and Venda
- Venda not significant for computer
- include hhincome woman best_edu

```
summary(lm4_0)

##
## Call:
## lm(formula = m4_0, data = adulthh)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.13357 -0.01363 -0.00520 -0.00345  0.99802
##
## Coefficients: (1 not defined because of singularities)
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     0.015337   0.006910   2.219  0.02646 *
## post_eventTRUE                 -0.015337   0.011607  -1.321  0.18638
## factor(a_lng)2                 -0.011977   0.007130  -1.680  0.09298 .
## factor(a_lng)3                 -0.010139   0.007031  -1.442  0.14933
## factor(a_lng)4                 -0.011884   0.007330  -1.621  0.10496
## factor(a_lng)5                 -0.001710   0.007348  -0.233  0.81596
## factor(a_lng)6                 -0.010081   0.007271  -1.386  0.16560
## factor(a_lng)7                 -0.010693   0.008477  -1.262  0.20712
```

```
## factor(a_lng)8                  0.118237   0.010196  11.597  < 2e-16 ***
## factor(a_lng)9                  0.025849   0.008567   3.017  0.00255 **
## factor(a_lng)10                 0.029305   0.007130   4.110 3.96e-05 ***
## factor(a_lng)11                 0.088444   0.007764  11.392  < 2e-16 ***
## factor(a_lng)12                 0.092771   0.021645   4.286 1.82e-05 ***
## post_eventTRUE:factor(a_lng)2   0.013953   0.011955   1.167  0.24318
## post_eventTRUE:factor(a_lng)3   0.014632   0.011794   1.241  0.21476
## post_eventTRUE:factor(a_lng)4   0.016932   0.012224   1.385  0.16603
## post_eventTRUE:factor(a_lng)5   0.013013   0.012301   1.058  0.29010
## post_eventTRUE:factor(a_lng)6   0.020708   0.012188   1.699  0.08932 .
## post_eventTRUE:factor(a_lng)7   0.024171   0.014170   1.706  0.08805 .
## post_eventTRUE:factor(a_lng)8  -0.118237   0.015523  -7.617 2.65e-14 ***
## post_eventTRUE:factor(a_lng)9  -0.023496   0.014036  -1.674  0.09413 .
## post_eventTRUE:factor(a_lng)10  0.010640   0.011962   0.890  0.37372
## post_eventTRUE:factor(a_lng)11  0.013933   0.013255   1.051  0.29321
## post_eventTRUE:factor(a_lng)12        NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1248 on 49276 degrees of freedom
##   (11933 observations deleted due to missingness)
## Multiple R-squared:  0.03597,Adjusted R-squared:  0.03554
## F-statistic: 83.57 on 22 and 49276 DF,  p-value: < 2.2e-16
```

4.1 LM4_1

```
summary(lm4_1)

##
## Call:
## lm(formula = m4_1, data = adulthh)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.14231 -0.01879 -0.01159  0.00079  1.01927
##
## Coefficients: (2 not defined because of singularities)
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    0.0318595  0.0071928   4.429 9.47e-06 ***
## post_eventTRUE                -0.0160615  0.0115788  -1.387  0.16540
## factor(a_lng)2                -0.0099108  0.0073031  -1.357  0.17477
## factor(a_lng)3                -0.0074446  0.0072117  -1.032  0.30194
## factor(a_lng)4                -0.0117912  0.0075144  -1.569  0.11662
## factor(a_lng)5                -0.0024480  0.0075095  -0.326  0.74444
```

```
## factor(a_lng)6                 -0.0090931  0.0074556  -1.220  0.22261
## factor(a_lng)7                 -0.0109606  0.0086869  -1.262  0.20705
## factor(a_lng)8                  0.1112536  0.0102699  10.833  < 2e-16 ***
## factor(a_lng)9                  0.0241967  0.0088206   2.743  0.00609 **
## factor(a_lng)10                 0.0327904  0.0073039   4.489 7.16e-06 ***
## factor(a_lng)11                 0.0832838  0.0079063  10.534  < 2e-16 ***
## factor(a_lng)12                 0.0929623  0.0216230   4.299 1.72e-05 ***
## a_edlitrden                    -0.0033726  0.0022402  -1.505  0.13220
## a_edlitwrten                   -0.0068192  0.0021992  -3.101  0.00193 **
## a_edlitrdhm                     0.0005543  0.0021514   0.258  0.79667
## a_edlitwrthm                    0.0021000  0.0021548   0.975  0.32978
## a_womanTRUE                    -0.0012292  0.0011597  -1.060  0.28918
## post_eventTRUE:factor(a_lng)2   0.0132434  0.0119238   1.111  0.26672
## post_eventTRUE:factor(a_lng)3   0.0145833  0.0117679   1.239  0.21526
## post_eventTRUE:factor(a_lng)4   0.0177728  0.0121993   1.457  0.14516
## post_eventTRUE:factor(a_lng)5   0.0129743  0.0122610   1.058  0.28998
## post_eventTRUE:factor(a_lng)6   0.0200726  0.0121631   1.650  0.09889 .
## post_eventTRUE:factor(a_lng)7   0.0253515  0.0141395   1.793  0.07299 .
## post_eventTRUE:factor(a_lng)8  -0.1117654  0.0154157  -7.250 4.23e-13 ***
## post_eventTRUE:factor(a_lng)9  -0.0181736  0.0140391  -1.294  0.19550
## post_eventTRUE:factor(a_lng)10  0.0095912  0.0119282   0.804  0.42136
## post_eventTRUE:factor(a_lng)11         NA         NA      NA       NA
## post_eventTRUE:factor(a_lng)12         NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1226 on 46438 degrees of freedom
##   (14767 observations deleted due to missingness)
## Multiple R-squared:  0.03921,Adjusted R-squared:  0.03867
## F-statistic: 72.89 on 26 and 46438 DF,  p-value: < 2.2e-16
```

4.2 LM4_5

```
summary(lm4_5)

##
## Call:
## lm(formula = m4_5, data = adulthh)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.71081 -0.02109 -0.00980 -0.00074  1.01644
##
## Coefficients:
```

```
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 -7.802e-04  1.838e-03  -0.424  0.67127
## post_eventTRUE              -1.186e-02  1.219e-03  -9.728  < 2e-16 ***
## tsongaTRUE                  -1.363e-02  2.430e-03  -5.609 2.04e-08 ***
## factor(a_edlitrden)2         1.540e-03  4.030e-03   0.382  0.70233
## factor(a_edlitrden)3         2.169e-04  5.283e-03   0.041  0.96724
## factor(a_edlitrden)4        -3.700e-03  6.823e-03  -0.542  0.58760
## factor(a_edlitwrten)2       -1.027e-02  4.020e-03  -2.555  0.01064 *
## factor(a_edlitwrten)3       -1.081e-02  5.229e-03  -2.067  0.03874 *
## factor(a_edlitwrten)4       -7.178e-03  6.701e-03  -1.071  0.28407
## factor(a_edlitrdhm)2        -2.615e-03  3.691e-03  -0.709  0.47857
## factor(a_edlitrdhm)3        -2.935e-03  5.139e-03  -0.571  0.56795
## factor(a_edlitrdhm)4        -8.026e-03  6.916e-03  -1.160  0.24587
## factor(a_edlitwrthm)2        8.491e-04  3.729e-03   0.228  0.81990
## factor(a_edlitwrthm)3        1.395e-03  5.147e-03   0.271  0.78640
## factor(a_edlitwrthm)4       -6.975e-03  6.920e-03  -1.008  0.31354
## a_womanTRUE                 -1.413e-03  1.147e-03  -1.232  0.21799
## hhincome                     2.753e-06  5.931e-08  46.412  < 2e-16 ***
## best_edu                     1.363e-03  1.165e-04  11.698  < 2e-16 ***
## post_eventTRUE:tsongaTRUE    1.207e-02  3.875e-03   3.114  0.00184 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1211 on 46446 degrees of freedom
##   (14767 observations deleted due to missingness)
## Multiple R-squared:  0.06175,Adjusted R-squared:  0.06139
## F-statistic: 169.8 on 18 and 46446 DF,  p-value: < 2.2e-16
```

4.3 LM2_5

```
summary(lm2_5)

##
## Call:
## lm(formula = m2_5, data = adulthh)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6364 -0.0768 -0.0356  0.0026  1.1055
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  7.624e-03  3.136e-03   2.431 0.015050 *
## post_eventTRUE              -5.432e-03  2.092e-03  -2.596 0.009433 **
```

```
## tsongaTRUE                -1.480e-02  4.167e-03  -3.550 0.000385 ***
## factor(a_edlitrden)2      -3.065e-02  6.901e-03  -4.441 8.96e-06 ***
## factor(a_edlitrden)3      -3.096e-02  9.031e-03  -3.428 0.000609 ***
## factor(a_edlitrden)4      -3.900e-02  1.166e-02  -3.344 0.000825 ***
## factor(a_edlitwrten)2     -1.746e-02  6.886e-03  -2.536 0.011224 *
## factor(a_edlitwrten)3     -2.105e-02  8.937e-03  -2.355 0.018517 *
## factor(a_edlitwrten)4     -1.871e-02  1.145e-02  -1.634 0.102274
## factor(a_edlitrdhm)2      -1.869e-03  6.322e-03  -0.296 0.767476
## factor(a_edlitrdhm)3      -4.239e-03  8.803e-03  -0.482 0.630138
## factor(a_edlitrdhm)4      -2.673e-02  1.188e-02  -2.251 0.024415 *
## factor(a_edlitwrthm)2      1.227e-03  6.383e-03   0.192 0.847597
## factor(a_edlitwrthm)3     -1.998e-03  8.820e-03  -0.227 0.820792
## factor(a_edlitwrthm)4     -3.575e-02  1.189e-02  -3.008 0.002631 **
## a_womanTRUE               -2.299e-02  1.960e-03 -11.729  < 2e-16 ***
## hhincome                   5.820e-06  1.022e-07  56.931  < 2e-16 ***
## best_edu                   5.835e-03  2.002e-04  29.143  < 2e-16 ***
## post_eventTRUE:tsongaTRUE  2.385e-02  6.684e-03   3.569 0.000359 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2097 on 47647 degrees of freedom
##   (13566 observations deleted due to missingness)
## Multiple R-squared:  0.129,Adjusted R-squared:  0.1286
## F-statistic: 391.9 on 18 and 47647 DF,  p-value: < 2.2e-16
```

5 Conclusions and Limitations

- need more local content
- need more research