

## Ghana Data GUPHS 2006

### Variable definitions and information, as used in apprenticeship paper, CSAE working paper WPS/2008-08

Note: All variables have been cleaned extensively from the raw data.

1. lw06: Monthly earnings in 2006 cedis. Calculated from the earnings information recorded for wage employees (based on whether they are paid daily, weekly, etc.) and profits reported by the self-employed. Double checked against revenues.
2. newraven: Raven's score out of 20. Each question weighted equally. Given value of zero if missing.
3. neweng: score on English test.
4. newrdg: score on reading test.
5. newmath: score on math/numeracy test. Second section weighted more heavily than the first section.
6. cityG1-cityG4: city dummies
7. income: dummy equalling 1 if reporting any earned income.
8. Student: dummy equalling 1 if a student. Cross-referenced from the individual and household questionnaires.
9. broeduc: closest brother's education in years. Replaced with average if no sibling.
10. siseduc: closest sister's education in years. Replaced with average if no sibling.
11. siblings: number of siblings.
12. appcost: self-reported fee for apprenticeship.
13. traintime: time spent training, vocational and other types, in years.
14. apptraintime: time spent training for people who only did apprenticeships.
15. educ: education in years. Checked extensively against information about highest education level completed. Technical and polytechnic count for 2 or 3 years, as reported by the individual. Koranic education receives zero years. If panel member is missing the education, it is taken from previous years of the survey.
16. educsq: education squared/100.
17. dadeduc: father's education in years. If panel member is missing this, it is taken from previous years of the survey.
18. momeduc: mother's education in years. If panel member is missing this, it is taken from previous years of the survey.
19. age: age in years. Calculated from approximate birthyear and birthmonth.
20. agesq: age squared/100.

21. male: dummy equalling 1 if male. Checked against past years for panel members and any discrepancy was checked by first name.
22. hhcreditaccess: dummy equalling 1 if hh reports borrowing in the household questionnaire. Also takes value of 1 if an individual reports borrowing in the individual questionnaire that is not picked up in the household questionnaire.
23. othcreditaccess: dummy equalling 1 if someone other than the individual is borrowing
24. hhothapp: number of other individuals also interviewed in the household who report being an apprentice; summed over the household id.
25. waterinhome: dummy equalling 1 if report having piped internal water in the house.
26. wevocskills: (uncleaned) describes type of skills acquired during training.
27. vocpast: dummy equalling 1 if did vocational training in the past.
28. apppast: dummy equalling 1 if former apprentice.
29. app\_now: dummy equalling 1 if current apprentice.
30. app\_ever: dummy equalling 1 if ever an apprentice (sum of apppast and app\_now).
31. lhours: natural log of reported hours worked per week. Missings replaced with  $\ln(40)$ .
32. kdistsec: distance to secondary school in km.
33. kdistprim: distance to primary school in km
34. hhhead: dummy equalling 1 if household head.
35. ehighest: description of highest education level achieved.
36. prim jssormid sss postsec: dummies for education years completed (primary, junior/middle, secondary, post-secondary, respectively)