Who Responds to Voluntary Cognitive Tests in Household Surveys? The Role of Labour Market Status, Respondent Confidence, Motivation and a Culture of Learning in South Africa

HENDRIK VAN BROEKHUIZEN AND DIETER VON FINTEL

Stellenbosch Economic Working Papers: 27/10

KEYWORDS: EDUCATION, BEHAVIOURAL ECONOMICS, VOLUNTARY ASSESSMENT, NUMERACY, SURVEY DESIGN, SURVEY NON-RESPONSE, SAMPLE SELECTION BIAS, RESPONDENT CONFIDENCE, MOTIVATION, CULTURE OF LEARNING, SOUTH AFRICA
JEL: C81, C83, D03, I21

HENDRIK VAN BROEKHUIZEN
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: HENDRIKVANB@SUN.AC.ZA

DIETER VON FINTEL
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: DIETER2@SUN.AC.ZA

UNIVERSITEIT STELLENBOSCH UNIVERSITY

BER
BUREAU FOR ECONOMIC RESEARCH

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

Who Responds to Voluntary Cognitive Tests in Household Surveys? The Role of Labour Market Status, Respondent Confidence, Motivation and a Culture of Learning in South Africa[1]

HENDRIK VAN BROEKHUIZEN AND DIETER VON FINTEL

ABSTRACT

Both South Africa's labour market and education system were directly influenced by the separate development policies of the apartheid regime. To this day, great inequalities persist in both domains. South Africa's performance in standardized international test scores (such as TIMMS) is poor even relative to most developing countries. Furthermore, the better quality of outcomes in former white schools still leaves learners from former black schools at a disadvantage that feeds through to severe labour market inequalities. This study is the first in a series of papers that attempts to understand the role of school quality on labour market outcomes. Here we scrutinize the measurement of numeracy test scores in the National Income Dynamics Survey (NIDS) of 2008, particularly in light of potential sample selection issues. While this survey measures standard welfare and labour market indicators, it is one of the first in South Africa to also ask respondents to complete a concurrent numeracy test. Response rates on this module were particularly low, given that the test was taken on a voluntary basis. We develop a basic empirical model to understand who is likely to take the test. We postulate that discouraged workers' low propensity to take the test is correlated with their reduced motivation to undertake job search, that the searching unemployed are highly motivated to take the test (as they wish to gauge their ability or practice assessments while embarking on the job search process), the poorest among the self-employed face severe time opportunity costs (as their low incomes are less secure than those of salaried workers) and the richest amongst the employed exhibit an income effect (in that the time opportunity costs of their high incomes reduce their willingness to respond to the numeracy test). Furthermore, locational effects suggest that those residing in geographical "points of entry" into the labour market are also more likely to take the test. The young (who are still in education) and the most educated (in the whole population) also tend to answer the test more readily. The latter observations indicate that some form of confidence in respondents' own abilities drives their response patterns. To explain these observed features, we construct composite indices of motivation/emotional well-being and individuals' confidence in their writing abilities using multiple correspondence analysis. While each of these psychological and behavioural factors is a strong predictor of test response,

[1] This document was produced within the Social Policy Research Group in the Department of Economics at Stellenbosch University with the financial assistance of the PSPPD (Programme to Support Pro-Poor Policy Development in South Africa), a partnership programme between The Presidency, Republic of South Africa, and the European Union (EU). The contents do not reflect the position of The Presidency or the EU.

they do not entirely eliminate the independent contributions of each of the observed influences mentioned above. Coefficient magnitudes of each of the sociodemographic variables are, however, reduced, indicating that the particular behavioural influences introduced in later models tell some of the story. Additional uncaptured behavioural and motivational factors are therefore investigated. Firstly, we investigate the role of survey fatigue (by controlling for the time it took to complete the survey before the test was administered), which plays an important role in the black and coloured subpopulations. It furthermore explains why the wealthiest amongst the formally employed are less likely to complete the numeracy test. However, surprisingly, "pseudoaltruistic" effects appear amongst the (wealthier) white population, in that the longer the duration of the preceding questions, the more likely they are to care about answering the test. However, this result cannot be generalized to the whole white population, as response rates were very low among this group. Secondly, (household) peer effects are strong throughout the population, suggesting that a culture of learning is pivotal in understanding response patterns. The results of this paper suggest that broad sociodemographic and labour market features remain important determinants of test response, even after controlling for behavioural features. This suggests that subsequent labour market work must take these drivers into account to avoid the risk of sample selection bias.

Keywords: education, behavioural economics, survey design, voluntary assessment, numeracy, survey non-response, sample selection bias, respondent confidence, motivation, culture of learning, South Africa

JEL codes: C81, C83, D03, I21

1 Introduction and Background

Measurements of cognitive ability have played an increasingly important role in empirical economic analysis. On a macroeconomic scale, it has been shown that a nation's average cognitive skill level predicts economic growth more closely than a simple measure of average educational attainment (Hanushek & Woessmann, 2008; Hanushek & Woessmann 2010a). Furthermore, in microeconomics, the well-known ability bias in earnings functions arises because the omission of any measure of ability distorts the measured marginal return of education (Belzil & Hansen, 2002). At both levels, it is evident that the value of education in determining welfare is distorted, because individuals and nations with different (average) abilities have varying capacities to convert "time in education" into productivity.

Cognitive ability is often proxied by scores from numeracy and literacy tests. The education production function literature seeks to determine under which circumstances and with which resources these test scores could improve. By implication, pupils' results represent the outputs and quality of the schooling infrastructure, teaching materials and teachers that promote learning in individuals. Indeed, large cross-country differences occur by these measures. For instance, South Africa ranks poorly, even within the group of developing countries. Within this country itself, large variations exist: the former white schooling system performs close to the international average, while the former black schooling system lags behind substantially (Van der Berg, 2007). Despite fiscal equalisation between the former education systems, these inequalities have persisted well beyond the abolition of the apartheid regime's separate education policies. Education production functions attempt to explain these disparities in performance within countries and groupings.

Some of the more well-known surveys that are generally used include the Trends in International Mathematics and Science Studies (TIMMS) and Programme for International Student Assessment (PISA). Usually the basis of such surveys has a very specific design, with school level data, classroom level information and pupil level variables being collected to represent a specific school-going cohort. Each pupil that is sampled from a carefully constructed sampling frame has to complete the relevant standardised academic assessment. These scientifically determined designs are implemented to obtain reflective indicators of student performance (and by implication ability) across regions, education systems and multiple other groupings. Even these purpose-designed surveys, however, reportedly suffer from comparability issues and sample selection bias (Hanushek & Woessmann, 2010b). The sample may not be representative of the test cohort in the population, because only students who are actually in school are tested[2], some schools (usually those that are located in remote areas) are sometimes excluded from the sample and often serious non-response arises. Each of these features likely biases numeracy scores upwards, as it is generally the case that potentially poorer performing students are excluded from estimates.

[2] This could be a problem where enrollment rates are low; the students who are not in school (and by implication not in the sample) are usually from households with lower socioeconomic status and are consequently less literate and numerate.

The main issue at stake is that, in the attempt to measure the ability of a predefined population through conducting numeracy tests, it is often low ability individuals that are omitted from the sample. This means one cannot hope to measure the whole ability distribution, because much of the bottom tail refuses to or is not able to participate in the test. Essentially the sample is non-randomly selected on the outcome variable of interest, so that a systematic part of the distribution remains unobserved. For instance, suppose that ability measures are to be included in wage functions to correct for omitted variable bias; should the least able individuals in the population (for one or more of the reasons cited above) not be included in the test writing sample, they will also not be included in the wage function. By attempting to solve one source of bias (an omitted variable), another is introduced (censoring the sample on a censored explanatory factor). If this is a problem in targeted (compulsory) numeracy tests conducted using a school sampling frame, it is by assumption a larger issue when these tests are completed on a voluntary basis. The National Income Dynamics Study (NIDS) which is used in this study has some important features that distinguish it substantially from surveys such as TIMMS and PISA. Firstly, the primary aim of the survey is not to measure student performance, but to enumerate welfare, labour market and other household information, so that the numeracy test does not constitute a priority module that is emphasised by survey enumerators. Secondly, the data is not collected within an educational environment, but is targeted at entire households with individuals from a wide age spectrum, whose members may or may not be confident at completing educational assessments. Thirdly and most importantly, the test is completed on a voluntary basis, so that non-response is very high. While typical international assessments use a response rate of 85% as a benchmark for reliability (Hanushek & Woessmann, 2010b), the response rate for the numeracy test in NIDS is almost the complement of that. As a result, it is highly doubtful whether this response pattern is randomly determined, and whether this allows for this measure to be used in subsequent research questions without taking cognisance of this fact in the methodology.

Very little empirical research has been conducted in South Africa or elswhere to examine the effects that sociodemographic, questionnaire-specific and environment-specific factors have on individuals' propensities to participate in *voluntary* assessment. This paper utilises the National Income Dynamics Study (NIDS) that was enumerated by the Southern African Labour and Development Research Unit (SALDRU) in 2008 for the Presidency of South Africa's Programme to Support Pro-Poor Policy Development (PSPPD). Currently the first wave of the data is available, though an individual level panel will be constructed from surveys in 2 year intervals. This data source is rich in its coverage: it contains detailed household expenditure and income items, as well as labour market variables. Of particular interest to this study is the module that asks individuals to volunteer to complete a numeracy test (Griffin et al., 2010). Given that South Africa suffers from severe inequalities in outputs within its educational system, and also performs poorly at all levels of education within even the group of developing countries (Van der Berg, 2007), the analysis of numeracy scores (along with their causes and effects) are of central concern in assessing the linkages between the education system and other inequalities in society. NIDS is one of the first datasets that incorporates this information explicitly in a more general household survey: it allows numeric ability to be connected to labour market outcomes at the individual level and as a result also sheds some insight into the role of scholastic achievement in broader societal inequalities (rather than focussing more narrowly on schools). Given that the numeracy

scores are not only the result of natural ability, but are also connected to varying school quality, such analyses allow a clearer understanding of how South Africa's diverse quality of education has influenced society more broadly.

This particular study, however, abstracts from these concepts. Rather, the interest is in understanding the selection mechanism underlying who wrote the NIDS numeracy test. While the ultimate objective is to use this information to correct for sample selection issues in further estimates, the patterns of response are of interest in themselves. Foremost, it is evident that a racial bias in response rates exists. Indeed, many important indicators in South Africa are skewed along the racial dimension: a greater proportion of blacks obtain poor quality education compared to whites; blacks have higher unemployment rates than whites and are therefore likely to be closer to the edge of the formal labour market. Each of these socioeconomic stratifications could influence test response rates in adverse ways. Poor education could influence the confidence that individuals have in their abilities, and consequently influence their willingness to participate in the numeracy test. High prevalence of discouraged workers amongst certain groups is indicative of low levels of labour market motivation, and by inference also of the willingness to participate in tests. On the other hand, the searching unemployed may exhibit higher levels of motivation to co-operate in the testing module, as they hope to gain in self-knowledge regarding their abilities[3] or perhaps already possess the competency to complete assessments (because the job search process may necessitate them to complete forms and reveal their ability in the first place). Furthermore, the young are strongly represented in the sample of test respondents, as older individuals may not recall their competencies from their time at school and therefore do not possess the confidence to complete the test.

The focus of this study is to uncover how various labour market activities influence co-operation in voluntary testing. Each of these statuses is associated with different psychological influences, perceptions of their well-being and also other opportunity costs (such as time). We postulate that some of these response patterns are linked to individuals' confidence in their own abilities. In addition to identifying which broad socioeconomic characteristics determine why individuals respond to tests, we also take heterogeneous psychological and labour market correlates into account. This enlightens our understanding of why individuals answer tests, how survey design can be altered to improve response rates, and (more interestingly) how various psychological states influence decisions. In this particular case the decision is to participate in a test, which could potentially unveil information to respondents on how able they really are to perform in the labour market. Should they not be confident about their abilities, they may be less likely to want to know this, and should they hope to either find a job or look for a better job (in other words, these are individuals who are also making marginal labour market decisions), they may be more likely to want to agree to measure their abilities.

Section 2 considers the problem of item non-response and sample selection bias in a broader context than numeracy testing. The sources and possible solutions to the issue are discussed at length. Section 3 establishes the possible correlates of test non-response. While confidence in abilities and labour market

[3] They may wish to "test" themselves to gauge their chances in the labour market, based on their assessment of how easily they felt they could complete the test.

issues are not dealt with separately, each of the relevant incentives to respond are related to these concepts where they are relevant. This paper contributes to the literature by explicitly considering how labour market status and psychological aspects influence respondents' motivation to participate in voluntary assessments. Section 4 introduces the data and descriptive analyses. Section 5 presents the results of probit models explaining test response, while section 6 concludes.

2 Item response in theory

Survey response patterns form part of a broad question that influences the reliability of questionnaire-based inference. For instance, sensitive items such as respondents' incomes typically elicit lower response rates amongst wealthier sub-populations, which has implications for estimates of poverty, inequality and returns to education (Ardington et al., 2006). Usually this response pattern is not distributed randomly across the population. We place our discussion within the context of the assumption that response to numeracy tests is also not a random phenomenon, and consequently attempt to understand and model the factors that influence this decision.

The literature on survey design distinguishes between two main types of survey nonresponse.[4] Unit nonresponse occurs when a unit (normally an individual or a household) in the eligible survey sample fails to respond to any of the items in the survey questionnaire.[5] By contrast, item nonresponse occurs when a unit fails to respond only to certain survey items (Gilley & Leone, 1991, p.282). Nonresponse is primarily the result of the survey enumerator's failure or inability to establish contact or effectively communicate with the response unit, the unit's inability or refusal to participate in the survey or respond to a survey item, the loss of data pertaining to the unit's survey answer(s), or a combination of these factors (Sherman, 2000, p.362). The interest of this paper is understanding respondents' decisions in refusing to answer a numeracy test, while we take the other aforementioned issues as given. In this context, item nonresponse is best understood by taking into account the characteristics of those individuals who do respond to surveys.[6] Porst and von Briel (1995, p. 9,10) identify three broad types of survey respondent. Firstly, "altruistic" individuals respond out of a sense of social obligation or in the belief that the survey will be of some benefit to society. Secondly, some individuals are intrigued or interested in some element of the survey, be it the survey topic, the interviewer, or simply the fact that they find the interview to be an enjoyable experience. Thirdly, individuals may respond for a broad array of personal reasons including an inability to refuse the interview, boredom (or having "nothing better to do"), loneliness, or because they feel marginalized and see the survey as an opportunity for their voice to be heard. Each of these classifications is based on the intrinsic motivation of the individual. Below we also discuss extrinsic factors that may (de)motivate individuals to respond to specific survey items.

[4] Additionally, longitudinal studies are also subject to non-response due to attrition. That is, respondents in one wave of a panel study may not be willing, able, or available to respond to follow-up waves of the survey (Marcus & Schütz, 2005, p.1).

[5] Here, an item may refer to a specific question in the questionnaire or a subsection of questions.

[6] Unless stated otherwise, *individuals*, *subjects*, and/or *respondents* are hereafter used interchangeably to refer to individuals who have already chosen to participate in a survey and for whom WTR therefore relates only to specific questionnaire items or sections of items.

Given the broad reasons why individuals may choose to participate in a survey, their subsequent *willingness to respond (henceforth WTR)* to any given questionnaire item can be expressed as an additive and interactive function of a set of survey-specific (for example questionnaire duration, item framing, question difficulty), person-specific (for example age, gender, personal interests and reservations, educational attainment), and environment-specific (for example household and family characteristics, physical and social environment) factors (Singer, 2002, p.165). Not only is it possible for some of these factors to be interrelated, but there may also be some measure of overlap between the three catagories. Furthermore, the weight associated with each individual component and, to a somewhat lesser extent, the direction in which it influences WTR, is unique to each respondent (Singer, 2002, p.166).

While such a conceptualization of item response is intuitively appealing, it may not always be possible to categorize the factors that influence item response as purely person-specific, survey-specific, or environment-specific. However, the current framework does allude to three critical behavioural underpinnnings of item response. Firstly, individuals base their response decisions on composite information. In other words, when deciding whether or not to respond to a certain questionnaire item, subjects do not simply consider the information emanating from the item in isolation, but instead attempt to consolidate information from a variety of sources in order to inform their response decision(s). Secondly, response decisions are based on *real* rather than *nominal* information. In other words, individuals collect, consolidate, interpret, and internalise all *de facto* (nominal) information pertaining to a questionnaire item and base their decisions on the resultent *ad hoc, ad hominem* (real) information.[7] Factors such as the respondent's age, the time spent on the survey leading up to the item in question, the interviewer's body language during the survey interview, and other "cold, hard facts" surrounding the questionnaire item constitute *nominal* information. This is then uniquely processed into *real* information in response to an individual's attitude, motivation and psychological disposition towards the nominal information. Thirdly, and following directly from the second point, even if two individuals were faced with exactly the same nominal information pertaining to a questionnaire item, there is no reason to assume *a priori* that they will face the same (or even similar) real information and, therefore, they may have a different WTR to an item, despite (for instance) being of the same age and race group and living in the same circumstances.[8]

One can simplify this conceptual framework of WTR by arguing that individuals actually face and base their response decisions on only two types of information: information that incentivises response and information that disincentivises response. It is, for example, commonly acknowledged that survey questionnaires impose a burden on survey respondents and that this burden disincentivises sustained item response. Bradburn (1978, p. 36) emphasizes four distinct dimensions of this respondent burden: Firstly interview duration, secondly the amount of effort that the respondent must expend to answer items in the questionnaire, thirdly the frequency with which the respondent is interviewed, and finally

[7] Of course, whether or not a given piece of nominal information can be regarded as relevant to a certain response decision is itself subject to the individual's subjective judgement.

[8] Because respondents differ from one another, the way in which they collect, consolidate, interpret, internalise, and discount the importance of pieces of nominal information are also likely to differ.

the degree of stress the respondent experiences when asked to answer questions of a psychologically disturbing nature or questions that constitute an invasion of privacy. These dimensions are interrelated and, *ceteris paribus*, an increase in one or more of the four should increase the burden on survey respondents and consequently disincentivise co-operation. Many of these factors are related to the intrinsic motivation of the individual (regardless of the nature of the item that is asked), and consequently we attempt to model many intrinsic factors below.

On the other side of response burden, there are also potentially a number of internal and external incentives that may serve to mitigate the extent of the burden, or even cause their effects to operate in an entirely different direction. It is, for example, conceivable that respondents who care a great deal about the topic of a given section in a survey may be willing to invest a far greater amount of both time and effort to answer the questions in it (Sharp & Frankel, 1983, p.38). Similarly, a meta analysis conducted by Singer *et al* (1999) shows that external incentives such as monetary compensation may induce survey participants to answer questionnaire items which they otherwise may not have.

Acknowledging the central role that response burden and incentives play in informing individual's response decisions, the willingnes of individual *n* to respond to item *i* can be expressed by the following function:

$$WTR_{ni} = f(Incentives_{ni}, Response\ burden_{ni})$$

where WTR is increasing in incentives and decreasing in response burden, both of which (as indicated by the subscripts) are themselves dependent on the nature and disposition of the individual and the nature of the item. Respondents will only answer items if their $WTR > 0$ such that they have a net incentive to respond. This assertion is true by construction, irrespective if individuals behave rationally or irrationally. That is, if an individual responds to an item, it must be the case that the incentives to answer the item outweighed the concurrent disincentives.

3 Correlates of test response

Given the heterogeneity of survey respondents and the specificity of a voluntary numeracy test, each sampled individual in the current context responds according to his or her own motivations and environment. The goal of this paper is to identify the broad correlates of response to numeracy testing in the South African context. This is analysed in light of South Africa's diverse demographic profile, its historically stratified education system and the impacts of individual motivation. However, the literature on test response is first consulted. Specifically, this section provides a brief overview of the roles that material compensation, response effort, questionnaire length and labour market status, respondent confidence, personality and emotional well-being, and household effects, social norms, and sociodemographics play in influencing item response decisions.[9]

[9] The attributes and behaviours of survey interviewers are also potentially critical determinants of individuals' response decisions. However, this correlate of item response is not discussed here. For an in-depth analysis of survey interviewer influence and interviewer-interviewee interaction, see Groves et al. (1992).

3.1 Material compensation

Recognizing the extent of the burden that survey questionnaires impose on individuals, it is not uncommon to offer material incentives for co-operation (including monetary compensation and/or gifts). The literature on the effects of such incentives in face-to-face household surveys is vast (see, for example, Singer et al (1999) for a meta-analysis) and finds that, in general, material incentives are highly effective instruments for incentivising response. Moreover, the evidence from survey experiments suggests that material compensation is especially effective at increasing the response rate on items such as literacy tests, which individuals find particularly burdensome (Singer, 2002, p.5; Berlin et al., 1992, p.398). In other words, external rewards may motivate participation when intrinsic motivation is low.

In light of the fact that material incentives may compensate for the extent of the burden imposed by (or the lack of any other incentive to respond to) a certain questionnaire item, one would expect its use to incentivise participation in voluntary assessment tests to be both prevalent and prolific. In fact, Berlin *et al* (1992, p. 398) show that monetary incentives are particularly effective at inducing those respondents with lower levels of literacy to respond to literacy tests. This group is typically among the poorer sections of society, so that a monetary reward would offer the greatest (relative) incentive to participate. Furthermore, the illiterate are likely not to be confident at completing a test. In this context, external rewards are likely to have a great impact in overcoming intrinsic disincentives. In the absence of any external material incentive, it should therefore not be surprising that respondents with lower levels of educational attainment may opt not to undertake voluntary assessment. While literacy and educational attainment have improved to high levels over time in South Africa (Louw et al., 2007), this issue can nevertheless not be ignored. Indeed, the distribution of literacy is non-random, so that older generations (particularly from previously disadvantaged communities) would potentially be less likely to complete numeracy tests in the absence of external incentives.

The data used in this study were collected without offering material incentives. However, the absence of monetary rewards allows for the analysis of what is known in the survey literature as the *zero-incentive group* – those individuals who respond to an item without the need for any additional explicit material incentive (Singer, 2002, p.170). Given the costs of rewarding potential respondents, it is of particular interest to understand why these individuals have, of their own accord, sufficient incentive to respond to tests items whereas others do not. The sections that follow consider some of the factors that influence numeracy test response. Using data consituted by only a zero incentive group allows for the clearer identification of these impacts.

3.2 Response Effort

As discussed in Section 2 above, response burden is rationally assumed to be increasing in response effort (Bradburn, 1978, p.37). Therefore, it is to be expected that the non-response rate will be increasing in the amount of effort required to respond to a section in a questionnaire. Of course, as indicated above, response effort depends on both the item and the respondent in question. Answering certain questionnaire items may require higher levels of effort, or longer sustained periods of effort, than other items.

The majority of items in standard household survey questionnaires mainly involve some form of (memory) recall effort. However, some items may also require retrieval effort (having to consult personal documents and records to respond to items), physical effort (for instance, having one's physical measurements taken), subjective estimation (such as ranking one's own ability to read), deductive and/or inductive reasoning (for example having to estimate one's rank in a country's income distribution), or other forms of mental processes (Blair & Burton, 1987, p.282). In this sense, literacy and/or numeracy tests arguably require a significant amount of response effort from respondents over and above that required by more standard questionnaire items. The NIDS numeracy test, for example, was based on South Africa's national schooling curriculum and was designed to assess questionnaire respondents' levels of numeric, algebraic, measurement, spacial, and data competency (Griffin et al., 2010, p.2). The test is not only intelectually challenging, but required at least ten minutes of willing respondents' time. When compared to other items in the NIDS questionnaires, it would thus have been rational for individuals to expect the response effort associated with participation in the numeracy test module to be comparatively high. This is particularly so for those individuals who were also asked to answer questions in the module targeted at general household issues and those who stood proxy for absent household members.

The extent of the response effort associated with a numeracy test (or any other item) is, of course, not the same for each respondent. Younger respondents would probably find it easier to recall what they learned about mathematics at school or at university (Glazerman et al., 2000, p.20). Similarly, respondents who went to good schools and/or attained high levels of education would be better prepared, and therefore potentially more inclined, to take a numeracy test than those who did not (Chevalier et al., 2008, pp.8,9). The point here is that the perceived level of response effort associated with any item is both subjective and relative. If an individual expects the response effort of answering an item to be relatively high, it will increase the corresponding response burden and, in the absence of sufficient incentives to counteract the burden, lead the respondent not to answer the item.

3.3 Questionnaire length and labour market status[10]

The opportunity costs associated with answering survey questionnaires is one of the most significant components of survey respondent burden (Bradburn, 1978, p.36). This opportunity cost is, of course, greater for some individuals than others. For self-employed or actively searching unemployed individuals, for example, the opportunity cost of responding to all items in a lengthy survey questionnaire may be much greater than for discouraged workers or economically inactve respondents (Green, 1996, p.174). In the case of the self-employed, their labour market income is not guaranteed in the same way as salaried workers and depends on their own time and effort in the workplace. The searching unemployed would presumably rather spend their time and effort in finding a secure job rather than answering surveys. For these groups, there may consequently be an incentive to complete the survey as quickly as possible or to simply leave out items that may be particularly time-consuming.

[10] Here, *questionnaire length* refers to the interview duration (when applicable), the time it takes for the respondent to actually answer survey questions and/or fill in the survey questionnaire, and the number of items and subsections enumerated in the survey questionnaire.

The fact that respondent burden generally increases with the length of a questionnaire, is not only the result of the opportunity cost of response. To understand this, it is necessary to consider both the depth of the survey questionnaire (the number of items and subsections in the survey) and the duration in terms of time that it takes to complete the questionnaire. All other things constant, the greater the number of items in a questionnaire and the longer it takes to respond to them, the more individuals will become fatigued as they progress through the questionnaire. This is due to the repeated and sustained exertion of response effort (Axinn & Pearce, 2006, p.42). Of course, once response fatigue sets in, the response effort associated with answering the remaining items in the questionnaire also increases. Therefore, it matters not only how many items there are in a questionnaire, but also in what order they appear. In fact, Axinn and Pearce (2006, p. 42) argue that, because more cognitively challenging items cause greater response fatigue, it is preferable to put them earlier rather than later in a questionnaire. In the dataset under consideration the exact opposite was true, so that we can clearly gauge the impact of response fatigue by considering the different times respondents spent on answering preceding survey questions and their subsequent response rates in the numeracy test.

Given these considerations, one would expect enumeration of challenging items such as literacy and/or numeracy test modules at the end of a survey questionnaire to have some potentially perverse effects. Firstly, late placement creates the possibility that those respondents who would be inclined to participate in voluntary assessment early on in the questionnaire, will have become too fatigued to do so by the time that the opportunity to participate presents itself. Secondly, since one would rationally expect the marginal opportunity cost of survey participation to increase with the length of the questionnaire, sustained item response towards the end of a questionnaire may appear relatively costly in comparison to response earlier on in the questionnaire. Late placement could thus serve as a further deterrent to item response, particularly for those individuals for whom the opportunity costs of responding to a survey questionnaire was already high in the first place. Therefore, given the relationship between respondent fatigue and response effort, one would expect that, *ceteris paribus*, the response burden associated with taking a numeracy or literacy test would be highest for those respondents for whom the questionnaire duration preceding the test module is the longest.

3.4 Respondent Confidence[11]

Respondent confidence is a potentially critical correlate of response decisions on cognitively demanding questionnaire items since it has a significant impact on the way in which individuals interpret, internalise, and, consequently, respond to other person-specific, survey-specific, and environment-specific factors. Although experimental studies have confirmed the existence of a generally positive relationship between individuals' true abilities and their confidence therein, Dunning et al (1989, p. 1082) find that overconfidence is a much more prevalent behavioural trait than underconfidence, irrespective of the actual underlying level of ability (Dougherty, 2001, p.579). An important implication from this finding is that individuals with high levels of confidence in their ability may have an incentive to behave in a similar fashion to individuals who are actually positioned higher-up in the true ability distribution (Burks et al., 2010, p.1). In terms of behavioural decision-making, therefore, it seems possible that self-confidence could compensate for ability shortfalls.

[11] Here, confidence refers to the level of confidence in one's own cognitive ability.

Confidence may be absolute in the sense that respondents think they are more competent than they actually are, or relative in the sense that they believe that they are more competent than other respondents in their peer group. This latter form of confidence is of particular interest since it reflects a dimension of incomplete knowledge regarding the ability distribution among one's peers and how this incomplete knowledge influences the subjective judgement of one's own ability. In the absence of any *a priori* test of ability (because the numeracy score of non-respondents remains unobserved) and confidence, it may also be easier to measure relative confidence by comparing respondents' subjective judgements of their own abilities against those of other respondents who are similar in terms of other observable characteristics.

In contrast to the assertions of the theory on self-preserving preferences, Burks *et al* (2010, p. 4) find that relatively overconfident individuals are far more likely to seek information about their true abilities than relatively underconfident individuals. However, such behaviour does in fact appear to be rational. By definiton, relatively overconfident individuals should expect any new information about their actual ability to be a confirmation of their own subjective judgement thereof. In this sense, an opportunity to unveil information about one's ability thus presents an opportunity for affirmation of one's perception thereof. Of course, a similar implication also holds true, albeit in a more undesirable sense, for underconfident individuals.

This finding has some important implications for item response decisions on challenging survey questionnaire items such as numeracy or literacy test modules. Firstly, it suggests that, *ceteris paribus*, the expected response effort associated with any given questionnaire item should be lower for those respondents who have greater confidence in their ability to respond to it than for those who lack such confidence. Secondly, it suggests that respondents who have confidence in their (numerical) literacy or their ability to write tests will have an additional incentive to respond to tests since they expect to derive some form of utility from confirming information about their ability. By contrast, respondents who lack confidence may not wish to expose themselves to any expected unpleasant truths regarding their cognitive ability. Therefore, respondent confidence has a certain duality in the sense that, while it may create a strong incentive to respond to a certain item for some respondents, it may contribute to the associated response burden for others.

The previous point becomes even more poignant when one considers the effects that respondent confidence may have on response burden through its effects on the perceived psychological stress and disclosure risk associated with responding to a certain questionnaire item (Couper & Singer, 2009, p.23). It should be clear from the discussion above that the potential for psychologically distressing effects and, consequently, the response burden associated with responding to a challenging item could appear greatest for less confident respondents. For them, the perceived likelihood of receiving some form of negative signal about their own ability is so high that it actually increases the psychological distress associated with participating in voluntary assessment. It may also be the case that the response burden associated with a given level of disclosure risk would be higher for less confident individuals. Put differently, since highly confident individuals may expect any signal about their cognitive ability that is unveiled through participation in voluntary assessment to be positive, their decision to respond to a test module may not be severely deterred by the possibility that other parties may thereby learn of their

actual underlying ability. By contrast, even if underconfident individuals were not excessively concerned that participation in an assessment module would justify their lack of confidence to themselves, they may not wish for such information to become known to others. This point links back to the theoretical discussion in Section 2 above. While respondents may face the same nominal disclosure risk, their levels of confidence influence the perceived real disclosure risk associated with an item and, since respondents base their decisions on this real information, the level of respondent confidence will invariably influence the response burden and subsequent WTR associated with a questionnaire item.

How does confidence then explain numeracy test reponse rates? Is it possible that it decreases the amount of effort individuals need to put into answering the item, or does it simply make respondents care more about answering the test because it may send signals to potential employers or because it serves as a potential 'self-test' for those seeking information about their own ability? In the context of the data under study, individuals' results are not revealed to them or any other person; they are only (anonomously) linked to other person level data. Given that this procedural information is known to respondents before they complete the numeracy test, it is likely that they were either really more confident in answering tests, or that they wish to confirm the perceptions of their own ability by judging how easily they could answer it (rather than judging this by an absent result).

3.5 Personality, Emotional well-being and Motivation

Research has shown that personality traits are among the most significant correlates of survey response decisions (Marcus & Schütz, 2005, p.960). This is not only because personality traits are closely related to individuals' confidence in their own cognitive abilities, but also because they are reflective of underlying values, norms, and motives. For example, while some individuals may respond better to internal loci of motivation, others may be externally motivated. As a result, it is conceivable that certain respondents may attempt to exhibit highly confident behaviour in order to send signals of apparent high underlying ability to other household members, the survey interviewer, or any other party who may have access to the survey questionnaire results (Burks et al., 2010, p.16). Alternatively, it has been shown that some individuals often derive satisfaction merely by successfully completing a task, almost irrespective of the nature of that task, and may therefore have an incentive to participate in voluntary assessment (Brüggen et al., 2007, p.155).

Personality traits do not simply influence respondents' WTR in isolation. It has also been found that respondents' emotional states, especially when some form of face-to-face interaction with an interviewer is involved, have a significant impact on their willingness to participate in the survey questionnaire and their WTR to whatever questionnaire items are applicable to them (Groves et al., 1992, p.479). In the same manner that emotions often dictate consumer decisions, emotional well-being may also influence survey respondents' response behaviour. In fact, in an experiment conducted by Allen at al. (1992, p. 493), it was found that respondents' own self-reported assessments of their emotional states were powerful predictors of their subsequent behaviour. Similarly, Groves et al. (1992, p.485) find that feelings of happiness generally lead to positive survey response decisions whereas feelings of anger and sadness have the opposite effect. The implication is that positive emotional states may serve to mitigate the extent of the response burden associated with a given questionnaire item. By contrast, individuals who suffer from depression or feel like everything is an effort may not be as

inclined to participate in voluntary assessment tests which already require relatively high levels of expected response effort.

3.6 Household effects, social norms, and sociodemographics

Respondents' values, beliefs, interests and attributes are not simply innate components of their personalities. Instead, factors like confidence, motivation, and other behavioural determinants are invariably linked to (and formed by) personal experiences and physical and social environments. In fact, all decisions, including those relating to item response, are functions of certain household and social norms (Brüggen et al., 2007, p.154). These norms govern the formation of intrinsic value judgements on which subsequent behaviours are based and, therefore, alter the feasible choice sets that individuals face (Sunstein, 1996, p.910). Since personal decisions and actions are powerful expressions of social cohension, individuals often *"…act in a manner that is consistent with the [familial or peer-based] social group with which they identify."* (Childers & Rao, 1992, p.198). Gino *et al.* (2009, p. 394) refer to such groups as individuals' "*in-groups*"[12]. In essence, individuals' perceptions of acceptable and appropriate behaviours are formed through observation of what other members in their *in-group* are saying and doing and, in the long run, they themselves internalise the underlying norms that drive those behaviours (Lindbeck, 1997, p.370).

The structure and extent of close interaction between members of an individual's household makes it a good example of an *in-group*. Assuming that individuals align their behavour firstly with those in this close *in-group* and thereafter with others, it is therefore possible that household members will have an incentive to adhere to relatively homogenous underlying values and norms, even when the members themselves differ greatly in terms of personalities and interests. By implication, one may expect to observe certain household or familial peer effects[13] on individuals' WTR. In other words, an individual's WTR to an item may be influenced by other eligible household members' WTR to that item. There are two potential reasons why this could happen. Firstly, a household member's response decision, if observable, may send a signal of the appropriate *in-group* behaviour in that specific context to other household members, thereby influencing their WTR to the item. Alternatively, if household members have sufficiently internalised household norms and values, they may, *ceteris paribus*, have similar incentives and, therefore, similar WTR to an item, even if they do not directly observe each others' response decisions. Irrespective of which of these scenarios apply, the expected outcome is that, on average, individuals will be more likely to respond to a questionnaire item if other houehold members also do so.

While household values and social norms clearly have important implications for survey respondent behaviour in general, they are particularly relevant for response decisions relating to cognitively challenging questionnaire items such as numeracy or literacy tests, where the amount of intrinsic

[12] An "in group" can be seen as the reference group in relation to who a person *wants* to be or associate with, while a peer group has a more *de facto* value, in the sense that those are the people one *does* share similarities with.

[13] The term "peer effect" is used in a broad sense here, as one member of a household may not be a peer of another in the usual sense. For instance, a child is not the peer of a pensioner. However, we understand household peer effects to be the impact that this particular sphere of influence has on individual members' behaviour.

motivation required to counteract the associated expected response burden is relatively high. As mentioned in Section 2 above, individuals have different reasons for choosing to participate in survey quesionnaires. Consequently, respondents may also choose to participate in voluntary assessment for different reasons. Those who associate with *in-groups* with a strong sense of social obligation may place great value on task completion and see participation in an assessment module as the "right thing to do". Individuals from *in-groups* that value a culture of learning, critical thinking and self-reflection may see participation in voluntary assessment as a valuable opportunity for self-assessment. Similarly, individuals who belong to *in-groups* that associate cognitive assessment with feelings of anxiety may have a disincentive to participate in numeracy or literacy tests.

Unfortunately, respondents' personal, household and other *in-group* values and norms may not always be directly observable. However, research has shown that these norms and values are often correlated with the sociodemographic factors that characterise individuals' physical and social environments (Clark, 1983, pp.1-2). Green (1996, p. 180), for example, argues that individuals who earn close to the median income are the most "socialized" towards "pseudoaltruism" and are, as such, inclined to act in an altruistic manner because it brings them personal gratification. Consequently, one would expect that those respondents positioned in the middle of the income distribution of their reference group will have the greatest social incentive to participate in voluntary assessment modules. By contrast, respondents from previously disadvantaged or marginalised communities may have several reasons for not wanting to participate in assessment tests. Firstly, if such individuals received education of an inferior quality relative to those from more affluent backgrounds, they may lack the necessary confidence to take numeracy or literacy tests. Secondly, individuals from disadvantaged communities may attach some stigma to voluntary assessment, especially if participation in such tests fall outside of the social norm. Finally, given their feelings of marginalisation, these individuals may feel that participation in a voluntary assessment module would do little to serve the interests of their broader social reference group.

From the discussion above, it is clear that sociodemographic variables such as gender, race, geographical location, labour market status, position in the income distribution, and educational attainment may be significant determinants of individuals' WTR to literacy and/or numaracy tests. The rest of this paper investigates the role of each of these factors.

4 Data, descriptive statistics and methodology

The numeracy module we analyse formed one part of the NIDS household survey, which covers a broad set of topics and is nationally representative. This data was collected by the Southern African Development Research Unit (SALDRU) as part of the Programme to Support Pro-Poor Policy Development in South Africa (PSPPD) of National Presidency. Eligibility to answer the test was originally intended to target the population between 12 and 72 years old (Griffin et al., 2010). However, the analysis below only considers those aged between 15 and 59. The lowerbound corresponds to that of the definition of the population of working age[14], while the upperbound reflects the age which

[14] This choice is based on the fact that we use many labour market explanations to understand test response, and because this model will be used to inform subsequent labour market analysis.

enumerators adhered to in reality (as is clearly evident in descriptive statistics). The test appeared close to the end of the questionnaire, and as with other items in the survey, was answered on a voluntary basis. Respondents were asked to write a 10 minute numeracy test, consisting of 15 multiple choice questions that evaluated items related to the South African school curriculum[15]. Repondents were given the option to choose between four different test levels, each linked to the highest level of school mathematics they had attained[16]. Results remained confidential and were not revealed to the participants afterwards.

Table 1 indicates that response rates for the numeracy test were not only low overall (at 22%), but that a racial bias is evident in answering the test. The most striking feature is the particularly poor response of Asians and whites, who constitute the wealthier pockets of South African society, and have also traditionally been schooled in the parts of the education system that have been better resourced (both in terms of infrastructure and in skilled teachers). To understand the impacts of numeracy scores on other outcomes, it is therefore firstly imperative to uncover the process underlying which types of individual were likely to participate in the numeracy test, as severe sample selection bias is likely to have occurred.

Table 1 Racial distribution of individuals who wrote the numeracy test (age 15-59). [Source: Own calculations from NIDS (2008)]

	Black	Coloured	Asian	White	Total
No	9,326	1,916	254	748	12,244
	77.00%	78.20%	93.38%	86.57%	78.00%
Yes	2,829	534	18	116	3,497
	23.27%	21.80%	6.62%	13.43%	22.22%
Total	12,155	2,450	272	864	15,741
	100%	100%	100%	100%	100%

NOTES: Figures are unweighted.

It is furthermore evident that an ability bias in response rates arises – despite the fact that the reason the test was conducted was to measure ability in the first place. Table 2 illustrates, by means of a fairly rudimentary Mincerian wage function (without accounting for sample selection or other issues that usually cause bias[17]), that individuals who wrote the test did not necessarily earn statistically different wages on average from those who did not. This is evident in the statistically insignificant dummy variable for which indicates who in the sample responded to the numeracy test. However, the interaction of this variable with education suggests that the marginal *returns* to an additional year of

[15] The coverage included numeration, algebra, measurement, space and data (Griffin et al., 2010, p.2)

[16] Despite this intended channeling of respondents to appropriate test levels, the data reveals that they were given free choice as to which level of the test they wished to take. Some individuals took tests that were too difficult according to the appropriate benchmark, while others took tests that were easier.

[17] Even though one of the ultimate aims of enumerating numeracy tests is to measure ability (to, in turn, reduce bias in returns to education estimates), we simply use differences in returns to education (without controlling for ability) to measure ability differences between groups. These different returns may, however, be driven by a multitude of other unmeasured factors.

education are approximately 3% higher for those who wrote the test relative to those who did not write the test (at a 10% level of statistical significance). This indicates that test respondents exhibit greater abilities to capitalise on additional human capital investments. This group, then, supposedly answered the numeracy test as a result of the confidence that they had in their known abilities. Measurements of ability in the form of the numeracy score are therefore potentially biased, as they do not capture the bottom tail of the ability distribution (at least for those already in the workplace). Despite the above evidence suggesting that response was low amongst traditionally poorer race groups, this result also suggests that groups with higher labour market returns were likely to complete the numeracy test. Hence, there is no linear one-dimensional socioeconomic measure that indicates whether individuals were likely to respond, as both those from relatively poor groups and those that did well in the labour market were in some sense likely to voluntarily take the numeracy test.

Table 2 Simple Mincerian Wage Function. [Source: Own Calculations from NIDS (2008)]

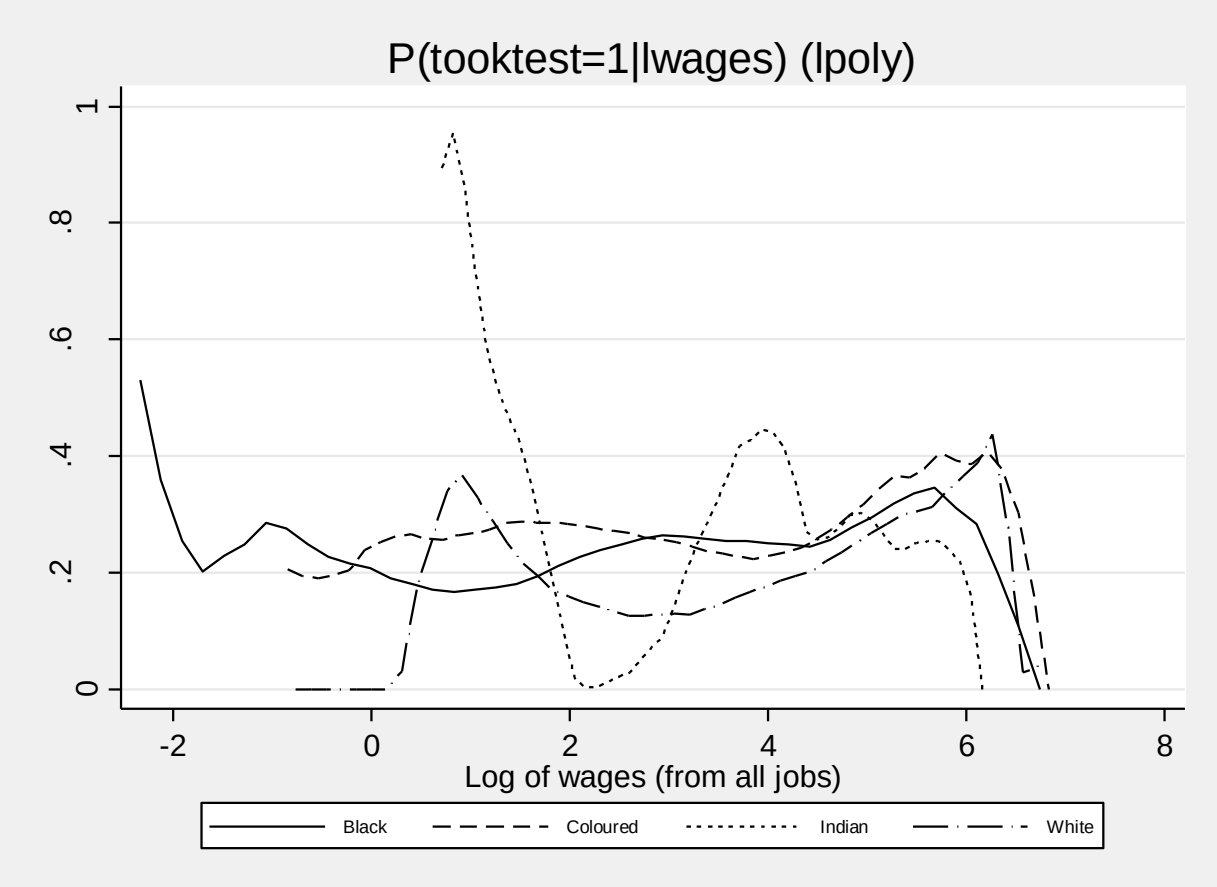
	log(Wages)	
Wrote Test	-0.304	
Education	0.122	***
Education * (Wrote Test)	0.032	*
Age	0.064	***
Age^2	-0.001	***
Female	-0.316	***
Coloured	0.044	
Indian	0.562	***
White	0.730	***
Constant	-0.061	
N	4695	
R-squared	0.297	
P(F>f)	0.000	

NOTES: *Significant at the 10% level **Significant at the 5% level ***Significant at the 1% level. Estimates are weighted. Significance levels are based on robust standard errors. “Wrote Test” is a dummy variable indicating whether an individual completed the numeracy test.

Given the broad socioeconomic dimension of response identified above, NIDS is well-suited to uncover the correlates of willingness to respond. Figure 1 considers the response rates of the employed according to their position in the wage distribution. For all race groups (except coloureds) response is particularly high towards the bottom of the wage range. The different absolute position of this bottom range for each race suggests that relative wealth (within a reference race group) determines response rather than absolute wealth. In particular, those who are relatively close to the bottom edge of the labour market (where “bottom” is defined within the race group), are more prone to answering the numeracy test. This is more pronounced for the black population, suggesting that the poorest among the employed are more likely to participate relative to somewhat richer groups. However, response again increases for the very rich in all groups. High response rates for the rich could potentially be explained by “pseudoaltruistic” behaviour, while the poorest wage earners find themselves at the edge

of the labour market. It is possible that some form of motivation among this group exists to establish or confirm their abilities (at least to themselves), in the hope that they could "move up the job queue".

Figure 1 Probability of taking test by wage level. [Source: Own calculations from NIDS (2008)]



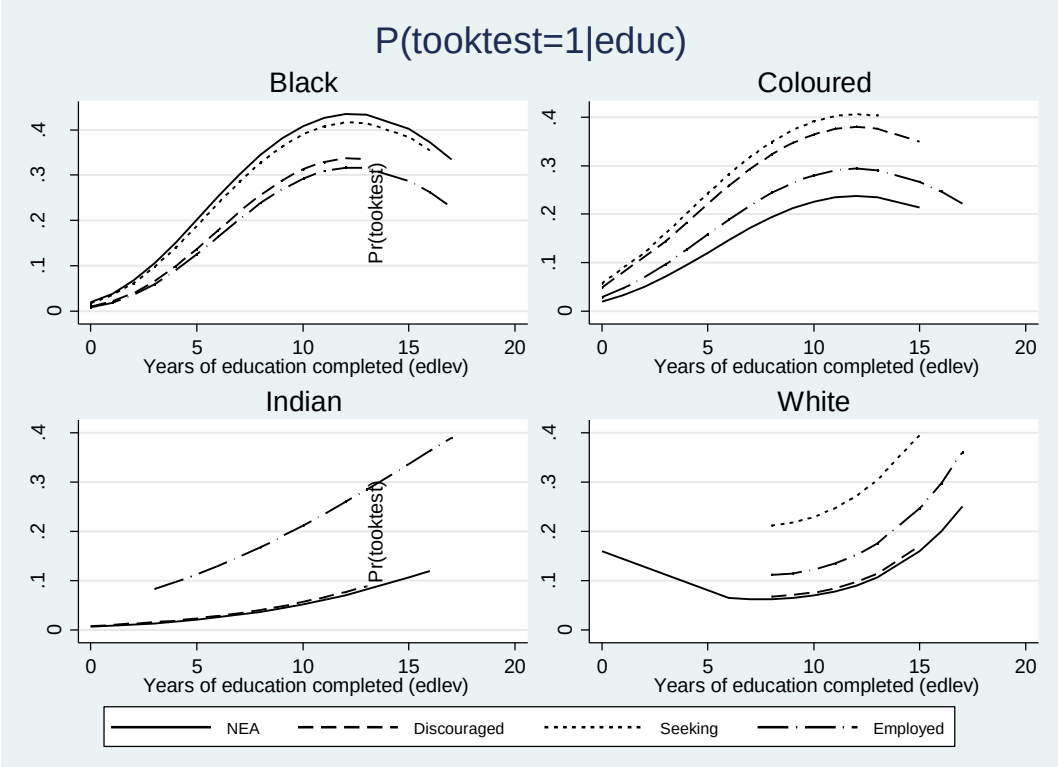
NOTES: The local polynomial smoother used the Epanechnikov kernel with the default bandwidth. Results are weighted

To broaden the analysis beyond just the employed, we also consider those in other sections of the labour market. Figures 2 and 3 highlight that for all race groups (except Indians), the searching unemployed are the most likely to complete the test. This group is the nearest to being employed (in a latent sense, and compared to the not economically active and discouraged workers). Interestingly, the economically inactive and discouraged workers exhibit a relatively low willingness to answer numeracy tests (though to varying degrees among the different race groups). This indicates that the cost of time is not the only factor driving the participation decision, but that motivation (which drives both labour market and test participation) is at play, so that it may become necessary to measure this explicitly in the subsequent models.

The literature has conjectured that the searching unemployed should presumably be less likely to participate in voluntary tests (Green, 1996, p.173), though this is not supported by the NIDS data.

However, the searching unemployed in South Africa may be unique for a number of reasons. Firstly, it may take longer for the job seekers to be absorbed into the labour market given the high unemployment rate in South Africa – as a result the cost of time amongst the searching unemployed is not as high compared to other countries where the transition into the workplace is faster. Secondly, given the high prevalence of discouraged workers in this tight labour market, it is evident that embarking on the job search process requires higher than average levels of motivation (relative to other unemployed individuals) (Kingdon & Knight, 2000, pp.4-9). It is, furthermore, possible that the searching unemployed are the overconfident individuals referred to in section 3.4. By implication these stronger levels of motivation also prompt the voluntary participation in the numeracy test. Thirdly, given that these individuals are in the process of looking for employment, they may value opportunities to "practise" responding to challenging questions. Participation in a numeracy test may be a low-risk simulation of what they could expect to experience during a job interview or a psychometric test. Fourthly, it may be possible that – since the searching unemployed have become accustomed to filling out application forms and submitting their CVs to companies – they may find answering test modules less burdensome than other unemployed and inactive individuals. The latter statement is, however, less likely for individuals in the market for unskilled work, where job search is less paper intensive.

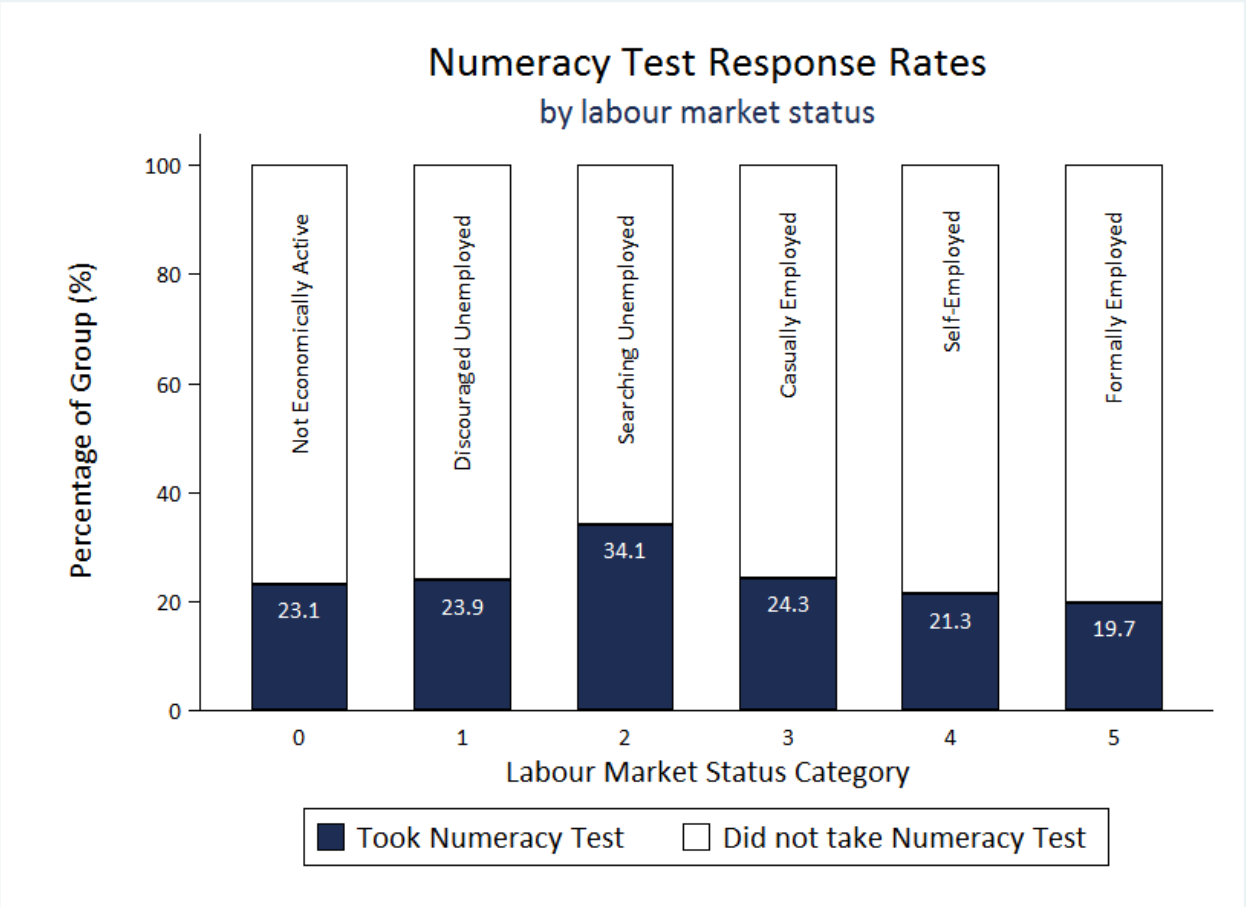
Figure 2 Probability of taking test, by labour market status and educational attainment. [Source: Own Calculations from NIDS (2008)]



NOTES: these figures were generated from the predictions of a probit model, controlling for a quadratic in education interacted with labour market status

Figure 3 reveals that the self-employed have only a slightly higher numeracy test response rate compared to the formally employed (who are the least likely to answer the test). For the latter, a high time cost in answering the survey exists. The former should, however, by expectation have lower response rates, given that they do not have guaranteed salaries and tend to invest time in succeeding in their own income generating ventures.

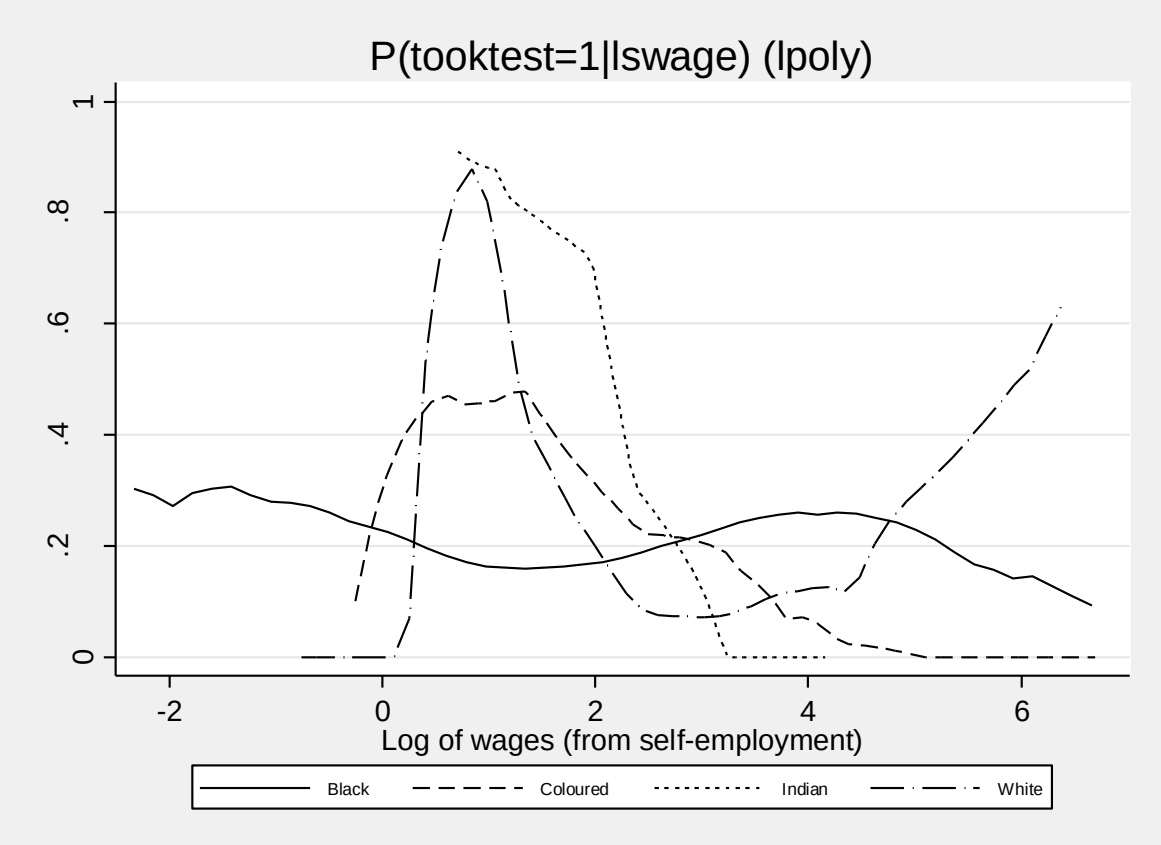
Figure 3 Numeracy Test Response Rates by Labour Market Status [Source: own calculations from NIDS (2008)]



NOTES: Figures are weighted

Figure 4 shows that the black population earning incomes only from self-employment do not exhibit the same behaviour as evidenced in the pattern observed when taking all labour market income sources into account (in other words, earnings from self-employment, formal work and casual work): those at the bottom of the self-employed wage distribution are unlikely to answer the test. It is evident that these individuals face higher costs to take tests, as they have low earnings, but have to generate it themselves. They do not find themselves in a job queue in which they could prove their ability to a prospective employer to become upward mobile. In contrast, casual workers may not face the same high stakes to work longer hours, and may be more likely to respond to the numeracy test.

Figure 4 Probability of taking test by wage level (self-employed only). [Source: Own calculations from NIDS (2008)]



To measure these socioeconomic dimensions, we construct a categorical variable that captures (as far as possible) the latent labour market status of individuals. Standard labour market classification splits the working age population[18] into the inactive, discouraged workers, the searching unemployed and those working. For each of these groups the cost of embarking on an intrusive and lengthy question in the survey instrument may differ. However, as highlighted above, the employed are also not a homogenous group, so that clearer categorical sub-definitions are introduced for this grouping. Costs of completing a numeracy test may differ for high and low earners, as is evident in the preceding discussion. The latter presumably work fewer hours and may have a greater incentive to prove their ability (even though potential employers do not view their results) relative to the former. As a result, they would be willing to substitute some leisure time to reveal their ability by taking the numeracy test. We therefore

[18] This paper only includes individuals that are between the ages of 15 and 59, though the International Labour Organisation classification extends this to those who are up to 65 years old. NIDS questionnaires indicate that adult numeracy tests were only supposed to be administered to individuals of the narrower age range. While the NIDS technical notes indicate that the initial design was to include individuals up to the age of 72 (Griffin et al., 2010, p.1), descriptive statistics show that response was an exception beyond the age of 59. As a result, the final questionnaire rather than the initial guideline was more likely followed by interviewers.

construct a dummy variable that categorises the employed into wage quintiles (by race)[19]. This introduces a crude proxy for the "distance from unemployment or inactivity" and how far individuals are from the "edge" of the labour market. We hypothesise that those at the "edge" of the labour market are more likely to answer the test relative to other groups: the searching unemployed are first of all likely to wish to reveal or practise their abilities and also have lower time costs of participation relative to the employed (though at least some positive time cost of job search should be taken into consideration). Those who earn lower wages are also more likely to wish to reveal or establish an indication of their abilities in the subconscious hope of becoming upwardly mobile. We also expand the extended labour market variable by interacting the wage quintiles with the corresponding employment types: casual employment, self-employment and formal employment, as different incentives are likely to exist within each category.

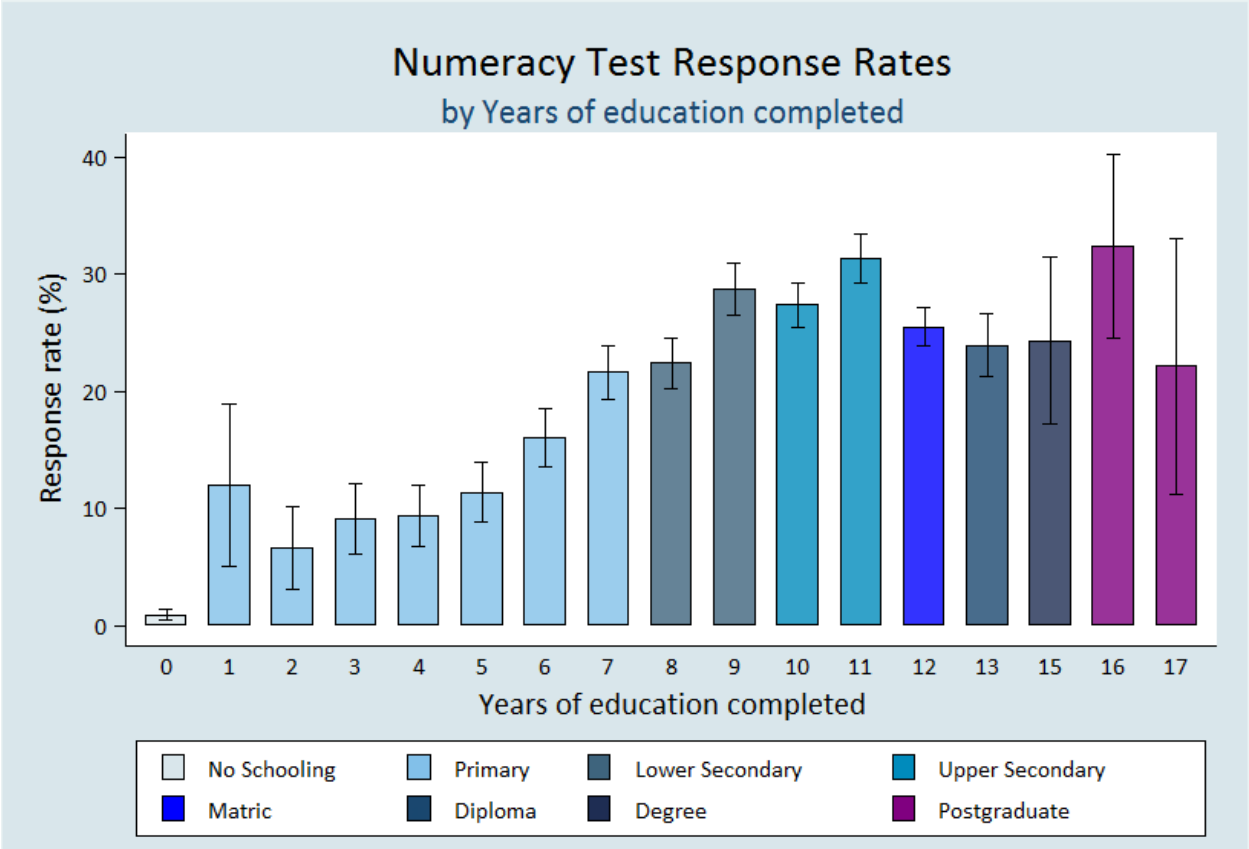
In addition, we control for area of residence. NIDS classifies urban areas into respective formal and informal sub-regions, while rural areas are split into formal and tribal authorities. This classification also captures whether individuals are integrated into the formal labour market, or whether they are at the margin. We also control for education levels to capture a further socioeconomic element. Given that South Africa's younger generations possess higher levels of educational attainment than older cohorts, it is also likely that they may exhibit higher levels of confidence in answering numeracy tests (as a result of their apparently higher levels of human capital). However, the young are also more likely to be at the margin of the labour market. While these issues have already been accounted for in the variables mentioned above, we introduce a full set of explanatory factors, including the age of respondents.

While educational attainment may only be a rough proxy for confidence, it is also a signal of ability. In Figure 2 it is evident that higher levels of education increase the probability of taking the test. The relationship is, however, concave for the black and coloured populations. This suggests that individuals with the highest levels of education do not base response decisions solely on confidence and ability (which should presumably be highest amongst this group), but that they also experience high opportunity costs of time relative to less educated individuals (given that leisure time is presumably more costly for the most educated). This is also evident in Figure 5, where response rates flatten off and even decline after the completion of secondary school. Greater variation exists within the most highly educated groups, suggesting that behaviour is not uniform among this group. For the white and Indian populations, in particular, the relationship is convex (Figure 2), suggesting that the most educated within this category are *more* likely to participate. Typically this is the result of strong feelings about the value of testing and the use of social surveys in general. This is an indication that "pseudoaltruism" and intrinsic motivation are of greater relevance to these groups. The international literature suggests,

[19] We choose racial quintiles, since time costs of participation are likely to be calculated relative to the wages that individuals could achieve within their own reference group. For instance, a white individual in the bottom (white) quintile (but in the top quintile of the entire population) would weigh up the cost of survey participation against job search (or leisure) to obtain the wage of similar white compatriots in higher quintiles. Figure 1 illustrates that relative wage levels within groups have a greater influence than the absolute level thereof. This strategy also introduces more variation in the data, as the bulk of white respondents in the survey are part of the top overall quintile – yet many within this group chose to complete the numeracy test.

however, that the pattern in the former groups is generally true, with response increasing until tertiary education is reached, after which it tapers off (Green, 1996, p.174).

Figure 5 Numeracy Test Response Rates by Education Level. [Source: Own calculations from NIDS (2008)]



NOTES: Figures are weighted and bars are surrounded by 95% confidence intervals

However, we also wish to explicitly control for confidence in abilities rather than proxy for it by education, which evidently also captures other forms of motivation and incentives. To do this, the subjective information that is collected in NIDS is exploited. Respondents were asked to rate both their ability to read and write on a scale from very poorly to very well. Two variants of the question are enumerated for each of reading and writing: firstly respondents are quizzed on their perceived ability in using their home language, and secondly with respect to English. The first is a measure of how competent they feel they are in overall communication, as this is likely to be the first language that individuals receive instruction in. The second measure is of importance, as the language in which the NIDS numeracy test was conducted was English, which is also the *lingua franca* of South Africa. Should non-English speakers not be confident in their English abilities, it would influence their willingness to participate in the test. Along other dimensions, reading is considered the more basic form of communication relative to writing (Crowhurst, 1991, p.316), so that we can split our analysis into confidence in more basic abilities (reading) and more advanced cognitive abilities (writing). However, we are not interested in individuals' absolute confidence in their ability, but how they rate their abilities

relative to a typical reference group. To do this, two absolute confidence in literary ability indices (one for writing and one for reading)[20] are first created by multiple correspondence analysis (MCA) (see Booysen et al. (2008) for an application), which summarises the variation found in a set of categorical variables in a similar fashion to a principal components' analysis[21]. Indices generated from the first component of the MCA were regressed on a number of explanatory variables, including demographic features and education levels. The fit of these regressions was particularly high, so that absolute confidence was particularly well-explained by the explanatory factors. Residuals from these regressions now represent the deviation of individuals' confidence from the average level of confidence in their reference group (represented by the levels of the covariates). As a result, they represent relative confidence. Positive residuals signify overconfidence in relative literary abilities, while negative residuals indicate the converse.

Similarly, we also construct an index of (absolute) emotional well-being and motivation, with a number of ratings combined into the composite variable[22]. These include how often individuals felt bothered, were depressed, considered it an effort to complete tasks, felt restless or lonely or were lethargic. Given the constituent variables, the index captures intrinsic motivation, rather than individuals' confidence in their own abilities (as above). The loadings on the index suggest that the resulting index reflects a measure of well-being and motivation.

Yet another variant of confidence is captured in the influence of family or household members. Should a large proportion of eligible test writers in a household choose to write the test, the potential stigma of writing the test is greatly reduced. In contrast, individuals may in fact experience pressure to take the test. Furthermore, should a culture of learning prevail within a household, familial peer effects will not only influence scholastic achievement levels, but in the first instance whether individuals are likely to participate in voluntary tests. We explicitly control for these effects by introducing a variable that captures the proportion of eligible individuals (other than the relevant respondent) who chose to take the test.

The discussion above alludes to the large opportunity costs of completing survey questionnaires. NIDS is a rich dataset with a wide variety of modules. Household heads were asked to complete a questionnaire relating to overall household circumstances before individual level information was collected. Typically adults were also asked to provide information on behalf of children and adults that were absent from the household. Furthermore, individuals were asked to provide extensive information on health issues.

[20] Henceforth we refer to this more generally as confidence, though to be clear, it does only measure literacy confidence.

[21] At first an index with each of the four categorical ratings was created (reading English, writing English, reading home language and writing home language). The index however accounted for only a small proportion of the total variation in the data. Hence, two sub-indices were created: one considered both the reading variables and the other both the writing variables. The resulting indices yielded satisfactory proxies for the constituent variables. The loadings on the categories indicated that the index captured measures of absolute "non-confidence", and was subsequently multiplied by minus one to obtain a measure of absolute confidence. The reading index indicates basic skill, while the writing index represents more advanced skills.

[22] We do not convert the well-being/ motivation index to a relative measure, as the regression models used to explain absolute well-being were of a particularly poor fit.

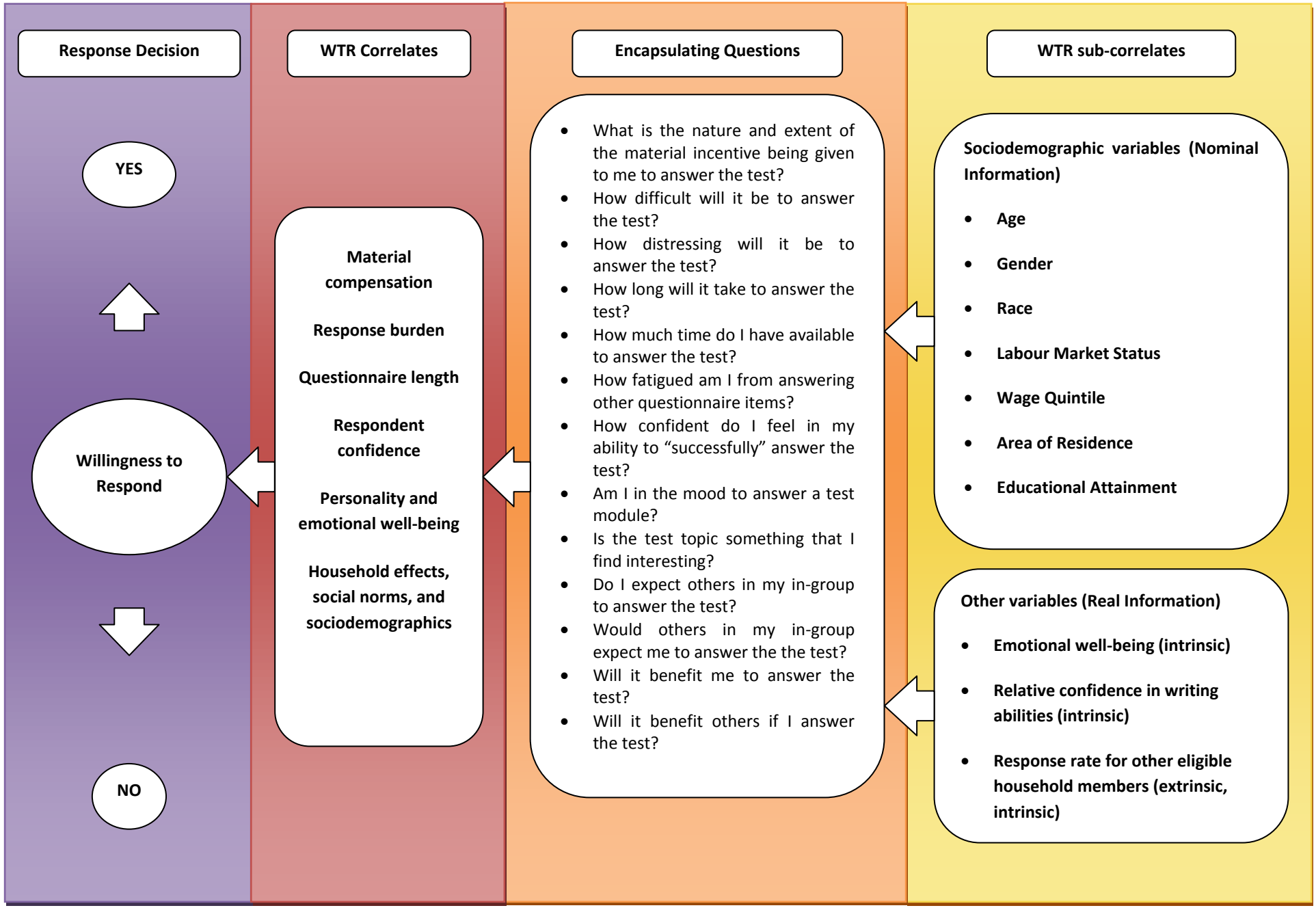
This included taking biometric measurements, which was done just before the numeracy test was administered to respondents[23]. It is clear that respondent fatigue is a real threat in this survey, and that time opportunity costs are potentially high. For adult respondents we calculate (as close as possible) the interview duration before the numeracy test was administered. The time spent on completing the household questionnaire was added to the time spent on adult level responses (of relevant individuals)[24]. Ten minutes was subtracted from this figure for those who indeed wrote the test – this was the recommended length of the test, and the duration variable was purged from this time to eliminate the built-in positive correlation between the longer duration resulting from taking the test and the test time itself. The resultant variable is therefore the time that elapsed before the test was administered, which serves as the basis for individuals' decision to continue with this extended part of the survey.

To align the features outlined above with the theoretical discussion of item response and item response correlates in Sections 2, 3, and 4, the decision to participate in the NIDS numeracy test module is summarized in a decision tree shown in Figure 6. The figure relates each of the explanatory variables outlined in this section to the broad correlates of item response as discussed in Section 3 through a series of encapsulating questions that ultimately serve to inform individuals' response decisions. Based on this decision tree, probit models were constructed to capture each of the features outlined above. Given that the descriptive evidence reveals slightly different behavioural patterns for each race group, separate models are also built along these lines.

[23] Initially a dummy variable to indicate whether respondents' biometric details were measured was included in our models to capture the effects of respondent fatigue. While this is a strong predictor of whether individuals continue with the numeracy test, it is a proxy for duration effects which are better captured with continuous measures.

[24] While one should also add time spent on child questionnaires and those for proxy adults, the end times of these particular variables were poorly recorded, so that this was not feasible.

Figure 6 NIDS numeracy test response decision tree



5 Results

Tables 4 to 7 in the appendix present the results of the probit analysis. The first set of results considers all races in one sample, while the subsequent output represents separate samples for each group[25]. In each case, the first column considers only the extended labour market status as a determinant of taking the numeracy test. The second column proceeds to control for educational attainment, which is correlated with the first set of variables *a priori*, but more closely proxies for the potential ability confidence which respondents would display in answering the test. In addition, this model accounts for age and location effects, which are also correlated with labour market status. Thirdly, the educational proxy is removed and the explicit confidence measures are introduced. We only control for relative writing confidence, because the coefficients on relative reading confidence were consistently statistically insignificant, small in magnitude and collinear with the other confidence measures (which distorted conclusions). Respondents' confidence in their own writing abilities is more decisive in determining whether they write the test, as this skill is slightly more advanced than reading and requires processing of thoughts over and above that required by reading (Crowhurst, 1991, p.316). Fourthly, duration and household peer effects are added to the model, while the fifth model includes all variables simultaneously. In the larger models we attempt to see whether real information (in other words, individual responses to confidence, emotional well-being and other opportunity costs) alters results on nominal information (such as the current labour market status and location of individuals). If this is the case, it means that nominal factors are also influenced by individual psychological responses; this explains why individuals that are otherwise homogenous may nevertheless exercise different choices in responding to the survey.

5.1 Whole population

For the entire population (Table 4) it is evident that the searching unemployed are either just as likely as or more likely to take the test than the economically inactive (the reference group in the model). This is the only statistically significantly positive variable among this set of covariates and the result is consistent across all specifications. Therefore those who are seeking to enter the workplace are most likely to participate in the test, which confirms the descriptive evidence presented above. This result is robust, even when controlling for confidence and emotional well-being. Consequently we cannot assert that the searching unemployed are more motivated, but it suggests that lower relative response burdens and willingness to practise answering challenging questions may be of greater relevance for this group.

Discouraged workers and the self-employed in the lower quintiles[26] are also less likely to complete the test than inactive individuals. In the first instance, discouraged workers have the least incentive to want to reveal their abilities or practise their cognitive skills, as they have (temporarily) given up on the job search process. Indeed, in the specifications where the emotional well-being index becomes statistically

[25] The Indian subsample was not analysed separately due to restrictive sample sizes in NIDS.

[26] The insignificance of the coefficients on the first and third wage quintiles for this group is rather a function of the low concentration of self-employed individuals in the bottom quintile, rather than having any real economic meaning.

significant, the significance of the discouraged worker dummy disappears, suggesting a strong correlation between this status and motivation. Secondly, the self-employed with the lowest wages are more averse to embarking on additional time consuming tasks, as their incomes are less secure than those of salaried workers. This effect, however, still persists strongly once interview duration is controlled for, so that time opportunity costs do not tell the entire story for this group. However, the magnitude of the coefficient declines, which suggests that this explanation does carry some weight.

No clear patterns emerge for casual workers, while (predominantly) salaried employees all have a relative aversion to answering the test. The coefficients are largest (in absolute values) for those in the lowest quintiles, which is closer to the situation experienced by the self-employed, and masks the tendencies identified in Figure 1. However, once controlling for interview duration, the highest quintiles respond no differently than the economically inactive, suggesting that the time taken to write the test is the greatest deterrent for high earners. Time opportunity costs are therefore relevant for the wealthiest.

The aforementioned results remain robust once controlling for location, education and age. These variables, though correlated with labour market outcomes, make an independent contribution in explaining test response. Those living in tribal authorities are least likely to complete tests, while inhabitants of urban and rural formal areas are equally likely to respond. Individuals in urban informal settlements have the highest response rates. Despite having controlled for labour market status, this serves as an indicator that individuals at the margin of the labour market are more willing to complete a test of ability. Given that informal settlements in urban areas absorb much of the migrant labour from rural areas, this group represents individuals who are intent on either finding jobs soon in their new location or who wish to seek out better opportunities. In contrast, those living in tribal authorities are the most divorced from the formal economy, so that they do not have the same incentive to display their abilities.

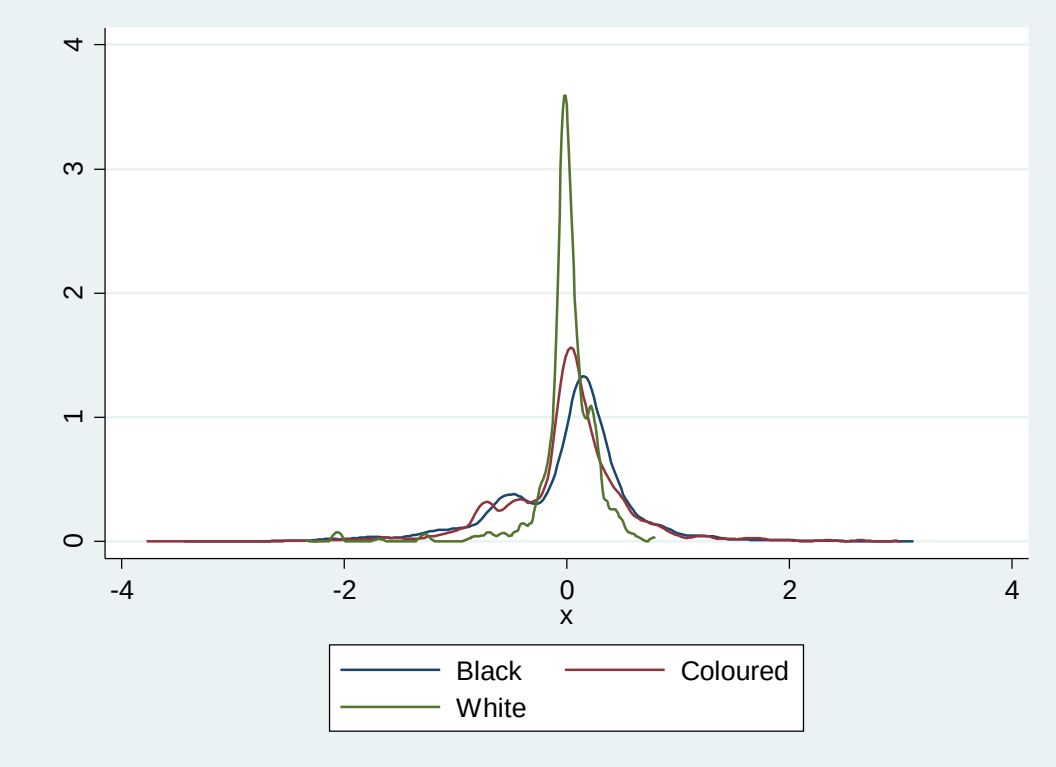
Education initially appears to be a proxy for confidence in abilities (column 2) – the higher attainment, the greater is the coefficient. However, controlling for the relative confidence measure (column 5) actually increases the impact rather than reducing it. Therefore educational attainment does not appear to model the same process that perceived relative ability does. An alternative explanation is that individuals with higher levels of education are more likely to care about the value of numeracy and are more co-operative in contributing to the results of such studies.

The age profile is convex, regardless of the model specified. This suggests that the young are likely to participate, but as individuals age this propensity declines and gradually increases again. The estimates reveal that test participation initially declines, but increases again after the age of 50. Again, the young are likely to be new entrants to the labour market, who wish to display their skills and possibly gain momentum in the job search process. Furthermore, the youngest are likely to still be in school, and most accustomed to answering tests, suggesting that they should have higher confidence in their cognitive abilities.

Absolute emotional well-being positively influences test participation, except when a full set of controls is introduced. This therefore serves as an indicator of intrinsic motivation, but is likely to be influenced by (or correlated with) other indicators, such as household peer effects and interview duration.

Relative writing confidence has no statistically significant influence on participating in the test, except if we control for education. This result is clarified by noting that the impact was negative (and insignificant) for the white subpopulation (Table 7) and insignificant for the coloured subpopulation (Table 6). However, for the black population the positive impact is consistently large and significant, regardless of specification (Table 5). A possible explanation for the poor explanatory power for whites is the narrower distribution of confidence amongst this population (Figure 6): it is highly concentrated around zero (which represents realistic perceptions relative to each individual's reference group). In contrast, the other race groups have widely dispersed confidence levels, with a tendency towards asymmetry (where underconfidence dominates, in contrast to what has been commonly found in the literature (see Holzberg (1989, p. 1082)). The hypothesis that relative writing confidence influences test participation is not dismissed by the apparent statistical insignificance of the estimates, but requires more nuanced analysis by race group.

Figure 6 Density of Relative Writing Confidence Measure: By Race. [Source: Own Calculations from NIDS (2008)]



The time that elapsed before the test was taken has a statistically strong influence on whether individuals agree to be tested, regardless of other controls introduced. The relationship is negative, but diminishes in size as the test time increases. The marginal impact only reaches zero after 322 to 358

minutes (depending on the specification), so that the effect is for all intents and purposes negative across the entire distribution. This confirms *a priori* expectations that long interviews cause respondent fatigue and discourage individuals to complete the test. However, different racial patterns exist in this regard and will be analysed below.

Household peer effects are strong and persist regardless of specification. This suggests that despite the many covariates that measure intrinsic motivation of individuals to participate, environmental factors play a dominant role. Therefore, even individuals that are not individually "confident" by any measures explained above, will participate in the test if a culture of learning[27] exists within the household or community, or where there is significant peer pressure to do so (or where the peer pressure not to participate is absent). This is an important result, as it can be extrapolated to other situations, where a culture of learning may be able to improve participation in educational and training programmes.

5.2 Racial Patterns and Anomalies

Tables 5 to 7 reproduce the same results, but this time delimiting the sample to specific population groups. The reason for this approach is because of the apparently different processes that prevail in the various groups (as is evident in Figures 1 to 4). We do not analyse the Indian population separately, as sample sizes become restrictively small.

It is first of all evident in Table 5 that the results for the black population strongly mirror those of the overall population, which is not surprising given that this is the majority group. However, one aspect that stands out is the consistently stronger influence of writing confidence on test participation. Given the asymmetry and high dispersion of writing confidence identified in Figure 6, it is evident that this population group leans towards being pessimistic about its writing abilities relative to its peers (as represented by the covariates used to construct the reference group for the relative confidence measures), and some individuals are more extreme in this position. Hence, the greater variation in the data allows us to clearly discern a confidence impact in participation; those of relatively high rank by this measure are in smaller concentration than those of low rank and are likely to have had exposure to better learning stimuli, which in turn prompt confidence in answering tests. The first suspicion would be that the relatively fewer black individuals of very high confidence levels are those that have been exposed to better quality (formerly white) schools. This would also explain why the confidence effect is not evident for (particularly) the white estimates of test participation, as most individuals would have received better quality education. However, it is evident in Table 3 that confidence has a negative relationship with school quality, so that this is not a satisfactory explanation. Conditioning on education levels does not alter the picture. Further investigation is warranted to understand why confidence is only significant for the black population.

[27] While we cannot explicitly measure a culture of learning, high response rates at the household level are assumed to proxy for this factor.

Table 3 Mean Confidence Level by Level of School Quality [Source: Own calculations from NIDS (2008)]

School Quality Index	Mean Confidence Level	Std. Dev.	Freq.
0	0.039563	0.459395	672
0.05	0.018334	0.430944	1384
0.1	0.029315	0.436327	1237
0.15	-0.02583	0.466249	467
0.2	-0.08552	0.462607	211
0.25	-0.09147	0.537654	247
0.3	-0.05922	0.329685	104
0.35	-0.10367	0.318967	24
0.4	-0.01294	0.069687	2
0.45	-0.47823	0.097886	2

NOTES: Confidence Levels are defined by the relative confidence index created from the residuals of the regression explaining the first component of the Multiple Correspondence Analysis, as discussed above. A sub-sample of NIDS also indicated which school individuals matriculated from. A school quality index was created by SALDRU and Cobus Burger of Stellenbosch University, based on historical matric results of the various schools. This measure has a mean of zero and a standard deviation of one in the matric population. This is not the case in the matched sample. The index was rounded to the second decimal place by the data compilers so that individual schools could not be identified on the basis of scores.

The coloured population has slightly different results to the black population. Firstly, discouraged workers are more likely to participate than the economically inactive. Again, the searching unemployed have a strong incentive to write the test. Interestingly it is the wealthy formally employed individuals who are more likely to write the test (in contrast to the black population). Furthermore, those living in urban informal areas are also less likely to participate than those in rural formal areas. Here the difference arises because informal urban areas are not a point of entry into the labour market for the coloured population (as it is for many black migrants). Rather, coloured individuals in urban areas usually live in formal dwellings. Those in informal dwellings are therefore more removed from the labour market than is the case for the black population[28]. Apart from this, most results agree with the black model. In particular, strong household peer effects dominate. However, age, confidence and emotional well-being do not enter with any significance.

For the white estimates (Table 7), very few of the labour market categories exhibit a different propensity to participate compared to the economically inactive. Only the wealthiest amongst the casually employed have a lower likelihood of writing the test. Most of the coefficients in the white model appear to be imprecisely estimated. This may be a function of the small sample size, or the lack of overall variation in most of the covariates within this group. Indeed, many of the cells of categorical variables are omitted from estimates, as few white individuals fall within the relevant classifications. However, some notable effects appear. White individuals with secondary education have a lower propensity to participate relative to those with post-school training. Education is highly correlated with labour market status, hence the insignificance of those results. In the case of the black population both sets of variables provided independent explanations of test participation. However, here only the fact that

[28] Also note that few Coloured individuals live in tribal authorities, so that this coefficient is not precisely estimated.

individuals are highly educated plays a role. The most notable other determinant for the white population is the duration of the interview. However, instead of a concave relationship (as for the other population groups), the relationship is positive until a turning point of 357 or 324 minutes, depending on the specification. This suggests that survey fatigue has no role to play within the white population group, which is contrary to expectations. Rather, individuals that have already invested large amounts of their time in completing the first part of the questionnaire are more likely to continue with the numeracy test. This suggests that, within the white population, a group of respondents was likely to complete a long survey regardless of fatigue and did not consider the additional time as a substantial cost. Rather, they viewed the time already invested in completing the first sections of the survey as wasted if the test was not also completed. Hence, these individuals are likely to value social surveys or educational instruments highly and wish to provide as complete information to the data collectors as possible. The marginal time cost of completing the test was small in relation to the perceived marginal benefits (each in turn a function of the substantial time already spent on the questionnaire) of completing the numeracy test. Individuals who did not complete the test were those that did not value completing the first parts of the survey in the first instance, suggesting that they did not place much emphasis on surveys in the first place. Hence, a pseudoaltruistic and social consciousness element arises here, in contrast to the survey fatigue generally expected from the wealthier section of the population. However, given the low response rates for the white population overall, this pseudoaltruistic group is only small so that one cannot generalise this property to the whole of this group.

6 Conclusion

While Hanushek & Woessmann (2010b) find that sample selection issues do not bias further analysis[29] using cognitive assessment scores, the international school surveys that they use usually have response rates in excess of 85%. Given the low response rates evidenced in the NIDS numeracy test and the clear socioeconomic dimensions along which these response patterns have been formed, it is necessary to take cognizance of the fact that, in this case, sample selection bias may be a serious concern. In particular, should the numeracy variable be used to proxy for ability in micro level earnings functions, the sample will be censored along multiple non-random dimensions. Voluntary assessments are likely to suffer a greater extent of non-response. Furthermore, because the NIDS numeracy test was not conducted in schools, household members that were not accustomed to testing formed part of the sample. As a result, low confidence levels coupled with the option not to write the test affect the decision to respond in a non-random manner. This study has investigated some of these issues by considering various dimensions along which response patterns formed.

Firstly, it appears that individuals who took the test generally have higher returns to education – when introducing ability as a variable to reduce the upward omitted variable bias on the returns to education, the correction brought about by this proxy could be offset because the sample is limited to higher ability, better earning individuals. Secondly, a further censoring process in another part of the distribution arises: in this case, individuals seeking to enter the workplace exhibit higher propensities to take the test, even though this subpopulation may not fall within the top of the ability distribution.

[29] Their focus is on educational production functions and cross-country growth regressions.

Hence, the selection process is non-linear and requires careful modelling. However, as Wooldridge (2009, p.323) notes, the selection on the independent variable may be exogenous, so that selection bias could potentially be obsolete.

This paper has shown that labour market correlates and heterogeneous psychological responses are effective at modelling this selection process. In particular, we conjecture that the disproportionate response among the searching unemployed is the result of higher levels of motivation (relative to the discouraged unemployed). However, controlling for emotional well-being and confidence in individuals' own writing abilities, the significantly higher relative response remains for this group (albeit with a smaller magnitude). Hence, it is likely that other uncaptured decision processes and types of motivation are still caught up in the searching unemployed dummy. It is possible that this group wishes to practice filling out forms and taking assessments, as they also do when embarking on the job search process. A similar analysis of location holds, where those residing at typical entry points into the labour market (informal urban areas in the case of the black population) are more likely to participate in the test. This is true, even when controlling for labour market status. Furthermore, the self-employed in the lowest wage earning quintiles are least likely to participate in the test, possibly as a result of time opportunity costs involved in establishing secure business operations. However, controlling for duration of the survey does not remove this impact, even though it reduces the size of the coefficient. Hence, other unobserved factors are still at play.

While confidence in individuals' writing abilities and emotional well-being or motivation are strong predictors of test participation (for the black population), these factors do not "explain away" the different propensities to participate in the test for individuals within various labour market categories. This firstly highlights that the underlying types of motivation and confidence that encourage individuals to search for employment (in particular) are not the same as those that cause them to write numeracy tests. Some overlap does exist (as coefficient magnitudes do change somewhat), but it does not explain the entire decision process. Rather, additional factors are of importance. We investigate whether household peer effects (that by implication represent a "culture of learning" within the household) can fill this gap in the explanation. Indeed, these effects improve test participation rates, but they still do not account for the significance of the other variables. Hence, a culture of learning does prevail in some households above others, and improves test response. However, this same culture of learning still does not explain why some types of labour market participants are more prone to answering the numeracy test. We furthermore consider opportunity costs of time. Presumably the employed and the searching unemployed have less time to participate in numeracy tests. While the interview duration explains the aversion of the formally employed to taking the test, it does not account for the enthusiasm of the searching unemployed. Hence, even controlling for a number of decision-related factors, this group remains an anomaly.

Interestingly, survey fatigue sets in for the black and coloured populations, but the opposite holds for the (generally wealthier) white population. The longer the time spent on completing the survey in the white population, the more likely they were to also take the test. This suggests that within the white group there is a particular interest in co-operating with educational projects and that a form of "pseudoaltruism" is present for a subset of respondents. This is also reflected in the results for

education, which suggest that more educated individuals are also most likely to take the test. Because confidence in writing abilities is controlled for, we cannot conclude that more educated individuals participate as a result of their relatively higher perceived abilities. Therefore the remaining explanation is that a form of “pseudoaltruism” drives this decision.

This paper has shown that psychological motivations (real information) determine whether individuals are likely to respond to a test. However, differences along labour market lines persist (even if smaller in magnitude), regardless of the behavioural characteristics we control for. This suggests that, while some overlapping behavioural patterns determine labour market choices and the choice to participate in a learning experience, these processes are not entirely the same. Since both the choice to participate in the labour market and in education are important objectives for South Africa’s development, further research into understanding the underlying motivations behind both is called for. More broadly, while voluntary test response constitutes a very specific decision, the motivational patterns uncovered here could inform our understanding of individuals’ willingness to participate in more general educational programmes and improve our knowledge of a culture of learning amongst South Africans.

7 Bibliography

Allen, C.T., Machleit, K.A. & Schultz Kleine, S., 1992. A Comparison of Attitudes and Emotions as Predictors of Behavior at Diverse Levels of Behavioral Experience. *The Journal of Consumer Research*, 18(4), pp.493 - 504.

Ardington, C., Lam, D., Leibbrandt, M. & Welch, M., 2006. The sensitivity to key data imputations of recent estimates of income poverty and inequality in South Africa. *Economic Modelling*, 23, pp.822-35.

Axinn, W.G. & Pearce, L.D., 2006. *Mixed method data collection strategies*. New York: Cambridge University Press.

Belzil, C. & Hansen, J., 2002. Unobserved Ability and the Return to Schooling. *Econometrica*, 70(5), pp.2075-91.

Berlin, M. et al., 1992. An Experiment in Monetary Incentives. In *Proceedings of the American Statistical Association, Survey research Section*. Alexandria, 1992.

Blair, E. & Burton, S., 1987. Cognitive Processes Used by Survey Respondents to Answer Behavioral Frequency Questions. *The Journal of Consumer Research*, 14(2), pp.280 - 288.

Booysen, F. et al., 2008. Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African Countries. *World Development*, 36(6), pp.1113-30.

Bradburn, N.M., 1978. Respondent Burden. In L. Reeder, ed. *Health Survey Research Methods: Second Biennial Conference*. Williamsburg, Va. Washington, D.C.: US Government Printing Office.

Brüggen, E.C., De Ruyter, J.C., Wetzels, M.G.M. & Schillewaert, N., 2007. Waarom doen respondenten mee aan onderzoek in online panels? : een onderzoek naar de relatie tussen motieven en respons. In *Ontwikkelingen in het marktonderzoek : jaarboek 2007*. Maastricht: De Vrieseborch. pp.153 - 169.

Burks, S.V., Carpenter, J.P., Goette, L. & Rustichini, A., 2010. *Overconfidence as a Social Signalling Bias*. IZA Discussion Paper No. 4840. Bonn, Germany: Institute for the Study of Labor.

Chevalier, A. et al., 2008. *Students' Academic Self-Perception*. CEE DP 90. London: Centre for the Economics of Education London School of Economics.

Childers, T.L. & Rao, A.R., 1992. The Influence of Familial and Peer-Based Reference Groups on Consumer Decisions. *The Journal of Consumer Research*, 19(2), pp.198 - 211.

Clark, R.M., 1983. *Family Life and School Achievement: Why Poor Black Children Succeed or Fail*. Chicago: Chicago University Press.

Couper, M.P. & Singer, E., 2009. The Role of Numeracy in Informed Consent for Surveys. *Journal of Empirical Research on Human Research Ethics*, 4(4), pp.17 - 26.

Crowhurst, M., 1991. Interrelationships between Reading and Writing Persuasive Discourse. *Research in the Teaching of English*, 25(3), pp.314 - 338.

Dougherty, M.R., 2001. Integration of the ecological and error models of overconfidence using a multiple-trace memory model. *Journal of Experimental Psychology: General*, 130, pp.579 - 599.

Dunning, D., Meyerowitz, J.A. & Holzberg, A.D., 1989. Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability. *Journal of Personality and Social Psychology*, 57(6), pp.1082 - 1090.

Gilley, O.W. & Leone, R.P., 1991. A Two-Stage Imputation Procedure for Item Nonresponse in Surveys. *Journal of Business Research*, 22, pp.281 - 291.

Gino, F., Ayal, S. & Ariely, D., 2009. Contagion and differentiation in unethical behavior: The effect of one Bad Apple on the Barrel. *Psychological Science*, 20(3), pp.393 - 398.

Glazerman, S., Schochet, P.Z. & Burghardt, J., 2000. *National Job Corps Study: The Impacts of Job Corps on Participants' Literacy Skills*. U.S. Department of Labor, Employment and Training Administration Office of Policy and Research Report. Mathematica Policy Research, Inc.

Green, K.E., 1996. Sociodemographic Factors and Mail Survey Response. *Psychology & Marketing*, 13(2), pp.171 - 184.

Griffin, P., Leibbrandt, M., Pavlovic, M. & Zuze, T.L., 2010. *Numeric Literacy in South Africa: Report on the NIDS Wave 1 Numeracy Test*. National Income Dynamics Study Technical Paper No. 5.

Groves, R.M., Cialdini, R.B. & Couper, M.P., 1992. Understanding The Decision to Participate in a Survey. *The Public Opinion Quarterly*, 56(4), pp.475 - 495.

Hanushek, E.A. & Woessmann, L., 2008. The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46(3), pp.607-68.

Hanushek, E.A. & Woessmann, L., 2010a. Education and Economic Growth. In D.J. Brewer & P.J. McEwan, eds. *Economics of Education*. Amsterdam: Elsevier. pp.60-67.

Hanushek, E.A. & Woessmann, L., 2010b. Sample Selectivity and the Validity of International Student Achievement Tests in Economic Research. *Economic Letters*, Forthcoming.

Kingdon, G. & Knight, J., 2000. *Are searching and non-searching unemployment distinct states when unemployment is high? The case of South Africa*. Centre for the Study of African Economies Working Paper Series Paper 112. Centre for the Study of African Economies.

Lindbeck, A., 1997. Incentives and Social Norms in Household Behavior. *The American Economic Review*, 87(2), pp.370 - 377.

Louw, M., van der Berg, S. & Yu, D., 2007. Convergence of a Kind: Educational Attainment and Intergenerational Social Mobility in South Africa. *South African Journal of Economics*, 75(3), pp.548-71.

Marcus, B. & Schütz, A., 2005. Who Are the People Reluctant to Participate in Research? Personality Correlates of Four Different Types of Nonresponse as Inferred from Self- and Observer Ratings. *Journal of Personality*, 73(4), pp.959 - 984.

Porst, R. & von Briel, C., 1995. *Waren Sie vielleicht bereit, sich gegenebenfalls noch einmal befragen zu lassen? Oder: Gründe für die Teilnahme an Panelbefragungen.* Arbeitsbereicht Nr. 95/04. Mannheim, Germany: ZUMA.

Sharp, L.M. & Frankel, J., 1983. Respondent Burden: A Test of Some Common Assumptions. *The Public Opinion Quarterly*, 47(1), pp.36 - 53.

Sherman, R.P., 2000. Tests of Certain Types of Ignorable Nonresponse in Surveys Subject to Item Nonresponse or Attrition. *American Journal of Political Science*, 44(2), pp.362 - 374.

Singer, E., 2002. The Use of Incentives to Reduce Nonresponse in Household Surveys. In R.M. Groves, D.A. Dillman, J.L. Eltinge & R.J.A. Little, eds. *Survey Nonresponse*. New York: Wiley. pp.163 - 78.

Singer, E. et al., 1999. The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys. *Journal of Offcial Statistics*, 15(2), pp.217 - 230.

Sunstein, C.R., 1996. Social Norms and Social Roles. *Columbia Law Review*, 96(4), pp.903 - 968.

van der Berg, S., 2007. Apartheid's Enduring Legacy: Inequalities in Education. *Journal of African Economies*, 16(5), pp.849-80.

Wooldridge, J.M., 2009. *Introductory Econometrics - A Modern Approach*. 4th ed. Cengage South-Western

8 Appendix

Table 4 Probit Models of Taking Test – whole population

		-1-	-2-	-3-	-4-	-5-
Labour Market Status (reference: Not Economically Active)	Discouraged Worker	-0.202***	-0.082	-0.193**	-0.163*	-0.026
	Searching Unemployed	0.057	0.130**	0.064	0.094*	0.172**
	Casual Employed - Quintile 1	-0.032	0.08	-0.004	-0.059	0.003
	Casual Employed - Quintile 2	-0.588*	-0.38	-0.568*	-0.643	-0.349
	Casual Employed - Quintile 3	-0.287*	-0.119	-0.265	-0.211	-0.06
	Casual Employed - Quintile 4	-0.053	0.043	-0.039	0.091	0.224
	Casual Employed - Quintile 5	-0.09	-0.029	-0.101	-0.081	0.013
	Self Employed - Quintile 1	-0.079	0.171	-0.048	-0.018	0.256
	Self Employed - Quintile 2	-0.848***	-0.622***	-0.885***	-0.946***	-0.718***
	Self Employed - Quintile 3	-0.221	0.059	-0.209	-0.147	0.196
	Self Employed - Quintile 4	-0.546***	-0.31	-0.567***	-0.413*	-0.083
	Self Employed - Quintile 5	0.097	0.312	0.104	0.111	0.375
	Other Employed - Quintile 1	-0.496***	-0.252**	-0.504***	-0.473***	-0.2
	Other Employed - Quintile 2	-0.435***	-0.191*	-0.435***	-0.370***	-0.08
	Other Employed - Quintile 3	-0.233**	-0.053	-0.222**	-0.095	0.12
	Other Employed - Quintile 4	-0.078	0.088	-0.065	0.035	0.247**
	Other Employed - Quintile 5	-0.173**	-0.108	-0.207**	-0.114	0.02
Race (Reference: Black)	Coloured	-0.064	-0.067	-0.065	0.002	0.03
	Indian	-0.328*	-0.439**	-0.314	-0.07	-0.156
	White	-0.314***	-0.452***	-0.312***	-0.181	-0.255**
Location (Reference: Rural Formal)	Tribal Authority		-0.165**			-0.175**
	Urban Formal		0.056			0.063
	Urban Informal		0.226**			0.206**
Education (Reference: No schooling)	Primary Education		1.222***			1.281***
	Lower Secondary Education		1.390***			1.427***
	Upper Secondary Education		1.523***			1.545***
	Matric		1.588***			1.660***
	Diploma/Certificate		1.660***			1.696***
	Bachelors Degree		1.793***			1.796***
	Postgraduate Degree		1.990***			2.015***

(continued on next page)

(continued)

		-1-	-2-	-3-	-4-	-5-
	Age		-0.076***			-0.088***
	Age^2		0.001***			0.001***
	Relative Writing Confidence			0.047	0.051	0.077*
	Absolute Emotional Index			0.066***	0.061***	-0.017
	Time Before Test				-0.006***	-0.007***
	(Time Before Test)^2				0.000***	0.000***
	Proportion of Eligible Household Members who took test				1.198***	1.257***
	Constant	-0.438***	-0.452	-0.438***	-0.633***	-0.485
	N	12213	12192	11983	11606	11606
	P(Chi2>c)	0.000	0.000	0.000	0.000	0.000
	Pseudo R2	0.019	0.098	0.02	0.114	0.193
Calculated Turning Points of Profiles	**Age**		50.145			50.360
	Time Before Test				321.519	357.941

NOTES: *Significant at the 10% level **Significant at the 5% level ***Significant at the 1% level. The sample only includes those aged between 15 and 59, who were eligible to take the adult numeracy test. Estimates are weighted. Significance levels are based on robust standard errors. Calculated turning points are based on the coefficient estimates: $-\beta_x/2\beta_{x^2}$

Table 5 Probit Models of Taking Test – Black population

		-1-	-2-	-3-	-4-	-5-
Labour Market Status (reference: Not Economically Active)	**Discouraged Worker**	-0.265***	-0.132	-0.258***	-0.250***	-0.115
	Searching Unemployed	0.009	0.095	0.022	0.045	0.123*
	Casual Employed - Quintile 1	-0.134	-0.019	-0.103	-0.116	-0.037
	Casual Employed - Quintile 2	-0.691*	-0.478	-0.662*	-0.737	-0.44
	Casual Employed - Quintile 3	-0.226	-0.053	-0.202	-0.166	-0.049
	Casual Employed - Quintile 4	0.049	0.224	0.078	0.269	0.477**
	Casual Employed - Quintile 5	-0.294	-0.226	-0.275	-0.245	-0.203
	Self Employed - Quintile 1	-0.340**	-0.065	-0.316**	-0.282*	0.022
	Self Employed - Quintile 2	-0.854***	-0.617**	-0.887***	-0.882***	-0.654**
	Self Employed - Quintile 3	-0.399**	-0.04	-0.408**	-0.28	0.14
	Self Employed - Quintile 4	-0.580**	-0.279	-0.595**	-0.406	-0.013
	Self Employed - Quintile 5	-0.013	0.357	-0.002	0.015	0.393
	Other Employed - Quintile 1	-0.591***	-0.297**	-0.548***	-0.524***	-0.209
	Other Employed - Quintile 2	-0.488***	-0.220*	-0.479***	-0.423***	-0.126
	Other Employed - Quintile 3	-0.324***	-0.129	-0.321***	-0.157	0.073
	Other Employed - Quintile 4	-0.233**	-0.036	-0.226**	-0.125	0.099
	Other Employed - Quintile 5	-0.213**	0.012	-0.250**	-0.149	0.126
Location (Reference : Rural Formal)	**Tribal Authority**		-0.243***			-0.281***
	Urban Formal		-0.002			0.006
	Urban Informal		0.188*			0.177
Education (Reference: No schooling)	**Primary Education**		1.184***			1.238***
	Lower Secondary Education		1.370***			1.391***
	Upper Secondary Education		1.436***			1.432***
	Matric		1.590***			1.655***
	Diploma/Certificate		1.514***			1.549***
	Bachelors Degree		1.503***			1.406***
	Postgraduate Degree		1.133***			1.281***

(continued on next page)

(continued)

		-1-	-2-	-3-	-4-	-5-
	Age		-0.070***			-0.076***
	Age2		0.001***			0.001***
	Relative Writing Confidence			0.062**	0.065*	0.077*
	Absolute Emotional Index			0.082***	0.078***	-0.008
	Time Before Test				-0.011***	-0.012***
	(Time Before Test)2				0.000***	0.000***
	Proportion of Eligible Household Members who took test				1.243***	1.329***
	Constant	-0.393***	-0.334	-0.394***	-0.486***	-0.305
	N	9602	9594	9435	9154	9154
	P(Chi2>c)	0.000	0.000	0.000	0.000	0.000
	Pseudo R2	0.016	0.114	0.02	0.127	0.225
Calculated Turning Points of Profiles	**Age**		59.053			62.284
	Time Before Test				283.509	300.414

NOTES: *Significant at the 10% level **Significant at the 5% level ***Significant at the 1% level. The sample only includes black individuals aged between 15 and 59, who were eligible to take the adult numeracy test. Estimates are weighted. Significance levels are based on robust standard errors. Calculated turning points are based on the coefficient estimates: $-\beta_x/2\beta_{x^2}$

Table 6 Probit Models of Taking Test – Coloured population

		-1-	-2-	-3-	-4-	-5-
Labour Market Status (reference: Not Economically Active)	Discouraged Worker	0.362*	0.3	0.401*	0.662***	0.581**
	Searching Unemployed	0.488**	0.403*	0.515**	0.634***	0.514**
	Casual Employed - Quintile 1	1.149***	1.205***	1.131***	1.371***	1.434***
	Casual Employed - Quintile 2	0.57	0.694	0.549	0.499	0.732
	Casual Employed - Quintile 3	0.068	0.313	0.063	0.289	0.453
	Casual Employed - Quintile 4	0.169	0.004	0.174	0.236	0.071
	Casual Employed - Quintile 5	0.524	0.152	0.201	0.292	0.211
	Self Employed - Quintile 1	0.639	0.398	0.621	0.971	0.772
	Self Employed - Quintile 2	-0.369	-0.116	-0.36	-0.281	-0.031
	Self Employed - Quintile 3	-0.33	-0.48	-0.363	-0.165	-0.294
	Self Employed - Quintile 4	-0.59	-0.712	-0.636	-0.269	-0.192
	Self Employed - Quintile 5	-1.119**	-1.165**	-1.121**	-1.141**	-1.072**
	Other Employed - Quintile 1	0.063	0.218	0.064	0.175	0.285
	Other Employed - Quintile 2	-0.218	-0.051	-0.214	-0.23	-0.096
	Other Employed - Quintile 3	0.430*	0.518**	0.438*	0.459*	0.482*
	Other Employed - Quintile 4	0.481**	0.513**	0.547**	0.517*	0.504*
	Other Employed - Quintile 5	0.093	-0.006	0.066	0.131	0.024
Location (Reference : Rural Formal)	Tribal Authority		1.112			0.821
	Urban Formal		0.079			-0.013
	Urban Informal		-0.732**			-0.841***
Education (Reference: No schooling)	Primary Education		1.219***			1.329***
	Lower Secondary Education		1.011***			1.167***
	Upper Secondary Education		1.467***			1.580***
	Matric		1.181***			1.407***
	Diploma/Certificate		1.617***			1.708***
	Bachelors Degree		1.952***			1.766***
	Postgraduate Degree		1.405*			1.863**

(continued on next page)

(continued)

		-1-	-2-	-3-	-4-	-5-
	Age		-0.037			-0.038
	Age^2		0			0
	Relative Writing Confidence			0.034	0.037	0.157
	Absolute Emotional Index			0.05	0.018	-0.021
	Time Before Test				-0.008	-0.007
	(Time Before Test)^2				0.000*	0
	Proportion of Eligible Household Members who took test				1.527***	1.560***
	Constant	-0.842***	-1.268**	-0.841***	-1.095***	-1.516**
	N	1800	1792	1769	1703	1703
	P(Chi2>c)	0.001	0.000	0.001	0.000	0.000
	Pseudo R2	0.033	0.094	0.038	0.182	0.234
Calculated Turning Points of Profiles	**Age**		64.172			75.622
	Time Before Test				107.371	107.426

NOTES: *Significant at the 10% level **Significant at the 5% level ***Significant at the 1% level. The sample only includes coloured individuals aged between 15 and 59, who were eligible to take the adult numeracy test. Estimates are weighted. Significance levels are based on robust standard errors. Calculated turning points are based on the coefficient estimates: $-\beta_x/2\beta_{x^2}$

Table 7 Probit Models of Taking Test – White population

		-1-	-2-	-3-	-4-	-5-
Labour Market Status (reference: Not Economically Active)	Discouraged Worker	0.052	0.069	0.192	0.069	0.278
	Searching Unemployed	0.368	0.699	0.347	0.361	0.755
	Casual Employed - Quintile 1	0.035	0.092	-0.012	-0.252	-0.014
	Casual Employed - Quintile 2					
	Casual Employed - Quintile 3					
	Casual Employed - Quintile 4					
	Casual Employed - Quintile 5	-1.141**	-0.972*	-1.064*	-1.174*	-1.179*
	Self Employed - Quintile 1	0.197	0.073	0.187	-0.048	-0.133
	Self Employed - Quintile 2	-0.569	-0.328	-0.597	-0.695	-0.345
	Self Employed - Quintile 3	-0.063	-0.13	0.039	-0.221	-0.095
	Self Employed - Quintile 4	0.121	-0.212	0.046	0.006	-0.262
	Self Employed - Quintile 5	0.561	0.186	0.55	0.367	0.083
	Other Employed - Quintile 1	-0.029	0.136	-0.314	-0.442	-0.351
	Other Employed - Quintile 2	0.038	0.29	0.015	-0.251	0.037
	Other Employed - Quintile 3	-0.022	0.215	0.038	-0.107	0.102
	Other Employed - Quintile 4	0.307	0.242	0.324	0.275	0.274
	Other Employed - Quintile 5	0.277	0.158	0.226	0.083	0.025
Location (Reference : Rural Formal)	Tribal Authority					
	Urban Formal		-0.544**			-0.423
	Urban Informal					
Education (Reference: No schooling)	Primary Education					
	Lower Secondary Education		-1.033*			-1.226**
	Upper Secondary Education		-0.604			-0.351
	Matric		-1.111***			-1.040**
	Diploma/Certificate		-0.454			-0.438
	Bachelors Degree		-0.285			-0.274
	Postgraduate Degree					

(continued on next page)

(continued)

		-1-	-2-	-3-	-4-	-5-
	Age		-0.072			-0.068
	Age^2		0.001			0.001
	Relative Writing Confidence			-0.093	-0.156	-0.352
	Absolute Emotional Index			0.012	0.062	0.051
	Time Before Test				0.017***	0.016**
	(Time Before Test)^2				-0.000***	0
	Proportion of Eligible Household Members who took test				0.26	0.186
	Constant	-1.024***	1.178	-0.985***	-1.586***	0.374
	N	609	601	581	560	557
	P(Chi2>c)	0.345	0.003	0.502	0.132	0.001
	Pseudo R2	0.022	0.086	0.026	0.081	0.141
Calculated Turning Points of Profiles	**Age**		35.599			35.236
	Time Before Test				357.131	324.156

NOTES: *Significant at the 10% level **Significant at the 5% level ***Significant at the 1% level. The sample only includes white individuals aged between 15 and 59, who were eligible to take the adult numeracy test. Estimates are weighted. Significance levels are based on robust standard errors. Cells that are left blank in addition to those in other tables indicate that there were no observations for the white population for that variable. Calculated turning points are based on the coefficient estimates: $-\beta_x/2\beta_{x^2}$