# Perceived change in happiness and depressive symptoms

A Research Report submitted in partial fulfilment of the Degree of

Master of Commerce (Economics/Economic Science)

in the School of Economics and Finance,

University of the Witwatersrand

by

Name: Imraan Yusuf

Student No: 493462

Supervised by: Professor Umakrishnan Kollamparambil

| | |
|---|---|
| Word Count – total (excluding cover page) | 15,412 |
| Word Count – main write-up | 11,504 |
| Word Count – references | 2,720 |
| Word Count – annexes | 1,188 |

Date: 14 June 2021

# Perceived change in happiness and depressive symptoms

**Imraan Yusuf**

**Abstract**

Depression / depressive symptoms have been widely studied; however, the literature has not sufficiently addressed the effect of contrast on depressive symptoms. This is particularly important in the South African context, given the high levels of depressive symptoms, when compared to other countries. This study aims to understand and quantify the effect of contrast on depressive symptoms by making use of ordinary least squared, System-GMM, and heterogeneity based instrumental variable regression. The latter two methods are used in order to remove differing forms of endogeneity bias. The analysis suggests that those who consider they are more happy than they were ten years ago have lower depressive symptoms than those who do not consider their level of happiness has changed. This is consistent for those individuals who have correct recall and those that have incorrect recall of contrast. Interestingly, the analysis finds that those who consider they are less happy have higher depressive symptoms, however, only when recall is correct. These findings highlight the importance of memory and perceptions, on depressive symptoms. They provide evidence in support of new (and old) types of memory training techniques which could be used to combat depression.

JEL: I31 / Y40

Keywords: depression, depressive symptoms, contrast, memory, happiness

# Contents

# 1. Introduction

The South African Depression and Anxiety Group (2010) defines depression as a whole-body disease (i.e. body, mood and thoughts). Depression is not something that can be "wished away" and it affects millions of people. Without treatment, it can last a long period of time (World Health Organisation, 2010; World Federation for Mental Health, 2012). On a worldwide basis, more than 264 million people suffer from depression, it is the leading cause of disability and a substantial contributor of the global burden of disease, and it can lead to suicide (World Health Organisation, 2020).

From an economic point of view depression may increase health costs and result in a reduction in total productivity in the work force (Donohue and Pincus, 2007). In relation to productivity losses, mental health is an important factor that plays a role in determining the number of working days lost, while there are also non-cognitive skills that are affected by mental health problems (Currie, 2020; Edin et al., 2017). In Europe, there is an increasing trend of depression and absenteeism from work. Those with mental health problems are more likely to retire early (McDaid, 2007). In relation to the healthcare system, a material portion of the social welfare payments are attributed to mental health problems (McDaid, 2007). This suggests that an improvement in mental health may result in an increase in productivity in the workplace and a reduction in welfare payments.

From a South African perspective, mental health and depression are important for multiple reasons. Firstly, because of the toll that apartheid has taken on South African citizens; secondly, because of the potential loss in earnings; and lastly, because of the already widespread nature of depression (Lund et al., 2013). SADAG (2017) points out that in South Africa the lifetime prevalence of depression is 9.7%; in rural urban based studies, a rate of depression of 27% has been found, whereas in urban areas, the prevalence of depression is estimated to be around 25%. From an economic point of view, SADAG (2017) highlights a R218bn loss in productivity in South Africa, as a result of depression, with Bateman (2015) estimating that average lost earnings for those with major depression and anxiety disorders at approximately R50,000-00 per annum.[1]

The ability to understand and mitigate the factors that facilitate depression may lead to an increase in economic activity and overall well-being of society. There has been substantial

[1] It is unclear which year the SADAG (2017) information is based on; however, it is still a sizeable loss.

input into the determinants of depression. Current literature has considered the effect of marriage, age, education, gender, race, household expenditure and income, lifetime assets, relative perceived socio-economic status, working hours, and perceived health, on depression (Akhtar-Danesh and Landeen, 2007; Ardington and Case, 2010; Everson et al., 2002; Mirowsky and Ross, 1992; Pearlin and Johnson, 1997). These types of analyses are particularly important in the South African context due to mean depression scores being higher in South Africa than other countries (Ardington and Case, 2010).

While the current literature is vast and considers many socioeconomic factors, there are still gaps within the literature. One of the gaps is understanding the relationship between perceived change in happiness (i.e. contrast) and depressive symptoms. This gap, which is filled by this paper, is both on a national and international level, given that the current relationship has not been captured previously. Notably, there is a difference between depressive symptoms and depression, with depression being concluded via clinical diagnosis. Depressive symptoms are used to highlight potential depression cases (Baron et al., 2017). Nonetheless, a reduction in depressive symptoms could be used to alleviate depression.

Contrast in this instance, is defined as how an individual perceives their level of happiness has changed over time. An individual may consider that their level of happiness has either increased, stayed the same, or decreased from a prior level of happiness. For those individuals that have considered their level of happiness has *decreased (increased)*, the expectation would be for these individuals to have *higher (lower)* depressive symptoms. These individuals recall memories such that the current life satisfaction is *inferior (superior)* to that of the past. The perception of *reduced (increased)* life satisfaction may result in an *increased (decreased)* of depressive symptoms. This relationship is facilitated through the role of memory.

The effect of contrast on depression is explored via three econometric methods, each with increasing efficiency. Specifically: ordinary least squares (OLS), the baseline model; System-GMM, the fixed effects removal model; and heterogeneity based instrumental variable (HBIV) regression – all using the National Income Dynamic Study (NIDS). The HBIV regression, like general instrumental variable (IV), regression considers a form of endogeneity that the System-GMM does not (i.e. time variant endogeneity), increasing the efficiency of the analysis. In addition, HBIV provides additional aid when there are potential weak instruments being used, which is not controlled for by System-GMM (Ćorić and Šimić, 2021).

The analysis suggests that those that perceive that they are more happy than they were ten years ago, have lower depressive symptoms than those that do not consider their happiness has changed over time. This is consistent with the hypothesis. The opposite is found for those who are less happy (with the direction being consistent with the hypothesis), however, the coefficients are not statistically significant in this case; therefore, conclusions cannot be drawn.

When considering the analysis above, it is important to understand that recall (of past levels of happiness), may not be accurate. The way we perceive change in happiness may not be how levels of happiness have changed. In the current study, incorrect recall (i.e. incorrect memory) is approximately 57%, suggesting that a substantial portion of individuals do not correctly recall how their level of happiness has changed over time; also, more generally found in Prati and Senik (2020).[2]

When considering the incorrect/correct recall subsets, the effect of more happy is similar and statistically significant for both correct and incorrect recall. However, the effect of less happy is only statistically significant for the correct recall sample (for the HBIV analysis). This suggests that perceptions (and memory) play a role in reducing depressive symptoms when an individual considers they are more happy, however, perceptions (and memory) do not play a role for those that consider they are less happy. This provides a basis for recommendations when it comes to combatting depressive symptoms. Notably, all results should be considered in line with the limitations provided, and there is potential for further research in this field.

The remainder of this paper is structured as follows: section 2 provides a review of the relevant literature, providing the link between contrast and depression; section 3 discusses the data being used, and provides some descriptive statistics; section 4 discusses the methodology for analysis, considering the regression methods mentioned above; section 5 provides the results and a discussion; section 6 points out the limitations and further research that can be conducted; and section 7 concludes.

[2] Generated based on the sample of individuals used in the analysis. Panel weighted average for national inference is 58%.

# 2. Literature review

The literature review first highlights the determinants of depression to date, and then builds the relationship between perceived change in happiness and depression. While most of the theory below relates directly to depression, one can draw the parallel to depressive symptoms.

### a. General literature review

There exists a vast array of literature on the determinants of depression, with many interesting findings in relation to socioeconomic factors. Depression has been considered through the effect of marriage. Married individuals are found to be less prone to depression. This relationship is facilitated through the differing social structure between married and unmarried individuals (Pearlin and Johnson, 1977).

Age has also been considered as a facilitating factor of depression in a study conducted in the United States (Mirowsky and Ross, 1992). The findings indicate an increase in age results in a reduction in depression due to life-time gains over the period, whereas the subsequent increase in depression (i.e. a quadratic relationship) is as a result of "*losses in marriage, employment and economic well-being*" (Mirowsky and Ross, 1992). Notably, in the South African context Burger, et al. (2017) finds the opposite relationship with depressive symptoms increasing with age and then falling.

A few studies have considered several socioeconomic factors, as opposed to focusing on one (Akhtar-Danesh and Landeen, 2007; Ardington and Case, 2010; Everson et al., 2002; Miech and Shanahan, 2000). In doing so, the following socio-economic and socio-demographic factors are assessed as determinants of depression (or pointed out as previously derived determinants): education, gender, race, household expenditure per member, income, number of assets, relative perceived socio-economic status, age, and immigrant status.

In addition to the above, long working hours have been found to play a role in determining depressive symptoms, potentially through the ambit of "*unhealthy lifestyles and stress-related physical diseases*" (Virtanen et al., 2011; Virtanen et al., 2012). Mui (1996) found that perceived health plays a role in determining depression for the elderly.

While majority of the findings above relate to international studies, there have also been several local studies. Tomlinson, et al. (2009) points out that depression is more prevalent in females and those with low levels of education. For the elderly in South Africa, Peltzer and Phaswana-

Mafuya (2013), find that functional disability, limited quality of life and chronic conditions are related to depression. Similarly, Burger, et al. (2017) finds that negative life events play a role in determining depressive symptoms. Tomita and Burns (2013) finds a negative association between social capital and depression, where the relationship is due to social trust and neighbourhood preferences.

### b. Contrast and depression

While the current literature on depression is quite broad, there exists a gap when it comes to understanding the relationship between contrast and depressive symptoms.[3] Therefore, there are no prior studies which explain the theoretical relationship between these two factors. As such, this section considers the general depression theory, in order to build the theoretical relationship between contrast and depression. Contrast in this instance, is defined as how an individual perceives their level of happiness has changed over time (increased, decreased or stayed the same).

The way an individual perceives change in happiness is directly related to the way we recall (via memory) levels of prior happiness.[4] If recall is biased (i.e. as a result of memory), an individual's perception of change in happiness is different from the actual change in happiness (Marsh et al., 2019). In this context, memory is the intermediating factor in the relationship between contrast and depression. The remainder of this section builds on this relationship.

There are *three* potential ways in which contrast and depression are related. All are through the intermediating factor of memory. Memory affects the way an individual can recall past levels of happiness and thus informs the change in happiness than an individual perceives.

Firstly, memory has been related to depression through depression theory (with limited empirical work). Secondly, contrast has been incorporated as a determinant of subjective well-being (through theory), which is related to depression. Thirdly, the effect of depression has been assessed on memory (i.e. a possible endogeneity effect for the current analysis). Each of these points are discussed, in turn, below.

---

[3] Alternatively, there is potential for analysis to be conducted to assess whether perceived change in happiness has a causal effect on *possible* depression (i.e. whether contrast affects the likelihood of depression). However, while it is possible to categorise an individual as depressed or not from the measure of depression, a clinical categorisation of everyone is not available. Therefore, a reliable analysis cannot be completed.

[4] The term "contrast" is used in literature either to indicate the "relative" differences from one's benchmark group (as in relative income) or as "change from one's own past". In this study, the term is used to refer to the latter.

In relation to memory and depression, theory suggests that the way an individual perceives their own life and recalls memories (or happiness, through memory) may be different from what occurred. Beck (1967) cognitive theory of depression suggests that those that are prone to depression show negative "*self-schemata*" around "*inadequacy, loss, failure and worthlessness*" (Alloy et al., 1997). Self-schemas (or self-image) are defined as personal cognitive structures that are based on generalisations or experience (Oxford Reference, 2020). These depressive self-schemata guide the way people perceive, interpret and recall memories, which results in a negative personal view (Alloy et al., 1997).[5] In this context, perceptions (of change in happiness) are important as a determinant of depression.

Relatedly, there is a theory of hopelessness, which is provided by Abramson, et al. (1989) and Alloy, et al. (1988). This states that the "*hypothesised cognitive vulnerability in the hopelessness theory operates to increase risk for depression through its effects on processing or appraisals of personally relevant life experiences*" (Alloy et al., 1997). This can be thought of as the perception of life events, which play a role in determining happiness (and therefore, perceived change in happiness), also play a role in determining depression. From a theoretical perspective, it is for the above reasons that perceived change in happiness (via memory) is likely to play a determinative role in assessing the likelihood of depression.[6]

Along the lines of empirical work, Gibbs and Rude (2004) assess whether overgeneral memory predicts depressive symptoms through the occurrence of stressful life events, finding a positive relationship between the two. In this context, overgeneral memory facilitated the way stressful life events were interpreted with their ability to predict depressive symptoms. A parallel can be drawn between memory, perceived change in happiness and depressive symptoms.

Lastly, forgetting (i.e. subjective memory) is important for emotion regulation (Nørby, 2015). Forgetting reduces what would be the effect of negative memories, and as such, influences perceived change in happiness. While Nørby (2015) makes the connection between memory and subjective well-being, it is possible for the same *type* of relationship (i.e. through forgetting) to exist between memory and depression, considering that depression has been related to subjective well-being, and that memory generally has a theoretical relationship with depression (OECD, 2013).

[5] Note, in this context, depressive self-schemata increase vulnerability towards depression.
[6] While the two theories above relate directly to memory, it is likely that the ability to recall may play a determinative factor, since it is indeed, the perceptions that one has that lies in the crux of the relationship.

While the above relates memory to depression, from a theoretical perspective, memory has also been related to subjective well-being. Adaptation level theory suggests that chasing happiness is self-defeating given that it is not possible to change one's level of happiness (Brickman and Campbell, 1971). This theory consists of contrast and habituation, of which, contrast is currently relevant (Brickman et al., 1978). Contrast is based on the concept of the happiness being derived from a previous reference point of happiness (i.e. memory of prior events). The reference point is based on the ability of the individual to remember the event, for without memory, the comparison is not possible (Tversky and Griffin, 1991). Therefore, memory plays a facilitating role in determining subjective well-being.

Notably, there is a co-determinant nature of depression and subjective well-being, suggesting that the relationships that affect subjective well-being may also affect depression (Lee and Hwang, 2018; Seo et al., 2018). In the current context, memory may facilitate contrast, which in turn plays a role in determining depressive symptoms.

While both avenues already discussed highlight contrast as a determinant of depression, there is also a potential reverse relationship. Memory, more generally, has been assessed to a substantial extent in empirical literature, however, the link of causality has been different. Most studies have assessed the effect of depression on memory (Marsh et al., 2019; Pokorski and Siwiec, 2008; Popovski and Bates, 2005; Roth and Rehm, 1980; Söderlund et al., 2014; Zuroff et al., 1983).

Generally, depressed individuals remember more negative memories than non-depressed individuals, which is facilitated through a reduction in remembering positive memories (Marsh et al., 2019; Zuroff et al., 1983). In addition, mild depression results in more realistic recall and future expectations, whereas severe depression is related to negatively biased past recall and future expectations (Anderson and Evans, 2015; Coyne and Gotlib, 1983; Eich et al., 1994; MacLeod et al., 1993; Strunk and Adler, 2009).[7] As is evident, depression may affect memory, which may affect contrast; the reverse relationship of which the current study seeks to assess. Various other studies have found (or failed to find) a relationship between memory and depression (Pokorski and Siwiec, 2008; Popovski and Bates, 2005; Roth and Rehm, 1980; Söderlund et al., 2014).

[7] References extracted from Marsh, et al. (2019).

To build on the above, memory has also been assessed from a general viewpoint (i.e. independent of depression). Marsh, et al. (2019) indicate that people tend to have a positive bias in both remembering memories and future simulation of events (Sharot, 2011; Sharot et al., 2007; Walker et al., 2003). In addition, positivity bias has also been shown across multiple sources over time (Berntsen, 1996; Clark et al., 2013; Ditta and Storm, 2016; Thompson et al., 1996; Suedfeld and Eich, 1995; Waldfogel, 1948; Walker et al., 2003). However, Prati and Senik (2020) show that while this exists for happy people, the opposite is true to those that are unhappy; aligning with the findings of depression above. Lastly, some studies have also shown that individuals align more positive events than negative ones, when thinking about the future (Newby-Clark and Ross, 2003; Weinstein, 1980).[8] These biases point out that recall is unlikely to perfect, providing further evidence for important role of perceptions.

Additionally, it is important to consider the anchoring effect of memory, in a slightly different light. If an analysis were to use actual change in happiness as reported by an individual, this would be based on the assumption that an individual considers that the scale of happiness between two periods in time is the same, however, this may not be the case (Kaiser, 2020). Individuals may update their beliefs on the scale with new information. Using perceived change in happiness takes into account the potential bias that may exist when scale changes occur.

Based on the above, depression may affect the way an individual recalls information, and as such, influence the way an individual's change in happiness is perceived. Furthermore, memory biases exist even in the absence of depression. Even though the theory is directly relating depression and memory, it is through the way an individual recalls information, therefore, suggesting that the ability to recall information (or the recall bias) may impact contrast. For instance, if an individual supresses negative memories, the ability to recall information is likely to be less than optimal.

To sum up, when perceived level of happiness has decreased (increased), the expectation would be for higher (lower) depressive symptoms. Memory recall is such that current happiness is inferior (superior) to past levels of happiness. In addition, there exists potential endogeneity bias through the reverse relationship and omission of ability to recall (or the recall bias).

[8] References extracted from Marsh, et al. (2019).

# 3. Data

### a. Overview

The NIDS database is a government funded, panel survey. The survey is conducted via the Southern Africa Labour and Research Unit (SALDRU), which is located at the University of Cape Town. The database includes a wide range of questions which relate to "*poverty and well-being, household composition and structure, fertility and mortality, migration, labour market participation, and economic activity, human capital formation, health, education, vulnerability and social capital*" (Datafirst, 2020).

The NIDS database started with wave 1 in 2008 and has since collected data on an additional 4 waves of data (in 2010-2011, 2012, 2014-2015, 2017). The first wave of data targeted 28,000 individuals across 7,300 different households. The data are collected for adults and children and where individuals are not available, proxy questionnaires are used. The relevant information relates to the data collected from adults, since children and proxy questionnaires do not include any information on depression. To supplement the individual level data, household level data are also collected (Brophy et al., 2018).

In using the NIDS database, it is important to understand the following: the main questions being used to assess the relationship; the representativeness of the dataset for national inference purposes along with the summary statistics of all variables; and the robustness of the measures used in the analysis.

### b. Main questions

The questions used in order to generate the main outcome variable are included in Figure 1 below. These questions are used to generate the Centre for Epidemiologic Studies Short Depression Scale (i.e. the depression score or CESD-10 score) that ranges from 0 to 30; as is used in Burger, et al. (2017). A common understanding about the depression score is that higher values relate to a higher risk/severity of depression, therefore, the depression score provides a useful screening device for depressive symptoms (Baron et al., 2017).

If two or more questions are unanswered, the depression score is not used (i.e. these are dropped from the analysis). Scoring of each question is completed as follows: all questions get a score from 0 to 3 based on the response, with 0 being allocated to the first option; scoring is reversed for question 5 and 8 since these are positive questions (Baron et al., 2017).

**Figure 1: CESD-10 questions**

**Section K: Emotional health**

*INTERVIEWER READ OUT*: We would like to know how your general well-being has been over the past week.

I am going to read a list of some of the ways you may have felt or behaved during the last week. Please state how often you have felt this way during the **past week.**

**Interviewer: Select one option on each line**

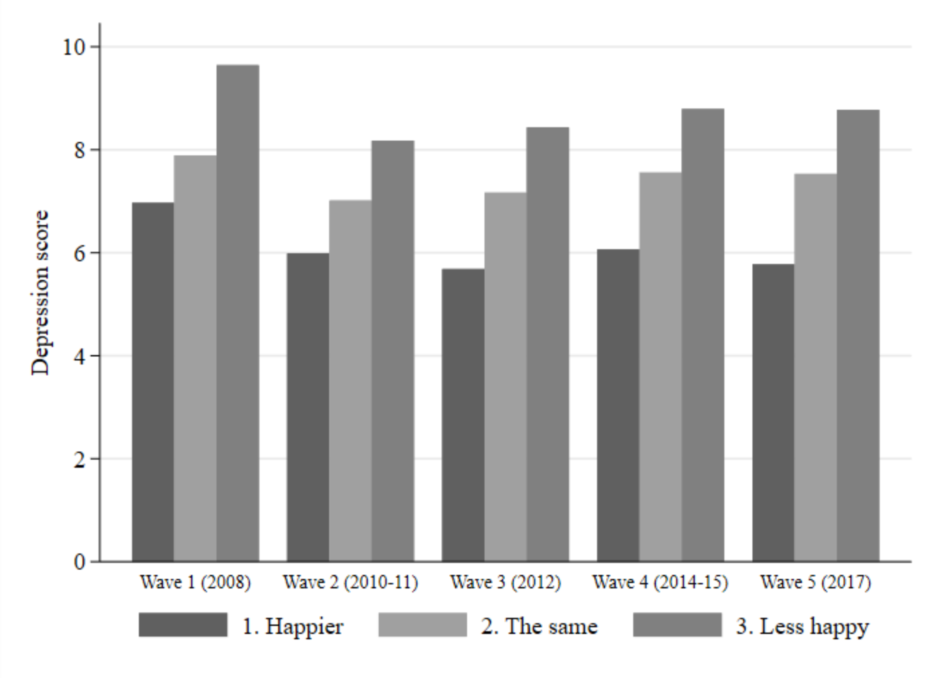
| | During the past week... | Rarely or none of the time (less than 1 day) | Some or little of the time (1-2 days) | Occasionally or a moderate amount of time (3-4 days) | All of the time (5-7 days) | Refused |
|---|---|---|---|---|---|---|
| **K1** *emobth* | I was bothered by things that usually don't bother me | 1 | 2 | 3 | 4 | -8 |
| **K2** *emomnd* | I had trouble keeping my mind on what I was doing | 1 | 2 | 3 | 4 | -8 |
| **K3** *emodep* | I felt depressed | 1 | 2 | 3 | 4 | -8 |
| **K4** *emoeff* | I felt that everything I did was an effort | 1 | 2 | 3 | 4 | -8 |
| **K5** *emohope* | I felt hopeful about the future | 1 | 2 | 3 | 4 | -8 |
| **K6** *emofear* | I felt fearful | 1 | 2 | 3 | 4 | -8 |
| **K7** *emoslp* | My sleep was restless | 1 | 2 | 3 | 4 | -8 |
| **K8** *emohap* | I was happy | 1 | 2 | 3 | 4 | -8 |
| **K9** *emolone* | I felt lonely | 1 | 2 | 3 | 4 | -8 |
| **K10** *emogo* | I could not "get going" | 1 | 2 | 3 | 4 | -8 |

*Source: NIDS questionnaire*

The question used in order to ascertain the perceived change in happiness across individuals is the following: "*Are you happier, the same or less happy with life than you were 10 years ago?*" (National Income Dynamics Study, 2017). Binary variables are generated to indicate when an individual is more happy or less happy with life.

As preliminary analysis between the two variables of interest, Figure 2 below suggests that there is likely to be a relationship between contrast and depressive symptoms. Those that perceive that they are less happy than they were previously, show a higher depression score, than those that perceive that they are happier than they were previously. Notably, these results are robust to each wave of data collected.

**Figure 2: Average CESD-10 score by perceived change in happiness, 2008-2017**




*Source: Analysis of NIDS data*

*Notes: Wave 5 panel weights used for all waves of data.*

In addition to the above, it is important to highlight that the depression score may be used to categorise individuals as depressed or not (Baron et al., 2017). Many different cut-off's have been used to classify possible depression, ranging from 8 to 16, however, the most common cut-off for the NIDS dataset has been 10 (Andreson et al., 1994; Baron et al., 2017; Tomita and Burns, 2013).[9] Ardington and Case (2010) point out that in context where the depression score has been validated, researchers will often use a cut-off value to discern between depressed and non-depressed individuals.

However, for the purposes of this study, the binary depression variable, although helpful, is less reliable than assessing the continuous depression variable. The reasons are two-fold. Firstly, the Baron, et al. (2017) study highlight differing values for some language groups in South Africa, which provide the highest accuracy of classification, however, this does not include all language groups in South Africa, therefore a full classification is not possible. In addition, there could be other characteristics which may affect the classification score that are unknown. Secondly, it is unlikely that the binary depression classification is as robust as

[9] Baron, et al. (2017) consider that the results in the paper indicate that the CESD-10 is an adequate tool for diagnosing depression. In addition, they find that the cut-off in the sample used was slightly higher than 10. The analysis was based on different language groups in South Africa, however, not all language groups were considered.

clinical diagnosis, especially considering the range of possible values that can be used to categorise depression (Oyenubi and Kollamparambil, 2020). Therefore, this can be, at most, considered a variable that considers *possible* depression.[10]

### c. Representativeness and summary statistics

In order to ensure that national level inference can be drawn from the NIDS database, a nationally representative sample was drawn in the first wave of data, however, there are two important considerations. First, due to response rates, some demographic groups have been underrepresented, whereas others have been overrepresented (Brophy et al., 2018). Second, there is non-random attrition exhibited in the data, from Whites, Indians/Asians and high-income earners (Brophy et al., 2018). This has led to the creation of panel weights in for national inference, and a top up sample included in the fifth wave of data (Branson & Wittenburg, 2019). Panel weights have been used in the current analysis, which will allow for nationally representative results.[11]

Table 1 below shows the descriptive statistics of the NIDS data, based on those individuals that are used in the regression analysis. The mean depression score for wave one was 7.89, which decreased to 6.84 in wave 5. A substantial portion of individuals have indicated that they are happier than they were ten years ago, which is in line with current literature on subjective well-being / happiness in South Africa (Kollamparambil, 2020). This figure increases to 60% in the fifth wave, from 45% in the first wave. Given the trend above, it is expected that the opposite would occur for the category of less happy, which is indeed the case. In the fifth wave, only 16% of individuals in the sample have indicated that they are less happy.

In addition to the above, most individuals are black, thus aligning with the South African demographic (Stats SA, 2020). More than 50% are female, and an average age of around 35-37 is exhibited, by wave of data. This is higher than the average age in South Africa given that only those older than 15 are considered in this analysis (i.e. they are the relevant sample). A

[10] In addition, a review of empirical techniques suggest that those considered, are not reliable. OLS does not consider the bounded nature of the binary dependent variable (Wooldridge, 2010). Probit and Logit do not account for endogeneity. Both IV Probit and IV Logit have limitations; the errors in "*the first stage regression cannot be normal or independently of the regressors*" as is required by the estimator and biased estimates even if a valid instrument is used, respectively (Carroll et al., 1998, cited in Rassen et al., 2009; Lewbel et al., 2012). Special regressor (SR) allows for endogenous binary variables (and dependent variables), however, a suitable SR is not available in the data (i.e. age satisfies all but one requirement; age has a quadratic relationship and not a linear one), and the limitations for weak instruments has not been fully assessed in the literature (Lewbel et al., 2012).
[11] The top-up sample cannot be used given the specification of the instrumental variable and the System-GMM analysis (i.e. these consider lags in the analysis, which the top-up sample does not have).

substantial portion of individuals are not economically active or employed (i.e. based on the employment variable), which is in line with the high levels of unemployment in South Africa (Stats SA, 2019). Majority of individuals are religious, and a limited, but material, percentage of individuals live an active lifestyle.

Weekly work hours are in the low- to mid-teens; however, this is likely a function of the high levels of unemployment in South Africa. Majority of individuals experience material levels of violence, which is an ongoing concern in South Africa (CSVR, 2007). On average, approximately 10 grades of schooling are completed (including Grade R), when considering up to high school education; and a limited percentage of the sample have tertiary education. South African's consider that they are in good health, and majority of individuals reside in Gauteng, Kwa-Zulu Natal, Eastern Cape and Western Cape (Stats SA, 2020).

**Table 1: Summary statistics of key variables**

| Explanatory variable | Mean (standard deviation) Wave 1 2008 | Mean (standard deviation) Wave 2 2010-11 | Mean (standard deviation) Wave 3 2012 | Mean (standard deviation) Wave 4 2014-15 | Mean (standard deviation) Wave 5 2017 | Min Wave 1 - 5 | Max Wave 1 - 5 |
|---|---|---|---|---|---|---|---|
| More happy | 45.35% | 45.33% | 51.48% | 60.86% | 60.02% | 0 | 1 |
| Less happy | 22.85% | 19.04% | 17.4% | 14.77% | 16.07% | 0 | 1 |
| Depression score | 7.89 (4.81) | 6.77 (4.19) | 6.63 (4.4) | 6.84 (4.18) | 6.68 (4.36) | 0 | 30 |
| Age | 34.57 (17.83) | 35.23 (17.76) | 35.94 (17.69) | 36.55 (17.51) | 37.27 (17.62) | 15 | 110 |
| Female | 57.18% | 55.26% | 55.49% | 53.97% | 53.93% | 0 | 1 |
| Black | 81.51% | 83.99% | 82.05% | 82.54% | 82.07% | 0 | 1 |
| Coloured | 8.07% | 7.79% | 9.03% | 8.51% | 8.91% | 0 | 1 |
| Indian | 2.13% | 2.14% | 1.96% | 2.15% | 2.1% | 0 | 1 |
| Employed | 42.63% | 36.52% | 40.25% | 45.99% | 44.81% | 0 | 1 |
| Married | 38.75% | 35.43% | 34.35% | 31.02% | 28.52% | 0 | 1 |
| Active lifestyle | 14.35% | 9.38% | 12.77% | 15.62% | 14.61% | 0 | 1 |
| Religious intensity - at least important | 88.95% | 90.5% | 90.63% | 92.25% | 90.31% | 0 | 1 |
| Perceived health – at least good | 84.6% | 91.26% | 89.78% | 89.13% | 89.88% | 0 | 1 |
| Perceived financial well-being - below average income | 47.29% | 43.98% | 46.87% | 42.41% | 38.41% | 0 | 1 |
| Two-year financial well-being | 3.65 (1.15) | 3.32 (1.22) | 3.4 (1.08) | 3.64 (1.14) | 3.58 (1.17) | 1 | 6 |
| Violence level | 2.93 (1.23) | 2.91 (1.11) | 3.17 (1.05) | 3 (1.03) | 3.18 (1.07) | 1 | 5 |
| Hours worked | 13.25 (20.4) | 11.65 (17.99) | 14.43 (20.54) | 16.68 (21.52) | 15.7 (20.88) | 0 | 168 |
| Household income (per capita) | 2768.17 (4747.2) | 3134.53 (8222.83) | 2973.8 (3970.38) | 3469.2 (6715.53) | 3530.85 (10984.65) | 0 | 899 418 |
| Years of education | 9.88 (4.32) | 9.86 (4.19) | 10.01 (4.03) | 10.31 (3.85) | 10.56 (3.67) | 0 | 13 |
| Tertiary education | 12% | 12.21% | 12.93% | 15.14% | 15.53% | 0 | 1 |

**Table 1: Summary statistics of key variables**

| Explanatory variable | Mean (standard deviation) Wave 1 2008 | Mean (standard deviation) Wave 2 2010-11 | Mean (standard deviation) Wave 3 2012 | Mean (standard deviation) Wave 4 2014-15 | Mean (standard deviation) Wave 5 2017 | Min Wave 1 - 5 | Max Wave 1 - 5 |
|---|---|---|---|---|---|---|---|
| Mpumalanga | 7.86% | 9.02% | 7.79% | 7.82% | 8% | 0 | 1 |
| Free State | 5.88% | 5.44% | 6.14% | 6.01% | 5.48% | 0 | 1 |
| Limpopo | 11.36% | 10.69% | 10.49% | 10.09% | 9.81% | 0 | 1 |
| Kwazulu Natal | 17.99% | 19.62% | 18.4% | 19.31% | 19.19% | 0 | 1 |
| Northern Cape | 2.34% | 2.4% | 2.14% | 2.25% | 2.36% | 0 | 1 |
| Eastern Cape | 12.37% | 10.96% | 12.01% | 12.01% | 11.85% | 0 | 1 |
| Western Cape | 9.84% | 8.55% | 11.16% | 11.29% | 11.63% | 0 | 1 |
| Gauteng | 25.82% | 26.54% | 25.69% | 24.98% | 25.72% | 0 | 1 |
| North West | 6.54% | 6.79% | 6.18% | 6.23% | 5.96% | 0 | 1 |

*Source: Analysis of NIDS data*
*Notes: Panel weights used to generate all mean values. Information based on adults only, no proxy or children information used. Only adults above 15 used in the analysis. Only adults with depression scores used in the analysis.*

### d. Robustness

In order to ensure the robustness of results, it is important to consider the robustness of the information being used from the NIDS database. The depression score is calculated from the CESD-10 questions in the NIDS database. If this depression score is not robust, the interpretation of results cannot be relied upon.

Baron, et al. (2017) have investigated the CESD-10 depression score as a useful tool in identifying depressive symptoms in South Africa. In doing so, a comparison to the Patient Health Questionnaire (PHQ) was completed. Conclusions indicate that the CESD-10 score is internally consistent and provide similar results to the PHQ. Baron, et al. (2017) considers that the "*CES-D-10 is a valid, reliable screening tool for depression in Zulu, Xhosa and coloured Afrikaans populations*", given that these are the focus demographics of the analysis.

In addition, Baron, et al. (2017) notes that the validity of the CESD-10 score has been assessed more generally. Findings indicate that the CESD-10 score has good psychometric properties in both the healthy and psychiatric individuals, as well as adolescents and elderly (Björgvinsson et al., 2013; Boey, 1999; Bradley et al., 2010; Irwin et al., 1999; Lee and Chokkanathan, 2008).[12]

[12] References from Baron, et al. (2017)

# 4. Methodology

The methodology section below considers Ordinary Least Squares, System-GMM and HBIV methods, which have successive improvements in terms of efficiency. Each consider a limitation that the previous is unable to.

### a. Ordinary least squares

Equation (1) below provides the baseline OLS model that can be used to describe the relevant relationship. This model seeks to describe the relationship between contrast and depressive symptoms, while controlling for a host of other factors that may be determinative of depression.

$$depression\_score_{it} = \beta_0 + \beta_1 morehappy_{it} + \beta_2 lesshappy_{it} + \beta_{3j} X_{itj} + \mu_{it} \quad (1)$$

The factor $depression\_score_{it}$, which is the dependent variable, accounts for the CESD-10 depression score. The two independent variables which are of importance are $morehappy_{it}$ and $lesshappy_{it}$, which are binary variables indicating whether an individual is more happy or less happy than they were ten years ago (i.e. perceived change in happiness). Excluded from the above model (due to collinearity) is the comparator group, which is those individuals that do not consider their level of happiness has changed in ten years.

The set of variables $X_{it}$ are the independent variables (i.e. exogenous explanatory regressors). Socio-economic factors such as education, gender, race, household income per member, the number of assets, relative perceived socio-economic status, age and immigrant status have been studied before, therefore providing an indication that socio-economic factors that are similar to the above should be included in the analysis (Akhtar-Danesh and Landeen, 2007; Ardington and Case, 2010; Everson et al., 2002; Miech and Shanahan, 2000). Marital status, age, working hours and perceived health have also been assessed as a determinant of depression (Mirowsky and Ross, 1992; Mui, 1996; Pearlin and Johnson, 1997; Virtanen et al., 2011; Virtanen et al., 2012). A full description of each variable included in the analysis is provided in Annex A.

While the OLS model above provides a baseline model for comparison, there are two major problems with using OLS. OLS does not control for possible endogeneity bias concerns that have been previously raised nor does it consider the panel nature of the data. In order to control for these factors specialised econometric methods are considered below.

### b. Endogeneity

Endogeneity may generally come in the form of omitted variable bias, simultaneity and measurement error (Wooldridge, 2010). The specific form of endogeneity in these models can be explained by the series of equations below.

$$depression\_score_{it} = \beta_0 + \beta_1 morehappy_{it} + \beta_2 lesshappy_{it} + \beta_{3j} X_{itj} + \mu_{1it} \quad \text{where } \mu_{1it} = \theta_1 M + v_1 \tag{2}$$

$$morehappy_{it} = \beta_0 + \beta_1 depression\_score_{it} + \beta_{3j} X_{itj} + \mu_{2it} \quad \text{where } \mu_{2it} = \theta_2 M + v_2 \tag{3}$$

$$lesshappy_{it} = \beta_0 + \beta_1 depression\_score_{it} + \beta_{3j} X_{itj} + \mu_{3it} \quad \text{where } \mu_{3it} = \theta_3 M + v_3 \tag{4}$$

The interpretation of the equations above are as per previously discussed, however, the only difference is now a more complex error term and two additional equations which consider the potentially reverse relationship. From equation (2), equation (3) and equation (4) above, each of the error terms include a common factor (i.e. $M$).

These common factors could result in the following form of endogeneity. For example, let us consider that $M$ increases in equation (3), which is positively related to the dependent variable, causing an increase in the likelihood of $morehappy_{it.}$ On the other hand, $M$ has also increased in equation (2), therefore increasing depressive symptoms. However, since $M$ is omitted from the analysis, and that $M$ is also included in equation (3), econometric analysis includes a positive bias to the $\beta_1$ in equation (2). This positive bias comes from the statistical relationship between the two variables, given the omission of $M$. The common factor can take the form of the following (with the list not being exhaustive):

- Ability to recall: As pointed out above, contrast is related to depression through memory, however, the ability to recall is also likely to affect depression. If an individual has chronic memory issues (perhaps Alzheimer's or Dementia), this may affect the likelihood of depressive symptoms in an individual, possibly by affecting the ability to recall (Bennet and Thomas, 2014; Cipriani et al., 2015). Additionally, given that contrast is defined as perceived change in happiness, it is likely to be affected by chronic memory issues, since perceptions of past levels of happiness are directly related to memory. While chronic memory issues are provided as an example, this does not necessitate that this is the only way in which the ability to recall may be a common factor (i.e. a memory bias could also facilitate as a common factor).

- Emotional state when answering questions: A heightened emotional state may affect the way the participants have answered the questions, especially considering that both the contrast questions and the CESD-10 questions both relate to past events. For instance, if an individual is sad, the individual may recall more negative memories, and as such, answer both questions with a bias. Deptula, et al. (1993) show that negative affective states relate negatively to recall (amongst the elderly); and Leight and Ellis (1981) point out that many studies have shown that emotional state is a significant variable when it comes to memory and cognitive activities. In addition, individuals generally have a positive memory bias, which may affect the way all questions were answered (Marsh et al., 2019).
- Current social support: Current social support is likely to aid in the onset of depression, reducing the likelihood of an individual being depressed (Lin and Dean, 1984; Paykel, 1994; Revenson et al., 1991; Stice et al., 2004).[13] In addition, social support affects subjective well-being / happiness, therefore, current social support is likely to affect perceived change in happiness (Brannan et al., 2012; Li et al., 2014; Siedlecki et al., 2014). For instance, if current social support is quite high, this is likely to increase subjective well-being / happiness, and therefore, influence perceived change in happiness.

### c. System-GMM

System-GMM, as proposed by Blundell and Bond (1998), makes use of the specification by Anderson and Hsiao (1981) and Arellano and Bond (1991), by increasing the efficiency of the estimation method. System-GMM implements both a levels and differences equation when instrumenting, with lagged levels being used to instrument the difference equation and lagged differences being used to instrument the levels equation (Ćorić and Šimić, 2021). While it is not the purpose of the current analysis, this also provides an added benefit of the inclusion of a dynamic term, which is the lag of the binary depression variable.

System-GMM uses instrumental variable techniques to remove any fixed effects bias. This methodology considers the panel nature of the dataset. Both are not considered by OLS. Generally, instruments are lags of the dependent variable, however, more than one type of instrument can be used. Given the choice on instrument highlighted below (i.e. the lag of the

[13] While there are variables relating to trust in the NIDS database, information relating to social support is not available.

endogenous variable), no additional instruments need to be added to the system-GMM methodology. As highlighted by Blundell and Bond (1998), additional variables can be considered endogenous, and controlled for. Therefore, both contrast variables are treated as endogenous variables.[14]

In the current analysis, the endogeneity bias *may* be removed via System-GMM; however, this is only the case if the factor *M* (see equation (2), equation (3) and equation (4) above) is time invariant. This would allow for a causal interpretation of the endogenous binary regressors. Given the existence of panel data, there are *also* likely to exist different levels of coping mechanisms for depressive symptoms, that are fixed for individuals over time. While there could be more arguments for the factor *M* to be time invariant, this is unlikely the case for the last two examples provided in the sub-section above.[15]

Additionally, System-GMM estimates can be biased when instruments are weak (Ćorić and Šimić, 2021). Given the findings highlighted below, in terms of weak instruments, this poses a further problem. One that cannot be tested since the weak instrument test is currently unavailable for System-GMM (Ćorić and Šimić, 2021). Therefore, while System-GMM can be implemented, a more robust method is required for refined results (see HBIV section below).

While fixed effects could be theorised, it is also important to consider whether they do exist on an empirical basis, therefore suggesting that the System-GMM approach is relevant. The pooled OLS analysis is not useful in this regard, however, a comparison between panel data fixed effects and random effects is useful.[16] In this regard, both the fixed effects and random effects panel data regression are completed, using all covariates. The Hausman (1978) test is conducted to assess whether the coefficients are statistically significant. Overall, the null hypothesis is rejected, suggesting the existence of fixed effects.[17] The regression and test results are available in Annex C.

System-GMM requires several post-estimation criteria to be met. Table 2 below provides the post-estimation statistics. The Arellano-Bond tests suggest first order autocorrelation exists (i.e. null hypothesis rejected at the 1% level), however, not second order autocorrelation (i.e.

[14] As with the SR method above, wave is used as the time variable instead of the year.
[15] In addition, other types of fixed effects could be included, such as individual level fixed effects such as ability to deal with depression, personality types, etc.
[16] This is also useful when considering the fixed effects specification for HBIV.
[17] Regression statistics are omitted since the purpose of the regression was only to test for fixed effects.

null hypothesis not rejected at the 10% level).[18] This is consistent with the model specification and does not suggest the reduction in the number of internal instruments is necessary (Roodman, 2009). The Hansen-J test (preferred to the Sargan test in two step estimation) for instrument validity is not rejected at the 10% level, therefore suggesting that the list of instruments is exogenous (Roodman, 2009).[19] The Difference in Hansen tests provide similar results and interpretations. Lastly, the Wald test suggests that the model has a good fit (i.e. at the 1% level).[20]

**Table 2: Post estimation statistics for System-GMM regression**

| | Full sample | Incorrect recall | Correct recall |
|---|---|---|---|
| AR(1) p-value | 0.000 | 0.000 | 0.000 |
| AR(2) p-value | 0.756 | 0.469 | 0.937 |
| Hansen-J test p-value | 0.600 | 0.131 | 0.747 |
| Difference in Hansen p-value | 0.668 | 0.731 | 0.746 |
| Wald test p-value | 0.000 | 0.000 | 0.000 |

Notes: Standard errors are clustered around the internal NIDS cluster variable. Panel weights from the NIDS wave 5 data are used for national inference and to control for attrition. Differences and lags of the lags of the dependent variable and endogenous regressors are used as instruments. All other explanatory regressors are also included as instruments. The two-step System-GMM estimator being used since we cannot assume that the errors are homoscedastic.

### d. HBIV

The HBIV method was first proposed by Lewbel (2012).[21] This is similar to the general instrumental variable (IV) approach, however, provides some additional benefits and has some additional requirement (all of which are outlined below).[22] Notably, whereas System-GMM fails to control for time variant endogeneity bias and weak instruments, the HBIV method does not include this shortfall (Ćorić and Šimić, 2021).

In order to find an instrument for the HBIV analysis, both instrumental relevance and instrumental validity had to be considered (for more on this, see Annex B). For this analysis, the first lag of the endogenous variables are used as instruments; a common approach (Laverde-Rojas et al., 2019). The correlation coefficient between the more happy variable and its first lag is 0.1141; and the correlation coefficient between the less happy variable and its first lag is

[18] The null hypothesis for this test is that there is no serial correlation in the error term.
[19] The null hypothesis for this test is that the instruments are valid.
[20] The null hypothesis for this test is that all coefficients are jointly zero.
[21] Rank Order IV has also been considered; however, this has not been used since the approach is not suitable when the endogenous regressors are binary. In order to confirm this, the creators of the method were contacted. The creators confirmed that the ROIV method is not applicable for binary endogenous regressors. Applicability of the method was not available from online sources (Vella & Verbeek, 1997).
[22] IV methods can be "*ill-behaved*' when instruments are weak (Wooldridge, 2010). Given the potential for weak instruments in the current analysis, the general IV method cannot be used.

0.0649. As is clear, the endogenous regressors and instruments are positively correlated, although the correlation is weak. This finding suggests that even though instruments can be included, the coefficients would be bias towards the OLS coefficients given that the instruments are weak (Wooldridge, 2010).

The HBIV method can be used when there are no external instruments available, or when the efficiency of IV regression needs to be enhanced. Rashad and Markowitz (2007), use the HBIV method in the study of obesity and health insurance, in order to enhance the efficiency of the IV regression due to the potential downfalls of having weak instruments. Therefore, the usage of the HBIV methodology is required in the current analysis.

The HBIV method identifies coefficients by "*having regressors that are not correlated with the product of heteroskedasticity errors*" (Laverde-Rojas et al., 2019). Implementation is completed via regressing all exogenous variables on each endogenous regressor. This allows for the extraction of estimated errors from the first stage regression. Instruments are then constructed by using the product of the first stage regression errors and the demeaned exogenous variables (Lewbel, 2012).

Based on the identification method explained above, identification is only achieved if there is a heteroscedasticity present in the model (Laverde-Rojas et al., 2019). The Breusch-Pagan (1979) test, with a null hypothesis of homoskedasticity, is sufficient to test for heteroskedasticity (Lewbel et al., 2012). The Breusch-Pagan/Godfrey/Cook-Weisberg post-estimation test is conducted, after general IV regression with the external instruments (i.e. the lagged endogenous variables). The null hypothesis is rejected at the 1% level. This indicates that heteroscedasticity is available in the model to be exploited.[23]

In addition to the above, the removal of the endogeneity bias is dependent on the structure of the type of endogeneity that is present in the model. More specifically, the type of endogeneity that exists, must be common to both the model of interest, and the model of the reverse relationships. This is exactly the type of endogeneity bias that has been explained in equation (2), equation (3) and equation (4) above, where the common factor that causes the endogeneity

[23] While the test could be run on each first stage regression separately, via OLS; it was more efficient to run the test once the IV regression was completed. A full sample (test value of 2916.36) and a sample limited to the fifth wave of the data (test value of 848.95) was used to conduct the test, with the null hypothesis being rejected in both instances.

bias is *M*. The existence of this common factor aids in the result of identifying the structural parameters of the equation (Lewbel, 2012).

The HBIV method requires post-estimation statistics for the Kleibergen-Paap test for weak identification (i.e. the instrumental relevance assumption above).[24] The null hypothesis on this test is that the instruments are weak.[25] As is evident from Table 3 below, all of the models assessed reject the null hypothesis of weak instruments at the 5% level. This suggests that the instruments are relevant.

**Table 3: Post-estimation statistics for HBIV regression**

| | **Kleibergen-Paap** | | **Hansen-J** | | **Difference in Sargan** | |
|---|---|---|---|---|---|---|
| | **Test value** | **The 5% critical value** | **Test value** | **P-value** | **Test value** | **P-value** |
| Full sample | 45.61 | 21.03 | 61.31 | 0.1311 | 2.667 | 0.2636 |
| Incorrect recall subset | 22.38 | 20.86 | 35.96 | 0.1436 | 0.587 | 0.7457 |
| Correct recall subset | 20.96 | 20.86 | 37.26 | 0.1133 | 1.628 | 0.4431 |

Notes: Standard errors are clustered around the internal NIDS cluster variable. Panel weights from the NIDS wave 5 data are used for national inference and to control for attrition. For the full sample, external instruments include the first lag of the endogenous regressors, whereas internal instruments are generated within the model and include all other regressors in the model. The subset samples differ in that the internal instruments based on the regressors included in the model, are only those that are considered purely exogenous. Those considered purely exogenous include the following: age, age squared, female dummy variable, black dummy variable, violence level in community, province dummy variables, and year dummy variables. For completeness, the full sample analysis is also conducted using only those regressors considered purely exogeneous, as instruments; in this case, the test value for the Kleibergen-Paap test is 54.13 and the 5% critical value is 20.86; the test value for the Hansen-J test is 28.45 and the p-value is 0.4406 and the test value for the Difference in Sargan test is 0.035 and the p-value is 0.9828. All results are consistent with the full sample analysis including all potential instruments.

Lastly, due to the creation of internal instruments by the HBIV method, the validity of instruments can be tested (i.e. the instrumental validity assumption above). Table 3 above also provides the post-estimation statistics for the Hansen-J and Difference in Sargan test. The Hansen-J test fails to reject the null hypothesis (at the 10% significance level) that the instruments are valid for all models. The Difference in Sargan test, which tests the validity of only the two external instruments is also completed. This test does not reject the null hypothesis at the 10% significance level for all models. Both tests suggest that the instruments are valid.

[24] In addition to the tests mentioned in this section, the Kleibergen-Paap test of under-identification is also completed. In each regression, the null hypothesis of under-identification is rejected at the 1% level.
[25] Given the inclusion of clustered standard errors, Cragg-Donald test is not reliable, therefore, the Kleibergen-Paap test is completed. As commonly done, focus is on the maximal IV relative bias.

# 5. Results and discussion

Results are presented for each of the selected methodologies mentioned above. First, results are provided for the full sample, secondly, results are provided for two subsets as a robustness check.

### a. Full sample

Table 4 below shows the results for all methods discussed. OLS (1) differs from OLS (2) in that the former only includes the endogenous regressors, whereas the latter includes all covariates.[26] System-GMM and HBIV analysis are also included. Standard errors are clustered around the internal NIDS cluster variable, and fixed effects specification are included for the HBIV methods.

The relevant relationship is between contrast and depressive symptoms. This is captured by the two dummy variables more happy and less happy. In relation to the binary variable for more happy, the coefficient is negative across all models (and statistically significant at least at the 5% level), suggesting that individuals who consider that they are more happy now than they were ten years ago, have a slightly lower depression score, and as such, lower depressive symptoms, than those who consider that their level of happiness has not changed over the last ten years. This is consistent with Figure 2 above.

This relationship is consistent with the literature around depression. Specifically, the cognitive theory of depression and the theory of hopelessness indicate that memory plays a role in determining depression, especially the perception of life events, which in turn relate to change in happiness (Abramson et al., 1989; Alloy et al., 1987; Alloy et al., 1988; Beck, 1967). The direction of the effect of the more happy variable is also consistent with the hypothesis, specifically, those who are perceive they are more happy, consider their happiness has increased over time, and as such, have lower depressive symptoms (Seo et al., 2018).

The coefficient on OLS (2) is smaller than the coefficient on OLS (1) suggesting that the inclusion of covariates has reduced omitted variable bias. System-GMM provides a much larger negative coefficient, suggesting that the fixed effects provided a large positive bias to the results. However, when turning to HBIV methodology, which also controls for the time

[26] Both regressions are pooled OLS regression.

variant endogeneity (as specified by the common factor) and aids when instruments are weak, we find that the coefficient reduces in absolute size. This suggests that time variant and time invariant effects have differing signs in terms of the endogeneity bias they generate or that weak instruments were a concern. HBIV analysis suggests that the effect of more happy on depressive symptoms is -1.526.

**Table 4: Full sample regression estimates**

| Dependent variable: depression score | **OLS (1)** | **OLS (2)** | **System-GMM** | **HBIV** |
|---|---|---|---|---|
| More happy | -1.392***<br>(0.0738) | -0.986***<br>(0.0827) | -5.362**<br>(2.597) | -1.526***<br>(0.347) |
| Less happy | 1.372***<br>(0.115) | 1.094***<br>(0.112) | 2.557<br>(3.662) | 0.482<br>(0.471) |
| Depression score – lag | | | -0.0112<br>(0.0156) | |
| Health perception – at least good | | -1.322***<br>(0.132) | -0.324<br>(0.275) | -0.880***<br>(0.156) |
| Religious intensity – at least important | | -0.261**<br>(0.119) | 0.198<br>(0.287) | -0.306*<br>(0.177) |
| Age | | 0.125***<br>(0.0101) | 0.0824***<br>(0.0260) | -0.0237<br>(0.156) |
| Age – squared | | -0.00128***<br>(0.000123) | -0.000878***<br>(0.000250) | -0.00122**<br>(0.000522) |
| Female | | 0.238***<br>(0.0559) | 0.358***<br>(0.0989) | |
| Black | | 0.999***<br>(0.159) | 0.999***<br>(0.204) | |
| Hours worked | | -0.00294<br>(0.00300) | -0.00234<br>(0.00430) | -0.00432<br>(0.00455) |
| Active lifestyle | | -0.701***<br>(0.0939) | -0.557***<br>(0.164) | -0.632***<br>(0.164) |
| Perceived income – below average | | -0.0383<br>(0.0859) | -0.625<br>(0.412) | -0.102<br>(0.124) |
| Financial well-being – 2-year expectation | | -0.161***<br>(0.0461) | 0.0256<br>(0.0922) | -0.145**<br>(0.0650) |
| Employed | | -0.216*<br>(0.122) | -0.170<br>(0.201) | -0.143<br>(0.223) |
| Neighbourhood violence level | | 0.201***<br>(0.0414) | 0.101<br>(0.0744) | 0.0958<br>(0.0657) |
| Married | | -0.562***<br>(0.0771) | -0.183<br>(0.227) | -0.330*<br>(0.193) |
| Household income per capita – log | | -0.118***<br>(0.0386) | 0.0852<br>(0.0770) | -0.0304<br>(0.0671) |
| Years of education | | -0.0437***<br>(0.0103) | -0.0518***<br>(0.0186) | -0.0849<br>(0.0577) |
| Tertiary education | | -0.183*<br>(0.0990) | -0.232<br>(0.161) | -0.365<br>(0.250) |
| Observations | 61,953 | 55,207 | 36,603 | 36,217 |
| R squared | 0.059 | 0.122 | - | 0.047 |

**Table 4: Full sample regression estimates**

| Dependent variable: depression score | **OLS (1)** | **OLS (2)** | **System-GMM** | **HBIV** |
|---|---|---|---|---|
| Year dummy variables | No | Yes | Yes | Yes |
| Province dummy variables | No | Yes | Yes | Yes |
| Robust standard errors | Yes | Yes | Yes | No |
| Fixed effects specification | No | No | - | Yes |
| Clustering | Yes | Yes | Yes | Yes |
| Panel weights | Yes | Yes | Yes | Yes |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Standard errors are clustered around the internal NIDS cluster variable. Panel weights from the NIDS wave 5 data are used for national inference and to control for attrition. For System-GMM, differences and lags of the lags of the dependent variable and endogenous regressors are used as instruments. All other explanatory regressors are also included as instruments. The two-step System-GMM estimator being used since we cannot assume that the errors are homoscedastic. For HBIV, external instruments include the first lag of the endogenous regressors, whereas internal instruments are generated within the model and include all other regressors in the model. For HBIV, as a robustness check, only those purely exogenous variables are included as instruments to understand whether there is any substantial difference in the coefficients. This regression includes fewer instruments. This comes at a loss of efficiency, however, is used as a robustness test against the full model (Lewbel, 2012). The analysis does not show a material difference in any of the coefficients. The co-efficient for more happy is –1.4844 (significant at the 1% level) and the coefficient for less happy is 0.3569 (not significant at the 10% level).

Unfortunately, none of the specialised models show a statistically significant coefficient for the less happy variable. However, both OLS models do show statistically significant coefficients, at the 1% level. Notably, even though System-GMM and HBIV do not have a statistically significant coefficient, the coefficient is always positive. In this regard, the expectation from the hypothesis is met (however, a conclusion is not possible), and the finding is consistent with Figure 2 above.

The discussion of the remaining regressors is limited to those that are statistically significant in either the System-GMM analysis or the HBIV analysis. For those individuals that perceive they have at least good health, the analysis suggests that they have lower depressive symptoms than those that do not. Similarly, those who have an active lifestyle have a lower depression score (and as such, lower depressive symptoms), than those that do not; which is consistent with the findings of Burger, et al. (2017) and Cooney, et al. (2013). Specifically, exercise may act as a diversion to negative thoughts, provide a social contact for social support, and release helpful hormones (Chen, 2013; Cooney et al., 2013; LePore, 1997).

In terms of race, consistent with Ardington and Case (2010) and Burger, et al. (2017), the analysis suggests that black individuals have higher depressive symptoms than other races. As also highlighted in Burger, et al. (2017) and Tomlinson, et al. (2009), females have higher depressive symptoms than males. These differences are likely attributable to the differing socioeconomic status when it comes to race and gender (Ardington and Case, 2010).

In addition to the above, both the System-GMM and the HBIV method posit a quadratic relationship for age, with the squared age variable being negative. Additionally, the coefficient on the age variable is positive in the System-GMM regression (HBIV is not statistically significant). This indicates that depressive symptoms increase with age, until a maximum is reached, after which depressive symptoms decrease with age. This finding is *inconsistent* with finding of Mirowsky and Ross (1992), however, that analysis is based in the United States. The analysis is consistent with the findings of Burger, et al. (2017), a South African study.

Consistent with the findings of Ardington and Case (2010), Burger, et al. (2017), Hamad, et al. (2008) and Tomlinson, et al. (2009), the System-GMM regression suggests that education plays a role in reducing depression, although the role is limited. Education may provide the necessary mechanisms to fight diseases (depression included) given the information learnt (Ardington and Case, 2010). Being married is also associated with having lower levels of depressive symptoms, a similar result is found by Akhtar-Danesh and Landeen (2007), and Pearlin and Johnson (1997). This is realised through the social support structure of marriage.

Lastly, religious intensity and expected financial well-being are also important determinants of depressive symptoms. Those individuals that consider religion to be important (at the least) have lower depressive symptoms. Having a higher expected future well-being reduces depressive symptoms based on the HBIV regression results. This result is consistent with Andington and Case (2010), where reporting that an individual is low on the socioeconomic status ladder relates to higher depressive symptoms.

Overall, the results suggest that having a positive mentality (through perceived change in happiness, expected financial well-being and perceived health status), active lifestyle, investing in education, being married, and considering religion as important, plays a role in reducing depressive symptoms. The effects on females and black individuals require further research in order to understand what measures can be taken in order to remove the unfair disposition.

### b. Correct and incorrect recall

It is also important to understand the effect of perceived change in happiness on those who have correct recall compared to those who do not have correct recall. This will further aid in the understanding of the role of perceptions. Given the existence of the happiness / subjective well-being variable in the NIDS dataset, the correct change in happiness can be calculated and compared to the perceived change in happiness.[27] Notably, this can only be calculated for the individuals in wave five of the NIDS survey. Therefore, in order to run panel data analysis, a *large* assumption is required. Specifically, those that have correct recall in wave five, also have correct recall in all other waves (and similarly for incorrect recall).

Table 5 presents the results from OLS, System-GMM and HBIV regression for correct recall and incorrect recall. All covariates are included in the analysis; however, the coefficients are omitted from the table below. System-GMM does not show a statistically significant coefficient on either endogenous regressor. The remainder of the results indicate that the *absolute* effect on those who have correct recall is slightly larger than those who have incorrect recall. However, this difference is not statistically significant for the HBIV method (for more happy). These results suggest that perceptions play an important role and further illustrate the hypothesis, as identified by the literature (i.e. outlined above in the main analysis).

One of the most important findings relates to the statistical significance of the less happy variable. In this instance, perceptions do not play a major role as a determinant of depressive symptoms, since the incorrect recall is *not* statistically significant, however, correct recall is statistically significant at the 10% level. This highlights that perceptions may provide a positive effect (i.e. via more happy), however, there is little evidence that they may have a negative effect (i.e. via perceived incorrectly less happy).[28]

**Table 5: Correct and incorrect recall regression estimates**

| Dependent variable: depression score | **OLS Incorrect recall** | **OLS Correct recall** | **System-GMM Incorrect recall** | **System-GMM Correct recall** | **HBIV Incorrect recall** | **HBIV Correct recall** |
|---|---|---|---|---|---|---|
| More happy | -0.777*** | -1.093*** | -1.871 | -3.009 | -1.565*** | -1.690*** |
| | (0.106) | (0.148) | (3.817) | (2.188) | (0.498) | (0.509) |
| Less happy | 0.891*** | 1.429*** | 1.939 | 1.397 | 0.429 | 1.403* |
| | (0.141) | (0.190) | (2.540) | (2.677) | (0.696) | (0.801) |
| Observations | 19,659 | 15,065 | 13,629 | 10,522 | 13,536 | 10,416 |
| R squared | 0.115 | 0.143 | - | - | 0.030 | 0.077 |
| All other covariates included | Yes | Yes | Yes | Yes | Yes | Yes |
| Year dummy variables | Yes | Yes | Yes | Yes | Yes | Yes |

[27] The relevant question is "*Using a scale of 1 to 10 where 1 means 'Very dissatisfied' and 10 means 'Very satisfied', how do you feel about your life as a whole right now?*"
[28] OLS is not considered given the shortcomings of this approach.

**Table 5: Correct and incorrect recall regression estimates**

| Dependent variable: depression score | OLS Incorrect recall | OLS Correct recall | System-GMM Incorrect recall | System-GMM Correct recall | HBIV Incorrect recall | HBIV Correct recall |
|---|---|---|---|---|---|---|
| Province dummy variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Robust standard errors | Yes | Yes | Yes | Yes | No | No |
| Fixed effects specification | No | No | - | - | Yes | Yes |
| Clustering | Yes | Yes | Yes | Yes | Yes | Yes |
| Panel weights | Yes | Yes | Yes | Yes | Yes | Yes |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Standard errors are clustered around the internal NIDS cluster variable. Panel weights from the NIDS wave 5 data are used for national inference and to control for attrition. For System-GMM, differences and lags of the lags of the dependent variable and endogenous regressors are used as instruments. All other explanatory regressors are also included as instruments. The two-step System-GMM estimator being used since we cannot assume that the errors are homoscedastic. For HBIV, external instruments include the first lag of the endogenous regressors, whereas internal instruments are generated within the model and include all other purely exogenous regressors in the model. Those considered purely exogenous include the following: age, age squared, female dummy variable, black dummy variable, violence level in community, province dummy variables, and year dummy variables.

Considering the findings above, the analysis suggests that perceived change in happiness (at least for those that consider they are more happy), may be used as a tool in combatting depressive symptoms. For instance, for those individuals who are faced with depressive symptoms, a consistent and combined focus on neutral or less than ideal memories of the *past* and positive memories of the *present* (i.e. possibly through a diary), may provide a reduction in depressive symptoms. However, given the potential for the moderate magnitude of this effect, this may be well suited in combination with other types of therapy.

To the extent that memory training techniques (i.e. through the ability to influence contrast) could be used to combat depressive symptoms, the current analysis provides an indication that this could be useful. However, Scogin, et al. (1985) assessed the relationship between memory training and depression and found no significant relationship. It is possible that the type of memory training is important in combatting depression, hence the difference in results. It seems to be the case, given that Dalgleish, et al. (2013) indicate that "*richly elaborated, concrete positive, or self-affirming memories in those with a history of depression can have a self-reported beneficial effect on mood.*" Dalgleish, et al. (2013) devise a new method of memory training for depressed individuals to be more easily able to access positive memories. The analysis highlights that this type of memory training aids in the "*recall of positive self-affirming memories*". This highlights that firstly, memory training is possible and secondly, memory training has been previously successful. The memory training techniques for perceived change in happiness (as above) could be useful in combatting depression.

# 6. Limitations and further research

### a. Limitations

While the results from the research above have been useful, there are several limitations that should be noted along with the results. First, Watkins and Teasdale (2001) have pointed out that distractions may influence over-general memory that depressed individuals exhibit. This may affect the answers provided by depressed individuals. Second, the analysis above comparing correct recall and incorrect recall assumes the classification of wave five for all other waves. Reliability of results is based on this assumption.

Third, the analysis assessed the effect of perceived change in happiness over a substantial amount of time (i.e. ten years). The analysis does not account for the change in happiness over differing time horizons, therefore, the interpretations of the analysis should not be over generalized. Fourth, the HBIV approach, while extremely useful, is less reliable than alternative approaches, given the reliance upon higher moments in the restriction conditions (Lewbel, 2012).

Fifth, the type of endogeneity bias mentioned above, is from the existence of omitted common factors that results in bias coefficients. However, to the extent that the there exists endogeneity through other means (i.e. non common factor), the endogeneity bias may persist.

### b. Further research

While the limitations should be considered when interpreting the results, they are also useful in understanding the type of further research that can be conducted in this field. Firstly, further research could focus on the effect of perceived change in happiness while also controlling for the actual change in happiness as reported by the individual. This could aid in removing any potential omitted variable bias from ability to recall, directly. Secondly, the time horizon for perceived change in happiness is substantial in the current analysis (i.e. ten years). In order to understand how short-term memory (or short-term contrast) may affect the likelihood of depression, further analysis could be conducted with varying time lengths. Thirdly, a database that includes a clinical diagnosis of depression can be used in order to assess the direct relationship between contrast and depression. Fourth, additional instruments could be retrieved in different datasets in order to strengthen the analysis.

# 7. Conclusion

This new study seeks to understand the relationship between perceived change in happiness and depression; one that has not been considered sufficiently in the literature. In doing so, System-GMM and HBIV are used in order to control for differing forms of endogeneity bias, which is a concern in many econometric studies. OLS is used as the baseline analysis.

The analysis shows that there is a negative relationship between those who perceive they are more happy (than ten years ago) and depressive symptoms, when comparing to a those who do not perceive a change in happiness. Unfortunately, the results for those who consider they are less happy are inconclusive. When splitting the regression between those that have correct recall and those that have incorrect recall, the same finding is evident for those who perceive they are more happy. Additionally, for those who are less happy and have correct recall, the analysis shows a positive effect on depressive symptoms, however, the no effect for those with incorrect recall.

The findings suggest that perceptions (and memory) play a material role in determining depressive symptoms (although only for those who have positive perceptions). These findings can be used to inform decisions on how depressive symptoms should be combatted. One potential avenue of thought is the use of memory training. As highlighted above, from a policy perspective, memory training that aids in increasing the accessibility of positive memories may be useful in alleviating depressive symptoms.

The remainder of the coefficients are as highlighted in the literature, with health perception, religious intensity, active lifestyle, expected financial well-being, employment, marriage, income and education, being negatively related to depressive symptoms and being female, black and the neighbourhood violence level being positively related to depressive symptoms. Lastly, as expected, age is shown to have a quadratic relationship with depressive symptoms.

Overall, the analysis reiterates the findings in literature, while providing an additional relationship that can be used to combat depression. This is of course an important finding, given the high levels of depression already exhibited in South Africa.

# References


Abramson, L. Y., Metalsky, G. I. & Alloy, L. B., 1989. Hopelessness depression: A theory based subtype of depression. *Psychological Review,* Volume 96, pp. 358-372.

Akhtar-Danesh, N. & Landeen, J., 2007. Relation between depression and sociodemographic factors. *International Journal of Mental Health Systems*, 1(4), pp. 1-9.

Alloy, L. B., Abramson, L. Y., Metalsky, G. I. & Hartlage, S., 1988. The hopelessness theory of depression: Attributional aspects. *British Journal of Clinical Psychology,* Volume 27, pp. 5-21.

Alloy, L. B. et al., 1997. Self-referent Information-processing in Individuals at High and Low Cognitive Risk for Depression. *Cognition & Emotion,* 11(5/6), pp. 539-568.

Anderson, R. J. & Evans, G. L., 2015. Mental time travel in dysphoria: Differences in the content and subjective experience of past and future episodes. *Consciousness and Cognition,* Volume 37, pp. 237-248.

Anderson, T. W. & Hsiao, C., 1981. Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association,* 76(375), pp. 598-606.

Andresen, E. M., Malmgren, J. A., Carter, W. B. & Patrick, D. L., 1994. Screening for Depression in Well Older Adults: Evaluation of a Short Form of the CES-D. *American Journal of Preventative Medicine,* 10(2), pp. 77-84.

Ardington, C. & Case, A., 2010. Interactions between Mental Health and Socioeconomic Status in the South African National Income Dynamics Study. *Journal for studies in Economics and Econometrics,* 34(3), pp. 69-85.

Arellano, M. & Bond, S., 1991. Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies,* 58(2), pp. 277-297.

Baron, E. C., Davies, T. & Lund, C., 2017. Validation of the 10-item Centre for Epidemiological Studies Depression Scale (CES-D-10) in Zulu, Xhosa and Afrikaans populations in South Africa. *BMC Psychiatry,* 17(6), pp. 1-14.

Bateman, C., 2015. Mental health under-budgeting undermining SA's economy. *South African Medical Journal,* 105(1), pp. 7-8.

Beck, A. T., 1967. *Depression: Clinical, experimental, and theoretical aspects.* New York: Harper & Row.

Bennet, S. & Thomas, A. J., 2014. Depression and dementia: Cause, consequence or coincidence?. *Maturitas,* 79(2), pp. 184-190.

Berntsen, D., 1996. Involuntary autobiographical memories. *Applied Cognitive Psychology,* Volume 10, pp. 435-454.

Björgvinsson, T., Kertz, S. J. & Bigda-Peyton, J. S., 2013. Psychometric Properties of the CES-D-10 in a Psychiatric Sample. *Assessment SAGE Journals,* 20(4), pp. 429-436.

Blaauw, D. & Pretorius, A., 2013. The Determinants of Subjective Well-being in South Africa – An Exploratory Enquiry. *Journal of Economic and Financial Sciences,* 6(1), pp. 179-194.

Blundell, R. & Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* Volume 87, pp. 115-143.

Boey, K. W., 1999. Cross-validation of a short form of the CES-D in Chinese elderly. *International Journal of Geriatric Phychiatry,* 14(8), pp. 608-617.

Bradley, K. L., Bagnell, A. L. & Brannen, C. L., 2010. Factorial Validity of the Center for Epidemiological Studies Depression 10 in Adolescents. *Issues in Mental Health Nursing,* 31(6), pp. 408-412.

Brannan, D. et al., 2012. Friends and family: A cross-cultural investigation of social support and subjective well-being among college students. *The Journal of Positive Psychology,* 8(1), pp. 65-75.

Branson, N. & Wittenburg, M., 2019. *Longitudinal and Cross-Sectional Weights in the NIDS Data 1-5,* Cape Town: Southern Africa Labour and Development Research Unit.

Breusch, T. & Pagan, A., 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica,* 47(5), pp. 1287-1294.

Brickman, P. & Campbell, D., 1971. Hedonic relativism and planning the good society. *Adaptation level theory,* pp. 287-301.

Brickman, P., Coates, D. & Janoff-Bulman, R., 1978. Lottery winners and accident victims: Is happiness relative?. *Journal of personality and social psychology,* 36(8), pp. 917-927.

Brophy, T. et al., 2018. *National Income Dynamics Study Panel User Manual,* Cape Town: Southern African Labour and Development Research Unit.

Burger, R., Dorrit, P. & von Fintel, M., 2017. The relationship between negative household events and depressive symptoms: Evidence from South African longitudinal data. *Journal of Affective Disorders,* Volume 218, pp. 170-175.

Chen, M. J., 2013. The neurobiology of depression and physical exercise. In: *Handbook of Physical Activity and Mental Health.* London: Routledge, pp. 169-184.

Cipriani, C. et al., 2015. Depression and dementia. A review. *European Geriatric Medicine,* 6(5), pp. 479-486.

Clark, I. A., Mackay, C. E. & Holmes, E. A., 2013. Positive involuntary autobiographical memories: You first have to live them. *Consciousness and Cognition,* 22(2), pp. 402-406.

Cooney, G. et al., 2013. Exercise for depression. *Cochrane Database of Systematic Reviews,* Issue 9, pp. 1-132.

Ćorić, B. & Šimić, V., 2021. Economic disasters and aggregate investment. *Empirical Economics,* pp. 1-38.

Coyne, J. C. & Gotlieb, I. H., 1983. The role of cognition in depression: a critical appraisal. *Psychological Bulletin,* 94(3), pp. 472-505.

CSVR, 2007. *The Violent Nature of Crime in South Africa,* Johannesburg: Centre for the Study of Violence and Reconciliation.

Currie, J., 2020. Child health as human capital. *Health Economics,* Volume 29, p. 452–463.

Dalgleish, T. et al., 2013. Method-of-Loci as a Mnemonic Device to Facilitate Access to Self-Affirming Personal Memories for Individuals With Depression. *Clinical Psychological Science,* 1(2), p. 156–162.

Datafirst, 2020. *National Income Dynamics Study.* [Online]
Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/NIDS/about

Deptula, D., Singh, R. & Pomara, N., 1993. Aging, emotional states, and memory. *The American Journal of Psychiatry,* 150(3), pp. 429-434.

Ditta, A. S. & Storm, B. C., 2016. Thinking about the future can cause forgetting of the past. *The Quarterly Journal of Experimental Psychology,* 69(2), pp. 339-350.

Donohue, J. M. & Pincus, H. A., 2007. Reducing the Societal Burden of Depression. *Pharmacoeconomics,* 25(1), pp. 7-24.

Edin, P.-A., Fredriksson, P., Nybom, M. & Öckert, B., 2017. The Rising Return to Non-Cognitive Skill. *Working paper: IZA Institute of Labour Economics,* pp. 1-35.

Eich, E., Macaulay, D. & Ryan, L., 1994. Mood dependent memory for events of the personal past. *Journal of Experimental Psychology: General*, 123(2), pp. 201-215.

Everson, S. A., Maty, S. C., Lynch, J. W. & Kaplan, G. A., 2002. Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. *Journal of Psychosomatic Research,* Volume 53, pp. 891-895.

Gibbs, B. R. & Rude, S. S., 2004. Overgeneral Autobiographical Memory as Depression Vulnerability. *Cognitive Therapy and Research,* 28(4), pp. 511-526.

Hamad, R., Fernald, L. C. H., Karlan, D. S. & Zinman, J., 2008. Social and economic correlates of depressive symptoms and perceived stress in South African adults. *Journal of Epidemiology and Community Health*, 62(6), pp. 538-544.

Hausman, J. A., 1978. Specification tests in Econometrics. *Econometrica*, 46(6), pp. 1251-1271.

Irwin, M., Artin, K. H. & Oxman, M. N., 1999. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). *Archives of International Medicine,* 159(15), pp. 1701-1704.

Kaiser, C., 2020. Using memories to assess the intrapersonal comparability of wellbeing reports. *Working paper: EconStor,* pp. 1-56.

Kollamparambil, U., 2020. Happiness, Happiness Inequality and Income Dynamics in South Africa. *Journal of Happiness Studies,* Volume 21, pp. 201-222.

Laverde-Rojas, H., Correa, J. C., Jaffe, K. & Caicedo, M. I., 2019. Are average years of education losing predictive power for economic growth? An alternative measure through structural equations modeling. *PLoS ONE,* 14(3), pp. 1-21.

Lee, A. E. Y. & Chokkanathan, S., 2008. Factor Structure of the 10-item CES-D Scale Among Community Dwelling Older Adults in Singapore. *International Journal of Geriatric Pshyciatry,* 23(6), pp. 592-597.

Lee, C. S. & Hwang, Y. K., 2018. Structural Relationship between Depression and Happiness in Korean High School Students. *International Journal of Pure and Applied Mathematics,* 118(24), pp. 1-16.

Leight, K. A. & Ellis, H. C., 1981. Emotional Mood States, Strategies, and State-Dependency in Memory. *Journal of Verbal Learning and Verbal Behaviour,* Volume 20, pp. 251-266.

LePore, S. J., 1997. Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. *Journal of Personality and Social Psychology,* 73(5), pp. 1030-1037.

Lewbel, A., 2012. Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. *Journal of Business & Economic Statistics,* 30(1), pp. 67-80.

Lewbel, A., Dong, Y. & Yang, T. T., 2012. Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics,* 45(3), p. 809–829.

Li, B. et al., 2014. Positive psychological capital: A new approach to social support and subjective well-being. *Social Behavior and Personality: an international journal,* 42(1), pp. 135-144.

Lin, N. & Dean, A., 1984. Social Support and Depression: A Panel Study. *Social Psychiatry ,* Volume 19, pp. 83-91.

Lund, C. et al., 2013. Mental illness and lost income among adult South Africans. *Soc Psychiatry Psychiatr Epidemiol,* Volume 48, p. 845–851.

Macloed, A. K., Rose, G. & Williams, J. M., 1993. Components of hopelessness about the future in parasuicide. *Cognitive Therapy Research,* Volume 17, pp. 441-455.

Marsh, L., Edginton, T., Conway, M. A. & Loveday, C., 2019. Positivity bias in past and future episodic thinking: Relationship with anxiety, depression, and retrieval-induced forgetting. *Quarterly Journal of Experimental Psychology,* 72(3), pp. 508-522.

McDaid, D., 2007. The economics of mental health in the workplace: what do we know and where do we go?. *Epidemiologia e Psichiatria Sociale,* 16(4), pp. 294-298.

Miech, R. A. & Shanahan, M. J., 2000. Socioeconomic Status and Depression Over the Life Course. *Journal of Health and Social Behavior,* Volume 41, pp. 162-176.

Mirowsky, J. & Ross, C. E., 1992. Age and Depression. *Journal of Health and Social Behaviour,* Volume 33, pp. 187-205.

Mui, A. C., 1996. Depression among Elderly Chinese Immigrants: An Exploratory Study. *Social Work,* 41(6), pp. 633-645.

National Income Dynamics Study, 2017. *Adult (W15+) Questionnaire Wave 5 2017.* Cape Town: Southern Africa Labour and Development Research Unit.

Newby-Clark, I. R. & Ross, M., 2003. Conceiving the Past and Future. *Personality and Social Psychology Bulletin,* 29(7), pp. 807-818.

Nørby, S., 2015. Why forget? on the adaptive value of memory loss. *Perspectives on Psychological Science*, 10(5), pp. 551-578.

OECD, 2013. *OECD Guidelines on Measuring Subjective Well-being,* Paris: OECD Publishing.

Oxford Reference, 2020. *self-schema.* [Online]
Available at:
https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100453465
[Accessed 1 May 2020].

Oyenubi, A. & Kollamparambil, U., 2020. *COVID-19 and Depressive Symptoms in South Africa*, Cape Town: Southern Africa Labour and Development Research Unit - NIDS CRAM.

Paykel, E. S., 1994. Life events, Social support and depression. *Acta Psychiatrica Scandanavica,* 89(s377), pp. 50-58.

Pearlin, L. I. & Johnson, J. S., 1977. Marital Status, Life-Strains and Depression. *American Sociological Review,* Volume 42, pp. 704-715.

Peltzer, K. & Phaswana-Mafuya, N., 2013. Depression and associated factors in older adults in South Africa. *Global Health Action,* Volume 6, pp. 1-9.

Pokorski, M. & Siwiec, P., 2008. Depression and Memory: A Comparative Study of Young and Old Women. *Journal of Physiology and Pharmacology,* 59(6), pp. 573-578.

Popovski, M. & Bates, G. W., 2005. Autobiographical Memory and Dysphoria: The Effect of Mood, Gender, and Cue Type on Generality and Latency. *North American Journal of Psychology*, 7(3), pp. 505-518.

Prati, A. & Senik, C., 2020. Feeling good or feeling better?. *Working paper: RePec*, pp. 1-34.

Rashad, I. & Markowitz, S., 2007. Incentives in Obesity and Health Insurance. *Working paper: NBER No. W13113*, pp. 1-31.

Rassen, J. A. et al., 2009. Instrumental Variable Analysis for Estimation of Treatment Effects With Dichotomous Outcomes. *American Journal of Epidemiology,* 169(3), p. 273–284.

Reed, W., 2015. On the Practice of Lagging Variables to Avoid Simultaneity. *Oxford Bulletin of Economics and Statistics,* Volume 77, pp. 897-905.

Revenson, T. A., Schiaffino, K. M., Majerovitz, S. D. & Gibofsky, A., 1991. Social support as a double-edged sword: The relation of positive and problematic support to depression among rheumatoid arthritis patients. *Social Science and Medicine,* 33(7), pp. 807-813.

Roodman, D., 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal,* 9(1), pp. 86-136.

Roth, D. & Rehm, L. P., 1980. Relationships Among Self-Monitoring Processes, Memory, and Depression. *Cognitive Therapy and Research,* 4(2), pp. 149-157.

Scogin, F., Storandt, M. & Lott, L., 1985. Memory-Skills Training, Memory Complaints, and Depression in Older Adults. *Journal of Gerontology,* 40(5), p. 562–568.

Seo, E. H. et al., 2018. Life satisfaction and happiness associated with depressive symptoms among university students: a cross-sectional study in Korea. *Annals of General Psychiatry,* 17(1), pp. 1-9.

Sharot, T., 2011. *The Optimism Bias: A Tour of the Irrationally Positive Brain.* New York: Vintage Books.

Sharot, T., Riccardi, A. M., Raio, C. M. & Phelps, E. A., 2007. Neural mechanisms mediating optimism bias. *Nature,* 450(7166), pp. 102-105.

Siedlecki, K. L., Salthouse, T. A., Oishi, S. & Jeswani, S., 2014. The Relationship Between Social Support and Subjective Well-Being Across Age. *Social Indicators Research,* Volume 117, p. 561–576.

Söderlund, H. et al., 2014. Autobiographical Episodic Memory in Major Depressive Disorder. *Journal of Abnormal Psychology,* 123(1), pp. 51-60.

South African Depression and Anxiety Group, 2017. *Depression Brochures.* [Online] Available at: http://www.sadag.org/index.php?option=com_content&view=article&id=1877&Itemid=142 [Accessed 22 April 2020].

Southern Africa Labour and Development Research Unit, 2018. *National Income Dynamics Study (NIDS) Wave 1, 2008.* Cape Town: Southern Africa Labour and Development Research Unit.

Southern Africa Labour and Development Research Unit, 2018. *National Income Dynamics Study 2014-2015, Wave 4.* Cape Town: Southern Africa Labour and Development Research Unit.

Southern Africa Labour and Development Research Unit, 2018. *National Income Dynamics Study 2017, Wave 5.* Cape Town: Southern Africa Labour and Development Research Unit.

Southern Africa Labour and Development Research Unit, 2018. *National Income Dynamics Study Wave 2, 2010-2011.* Cape Town: Southern Africa Labour and Development Research Unit.

Southern Africa Labour and Development Research Unit, 2018. *National Income Dynamics Study Wave 3, 2012.* Cape Town: Southern Africa Labour and Development Research Unit.

Stats SA, 2019. *Quarterly Labour Force Survey* , Pretoria: Statistics South Africa.

Stats SA, 2020. *Mid-year population estimates,* Pretoria: Statistics South Africa.

Stice, E., Ragan, J. & Randall, P., 2004. Prospective Relations Between Social Support and Depression: Differential Direction of Effects for Parent and Peer Support?. *Journal of Abnormal Psychology,* 113(1), p. 155–159.

Strunk, D. R. & Adler, A. D., 2009. Cognitive biases in three prediction tasks: a test of the cognitive model of depression. *Behaviour Research and Therapy,* 47(1), pp. 34-40.

Suedfeld, P. & Eich, E., 1995. Autobiographical memory and affect under conditions of reduced environmental stimulation. *Journal of Environmental Psychology,* Volume 15, pp. 321-326.

Thompson, C. P., Skowronski, J. J., Larsen, S. & Betz, A., 1996. *Autobiographical memory: Remembering what and remembering when.* New York: Erlbaum.

Tomita, A. & Burns, J. K., 2013. A multilevel analysis of association between neighborhood social capital and depression: Evidence from the first South African National Income Dynamics Study. *Journal of Affective Disorders,* 144(1-2), pp. 101-105.

Tomita, A. & Burns, J. K., 2013. Depression, disability and functional status among community-dwelling older adults in South Africa: evidence from the first South African National Income Dynamics Study. *International Journal of Geriatric Psychiatry,* Volume 28, p. 1270–1279.

Tomlinson, M. et al., 2009. The epidemiology of major depression in South Africa: Results from the South African Stress and Health study. *South African Medical Journal,* 99(5), pp. 368-373.

Tversky, A. & Griffin, D., 1991. *Endowment and contrast in judgments of well-being.* Oxford: Pergamon Press.

Vella, F. & Verbeek, M., 1997. *Using rank order as an instrumental variable: an application to the return to schooling.* Leuven: Katholieke Universiteit.

Virtanen, M. et al., 2011. Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. *Psychological Medicine,* Volume 41, pp. 2485-2494.

Virtanen, M. et al., 2012. Overtime Work as a Predictor of Major Depressive Episode: A 5-Year Follow-Up of the Whitehall II Study. *PloS ONE,* 7(1), pp. 1-5.

Waldfogel, S., 1948. The frequency and affective character of childhood memories. *Psychological Monographs: General and Applied*, 62(4), pp. i-39.

Walker, W. R., Skowronski, J. J. & Thompson, C. P., 2003. Life is pleasant and memory helps to keep it that way. *Review of General Psychology,* Volume 7, pp. 203-210.

Wang, Y. & Bellemare, M. F., 2019. Lagged Variables as Instruments. *Working Paper,* pp. 1-38.

Watkins, E. & Teasdale, J. D., 2001. Rumination and overgeneral memory in depression: Effects of self-focus and analytical thinking. *Journal of Abnormal Psychology,* 110(2), pp. 353-357.

Weinstein, N. D., 1980. Unrealistic optimism about future life events. *Journal of Personality and Social Psychology,* Volume 39, pp. 806-820.

Wooldridge, J. M., 2010. *Econometric Analysis of Cross Section and Panel Data.* 2nd ed. Cambridge: MIT Press.

World Federation for Mental Health, 2012. *Depression: A Global Crisis,* Occoquan: World Federation for Mental Health.

World Health Organisation, 2020. *Depression.* [Online]
Available at: https://www.who.int/news-room/fact-sheets/detail/depression
[Accessed 19 April 2020].

Zuroff, D. C., Colussy, S. A. & Wielgus, M. S., 1983. Selective Memory and Depression: A Cautionary Note Concerning Response Bias. *Cognitive Therapy and Research,* 7(3), pp. 223-232.

# Annex A: Variables and descriptions

**Table 6: Variables and descriptions**

| Variable | Definition |
|---|---|
| Depression score | Described in main text. |
| Memory / Contrast | Described in main text. |
| Perceived health – at least good | Binary variable of perceived health takes on: excellent (1), very good (1), good (1), fair (0) and poor (0). |
| Religious intensity – at least important | Binary variable of religious intensity takes on: not important at all (0), unimportant (0), important (1), and very important (1). |
| Perceived income – below average | Binary variable for perceived income takes on: much above average income (0), above average income (0), average income (0), below average income (1), and much below average income (1) |
| Province | Binary variables for the province the individual resides in. |
| Household income (per capita) | The log of household income per capita; derived from the total household income and number of people in the household. |
| Age | The age of an individual – only including 15 years or older. |
| Age squared | The square of the age of the individual. |
| Female | Binary variable for female (1). |
| Black | Binary variable for race being black (1) . |
| Hours worked | The number of hours that an individual works per hour, which is generated from the number of hours worked in primary and secondary employment, including self-employment, agricultural work and helping others with business activities. Updates were made in those instances where self-employment and agricultural employment were included in primary and secondary employment hours. |
| Active lifestyle | An active lifestyle is defined as whether the individual works out three or more times a week. |
| Two-year financial well-being | Subjective view on financial well-being from two years from now, defined on a scale of 1 (lowest) to 6 (highest). |
| Employed | Binary variable for employment (1). |
| Level of violence | Overall level of violence in the area which an individual resides on a scale of 1 to 5. This is generated from the questions in Wave 2 to Wave 5 that relate to violence in the neighbourhood (6 questions). For Wave 1 this is based on the two questions available in the data. An average is generated from these questions and then rounded off to generate a violence score. Binary variables are generated for each category. |
| Married | Binary variable for whether the individual is married (1). |
| Number of years of schooling | The number of years of schooling an individual has completed (0-13 years) |
| Tertiary education | Binary variable for any level of tertiary education |
| Year | Binary variables for each wave of data collected. |

# Annex B: Theoretical motivation for the instruments

A common understanding amongst many is that a good instrument, to counteract endogeneity, is seldom easy to find (Wooldridge, 2010). The NIDS database does not provide a myriad of suitable instruments, however, there does exist one suitable instrument. The lag of the generated contrast variables are used as instruments. Using the lag of an endogenous variable as an instrument has been common in research, with the limitations related to the general requirements of IV regression analysis (Reed, 2015; extended in a working paper by Wang and Bellemare, 2019). Specifically, it is required that the instrument is both relevant and valid.

Instrumental relevance indicates that the instrument and endogenous regressor should be correlated. From a theoretical point of view, this is expected. Blaauw and Pretorius (2013) point out that the general relationship between age and subjective well-being is quadratic, with subjective well-being first decreasing with age, then increasing after the minimum value is reached. While Blaauw and Pretorius (2013) specifically discuss subjective well-being, the parallel to happiness can also be drawn.

Although the relationship is quadratic, this is still likely to evidence the relationship between contrast and its lagged value. For instance, assume that an individual is moving along the downward sloping region of the age and subjective well-being relationship. Given that subjective well-being is decreasing over time (with age), the lag of contrast should be correlated with the contrast. This is due to the change in happiness moving in the same direction over time (i.e. a negative first derivative). The same is true on the upward sloping part of the subjective well-being and age curve.

However, in the case where the individuals are at the turning point of the relationship between age and subjective well-being, there should be a point where contrast changes. While this may break the expected relationship between contrast and its lag, it is unlikely to do so for most individuals, especially considering the wide spectrum of ages that are included in the NIDS database. Notably, this relationship only holds under the strict assumption that all else is held constant.

In addition to instrumental relevance, instrumental validity is required. Instrumental validity means that while the instrument should be related to the endogenous regressor, the instrument should not have a causal relationship with the dependent variable (Wooldridge, 2010).

Therefore, the lags of contrast should not be a determinative factor for depressive symptoms; the underlying relationship should be contemporaneous, and not facilitated through time.

It is difficult to understand why the relationship should not be contemporaneous in nature. The relationship between contrast and depression has been explained above through the mediating factor of memory. Memory changes over time, with new memories being included as time passes; any current level of memory should consider all previous memories. Therefore, any relationship between contrast and depression, should be contemporaneous. Old levels of memories, and as such, contrast, should be overwritten by the current bank of memories.

While one could argue that memory may not be the same between two points in time due to forgetfulness (i.e. in addition to gaining new memories), this would also not facilitate the non-contemporaneous relationship. Simply, depression is unlikely be related to something that is no longer existent (i.e. since the change is due to forgetfulness).

# Annex C: Fixed effects and random effects results

**Table 7: Fixed effects and random effects results**

| | Fixed effects | Random effects |
|---|---|---|
| More happy | -0.996***<br>(0.0663) | -1.024***<br>(0.0517) |
| Less happy | 0.892***<br>(0.0832) | 1.138***<br>(0.0640) |
| Health perception – at least good | -1.021***<br>(0.0756) | -1.412***<br>(0.0581) |
| Religious intensity – at least important | -0.173*<br>(0.0946) | -0.205***<br>(0.0759) |
| Age | -0.291***<br>(0.0740) | 0.104***<br>(0.00515) |
| Age – squared | -0.000773***<br>(0.000212) | -0.000951***<br>(5.69e-05) |
| Female | | 0.232***<br>(0.0315) |
| Black | | 1.006***<br>(0.124) |
| Hours worked | 0.000504<br>(0.00176) | -0.00172<br>(0.00141) |
| Active lifestyle | -0.655***<br>(0.0813) | -0.678***<br>(0.0592) |
| Perceived income – below average | 0.0929<br>(0.0691) | 0.150***<br>(0.0570) |
| Financial well-being – 2-year expectation | -0.148***<br>(0.0321) | -0.163***<br>(0.0266) |
| Employed | -0.238***<br>(0.0773) | -0.180***<br>(0.0618) |
| Neighbourhood violence level | 0.109***<br>(0.0324) | 0.198***<br>(0.0270) |
| Married | -0.317***<br>(0.0807) | -0.523***<br>(0.0393) |
| Household income per capita – log | -0.0830**<br>(0.0399) | -0.0814***<br>(0.0256) |
| Years of education | 0.0127<br>(0.0239) | -0.0303***<br>(0.00651) |
| Tertiary education | 0.00290<br>(0.118) | -0.119*<br>(0.0613) |
| Constant | 20.31***<br>(2.388) | 6.421***<br>(0.301) |
| Hausman test results | Chi squared test statistic of 132.61, with a p-value of 0.000, therefore the null hypothesis of no systematic difference in coefficients is rejected. | |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Panel weights not used; random effects models do not allow for weights. Year/province dummy variables included. Each regression has 85,576 observations. Standard errors are clustered. When non-clustered standard errors are included, the following results are attained: Chi squared test statistic of 322.95, with a p-value of 0.000, therefore the null hypothesis of no systematic difference in coefficients is rejected.