The Journal of Development Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/fjds20

Mental and Physical Health Effect of Rural-Urban Migration in South Africa: A Quasi-Experimental Impact Evaluation Study

Bianca Capazario & Umakrishnan Kollamparambil

To cite this article: Bianca Capazario & Umakrishnan Kollamparambil (2022) Mental and Physical Health Effect of Rural-Urban Migration in South Africa: A Quasi-Experimental Impact Evaluation Study, The Journal of Development Studies, 58:9, 1732-1749, DOI: 10.1080/00220388.2022.2048654

To link to this article: https://doi.org/10.1080/00220388.2022.2048654

The Journal of Development Studies, 2022
Vol. 58, No. 9, 1732–1749, https://doi.org/10.1080/00220388.2022.2048654

Mental and Physical Health Effect of Rural-Urban Migration in South Africa: A Quasi-Experimental Impact Evaluation Study

BIANCA CAPAZARIO & UMAKRISHNAN KOLLAMPARAMBIL
School of Economics and Finance, University of Witwatersrand, Johannesburg, South Africa

(Original version submitted January 2021; final version accepted February 2022)

ABSTRACT *Using the National Income Dynamics Study (NIDS) longitudinal dataset, this study undertakes a difference-in-differences (DiD) evaluation of the impact of rural-urban migration on mental and physical health in South Africa. The contribution of the study is in considering sample selection bias as well as the causal direction of the relationship through the use of propensity score matching techniques and restricted sample DiD estimation. This study finds that the rural-urban migrants, within the South African NIDS sample, experience a decline in reported physical and mental health outcomes. The study identifies social isolation and difficult living conditions as some of the factors behind the adverse health outcomes. The findings underscore the fact that while favourable economic outcomes will likely occur as a result of migration efforts (such as employment opportunities and increased income), it comes at a cost of both physical and mental health.*

KEYWORDS: Impact evaluation; mental health; subjective health; rural-urban migration; South Africa

1. Introduction

The spatial, economic, and social organisation of South Africa has been moulded by the historical segregation of race and ethnic groups under the Apartheid government (Davies, 1981). These racist policies resulted in the deepening divide between the labour market, education and health-care systems in the urban and rural areas. As a result, many rural households were forced to find access to urban economies (Collinson, 2010). Often times, this is done without the ability to formally relocate, leading to oscillating trends of migration that characterise the South African context (Posel, 2004; Smith, 2003; Wilson, 2001).

While urbanisation may present positive effects of improved levels of education, health care access, and employment opportunities; rural-urban migrants still face the more challenging task of finding both adequate accommodation and absorption into the urban labour force. The migration process may bring about separation from family and the need to make lifestyle adjustments, which in-turn introduce new physical and mental health risks to the migrating population. Overcrowding and poor quality of housing has led to a rise in disease in South African informal settlements; this is further perpetuated by increased pollution levels, lack of running water supplies, and poor sanitation conditions (Gong, 2012; Smith, 2003).

Correspondence Address: Umakrishnan Kollamparambil School of Economics and Finance, University of Witwatersrand, Johannesburg 2050, South Africa. Email: uma.kollamparambil@wits.ac.za

Supplementary Materials are available for this article which can be accessed via the online version of this journal available at http://dx.doi.org/10.1080/00220388.2022.2048654.

These environmental risk factors are typically inversely related to positive health outcomes (Kristiansen, Mygrind, & Krasnik, 2007). Furthermore, rural-urban migrants of South Africa are often observed to retain ties with their *homelands*, and like most African countries, an increase in the mobility between sectors is strongly correlated with high and increasing rates of contracting communicable and non-communicable diseases (Posel, 2004; Vearey, Modisenyane, & Hunter-Adams, 2017).

On one hand, rural living conditions may lead to the poor underlying health levels of rural dwellers; in conjunction with a lack of health-care resources, this may prompt rural-urban migration. On the other hand, migration that is prompted by economic endeavours may lead to the exposure of several health diminishing factors (Vearey et al., 2017). Despite its necessity and relevance, South African literature is lacking in robust empirical analysis of the effects of rural-urban migration and its possible association with changes in health outcomes for the migrating population. The gaps in the existing South African body of knowledge typically extend toward the following; a) Studies do not include subjective measures specifically related to depression risk and physical health outcomes; b) Studies do not make use of a counterfactual, by way of constructing appropriately matched control and treatment cohorts, which is necessary to evaluate the impact of rural-urban migration on the reported health outcomes; or c) Studies fail to break the chain of simultaneity inherent to migration and health investigations.

This study will address these gaps in the South African body of knowledge, specifically through subjective measures of mental and physical health, using a longitudinal dataset that enables a difference-in-differences (DiD) impact evaluation framework. Further, the contribution of the study is in considering sample selection bias as well as the causal direction of the relationship through the use of propensity score matching techniques and restricted sample analysis.

2. Literature review

While rural-urban migration may foster favourable economic outcomes, such as new employment opportunities, higher income level attainment, and better access to education; the impact on health outcomes is more ambiguous. Positive physical health outcomes may manifest through the following mechanisms: (a) urban economies offer better access to healthcare services, which enables early diagnosis and treatment of illnesses; (b) favourable urban employment opportunities may translate into higher earnings and better allow migrants to afford a higher standard of living.

However, migration theories explain that positive self-selection (Healthy Migrant hypothesis) (Evans, 1987; Lu, 2008; Lu & Qin, 2014; Oyebode et al., 2015; Tong & Piotrowski, 2012) and tendency for unhealthy migrants to return to rural origins (Salmon Bias) (Lu, 2008) may present an upward bias in the positive health results seen in literature. To this point, it is important to address all sources of potential bias when estimating the impact of migration on health.

On the other hand, the reasons for a decline in migrant physical health outcomes are manifold; a) The adoption of a *Western lifestyle,* in light of attaining higher income levels. *Western diets* are associated with decreased consumption of less fruits and vegetables, and increased consumption of sugar, fat, and processed meat (Salant et al., 2003); (b) The negative effects associated with housing displacement, individuals who fail to attain adequate housing expose themselves to a range of environmental risk factors, namely; overcrowding, lack of water and sanitation, and other poor living conditions (Kristiansen et al., 2007); (c) Mobility increases risk of contracting communicable diseases (Posel, 2004; Vearey et al., 2017); (d) The adoption of unhealthy coping mechanisms, increased feelings of anonymity and detachment from surrounding society and the associated responsibilities thereof, may lead to the adoption risky behaviours such as the use of drugs, excessive alcohol consumption or risky sexual engagements (Kristiansen et al., 2007).

Also, migration is often categorised as a critical life event which can lead to increased levels of stress, which may manifest in not just physical, but also mental health problems (Berry, 1997; Elo et al., 2003). Newly settled migrants are susceptible to adjusting their risk perceptions. This is often compounded by overwhelming feelings of loss and dealing with the psycho-social issues (relating to unemployment or loneliness) which may further compromise the migrants' inclination to relate current risk behaviour to future health (Kristiansen et al., 2007). Acculturative stress also may compromise good mental health outcomes (Salant et al., 2003), however, this is not fully supported by literature. Acculturation may reduce mental distress over time, given that mental health may be compromised by the stress of adapting to unfamiliar societies. But in the same light, highly acculturated individuals may experience poor mental health effects due to prolonged periods of exposure to stressful conditions (Escobar et al., 2000; Fennelly, 2007; Koneru, De Mamani, Flynn, & Betancourt, 2007; Salant et al., 2003). On the other hand, McCullough, Kurzban, & Tabak (2013) points to the social revenge mechanism as a driver of migration which could result in improved mental health when the benefit from moving away from harmful relationships outweigh the benefits of the relationship.

Existing empirical literature from international and South African studies have assessed the interaction between migration and health through the lens of various health outcomes, namely; (a) subjective measures of physical health (Biao, 2007; Chen, 2011; Hu, Cook, & Salazar, 2008; Lu, 2010; Lu & Qin, 2014; Zhang, Liu, & Wu, 2015); (b) subjective wellbeing and mental health outcomes (Chen, 2011; Lu & Qin, 2014; Mulcahy & Kollamparambil, 2016; Wolf et al., 2017; Zhang et al., 2015) and (c) specific health indicators extending towards blood pressure, Body Mass Index (BMI), Human Immunodeficiency Virus (HIV), (Coffee, Lurie, & Garnette, 2007; Ebrahim et al., 2010; Ljungvall & Gerdtham, 2010; Lu, 2008, 2010; Salmond, et al., 1985). There is nevertheless a gap in literature, in terms of analysis using the impact evaluation framework, in the study of physical and mental health, with due consideration to various sources of estimation bias.

China has a comparable history to that of South African in terms of restricting the geographical mobility of its population. Findings in existing Chinese literature show that because sick migrants tend to return to traditional dwelling areas, rural-urban migrant studies often report a physical health advantage (Biao, 2007; Chen, 2011; Hu et al., 2008). Logit and ordinal logit regressions employed to assess the association between dichotomous reported health outcomes and rural-urban migration show a limited support of the healthy migrant paradox (Lu & Qin, 2014; Zhang et al., 2015). However, no attempts to break simultaneity are made, and association of health outcomes between urban residents and the migrant cohort is not significant for majority of the model specifications.

The panel data study by Mulcahy and Kollamparambil (2016) evaluates the impact of rural-urban migration on subjective well-being in South Africa. It is found that subjective wellbeing decreases after migration efforts as a result of false expectations, the emotional burden of moving away from family, and a reduction in social capital. The econometric specifications combat self-selection bias and issues of attrition. However, the study does not extend towards self-reported physical health outcomes nor does it evaluate specific mental health outcomes such as depressive symptoms.

3. Data and descriptive statistics

3.1. Data

This study employs data from the South African National Income Dynamics Study (NIDS), surveys conducted between the years 2008 and 2017. This is a panel dataset which tracks household and individual level data over five-waves, collected at approximately two-year intervals. Broadly, the analysis of this study is based on a five-period quasi-experimental data design. For purposes of the analysis, the treatment group comprises of rural-urban migrants who migrate

from rural to urban areas between waves two and three and thereafter do not change location. The rural dwellers who do not migrate form the control group. The post-treatment period comprises of waves three, four and five.

This paper makes use of subjective physical health and self-reported depressive symptoms as the outcome variables. The physical outcome measure is based on the survey question "*How would you describe your health at present?*" with response options provided as: *excellent, very good, good, fair* or *poor*. For purposes of this study this self-reported health status is converted into a dichotomous variable taking a value of one if the individual has expressed his/her health as being good, very good or excellent; zero if the individual has rated their health as being fair or poor. The depressive symptoms variable, used to measure mental health outcome, is comprised of the ten items on the Center for Epidemiological Studies Depression (CESD-10) scale. Each question could be responded to as "*not at all*", "*several days*", "*more than half the days*" or "*nearly every day*". The responses are coded from 0 to 3, creating the outcome variable of CESD-10 scale with a range of 0–30, with increasing values indicating higher risk of depression. Andresen, Malmgren, Carter, and Patrick (1994) recommend cut-offs that range from 8 to 10 for indications on risk of positive screening for depression. However, a number of studies that have used the CESD-10 provided in the NIDS data use a cut of 10 (Ardington & Case, 2010; Asante, Andoh-Arthur, Asante, & Andoh-Arthur, 2015; Kilburn, et al., 2018; Peltzer et al., 2013; Tomita & Burns, 2013). Furthermore, recommended cut-offs can vary by region and in a validation study done in South Africa, Bhana et al. (2019) recommend cut-offs of 11–13, depending on language. This analysis prefers the use of depressive score to keep the controversy regarding the appropriate threshold at bay. Nevertheless, as a robustness check, the study also presents the impact evaluation results using a cut-off of 10 and above.

The study is restricted to analysing rural-urban migration, which is defined in terms of geographical location; rural or urban. The rural category extends towards farmlands or traditional areas of dwelling. Individuals who have moved more than once during the five-waves of data collection are dropped from the sample so as to control for oscillating trends of migration (Collinson, Tollman, Kahn, & Clark, 2003). To avoid migration for purposes of accessing urban healthcare facilities due to old age, the sample is restricted to individuals within the age-group of 16–65 years. Also, by considering individuals who are of working age, one is able to select those who engage in migration activities of their own free-will (Kollamparambil, 2017). In this way, the sample reflects individuals that are more likely to have engaged in migration activities for economic reasons, rather than age related healthcare reasons. Various other control variables- guided by literature on subjective physical (Morudu & Kollamparambil, 2020; Staudinger, Fleeson, & Baltes, 1999; Callan, Hyunji, & William, 2015) and mental health (Burger, Posel, & von Fintel, 2017; Kollamparambil, 2021)- are harvested from NIDS (Appendix Table A1 for details).

3.2. Descriptive statistics

Table 1 provides detailed descriptions of the characteristics of the individuals who form part of both panel datasets. The table illustrates differences between rural dwellers (control group) and rural-urban migrants (treatment group). Given the differences in sample size based on data availability between physical and mental health, we present separate descriptive statistics for each.

Majority of the migrating population are between twenty and forty years of age for both samples. Majority of the treatment group have gone through high school and attained a matriculation certificate (67% and 68% respectively). Migrant populations tend to have better employment outcomes, compared to that of their rural dweller counterparts; with over 50% of migrants finding employment, while under 40% of rural dwellers are employed in both samples. Average household size is lower for migrants in both samples.

Table 1. Descriptive statistics

	Physical health sample			Depression risk sample		
	Rural dwellers	Rural-Urban migrants	*t*-test	Rural dwellers	Rural-Urban migrants	*t*-test
Age category						
18–19	10.3% (559)	16% (36)	**	11% (380)	16.6% (20)	**
20–40	35.4% (1910)	60% (208)	***	37% (1263)	60% (72)	***
40–60	42.8% (2315)	22% (63)	***	41% (1412)	27.6% (26)	***
60+	11.4% (616)	2% (6)	***	11% (361)	2% (2)	***
Race						
African	95.2% (5142)	92.6% (287)	–	96% (3282)	93% (112)	–
Coloured	2.6% (143)	6.4% (19)	–	2.4% (81)	6% (7)	–
Indian	1.3% (72)	1% (7)	–	1% (35)	1% (1)	–
White	0.8% (43)	0% (0)	–	0.5% (18)	0% (0)	–
Gender						
Female	69% (3722)	71% (199)	–	68% (2324)	72% (86)	–
Education						
None	17% (925)	6.4% (13)	***	14.3% (490)	5.8% (7)	***
Primary School	28% (1493)	7.3% (15)	***	27% (924)	7.5% (9)	***
High School	49% (2600)	16% (33)	***	50.6% (1728)	13.3% (16)	***
Matric	6% (291)	68% (138)	***	6% (202)	68.4% (82)	***
Bachelors/Diploma	1% (58)	1% (2)	–	1.4% (49)	1.7% (2)	–
Postgraduate	0.4% (24)	1.5% (3)	–	0.6% (19)	3.3% (4)	–
Other	0.1% (7)	0% (0)	–	0.1% (4)	0% (0)	–
Economic activity						
Employed	36% (1949)	51% (105)	***	37% (1254)	53% (64)	***
Household						
Average household size	(6.26)	(4.6)	***	(6.20)	(4.3)	***
Sample size (N)	5400	204		3416	120	

Source: NIDS Dataset.

A higher proportion of the migrant treatment group rated their health as good prior to migration, compared to rural dwellers (Figure 1). However, this changes dramatically post migration from wave 3. A lower proportion of the migrant group reported good-health post migration. This proportion also seems to be declining for the migrant group in the post migration waves.

The average depression risk scores (Figure 2) illustrate significant increase among the treatment group of migrants post migration from wave 3. In the waves prior to migration (wave two) the rural-urban migrant cohort exhibits a lower average depressive symptom scores compared to their rural dweller counterpart. After migration, the rural-urban migrant cohort experienced a sharp increase in the average depressive symptoms scores and, the migrant group records a higher average score compared to their rural dweller counterpart.

4. Methodology

A quasi-experimental design based DiD model is used to explore the effect of rural-urban migration on the health outcomes of migrants. The double differencing in DiD allows to eliminate the individual-specific bias, as well as time-trend bias (Angrist & Pischke, 2008). However, failing to select appropriate control groups or not creating an appropriate counterfactual may yield biased results in the presence of self-selection (Wapenaar & Kollamparambil, 2019). This necessitates the employment of Propensity Score Matching (PSM) to account for observed confounding factors. When this is done adequately one may infer that, based on observable

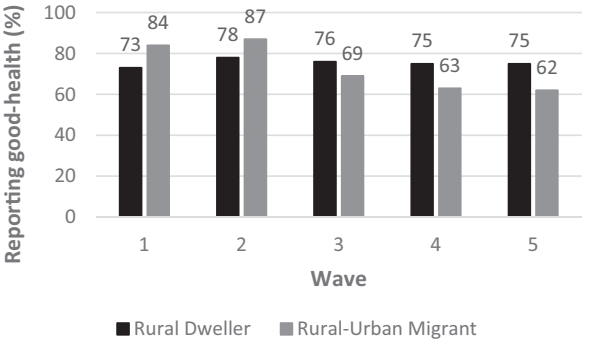

Figure 1. Self-reported good health.

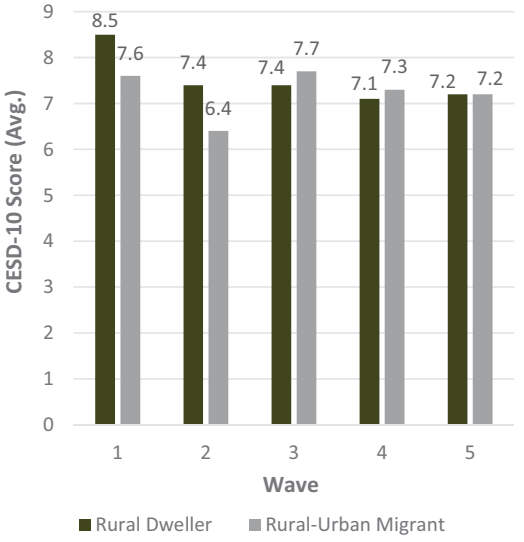

Figure 2. Average CESD-10 score.

characteristics, individuals who are similar enough, but differ in terms of treatment alone (migration), may experience a difference in health status, which may then be attributed to the effects of the intervention. Lastly, restricted sample estimation is undertaken to control for possible bias arising through simultaneity-led endogeneity. The baseline DiD model is written as:

$$Y_{it} = \alpha_0 + \beta_i X_{it} + d_1 Treat_i + d_2 Post_t + d_3 (Treat)_i (Post)_t + \varepsilon_{it} \quad (1)$$

Where;

1. Y_{it} is indicative of the health outcome of interest
2. α_0 is the constant term,
3. $\beta_i X_{it}$ is the vector of controls included as per respective literature on the predictors of subjective physical health and mental health.
4. $d_1 Treat_i$ is a dummy equal to 1 for the treatment group of migrants across all waves
5. $d_2 Post_t$ is a dummy equal to 1 for the time periods after treatment has occurred across treatment groups
6. d_3 is the coefficient of interaction term indicating impact or a difference-in-differences
7. ε_{it} represents the error term

The control variables included in the subjective physical health model is guided by studies like Staudinger et al. (1999), Callan et al. (2015) and, Morudu and Kollamparambil (2020). These include demographic controls (age, gender, race, marital status), individual self-regulatory characteristics (alcohol use, smoking), socio-economic status (per capita household income, employment status, educational attainment, private medical insurance) and environmental factors (dwelling condition). In addition to these, studies like Kollamparambil (2021) and Burger et al. (2017) have shown that other factors like religiosity, safety perception, future income expectations are key drivers of mental health in South Africa. As such, the mental health model is augmented by these variables. The detailed variable definitions are provided in Appendix Table A1.

While linear DiD estimations are selected for the physical and mental health models, given the binary and continuous nature of the respective outcome variables, as a robustness test, non-linear DiD result is also presented for mental health outcome variable based on binary form using a cut off of CESD-10 ≥ 11.

The DiD methodology rests on two vital assumptions. The first is the "parallel-lines" trends assumption which is tested in accordance with the Autor (2003) method:

$$Y_{it} = \mu_0 + \lambda_0 time + \gamma X_{it} + Treat_{it}\mu + Treat{*}time.\lambda + \eta_{it} \tag{2}$$

Where,

time is the time trend ranging from years 1 to 5

$Treat_{it}$ is a dummy equal to 1 for the treatment group of migrants across all waves

Equation 2 includes the interaction between treatment variable and the dummy variable for each year. Accepting the null H0: $\lambda = 0$ (for the pre-treatment periods, 1 and 2 waves) implies accepting that the parallel-trend assumption is not violated whenever one assumes no "anticipation effects". For estimation purposes, the benchmark is taken as the interaction variable (time2* treatment) immediately prior to migration. If the coefficient of the interaction terms (time1*treatment) prior to the treatment is statistically equal to zero one can reasonably expect the parallel trend to hold (Angrist & Pischke, 2008). The results in Supplementary Tables S1 and S2 indicate that this is the case.

The second assumption required for DiD is that there is random assignment into treatment and control groups. This brings about the need to control for all observed confounding factors that may contribute towards self-selection into treatment. Towards this end, we use Propensity Score Matching (PSM) based DiD model. Matching extensions allow for the control of observable confounding factors, therefore eliminating potential bias introduced through self-selection (Gertler, 2016).

Model I presents the base DiD estimation, and Model II is the clustered standard errors estimation of Model I. A fixed effects estimation (Model III) is also presented that accounts for unobserved individual heterogeneity, as a comparison to Model I.

Model IV extends the DiD by way of introducing Kernel based Propensity Score Matching (PSM) which acts to find best matches between control and treatment groups based on observed characteristics. This will allow comparisons between likened individuals to be made. PSM rests on the assumption of significant overlap (Angrist & Pischke, 2008). Motivation for meeting this requirement is presented in Supplementary Table S3 and Figures S1–S2.

Standard DiD also suffers from the potential for self-selection which leads to imbalances in characteristics between the treatment and control groups (Rosenbaum & Rubin, 1983). This necessitates measures be taken to account for potential structural differences between the two cohorts. The propensity score matching based DiD is to overcome the shortfalls of the standard DiD. Matching establishes appropriate counterfactuals for the treated on the basis of pre-treatment observable characteristics. To overcome the 'curse of dimensionality' induced by the use of a range of explanatory variables, a balancing score is created for each individual which reduces the vector of covariates to a single value that measures the likelihood of treatment

status (Wapenaar & Kollamparambil, 2019). Balancing tests were conducted following estimation of the propensity score (Supplementary Table S4). The matching process led to a fall in mean bias due to covariate imbalances from 20.4 per cent to 6.6 per cent. The DiD equation 1 is re-estimated with the weights derived from the PSM matching in Model IV.

The PSM matched model (Model IV) is an improvement from Model I, but nonetheless suffers from possible bias arising from simultaneity. In order to account for this, Model V restricts the pre-migration sample to healthy individuals (good reported physical health and CESD-10 score less than 13). This reduces the possibility that migration was undertaken for purpose of accessing healthcare facilities. The PSM-DiD of this restricted sample is presented as Model V.

Lastly, a test of attrition using logit models (Fitzgerald, Gottschalk, & Moffitt 1998) also shows that there is no relationship of subjective physical health, mental health and treatment with the probability of not being interviewed in the subsequent NIDS waves (Supplementary Table S5).

5. Results

The estimation results for the reported physical health and the depressive symptoms models are presented in Tables 2 and 3 respectively. The migration effect, after controlling for various economic, behavioural, environmental, and demographic factors, is indicated by the interaction between the "*Post*" and "*Treat*" variables, the "*Post***Treat*" variable. *Model I* is a baseline DiD estimation, *Model II* is the clustered standard errors estimation, *Model III* is the fixed effects estimation accounting for unobserved individual heterogeneity, *Model IV* is the PSM-DiD estimation and *Model V* is the restricted sample regressions for reported physical health and risk of depression outcomes, after controlling for possible selection and endogeneity bias.

5.1. Subjective physical health

The fixed effects as well as the quasi-experimental results illustrate a negative effect of migration on reported health outcomes across all models (Table 2). The significance of the coefficient of the *post*treat* variable increases in models that account for selection and endogeneity bias. Clustered standard errors do not change the significance of coefficients between Models I and II. The estimated reduction in probability of self-reported good health after migration is between 5%-9% in the models (Table 2). Accounting for unobserved individual heterogeneity through the fixed effects model (Model III), the sectional bias (Model IV) and finally both selection as well as endogeneity bias (Model V) yields significant negative effect of migration on self-reported physical health. While model IV yields the highest impact with 9% reduction in probability of self-reported good health after migration, adjusting the model to incorporate endogeneity bias in Model V, brings down the impact of migration to 5% reduction in probability of self-reported good health.

The negative association between physical health status and rural-urban migration may be explained through various mechanisms. The persisting challenges in accessing quality health care is often used to explain this negative relationship between health and rural-urban migration (Baron-Epel & Kaplan, 2009; Dias, Gama, & Martins, 2013; Dias, Gama, & Rocha, 2010), however, this is controlled for using the *Medical Aid* variable and matching of the treatment and counterfactual groups through PSM. Therefore, the mechanisms more appropriate in explaining this negative relationship are; (a) The negative externalities stemming from habitation challenges faced by the migrants (Gong, 2012; Smith, 2003), and (b) The adoption of poor lifestyle choices or risky behaviours (Kristiansen et al., 2007).

Rapid urbanisation could result in poor living conditions for the treatment cohort. In this sample, close to 10% of the migrating population is shown to move from living in a formal dwelling space to an informal dwelling space after rural-urban migration. Put differently, 10%

Table 2. Subjective physical health[@] DiD results

Variables	(I) DiD	(II) DiD#	(III) Fixed effects○#	(IV) DiD& PSM	(V) DiD& PSM (Restricted Pre-migration sample)
Migration effects					
Post	–0.0654***	–0.0654***	–0.0386*	0.00758	–0.018
	(0.0208)	(0.0194)	(0.0208)	(0.0195)	(0.015)
Treat~	0.0879**	0.0879***		0.0627***	0.029***
	(0.0396)	(0.0212)		(0.0118)	(0.012)
***Post*treat**	**–0.0599**	**–0.0599**	**–0.0680***	**–0.0828*****	**–0.0465*****
	(0.0533)	**(0.0373)**	**(0.0389)**	**(0.0173)**	**(0.017)**
Economic factors					
Ln(rhipc)	–0.00431	–0.00431	–0.00127	0.00591	0.010**
	(0.00601)	(0.00604)	(0.00902)	(0.00503)	(0.0105)
Unemployment	–0.0271**	–0.0271**	0.0408***	–0.0502***	–0.039***
	(0.0109)	(0.0114)	(0.0146)	0.00591	0.010**
Medical Aid	0.0125	0.0125	0.0116	–0.00700	–0.005
	(0.0103)	(0.0122)	(0.0169)	(0.00714)	(0.007)
Behavioural factors					
No Alcohol consumption	0.0242	0.0242*	–0.0482***	0.0241*	0.006
	(0.0150)	(0.0138)	(0.0164)	(0.0130)	(0.0278)
Regular smoker	–0.0952***	–0.0952**	–0.0416	–0.0245	–0.004
	(0.0325)	(0.0376)	(0.0342)	(0.0298)	(0.0765)
Environmental factors					
Formal dwelling	0.0406***	0.0406***	0.0119	0.0342***	0.029***
	(0.0110)	(0.0116)	(0.0155)	(0.00940)	(0.009)
Mother Education	0.0222***	0.0222***	0.00558	0.00507*	0.002
	(0.00294)	(0.00321)	(0.00474)	(0.00263)	(0.003)
Demographic factors					
Age	–0.00343***	–0.00343***	–0.00104	–0.00254***	–0.002***
	(0.000438)	(0.000475)	(0.000673)	(0.000431)	(0.000)
Education	0.0475***	0.0475***	0.0100	0.0465***	0.048***
	(0.0123)	(0.0132)	(0.0192)	(0.0119)	(0.012)
Female	–0.0295***	–0.0295***		–0.0414***	–0.045***
	(0.0111)	(0.0110)		(0.00975)	(0.009)
Married	0.0247**	0.0247*	0.0342	–0.0330***	0.033***
	(0.0113)	(0.0135)	(0.0297)	(0.0119)	(0.012)

(*continued*)

Table 2. (*Continued*)

Variables	(I) DiD	(II) DiD#	(III) Fixed effects○#	(IV) DiD& PSM	(V) DiD& PSM (Restricted Pre-migration sample)
African	0.0137	0.0137		0.153***	0.157***
	(0.0240)	(0.0275)		(0.0191)	(0.018)
Year	0.0128***	0.0128***	0.0167***	–0.00296	0.003
	(0.00363)	(0.00330)	(0.00378)	(0.00310)	(0.003)
Constant	–24.74***	–24.74***	–32.80***	4.942	–6.489
	(7.291)	(6.634)	(7.585)	(6.290)	(6.059)
Observations	5,604	5,604	5,604	4,033	2,389
R-squared	0.067	0.067	0.014	0.092	0.110
Number of pid			1,979		

Standard errors in parentheses, #Clustered standard errors, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

@The binary dependent variable: taking a value of 1 when physical health is excellent, good, and fair, and a value of 0 when physical health is poor.

○ The fixed effects estimation accounts for unobserved time invariant individual heterogeneity. However, the model excludes the coefficients of the treatment variable as well as time invariant variables like Female and African.

~The "Treat" variable is reflective of the group who engages in migration efforts.

of the treatment group who were living in a formal dwelling space prior to migration now find themselves in an informal dwelling space post migration. "Formal dwelling" is shown to be significantly and positively related to good physical health outcomes. This presents empirical evidence to support mechanism (a) outlined above. In addition, risky behavioural factors are seen to have a significant impact on health outcomes. Alcohol consumption and smoking habits have mostly adverse effects on reported physical health outcomes.

Unemployment is shown to have significant and negative impact on health across all models in Table 2. There is empirical support for unemployment and lower earnings acting as a driving force behind obesity (Ebrahim et al., 2010; Henry & Kollamparambil, 2017; Ljungvall & Gerdtham, 2010); and lack of nutrition due to poorer eating habits.

Income is shown to be positive and significant in Model III. Health is also seen to decrease with age. Female individuals are more susceptible to poor health outcomes, in the final chosen model.

5.2. *Risk of depression*

The relationship between rural-urban migration and the risk of experiencing symptoms of depression, as denoted by the *post*treat* variable, remains consistently positive and significant throughout all models (Table 3). The estimated increase in the average depressive score after migration is between 0.9 and 1.6 in models I-V (Table 3). While the base models (I -III) have a marginal effect greater than one, after accounting for possible selection and endogeneity bias, the effect of migration on depression risk is shown to decrease to 0.9 in model V; implying that the alternative model specifications overstate the effects of migration on the manifestation of depressive symptoms. Nevertheless, the results are statistically significant at 95% confidence level, providing strong evidence of a psychological cost associated with rural-urban migration.

The results are in line with existing literature which suggests that the mental well-being of migrants may suffer as a result on migration efforts (Chen, 2011; Lu, 2010; Mulcahy & Kollamparambil, 2016; Zhang et al., 2015). Related studies presume that the decreased mental health among migrants is due to a reduction in social support and family separation (Lu, 2010; Mulcahy & Kollamparambil, 2016). Evidence of reduced social support is found in comparing the average household size of the treatment group in the period prior and post migration efforts (Table 1).

As a further robustness test, the depression risk variable is considered in binary form, taking value 1 if the CESD-10 score is 10 or great, and zero otherwise. A non-linear DiD model based on the binary mental health outcome is presented in Supplementary Table S6. The results are consisent with the linear DiD model (Table 3).

5.3. *Robustness checks*

As additional checks, DiD estimations were undertaken including; a) circular migrants in the sample and, b) without sample restriction to individuals between 18 and 65 years of age. The inclusion of circular migrants yielded 44 new observations in the treatment group, whereas removing age restrictions yielded 1518 total observations (control group and treatment group), of which 6 observations belonged to the treatment group.

The subjective physical health and mental health estimations for these two samples as well as their descriptive statistics are included in the supplementary file (Supplementary Tables S7–S10). The results indicate that the increase in depression score amongst migrants is robust to both the samples. The subjective physical health declined for the treatment group in both estimations, however the coefficients are not significant in both. This could be indicative of the salmon bias pattern, highlighted by Lu (2008) whereby the migrants return to their rural origins when their health deteriorates, which often also closely correlate with old age. It is therefore

Table 3. Depression score[@] DiD results

Variables	(I) DiD	(II) DiD#	(III) Fixed effects○#	(IV) DiD& PSM	(V) DiD& PSM (Restricted Pre-migration sample)
Migration effects					
Post	0.250	0.250	–0.0650	–1.252**	–0.472*
	(0.291)	(0.285)	(0.0286)	(0.494)	(0.275)
Treat~	–1.054	–1.054**		–0.675**	0.397
	(0.697)	(0.535)		(0.278)	(0.951)
***Post*treat**	**1.569***	**1.569***	**1.270****	**1.101****	**0.869****
	(0.841)	**(0.839)**	**(0.0762)**	**(0.432)**	**(0.464)**
Economic factors					
Ln(rhipc)	–0.322***	–0.322***	–0.0886	0.317**	–0.084
	(0.0924)	(0.0920)	(0.0247)	(0.133)	(0.437)
Unemployed	0.118	0.118	0.168	0.370	0.132
	(0.157)	(0.156)	(0.0450)	(0.245)	(0.247)
Medical Aid	0.337	0.337	–0.622*	0.519	0.498
	(0.335)	(0.330)	(0.0844)	(0.724)	(0.657)
Expected income 5 years	1.60e–08	1.60e–08	2.05e–10	6.45e–08***	4.14e–08***
	(3.46e–08)	(1.11e–08)	(1.62e–08)	(1.86e–08)	(1.13e–07)
Behavioural factors					
No Alcohol consumption	0.170	0.170	0.420**	1.218***	1.174***
	(0.212)	(0.210)	(0.0187)	(0.368)	(0.361)
Regular smoker	–0.273	–0.273	0.0992	–3.303***	–3.496***
	(0.506)	(0.460)	(0.0269)	(0.809)	(0.720)
Religion	0.279	0.279	0.393	1.461***	1.576***
	(0.265)	(0.276)	(0.191)	(0.377)	(0.374)
Environmental factors					
Formal dwelling	–0.356**	–0.356**	–0.381	–0.504**	–0.447**
	(0.160)	(0.166)	(0.0919)	(0.221)	(0.218)
Safety	0.123	0.123	0.0201	0.288	0.230
	(0.171)	(0.163)	(0.0237)	(0.255)	(0.252)
Mother Education	–0.121***	–0.121***	0.0201	–0.0741	–0.700***
	(0.0423)	(0.0443)	(0.0372)	(0.0648)	(0.181)
Household Size	–0.0213	–0.0213	–0.0337	0.00142	–0.014
	(0.0213)	(0.0239)	(0.0218)	(0.0338)	(0.036)

(*continued*)

Table 3. (*Continued*)

Variables	(I) DiD	(II) DiD#	(III) Fixed effects○#	(IV) DiD& PSM	(V) DiD& PSM (Restricted Pre-migration sample)
Demographic factors					
African	1.679***	1.679***		0.719	0.968**
	(0.372)	(0.392)		(0.517)	(0.495)
Age	0.0198***	0.0198***	0.0139	0.0600***	0.035***
	(0.00619)	(0.00652)	(0.00280)	(0.0108)	(0.011)
Female	0.361**	0.361**		–0.804***	–0.863***
	(0.158)	(0.156)		(0.257)	(0.257)
Married	–0.612***	–0.612***	–0.654	–0.369	0.106
	(0.163)	(0.169)	(0.171)	(0.288)	(0.290)
Education	–0.466***	–0.466***	0.0969	–1.405***	–1.793
	(0.175)	(0.179)	(0.0555)	(0.312)	(0.317)
Year	–0.206***	–0.206***	–0.194**	–0.184**	–0.252***
	(0.0519)	(0.0511)	(0.00875)	(0.0821)	(0.080)
Constant	421.5***	421.5***	399.0**	370.7**	98.14
	(104.1)	(102.6)	(17.49)	(163.3)	(174.0)
Observations	3,536	3,536	3,536	1,901	1190
R-squared	0.064	0.064	0.028	0.088	0.138
Number of pid			1,664		

Standard errors in parentheses, #Clustered standard errors, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

[@]Continuous dependent variable ranging from 0 to 30 is a measure of the depressive symptoms based on the CESD-10 criterion (see Appendix Table A1).

°The fixed effects estimation accounts for unobserved time invariant individual heterogeneity. However, the model excludes the coefficients of the treatment variable as well as time invariant variables like Female and African.

~The "Treat" variable is reflective of the group who engages in migration efforts.

not surprising that the statistically significant decline in physical health observed in the restricted sample is not visible in the samples where older individuals and circular migrants are included.

6. Limitations

The baseline DiD estimation does not employ sample weighting and as such the study cannot claim to be representative of the South African population as a whole. Nevertheless, the findings can be considered to be robust to the sample of study. Further, limitation is that this study is not able to control for HIV status and adherence to anti-retroviral treatment due to a lack of data availability. Given that HIV spread is accelerated by frequent migration between areas of different HIV prevalence's (Coffee et al., 2007), this is a limitation left to be addressed. It is also important to acknowledge that endogeneity may still be introduced into the DiD estimator through time variant unobservable variables, correlation between covariates, and/or omitted variables.

The use of subjective measure may also be regarded as a limitation in the sense that various external factors may influence results that are reported. These factors may subsequently lead to an under or over-stating of an individual's true health status (Ray, 1998). However, various empirical studies suggest that over-all health outcomes may be accurately predicted by self-reported health. In this way, subjective health reports are typically consistent with objective mental and physical health reports (Biddle, Kennedy, & McDonald, 2007; Chiswick & Miller, 2008; Ider & Benyamini, 1997).

Lastly, the study acknowledges that endogeneity bias emanating from unobservable time varying heterogeneity remains in the estimations and as such the causal interpretation of the results presented is limited by it.

7. Discussion

This study finds that the rural-urban migrants, within the South African NIDS sample, experience a decline in reported physical and mental health outcomes. Various mechanisms can explain the underlying reasons for this resultant deterioration in mental and physical health outcomes.

The stress of adapting to a new environment may place a significant strain on the mental wellbeing of those that engage in rural-urban migration. Likewise, the subsequent unhealthy living environment, adoption of unhealthy eating or exercising habits is likely brought on despite realising higher disposable incomes levels in urban areas (Kristiansen et al., 2007).

Evidence for higher prevalence's in depression symptoms experienced by rural-urban migrants in this study has been explained in literature by the experience of social isolation by new migrants. This is potentially brought about by feelings of loneliness and separation from family (Bhugra, 2003; Chen, 2011; Mulcahy & Kollamparambil, 2016; Qiu, et al., 2011; Zhang et al., 2015). Evidence for possible isolation and separation stress can be inferred by the decline the average household size for the treatment cohort observed in this study.

Rural-urban migrants are also found in literature to engage in risky sexual behaviour, and in turn is associated with the spread of HIV and other sexually transmitted diseases (Coffee, Lurie, & Garnette, 2007) that leads to adverse health outcomes. Rural-urban migration is also associated with increases in obesity levels, which in-turn drive other health risk factor changes (Ebrahim et al., 2010; Ljungvall & Gerdtham, 2010).

Given the analysis performed in this study, it is clear that rural-urban migration has non-negligible, and significantly adverse effects on both the mental and physical health outcomes for the migrating population. While favourable economic outcomes will likely occur as a result of

migration efforts (such as employment opportunities and increased income), it comes at a cost of both physical and mental health.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

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 Preventive Medicine*, *10*(2), 77–84.

Angrist, J. D. & Pischke, J. S., 2008. *Mostly harmless econometrics: An Empiricist's companion*. Princeton: Princeton University Press.

Ardington, C. & Case, A., 2010. Interactions between mental health and socioeconomic status in the South African national income dynamics study. *Studies in Economics and Econometrics*, *34*(3), 69–85.

Asante, K. O., Andoh-Arthur, J., Asante, K.O. & Andoh-Arthur, J. 2015. Prevalence and determinants of depressive symptoms among university students in Ghana. *Journal of affective disorders*, *171*, 161–166. doi:10.1016/j.jad.2014.09.025

Autor, D. H., 2003. Outsourceing at will: The contribution of unjust dismissal doctrine of the growth of employment outsourcing. *Journal of Labour Economics*, *21*(1), 1–42. doi:10.1086/344122

Baron-Epel, O. & Kaplan, G., 2009. Can subjective and objective socioeconomic status explain minority health disparities in Israel? *Social Science Medicine*, *69*(10), 1460–1467. doi:10.1016/j.socscimed.2009.08.028

Berry, J. W., 1997. Immigration, acculturation, and adaption. *Applied psychology*, *46*(1), 5–34.

Bhana, A., Mntambo, N., Gigaba, S.G., Luvuno, Z.P.B., Grant, M., Ackerman, D., Ntswe, E., Nomathemba, M. & Petersen, I., 2019. Validation of a brief mental health screening tool for common mental disorders in primary healthcare. *SAMJ: South African Medical Journal*, *109*(4), 278–283. doi:10.7196/SAMJ.2019.v109i4.13664

Bhugra, D., 2003. Migration and depression. *Acta Psyhiatrica Scandinavica*, *2003*(108), 67–72. doi:10.1034/j.1600-0447.108.s418.14.x

Biao, X., 2007. How far are the left-behind left behind? A preliminary study in rural China. *Population, Space and Place*, *13*(3), 179–191. doi:10.1002/psp.437

Biddle, N. K., Kennedy, S. & McDonald, J., 2007. Health assimilation patterns among Australian immigrants. *The Economic Records*, *83*(260), 16–30. doi:10.1111/j.1475-4932.2007.00373.x

Burger R., Posel D., & von Fintel M. 2017. The relationship between negative household events and depressive symptoms: Evidence from South African longitudinal data. *Journal of Affective Disorders. 218*: 170–175. doi:10.1016/j.jad.2017.04.031

Callan M., Hyunji, K., & William, M. 2015. Predicting self-rated mental and physical health: The contributions of subjective socioeconomic status and personal relative deprivation. *Frontiers in Psychology*, *2015*, 1415–1415. doi:10.3389/fpsyg.2015.01415

Chen, J., 2011. Internal Migration and health: Re-examining the healthy migrant phenomenon in China. *Social Science & Medicine*, *72*(8), 1294–1301. doi:10.1016/j.socscimed.2011.02.016

Chiswick, B. Y. & Miller, P., 2008. Immigration selection system and immigration health. *Contemporary Economic Policy*, *26*(4), 555–578. doi:10.1111/j.1465-7287.2008.00099.x

Coffee, M., Lurie, M. N. & Garnette, G. P., 2007. Modelling the Impact of Migration on the HIV Epidemic in South Africa. *AIDS*, *21*(3), 343–350. doi:10.1097/QAD.0b013e328011dac9

Collinson, M., Tollman, S., Kahn, K. & Clark, S., 2003. Highly prevalent circular migration: Households, mobility and economic status in rural South Africa. In *Africa on the move: African migration and urbanisation in comparative perspective* (pp. 194–216). Johannesburg: Agincourt Health and Population Unit: University of the Witwatersrand.

Collinson, M. A., 2010. Striving against adversity: The dynamics of migration, health and poverty in rural South Africa. *Global Health Action. 2010*, 80.

Davies, R. J., 1981. The spatial formation of the South African City. *GeoJournal*, *2*(2), 59–72. doi:10.1007/BF00196325

Dias, S., Gama, A. & Martins, M. O., 2013. Health status and preventative behaviors of immigrants by gender and origin: A Portuguese cross-sectional study. *Nursing & Health sciences*, *15*(3), 309–317. doi:10.1111/nhs.12032

Dias, S., Gama, A. & Rocha, C., 2010. Immigrant women's perceptions and experiences of health care services: Insights from a focus group study. *Journal of Public Health*, *18*(5), 489–496. doi:10.1007/s10389-010-0326-x

Ebrahim, S., Kinra, S., Bowen, L., Andersen, E., Ben-Shlomo, Y., Lyngdoh, T., Ramakrishnan, L., Ahuja, R. C., Joshi, P., Das, S. M. & Mohan, M., 2010. The effect of rural-to-urban migration on obesity and diabetes in India: A cross-sectional study. *PLoS One*, *7*(4), 1–12. doi:10.1371/journal.pmed.1000268

Elo, A. L., Leppanen, A. & Jahkola, A., 2003. Validity of a single-item measure of stress symptoms. *Scandinavian Journal of Work, Environment & Health*, *29*(6), 444–451. doi:10.5271/sjweh.752

Escobar, J. I., Nervi, C. H., Gara, M. A., Escobar, J.I., Nervi, C.H. & Gara, M.A. 2000.. Immigration and mental health: Mexican Americans in the United States. *Harvard review of psychiatry*, *8*(2), 64–72. doi:10.1080/hrp_8.2.64

Evans, J., 1987. Introduction: Migration and Health. *The International Migration Review*, *21*(3), 5–14. doi:10.2307/2546605

Fennelly, K., 2007. Health and well-being of immigrants. In E. Walker & E. Barnett, eds. *Immigrant Medicine* (pp. 19–26). Philadelphia: Elsevier.

Fitzgerald, J., Gottschalk, P. & Moffitt, R.A., 1998. An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics.

Gertler, P. J., 2016. *Impact evaluation in practice*. Washington, D.C.:The World Bank.

Gong, P. 2012. Urbanisation and Health in China. *Lancet*.

Henry, J. & Kollamparambil U. (2017). Obesity-based labour market discrimination in South Africa: A dynamic panel analysis, *Journal of Public Health*, *25(6)*, 671–684. doi:10.1007/s10389-017-0822-3

Hu, X., Cook, S. & Salazar, M. A., 2008. Internal migration and health in China. *The Lancet*, *372*(9651), 171–1719.

Ider, E. L. & Benyamini, Y., 1997. Self-related health and mortality: A review of twenty-seven community studies. *Journal or Health and Social Behavior*, *38*(1), 21–37. doi:10.2307/2955359

Kilburn, K., Prencipe, L., Hjelm, L., Peterman, A., Handa, S., & Palermo, T., 2018. Examination of performance of the Center for Epidemiologic Studies Depression Scale Short Form 10 among African youth in poor, rural households. *BMC Psychiatry*, *18*(1), 1–13. doi:10.1186/s12888-018-1774-z

Kollamparambil, U., 2017. Impact of internal in-migration on income inequality in receiving areas: A district level study of South Africa. *The Journal of Development Studies*, *52*(12), 2145–2163. doi:10.1080/00220388.2016.1277019

Kollamparambil, U. 2021. Socio-economic inequality of wellbeing: A comparison of Switzerland and South Africa. *J Happiness Stud 22*, 555–574.

Koneru, V. K., De Mamani, A. G., Flynn, P. M. & Betancourt, H., 2007. Acculturation and mental health: Current findings and recommendations for future research. *Applied and Preventive Psychology*, *12*(2), 79–96. doi:10.1016/j.appsy.2007.07.016

Kristiansen, M., Mygrind, A. & Krasnik, A., 2007. Health effects of migration. *Danish Medical Bulletin*, *54*(1), 46–47.

Ljungvall, A. & Gerdtham, U. G., 2010. More equal but heavier: A longitudinal analysis of income-related obesity inequalities in an adult Swedish cohort. *Social Science & Medicine*, *70*(2), 221–231. doi:10.1016/j.socscimed.2009.10.014

Lu, Y., 2008. Test of the 'healthy migrant hypothesis': A longitudinal analysis of health selectivity of internal migration in Indonesia. *Social Science & Medicine*, *67*(8), 1331–1339. doi:10.1016/j.socscimed.2008.06.017

Lu, Y., 2010. Rural-urban migration and health: Evidence from longitudinal data in Indonesia. *Social Science & Medicine*, *70*(3), 412–419. doi:10.1016/j.socscimed.2009.10.028

Lu, Y. & Qin, L., 2014. Healthy migrant and salmon bias hypotheses: A study of health and internal migration in China. *Social Science & Medicine*, *102*, 41–48. doi:10.1016/j.socscimed.2013.11.040

McCullough, M.E., Kurzban, R. & Tabak, B.A., 2013. Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, *36*(1), 1–15. doi:10.1017/S0140525X11002160

Morudu P, & Kollamparambil U (2020) Health shocks, medical insurance and household vulnerability: Evidence from South Africa. *PLoS One 15*(2): e0228034. doi:10.1371/journal.pone.0228034

Mulcahy, K. & Kollamparambil, U., 2016. The impact of rural-urban migration on subjective well-being in South Africa. *The Journal of Development Studies*, *52*(9), 1357–1371. doi:10.1080/00220388.2016.1171844

Oyebode O., Pape, U. J., Laverty, A. A., Lee, J. T., Bhan, N., & Millett, C., 2015. Rural, urban and migrant differences in non-communicable disease risk-factors in middle income countries: A cross-sectional study of WHO-SAGE data. *PLoS One*, *10*(4), 47. doi:10.1371/journal.pone.0122747

Peltzer, K., Pengpid, S., Tiembre, I., Peltzer, K., Pengpid, S., & Tiembre, I., 2013. Mental health, childhood abuse and HIV sexual risk behaviour among university students in Ivory Coast. *Annals of general psychiatry*, *12*(1), 18. doi:10.1186/1744-859X-12-18

Posel, D., 2004. Have migration patterns in post-apartheid South Africa Changed. *Journal of Interdisciplinary Economics*, *15*(3–4), 277–292. doi:10.1177/02601079X04001500303

Qiu, P., Caine, E., Yang, Y., Chen, Q., Li, J., & Ma, X., 2011. Depression and associated factors in internal migrant workers in China. *Journal of Affective Disorders*, *134*(1–3), 198–207. doi:10.1016/j.jad.2011.05.043

Ray, D., 1998. Development economics. Princeton: Princeton University Press.

Rosenbaum, P.R. & Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Salant, T., Lauderdale, D. S., Salant, T., and Lauderdale, D. S. 2003. Measuring culture: A critical review of acculturation and health in Asian immigrant populations. *Social science & medicine*, *57*(1), 71–90. doi:10.1016/S0277-9536(02)00300-3

Salmond, C. E., Joseph, J. G., Prior, I. A., Stanley, D. G., & Wessen, A. F. 1985. Longitudinal analysis of the relationship between blood pressure and migration: The Tokelau Island migrant study. *American Journal of Epidemiology*, *122*(2), 291–301. doi:10.1093/oxfordjournals.aje.a114101

Smith, D. M., 2003. *The apartheid city and beyond: Urbanization and social change in South Africa*. Stellenbosch: Routledge.

Staudinger, U. M., Fleeson, W., & Baltes, P. B. (1999). Predictors of subjective physical health and global well-being: Similarities and differences between the United States and Germany. *Journal of Personality and Social Psychology*, *76*(2), 305–319. doi:10.1037/0022-3514.76.2.305

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*, *28*(12), 1270–1279. doi:10.1002/gps.3954

Tong, Y. & Piotrowski, M., 2012. Migration and health selectivity in the context of internal migration in China, 1997–2009. *Population Research and Policy Review*, *31*(4), 497–543. doi:10.1007/s11113-012-9240-y

Vearey, J., Modisenyane, M. & Hunter-Adams, J., 2017. Towards a migration-Aware health system in South Africa: A strategic opportunity to address health inequity. *SAHR*, *2017*(1), 89–98.

Wapenaar, K. & Kollamparambil, U., 2019. Piped water access, child health and the complementary role of education: Panel Data Ecidence from South Africa. *The Journal of Development Studies*, *55*(6), 1182–1200. doi:10.1080/00220388.2018.1487056

Wilson, F., 2001. Minerals and migrants: How the mining industry has Shaped South Africa. *Daedalus*, *130*(1), 99–121.

Wolf, S., Wolf, S., Hahn, E., Dettling, M., Nguyen, M.H., Wingenfeld, K., Stingl, M., Hanewald, B. and Ta, T.M.T., et al. 2017. Migration-related stressors and their effect on the severity level and symptom pattern of depression among Vietnamese in Germany. *Depression research and treatment*, *2017*, 1–9. doi:10.1155/2017/8930432

World Health Organization. 2020. *Health Impact Assessment*. [Online] Retrieved from https://www.who.int/hia/evidence/doh/en/.

Zhang, L., Liu, S. & Wu, S., 2015. Internal migration and the health of the returned population: A nationally representative study of China. *BMC public health*, *15*(1), 719. doi:10.1186/s12889-015-2074-x

Appendix

Table A1. Variable description

Variable	Description	Type
Physical health: as an independent variable	Subjective physical health: 1= if individuals rate their health as; Excellent (1), Very good (2); Good (3); or Fair (4) 0= if individuals rate their health; Poor (5)	Dummy
Depression risk: as an independent variable	Depression risk is based on the CESD-10 criteria, where responses to 10 questions (with responses on a scale of 0–3) are summed to produce a score ranging 0–30, with increasing scores indicative of higher risk of depression.	Continuous
Safety	This variable determines if any of the following occurrences are fairly common or very common in their surrounding neighbourhood, indicating that their safety is potentially compromised; (i) theft, (ii) domestic violence; (iii) gang activity; (iv) general violence; (v) drug related activity. 1= Safety is potentially compromised 0= Safety is not potentially compromised	Dummy
Medical aid	Is the individual is covered by medical insurance? 1= Yes 0= No	Dummy
Alcohol consumption	1= if the individual consumes alcohol; 0= otherwise	Dummy
Ln(rhipc)	Log transformation of real household income per capita	Continuous
Formal dwelling	1= if currently living in a formal dwelling.	Dummy

(*continued*)

Table A1. (*Continued*)

Variable	Description	Type
	0 = otherwise	
Education	1= if completed 12 years of schooling 0= otherwise	Dummy
Regular smoker	1= if Regular smoker. 0= otherwise	Dummy
Mother education	0= No formal schooling 1= Primary Schooling 2= High School 3= Undergraduate 4= Postgraduate	Ordinal
Religion	1= if very important or Important to the individual 0= Unimportant or Not important at all	Dummy
Year	Indicative of the year, namely, 2008, 2010, 2012, 2014, 2017.	Ordinal
Age	Current age of the respondent, in years	Continuous
Female	Gender of the respondent. 1= Female 0= Male	Dummy
Married	Marital status of the respondent. 1= Married or living with a partner 0= Widowed, Divorced or separated or Never married	Dummy
Unemployed	1= Unemployed; or not economically active 0= Employed	Dummy
African	1= Black African 0= otherwise	Dummy
Treat	Indicating treatment group assignment (treatment being those who engage in rural-urban migration as per the quasi-experimental definition) 1= Treatment group 0= Control group	Dummy
Post	Indicating the period post intervention. 1= Period(s) post intervention 0= Period(s) prior to intervention	Dummy
Diff in diff	The interaction term between the "treat" and "post" dummy variables	Dummy