ESSAYS ON LABOR MARKET OUTCOMES IN SOUTH AFRICA

by

Zoë McLaren

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Public Policy and Economics)
in The University of Michigan
2010

Doctoral Committee:

Professor James A. Levinsohn, Co-Chair
Professor Jeffrey A. Smith, Co-Chair
Professor David A. Lam
Professor John E. DiNardo

For Joti and Sheela

ACKNOWLEDGEMENTS

I would like to thank the members of my committee – Jim Levinsohn, Jeff Smith, David Lam and John DiNardo – for their mentorship, guidance and support. I also owe a debt of gratitude to Matt Rutledge, Osborne Jackson, Mike Gough, Brian Rowe, Elias Walsh and Brian Kovak for their advice and friendship over the years. Their invaluable suggestions and encouragement improved the work in this dissertation and helped keep me motivated.

This work would not have been possible without the committment and support from South Africans who shared their data and their expertise with me: Pali Lehohla, Peter Buwembo, Seble Worku and Jacques deKlerk at Statistics South Africa, Olive Shisana and Khangelani Zuma at the Human Sciences Research Council, and Yogan Pillay, Nhlanhla Ntuli, Mokgoropo Makgaba and Motshana Phohole at the South African Department of Health. I would also like to thank Murray Leibbrandt, Nicola Branson and Cally Ardington at the University of Cape Town for welcoming me into their academic home during the time I spent in South Africa.

In addition to those already mentioned, this dissertation has benefited from helpful comments from Raj Arunachalam, Charlie Brown, Taryn Dinkelman, Jessica Goldberg, Amar Hamoudi, Justin McCrary, Nicoli Nattrass, Edward Norton, Duncan Thomas, Rebecca Thornton and seminar participants at the University of Michigan and the University of Cape Town.

I am grateful for financial support from the William and Flora Hewlett Foundation Dissertation Fellowship, Rackham Graduate School and the Nonprofit and Public Management Center.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

This dissertation presents new evidence on the causes of persistently high unemployment in South Africa. South Africa has one of the highest unemployment rates in the world. Nearly one in four South African labor force participants are unemployed. I use three new sources of data to examine how individuals and households alter their labor market behavior to cope with negative shocks. My findings will inform policy aimed at reducing unemployment and inequality in South Africa. This volume consists of three distinct essays: the first two focus on the contribution of HIV/AIDS to unemployment and the third focuses on the role of the intra-household transfers in keeping unemployment high.

The first chapter examines the labor market impact of the largest AIDS treatment program in the world which enrolled over 500,000 patients between 2004 and 2008. I use geographic and temporal variation in the program rollout to identify the causal impact of the program. I find that for men the likelihood of labor force participation and employment both rise after an AIDS treatment clinic opens less than 15 miles away, but there are no discernable effects for women. Over time, as a greater proportion of the population begins receiving treatment, labor force participation falls and employment rises for both men and women. These results demonstrate that AIDS treatment may be undersupplied if the positive employment effects are not taken into account when designing health policy.

The second chapter, co-authored with James A. Levinsohn, Olive Shisana and Khangelani Zuma, estimates the causal effect of HIV status on employment outcomes in South Africa using two econometric methods based on the propensity score. Because individuals with HIV are more likely to fall into poverty, and the poor may be at higher risk of contracting HIV, simple estimates of the effect of HIV status on economic outcomes will tend to be biased. We rely on rich data on sexual behavior and knowledge of HIV from a large national household-based survey, which included HIV testing, to control for systematic differences between HIV-positive and HIV-negative individuals. This paper provides the first nationally representative estimates of the impact of HIV status on labor market outcomes for southern Africa. We find that being HIV-positive is associated with a seven percentage point increase in the likelihood of being unemployed. South Africans with less than a high school education are close to 11 percentage points more likely to be unemployed if they are HIV-positive. Despite high unemployment rates, being HIV-positive confers a disadvantage and reinforces existing inequalities in South Africa.

The third chapter examines whether a negative shock to household employment and the corresponding fall in the reservation wage leads unemployed household members to resume job search or obtain employment. Using nationally representative panel data, I find that men are more likely to increase search activity following a negative employment shock to the household, but only more likely to obtain employment 12-18 months after a negative shock or when the household experiences two consecutive shocks. There is no change in labor force participation for women, however, they are more likely to obtain employment 12-18 months after a negative shock. My results represent suggestive evidence that structural and frictional factors constrain household responses to negative shocks.

CHAPTER II

The Effect of Access to AIDS Treatment on Employment Outcomes in South Africa

2.1 Introduction

Researchers have established a robust causal link between health status and economic outcomes. Poor physical health can negatively affect economic outcomes by diminishing physical strength and stamina, impairing mental health, reducing productivity and causing absenteeism. South Africa is at risk to suffer a severe negative economic impact due to HIV/AIDS because its high HIV prevalence rate adds another obstacle to reducing high unemployment and high inequality. HIV strikes individuals in their prime productive years and has the potential for devastating effects because the disease is chronic, severely incapacitating in late stages, and eventually fatal. The morbidity and mortality associated with HIV/AIDS have been shown to affect household employment outcomes, consumption, savings behavior, educational attainment and investment in children (Bachmann and Booysen, 2003; Collins and Leibbrandt, 2007; Cohen, 2002; Beegle, 2005).

The labor market effects of HIV/AIDS are of particular importance because income levels constrain many household decisions. Households with HIV-positive (HIV+) members face increased expenditures for medical treatment and the specter of funeral costs, but their income is reduced due to illness or the need to care for sick household members. The direct and indirect costs of HIV contribute to the socio-economic vulnerability of these households, potentially reinforcing existing inequalities.

This study examines the economic impact of the government provision of free antiretroviral (ARV) drug treatment for AIDS in South Africa. ARV treatment offers promise as an effective policy intervention to improve the lives of the nearly 6 million South Africans who are HIV+. AIDS treatment dramatically improves health and increases life expectancy by about seven years (Bachmann, 2006). South Africa's ARV treatment program is the largest in the world even though it has only treated a fraction of the HIV+ population in South Africa (UNAIDS, 2008). The government began providing free ARV drugs in July 2004 – much later than most developing countries – but scaled up rapidly, increasing patient enrollment to about 350,000 by September 2007.

South Africa has high unemployment rates compared to other middle-income countries, which may dampen any impact of ARVs on labor market outcomes. Unemployment peaked at 30.4 percent for ages 16 to 64 in September 2002 and still stood at 22.7 percent in September 2008 (Statistics South Africa). The unemployment rates for Black Africans, who constitute 80 percent of the population, are almost twice as high as those for Coloureds, Indians and Whites (the other three races in South Africa), which underscores the degree of racial inequality.[1]

[1]Coloureds are a mixed-race group descended from Whites, Malaysian slaves, and indigenous people who intermarried in the 1600s. Indians are descendents of immigrants from India.

A large literature focuses on the impact of health shocks on consumption and employment outcomes. However, many studies are plagued by reverse causality, selection bias and/or omitted variable bias.[2] These issues are particularly salient in the health literature because economic outcomes provide inputs for health status, and many unobserved factors that influence productivity also influence health. A few studies have overcome issues of endogeneity to obtain causal estimates. For example, Thomas et al. (2006) find that Indonesian adults randomly assigned to receive a weekly iron supplement were more likely to be working and appeared to have higher productivity. In another experimental study, Miguel and Kremer (2004) find that providing deworming drugs to Kenyan schoolchildren reduced absenteeism by one quarter. Mohanan (2008) uses exogenous variation in health due to bus accidents in India and finds that injuries lead to increased debt and reduced spending on festivals and education.

Despite the importance of determining the effect of HIV on labor market outcomes, our understanding of the nature of the causal relationship in Africa is limited to evidence from a few studies. Recent work has found a sizable negative impact on employment outcomes which likely accumulates as the disease progresses and may have especially devastating effects on the poorest households. Fox et al. (2004) find that tea plantation workers in Kenya who subsequently die of AIDS exhibit a 15 percent reduction in income compared to other workers. Murray et al. (2005) find a 30 percent higher rate of workplace injuries among recently-diagnosed HIV+ miners in South Africa compared to HIV- miners. Panel data shows that these two groups had similar baseline levels of workplace injuries. In their study of mine workers in

[2]See Strauss and Thomas (1998) for further discussion.

Botswana, Habyarimana et al. (2008) find that HIV+ workers experience a five-fold increase in absenteeism when they progress to AIDS. Using propensity score methods and nationally representative survey data, Levinsohn et al. (2009) find that being HIV+ leads to a 7 percentage point reduction in the likelihood of being employed in South Africa.[3]

Though the medical literature has documented strong evidence that ARVs improve the health of HIV+ individuals in both developed and developing countries, our knowledge of whether and how these improvements in health translate into economic outcomes in Africa comes from only small-scale studies.[4] ARV treatment is associated with increases in labor force participation and productivity. Thirumurthy et al. (2008) compare a sample of HIV+ individuals who participated in an ARV treatment program in Kenya to a random sample from the surrounding area. They observe increases of 8.5 percentage points in labor force participation and 4.6 percentage points in weekly hours worked for individuals on ARV treatment, with the largest increases observed within the first six months of ARV treatment initiation. They find that men are more likely to exhibit statistically significant changes in hours worked whereas women are more likely to change their labor force participation rates. Using time since ARV initiation as an instrument for HIV-related health status, Habyarimana et al. (2008) show that mine workers in Botswana experience a large reduction in absenteeism as health status improves in the 6-12 months following treatment inception, and these gains persist for at least four years. Larson et al. (2008) conservatively estimate that one year after initiating ARV treatment,

[3] See Levinsohn et al. (2009) for a more detailed discussion of these and other related studies.

[4] See Hammer et al. (1997), Hogg et al. (1998), Palella et al. (1998), Floridia et al. (2002), Laurent et al. (2002), Marins et al. (2003), Koenig et al. (2004), Coetzee et al. (2004), Wools-Kaloustian (2006). ARVs are effective at raising white blood cell levels and reversing extreme weight loss, thereby improving overall health and increasing life expectancy by about seven years from the initiation of treatment.

tea pluckers in Kenya work twice as many days per month as they would in the absence of treatment. Using nearest-neighbor matching with panel data, Larson et al. (2009) find that after one year on ARVs, HIV+ women worked 30 percent fewer days plucking tea than the HIV- comparison group. There was no significant difference for men. Each of these studies found statistically significantly better economic outcomes for individuals who were on an ARV treatment regimen, however, the identification strategies used by Habyarimana et al. (2008) and Larson et al. (2009) produce the most convincing causal effects.

In 2004, the South African government began the rollout of free ARVs in public health clinics. It was an ambitious government program that aimed to have one ARV clinic in each of 53 districts within the first year and a clinic in each of 253 municipalities within five years (Mbewu and Simelela 2003).[5] For most South Africans, this represented the first time that ARV treatment was accessible because ARVs were not widely available in South Africa prior to this program.[6] In the first three years of the program, 339 clinics opened and approximately 350,000 patients were enrolled on ARVs. A rollout of this size and scope presents an unprecedented opportunity to examine the effect of access to AIDS treatment at the national level.

This study is the first evaluation of the largest AIDS treatment program in the world.[7] It is also the first study, to my knowledge, to use such rich, nationally representative microeconomic data to address the impact of HIV/AIDS at the national level. I provide new evidence on the impact of AIDS treatment induced improvements

[5]Table 2.1 compares the Census geographic units that are relevant to this study.

[6]While some private treatment options existed, they were generally conditional on employment, costly, or both.

[7]No internal Department of Health evaluations have been undertaken and the data have not previously been released to outside researchers.

in health status on economic outcomes, including spillover effects within households and neighborhoods. The centralized accreditation process used to determine the timing of clinic opening provides an exogenous source of geographic and temporal variation in ARV access. This allows me to identify a plausibly causal effect of improved access to treatment on employment outcomes.

I combine newly available data from ARV treatment clinics with detailed, nationally representative economic outcome data. I take advantage of geographic and temporal variation in the ARV program rollout to determine the effect of access to ARV treatment on labor force participation and employment. I use detailed geographic coordinates to link data from the semi-annual South African Labour Force Survey (LFS) to data on patient enrollment at public ARV clinics. Having seven consecutive waves of the LFS data spanning September 2004 through September 2007 enables me to perform fixed effects (FE) estimation at the neighborhood level.

I find that after a clinic opens between 3 and 15 miles away, labor force participation rises by 2 percentage points for Black African men and employment rises by 3.3 percentage points. There are no discernible effects of the distance to the nearest clinic for women. As clinics grow over time, I find that a one percentage point increase in the fraction of the neighborhood population receiving treatment decreases labor force participation by 0.5 percentage points and raises employment by 0.4 percentage points for men. Women exhibit similar patterns.

Considering that my estimates average over the entire Black African population, including HIV+ individuals in the latent stage of HIV infection and HIV- individuals, the labor market impact for households that obtain ARV treatment is substantial. If these positive labor market effects aren't taken into account when designing health policy in South Africa, ARV treatment may be under-supplied.

The rest of this paper is organized as follows: Section 2 lays out the context for the ARV treatment rollout, Section 3 describes the data used for the analysis, Section 4 presents methods and results and Section 5 concludes.

2.2 Context

This study examines the impact on employment outcomes of an improvement in health at an advanced stage of HIV infection. In South Africa, as in most developing countries, HIV+ individuals do not start highly-active anti-retroviral treatment (HAART) until approximately 8 to 10 years after initial infection, when a patient progresses to full-blown AIDS. ARVs are prescribed once a patient's white blood cell count (i.e. CD4 count) drops below 200 cells/cubic millimeter or they exhibit opportunistic infections or cancers characteristic of a depleted immune system.[8] Once patients are deemed eligible for ARVs, they must complete a 2-to-4-week adherence trial to ensure they are able to follow the daily medication regimen; only then will ARVs be prescribed. Patients exhibit dramatic improvements in health once the treatment is initiated. For example, Coetzee et al. (2004) found that within three months of treatment initiation, the percent of patients with a viral load below 400 copies/ml had increased from zero percent to 88 percent, which is accompanied by improvements in overall health.

[8]The list of diseases includes tuberculosis, Kaposi's sarcoma, fungal infections and some viral infections that a full-strength immune system would be able to fight. See World Health Organization (2005) for additional information.

2.2.1 Conceptual Framework

The dramatic improvement in health brought about by ARV treatment should raise the productivity of sick workers which would increase labor force participation, search activity, employment and wage income. Improved access to ARVs can also affect employment outcomes through changes in expectations for individuals who may require treatment in the future. ARV treatment extends life expectancy by approximately seven years: individuals are likely to survive about eight years after treatment initiation compared to one year or less without treatment (Coetzee et al., 2004). Because the increase in life expectancy comes within ten years for anyone who is already HIV+, it would have a larger impact on behavior than a similar change in life expectancy that is only realized in old age. The cost of losing twenty years of life is higher in youth than in old age and with discounting, even a similar cost would mean less further into the future.

Increased access to treatment will produce a temporal spillover effect as individuals who know or suspect they are HIV+ increase labor market attachment in response to improved access to treatment they will need in the future. On the other hand, intra-household effects could lead to the crowding out of the labor force participation of some members of the household. We might see no change in household-level employment rates, for example, if the primary breadwinner returning to the labor force after being ill allows a secondary earner to leave the workforce with no reduction of household income. Positive effects will spill over within the household if family or extended family members, usually women or young adults, are released from caretaking responsibilities and are therefore able to pursue employment outside the home.

The size and direction of the impact observed depends on the relative magnitudes of these effects.

Because my data are nationally representative, I am able to examine the equilibrium effects of changes on the labor supply-side. A priori and without demand-side data, it is unclear whether more healthy workers in the local labor force would increase unemployment or put downward pressure on wages. I expect to see the largest effect on participation and search activity because these changes can happen virtually overnight. The effect on employment is likely to be smaller because of the lags involved in obtaining employment and the labor demand-side factors that may limit hiring.

There are two reasons there may be a lag between obtaining access to ARV treatment and any change in labor market outcomes. First, it takes 3 to 6 months after enrollment in a treatment program for a patient to realize the initial health improvements from ARVs and longer for the full health benefits to come to fruition. Second, though individuals can adjust their labor supply almost immediately, it generally takes time to obtain employment, especially in the South African labor market.[9]

2.2.2 Description of the Rollout

The South African government began providing ARVs in public health clinics in July 2004, after a long delay caused by lack of political support. ARV provision was a highly politicized issue because the President and Health Minister, despite ample scientific evidence to the contrary, had perpetuated myths that HIV did not cause AIDS

[9] McLaren (2010) shows that in the South African labor market, it can take 6-12 months for an increase in individual labor force participation to result in a change in employment status.

and that ARVs were ineffective (Nattrass, 2007). Meanwhile, South Africa's northern neighbor, Botswana, launched its national ARV treatment program in December 2001, two and a half years before South Africa would begin enrolling patients (Chigwedere et al. 2008). The South African treatment program was initiated following a civil disobedience initiative by the Treatment Action Campaign, a non-governmental organization which demanded that the government provide life-saving ARV drugs to its citizens (Nattrass, 2007).

The government's stated goal was to provide "equitable" access to AIDS treatment within every locality so that every South African would have access to these services and historically under-served districts would receive the same standard of care as more advantaged districts (Mbewu and Simelela, 2003). The Department of Health (DOH) conducted a centralized accreditation process to ensure that the proposed ARV treatment facilities were equipped to provide a standard quality of care. The minimum requirements for accreditation included having an on-site team of clinicians, nurses and nutritionists, access to care 24 hours a day, access to lab services, access to a pharmacy with secure drug storage facilities, adequate on-site consultation space and a patient tracking and monitoring system.[10]

Provincial governments were tasked with heading up the implementation of the clinic rollout and were granted a fair amount of autonomy. Provincial departments of health identified facilities that would apply for accreditation and shepherded clinics through the process. Provinces selected facilities with the following goals in mind:

[10]Though this appears to be an extensive list, the requirements "largely coincide with current standard operating procedures and practices at public health care facilities in South Africa" and the government pledged that "additional financial and technical resources [would] be deployed to service points in resource-constrained or underserved areas" to help them meet the requirements (Mbewu and Simelela, 2003).

establishing one clinic per district (and once that had been achieved, one clinic per sub-district); selecting facilities that could meet the requirements for accreditation within one year; opening additional clinics in districts with large AIDS populations; and choosing geographic locations that would keep transportation costs low for patients in the service area.[11] The national DOH evaluated applications from these clinics, provided feedback and made on-site visits to determine whether the criteria for accreditation had been fulfilled. There was high demand for accreditation once the process began and the small DOH accreditation team struggled to keep up with the demand.[12] The accreditation process created the main bottleneck in delivering ARV treatment to a particular location. In most cases, proposed ARV treatment centers were already providing other HIV/AIDS care on-site so they were able to start enrolling patients from this pool in the ARV program immediately following accreditation.[13]

The pattern of clinic openings provides additional evidence that the accreditation process was the primary determinant of the pace of the rollout. Clinics opened around the country at a fairly constant rate over the period in question (see Figure 2.1). There was variance across the nine provinces in the number of clinics opening each month with Kwa-Zulu Natal opening half of its clinics by early 2005 while the Free State had only opened half of its clinics by mid-2006 (see Figure 2.2). On-site accreditation visits were conducted individually and not in one particular geographic area at a time, however, some temporal clustering of clinic openings by

[11] Mbewu and Simelela, (2003); Personal communication with Andronica Ratshefola, former Assistant Director for the Comprehensive HIV and AIDS Care, Management and Treatment Program, 29 July 2009.

[12] Ratshefola, personal communication, 29 July 2009.

[13] Other HIV/AIDS care included HIV testing, counseling and treatment of tuberculosis and opportunistic infections. It is important to note that adding ARV drugs to this treatment regimen considerably improved the effectiveness of care.

province is evident from the data.[14] The nine provinces vary widely in their resource endowment, management competence and political efficacy, so it is not surprising that there was wide variation in effectiveness at the provincial level for this program. Gauteng, Western Cape and Northwest Province scaled up quite quickly in the early periods; in fact, the Western Cape had begun planning for clinics prior to the official commencement date of the program (Venter 2006). Venter (2006) suggests that poor management and lack of political resolve accounted for some provinces lagging behind in the rollout.

During the period of study, the demand for ARV treatment far exceeded the supply and there was queueing for appointments at every clinic site.[15] Enrollment capacity was determined by the national DOH guidelines for the maximum allowable ratio of enrolled patients to health workers (Mbewu and Simelela, 2003). Staff numbers were set during the accreditation process and were therefore unlikely to change from month to month in response to any changes in demand.[16] Monthly enrollment rates at each clinic corroborate that: patients were enrolled at each clinic at a fairly constant rate over the period of study.

The best measure of access to ARV treatment is the distance from a neighborhood to the nearest open clinic since the cost of travel tends to be directly related to the distance travelled in South Africa. One legacy of the Apartheid era is that transportation costs are generally high because there are limited public transportation

[14]This is partially due to provincial plans that set monthly or quarterly goals for accreditation.

[15]Personal communication with Dr. Francois Venter, President of the Southern African HIV Clinicians Society and Clinical Director of the University of Witwatersrand Reproductive Health and HIV Research Unit, 27 July 2009.

[16]For these same reasons, the supply of care at ARV treatment sites was also likely to be unresponsive to changes in demand for other HIV care, which could arise due to differential HIV testing behavior once ARVs were available, for example.

options.[17] A change in the distance to the nearest AIDS treatment clinic should affect the likelihood that an individual seeks treatment. Figure 2.3 shows that though there was substantial variation in the distance to the nearest clinic in September 2004 near the beginning of the rollout; by March 2006 half the neighborhoods in the sample were less than 8.5 miles from the nearest clinic. South Africans may obtain ARV treatment at any public clinic in the country, but most will attend the nearest clinic to keep travel costs down.[18] In fact, travel costs were cited as a primary obstacle to obtaining access.[19] Travel costs can add up since multiple visits are required to determine eligibility for ARV treatment and to monitor adherence and toxicity. ARV patients must return to the clinic about four times in the first two months, and then either monthly, or quarterly if there are no complications.[20]

The number of patients receiving care through public clinics dwarfed the number of patients receiving care through private clinics (see Figure 2.4). Private sector enrollment grew at a constant rate between 2004 and 2007. About 45,000 patients were in treatment in private clinics in October 2004, and the number of patients had only increased to 60,000 by mid-2005 and 67,600 by mid-2006 (Johnson, 2006; Johnson and McLeod, 2007). On the other hand, the public sector grew exponentially. It surpassed private enrollment in mid-2005 and by 2008 was almost six times as large. Because private insurance is costly and often tied to employment, public clinics serve a clientele that is more likely to be poor, unemployed and non-White.

[17] The Apartheid government took measures to limit the mobility of non-Whites.

[18] Though concerns about stigma may lead some individuals to avoid seeking treatment at the closest clinic, Thornton (2008) finds that a small cash incentive was enough to overcome these types of psychological costs.

[19] Ratshefola, personal communication, 29 July 2009.

[20] Personal communication with Dr. Matthew Fox, Assistant Professor of International Health, Boston University, 26 August 2009.

2.3 Data

I create a new data set by linking detailed economic data from the South Africa Labour Force Survey (LFS) with Department of Health data from clinics providing ARV treatment. I obtained geographic coordinates of the ARV clinics and LFS neighborhoods which enables me to calculate the distance from a neighborhood to the nearest clinic as new clinics opened.

2.3.1 Labour Force Survey data

The LFS is the most comprehensive source of national microeconomic data in South Africa and is therefore particularly suited to an analysis of the economic impact of the ARV rollout. The survey collects detailed information about the labor market situation of individuals aged 15-65 years, and basic information about children and seniors, in a nationally representative sample of approximately 30,000 households. The questionnaire includes questions about demographic characteristics; biographical information; activities related to work; unemployment and non-economic activities; agricultural activities and uncompensated activities. The LFS is equivalent to the U.S. Current Population Survey, but is conducted only twice-yearly, in March and September. I use seven waves of data for my analysis, from September 2004 (LFS wave 10) through September 2007 (LFS wave 16).[21] I use sampling weights provided by Statistics South Africa (StatsSA) in generating all my results. These weights correct for dwelling-unit non-response and are benchmarked to population estimates.[22] Though the LFS data were designed as a rotating panel, with approximately 20 per-

[21] These waves use a master sample based on the 2001 Census that was drawn for LFS wave 10 (September 2004). The survey continued to be conducted after September 2007; however, it used a new sampling frame and altered a number of the components of the survey. I do not have detailed geographic data for waves beyond September 2007 so I cannot incorporate them into the analysis at this time.

[22] The overall dwelling unit non-response rate is approximately 10 percent. The author imputed instances of item non-response to zero.

cent of the households refreshed every wave, it is not possible to link observations between survey waves to create a true panel of individuals or households. The survey was not designed to track individuals across waves. However, neighborhoods can be linked over time using geographic coordinates. Employment status is derived according to standard International Labour Organization (ILO) definitions using responses to a series of questions in the survey.[23]

I link the LFS data with proprietary geographic information at the primary sampling unit (PSU) level. There are 2,996 PSUs in my sample, of which 441 are within one of the seven metropolitan areas. The PSUs range in size from 0.0075 square miles in dense urban areas to 1,460 square miles in the Karoo desert, but most are small. The distribution has a median area of 0.15 square miles, which roughly corresponds to the sixteen city blocks in Manhattan, and an average area of 6.4 square miles.[24]

2.3.2 Clinic enrollment data

The measures of access to treatment used in this analysis are derived from clinic-level reports of patient enrollment in ARV treatment. I use proprietary data collected by the South African Department of Health from all government-sponsored ARV clinics that opened between July 2004, when the rollout began, and September 2007. The data contain monthly reports of the cumulative number of patients initiated on treatment at each clinic. Data on attrition due to death or loss to follow-up were not collected. However, because demand outstripped supply and supply was very inelastic over this period, the enrollment numbers are a good measure of the sup-

[23]The only difference from U.S. Bureau of Labor Statistics (BLS) definitions is that seasonal workers in the off-season are also considered out of the labor force (also known as not economically active), rather than as unemployed.

[24]The distribution is highly skewed: only ten neighborhoods are larger than 5,000 square miles.

ply of treatment available at a particular time. I classified clinics as "open" once they had enrolled at least one patient.[25] The first fifty-three clinics opened in July 2004, and by September 2007 there were 339 clinics open in South Africa. Figure 2.5 shows that clinics were clustered in the main metropolitan areas, but otherwise were geographically widely dispersed. Approximately two-thirds of the clinics were in hospital facilities and one-third were in smaller health centers (see Table 2.2). Cumulative patient enrollment numbers increased from 5,574 in July 2004 to 357,292 in September 2007 .

2.4 Results

2.4.1 Empirical strategy 1

I use seven waves of the LFS to perform fixed effects (FE) estimation at the PSU level to control for bias due to time-invariant unobserved variables.[26] I estimate the following equation separately by gender for individual i in neighborhood j at time t:

$$Y_{ijt} = \beta_0 + \beta_1 nearest0to3mi_{ijt} + \beta_2 nearest3to15mi_{ijt} + \phi' X_{ijt} + \delta_t + \alpha_j + \epsilon_{ijt}, \quad (2.1)$$

where Y_{ijt} is an indicator variable for the outcome of interest, either labor force participation or employment, $nearest0to3mi_{ijt}$ is an indicator for the nearest open clinic being within 3 miles of the neighborhood centroid and $nearest3to15mi_{ijt}$ is an indicator for the nearest clinic being between 3 and 15 miles away.[27] The vector X includes the following covariates: age, age squared, years of primary education, years of secondary education, have completed a Matric (high school), have some post-Matric education, ever been married, spouse lives in household, number of adults in the

[25] To my knowledge, clinics were never closed. However some failed to enroll a positive number of patients in a subsequent month (as reported in the DOH data), which is likely due to non-reporting.

[26] Fixed-effects estimation relies on the assumption that the mean of the outcome variable is stable over time. This is likely to hold in repeated cross sections and rotating panels (Heckman and Robb, 1985). One exception is if there is selective attrition on unobservables, for example, due to differential AIDS death rates.

[27] These cut-off points correspond to the tertiles of the initial distribution of distance to the nearest clinic.

household and number of children (aged 14 and under) in the household. I include a set of interactions between survey wave and district to control for district-specific time effects.[28] I restrict the sample to individuals between age 25 and 44 because the HIV prevalence rate peaks for men and women in this age group. Standard errors are clustered at the main place level, which is slightly larger than the neighborhood.

2.4.2 Empirical strategy 2

ARV treatment was highly rationed in South Africa because demand for treatment far exceeded the supply throughout the period of analysis (the first three years of the rollout). Access to treatment can be measured by the ratio of treatment enrollment slots to the population at risk for needing treatment. South Africans may seek ARV treatment initiation at any clinic but are likely to use nearby clinics, though not necessarily the nearest clinic, to minimize transportation costs. I take into account the availability of treatment slots and the competition from other at-risk populations in the vicinity of each neighborhood. I use data on treatment enrollment by clinic to calculate the likelihood that an HIV+ individual in a particular neighborhood (PSU) is able to obtain a treatment enrollment slot.

Although only a fraction of HIV+ individuals have progressed beyond the latent stage of the disease and are in need of ARV treatment, I use the size of the HIV+ population as a good estimate of the relative size of the population at risk for needing ARV treatment. The vast majority of patients at public ARV clinics are Black African. I therefore calculate the size of the HIV+ population in a particular neighborhood by multiplying the Black African population in a neighborhood by the

[28] Figure 2.5 shows the 53 districts outlined in grey.

probability of being HIV+ in the municipality (m),

$$HIV^{+}Popn_j = BlackAfricanPopn_j * ProbHIV_m^{+}, \tag{2.2}$$

where the probability of being HIV+ is estimated by the HIV prevalence rate which is obtained from the 2005 SABSSMII survey.

I do not know which individuals in a neighborhood are obtaining ARV treatment nor do I know at which clinic individuals choose to obtain treatment. I therefore rely on the same assumption used in the first empirical strategy: that individuals are more likely to obtain treatment from a particular clinic the closer that clinic is. The demand for treatment is therefore inversely related to the distance to the nearest clinic,

$$DemandForTreatment_{jk} = \frac{HIV^{+}Popn_j}{Distance_{jk}} * 1(Distance_{jk} < radius). \tag{2.3}$$

I also assume that there is a maximum distance beyond which individuals will not travel to access a clinic. This enters the equation above as the $radius$, set to 50 miles in my analysis, which is the distance between a neighborhood and clinic beyond which demand is set to zero.[29] I use this estimated demand function as a weight in calculating the likelihood of an individual obtaining a treatment slot, where the denominator normalizes the weights to sum to one over all neighborhoods (J) for each clinic (k):

$$\frac{Demand_{jk}}{\sum_{j=1}^{J} Demand_{jk}}. \tag{2.4}$$

Recall that this $Demand_{jk}$ is zero for all neighborhood-clinic pairs that are beyond

[29]The results do not change if I use log distance in Equation 2.3. Results do not change when the radius varies between 10 miles and 50 miles.

the distance specified by the radius.

Using the weight in Equation 3.2, I allocate patients enrolled at clinic k in time t to neighborhoods $j = 1, ..., J$ with probability proportional to the demand for treatment. The allocated patients are summed over all clinics to calculate the total number of treatment slots available to a particular neighborhood (PSU):

$$PSUTreatmentSlots_{jt} = \sum_{k=1}^{K} EnrolledPatients_{kt} * \left(\frac{Demand_{jk}}{\sum_{j=1}^{J} Demand_{jk}} \right). \tag{2.5}$$

I then calculate the likelihood of being treated in a particular neighborhood in time t as the number of treatment slots available to that neighborhood in Equation 2.5 scaled by the neighborhood population,

$$ProbabilityOfBeingTreated_{jt} = \frac{PSUTreatmentSlots_{jt}}{BlackAfricanPopn_j}, \tag{2.6}$$

where I use the PSU population from the 2001 South African Census so that changes in the probability of being treated come from changes in the availablility of treatment slots, and not from changes in the PSU population.

I estimate a similar fixed-effects regression as in Equation 2.1, where the treatment variables of interest are replaced with $ProbabilityOfBeingTreated_{jt}$:

$$Y_{ijt} = \beta_0 + \beta_1 ProbabilityOfBeingTreated_{jt} + \phi' X_{ijt} + \delta_t + \alpha_j + \epsilon_{ijt}, \tag{2.7}$$

and the rest of the specification is identical.

2.4.3 Specification tests

Because I am identifying off changes within a neighborhood over time we would be concerned if neighborhoods with better access in September 2004 have different changes over time in underlying characteristics compared to neighborhoods with less access in September 2004. I compare those with better access to AIDS treatment clinics in September 2004 with those who have less access, splitting the sample at the median distance to the nearest clinic. I perform fixed effect regressions similar to my estimation equation, but I regress each X covariate on the other covariates, time dummies, and the set of time dummies interacted with an indicator variable for having better access to ARV treatment:

$$X_{k\_ijt} = \beta_0 + Near * \gamma_t + \gamma_t + \phi' X_{-k\_ijt} + \alpha_j + \epsilon_{ijt}. \tag{2.8}$$

Table 2.3 reports some descriptive statistics for the sample of Black men used in my analysis and results from these diagnostic regressions. The first two columns contain raw sample means of individual covariates for individuals who were nearer to clinics at baseline with those who were farther, respectively, and the third column presents the t-statistic on the difference in means. The fourth and fifth columns contain the change in the value of each characteristic between the baseline wave (September 2004) and the final wave in the sample (September 2007). The sixth column contains the F-statistic for the joint test of the null of zero coefficients on the full set of time dummies interacted with the dummy for more access ($Near$), and the seventh column contains the p-value for this F-statistic. The t-statistic on the difference between these raw means indicates a statistically significant difference on most covariates. However, the results from the F-test show that there are no significant differences over time between neighborhoods with more access and those with less, providing additional evidence that the variation in the timing of clinic opening was exogenous. When the analysis is repeated in Table 2.4 for the sample of Black women, there are only statistically significant differences over time in the proportion of individuals who lived in the same residence six months ago. Considering the number of coefficients involved in these tests, there are no more statistically significant tests than we might expect by chance. Additionally, it is clear from the values in columns 4 and 5 that these differences are generally small.[30]

2.4.4 Results

Table 2.5 presents the first set of main results examining the impact of the distance to the nearest clinic on labor force participation for the sample of Black men. Column 1 contains results from regressions of an indicator for being a labor force participant on two measures of treatment access. The specifications in columns 2, 3 and 4 add a set of individual characteristics, district-wave interactions and neighborhood-level fixed effects, respectively. The unconditional estimate in column 1 indicates that Black men who live close to clinics are statistically significantly more likely to be labor force participants than those who live between 15 and 75 miles from a clinic.[31] Adding the additional controls and district-wave interactions adjusts for potential sources of endogeneity, reducing the size of the coefficients and their significance. The preferred specification in column 4 shows that Black men are 2 percentage points more likely to be labor force participants after a clinic opens between 3 and 15 miles of their residence compared to those who are more than 15 miles away. This effect is significant at the 90 percent confidence level. Surprisingly, I do not find a significant

[30]Results are nearly identical if the predicted proportion of the neighborhood being treated is used as the measure of access (not shown).

[31]Neighborhoods that are still more than 75 miles from the nearest clinic in September 2007 are excluded from the analysis in all waves because they are particularly remote and may differ systematically from other neighborhoods. Only 0.8 percent of the sample is dropped due to this restriction.

impact when a clinic opens less than 3 miles away. The point estimate is smaller – 0.7 percentage points – but not statistically significantly smaller (F=1.06, p=0.30).

The signs and magnitudes of the coefficients on the individual characteristics are what we would expect in the context of the South African labor market. The impact of having completed high school (the Matric examination) is a statistically significant 2.4 percentage points and about one third of the magnitude of the impact of having completed some post-Matric schooling. Individuals who have completed at least one year of schooling beyond the Matric (either a technical or university degree, or higher) have substantially higher labor force participation. Having a spouse in the household, conditional on marital status, is associated with increased labor force participation for men, but additional household members decreases the likelihood of participation.

My analysis focuses on results from the preferred regression specification that includes individual characteristics, district-time interactions and neighborhood fixed effects. For comparison, Table 2.6 repeats results from columns 3 and 4 of Table 2.5 in the first two columns and presents results for the likelihood of employment in the second two columns. For Black men, the impact of distance to the nearest clinic is similar for employment and labor force participation. There is a positive, but not significant, increase in employment of 1.8 percentage points when the distance to the nearest clinic is less than 3 miles. However, the likelihood of employment rises by 3.3 percentage points when the distance to the nearest clinic is between 3 and 15 miles. This result is significant at the 95 percent confidence level. There is a bigger impact of schooling on the likelihood of employment than on the likelihood of participation, as we would expect, and we observe similar patterns for the household composition characteristics for both participation and employment.

For Black women, labor force participation and employment appear to be unaffected when the distance to the nearest clinic is below 15 miles; the point estimates are small and not statistically significant (see Table 2.7). Though the point estimates are positive and significant in the specification in column 1, when the fixed effects are added in column 2 the magnitude falls dramatically and the point estimates are no longer significant. Women exhibit a similar pattern in terms of employment. It appears that the likelihood of employment actually decreases by about 1 percentage point when the distance to the nearest clinic is between 3 and 15 miles, however, this point estimate is not significant. The returns to schooling in terms of participation and employment are larger for women than for men, especially for the impact of having some post-Matric education on the likelihood of employment. Women are less likely to participate or be employed if their spouse resides in the household. Having an additional adult in the household reduces the likelihood of participation and employment for women, but to a lesser degree than for men.

Table 2.8 presents the first set of results from Empirical Strategy 2 that examines the impact of the fraction of the neighborhood receiving treatment on the likelihood of labor force participation for Black men. Column 1 shows that when one percentage point more of the population obtains access to treatment, men are 0.2 percentage points more likely to be participants. However, including district*wave interactions reduces the point estimate to zero, and including neighborhood fixed effects leads to an estimated reduction in the likelihood of participation of 0.5 percentage points. It is worth noting that the coefficients on the other covariates are virtually identical to the estimates from Empirical Strategy 1.

As before, Table 2.9 repeats columns 3 and 4 of the previous table for comparison, and presents results for the impact of the fraction of the neighborhood receiving treatment on employment. The preferred specification in column 4 shows that men are 0.4 percentage points more likely to be employed when one percentage point more of the neighborhood population obtains access to treatment. Women exhibit a similar pattern (see Table 2.10). They are slightly, though not statistically significantly, less likely to participate as the fraction treated grows, but they are 0.7 percentage points more likely to be employed when one percentage point more of the neighborhood population is treated.

2.5 Discussion

This paper provides new evidence on the impact of improved access to AIDS treatment on employment outcomes in South Africa. Using geographic and temporal variation in the rollout of ARV clinics, I find that having a clinic between 3 and 15 miles away increases labor force participation by 2 percentage points and employment by 3.3 percentage points for Black men. There are no discernible effects of the distance to the nearest clinic for women. As clinics grow over time, I find that a one percentage point increase in the fraction of the neighborhood population receiving treatment decreases labor force participation by 0.5 percentage points and raises employment by 0.4 percentage points for men. Women exhibit similar patterns.

Empirical Strategy 1 picks up the impact of having a clinic whereas Empirical Strategy 2 also picks up the effect of existing clinics growing over time. For Empirical Strategy 1, I expected the coefficient on the indicator for a clinic being within 3 miles to be larger than that for being between 3 and 15 miles. One explanation for the opposite result is that the most meaningful change in access for men is when the distance to the nearest clinic falls below 15 miles. It may reflect that neighborhoods with a clinic less than 3 miles away likely already have a clinic between 3 and 15 miles away and have already adjusted their labor market attachment accordingly. Also, nearer clinics may be newer clinics that have treated fewer people.

It is not surprising that improved access to ARV treatment, measured by the fraction of the neighborhood treated, leads to an increase in employment for men and women, however, the decrease in labor force participation for both sexes is unexpected. The 0.5 percentage point decrease in the likelihood of labor force participation for men may reflect a return to equilibrium for men who increased their participation in response to a clinic opening, but were unable to find work. It does not appear to reflect an increase in investment in education or training since enrollment in educational programs (broadly defined) does not rise over this period (results not shown). Further analysis is required to investigate these possibilities.

I observe a slightly larger impact on the likelihood of employment for women than for men, but the difference by gender is not statistically significant. I expected to find a larger effect for women than men for three reasons. First, the HIV prevalence rate among women is almost twice the rate among men (Shisana, 2005). Second, women generally outnumber men more than two-to-one at treatment facilities, exceeding the ratio that can be explained by differences in prevalence alone (Nattrass, 2008; Muula et al., 2007). Third, women are more likely to be exposed to spillover effects within the household because of their traditional role as care-givers.

One limitation of this study is that I cannot examine outcomes separately by HIV status because my data lack information on individual HIV status.[32] I focus my analysis on the sample of households most likely to contain an HIV+ person based on the age profile of HIV prevalence, but cannot differentiate between households with an HIV+ member and those without. It is difficult to find household and individual characteristics that reliably predict HIV status other than race (Levinsohn et al., 2009). The treatment effect is heterogeneous along many dimensions: HIV status, the stages of HIV infection, an individual's perceived risk of acquiring HIV, and the number of household members that are infected. Individuals who have converted to AIDS should have a substantially larger treatment effect than those who believe they are unlikely to be HIV+ (regardless of their actual HIV status) and have no HIV+ household members.

My results underscore the potential benefits of providing targeted labor market interventions alongside the rollout of AIDS treatment. Considering that my estimates average over the entire population, including HIV- individuals and HIV+ individuals in the latent stage of HIV infection, the implied labor market impact for households that obtain ARV treatment is substantial. If these positive labor market effects aren't taken into account when designing health policy in South Africa, ARV treatment may be under-supplied.

[32] The design does, however, allow me to incorporate general equilibrium and spillover effects, which are not generally discernible from the typical ARV studies that collect information from ARV patients only, but not their households, neighborhoods or local labor markets.

2.6 References

Bachmann, Max O. 2006. "Effectiveness and cost effectiveness of early and late prevention of HIV/AIDS progression with antiretrovirals or antibiotics in Southern African adults." *AIDS Care*, 18: 109-120. (Erratum in *AIDS Care*, 2006, 18: 415-416.)

Bachmann, Max O. and Frederick L.R. Booysen. 2003. "Health and economic impact of HIV/AIDS on South African households: a cohort study," *BioMed Central Public Health*, 3 (14).

Beegle, Kathleen. 2005. "Labor Effects of Adult Mortality in Tanzanian Households," *Economic Development and Cultural Change*, 53 (3), 655-83.

Chigwedere, P, George R. Seage III, Soa Gruskin, Tun-Hou Lee and M. Essex. 2008. "Estimating the Lost Benefits of Antiretroviral Drug Use in South Africa." *Journal of Acquired Immune Deficiency Syndrome*, 49(4).

Coetzee, D, K. Hildebrand, A. Boulle, G. Maartens, F. Louis, V. Labatala, H. Reuter, N. Ntwana and E. Goemaere. 2004. "Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa." *AIDS*, 18: 887-95.

Cohen, Desmond. 2002. "Human capital and the HIV epidemic in sub-Saharan Africa," ILO Programme on HIV/AIDS and the World of Work Working Paper, June.

Collins, Daryl L. and Murray Leibbrandt. 2007. "The financial impact of HIV/AIDS on poor households in South Africa," *AIDS*, 21 (Suppl 7), S75-S81.

Floridia, M., V Fragola, CM Galluzzo, G Giannini, MF Pirillo, M Andreotti, C Tomino and S Vella. 2002. "HIV-Related Morbidity and Mortality in Patients Starting Protease Inhibitors in Very Advanced HIV Disease," *HIV Medicine*, 3(2): 75-84.

Fox, Matthew P., Sydney Rosen, William B. MacLeod, Monique Wasunna, Margaret Bii, Ginamarie Foglia, and Jonathon L. Simon. 2004. "The impact of HIV/AIDS on labour productivity in Kenya," *Tropical Medicine and International Health*, 9 (3), 318-24.

Habyarimana, James, Bekezela Mbakile and Cristian Pop-Eleches. 2008. "HIV/AIDS, ARV Treatment and Worker Absenteeism: Evidence from a Large African Firm." Georgetown University: Mimeo.

Hammer, S.M., K. Squires, M. Hughes, J. Grimes, L. Demeter, J. Currier, J. Eron, J. Feinberg, H. Balfour, L. Deyton, J. Chodakewitz and M. Fischl. 1997. "A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less." *New England Journal of Medicine*, 337:725-733.

Heckman, J. J. and R. Robb Jr. 1985. "Alternative methods for evaluating the impact of interventions: An overview." *Journal of Econometrics* 30(1-2): 239-267.

Hogg, Robert S., K.V. Heath, B. Yip, K.J. P. Craib, M.V. OShaughnessy, M.T. Schechter, J. S. G. Montaner. 1998. "Improved survival among HIV-infected individuals following initiation of antiretroviral therapy." *Journal of the American Medical Association*, 279: 450-454.

Johnson, LF. 2006. "Estimated numbers of patients on antiretroviral treatment in the South African private health sector." *AIDS Analysis Africa Online*, 2006, Jan/Feb: 4-6.

Johnson, Leigh F. and Heather D. McLeod. 2007. "Steady growth in antiretroviral treatment provision by disease management and community treatment programmes." *South African Medical Journal*, 97 (5), May.

Koenig, Serena P., Fernet Leandre, and Paul E. Farmer. 2004. "Scaling-up HIV treatment programmes in resource-limited settings: the rural Haiti experience." *AIDS*, 18: S21-S25.

Larson, B., M. Fox, S. Rosen, M. Bii, C. Sigei, D. Shaffer, F. Sawe, K. McCoy, M. Wasunna and J. Simon. 2009. "Do the socioeconomic impacts of antiretroviral therapy vary by gender? A longitudinal study of Kenyan agricultural worker employment outcomes." *BMC Public Health*, 9:240.

Larson, B., M. Fox, S. Rosen, M. Bii, C. Sigei, D. Shaffer, F. Sawe, M. Wasunna and J. Simon. 2008. "Early effects of antiretroviral therapy on work performance: preliminary results from a cohort study of Kenyan agricultural workers." *AIDS*, 22(3).

Laurent, C., N. Diakhate, N. F.N. Gueye, M.A. Toure, P.S. Sow, M.A. Faye, M. Gueye, I. Laniece, C.T. Kaned, F. Liegois, L. Vergne, S. Mboup, S. Badiane, I. Ndoye and E. Delaporte. 2002. "The Senegalese government's highly active antiretroviral therapy initiative: an 18-month follow-up study." *AIDS*, 16: 1363-1370.

Levinsohn, James, Zoë McLaren, Olive Shisana and Khangelani Zuma. 2009. "HIV and Labor Market Participation in South Africa." University of Michigan: Mimeo.

Marins, J.R.P., L.F. Jamal, S. Y. Chen, M.B. Barros, E.S. Hudes, A.A. Barbosa, P. Chequer, P.R. Teixeira and N. Hearst. 2003. "Dramatic improvement in survival among adult Brazilian AIDS patients." *AIDS*, 17:1675-1682.

Mbewu, Anthony and Nono Simelela. 2003. "Operational Plan for Comprehensive HIV and AIDS Care, Management and Treatment for South Africa." South African Government Information Service. (Downloaded 19 November 2003 from http://www.info.gov.za/issues/hiv/careplan.htm)

McLaren, Zoë. 2010. "How does intra-household job loss affect labor market outcomes of the unemployed in South Africa." University of Michigan. Mimeo.

Miguel, Edward and Michael Kremer. 2004. "Worms: Identifying impacts on education and health in the presence of treatment externalities." *Econometrica*, 72(1), January, 159-217.

Mohanan, Manoj. 2008. “Causal Effects of Health Shocks on Consumption and Debt: Quasi-Experimental Evidence from Bus Accident Injuries.” Harvard University. Mimeo.

Murray, Jill, Pam Sonnenberg, Gill Nelson, Stuart Shearer, Andre Bester, Arthur Begley, and Judith R. Glynn. 2005. “Effect of HIV on work-related injury rates in South African goldminers,” *AIDS*, 19:2019-24.

Muula, Adamson S., Thabale J. Ngulube, Seter Siziya, Cecilia M Makupe, Eric Umar, Hans Walter Prozesky, Charles S. Wiysonge and Ronald H. Mataya. 2007. “Gender distribution of adult patients on highly active antiretroviral therapy (HAART) in Southern Africa: a systematic review.” *BMC Public Health*, 7(63).

Nattrass, N. 2008. “Gender and Access to Antiretrovirals in South Africa.” *Feminist Economics*, 14(4), 19-36.

Nattrass, N. 2007. *Mortal Combat: AIDS Denialism and the Fight for Antiretrovirals in South Africa*, University of KwaZulu-Natal Press, Durban.

Palella, F.J., K. Delaney, A. Moorman, M. Loveless, J. Fuhrer, G. Satten, D. Aschman and S. Holmberg. 1998. “Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection.” *New England Journal of Medicine*, 338: 853-860.

Shisana, O, T Rehle, LC Simbayi, W Parker, K Zuma, A Bhana, C Connolly, S Jooste, and V Pillay. 2005. *South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey.* Human Sciences Research Council Press, Cape Town.

Strauss, John and Duncan Thomas. 1998. "Health, Nutrition, and Economic Development," *Journal of Economic Literature*, 36(2): 766-817.

Thirumurthy, Harsha, Joshua Graff Zivin and Markus Goldstein. 2008. "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya." *Journal of Human Resources*, 43:511-552.

Thomas, Duncan; Frankenberg, Elizabeth; Friedman, Jed; Habicht, Jean-Pierre; Hakimi, Mohammed; Ingwersen, Nicholas; Jaswadi; Jones, Nathan; McKelvey, Christopher; Pelto, Gretel, et al. 2006. "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence," California Center for Population Research. On-Line Working Paper Series. Los Angeles, CA.

Thornton, Rebecca L. 2008. "The Demand for, and Impact of, Learning HIV Status," *American Economic Review*, 98 (5): 1829-63.

UNAIDS. 2008. *Report on the global AIDS epidemic.* Joint United Nations Programme on HIV/AIDS (UNAIDS).

Venter, Francois. 2006. AIDS Analysis Africa Online. Jan/Feb 2006. Downloaded at http://www.redribbon.co.za/aaao/.

Wools-Kaloustian, Kara, Silvester Kimaiyo, Lameck Diero, Abraham Siika, John Sidle, Constantin T. Yiannoutsos, Beverly Musick, Robert Einterz, Kenneth H. Fife and William M. Tierney. 2006. "Viability and effectiveness of large-scale HIV treatment initiatives in sub-Saharan Africa: experience from western Kenya." *AIDS*, 20:41-48.

World Health Organization. 2005. "Interim WHO Clinical Staging of HIV/AIDS and HIV/AIDS Case Definitions for Surveillance."

2.7 Tables and Figures

Table 2.1: Geographic units in South Africa Census

Unit	**Number in SA**	**Number in sample**	**Min** size	**Max** sizs	**Mean**	**Median**
PSU	80,787	2,996	0.0075	1,460	20	6.4
Main Place	2,674	1,376	.0075	22,023	264	15.6
Municipality	253	253	92	11,351	1,860	732
District	53	53	635	48,939	8,118	5,910
Province	9	9	6,568	139,703	54,884	47,842

Table shows areas measured in square miles.

Table 2.2: Type of ARV clinic by province for clinics open by September 2007

Province	**% of Popn.**	**Health Center**	**Hospital**	**Total**
Eastern Cape	14	7	33	40
Free State	6	25	9	34
Gauteng	20	22	27	50
Kwa-Zulu Natal	21	13	57	70
Limpopo	12	1	35	36
Mpumalanga	7	0	19	19
Northern Cape	2	6	5	11
North West	8	2	21	23
Western Cape	10	27	30	58
Total	100	103	236	339
Percent		30.4	69.6	100

Column 2 shows percent of South African population residing in each province.

Table 2.3: Comparing characteristics at baseline and over time for Black men

Variable	Nearer baseline	Farther baseline	t-Stat on diff.	Nearer change	Farther change	F-stat (wave)	P-val
Age	33.06	32.86	1.18	-0.12	-0.12	0.24	0.96
Yrs. of primary educ.	5.75	6.38	-7.88	0.30	0.02	1.30	0.25
Yrs. of secondary educ.	2.19	2.93	-8.22	0.30	0.17	0.37	0.90
Completed Matric (H.S.)	0.25	0.38	-6.00	0.03	0.02	1.17	0.32
Some post-Matric educ.	0.02	0.03	-2.00	-0.01	0.00	1.09	0.37
Never held a job	0.26	0.22	2.00	-0.03	-0.06	1.25	0.28
Ever married	0.45	0.46	-0.50	-0.06	-0.02	0.92	0.48
Spouse resides in hhold	0.33	0.37	-2.00	-0.03	-0.03	0.90	0.49
Number of adults in hhold	3.17	2.81	4.00	-0.04	-0.01	1.22	0.29
Number of kids in hhold	1.51	1.09	7.17	-0.02	-0.05	0.70	0.65
Senior in hhold	0.23	0.11	12.00	0.01	0.00	0.43	0.86
Lived here 6 months ago	0.97	0.96	1.00	-0.01	-0.04	0.71	0.64

Notes: Nearer defined as closer than median (7.5 miles) from nearest clinic in first wave of sample. Change defined as change in value of variable between first wave (Sept 2004) and last wave (Sept 2007) of sample. F-stat for joint test of the null of zero coefficients on the full set of wave dummies interacted with treatment dummy. Distance calculated from centroid of neighborhood to clinic location. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Standard errors clustered at the main place level.

Table 2.4: Comparing characteristics at baseline and over time for Black women

Variable	Nearer baseline	Farther baseline	t-Stat on diff.	Nearer change	Farther change	F-stat (wave)	P-val
Age	33.42	32.97	3.00	-0.02	-0.05	0.22	0.97
Yrs. of primary educ.	5.77	6.37	-10.00	0.30	0.12	0.83	0.54
Yrs. of secondary educ.	2.18	2.88	-10.00	0.30	0.29	0.32	0.93
Completed Matric (H.S.)	0.24	0.33	-10.00	0.03	0.06	0.51	0.80
Some post-Matric educ.	0.01	0.03	-1.00	0.00	0.00	1.91	0.08
Never held a job	0.43	0.39	1.50	-0.05	-0.10	0.94	0.47
Ever married	0.56	0.54	1.00	-0.08	-0.04	1.81	0.09
Spouse resides in hhold	0.33	0.42	-4.50	-0.05	-0.03	1.87	0.08
Number of adults in hhold	3.33	3.08	3.57	-0.09	0.05	1.60	0.14
Number of kids in hhold	2.44	1.88	8.14	-0.05	-0.09	0.55	0.77
Senior in hhold	0.22	0.13	9.00	0.02	0.01	0.22	0.97
Lived here 6 months ago	0.98	0.98	0.00	-0.02	-0.03	2.83	0.01

Notes: Nearer defined as closer than median (7.5 miles) from nearest clinic in first wave of sample. Change defined as change in value of variable between first wave (Sept 2004) and last wave (Sept 2007) of sample. F-stat for joint test of the null of zero coefficients on the full set of wave dummies interacted with treatment dummy. Distance calculated from centroid of neighborhood to clinic location. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Standard errors clustered at the main place level.

Table 2.5: The effect of the distance to the nearest clinic on the likelihood of labor force participation of Black African men

Variable	(1)	(2)	(3)	(4)
Distance to nearest clinic $<$ 3 miles	0.100***	0.035***	0.033***	0.007
	(0.015)	(0.011)	(0.011)	(0.013)
Distance to nearest clinic 3-15 miles	0.047***	0.008	0.009	0.020*
	(0.012)	(0.010)	(0.009)	(0.012)
Completed the Matric (High school)		0.024***	0.025***	0.024***
		(0.009)	(0.009)	(0.008)
Completed some post-Matric education		0.041***	0.049***	0.073***
		(0.014)	(0.014)	(0.011)
Spouse resides in household		0.032***	0.033***	0.051***
		(0.009)	(0.007)	(0.008)
Number of adults in household		-0.031***	-0.030***	-0.025***
		(0.002)	(0.002)	(0.002)
District*wave interactions	No	No	Yes	Yes
PSU Fixed Effects	No	No	No	Yes
Number of obs.	61,593	61,593	61,593	61,593
R^2	0.01	0.11	0.15	0.25

Notes: Standard errors in parentheses are clustered at the main place level. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Omitted category is individuals who are 15-75 miles from the nearest clinic. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 2.6: The effect of the distance to the nearest clinic on the likelihood of labor force participation and employment for Black African men

Dependent variable:	Participation		Employment	
Variable	(1)	(2)	(3)	(4)
Distance to nearest clinic < 3 miles	0.033***	0.007	0.012	0.018
	(0.011)	(0.013)	(0.012)	(0.014)
Distance to nearest clinic 3-15 miles	0.009	0.020*	0.009	0.033**
	(0.009)	(0.012)	(0.011)	(0.011)
Completed the Matric (High school)	0.025***	0.024***	0.060***	0.047***
	(0.009)	(0.008)	(0.010)	(0.010)
Completed some post-Matric education	0.049***	0.073***	0.125***	0.115***
	(0.014)	(0.011)	(0.018)	(0.020)
Spouse resides in household	0.033***	0.051***	0.057***	0.090***
	(0.007)	(0.008)	(0.010)	(0.011)
Number of adults in household	-0.030***	-0.025***	-0.057***	-0.049***
	(0.002)	(0.002)	(0.002)	(0.002)
PSU Fixed Effects	No	Yes	No	Yes
Number of obs.	61,593	61,593	61,593	61,593
R^2	0.15	0.25	0.23	0.34

Notes: Standard errors in parentheses are clustered at the main place level. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Omitted category is individuals who are 15-75 miles from the nearest clinic. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 2.7: The effect of the distance to the nearest clinic on the likelihood of labor force participation and employment for Black African women

Dependent variable:	Participation		Employment	
Variable	(1)	(2)	(3)	(4)
Distance to nearest clinic < 3 miles	0.074***	0.002	0.034**	-0.000
	(0.013)	(0.017)	(0.016)	(0.018)
Distance to nearest clinic 3-15 miles	0.017*	-0.002	0.009	-0.011
	(0.010)	(0.013)	(0.010)	(0.011)
Completed the Matric (High school)	0.066***	0.057***	0.108***	0.082***
	(0.009)	(0.008)	(0.011)	(0.010)
Completed some post-Matric education	0.165***	0.127***	0.310***	0.207***
	(0.018)	(0.021)	(0.020)	(0.028)
Spouse resides in household	-0.023***	-0.058***	-0.038***	-0.067***
	(0.009)	(0.008)	(0.010)	(0.010)
Number of adults in household	-0.011***	-0.010***	-0.030***	-0.025***
	(0.002)	(0.002)	(0.002)	(0.002)
PSU Fixed Effects	No	Yes	No	Yes
Number of obs.	76,413	76,413	76,413	76,413
R^2	0.11	0.20	0.13	0.24

Notes: Standard errors in parentheses are clustered at the main place level. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Omitted category is individuals who are 15-75 miles from the nearest clinic. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 2.8: The effect of the fraction of the neighborhood receiving treatment on the likelihood of labor force participation of Black African men

Variable	(1)	(2)	(3)	(4)
Fraction of neighborhood treated	0.002*	0.003***	0.000	-0.005**
	(0.001)	(0.001)	(0.001)	(0.002)
Completed the Matric (High school)		0.024***	0.025***	0.024***
		(0.009)	(0.009)	(0.008)
Completed some post-Matric education		0.042***	0.049***	0.073***
		(0.014)	(0.014)	(0.011)
Spouse resides in household		0.034***	0.034***	0.051***
		(0.009)	(0.007)	(0.008)
Number of adults in household		-0.031***	-0.030***	-0.025***
		(0.002)	(0.002)	(0.002)
District*wave interactions	No	No	Yes	Yes
PSU Fixed Effects	No	No	No	Yes
Number of obs.	61,590	61,590	61,590	61,590
R^2	0.00	0.11	0.15	0.25

Notes: Standard errors in parentheses are clustered at the main place level. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Omitted category is individuals who are 15-75 miles from the nearest clinic. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 2.9: The effect of the fraction of the neighborhood receiving treatment on the likelihood of labor force participation and employment of Black African men

Dependent variable:	Participation		Employment	
Variable	(1)	(2)	(3)	(4)
Fraction of neighborhood treated	0.000	-0.005**	-0.001	0.004**
	(0.001)	(0.002)	(0.001)	(0.002)
Completed the Matric (High school)	0.025***	0.024***	0.060***	0.047***
	(0.009)	(0.008)	(0.010)	(0.010)
Completed some post-Matric education	0.049**	0.073**	0.125***	0.115***
	(0.014)	(0.011)	(0.018)	(0.020)
Spouse resides in household	0.034**	0.051**	0.057***	0.090***
	(0.007)	(0.008)	(0.010)	(0.011)
Number of adults in household	-0.030**	-0.025**	-0.057***	-0.049***
	(0.002)	(0.002)	(0.002)	(0.002)
PSU Fixed Effects	No	Yes	No	Yes
Number of obs.	61,590	61,590	61,590	61,590
R^2	0.15	0.25	0.23	0.34

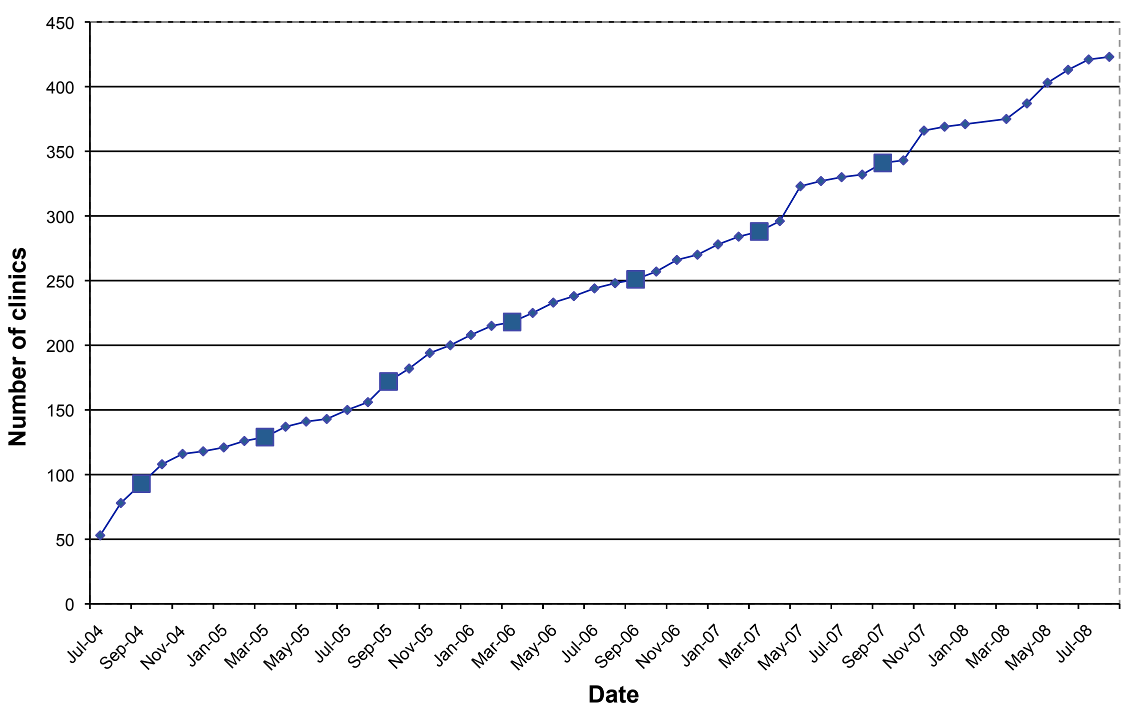
Notes: Standard errors in parentheses are clustered at the main place level. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Omitted category is individuals who are 15-75 miles from the nearest clinic. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 2.10: The effect of the fraction of the neighborhood receiving treatment on the likelihood of labor force participation and employment of Black African Women

Dependent variable:	Participation		Employment	
Variable	(1)	(2)	(3)	(4)
Fraction of neighborhood treated	0.003***	-0.003	0.003***	0.007***
	(0.001)	(0.003)	(0.001)	(0.002)
Completed the Matric (High school)	0.066***	0.057***	0.108***	0.082***
	(0.009)	(0.008)	(0.011)	(0.010)
Completed some post-Matric education	0.169***	0.127***	0.312***	0.207***
	(0.019)	(0.021)	(0.020)	(0.028)
Spouse resides in household	-0.022***	-0.058***	-0.038***	-0.067***
	(0.009)	(0.008)	(0.010)	(0.010)
Number of adults in household	-0.011***	-0.010***	-0.030***	-0.025***
	(0.002)	(0.002)	(0.002)	(0.002)
PSU Fixed Effects	No	Yes	No	Yes
Number of obs.	76,405	76,405	76,405	76,405
R^2	0.11	0.20	0.13	0.24

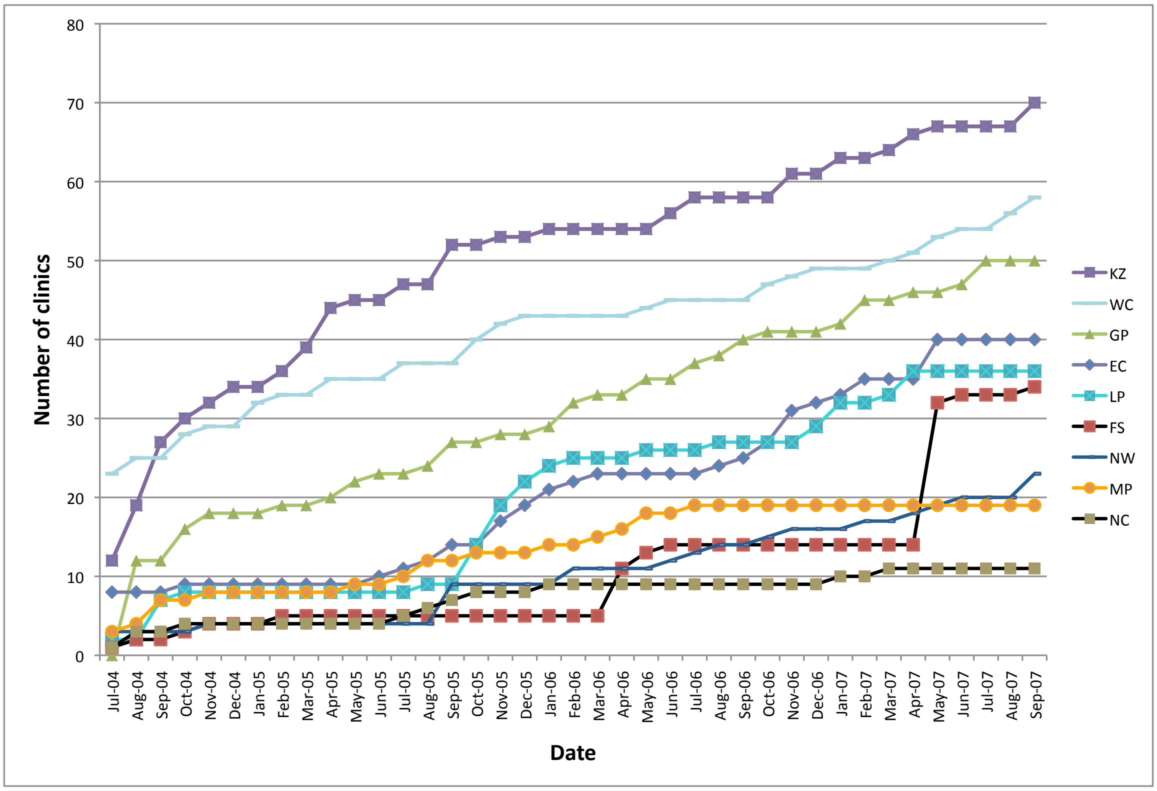
Notes: Standard errors in parentheses are clustered at the main place level. Sample includes individuals aged 25-44 who live in households containing a 25-44 year old (including self). Excludes health workers and neighborhoods that remain more than 75 miles from the nearest clinic in September 2007. Omitted category is individuals who are 15-75 miles from the nearest clinic. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Figure 2.1: Cumulative number of open ARV clinics over time



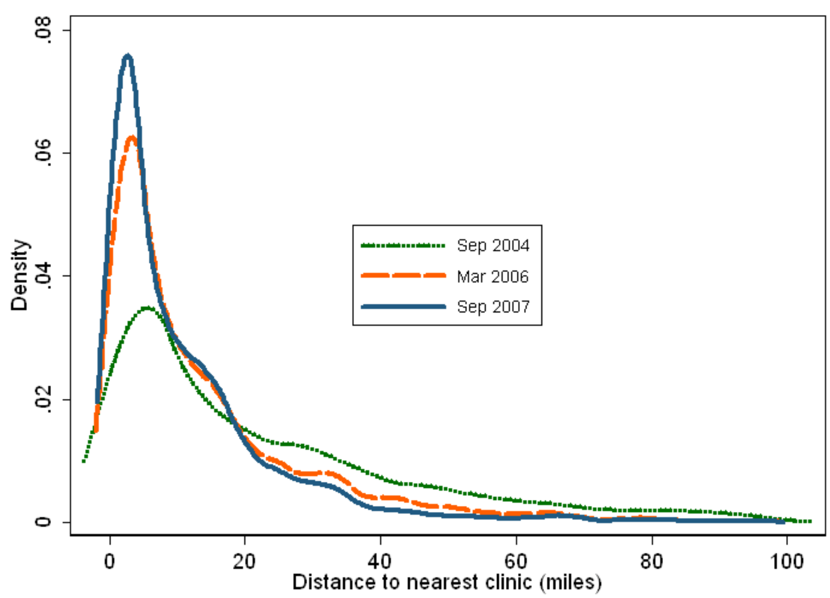
Squares represent dates of Labour Force Survey waves used in analysis.

Figure 2.2: Cumulative number of open AIDS treatment clinics over time by province



KZ = Kwa-Zulu Natal, WC = Western Cape, GP = Gauteng, EC = Eastern Cape, LP = Limpopo, FS = Free State, NW = Northwest Province, MP = Mpumalanga, NC = Northern Cape.

Figure 2.3: Density of distance to the nearest clinic over time



Author's calculations.

Figure 2.4: Cumulative number of AIDS patients enrolled in ARV treatment over time by sector

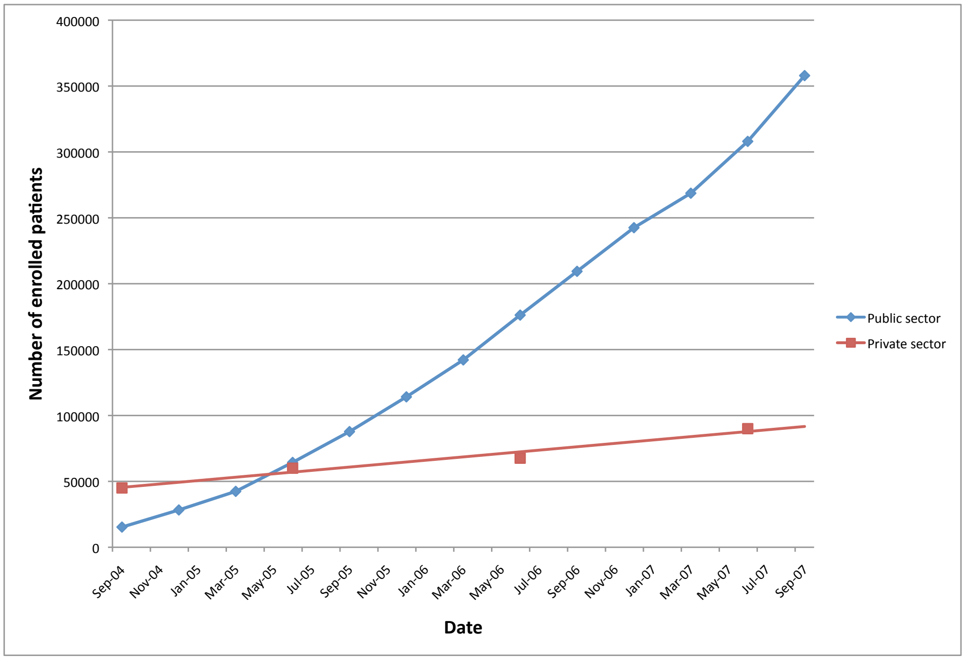
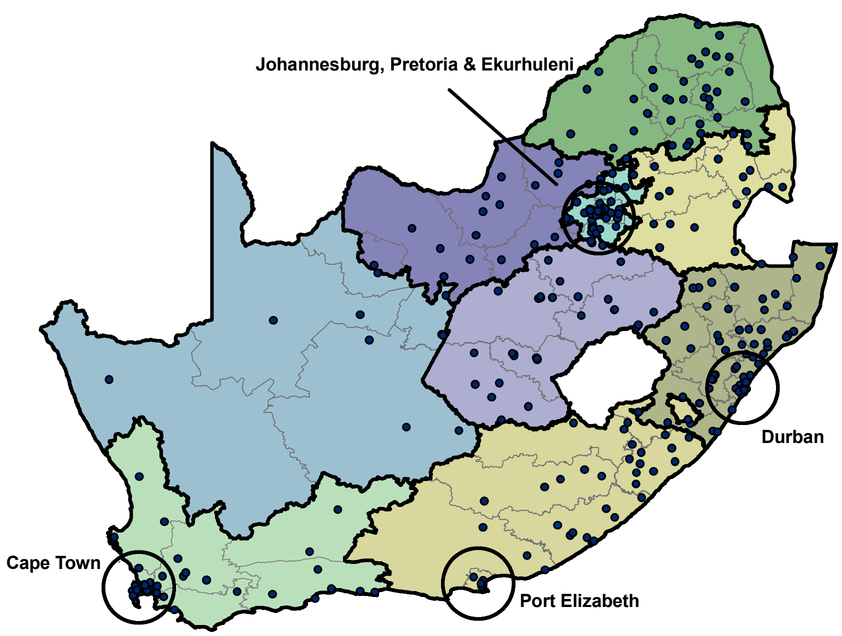


Figure 2.5: Geographic distribution of ARV clinics in September 2007



Map shows province boundaries (in black), district boundaries (in gray) and the six metropolitan areas (labeled).

CHAPTER III

HIV Status and Labor Market Participation in South Africa

3.1 Introduction

This paper employs an especially rich South African data set to estimate the causal impact of HIV status on an individual's labor market participation. As such, the paper contributes to a better understanding of the economic impact of HIV/AIDS.

The existence of an impact of HIV on labor market participation is hotly debated in South Africa. Some argue that HIV is a severe constraint on current economic growth because those who are too ill to work represent lost economic output. Others argue that because unemployment is so high (in the 30 to 40 percent range according to the broad definition[1]), HIV has minimal economic impact today. Proponents of this view note that the sex-age cohorts with the highest unemployment are essentially the same as those with the highest HIV prevalence rates.[2] Our estimates speak directly to this issue. This paper provides the first nationally representative estimates of the impact of HIV on labor market status, certainly for southern Africa but probably also more broadly for non-rich countries.

[1] The broad definition of unemployment includes individuals who desired employment but had no job search activity within the past month (i.e. discouraged workers).

[2] Another argument that HIV has a minimal impact on employment is that with anti-retroviral (ARV) therapy, HIV+ individuals are not likely to be too sick to work, and hence HIV has only a minimal impact on labor market participation. As discussed below, this argument, while germane to some countries, is probably not highly relevant to South Africa in 2005.

There are at least two reasons for the paucity or outright lack of evidence– stringent data requirements and econometric challenges.

Obtaining estimates of the impact of HIV on labor market participation clearly requires data on both labor market participation (which is generally available) and on an individual's HIV status (which is collected fairly rarely).[3] Our econometric methods also require extensive data on correlates of an individual's HIV status.

Even with our rich data set, econometric challenges arise due to econometric endogeneity. In our context, endogeneity arises because an individual's labor market status might affect their HIV status. For example, migrating for employment opportunities may put individuals at greater risk for HIV (reverse causality bias). Alternatively, there may be factors that affect both HIV status and labor market status, such as an individual's age or their willingness to invest in the future (selection bias, omitted variable bias).[4]

Two of the standard approaches to addressing these issues are infeasible in this instance. The traditional solution is to apply an instrumental variables approach. For the application at hand, this would require an instrument that is correlated with the individual-level HIV status and is orthogonal to shocks to labor market status (conditional on observed covariates. This is a tall order to fill. One class of correlates with HIV status – socioeconomic variables such as income, education, sex, age and even household size – are likely correlated with labor market status, while another

[3]There are other data sets available for which the techniques employed in this paper are potentially applicable. These include some of the Demographic Health Surveys (DHS) and the Botswana AIDS Impact Survey (see Levinsohn and McCrary (2008))

[4]These challenges are discussed in detail in Strauss and Thomas (1998).

class of correlates – health-related variables such as sexual practices and knowledge about HIV – are themselves likely to be correlated with the socioeconomic covariates and, by association, with labor market status.[5] (A closely related approach uses a control function with an exclusion restriction, and the same concerns that preclude good instruments similarly preclude a convincing exclusion restriction. See Heckman and Navarro-Lozano (2004).)

The other approach to addressing sample selection which has gained currency in development economics is to conduct interventions such that the bias vanishes by design.[6] In a study of HIV status and labor market participation, this would involve randomly assigning HIV status to individuals and comparing labor market status across the two groups. This is one context for which the experimental approach is impossible.

We turn to methods based on the propensity score to address the issues of selection and reverse causality. We investigate the impact of HIV status on labor market participation using two methods– propensity score reweighting and control functions.[7] While both have been used to address the endogeneity issue in the literature, there have been few direct comparisons of multiple estimators for the same problem. We use both methods as a marginal methodological contribution and, more importantly, as a robustness check.

---

[5] With country-level data, the fraction of the male population that has been circumcised has been used as an instrument (see Werker et al. (2006)). At the individual-level, this instrument is problematic. The protective effect may not transfer to women (Wawer et al. (2008)) and there is often virtually no variation within large categories of the male population like religion, race and ethnic/tribal affiliation (see McKelvey (2007)).

[6] For example, Thomas et al. (2006) use experimental methods to obtain a causal estimate of the effect of improved health (in this case, from increased iron intake) on labor market outcomes.

[7] These methods are drawn mostly from the program evaluation literature, but they have also been used in other fields of economics. For example, a similar control function approach is used in Olley and Pakes (1996) for solving the endogeneity problem in the context of production functions. Levinsohn and McCrary (2008) use the propensity score re-weighting technique to address a missing data problem in the context of HIV prevalence.

Propensity score reweighting and the control function approach both rely on the assumption that sample selection into being HIV+ is random, conditional on attributes that are observable to the econometrician (selection on observables). Given the richness of the data set we use, this assumption is weaker than it might be under other circumstances.[8] The data set contains detailed information on sexual practices and knowledge of HIV transmission in addition to household characteristics and other factors that might influence HIV status. We examine the validity of the assumption of selection on observables in section 5 by testing whether a parsimonious specification succeeds in balancing variables not used in the estimation (balancing tests) and experimenting with multiple specifications to confirm that the inclusion of additional controls does not change the results (robustness checks).

In the next section, we briefly discuss the related literature. Section 3 discusses the data that are employed. In section 4, the two approaches to estimation are presented. Section 5 presents results, while section 6 provides both caveats and conclusions.

3.2 Related Literature

While understanding the relationship between HIV status and labor market participation is important from multiple perspectives, there are few studies that present evidence on the impact of HIV and fewer still that address issues of endogeneity.

A number of studies develop macroeconomic simulation models to examine the

[8] The survey questionnaire included over 175 questions, many with multiple sub-sections, yielding over four hundred individual covariates.

effect of HIV/AIDS on economic growth, however these models may be extremely sensitive to assumptions about life-expectancy for HIV+ individuals (Cuddington and Hancock (1994), Kambou et al. (1992), and Arndt and Lewis (2000)). Instead of assuming an HIV/AIDS mortality rate, Bloom and Mahal (1997) estimate the impact of AIDS on growth by exploiting cross-country variation in HIV prevalence rates. They find that HIV has an insignificant effect on per capita GDP growth for the sample of 51 countries.

The HIV pandemic not only affects current economic growth, but can have lasting effects on growth rates into the future. Kalemli-Ozcan (2006) and Bell et al. (2003) model the impact of current HIV prevalence rates on future economic growth. Because HIV+ parents are likely to die before reaching old age, they may invest less in their children's human capital acquisition, lowering the stock of human capital and contributing to lower growth rates in the future. On the other hand, if the AIDS epidemic causes a reduction in fertility that dominates this human capital effect, higher future living standards may result (Young (2005)).

Most papers that examine the impact of HIV on economic growth look exclusively at the effect of AIDS mortality, overlooking the effects of the illness on employment and productivity. There are some exceptions. Murray et al. (2005) find that the rate of minor work-related injuries was 30 percent higher for HIV+ miners in South Africa than for HIV- miners. The correlation was observable within one year of sero-conversion which suggests that the acute initial infection or the psychological shock has an effect long before AIDS symptoms appear.[9]

[9]This is likely an underestimate of the impact of HIV because miners who are most affected are more likely to take on easier tasks at work or leave employment altogether.

Habyarimana et al. (2007) use ARV therapy inception date as an instrument for health status to examine the effect of health on productivity for mine workers in Botswana. They document an inverse-V-shaped pattern of absenteeism for HIV+ workers in the two years around the inception of anti-retroviral (ARV) therapy. Workers who subsequently enrolled in ARV therapy missed about five times as many days of work as non-enrolled workers in the year prior to ARV therapy inception, but absenteeism rates returned to pre-peak levels after a year of therapy. Non-enrolled workers appear to be a valid control group because the two groups had similar levels of absenteeism from 5 years to 1 year prior to therapy inception. These findings suggest that health can have large effects on employment outcomes, and that ARV therapy is effective in reducing the disparity in productivity between HIV+ and HIV- workers.[10] Fox et al. (2004) examine differences in on-the-job productivity between workers who subsequently died of AIDS and other workers on tea plantations in Kenya. They find that in the last year before death, AIDS victims are less productive (in this case, measured by the quantity of tea leaves picked), are more likely to be reassigned to less strenuous but less lucrative tasks and are more often absent from work.

One way to estimate the causal effect of HIV status on employment is to use a plausibly exogenous instrument for HIV status. Variation in circumcision has been used to instrument for HIV status because circumcision has been found to be associated with a reduced risk of HIV both in regression analyses (Weiss et al. (2000)) and in randomized controlled trials (Auvert et al (2005)). Werker et al. (2006) find that

[10] Thirumurthy et al (2008) also found that in western Kenya, labor supply increased within six months of initiating ARV therapy.

HIV/AIDS did not have a measurable effect on economic growth, savings, or fertility behavior in African countries but that there was weak evidence that HIV/AIDS reduced youth literacy, and increased malnutrition. McKelvey (2007) finds that across nine African countries and Haiti, HIV+ individuals are significantly less likely to have been employed or to have earned enough to contribute more than half of household expenditures. However, specification checks using two populations that would not be expected to benefit from circumcision (men under 20 years old and those who have never had sex) suggest that unobservable differences that vary with circumcision may be driving the results.

Transitory economic shocks can have short-term effects on an individual's propensity to engage in risky behavior, but potentially long-term effects on health. This is one avenue through which employment outcomes may impact HIV status, and one reason why reverse causality is a salient issue. Women who are economically vulnerable may become active with multiple sexual partners (Dinkelman et al. (2007)), or turn to sex-for-gift exchanges to smooth consumption (Dunkle et al. (2004); LeClerc-Madlala (2002)). Migration is another potential response to transitory shocks that can put both men and women at greater risk for HIV (Zuma et al. (2005)).

3.3 Data

Our data come from the nationally-representative South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey (SABSSM II) conducted in 2005 by the Human Sciences Research Council (HSRC), the Centre for AIDS Development, Research and Evaluation (CADRE) and the Medical Research Council.[11] The survey asked adult respondents questions about demographics, knowledge of HIV, sexual history, knowledge of voluntary counseling and testing (VCT) services, health, mental health, and drug and alcohol use. It also included a household module that asked for basic demographic data for all household members in addition to questions about household infrastructure and participation in government programs. While there were no questions directly addressing income or expenditure, the individual survey did query labor market participation.[12]

The sample consisted of 23,275 individuals in 10,584 households.[13] Not everyone in a household was sampled. Field workers randomly selected at most one person from each of three age groups (2-14 years, 15-24 years and 25 and up) to be interviewed. While this sampling strategy is an efficient way to obtain a measure of HIV prevalence, it is problematic for any analysis of household dynamics. For example, we are unable to directly measure the impact of having an HIV+ spouse on adult labor supply, since only one household member above the age of 25 was surveyed.

The survey also included an opt-out HIV test for respondents age 2 and older. The response rate for testing was 65.4 percent overall, and 73.3 percent in the adult sample (age 15 and older) used for analysis (see Table 3.1). HIV incidence data was collected from all HIV+ specimens using an enzyme immunoassay that measures the

[11] A detailed description of the survey methodology is found in Shisana et al. (2005), available online at: http://www.hsrcpress.ac.za/product.php?productid=2134&cat=0&page=1

[12] Respondents were asked to classify their "present employment situation" as one of the following thirteen categories: homemaker not looking for work, homemaker looking for work, unemployed not looking for work, unemployed looking for work, informal sector not looking for permanent work, old age pensioner, sick/disabled and unable to work, student/pupil/learner, self-employed full time (40 or more hours per week), self-employed part time (less than 40 hours per week), employed part time (if none of the above), employed full time or other.

[13] When matching individuals to households based on household identification numbers, approximately 14 percent of the individual observations (3413) could not be matched to household data. Household variables were imputed to zero for these unmatched observations. A specification check suggests that the data were consistent with a pattern of "missing at random"; a dummy variable for being unmatched was not significant in any specification. As with all variables that are imputed, an indicator variable was created that took a value of 1 if the value was imputed and 0 otherwise. These indicator variables are included in the empirical specifications. Individual survey records were matched to HIV incidence and viral load data using the barcode number that identified each specimen.

ratio of HIV antibodies to other antibodies to determine the elapsed time since HIV infection.

Table 3.2 and Figure 3.1 present pertinent descriptive statistics. Table 3.2 presents HIV prevalence rates by race. Table 3.2 indicates that while HIV impacts all races, Africans have, by far, the highest prevalence rates. In our sample, 17.34 percent of Africans are HIV+ while the population group with the next highest prevalence rate, Coloured, has a rate of 2.73 percent.[14] [15]

Figure 3.1 presents the age profile of HIV prevalence in South Africa– a profile broadly comparable to those of other African countries. HIV prevalence peaks between age 25-30 for women and between 30-40 for men. The age profile for women is about 5 percentage points lower than UNAIDS ante-natal clinic statistics, which may be due to differences in the sampling frame (UNAIDS (2005)).

Table 3 reports employment status by HIV status separately for men and for women. For our analysis, individuals were assigned to a labor market status using the broad definition of unemployment, which includes discouraged workers. Respondents were classified as unemployed if they report that they are unemployed (either looking for work or not) or if they are a housewife or homemaker who is looking for work. They were classified as not economically active if they reported being a student, an old age pensioner, too sick or disabled to work, or a housewife or home-

[14]These rates apply to our sample which only includes individuals aged 15 and over. Hence they differ from the national prevalence rates reported in Shisana et al. (2005).

[15] The prevalence rates in Table 3.2 do not correct for sample selection due to non-random opt-out of the HIV test. We have replicated our analysis on the impact of HIV on labor market participation with weights that do correct for non-random opt-out (as computed by HSRC) and the results are virtually identical. This suggests that selection into testing is independent of the effect of HIV on employment.

maker who is *not* looking for work. Otherwise, they were classified as employed.[16] Prevalence rates are higher for the unemployed relative to the employed, and this is especially pronounced for women. Almost 21 percent of unemployed women are HIV+. Students and other not economically active individuals (mostly old age pensioners) have lower than average prevalence rates, and these individuals are not included in our analysis. A similar proportion of each labor market group refused the HIV test.

3.4 Methodologies

3.4.1 Overview

Generating a plausible counterfactual is the core challenge to identifying a causal effect of HIV. For the situation at hand, this approach entails creating a counterfactual in which one could compare individuals who were virtually identical except for their HIV status and then compare differences in labor market status. The following subsections present two methods, each based on the propensity score, to generate unbiased estimates of the effect of being HIV+ on labor market status.

Our "treatment" (as the public health and program evaluation literatures refer to it) is HIV status, denoted by D, where $D = 1$ is HIV+ and $D = 0$ is HIV-. If HIV status were independent of the untreated value of our outcome of interest, labor market status (denoted Y_0), then there is no sample selection problem and we can use a simple "naive" estimator, a simple difference in means, to estimate the treatment effect. However, HIV status is not randomly assigned and is in all likelihood related to our outcomes of interest. We make the conditional independence assumption

[16] Those who responded being in the "other" category are grouped with the employed to generate conservative estimates.

(CIA) that the untreated value of labor market status is independent of HIV status conditional on a vector of covariates (denoted X) :

$$Y_0 \perp D|X. \tag{3.1}$$

3.4.2 The Propensity Score

In this context, the propensity score, denoted $P(HIV+)|X \equiv p(X)$, is the probability that an individual is HIV+ conditional on a vector of observable exogenous covariates (X). These covariates must be exogenous to HIV status in that they cannot be affected by HIV status (i.e. it would be inappropriate to use sexual behavior that may be changed if HIV status is known).[17] Rosenbaum and Rubin (1983) show that if the CIA is satisfied by conditioning on the vector X, it is satisfied by conditioning on the propensity score (i.e. the propensity score is a sufficient statistic for the X vector in the CIA). Hence, the propensity score makes the problem of finding a comparable control group tractable by reducing the dimensionality of the comparison, while still satisfying the CIA.

3.4.3 Propensity Score Reweighting

Propensity score reweighting uses the propensity score to create a counterfactual distribution of X in the HIV- (control) population so as to match the distribution of X in the HIV+ population. Essentially, HIV- observations with X characteristics that are most like HIV+ observations (i.e. that have a high $\hat{p}(X)$) receive the most weight, whereas HIV- observations that are very different from the HIV+ population

[17] It is important to note that these are not the same variables we would use as instruments in an instrumental variables (IV) regression. In that case we would want Zs correlated with HIV status but not with labor market status; in our case these Zs should not be included in the propensity score regression. Because actual HIV status is known by the econometrician and can therefore be controlled for in the regression, the covariates in X do not need to be good predictors of HIV status. In fact, selecting the attributes for X depending on predictive power can actually increase bias. See Heckman and Navarro-Lozano (2004).

receive less weight.

More formally, Dehejia and Wahba (1997) and DiNardo et al. (1996) show that

$$\Delta^{ATET} = \frac{1}{N^T}\Sigma_{i=1}^{N^T+N^C}\left(D_iY_i - (1-D_i)Y_i\frac{p(X_i)}{1-p(X_i)}\right) \tag{3.2}$$

is a consistent estimator for the average treatment effect on the treated (ATET), provided the CIA and the common support condition hold. In the calculation, HIV+ observations receive a weight of 1 (the first term in parentheses) and HIV- observations are weighted with $\omega_i = \frac{\hat{p}(X_i)}{1-\hat{p}(X_i)}$ (the second term in parentheses). This estimator does not impose a functional form on the relationship between HIV status and the outcome of interest (Dehejia and Wahba (1999)).

This approach requires overlap in the support of the propensity scores for the HIV+ and HIV- groups (the common support condition). In the Results section, we present density plots to verify the validity of this assumption.[18] Busso et al. (2009) show that reweighting outperforms propensity score matching in settings likely to be encountered in empirical work.

3.4.4 Control Functions

The control function approach is an alternative econometric strategy for addressing selection bias. The reason one obtains biased estimates from a simple regression of labor market status on HIV status is that the disturbance term in such a regression is correlated with HIV status (the independent variable). The essence of the control function approach is to control for the portion of the disturbance term that is corre-

[18] Frölich (2004) demonstrates that an estimated propensity score close to 1 can cause problems for estimating the ATET. Our values of the estimated propensity scores are not large enough for this to be a concern.

lated with HIV status. Once the portion of the disturbance term that is responsible for the correlation is expunged, the new error term is uncorrelated with HIV status, and the regression yields unbiased estimates of the impact of HIV on employment status.[19] As with propensity score reweighting, we maintain the assumption of selection on observables.[20]

The data generating process is given by:

$$Y = \Lambda\Big(\beta_0 + \gamma D + f(X) + \epsilon\Big) \tag{3.3}$$

where $f(X)$ is a function of observables X. Under the assumption of selection on observables, conditioning on $f(X)$ results in a disturbance term, ϵ, that is independent of D (HIV status) and hence, the estimate of the parameter of interest, γ is unbiased. In practice, a polynomial in the estimated propensity score as well as linear terms in X are used to flexibly model $f(X)$. We estimate:

$$Y = \Lambda\Big(\beta_0 + \gamma D + X'\phi + \sum_{i=1}^{k} \beta_i \hat{p}(X)^i + \epsilon\Big) \tag{3.4}$$

separately for HIV- and HIV+ groups and obtain predicted values, $\hat{Y}_0$ and $\hat{Y}_1$ respectively, from each regression. We calculate the ATET by averaging the difference in predicted values across HIV+ observations.

---

[19]The control function approach was developed in Heckman and Robb (1985) and has been used to estimate the impact of training on earnings (Heckman and Hotz (1989)), the returns to education (Card (1999)) and the capitalization of pollution into housing values (Chay and Greenstone (2005)) among other applications.

[20]A similar control function method can be used if there is selection on *un*observables (see Heckman and Navarro-Lozano (2004)). However, allowing for selection on unobservables requires an exclusion restriction– a variable that is correlated with HIV status but uncorrelated with labor market status.

3.5 Results

Estimating the impact of HIV status on labor market participation is a three-step process. The first step is to estimate the propensity score. The second step is to empirically examine the validity of the CIA and the common support condition to ensure that the appropriate observable variables are included in the propensity score regression. If the underlying assumptions hold, then one can proceed to the third step, estimating the impact of HIV on labor market participation using propensity score reweighting and a control function approach.

3.5.1 The Propensity Score

There are competing philosophies behind what constitutes a properly specified propensity score regression. One approach is to adopt a relatively parsimonious specification, albeit one still rich enough to plausibly satisfy the CIA, while another approach is to include most all plausible regressors. We adopt the former, but experiment with the latter in sensitivity analyses.

The plausibility of the selection on observables assumption clearly rests on having data that can, in our context, account for selection into HIV status. The SABSSM II data set has several *hundred* variables for most respondents. In addition to the usual demographic information, the survey also collected extensive information on sexual practices and knowledge about HIV transmission, which is exactly the type of information needed to account for selection into HIV status. Information on both sexual practices *and* knowledge about HIV transmission is important for predicting HIV status. A respondent who has multiple partners but is well informed on how HIV is transmitted and practices safe sex will have a low probability of being HIV+. Similarly, a respondent who knows very little about how HIV is transmitted but is abstinent will have a low likelihood of being HIV+. Conditional on comprehensive information about an individual's sexual practices and knowledge regarding HIV transmission, unobservables such as attitudes towards risk or moral beliefs may have little explanatory power. Hence, the selection on observables assumption seems especially appropriate given the specifics of our data. The selection on observables assumption is testable, and we investigate the reasonableness (or not) of this assumption by conducting balancing tests on covariates in the propensity score specification as well as plausible covariates that were not included in the specification (see Section 3.5.2).

We estimate the propensity score as the predicted value of a logit regression. Table 3.4 reports the estimated coefficients (not marginal effects) and standard errors from our base case.[21][22] We do not discuss the estimated coefficients because the main focus of this paper is on using the propensity score to correct for selection bias rather than on the correlates of seropositivity. Furthermore, it is quite difficult to have much intuition about the marginal impact of a single regressor conditional on the other 44.

[21]Elements of the X vector of covariates were imputed to zero for item non-response, and a dummy variable for imputation was included in the specifications. We include indicator variables for each province in South Africa. These coefficients are not reported in the table.

[22]We also examine particular subsamples of the data (e.g. by sex, by education, by area of residence). For each of these sub-samples, we re-estimate the propensity score. These results are available on request.

3.5.2 Examining the Validity of the Propensity Score Method

We next examine the validity of the CIA and common support assumption under our preferred (base case) propensity score specification. The former is done with balancing tests and the latter by examining empirical distributions.

The intuition behind balancing tests is appealingly clear. The idea of propensity score reweighting is to reweight the distribution of the observables (X's) of the HIV- population so as to match the distribution of the X's for the HIV+ population (see Equation 3.2). Balancing tests simply examine whether the difference in means of X's between HIV+ and HIV- populations is reduced when the observations are reweighted. If the reweighted means are similar (i.e. not significantly different) for HIV+ and HIV- populations, then the reweighting has achieved its goal, the data are balanced, and the CIA is appropriate.

Table 3.5 reports the results of balancing for those variables that enter the propensity score regression (i.e. internal balancing).[23] The results in Table 3.5 apply to the entire sample. Balancing tests were also conducted for each of the four subsamples with comparable results. Row 1 of Table 3.5 presents the results for age; HIV+ individuals were on average 3.038 years younger than the HIV- population. This difference had a t-statistic of 6.087 so the difference was highly significant. After reweighting, the difference falls to 0.254 years and is not significantly different from zero. None of the variables in the propensity score regression have a statistically significant difference between the HIV+ and HIV- populations in the reweighted data.

[23]We also conducted balancing tests for each of the indicator variables for missing data, but these are not reported in the table. All were balanced.

The propensity score weight, by design, attempts to minimize differences between HIV+ and HIV- groups for the variables included in the propensity score estimation (the X vector). A more stringent balancing test criterion is whether the propensity score weight also succeeds in balancing covariates that were not used in the estimation. External balancing tests were conducted on 56 variables. These external variables included measures of sexual activity and almost 50 variables that are plausibly related to economic well-being (e.g. source of water, type of cooking fuel, type of toilet, and measures of privation.) In the unweighted data, 20 of these variables had means that were significantly different (at the 95 percent significance level) between the HIV+ and HIV- populations. After reweighting, only three of those differences were still significant.

Based on internal and external balancing tests, we conclude that the propensity score specification is adequately reweighting the data to justify the CIA. While selection on observables is a strong assumption, we find that the richness of our data provides support for the assumption.

The balancing tests provide support for the CIA. The other assumption underlying the propensity score reweighting approach is the common support condition.[24] We examine the appropriateness of the support condition by comparing the empirical distributions of the propensity scores for the HIV+ and HIV- populations. These distributions are shown in Figure 3.2. In that figure, it is clear that the densities have a common support.

[24]The support condition is not required for the control function approach, however confirming that it holds ensures that we are not relying on solely on functional form assumptions for any values of the propensity score.

3.5.3 The Impact of HIV

Table 3.6 presents estimates of the causal impact of HIV on labor market participation. The table is organized such that each column presents estimates resulting from a different estimator and each row presents estimates using different samples of the data.

The first column presents naïve estimates of the impact of HIV status on labor market participation. Estimates in this column are simply the (unconditional) difference in the mean employment status for the HIV+ and HIV- populations. Variables are defined such that the 0.157 figure in the first cell implies that, on average, HIV+ individuals are 15.7 percentage points more likely to be unemployed.

The second column presents the coefficient on HIV status when the variables in the propensity score are included as controls in a simple logit regression of employment status on HIV status. This is a simple regression to run and is the same as the control function approach but it excludes the higher order terms of the propensity score. The logit result implies that for the full sample being HIV+ lowers the probability of employment by 6.8 percentage points.

The third and fourth columns present estimates of the causal impact of HIV on employment using propensity score reweighting and the control function approaches, respectively. Each parameter estimate in columns 3 and 4 is the marginal effect of HIV status from a logit regression. For column 3, it is a simple logit regression of labor market status on HIV status using propensity score reweighting while in column 4, the reported estimate is the coefficient on HIV status (γ) in the regression given by Equation 3.4. Using the entire sample, propensity score reweighting indicates that being HIV+ raises the probability of unemployment by 6.4 percentage points while the control function approach indicates an increase of 7.6 percentage points. Each of these impacts is precisely estimated.[25] Our estimates of the causal impact of HIV on employment imply that, all else equal, being HIV+ raises the probability of unemployment by about seven percentage points.

The estimates in the first row of Table 3.6 apply to the entire sample and, as such, potentially hide substantial underlying heterogeneity in the impact of HIV on labor market status. We investigate this heterogeneity in the remainder of Table 3.6 by restricting the sample to particular sub-populations. The second through seventh rows in Table 3.6 present results obtained from an analysis of sub-samples of the data, using propensity score estimates calculated within the sub-sample alone.

For men, propensity score reweighting and the control function approach give estimates of 0.062 and 0.092, respectively. The former is not precisely estimated while the latter still is. For females, the point estimates from propensity score reweighting and the control function approach are 0.057 and 0.067. These are about the same as the estimates obtained with the entire sample, and they remain precisely estimated at conventional levels of statistical significance.

The next two rows in Table 3.6 restrict the analysis to respondents aged 25 and older, and divide the sample between individuals whose education level is a Matric

[25] Both the reweighting and the control function approach use the *estimated* propensity score. Hirano et al. (2003) show that using the estimated propensity score rather than the true propensity score produces efficient estimates. They suggest bootstrapping to obtain standard errors. We do so.

or higher and those whose education level is less than a Matric.[26] The message here is clear: the impact of HIV on labor market status is severe for those with lower levels of education and is negligible for those with higher levels. The causal impact of HIV status on labor market status for those with less than a Matric is an increase in the likelihood of unemployment of 10.3 percentage points (with propensity score reweighting) and 11.8 percentage points (with the control function approach.) These are large and precisely estimated impacts.[27]

The last two rows highlight urban/rural differences in the impact of HIV on unemployment. Although both HIV and unemployment are more prevalent in rural areas, the causal impact of HIV on unemployment is larger in urban areas where being HIV+ lowers the probability of employment by about 8 percentage points.

3.5.4 Discussion

A concern about the role of selection bias motivated our choice of methodologies. Comparing the naïve estimates in column 1 of Table 3.6 to the causal estimates in columns 3 and 4 speaks to this issue. The naïve estimates are about twice as large as the causal estimates, and this highlights the importance of selection. Our results are consistent with the hypothesis that individuals who are HIV+ are more likely to be unemployed than the average South African, irrespective of their HIV status.

Because this study is probably the first to examine the causal impact of HIV on

[26] Respondents under 25 may not have completed their schooling. A Matric is about equivalent to a high school education.

[27] We repeated the analysis for subgroups defined by race as well as by groups defined by education and sex, age, and education and age. These results are not reported here, but are available upon request from the authors. The results in Table 3.6 capture the gist of these divisions. Women, and especially women with lower levels of education, experience larger causal impacts of HIV on labor market status.

employment outcomes, it is difficult to place the magnitude of the estimated impact in context. There are no other estimates available for comparison. There are at least two economic arguments for why one might have expected no causal impact. First, if unemployment were so pervasive that HIV+ individuals would be unemployed even in the absence of HIV, one would expect no impact. Second, if ARVs were sufficiently widely used, one might expect either no impact or a tiny impact.[28] Our estimates indicate that these arguments, while perhaps *ex ante* plausible, are on average simply incorrect. Being HIV+ lowers the probability of employment.

A counter-argument to the notion that HIV confers a negligible penalty in the face of extremely high unemployment is that in the presence of high unemployment, even a small disadvantage (e.g. the stigma sometimes associated with HIV) much less a large disadvantage (adverse physical effects of HIV) mean the difference between keeping a job and losing it. The results in Table 3.6 are consistent with this counter-argument.

One way to gauge the magnitude of our estimates (beyond noting that they are not zero) is to compare the estimated marginal effect of HIV status (D) with the estimated marginal effect of other respondent characteristics (X) in Equation (3.4). For men, the magnitude of the labor market advantage of being HIV- is approximately equal to the impact of 3 years of age, a Matric qualification (compared to no education), or the absence of a female pensioner in the household. For women it is equal to the impact of 1.5 years of age, some secondary education (compared to no education), or the absence of a male pensioner in the household.

[28] One might expect a tiny impact rather than no impact because even with ARVs, most HIV+ people eventually do become too ill to work.

Our results are conditional on the availability of ARVs as of 2005. We do not have data on which HIV+ respondents were on ARV therapy. Access to ARVs in South Africa is far from universal. One estimate is that in 2005, the year of our sample, only about 18 percent of those who needed ARVs were actually using them. (See Dorrington et al (2006).) As the availability of ARVs changes, the impact of HIV on labor market status, as we estimate it, will change. Because ARVs are more widely available today than they were in 2005, *ceteris paribus*, the impact of HIV on unemployment is lower today than it was in 2005. In addition, ARV usage is non-random, and this may in part contribute to the pattern of results in Table 3.6. For example, ARVs are much more likely to be employer-provided in the formal sector than they are in the informal sector. Our findings are consistent with the fact that women and the less-educated tend to be more heavily represented in the informal sector and in domestic (housekeeper) work and hence less likely to receive employer-provided ARVs.

We also found that being HIV+ had virtually no employment impact for better educated workers (and recall this result already accounts for the fact that highly educated individuals are less likely to be HIV+). ARVs may also be contributing to this finding. Employers have greater incentive to invest in ARVs for workers who are more difficult to replace such as highly educated workers. In sum, HIV appears to reinforce the already existing inequalities in South Africa.

3.5.5 Sensitivity Analyses

We investigate the sensitivity of our results both to alternative definitions of unemployment and to alternative specifications of the propensity score regression. As described in section 3, individuals were divided into three groups in our base case specification – unemployed, employed, and not economically active (NEA). The first two groups are included in our analysis while the third is not. In the base case results using the broad definition of unemployment, so-called discouraged workers – workers who were not actively seeking a job but desired employment – were classified as unemployed. The narrow definition of unemployment classifies discouraged workers as NEA. We repeat the analysis using this narrow definition of unemployment. Results are reported in Table 3.7. The results are very similar to those in Table 3.6. The estimated coefficients are slightly smaller than those reported in Table 3.6, with only 1 of the 15 coefficients (in the latter 3 columns) differing from their counterpart in Table 3.6 by more than two percentage points. We also repeated the analysis by combining the NEA with the employed to create a group that might (awkwardly) be called not-unemployed. Results are reported in Table 3.8. Compared to Table 3.6, coefficients tend to increase, usually by about two percentage points. Results tend to be more precisely estimated, and this is driven in part by the increase in the sample size when NEA individuals are included in the analysis. From Tables 3.7 and 3.8, we conclude that our findings are robust to alternative definitions of labor market status.

We also experimented with alternative specifications of the propensity score regression. The results obtained from regressions using additional variables in the calculation of the propensity score are virtually identical to results obtained from the preferred specification. These results are not reported here. When including ten variables with information on the household, no point estimate changed by more than two percentage points, and only three of the 15 point estimates changed by more than one percentage point.[29] Adding additional behavioral variables did not change any of the point estimates by as much as one percentage point.[30]

3.6 Conclusions

Identifying the causal impact of HIV on labor market status requires addressing the issue of selection into being HIV+. In the absence of plausible instruments, we exploit the richness of the data and assume that selection is on observables. External balancing tests support the validity of this assumption. Employing two estimation strategies, we find that being HIV+ causes, on average, an increase in the likelihood of unemployment of about seven percentage points. This penalty exists despite very high unemployment rates. The average impact hides important heterogeneity. HIV's causal impact on unemployment is larger (10 to 11 percentage points) for less educated South Africans. The results are robust to multiple alternative econometric specifications.

3.6.1 Caveats

These results are the first nationally representative estimates of the causal impact of HIV on employment in South Africa. While informative, they are not dispositive. Rather, the results should be interpreted with caution for at least five reasons.

[29] The additional household variables chosen to indicate socioeconomic level were: type of toilet facility, source of energy for cooking, access to electricity, presence of a land line, and a dummy for whether the household information was missing for the observation.

[30] The additional behavioral variables were condom use at last sex, number of current sexual partners, number of partners in the last year, whether the respondent had been tested for HIV before, whether they had received their test result, and whether they had heard of ARVs.

First, our analysis does not account for any general equilibrium effects. In particular, it would be misleading to think that if ARVs were made universally available or a cure for HIV/AIDS were found, that HIV+ individuals would see their likelihood of employment rise on average by about seven percentage points. Rather, the labor market would adjust and these adjustments would depend on supply and demand elasticities. Complicating this analysis, the ability of the labor market to immediately absorb additional healthy workers is questionable.

Second, data limitations preclude an analysis of the indirect labor market impact of having multiple HIV+ adult household members. Recall that the structure of the SABSSM II survey is such that only one adult age 25 or older is sampled from each household. It is unclear in which direction our results may be biased. An HIV- worker could be unemployed because they are caring for an HIV+ spouse, resulting in downward bias (i.e. the true impact is larger than our estimates suggest), or an HIV+ worker might be more motivated to obtain employment to financially support another HIV+ household member, resulting in upward bias.

Third, our results are conditional on the time profile of HIV and prevalence rates as of 2005. Given an approximately nine-year period (on average) of latent HIV infection before AIDS conversion, and HIV prevalence rates that increased from 5 percent in 1996 to 12 percent in 2001,[31] we would expect the "stock" of individuals with AIDS to increase quite sharply between 2003 and 2008. This implies that, *ceteris paribus*, the impact of HIV on unemployment would rise in coming years

[31] ASSA demographic model predictions cited in Nattrass (2004) (p. 42). The model on which these estimates are based predicted a 14 percent prevalence rate for 2004 – about the same as that of our 2005 data.

as the number of HIV+ individuals who are too ill to work increases. However, it is unclear how this effect would interact with any increase in the availability of ARVs.

Fourth, our results are not structural. As such, we are unable to convincingly address the particular avenues through which HIV impacts labor market status. Relatedly, we are unable to conduct detailed policy analysis. For example, increased access to ARVs and successful programs to de-stigmatize HIV might each increase the likelihood of employment, but our approach cannot conduct the counterfactual experiments to estimate the likely impacts of these potential policies.

Lastly, our results are generally not applicable to other countries. South Africa has a stunningly high rate of unemployment, high HIV prevalence rates and a troubled history with the distribution of ARVs – three factors that suggest that it may be misguided to generalize the results of this study to other countries.

3.7 References

Arndt, C. and J. D. Lewis. 2000. "The Macro Implications of HIV/AIDS in South Africa: A Preliminary Assessment." *South African Journal of Economics* 68 (5): 856-87.

Auvert, B., D. Taljaard, E. Lagarde, J. Sobngwi-Tambekou, R. Sitta, and A. Puren. 2005. "Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial." *PLoSMed* 2 (11): e298.

Bell, C., S. Devarajan, and H. Gersbach. 2003. "The Long-Run Economic Costs of AIDS: Theory and an Application to South Africa." World Bank Policy Research Working Paper, October, 31-52.

Bloom, D. E. and A. S. Mahal. 1997. "Does the AIDS epidemic threaten economic growth?" *Journal of Econometrics* 77 (1):105-124.

Busso, M., J. DiNardo, and J. McCrary. 2010. "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators." Working Paper.

Card, D. 1999. "The Causal Effect of Education on Earnings." *Handbook of Labor Economics* 3, 1801-1863.

Chay, K. Y. and M. Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy* 113 (2): 376-424.

Cuddington, J. and J. Hancock. 1994. "Assessing the impact of AIDS on the growth path of the Malawian economy." *Journal of Development Economics* 4, 363-8.

Dehejia, R. H. and S. Wahba. 1997. "Causal Effects in Non-Experimental Studies: Re- Evaluating the Evaluation of Training Programs." in Rajeev Dehejia, ed., *Econometric Methods for Program Evaluation*, Ph.D. Dissertation, Harvard University, chapter 1.

Dehejia, R. H. and S. Wahba. 1999. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94 (448): 1053-1062.

DiNardo, J., N. M. Fortin, and T. Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001-1044.

Dinkelman, T., D. Lam, and M. Leibbrandt. 2007. "Household and community income, economic shocks and risky sexual behavior of young adults: evidence from the Cape Area Panel Study 2002 and 2005." *AIDS* 21 (Suppl 7): S49-S56.

Dorrington, R., D. Bradshaw, L. Johnson, and T. Daniel. 2006 "The demographic impact of HIV/AIDS in South Africa: National and Provincial Indicators for 2006."

Technical Report. Joint publication of the Centre for Actuarial Research, the Burden of Disease Research Unit, and the Actuarial Society of South Africa (ASSA).

Dunkle, K., Rachel K. Jewkes, Heather C. Brown, Glenda E. Gray, James A. McIntyre, and Sioban Harlow. 2004. "Transactional sex among women in Soweto, South Africa: Prevalence, risk factors and association with HIV infection." *Social Science and Medicine* 59, 1581- 92.

Fox, M. P., S. Rosen, W. B. MacLeod, M. Wasunna, M. Bii, G. Foglia, and J. L. Simon. 2004. "The impact of HIV/AIDS on labour productivity in Kenya." *Tropical Medicine and International Health* 9 (3): 318-24.

Frölich, M. 2004. "Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators." *Review of Economics and Statistics* 86 (1): 77-90.

Habyarimana, James, Bekezela Mbakile and Cristian Pop-Eleches. 2008. "HIV/AIDS, ARV Treatment and Worker Absenteeism: Evidence from a Large African Firm." Georgetown University: Mimeo.

Heckman, J. and S. Navarro-Lozano. 2004. "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *The Review of Economics and Statistics* 86 (1): 30-57.

Heckman, J. J. and R. Robb Jr. 1985. "Alternative methods for evaluating the impact of interventions: An overview." *Journal of Econometrics* 30(1-2): 239-267.

Heckman, J. and V. J. Hotz. 1989. "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association* 84 (408): 862-874.

Hirano, K., G. W. Imbens, and G. Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71 (4): 1161-1189.

Kalemli-Ozcan, S. 2006. "AIDS, Reversal of the Demographic Transition and Economic Development: Evidence from Africa." NBER Working Paper, 121-81.

Kambou, G., S. Devarajan, and M. Over. 1992. "The economic impact of AIDS in an African country: simulations with a computable general equilibrium model of Cameroon." *Journal of African Economies* 1: 109-130.

LeClerc-Madlala, S. 2002. "Youth, HIV/AIDS and the importance of sexual culture and context." *Social Dynamics - A Journal of the Centre for African Studies*, University of Capetown, 28, 20-41.

Levinsohn, J. and J. McCrary. 2008. "Estimating Correct HIV Prevalence Rates: The Case of Botswana." Working Paper.

McKelvey, C. 2007. "Circumcision and the Labor Market Consequences of HIV in Developing Countries." Mimeo.

Murray, J., P. Sonnenberg, G. Nelson, S. Shearer, A. Bester, A. Begley, and J. R. Glynn. 2005. "Effect of HIV on work-related injury rates in South African goldminers," *AIDS*, 19:2019-24.

Natrass, N. 2004. *The Moral Economy of AIDS in South Africa.* Cambridge University Press.

Olley, S. and A. Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263-1298.

Rosenbaum, P. R. and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41-55.

Shisana, O, T Rehle, LC Simbayi, W Parker, K Zuma, A Bhana, C Connolly, S Jooste, and V Pillay. 2005. *South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey.* Human Sciences Research Council Press, Cape Town.

Strauss, John and Duncan Thomas. 1998. "Health, Nutrition, and Economic Development," *Journal of Economic Literature*, 36(2): 766-817.

Thirumurthy, Harsha, Joshua Graff Zivin and Markus Goldstein. 2008. "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya." *Journal of Human Resources*, 43:511-552.

Thomas, D.; Frankenberg, E.; Friedman, J.; Habicht, J.P.; Hakimi, M.; Ingwersen, N.; Jaswadi; Jones, N.; McKelvey, C.; Pelto, G. 2006. "Causal EFFect of Health on Labor Market Outcomes: Experimental Evidence," California Center for Population Research. On-Line Working Paper Series. Los Angeles, CA.

UNAIDS. 2005. "UNAIDS/WHO AIDS Epidemic Update." Technical Report.

Wawer, M., G Kigozi, D Serwadda, F Makumbi, Nalugoda, S Watya, D Buwembo, V Ssempijja, L Moulton, and R Gray. 2008. "Trial of Male Circumcision in HIV+ Men, Rakai, Uganda: Effects in HIV+ Men and in Women Partners." Paper presented at Conference on Retroviruses and Opportunistic Infections.

Weiss, Helen A., Maria A. Quigley, and Richard J. Hayes. 2000. "Male circumcision and risk of HIV infection in sub-Saharan Africa: a systematic review and meta-analysis." *AIDS* 14, 2361-70.

Werker, E., A. Ahuja, and B. Wendell. 2006. "Male Circumcision and AIDS: The Macroeconomic Impact of a Health Crisis." Harvard Business School Working Paper, (07-025).

Young, A.. 2005. "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations." *Quarterly Journal of Economics* 120 (2).

Zuma, K., M.N. Lurie, B.G. Williams, D. Mkaya-Mwamburi, G.P. Garnett, and

A. W. Sturm. 2005. "Risk factors of sexually transmitted infections among migrant and non-migrant sexual partnerships from rural South Africa." *Epidemiol. Infect.* 133, 421-428.

3.8 Tables and Figures

Table 3.1: Sample sizes

Sample sizes	
Total sample	23275
Aged 15 and up	16398
Agreed to HIV test	12032
Tested HIV+	1351
Infected for more than 6 months	1172

Table 3.2: Naïve HIV prevalence rates by race

Race	HIV prevalence rate (%)	N
African	17.34	9,664
Coloured	2.73	3,013
Indian	1.27	1,772
White	0.53	1,913
Total	14.14	16,362

[1] Sample includes all individuals aged 15 and over.

[2] Rates are calculated using sample weights only.

Table 3.3: Employment status by HIV testing status (for men and women)

Employment status	HIV+	HIV-	Not tested	Total	N
Men					
Employed (and other)	8.5	64.1	27.4	100	2,815
Unemployed (broad definition)	12.6	63.0	24.4	100	1366
Student (NEA)	1.9	68.3	29.8	100	1408
Other NEA	4.3	72.2	23.5	100	621
Missing data	23.4	58.5	18.2	100	131
Total	8.1	65.4	26.6	100	6,341
Women					
Employed (and other)	10.9	61.5	27.6	100	2,693
Unemployed (broad definition)	20.8	58.1	21.2	100	3,037
Student (NEA)	7.7	67.2	25.1	100	1,568
Other NEA	6.8	72.1	21.1	100	2,636
Missing data	5.8	67.3	27.0	100	123
Total	12.8	63.8	23.4	100	10,057

[1] Sample includes all individuals aged 15 and over. NEA: not economically active.

Table 3.4: Propensity score regression for HIV+ status (base case)

Variable	Coefficient
Age	0.133***
	(0.048)
Age-squared	-0.002***
	(0.000)
Female	0.466***
	(0.109)
African	2.965***
	(0.416)
Coloured	1.450***
	(0.456)
Indian	0.288
	(0.595)
Urban resident	0.303***
	(0.124)
Never married	0.376***
	(0.122)
Has had sex	2.565***
	(1.025)
Age at first sex	-0.176**
	(0.093)
Age at first sex squared	0.004*
	(0.002)
Completed primary education	0.245*
	(0.211)
Completed secondary education	0.346*
	(0.198)
Holds a matric qualification	0.119*
	(0.227)
Has some post-matric education	-0.508
	(0.270)
Used a condom at first sex	-0.293*
	(0.163)
Knows HIV transmitted through vaginal sex	-0.197
	(0.203)
Believes HIV not transmitted through witchcraft	-0.176
	(0.162)
Knows condoms prevent HIV transmission	-0.063
	(0.166)
Knows reducing number of partners reduces risk	0.190*
	(0.116)
Male of pension age in household	0.032
	(0.457)
Female of pension age in household	-0.218
	(0.186)
Constant	-8.565***
	(0.922)
Pseudo-R^2	0.1294
Observations	7467

Table 3.5: Balancing tests for variables used in estimation of propensity score

	Unweighted			**Reweighted**		
Variable	**Diff.**	**Std. E**	**t-Stat**	**Diff.**	**Std. E**	**t-Stat**
Age	-3.038	0.499	-6.087	-0.254	0.487	-0.523
Age-squared	-275.752	40.647	-6.784	-21.143	38.573	-0.548
Female	0.125	0.024	5.125	0.000	0.026	0.009
African	0.225	0.009	25.459	0.000	0.004	0.022
Coloured	-0.104	0.006	-17.260	0.000	0.004	0.006
Indian	-0.023	0.003	-8.478	-0.000	0.001	-0.047
Eastern Cape	-0.002	0.014	-0.125	-0.002	0.015	-0.118
Northern Cape	-0.014	0.003	-5.107	-0.000	0.002	-0.193
Free State	0.014	0.016	0.858	-0.000	0.017	-0.029
Kwa-Zulu Natal	0.084	0.021	3.972	0.002	0.024	0.100
Northwest Province	0.012	0.017	0.721	0.001	0.017	0.077
Gauteng	0.013	0.022	0.578	0.004	0.023	0.176
Mpumalanga	0.042	0.013	3.189	-0.003	0.015	-0.189
Limpopo	-0.036	0.012	-2.899	-0.002	0.012	-0.200
Urban dweller	-0.006	0.025	-0.244	0.002	0.027	0.094
Never married	0.123	0.024	5.011	0.013	0.027	0.491
Has had sex	0.025	0.012	2.055	0.002	0.012	0.139
Age at first sex	-0.621	0.398	-1.559	0.024	0.423	0.056
Age at first sex-squared	-21.616	8.736	-2.474	-0.314	9.238	-0.034
Completed primary education	0.035	0.022	1.575	-0.007	0.024	-0.310
Completed secondary education	0.074	0.024	3.056	0.008	0.026	0.322
Holds a matric qualification	-0.019	0.020	-0.952	0.000	0.021	0.022
Completed some post-matric educ	-0.080	0.011	-7.382	0.002	0.009	0.199
Used a condom at first sex	-0.040	0.017	-2.291	0.003	0.017	0.196
Knows HIV transmitted through vaginal sex	-0.012	0.012	-0.955	-0.002	0.013	-0.162
Believes HIV not transmitted through witchcraft	-0.016	0.018	-0.918	0.003	0.019	0.138
Knows condoms prevent HIV transmission	0.026	0.015	1.744	0.001	0.016	0.089
Knows reducing number of partners reduces risk	0.004	0.023	0.161	0.001	0.025	0.027
Male of pension age in household	-0.013	0.015	-0.847	-0.003	0.016	-0.161
Female of pension age in household	-0.023	0.012	-1.853	-0.002	0.013	-0.123

[1] Difference between HIV+ and HIV- subsamples. Includes all individuals aged 15 and older.

Table 3.6: Marginal effect of being HIV+ on likelihood of being *un*employed

Sample	Naïve		Logit		RW		CF		N
Full sample	0.157	***	0.068	**	0.064	***	0.076	***	7467
	(0.024)		(0.029)		(0.024)		(0.021)		
Male	0.074		0.059		0.062		0.092	**	3087
	(0.041)		(0.045)		(0.045)		(0.040)		
Female	0.166	***	0.065		0.057	**	0.067	***	4380
	(0.027)		(0.033)		(0.029)		(0.025)		
Matric and up	0.183	***	-0.016		-0.004		0.026		1803
	(0.054)		(0.047)		(0.068)		(0.055)		
Less than Matric	0.144	***	0.110	***	0.103	***	0.118	***	3702
	(0.032)		(0.035)		(0.031)		(0.030)		
Rural	0.121	***	0.056		0.048		0.064	**	2270
	(0.039)		(0.043)		(0.037)		(0.030)		
Urban	0.179	***	0.075	**	0.077	**	0.078	***	5197
	(0.030)		(0.036)		(0.032)		(0.027)		

[1] Full sample and results by sex include individuals aged 15 and older. Results by education level are restricted to individuals aged 25 and older.
[2] Naïve is OLS without controls. Controls used in logit are same as those used in propensity score estimation.
[3] Bootstrapped standard errors reported for RW and CF. *** significant at 99 percent level, ** significant at 95 percent level.

Table 3.7: Sensitivity check using narrow definition of unemployment

Sample	Naïve		Logit		RW		CF		N
Full sample	0.160	***	0.053		0.055	**	0.066	***	6907
	(0.025)		(0.029)		(0.024)		(0.022)		
Male	0.064		0.038		0.045		0.078		2907
	(0.041)		(0.043)		(0.045)		(0.040)		
Female	0.177	***	0.061		0.051		0.068	**	4000
	(0.028)		(0.036)		(0.030)		(0.027)		
Matric and up	0.169	***	-0.029		-0.016		0.032		1748
	(0.054)		(0.040)		(0.065)		(0.056)		
Less than Matric	0.147	***	0.092	**	0.086	***	0.102	***	3390
	(0.033)		(0.037)		(0.032)		(0.031)		
Rural	0.123	***	0.036		0.033		0.051		2064
	(0.041)		(0.048)		(0.041)		(0.033)		
Urban	0.183	***	0.060		0.073	**	0.072	***	4843
	(0.031)		(0.034)		(0.032)		(0.027)		

[1] Full sample and results by sex include individuals aged 15 and older. Results by education level are restricted to individuals aged 25 and older.
[2] Naïve is OLS without controls. Controls used in logit are same as those used in propensity score estimation.
[3] Bootstrapped standard errors reported for RW and CF. *** significant at 99 percent level, ** significant at 95 percent level.

Table 3.8: Sensitivity check including not economically active (NEA) individuals in the sample

Sample	Naïve		Logit		RW		CF		N
Full sample	0.222	***	0.082	***	0.095	***	0.106	***	12032
	(0.021)		(0.021)		(0.023)		(0.021)		
Male	0.146	***	0.046		0.064		0.104	***	4529
	(0.037)		(0.034)		(0.043)		(0.037)		
Female	0.248	***	0.092	***	0.101	***	0.092	***	7503
	(0.025)		(0.026)		(0.028)		(0.024)		
Matric and up	0.199	***	-0.002		0.012		0.021		2173
	(0.051)		(0.036)		(0.063)		(0.053)		
Less than Matric	0.206	***	0.088	***	0.098	***	0.102	***	5739
	(0.028)		(0.029)		(0.030)		(0.028)		
Rural	0.192	***	0.074	**	0.086	**	0.098	***	3956
	(0.034)		(0.034)		(0.035)		(0.030)		
Urban	0.244	***	0.090	***	0.107	***	0.109	***	8076
	(0.027)		(0.025)		(0.030)		(0.026)		

[1] Full sample and results by sex include individuals aged 15 and older. Results by education level are restricted to individuals aged 25 and older.
[2] Naïve is OLS without controls. Controls used in logit are same as those used in propensity score estimation.
[3] Bootstrapped standard errors reported for RW and CF. *** significant at 99 percent level, ** significant at 95 percent level.

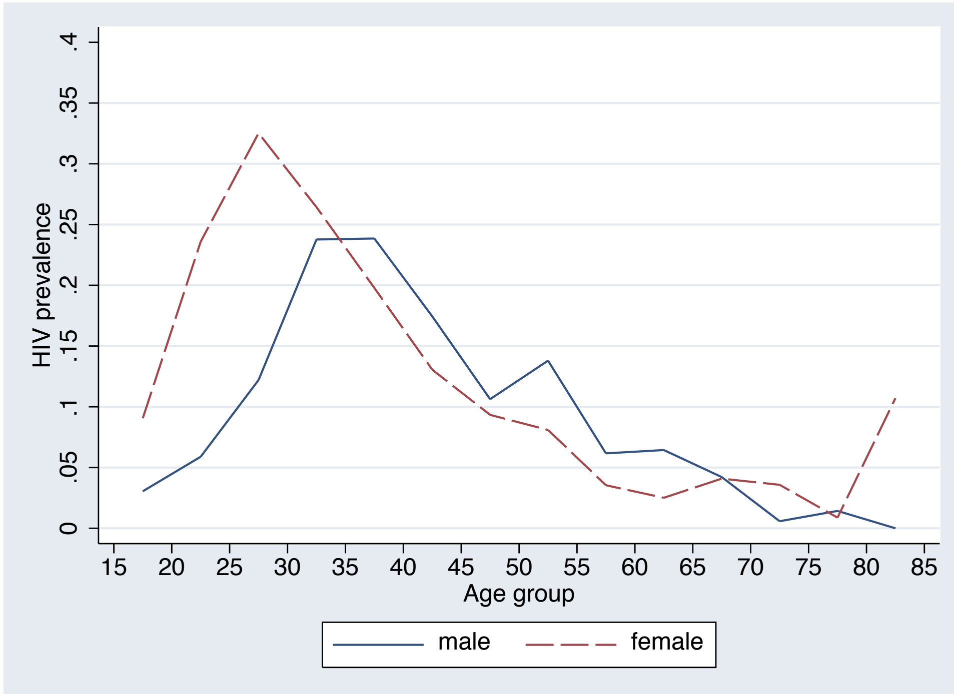


Figure 3.1: Age profile of naïve HIV prevalence

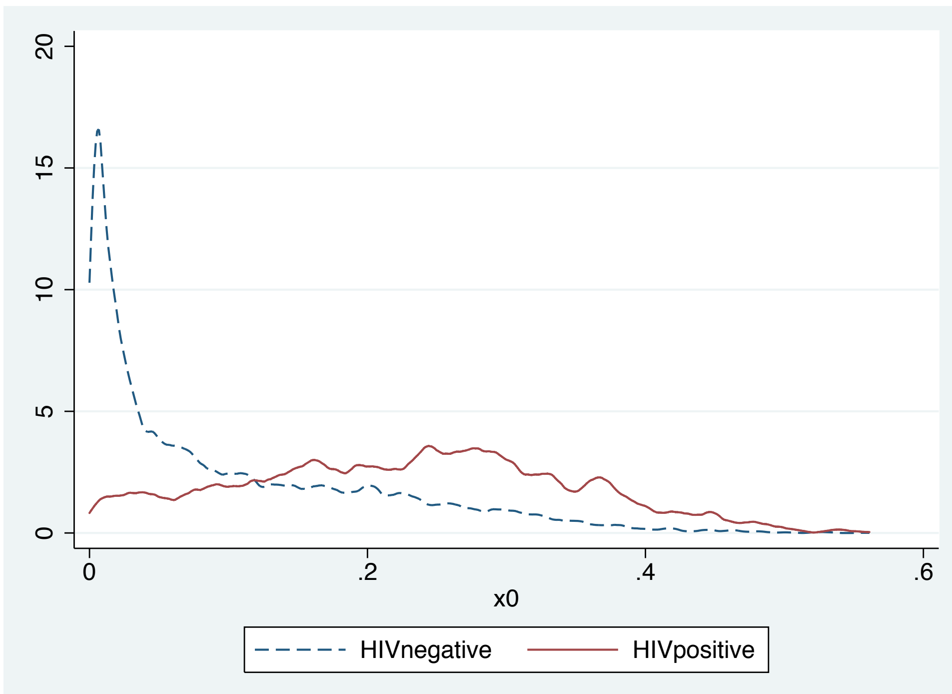


Figure 3.2: Empirical distributions of the propensity scores for the full sample

CHAPTER IV

How Does Intra-Household Job Loss Affect Labor Market Outcomes of the Unemployed in South Africa?

4.1 Introduction

South Africa has one of the highest unemployment rates in the world. More than ten years after the end of the Apartheid regime, the unemployment rate is still above 25 percent under the International Labor Organization (ILO) classification. By excluding discouraged workers, this official unemployment rate severely underestimates the true burden of unemployment, which is above 40 percent under the broad classification.[1] Black Africans and adults under 35 years old are disproportionately affected, and a large proportion of the unemployed have never held a job. New Labour Force Survey (LFS) panel data promises to offer insight into the unemployment problem by allowing researchers to track individuals who transition from unemployment to employment. A better understanding of the factors that determine transition rates will enable the government to target policy interventions more effectively.

After the fall of the Apartheid regime in 1994, Black African, Coloured and Indian workers who had previously been restricted from entering the cities or from obtaining certain types of employment joined the workforce en masse. In addition, Black

[1] South Africa Labour Force Survey Statistical Release P0210, September 2006.

Africans have flowed steadily into the labor force over the past decade and continue to do so. The new entrants to the labor force changed not only the unemployment rate, but also the composition of the unemployment pool because they had lower skill levels and less work experience due to Apartheid-era employment restrictions. These new entrants to the labor market still have not been fully absorbed into the workforce. While unemployment rates have declined slightly since 2003, there is no indication that they will be brought down to reasonable levels anytime soon.[2]

Unemployment appears to be particularly sticky: many individuals are unable or unwilling to extricate themselves from it. Moreover, in the Labour Force Survey (LFS) panel over 60 percent of the unemployed have never held a job before, including many adults under 35 years old. The stickiness of unemployment results in many South Africans being unemployed for long stretches of time, and the duration of unemployment is generally highest for Black Africans. Klasen and Woolard (2009) find that only about 3 percent of the unemployed are receiving unemployment support at any point in time. Having such a small portion of the unemployed receiving government income raises the question of how the unemployed are supporting themselves.

In the absence of a formal unemployment support system targeted at the unemployed, households and families are an important form of social insurance. Earned and unearned income of other household members is likely an important source of support for the unemployed during an extended period of job search or discouragement, acting as a private safety net. In Southern Africa, it is common to share resources among extended family networks. For example, the elderly share their

[2] South African Labour Force Survey Data.

pensions with kin networks in the expectation of being supported in times of need (Sagner and Mtati 1999). A number of studies have shown that pension income to one household member is generally shared within the household, which can alter labor supply (Bertrand et al. 2003), affect transfers from migrants (Jensen 2003), enhance household security (Ardington and Lund 1995) and change the allocation of household income to food, schooling and savings (Case and Deaton 1998). Duflo (2003) demonstrates that the allocation of pension income depends on the gender of the pension recipient.[3] Like pensioners, employed members of the household will generally take on some of the burden of financially supporting unemployed household members.

Just as changes in welfare and unemployment benefits have effects on labor market outcomes, changes in household income alter job search parameters, namely the value of the outside option and as a result, the reservation wage. In this paper, I focus specifically on the effect of a household member's job loss on the labor market outcomes of other members of the household. I attempt to identify what Woytinsky called the *added worker effect*: "the familiar story of the head of the family losing his job whereupon his wife and children also start looking for work" (Humphrey 1940). A decrease in household income will decrease the value of the outside option causing the reservation wage to fall, search effort to increase and unemployment duration to decrease.

A job loss or reduction in wage of one person affects the *desired* labor supply of other members of the household.[4] Maloney (1987) identifies the two pathways

[3]The tight kin network of exchanges and obligations has also been documented among black urban poor near Chicago, Illinois (Stack, 1974).

[4]This holds whether the household is modeled as a unitary decision-maker or as a set of actors with bargaining

through which the added worker effect operates: the income effect and the cross-substitution effect. A reduction in household income decreases per capita income for each member of the household. In response, individuals shift away from leisure, and increase their labor supply and search effort. This effect would be particularly pronounced if households prefer increased labor supply to dissaving or incurring debt, or if the latter options are infeasible (as is the case for most Black African households in South Africa). The positive cross-substitution effect arises when a household member takes over some home production activities after he or she loses a job or has their hours reduced (Maloney 1987). This lowers the shadow wage for home production for other unemployed or underemployed household members (i.e. the outside option, in the job search literature), which lowers the reservation wage and leads them to increase their labor supply.[5] The added worker effect may also operate through an increase in motivation to find employment due to pressure from or altruism towards other household members.

Using U.S. data and examining only husband-wife pairs, Heckman and MaCurdy (1980, 1982), Lundberg (1985), Spletzer (1997) and Stephens (2002) find evidence of a small but significant added worked effect. Using the Panel Study on Income Dynamics (PSID), Yeung and Hofferth (1998) find a marginally statistically significant added worker effect for black families in the U.S., noting that homeowners were more likely to increase work hours, and families in areas with high unemployment were least likely to do so. Layard et al. (1980) finds that the employment outcomes of husbands and wives are positively related rather than negatively related as the

power (Basu et al. 1999), provided that bargaining power over household resources is greater than zero for all household members.

[5] I do not attempt to distinguish the relative size of the income effect and the cross-substitution effect in either the theory or the empirical work in this paper. Both effects lower the reservation wage.

added worker effect predicts, which could be evidence of either assortative mating or being subjected to the same poor labor market. Also using the PSID, Maloney (1987, 1991) finds little evidence of the added worker effect, except in the case of men's *under*employment being positively related to women's labor supply, and he rejects the hypothesis that unobserved variables are obscuring the effect. Gruber and Cullen (2000) find that unemployment insurance crowds out the added worker effect by mitigating the loss in household income.

There is a more limited literature on the added worker effect internationally. Fernandes and de Felício (2002) find an added worker effect in Brazil that is larger than the U.S. estimates; they attribute the magnitude to liquidity constraints that prevent families from smoothing temporary income shocks. Similarly, other studies in Turkey (Baslevent and Onaran 2003) and Mexico (Parker and Skoufias 2004) have found the added worker effect to be more pronounced during times of economic crisis when credit constraints are most binding. Serneels (2002) finds no added worker effect in Ethiopia in terms either employment or desired labor supply and concludes that it is because families were able to sell assets or employ consumption smoothing.

The added worker effect is primarily defined as influencing *desired* labor supply, which may or may not translate into an observable change in employment. Maloney (1987) stresses that the added worker effect may not be evident in the data because actual hours of work are desired hours censored at zero, and I do not observe changes in desired hours below zero in the LFS data. This issue is further compounded in South Africa where structural and frictional causes of unemployment undoubtedly prevent job seekers from translating desired hours into employment in many cases.

In this paper, I examine whether a negative shock to household income increases the likelihood that unemployed individuals (discouraged or searching) obtain employment and whether discouraged workers begin or resume searching for employment in response to the shock. I focus on Black African households because Black Africans compose over 80 percent of the South African population, their unemployment rate is very high, and kin support networks are generally strong. I find clear evidence of an added worker effect among Black African men and some evidence of the effect among Black African women. Though men are less likely to obtain employment in the 6-12 months following a negative employment shock to the household, they are statistically significantly more likely to obtain employment 12-18 months after a negative shock and following two consecutive negative shocks. Women are slightly more likely to obtain employment only 12-18 months after a negative shock. Labor force participation is consistent with an added worker effect for men: male discouraged workers are marginally more likely to begin or resume a job search if they are members of a household that has experienced a recent job loss. However, if the household experienced one or more job losses 6-12 months ago and 12-18 months ago, then discouraged workers are less likely to start searching. Neither of these effects is observed for women.

The outline of my paper is as follows: Section 2 presents a simple model of job search and labor market status to elucidate the mechanisms of the added worker effect. Section 3 describes new Labour Force Survey data that became available in September 2006 that for the first time allows researchers to track a nationally-representative sample of individuals over time. To my knowledge, this is one of the first papers to make use of the LFS panel data. Section 4 presents some descriptive statistics that compare households that experience a transition into unemployment with those that do not. I also include transition matrices to compare employment outcomes for Black Africans and Whites. I present my regression methods in Section 5 and my results in Section 6. I discuss some of the implications of my results in Section 7. Section 8 concludes.

4.2 Theoretical Model

I present a model of search and employment decisions, based on McCall's model of intertemporal job search (McCall, 1970). An agent chooses her labor force participation status in each period to maximize the present discounted value of lifetime income:

$$\max\{PDV\} = \max E[\sum_{t=0}^{\infty} \beta^t y_t], \tag{4.1}$$

where y_t is income in period t and $0 < \beta < 1$ is the discount factor.

The agent is not employed in the initial period, so initially she chooses between searching and not searching (discouraged) and

$$V(w) = \max\left\{V_{discouraged}, V_{search}\right\}. \tag{4.2}$$

In each period spent searching, the agent pays search costs x and receives one wage offer drawn from the distribution $F(w)$. There is only one level of search effort in the model, and search costs (x) are identical across individuals.

$$V(w) = \max\left\{V_{discouraged}, -x + \max\{V_{accept}, V_{reject}\}\right\} \tag{4.3}$$

If the agent accepts a wage offer of w, then she receives a wage of w every period and retains her job indefinitely (there are no exogenous or endogenous separations in the model). This provides a payoff of $\frac{w}{1-\beta}$. The value of rejecting the present period's wage offer is

$$V_{reject} = b + \beta \int_0^\infty V(w')dF(w'), \tag{4.4}$$

the value of an individual's outside option in the present period, plus the discounted expected value of searching the following period.

The agent accepts the first wage offer above the reservation wage that she receives. An optimizing worker will set the reservation wage ($\bar{w}$) such that the value of accepting the reservation wage is equal to the value of rejecting it:

$$V(\bar{w}) = \frac{\bar{w}}{1-\beta} = b + \beta \int_0^\infty V(w')dF(w'), \tag{4.5}$$

and solving for the reservation wage gives us

$$\bar{w} = b(1-\beta) + \beta \left\{ E[w]|_{accept} + \int_0^{\bar{w}} F(w')dw' \right\}. \tag{4.6}$$

If the value of discouragement exceeds the value of searching ($V_{discouraged} > V_{search}$), then, holding the other parameters constant, the agent will choose to remain discouraged every period, obtaining a payoff of $\frac{b}{1-\beta}$. By the same logic, choosing to search in the first period entails that the agent will always choose to search, provided there are resources available to cover search costs (x).

Expanding all the terms in Equation (4.3) we have:

$$V(w) = \max\left\{ \frac{b}{1-\beta}, -x + \max\{\frac{w}{1-\beta}, b + \beta \int_0^\infty V(w')dF(w')\} \right\}. \tag{4.7}$$

A change in the value of the outside option

Within this framework, I ask how the agent's decisions are affected by a change in the value of the outside option (b). If the value of b falls, we obtain two key results: (1) the reservation wage will fall, leading to a shorter duration of unemployment and (2) the value of discouragement will fall more than the value of searching. Thus, if the value of the outside option falls, we would expect to see individuals moving from discouraged to searching, and from searching (or directly from discouraged) to employed.

Result 1

A decrease in the value of the outside option (b) leads to a decrease in the reservation wage. Total differentiation of the reservation wage in Equation (4.6) gives us

$$d\bar{w} = (1-\beta)db + \beta\left\{F(\bar{w})d\bar{w}\right\}, \tag{4.8}$$

and solving for $\frac{d\bar{w}}{db}$ gives us

$$\frac{d\bar{w}}{db} = \frac{(1-\beta)}{1-\beta F(\bar{w})} \in [0,1]. \tag{4.9}$$

As the reservation wage $\bar{w}$ falls, the probability of rejecting a particular wage offer, $F(\bar{w})$, falls leading to a shorter duration of unemployment. Individuals will transition from unemployment to employment at a higher rate and the unemployment pool will shrink.

Result 2

The second result implies that after a fall in the outside option (b), workers will be more likely to transition from discouraged to searching:

$$\frac{\partial V_{disc}}{\partial b} > \frac{\partial E[V_{search}]}{\partial b}. \tag{4.10}$$

If a worker was initially indifferent between being discouraged and searching, then after a decrease in the value of the outside option (b), she will prefer searching. The intuition behind this result is that if she remains discouraged, she is equally likely to receive b each period, whereas if she transitions to searching, the reservation wage will adjust (fall) to make her *less* likely to obtain b in the subsequent period (i.e. more likely to accept a wage offer). (In other words, the reservation wage adjusts to reduce the negative impact of a fall in b on the PDV of future income.)

My regression model examines the effect of a fall in household income due to an employment shock on the likelihood of transitioning from discouraged to searching, and from either discouraged or searching to employed.

4.3 The Data

The South Africa Labour Force Survey (LFS) has been conducted biannually by Statistics South Africa (StatsSA) since March 2000, with the first wave of each year in March, and the second wave in September. The sample consists of about 3000 primary sampling units (PSUs) drawn from the 1996 Census Master Sample. Each PSU contains between 100 and 250 dwelling units. StatsSA used a stratified clustered random sample (with probability proportional to size), which results in each wave being a nationally representative cross-section of the country's population. I use StatsSA-provided sampling weight in generating all my results. These weights correct for dwelling-unit non-response and are benchmarked to population estimates.[6]

Detailed information was collected about the labor market situation of individuals aged 15-65 years, focusing on the preceding seven days. Data collection was performed by field staff trained by StatsSA in a face-to-face interview with the head of household. The LFS questionnaire includes questions about demographic characteristics, biographical information, activities related to work, unemployment and non-economic activities, agricultural activities and uncompensated activities.[7]

The matched and edited panel data files were provided directly to me by StatsSA on 1 September 2006.[8] Using standard International Labour Organization (ILO) definitions, I classified individuals as employed (in either the formal or the informal sector), unemployed or not economically active (NEA) based on responses to a series of survey questions. The methodology is summarized in Table 4.1. Respondents were employed if they had performed a job activity in the past 7 days, or if they were absent from a job due to bad weather, or due to personal leave to care for their own illness or that of a family member. Respondents were unemployed if they could not find work, or if they had a job but were absent due to transport problems, a layoff or another reason not mentioned above. They had to be willing to accept a suitable job if it were offered and be ready to start work within one week to be classified as unemployed. The respondent was also considered unemployed if he or she had a job

[6]The overall dwelling unit non-response rate is approximately 10 percent. The author imputed instances of item non-response to zero.

[7]More information is available at http://www.statssa.gov.za.

[8]I merged in some additional variables from the cross-sectional LFS data using the household identification number and person code variables to identify individuals.

that started at a definite date in the future. Within the unemployed, a respondent was classified as searching if they had taken active steps to look for work or to start their own business in the four weeks prior to the interview, and classified as discouraged otherwise. I classified individuals as NEA if they had another primary activity (i.e. student, homemaker, retired) *and* they preferred not to work. Seasonal workers in the off-season were also considered NEA.

Because the StatsSA classification definition changed slightly from wave to wave in the panel, I standardized the criteria across waves to match the protocol used by StatsSA for wave 7 (March 2003) so that employment transitions would be true transitions and not an artifact of changes in classification criteria. No more than 3 percent of the observations in each wave are classified differently than under StatsSA's varying definition. I must stress that these individuals are not classified incorrectly – only the definition differed.

I included discouraged workers in my analysis even though they are not included in the official (ILO) definition of unemployment because discouragement is not an absorbing state in South Africa. The transition rate between discouragement and employment is over 10 percent for Black Africans. Discouraged workers may begin searching and obtain employment within the six months elapsing between survey waves. Also, the offer arrival rate for discouraged workers may not be zero. As the theoretical model points out, the unemployed will not search in periods where there are not enough resources to cover search costs. However, if funds become available (e.g. savings accumulate) they may resume searching.

4.3.1 Panel Data

The results in this paper are based on individual-level data from waves 4 through 9 of the LFS, as released by StatsSA. From September 2001 (wave 4) to March 2004 (wave 9), the sample involved a rotating panel design, with 20 percent of respondents being rotated out between waves. Thus, between wave 4 and wave 5, 80 percent of respondents were kept in the sample, and a new 20 percent were included. This continued through each wave until wave 9, with the exception that no one was rotated out between wave 5 and wave 6. Individuals are in the cross-sectional data for anywhere between one and six waves.

Considerable effort on the part of StatsSA provided me with a panel of individuals who were present in two or more cross-sectional waves. The LFS was not originally designed to be used as a panel: respondents did not have a unique identification number across time and the survey did not include any tracking or following rules for individuals. However, physical dwelling units (households) generally retained the same identification number from wave to wave, which provided one way to attempt to track individuals. Linking individual respondents' records across time was a complex and time-consuming task. First, StatsSA staff visually compared the paper copies of the surveys to determine whether (based on name, age, sex and education composition of the household) the same family was present in consecutive waves. Then, any records that were not linked by StatsSA staff underwent a computer matching process that created provisional matches based on household identification number and person code (within the household). These provisional matches were rejected under any of the following conditions: sex or race differs between waves, age differences greater than 5 years between waves, or education differences greater than 3 years.[9]

The cross-sectional sample size is approximately 30,000 households with about 100,000 individual observations; the panel sample size ranges between 45,000 and 71,000 individuals. Panel inclusion rates range between 47 percent for wave 9 and 66 percent for wave 5 out of an expected 80 percent maximum match rate due to the rotating panel design (see Table 4.2). To be included in the panel data, an individual had to have been present in two or more cross-sectional waves *and* have their records linked during the procedure described above. Not everyone was linked, and there were concerns that the probability of inclusion in the panel was not randomly distributed. A logistic regression analysis of panel inclusion shows that Black Africans who were married, older, urban residents or who had some primary or secondary education were more likely to be in the panel, as were men with at least some education beyond high school. Employed and unemployed Black Africans were less likely to be in the panel than those who were not economically active (see Table 4.3). I used inverse probability weighting (IPW) to re-weight the LFS panel sample to reflect the distribution of covariates in the LFS cross-section.

The primary sampling unit of the survey is a dwelling, not a family or an individual. The surveys do not follow people who changed their residence, nor do they keep track of mortality or births.[10] This selection is not trivial, and my estimates are unbiased only insofar as the likelihood of being in the panel is orthogonal to the rate of transition across labor market states, conditional on the included covariates.

[9]Detailed information about the matching process is available in StatsSA (2006).

[10]The LFS data contains some basic information on migrant workers who are part of the household, however, migrants did not complete the individual questionnaire if they were not present in the household at the time of the survey.

Using inverse probability weighting to correct for differential selection into the panel did not change the results substantially.

4.3.2 Panel Inclusion Rate

I carried out analyses to ensure that the record-linking and data editing process was reliable, and did not result in a large number of false matches. Because I am analyzing employment status changes, my results are more sensitive to false matches than if I were examining variables that did not tend to vary over time. To the extent that false matches appear in the data, the transition rates will be overestimated. However, if an individual was not matched, I cannot tell whether this was because he or she was only in one cross-sectional sample or because StatsSA was unable to locate the matching record in another wave. Individuals who were in multiple survey waves but who were not matched due to measurement error or other data discrepancies would be underrepresented in the data.

I examined the panel-inclusion rate by household to estimate an upper bound on the number of false matches. For the vast majority of households, either everyone in the household was included in the panel or no one was. It is clear from Figure 4.1 that very few households with more than one member had only a fraction of members matched. This is reassuring; if every member of the household were falsely matched, it would be highly unlikely that every match in the household would have survived beyond the provisional matching stage. Household non-response accounts for a substantial portion of the entirely un-matched households.[11] To address the issues created by non-random panel inclusion, I compared results from two samples

[11] The individual weights provided by StatsSA (and used in this analysis) correct for household non-response.

in the panel: one that contained individuals in households where 100 percent of members were in the panel and a second sample that included households where at least 50 percent of household members were included in the panel. Results from both samples were similar; I present only results using the second (larger) sample.

4.4 Descriptive Statistics

Tables 4.4 and 4.5 compare transition rates for men and women, respectively, between four employment categories: not economically active (NEA), discouraged, searching and employed (informal and formal sectors combined).[12] The value in each cell is the proportion of individuals in the row category in September 2003 (wave 8) who transitioned into the column category by March 2004 (wave 9); these transition rates are broadly representative of the rates between the other LFS panel waves. Two things are immediately evident from the transition matrices. First, discouragement is a particularly sticky employment category for Black Africans compared to Whites. Over 35 percent of Black Africans who are discouraged remain so six months later, whereas this figure is only about 13 percent for White men, and 20 percent for White women. Clearly the duration of unemployment varies by race. Second, a similar proportion of Black African men and women transition from discouragement to employed as transition from unemployed to employed within the six months that elapses between waves. Discouragement is not an absorbing state and is clearly distinct from NEA based on the transition rates. Following Flinn and Heckman (1983) and Gönül (1992) it makes sense to consider discouraged workers when examining the added worker effect in Black African households.[13]

[12]The other two racial categories, Indians and Coloureds, are omitted but their employment outcomes tend to fall between those of Black Africans and Whites.

[13]Kingdon and Knight (2006) performs a series of comparisons of searching and non-searching unemployed and concludes that search is hindered by constraints rather than tastes in South Africa. This provides additional support

In my central analysis I use the incidence of one or more job losses within the household to investigate the added worker effect. Table 4.6 compares individuals in households where at least one household member transitioned to unemployment with individuals in households where no one transitioned. The first two columns present sample means and the third column shows the t-statistic on the difference in means, conditional on the other covariates in the table.[14] The two samples appear to be quite similar with two exceptions: men in households experiencing job loss have more children living with them, and individuals in households experiencing a job loss are less educated. The fourth through sixth columns of the table show that women in households experiencing job loss have less post-high school education, but are similar otherwise.

4.5 Methods

I performed a logistic regression analysis to examine the probability of transition conditional on a set of covariates. For the first analysis, the sample consisted of individuals who were unemployed in time t-1, and the dependent variable is whether they were employed (in either the formal or informal sector) in time t (i.e. six months later). For the second analysis, the sample was restricted to individuals who were unemployed and not searching (i.e. discouraged) in time t-1, and the dependent variable is whether they were labor force participants (i.e. either searching or employed) six months later. I limit my sample to Black Africans and perform separate logits for men and women because I expect that changes in household characteristics

for considering discouraged workers as distinct from being out of the labor force.

[14] Age-squared was also included in all regressions.

will affect the labor supply of each gender differently. Regression analysis for the employment outcome is based on the following specification:

$$Y_{ijt} = \Lambda(\beta_0 + \beta_1 \text{loss}_{t-1} + \beta_2 \text{loss}_{t-2} + \beta_3 \text{loss}_{t-1} * \text{loss}_{t-2} + \phi' X_{ijt-1} + \delta_t + \alpha_j + \epsilon_{ijt}) \quad (4.11)$$

where Y_{ijt} is an indicator variable for employment status, $loss$ is an indicator for a household member experiencing a job loss, $\Lambda(\bullet)$ is the logistic function, X_{ijt} is a vector of individual and household characteristics, δ_t is a set of time dummies and α_j is a set of province dummies.[15]

The sample for the transition regression analysis is restricted to Black Africans who were unemployed in at least one panel wave and whose household had at least a 50 percent panel-inclusion rate. Observations from all five panel waves were pooled in the regression, and standard errors are clustered at the individual level. I used person weights provided by StatsSA in generating all my results to correct for stratified sampling and non-response. I correct for differential selection into the panel using inverse probability weights.

4.6 Results

4.6.1 Transition Regressions: Obtaining employment

I find clear evidence of an added worker effect among Black African men and some evidence of the effect among Black African women. Table 4.7 presents results from regressions of an indicator for employment status on a set of covariates for both men and women. The specification in the first column includes only the three listed measures of recent job loss within the household, the second column adds a

[15] See table notes for full list of controls.

set of individual and household characteristics and the third column, which is the preferred specification, adds measures of welfare receipt.[16] These three specifications are repeated for women in the fourth through sixth columns. Contrary to what theory predicts, neither sex responds to a negative employment shock in the household by increasing their likelihood of obtaining employment within 6-12 months. We can see from the third column that men are 6 percentage points less likely to obtain employment within the 6-12 months following an instance of household job loss but 4.3 percentage points more likely to obtain employment 12-18 months afterwards. When the household experiences two consecutive instances of job loss, men are 9 percentage points more likely to be employed. Women are 2.2 percentage points more likely to obtain employment 12-18 months after the household experiences a job loss. It appears that the added worker effect takes over a year to have its full impact.

I find evidence of gender-specific network effects. Both the number of employed men and the number of employed women increase the likelihood of obtaining employment for both sexes; the same-sex coefficients are almost twice the size of the opposite-sex coefficients. Lastly, the more adults in the household net of the number of employed adults, the less likely an unemployed individual is to obtain employment.

Because the survey is semi-annual, we would fail to observe any added worker effect that occurs entirely within 6 months of the job loss in the household. Such a short timeline would be consistent with one characterization of the added worker effect – that women in husband-wife pairs obtain a temporary second-best job quickly

[16] The child grant is available to primary caregivers with co-resident children under 18.

as a stop-gap measure while other household members seek better employment opportunities. Lundberg (1985) found that female labor supply increased for only two months following the shock and Spletzer (1997) found a small effect contemporaneous to the shock (rather than two months after). It is reasonable that with so much unemployment in South Africa, the added worker effect might have a smaller magnitude, but take longer to appear than it would in the U.S.

Layard et al. (1980) attribute their inability to find an added worker effect to the fact that it was swamped out by the discouraged worker effect that arises when poor labor market conditions cause market wages to fall for all members of the household, making potential added workers less likely to seek employment. The discouraged worker effect is particularly pronounced during times of economic crisis (Baslevent and Onaran 2003), but can also play a role when local unemployment rates rise (Lundberg 1985, Gruber and Cullen 2000). It is likely that in this case, the discouraged worker effect dominates the added worker effect in the short run.

I find supportive evidence for gender-specific network effects. More employed individuals in the household increase the amount of information available about local labor markets and job opportunities that filters from the workplace to the home (Dinkelman 2004). In addition, the number of firms is non-decreasing in the number of employed household members, increasing the probability that at least one of these firms would be hiring. The gender component of the network effects is unsurprising due to the fact that many occupations in South Africa remain gender specific.[17] My results are consistent with findings by Magruder (2010) that fathers serve as useful

[17] For example, construction, the industry with the highest employment growth rates during the panel, is predominantly male, whereas the many women work as domestic workers or teachers.

network connections for young adult sons but not daughters. However, this evidence of female network effects contrasts with his finding that mothers are not useful network connections for daughters. There may be greater variance in the network value of women than men due to different hiring patterns in gender-specific industries.

Theory predicts that individuals in households with more employed members are less likely to obtain employment in response to a shock because the proportional negative effect of a single job loss on the outside option is smaller. I find some evidence of this when I add the interaction (between number employed and household shock) to the regression specification. Employed workers may increase their work hours in response to a shock, exhibiting the added worker effect on the intensive margin and mitigating the effect we observe on the extensive margin. However, in a separate set of regressions, I find no evidence of a change in "usual" hours worked in response to a household job loss for either men or women (results not shown).

The best predictor of obtaining employment is having lost a job recently. For both men and women, those who lost a job between 6-18 months prior to the survey are more likely to obtain employment than those who lost a job more than 18 months earlier. This suggests a degree of negative duration dependence since those who have been unemployed for the shortest amount of time are most likely to obtain employment. However, this variable may also be picking up the effect of voluntary job separations or of selection on unobserved variables, such as taste for work.

Unemployed men in households that receive a government pension are less likely to obtain employment than those in households without pension income. This result is consistent with the findings of Bertrand et al. (2003), Edmonds et al. (2005), and Ranchhod (2009). Using the South African LFS panel data, Ranchhod (2009) examines the effect of a loss of pensioner income on labor force activity and household composition, identifying the loss of pension receipt using a survey question and using a death of a pensioner in the household. He finds that the probability of employment increased the most among women aged 36-50 and among adults of both sexes between 51 years old and the pension age (60 for women, 65 for men). His results persist when taking into account the possibility of endogenous household formation around a pension recipient. Bertrand et al. (2003) found that household members aged 16-50 reduced their labor supply in the presence of a pensioner, with the largest effect for the oldest prime-age male. Examining household formation in response to pension income, Edmonds et al. (2005) found that a female pension recipient increased the number of children under age 6, and women aged 18-23 (likely their mothers), while there were fewer women aged 30-39 present in the household.

4.6.2 Transition Regressions: Labor force participation

I now turn to results from my analysis of labor force participation in response to a job loss within the household (see Table 4.8). For men, labor force participation is consistent with the added worker effect model. Unemployed men are 11.7 percentage points more likely to transition into labor force participation in households that experience a negative household shock 6-12 months earlier, and 2.1 percentage points more likely for shocks 12-18 months earlier, however neither estimate is statistically significant. The probability of resuming labor force participation is reduced by a statistically significant 26.4 percentage points for men in households that experience negative shocks in two subsequent waves, which is substantial. However, less than

2 percent of the sample experiences two subsequent shocks. This suggests that in households where two or more individuals have lost a job recently, limited resources cannot be allocated to search efforts or if they are, they support the search efforts of the most recently unemployed. I do not observe any of these effects for women.

While there appears to be no statistically significant impact of employed men on the labor force participation of either gender, each additional employed woman in the household raises the likelihood of resuming labor force participation by 5.2 percentage points for women. Men in households with an additional adult of working age are 4.4 percentage points less likely to be labor force participants. This is likely related to the importance of network effects for men. There are clear incentives to delay investment in active job search if there are potential network benefits, or if there is a chance that another unemployed household member would make the investment in job search first (and obtain employment first).

Pension income supports the unemployed and raises reservation wages, but may also enable the unemployed to undertake an active job search. Individuals in pension-recipient households are slightly more likely to transition from discouraged into active searching, however this result is not statistically significant for men or women. Ardington et al. (2007) found that the pension allows household members to migrate to find work, specifically by providing resources to support a job search.

4.7 Discussion

The true magnitude of the added worker effect is likely larger than the estimates I have produced. The added worker effect is defined as a change in desired labor supply, but in my data I can only observe changes in search activity and employment. A number of factors falling under the category of frictional unemployment can prevent individuals from translating labor supply into employment, or even discourage increases in search activity. Obviously, in addition to desiring employment a job seeker must find an employer willing to hire them at or above their reservation wage – which may take time considering the slow employment growth in South Africa and the glut of unemployed individuals.

We would expect it to be easier to transition into search than into employment, however high costs associated with job seeking in South Africa may pose a significant barrier to an individual conducting an active search if he or she desires to do so. The majority of Black Africans live in residential neighborhoods established under apartheid that are far from business centers and that remain highly segregated (Christopher 2001, 2005). Even informal enterprises are clustered in inner-city zones, and sparser in Black African townships and informal settlement areas where there is less opportunity for economic growth (Rogerson, 1996). There are high transportation costs associated with job searching due to the spatial separation and the fact that public transportation infrastructure built under apartheid was designed to serve White areas. Transportation costs account for up to a 10 percent share of consumption for many South Africans (Klasen, 1997). Additionally, there are high screening costs (due to varying educational quality) and high dismissal costs (legal and administrative obstacles) that require job seekers to invest in the employer's interview and assessment process.

A number of studies including Wilson and Ramphele (1989) and Kingdon and Knight (2006) have found that poverty inhibits job search. The ability to search may reflect the availability of household resources to support the search, rather than imply that the searching worker is more determined to obtain employment (or that their unobserved qualities are better than those who are not searching). The finding that men are less likely to begin or resume search following two consecutive job losses in the household likely reflects the loss of household resources to fund an on-going job search, despite the fact that desired labor supply may have increased.

Two factors that may lead to an increased likelihood of obtaining employment are an increase in search effort and/or intensity and a decrease in the reservation wage, however, in this study, I am unable to distinguish the relative importance of each. The results from the search regressions demonstrate that search effort does increase following a household job loss, at least for men. The LFS questionnaire only asks about search behavior in the one month prior to each survey wave, so it appears that the increased effort may be sustained for a number of months between the household job loss and the survey date.[18] The theoretical model demonstrates that reservation wages will fall when the value of the outside option falls. Direct evidence of falling reservation wages would be evident if the accepted wage following a household job loss was lower than we would otherwise expect. However, my data do not allow me to create a reasonable counterfactual because there are no questions about reservation

[18]The LFS data is inadequate to address the effect on search intensity because individuals were only asked about the extensive margin of search behavior.

wages and too little information to control for past shocks to the reservation wage.

4.7.1 Network effects

Network effects appear to play an important role in obtaining employment. To my knowledge, no one has undertaken a comparison of the effectiveness of different job search methods in South Africa for different types of jobs or different types of job searchers. However, there are three reasons why networks are preferred to other job search methods: job search is costly, employers may prefer to use networks for hiring, and high crime rates reduce the effectiveness of other methods.

Network effects may crowd-out typical job search activities at least at first, especially if other activities are more costly. The network effect might lead an unemployed person to delay initiating a job search (and making a costly investment) if there are other employed household members who could provide network connections. This expectation may explain why, in larger households where a job loss has occurred, men are less likely to obtain employment after one period and more likely after two periods. Reliance on a network may take longer because the search is concentrated among fewer firms, but it is less costly.

Employers may rely heavily on the social networks of their employees to find qualified candidates because there are few good signals of candidate quality. Wittenberg (2002) suggests that this dependence on networks is due to the matric qualification (high school leaving exam) being a poor signal of educational quality because the quality of secondary schools varies so widely. He also posits that because such a large proportion of the jobless possess a matric, it is a useless sorting criterion. High firing costs in South Africa make it more risky to hire an employee whose productivity is uncertain and contribute to the importance of networks as a means to screen potential employees.

High crime rates in South Africa may make job searching especially difficult for young male Black Africans. Individuals and businesses may be more suspicious of non-Whites. For example, unwillingness to grant entry to non-Whites has been reported as one reason for household survey non-response among Whites in South Africa.[19] From the employers' point of view, the social networks of their employees may be a less risky pool of applicants in terms of crime.

4.7.2 Possible sources of bias

There are three possible sources of negative bias in my research design: local labor market shocks, anticipated household job loss and the substitution of a pensioner or student to provide additional income to the household. A negative local labor market shock will increase the probability of a household job loss while at the same time decreasing the probability that an unemployed individual will be able to find employment (i.e. the discouraged worker effect), thereby depressing the effect that I seek to measure. If a household job loss is anticipated by other members of the household (for example, if the employee plans to give notice or starts to experience a negative health shock), then the added worker effect would take effect before the actual job loss occurs. Thus, the change in desired or actual hours once the household job loss happens would be smaller than if the job loss were unexpected by other members in the household.

[19] Statistics South Africa, Report on Census Publicity Research Study, March 2005.

Endogenous household formation may lead to negative bias in the case of migrant workers or positive bias in the case of formation around a substitute income earner such as a pensioner or former student. In fact, the household composition variables on which I condition may themselves be endogenous. If members of the household become migrant workers to improve their employment prospects when another household member loses a job, they exit the sample, leading to an underestimate of the added worker effect. While the data set includes information on the number of migrants and the amount of their remittances, it does not conduct the entire survey with migrant workers or other past household members who are currently away from the household. This is more of a problem when using panel data because migrants would be less likely to be in two or more LFS waves and thus less likely to be included in the panel. Part of the response to a household job loss may be household formation around a pension recipient as a substitute for increased labor supply. Alternatively, children and youth may be pulled out of school in response to a household job loss, and/or encouraged to find employment. These individuals would not be in my sample if they went directly from NEA (a student) to employed between survey waves so I would be underestimating the effect of a household job loss. However, they would show up in my sample in the following period if they transitioned from NEA to unemployed (i.e. they were not able to find employment within 6 months of leaving school). An increase in income from a student pulled out of school or a pensioner brought into the household in response to a shock would lead to an underestimate of the effect of household job loss on the labor market outcomes of the unemployed.

Because the LFS questionnaire is not retrospective, we fail to observe spells of employment that occur during the six months between two survey waves. I estimated the likely volume of unobserved employment spells by comparing the number of individuals whose tenure is between 1-6 months with the number whose tenure is between 7-12 months. Approximately 50 percent of the transitions would be unobserved during the summer months in South Africa (September to March) when seasonal workers are employed, with between 33-40 percent unobserved transitions in the other half of the year.[20] This magnitude of unobserved employment spells would bias my results against finding an added worker effect.

I have identified two potential sources of positive bias: an exogenous change in union power and an exogenous change in wage level due to the minimum wage. Positive bias is more of a concern than negative bias at this point because I expect there to be a positive relationship between a household job loss and the probability of the unemployed obtaining employment. An exogenous increase in union power (or the initiation of a collective bargaining agreement) would reduce the probability that union employees (insiders) lose their job. At the same time, it could increase the wage level which would make it more costly for the firm to hire new workers (outsiders), thereby reducing hiring rates. An exogenous increase in the minimum wage would have a similar effect in that it could reduce hiring rates without affecting job separations, especially since labor laws in South Africa can make layoffs and dismissals fairly costly.

[20]Calculated from LFS cross-sectional data, waves 2-11.

4.8 Conclusion

I find clear evidence of an added worker effect among Black African men and some evidence of the effect among Black African women. Though men are less likely to obtain employment in the 6-12 months following a single negative employment shock to the household, they are statistically significantly more likely to obtain employment 12-18 months after a negative shock as well as after the household experiences two consecutive negative shocks. Women are slightly more likely to obtain employment only 12-18 months after a negative shock. Labor force participation is consistent with an added worker effect for men: male discouraged workers are marginally more likely to begin or resume labor force participation if they are members of a household that has experienced a recent job loss. However, if the household experienced one or more job losses 6-12 months ago and 12-18 months ago, then discouraged workers are less likely to start searching, likely due to credit constraints. Neither of these effects is observed for women. The increase in active labor force participation for men following a negative shock, appears to translate into an increased likelihood of obtaining employment.

It is clear that employment outcomes within the household are interrelated, and that one must take this into account when designing policy interventions and social insurance programs. The added worker effect literature has mainly examined married couples, but in this case I find an effect for working age males within an extended household. My findings shed light on how household formation provides insurance against negative shocks when the public safety net is limited, especially when households face binding credit constraints. I demonstrate one motivation for why households support unemployed members for extended periods of time: these members may be called upon to help provide for the household in times of economic hardship.

Future work in this area could address the role that falling reservation wages play in driving the added worker effect.[21] If individuals are optimizing, then lowering the reservation wage in response to a negative household shock will make them worse off in the long run than in the absence of the shock, even if they are able to find employment. However, the shock may lead workers (especially those who have never held employment) to acquire new information about the labor market and revise their expectations. Being employed and reversing any human capital depreciation that had occurred are positive outcomes in themselves.[22] A first step would be to examine the types of jobs and wages that are accepted under economic hardship and whether their quality differs from job offers accepted in the absence of a negative shock.

My findings demonstrate that despite an increase in labor force participation, it often takes more than a year to obtain employment when the household suffers an economic hardship. This suggests that structural and frictional factors constrain an individual's response to a household shock. Addressing some of these factors in the labor market will enable households to respond more quickly to an adverse shock and limit the negative repercussions. This is especially important for poorer households who are more credit constrained (and constrained in other ways) and thus have access to fewer methods of mitigating the effect of economic shocks. It would also

[21] And in particular, address the claim that reservation wages of young South Africans are "too high", though Nattrass and Walker (2005) find otherwise among Coloureds in Khayelitsha (Cape Town).

[22] It would be worth investigating the magnitude of the job offer arrival rate for on-the-job search.

increase the rate at which job seekers enter the workforce and shrink the pool of unemployment over time.

4.9 References

Ardington, C., A. Case and V. Hosegood. 2007. “Labor supply responses to large social transfers: Longitudinal evidence from South Africa.” Paper presented at the Twelfth Annual Conference of the African Econometric Society, Cape Town.

Ardington, E. and F. Lund. 1995 “Pensions and development: social security as complementary to programmes of reconstruction and development.” *Development Southern Africa* 12 (4): 557-577.

Baslevent, C. and O. Onaran. 2003 “Are Married Women in Turkey More Likely to Become Added or Discouraged Workers?” *Labour* 17: 439-458.

Basu, K., G. Genicot and J.E. Stiglitz. 1999. “Household labour supply, unemployment, and minimum wage.” World Bank Policy Research Working Paper 2049.

Bertrand, M., Mullainathan, S. and Miller, D. 2003. “Public Policy and Extended Families: Evidence from Pensions in South Africa.” *The World Bank Economic Review* 17 (1): 27-50.

Case, A. and A. Deaton. 1998. “Large Cash Transfers to the Elderly in South Africa.” *Economic Journal.* 108: 1330-1363.

Christopher, A. J. 2001. “Urban Segregation in Post-apartheid South Africa.” *Urban Studies.* 38(3): 449-466.

Christopher, A. J. 2005. “Does South Africa Have Ghettos?” *Tijdschrift voor Economische en Sociale Geografie*. 96(3): 241-252.

Dinkelman, T. 2004. “How Household Context Affects Search Outcomes of the Unemployed in KwaZulu Natal, South Africa: A Panel Data Analysis.” *South African Journal of Economics* 72 (3): 484-521.

Duflo, E. 2003. “Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa.” *The World Bank Economic Review* 17 (1):1-25.

Edmonds, E.V., K. Mammen and D.L. Miller. 2005. “Rearranging the family? Income support and elderly living arrangements in a low-income country.” *Journal of Human Resources*. 40(1): 186-207.

Fernandes, R. and F. de Felício. 2002. “The Entry of the Wife into the Labor Force in Response to the Husband’s Unemployment: A Study of the Added Worker Effect in Brazilian Metropolitan Areas.” *Economic Development and Cultural Change* 53: 887-911.

Flinn, C. and J. Heckman. 1983. “Are Unemployment and Out of the Labour Force Behaviourally Distinct Labour Force States.” *Journal of Labour Economics* 1(1): 28-42.

Gönül, F. 1992. "New Evidence on Whether Unemployment and Out of the Labour Force are Distinct States." *Journal of Human Resources* 27(2): 329-361.

Gruber, J. and J.B. Cullen. 2000. "Does Unemployment Insurance Crowd out Spousal Labor Supply?" *Journal of Labor Economics* 18(3): 546-72.

Heckman, J. and T. MaCurdy. 1980. "A Life Cycle Model of Female Labour Supply." *Review of Economic Studies* 47(1):47-74.

Heckman, J. and T. MaCurdy. 1982. "Corrigendum on a Life Cycle Model of Female Labour Supply." *Review of Economic Studies* 49(4): 659-60.

Humphrey, D.D. 1940. "Alleged Additional Workers in the Measurement of Unemployment." *Journal of Political Economy* 48: 412-9.

Jensen, R. 2003. "Do Private Transfers Displace the Benefits of Public Transfers? Evidence from South Africa." *Journal of Public Economics* 88: 89-112.

Kingdon, G. and J. Knight. 2006. "The measurement of unemployment when unemployment is high." *Labour Economics* 13: 291-315.

Klasen, S. 1997. "Poverty, Inequality and Deprivation in South Africa: An Analysis of the 1993 SALDRU Survey." *Social Indicator Research* 41: 51-94.

Klasen, S. and Woolard, I. 2009. "Surviving Unemployment without State Support: Unemployment and Household Formation in South Africa." *Journal of African Economies* 18(1):1-51.

Layard, R., M. Barton, and A. Zabalza. 1980. "Married Women's Participation and Hours." *Economica* 47 (February): 51-72.

Lundberg, S. 1985. "The Added Worker Effect." *Journal of Labor Economics* 3(1): 11-37.

Magruder, Jeremy R. 2010. "Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa." *AEJ: Applied Economics* 2(1): 62-85.

Maloney, T. 1987. "Employment Constraints and the Labor Supply of Married Women: A Reexamination of the Added Worker Effect." *Journal of Human Resources* 22(1): 51-61.

Maloney, T. 1991. "Unobserved Variables and the Elusive Added Worker Effect." *Economica* 58(230): 173-87.

McCall, J. J. 1970. "Economics of Information and Job Search." *Quarterly Journal of Economics* (84)1: 113-126.

Nattrass, N. and R. Walker. 2005. "Unemployment and Reservation Wages in Working-Class Cape Town." *South African Journal of Economics* 73(3):498509.

Parker, S.W. and E. Skoufias. 2004. "The added worker effect over the business cycle: evidence from urban Mexico." *Applied Economics Letters* 11(10): 625-630.

Ranchhod, V. 2009. "Household responses to adverse income shocks: Pensioner out-migration and mortality in South Africa." SALDRU Working Paper.

Rogerson, C.M. 1996. "Urban poverty and the informal economy in South Africa's economic heartland." *Environment and Urbanization* 8(1): 167-179.

Sagner, A. and R.Z. Mtati. 1999. Politics of pension sharing in urban South Africa. *Ageing and Society* 19: 393-416.

Serneels, P. 2002. "The Added Worker Effect and Intrahousehold Aspects of Unemployment." Centre for the Study of African Economies Working Paper, Series Ref: WPS/2002-14.

Spletzer, J. 1997. "Reexamining the Added Worker Effect." *Economic Inquiry* 35(2): 417-27.

Stack, C. 1974. *All Our Kin: Strategies for Survival in a Black Community.* Harper and Row, New York.

Statistics South Africa, National Statistics System Division. 2006. Measurement, Standards and Capacity Assessment, The South African Labour Force Panel Study Methodology Document.

Stephens, M. 2002. "Worker Displacement and the Added Worker Effect." *Journal of Labor Economics* 20: 504-537.

Wilson, F. and M. Ramphele. 1989. *Uprooting Poverty: The South African Challenge*. W.W. Norton & Co., London.

Wittenberg, M. 2002. "Job search in South Africa: A non-parametric analysis." *South African Journal of Economics* 70 (8): 1163-97.

Yeung, J.W. and S. Hofferth. 1998. "Family adaptations to income and jobloss in the US." *Journal of Family and Economic Issues* 19: 255-83.

4.10 Tables and Figures

Table 4.1: Employment status based on responses to survey questions

Performed job activity in last 7 days?					
Yes	No				
$\Downarrow$	$\Downarrow$				
Employed	Temporarily absent from work?				
	Yes	No			
	$\Downarrow$	$\Downarrow$			
	Employed	Able to start work in 1 week?			
		No	Yes		
		$\Downarrow$	$\Downarrow$		
		NEA	Desire employment?		
			No	Yes	
			$\Downarrow$	$\Downarrow$	
			NEA	Took steps to find work in last month?	
				Yes	No
				$\Downarrow$	$\Downarrow$
				Unemployed	**Discouraged**

Table 4.2: Labour Force Survey sample sizes by wave

Date	Wave	Sample size		% in panel
		Panel	X-section	
Sep 2001	4	60,639	106,439	56.97
Mar 2002	5	71,153	109,410	65.03
Sep 2002	6	68,050	102,480	66.40
Mar 2003	7	66,825	100,834	66.27
Sep 2003	8	58,756	98,748	59.50
Mar 2004	9	45,856	98,256	46.67

Due to rotating panel design where 20 percent of sample was rotated out each wave, the panel inclusion rate is out of a maximum of 80 percent.

Table 4.3: Likelihood of being included in panel data

Variable	Men	Women
Employed	-0.292***	-0.123***
	(0.048)	(0.041)
Unemployed (broad definition)	-0.164***	-0.099***
	(0.046)	(0.038)
Urban resident	0.098***	0.042
	(0.035)	(0.033)
Years of primary school completed	0.020**	0.017**
	(0.0087)	(0.0077)
Years of secondary school completed	0.048***	0.053***
	(0.014)	(0.013)
Completed high school	-0.075	-0.063
	(0.060)	(0.056)
Completed some post-high school education	0.289**	0.145
	(0.13)	(0.12)
Widowed	-0.217	-0.059
	(0.14)	(0.061)
Divorced	-0.213**	-0.115
	(0.11)	(0.078)
Never married	-0.131***	-0.095***
	(0.046)	(0.036)
Observations	22,923	26,279

Table presents logit coefficients with standard errors in parentheses. Sample includes Black Africans aged 16-64. Regressions include age group and province dummy variables. Wave 4 only; representative of results from other waves.

Table 4.4: Transition matrices for men by race

Black African men			Mar 2004			
Sept 2003	NEA	Discouraged	Searching	Employed	Total	*N*
NEA	75.41	8.74	9.36	6.49	100	*1,840*
Discouraged	19.34	35.98	25.29	19.39	100	*661*
Searching	14.80	17.03	45.65	22.52	100	*999*
Employed	4.32	4.68	9.10	81.90	100	*2,584*
Total	28.29	11.40	17.12	43.19	100	*6,084*

White men			Mar 2004			
Sept 2003	NEA	Discouraged	Searching	Employed	Total	*N*
NEA	76.83	0.31	1.60	21.25	100	*124*
Discouraged	11.14	12.74	17.55	58.57	100	*8*
Searching	8.26	5.62	37.58	48.55	100	*26*
Employed	4.63	0.99	1.57	92.82	100	*805*
Total	12.61	1.13	2.39	83.88	100	*963*

Sample includes ages 16-64. All values are weighted. Value in cell is proportion of individuals in row category in September 2003 who transitioned into column category by March 2004. Transition rates are broadly representative of other waves in panel. Source: Labour Force Survey panel data, wave 8 and wave 9.

Table 4.5: Transition matrices for women by race

Black African women			Mar 2004			
Sept 2003	NEA	Discouraged	Searching	Employed	Total	*N*
NEA	63.31	16.24	11.75	8.70	100	*2,411*
Discouraged	24.49	38.12	25.10	12.29	100	*1,266*
Searching	20.97	21.81	42.46	14.76	100	*1,192*
Employed	10.33	6.25	10.07	73.35	100	*2,316*
Total	32.30	18.20	19.00	30.49	100	*7,185*

White women			Mar 2004			
Sept 2003	NEA	Discouraged	Searching	Employed	Total	*N*
NEA	83.44	1.86	3.79	10.91	100	*289*
Discouraged	12.81	20.00	17.36	49.83	100	*14*
Searching	43.34	14.34	28.69	13.62	100	*26*
Employed	8.85	0.75	2.34	88.05	100	*619*
Total	31.55	1.61	3.53	63.31	100	*948*

Sample includes ages 16-64. All values are weighted. Value in cell is proportion of individuals in row category in September 2003 who transitioned into column category by March 2004. Transition rates are broadly representative of other waves in panel. Source: Labour Force Survey panel data, wave 8 and wave 9.

Table 4.6: Means of covariates: Did someone in the household transition to unemployment 6-12 months ago?

	Men				Women			
Variable	Yes	No	tStat on diff		Yes	No	tStat on diff	
Age	30.47	32.30	-1.85	*	31.98	32.83	0.41	
Urban resident	0.55	0.57	0.25		0.45	0.49	-1.55	
Number of adults in hhold	0.11	0.04	1.17		2.78	2.07	1.77	*
Number of kids in hhold	2.33	1.56	2.27	**	3.09	2.41	1.26	
Yrs. of primary education	6.29	6.06	1.68		6.14	5.98	1.50	
Yrs. of secondary education	2.34	2.28	0.10		2.35	2.33	-0.58	
Completed high school	0.18	0.23	-2.30	**	0.20	0.23	-1.07	
Some post-H.S. education	0.01	0.02	-3.00	***	0.01	0.02	-4.37	***
Pension recipient in hhold	0.20	0.15	0.56		0.19	0.20	-0.73	
Number of observations	1,996	40,182			2,188	44,661		

Sample includes Black Africans ages 16-64. All values are weighted. t-statistic on difference in means calculated conditional on the other covariates listed in the table plus age-squared. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 4.7: Likelihood of obtaining employment within past 6 months for unemployed Black Africans (mean derivatives)

Variable	(1) Men	(2) Men	(3) Men	(4) Women	(5) Women	(6) Women
$\geq$ 1 hhold member became unemployed 6-12 months ago	-0.012	-0.065**	-0.060	0.027***	0.008	0.009
	(0.014)	(0.032)	(0.031)	(0.009)	(0.023)	(0.023)
$\geq$ 1 hhold member became unemployed 12-18 months ago	-0.001	0.044***	0.043***	-0.011	0.022**	0.022**
	(0.013)	(0.015)	(0.015)	(0.009)	(0.011)	(0.011)
Hhold members became unemployed 6-12 & 12-18 months ago	0.018	0.097**	0.090**	-0.051	-0.026	-0.026
	(0.035)	(0.046)	(0.045)	(0.030)	(0.035)	(0.036)
Number of employed men in hhold		0.080***	0.080***		0.027***	0.026***
		(0.009)	(0.009)		(0.006)	(0.006)
Number of employed women in hhold		0.047***	0.044***		0.045	0.047
		(0.009)	(0.009)		(0.026)	(0.026)
Number of adults in household		-0.040***	-0.038***		-0.017***	-0.016***
		(0.006)	(0.006)		(0.004)	(0.004)
$\geq$1 became unemployed 6-12 months ago X Number employed		-0.025	-0.028**		-0.019**	-0.020**
		(0.013)	(0.013)		(0.009)	(0.009)
$\geq$1 became unemployed 12-18 months ago X Number employed		-0.029**	-0.028**		-0.010	-0.010
		(0.014)	(0.014)		(0.010)	(0.010)
Individual transitioned to unemployment 6-12 months ago		0.151***	0.148***		0.135***	0.135***
		(0.017)	(0.017)		(0.013)	(0.013)
Individual transitioned to unemployment 12-18 months ago		0.106***	0.104***		0.068***	0.067***
		(0.033)	(0.033)		(0.026)	(0.026)
Pension recipient in household			-0.053***			-0.017
			(0.012)			(0.009)
Individual and household controls	No	Yes	Yes	No	Yes	Yes
Welfare receipt controls	No	No	Yes	No	No	Yes
Number of observations	10,057	10,057	10,057	14,752	14,752	14,752
R^2	0.00	0.05	0.06	0.00	0.06	0.06

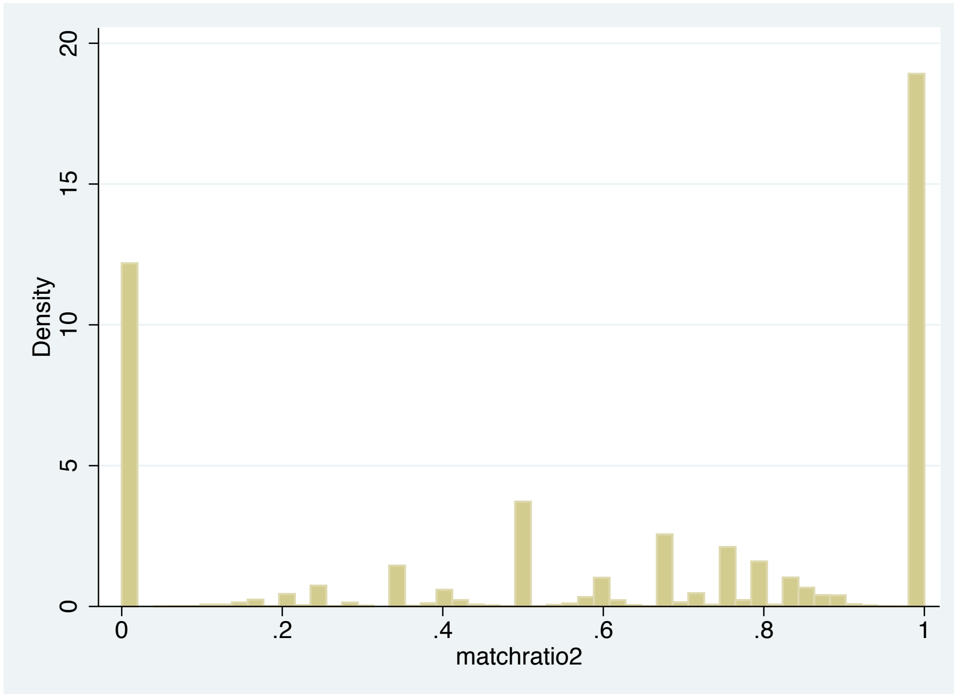
Sample includes ages 16-64, weighted. All waves pooled. Standard errors are clustered by individual. Second and third specifications include controls for urban resident, age, age squared, number of kids in household, never held a job before, never held a job before *age$\geq$ 35, province, wave, presence in $\geq$ three waves, missing hhold jobloss data for 12-18 months ago, and a constant. Third specification also includes a dummy for the household receiving a government child grant. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Table 4.8: Likelihood of becoming a labor force participant within past 6 months for unemployed Black Africans (mean derivatives)

Variable	(1) Men	(2) Men	(3) Men	(4) Women	(5) Women	(6) Women
$\geq$ 1 hhold member became unemployed 6-12 months ago	0.124***	0.115	0.117	0.097***	-0.011	-0.012
	(0.039)	(0.070)	(0.070)	(0.026)	(0.062)	(0.061)
$\geq$ 1 hhold member became unemployed 12-18 months ago	-0.002	0.023	0.021	0.044	0.045	0.044
	(0.033)	(0.033)	(0.033)	(0.023)	(0.026)	(0.026)
Hhold members became unemployed 6-12 & 12-18 months ago	-0.179	-0.263**	-0.264**	-0.044	0.035	0.037
	(0.103)	(0.127)	(0.127)	(0.069)	(0.089)	(0.089)
Number of adults in household		-0.045***	-0.044***		-0.015	-0.016
		(0.015)	(0.014)		(0.011)	(0.011)
Number of employed men in hhold		0.019	0.023		-0.007	-0.005
		(0.027)	(0.027)		(0.018)	(0.019)
Number of employed women in hhold		0.045	0.047		0.05**	0.052**
		(0.026)	(0.026)		(0.024)	(0.024)
$\geq$1 became unemployed 6-12 months ago X Number employed		0.043	0.045		-0.022	-0.022
		(0.034)	(0.034)		(0.025)	(0.025)
$\geq$1 became unemployed 12-18 months ago X Number employed		0.077**	0.080**		0.005	0.005
		(0.036)	(0.036)		(0.028)	(0.028)
Individual transitioned to unemployment 6-12 months ago		0.134***	0.134***		-0.022	-0.021
		(0.052)	(0.052)		(0.047)	(0.047)
Individual transitioned to unemployment 12-18 months ago		0.083	0.086		0.161**	0.161**
		(0.125)	(0.124)		(0.075)	(0.075)
Pension recipient in household			0.023			0.019
			(0.027)			(0.021)
Individual and household controls	No	Yes	Yes	No	Yes	Yes
Welfare receipt controls	No	No	Yes	No	No	Yes
Number of observations	2,212	2,212	2,212	4,109	4,109	4,109
R2	0.01	0.09	0.09	0.00	0.07	0.07

Sample includes ages 16-64, weighted. All waves pooled. Standard errors are clustered by individual. Second and third specifications include controls for urban resident, age, age squared, number of kids in household, never held a job before, never held a job before *age$\geq$ 35, province, wave, presence in $\geq$ three waves, missing hhold jobloss data for 12-18 months ago, and a constant. Third specification also includes a dummy for the household receiving a government child grant. *** - Significant at the 99% confidence level, ** - 95% level, * - 90% level.

Figure 4.1: Proportion of household members that are matched and included in LFS panel data



CHAPTER V

Conclusion

This dissertation provides evidence that employment dynamics in response to HIV/AIDS and intra-household transfers help explain persistently high unemployment in South Africa. The first two chapters, taken together, demonstrate that being HIV+ leads to worse employment outcomes but that providing ARV treatment for AIDS mitigates much of the negative impact. The third chapter determines that the labor market behavior of the unemployed is responsive to a reduction in intra-household transfers after a job loss within the household, which suggests that elevated reservation wages contribute to higher unemployment and longer unemployment durations in South Africa. Many households struggle to cope with both HIV/AIDS and unemployment. Better alignment and coordination of health and labor policy could improve the well-being of these households and reduce inequality in South Africa.