Can spatial mobility insure families against long-term impacts of economic shocks?

Evidence from drought and disability in South Africa

Taryn Dinkelman[1]

Dartmouth College

July 13, 2012

This paper combines variation in South Africa's *apartheid*-era mobility restrictions with variation in drought events to illustrate how spatial mobility insures against negative impacts of weather shocks on health. Using a triple difference strategy, I show that childhood drought exposure in the most restricted areas increases male disability rates by 20%. I estimate that drought induces four times more adult outmigration from areas facing weaker restrictions than from areas facing the harshest restrictions. Remittances link this differential outmigration with differential health impacts of drought. My results identify a specific channel through which barriers to spatial mobility impose long-run economic costs. [100 words]

Key words: spatial labor mobility, migration as insurance, economic shocks, long-term health outcomes, disability, drought, South Africa

JEL codes: O15, I15, J61, R23, N37

Geographic, institutional and political constraints on mobility often limit households' abilities to cushion income shocks through labor migration. Such constraints are key features of many developing countries. While barriers to spatial mobility clearly exacerbate income volatility for the poorest households and inefficiently allocate resources, little empirical evidence exists on how these barriers undermine the growth potential of economies.[2] This paper investigates one particular channel through which restrictions on spatial mobility impose long-run economic costs. I ask: do barriers to labor migration exacerbate negative effects of local weather shocks on long-term outcomes like health? Alternatively, can spatial mobility insure families against the long-run health effects of childhood exposure to local drought?

This migration mechanism is likely to be particularly important in rural areas where weather shocks strongly influence agricultural incomes. Indeed, prior research has shown how families use migration to deal with weather-induced income risk. In one of the first studies of migration as insurance, Rosenzweig and Stark (1989) structurally model how marriage migration in rural India implicitly insures households against agricultural income risk and find evidence for their model in the six ICRISAT villages. Paulson (2000) presents empirical evidence from Thailand that migrants select destinations for their potential to insure against location-specific income risk. Taking a different angle on labor mobility, Jayachandran (2006) shows that limited access to outside labor markets exacerbates negative wage effects of productivity shocks by swelling the ranks of local labor supply.[3] In this paper, I show that by helping to smooth income around the time of local shocks, migration also insures against negative long-term effects of these economic shocks.

Identifying the effects of migration on long-term outcomes is empirically challenging because the migration decision typically depends on expected returns to migration (Sjaastad 1962). Many of the relevant costs and benefits affecting this calculation are unobserved or unobservable,

---

[2] Rosenzweig (1988) writes: "Any barriers to the reallocation of labor resources accompanying economic development are potentially critical impediments to further income growth". Clemens (2011) argues that despite a lack of empirical evidence, barriers to emigration (in his case, international emigration) from poor countries are likely have first order effects welfare effects in these areas.
[3] Kaplan (2012) demonstrates that even in the US, families use changes in residence (specifically, young adults can choose to move back home) to insure against local labor market risk.

raising important endogeneity concerns.[4] Selection bias is likely to contaminate any comparisons of health outcomes across migrant and non-migrant families or across high- and low-migrant areas.[5] Likewise, it would be difficult to identify the long-term benefits to spatial mobility by studying policies that shift a country towards more spatially integrated labor markets. Such a design risks confounding spatial mobility changes with more general changes in market access.[6]

One solution to these identification challenges is to measure how the long-term effects of local shocks differ across more and less mobility-restricted areas, controlling for average differences between these areas. This is the basic research strategy of my paper. I examine what happens to long-run health outcomes and to migration in response to highly prevalent local weather shocks in a context where different areas of a country are subject to different degrees of mobility restriction. That is, given a set of externally-imposed limits on spatial mobility, I ask whether drought exposure in early childhood differentially affects later-life health outcomes (disability rates) among individuals born into differentially restricted areas. To provide direct evidence on the migration mechanism, I also ask whether outmigration responds differently to local drought across more and less mobility-restricted regions.

My analysis centers on South Africa during the *apartheid* period when a host of policies restricted African rights of movement, residence and employment in the modern sectors of the economy. During this era, the South African government consigned a majority of Africans to one of ten homeland areas, spatially isolated from the modern economy. As I explain in Section 1, legal restrictions on movement accumulated throughout *apartheid* and the oldest rural homelands– the TBVC states[7] – ended up facing the highest externally imposed barriers to permanent labor migration for the longest time. My basic identification strategy compares the effects of local drought in specific years across these TBVC and the non-TBVC homelands, controlling for differences between these areas in non-drought years and for year fixed effects.

---

[4] Uncertainty about returns in a distant location may also affect the calculation of expected returns to migration and hence migration choices, as in Bryan, Chowdhury and Mobarak (2012).
[5] De Brauw and Giles (2008) provide a good discussion of selection bias issues arising in micro-level migration studies.
[6] For example, see Redding and Sturm (2008) on the reunification of East and West Germany.
[7] TBVC stands for Transkei, Boputhatswana, Venda and the Ciskei. These areas were granted independence from South Africa in the latter years of *apartheid.*

To implement this strategy, I use 1996 Census data to measure outcomes for individuals who have ever lived in the former homelands and match this to annual measures of local weather shocks.

A crucial assumption of this research design is that (controlling for year and district fixed effects) there are no contemporaneous shocks to health outcomes or outmigration that are coincident with drought and that differ by TBVC and non-TBVC areas. Furthermore, I assume that TBVC and non-TBVC areas differ only in limitations on migration. There should be no additional unobservable differences across these areas that affect health or outmigration differently during drought. To support these assumptions, I show that socio-economic and demographic variables are on average the same across areas and I control for district fixed effects in all regressions to reduce concerns about constant unobserved differences between TBVC and non-TBVC districts. An additional strength of my design lies in the many separate natural experiments for economic shocks identified by drought events across years and districts. These multiple drought events minimize concerns that confounding shocks correlated with drought drive the results.

My estimates indicate that limits to free labor mobility exacerbate the negative health effects of early childhood drought exposure. For African males from TBVC areas, drought exposure at birth significantly raises the probability of serious disability by about 1 percentage point relative to males from non-TBVC areas. Vision and physical disabilities account for most of this disability result. Overall, drought exposure *in utero* and at ages one and two increases disability rates among males in all areas, implying that families were not able to fully insure against drought shocks in any areas. Drought-induced changes in fertility, mortality and sex selection do not account for these results.

These estimated disability effects are large and economically meaningful. Drought raises disability rates by 20% relative to the mean in TBVC areas, consistent with the literature on the fetal and childhood origins of health (e.g. Almond and Mazumdar 2011).[8] My estimates are

---

[8] Almond and Currie (2011a, 2011b) provide excellent reviews of this literature. Martorell et al (1994) and Martorell (1999) provide a general discussion of the link between nutrition, disease and adult health. Recent empirical studies

particularly striking given the youthfulness of the sample: old-age related disabilities are not driving these results for males between the ages of 10 and 48. The higher prevalence of vision and physical disabilities among drought-exposed cohorts in mobility-restricted areas undoubtedly imposes long-term costs on local economies by limiting the productivity of prime-age workers.

I show that migration is a key mechanism mitigating the long-term health impacts of drought. Using the same basic triple difference strategy, I estimate rates of permanent adult outmigration in drought years from TBVC and non-TBVC areas during *apartheid*, controlling for differences in outmigration in non-drought years and for year and district fixed effects.[9] Since there is no historical data on migration in South Africa, I present a novel way to use a cross-section of Census data to learn about migration histories. Using information about a person's "last move" and her current and prior districts, I construct a pseudo-panel dataset capturing each person's place of residence between 1948 and 1986 and the year that they moved out. I match this outmigration data to drought indicators at the district-year level and estimate how outmigration at the district-year level responds to these local shocks. I show that measurement error in migration status (which only arises for individuals who move more than once) biases outmigration and health estimates downwards with sufficiently small fractions of repeat movers, as is the case in South Africa.

My results confirm that the TBVC status of a district captures differences in external limits on free labor mobility. Drought induces significantly more adults to move away from non-TBVC areas: 0.08% more adults leave these districts during drought. This represents a more than 30% increase in outmigration relative to the annual mean. In contrast, the increase in outmigration from TBVC areas is substantially lower, at only 10% of the annual mean.

---

in the the early-life health shocks literature include: Almond (2006) on exposure to the flu epidemic, Chen and Zhou (2007), Meng and Qian (2009), Lindeboom et al (2010) and van den Berg (2011) for analyses of early-life famine exposure, and Maccini and Yang (2009) and Aguilar and Vicarelli (2012) for exposure to floods, Banerjee, Duflo, Postel-Vinay and Watts (2010) for exposure to agricultural income shocks. Jensen (2000), Hoddinott and Kinsey (2001) and Alderman, Hoogeevan and Rossi (2009) analyze immediate and medium-term negative impacts of drought on child nutrition and schooling outcomes but do not consider migration as a mitigating mechanism.
[9] The Census does not collect information on temporary migration.

The primary connection between permanent adult outmigration and infant health in households left behind during drought is through remittances.[10] Migrant incomes earned in distant, unaffected labor markets can be remitted back home to smooth incomes in anticipation of or in response to income shocks (Rosenzweig 1988). While historical data on remittances during *apartheid* does not exist, data from more recent years reflects the continued importance of these migrant and remittance linkages in South Africa. In 1996, over 30% of households in former homeland areas contained a migrant worker.[11] Since these migrant networks are entrenched in the former homeland areas, I use the same Census data to show that remittances respond to drought in 1995. Controlling for historical drought prevalence in each district and the presence of a migrant worker in the household, I estimate that households in former homeland districts are between 13% and 16% more likely to receive remittances in 1995 if there was a drought in their district in 1995. Remittance flows into households with migrant workers drive this effect. The effect is slightly stronger into households located in non-TBVC areas, reflecting the stronger historical migrant links between non-TBVC areas and the rest of the economy. These remittance results suggest that adult outmigration from homelands during *apartheid* likely facilitated flows of money in the opposite direction and especially in response to drought.

My paper makes four main contributions. First, I extend the development literature on migration as insurance by showing that limits on spatial mobility negatively affect long-term outcomes of considerable interest: the prevalence of serious health disabilities. My results motivate an even larger role for spatial mobility as insurance than the literature has previously emphasized. Second, I connect this migration as insurance literature with the economic literature on early-life health shocks by highlighting migration as one specific mechanism through which families mitigate the impact of negative economic shocks on children.[12] Third, I provide evidence on the

---

[10] Rosenzweig (1988) discusses remittances in his review of labor markets in developing countries; Yang (2011) provides an overview of the remittance channel in the context of international migration. There is evidence from the international migration literature that remittances smooth incomes in Botswana (Lucas and Stark 1988), Mexico (Amuedo-Durantes and Ponzo 2011) and the Philipines (Choi and Yang 2007).

[11] In household survey data from 2002, over three quarters of rural black households containing a migrant worker received remittances and over one third of these households reported remittances as the main source of income (Posel and Casale 2006).

[12] Strauss and Thomas (1995) discuss parental decisions about child health investments in response to shocks although there is as yet little work on these mitigating mechanisms in the early-life health shocks literature.

specific channels through which spatial mobility can insure against the impact of local income shocks: adult outmigration and remittance flows in response to drought. Fourth, I exploit exogenous variation in economic shocks and plausibly exogenous variation in mobility restrictions to establish the long-term health effects of limited mobility and to identify the migration channel. The South African case allows me to focus on mobility restrictions that are exogenous to the individual thereby getting around some of the difficult selection issues associated with migration. This context presents a unique opportunity to identify one specific channel through which barriers to spatial mobility impose long-run economic costs.

My results also have implications for policy. *Apartheid*-era restrictions on labor migration represent one example of spatial segregation. Similar legal restrictions have controlled internal migration in China, Israel, Malaysia, Russia, and the Ukraine in the past. Most countries restrict travel and employment for foreigners use external passport and visa controls. In other settings, the nature of the physical terrain or inadequate transport infrastructure spatially segregates labor markets. The empirical evidence from South Africa suggests that spatial integration of labor markets may generate significant welfare gains for poor countries.[13] By enabling family members to more easily work in distant labor markets and send remittances homewards, such spatial mobility may reduce the negative health impacts of highly prevalent environmental events like drought.

The paper begins by setting out some of the historical background on labor mobility restrictions in South Africa. Section 2 describes my identification strategy, Section 3 describes data, key variables and measurement error issues and Section 4 presents the main results for disability, outmigration, remittances, and composition effects. Section 5 concludes.

---

[13] This point is related to Burgess and Donaldson (2010), who suggest that price reductions and quantity volatility brought about by more open markets can theoretically mitigate the impact of negative weather events. However, when farming households subsist on agriculture or livestock, product market integration is unlikely to provide substantial protection against large negative income effects associated with drought. Labor migration may therefore help families respond to drought even if product markets are integrated.

1. **Labor markets and restricted mobility in South Africa during *apartheid***

Throughout the years of formal *apartheid* (1948-1994), Africans were never entirely free to move around the country for work or other reasons.[14] The white government implemented highly organized systems of documentation and control to allocate African labor to white firms, farms and households. For example, all Africans were legally required to carry pass books with up-to-date information on work and travel permissions and to produce them on demand. Job seekers were required to register at local labor bureaux for permission to work rather than being able to search freely in the labor market.

The creation of ten homeland "states" or Bantustans during the 1950s, 1960s and 1970s was a crucial part of this strategy to control mobility. These states provided a space to locate excess labor and non-labor population resources (women, children and the aged) far from urban centers, in rural parts of the country.[15] Several million Africans were resettled in these homelands during the 1960s (see Simkins 1983 for an overview). My empirical analysis focuses on how these homeland residents were differentially able to respond to local drought.

Africans had limited rights to live or look for work outside of an assigned homeland area. They were prohibited from migrating freely between homelands and between homelands and urban areas. Permissions for labor migration were typically granted in response to labor demand from the urban and white economies: for example, large numbers of unskilled workers were drafted into mine work (e.g. Wilson 1972), the manufacturing sector or domestic services at the whim of the white economy.[16] Living standards among those individuals remaining in the homelands were poor enough that that drought shocks could seriously affect the nutritional environment,

---

[14] Wolpe (1988), Simkins (1988), Lemon (1984) and Savage (1986) for discussion of the central policies of population control under *apartheid*

[15] Lemon (1984) writes "Probably no avowedly capitalist country controls its labor market to the same degree as South Africa….State restrictions on freedom of movement continue to hinder Africans in particular from selling their labor freely." Describing twentieth century population distributions, Simkins (1983) concludes that South Africa was under-urbanized relative to other countries at the same level of economic development in the early 1980s, largely due to the policies of spatial segregation and labor mobility restrictions.

[16] In Secretary for Bantu Administration and Development General Circular No. 25 (1967), "1. It is accepted Government policy that the Bantu are only temporarily resident in the European areas of the Republic, for as long as they offer their labour there. As soon as they become, for some reason or other, no longer fit for work or superfluous in the labour market, they are expected to return to their country of origin or the territory of the national unit where they fit in ethnically if they were not born and bred in the homeland."

compromising health capital accumulation. As late as the 1990s over half of rural African adults still consumed under 2,100 calories per person per day (Wilson 1996).[17]

The four earliest homelands (Transkei, Boputhatswana, Venda and Ciskei, or TBVC areas) were formalized by 1962. QwaQwa, KwaZulu, Gazankulu, Lebowa, Kangwane and KwaNdebele (the remaining non-TBVC homelands) were established much later in the period.[18] I treat the four TBVC areas as districts that faced the greatest restrictions on free movement of labor for the longest time. Not only were these states the oldest homelands, they were eventually granted political independence from South Africa. In certain years, TBVC residents even required passports to enter South Africa (Savage 1986).

Figure 1 shows these ten areas scattered throughout the country. I used maps like Figure 1 (as well as more detailed maps) and ArcGIS software to spatially identify which districts covered a "majority homeland or TBVC area" and assigned each district d a value of $TBVC_d$=1 or 0 based on this spatial match. The Data Appendix describes this assignment in more detail. The TBVC indicator is the broad measure I use through much of the analysis to represent external restrictions on spatial mobility. Since this indicator does not vary over time, I also experiment with a narrow definition of TBVC status that takes a value of one in TBVC districts only during their years of independence, and is otherwise zero. My outmigration results are robust to this more narrow definition.

2. Empirical strategy

The empirical challenge in this paper is how to identify whether spatial mobility helps families to mitigate the long-term health impacts of drought. I do this by estimating the impact of local shocks on health outcomes and on outmigration across differentially restricted labor markets in a difference-in-difference-in-differences research design.

---

[17] Using data from the South African National Income Dynamics Panel, Mariotti (2011) shows that an acute income shock (related to labor demand) in some of the homelands in the mid-1970s led to height improvements among African men.
[18] Transkei, Boputhatswana, Venda and the Ciskei were established in 1959, 1961, 1962 and 1961 respectively and legally granted independent status in (order) October 1976, December 1977, September 1979 and 1981. The remaining homelands were established: QwaQwa (1969), KwaZulu (1970), Lebowa and Gazankulu (1971), KaNgwane (1976) and KwaNdebele (1977).

For the health analysis, I use Census information on each person's prior district of residence to estimate the effect of drought exposure in the year and place of birth on disability later in life.[19] Using individual-level data on disability and drought exposure and controlling for district μ_d and birth year ϕ_t fixed effects, I estimate:

$$Y_{jdt} = \delta_0 + \delta_1 DROUGHT_{dt} + \delta_2 DROUGHT_{dt} * TBVC_d + \mu_d + \phi_t + \omega_{dt} \qquad (1)$$

where Y_{jdt} indicates the disability status (or number of disabilities) of person j born in district d in year t. $DROUGHT_{dt}$ indicates whether district d experienced a drought in year t and $TBVC_d$ indicates whether the district falls within the boundaries of a TBVC state or not.[20] I estimate this main specification using the full sample of males and females. I also estimate (1) for male and female subsamples because prior research documents the sensitivity of males to early-life nutritional insults (Almond and Mazumdar 2011, Almond 2006; see Almond and Currie 2011a for a review). Since exposure to nutrition deficiencies and disease shocks at various young ages can undermine long-term health, I also estimate an expanded specification that controls for drought exposure *in utero* and at each age up to age four and the interactions of each exposure variable with a $TBVC_d$ indicator.

Birth year fixed effects in equation (1) account for age effects in health outcomes and for any contemporaneous national shocks relevant for these outcomes. District fixed effects additionally control for constant unobservable differences between districts that may affect health. For example, some districts may be drought-prone, have different access to public health facilities, or different approaches to child health investments.[21] The parameters of interest are δ_1, the effect of drought in the year and district of birth in non-TBVC areas and δ_2, the differential effect of

---

[19] The assignment of each individual to their birth district is discussed in the Data Appendix.

[20] The Census data do not report month of birth information even though this was asked on the Census form. However, the annual measure of drought is a reasonable measure of exposure in early life, since precipitation deficits measured over a longer period of time more accurately reflect conditions of drought than acute rainfall shocks over shorter periods.

[21] In practice, all of the homelands suffered from a dire under-provision of public health-care facilities. Coovadia et al (2009) report that in the 1970s, the homelands had a 1:15 000 doctor:patient ratio. Brown (1987) describes how health expenditures were not prioritized in homeland areas.

drought exposure in the birth year in TBVC relative to non-TBVC areas. The within-district variation in birth timing relative to drought identifies δ_1 and δ_2.

If drought negatively affects health in both areas, disability rates will be higher after early-life drought exposure ($\delta_1 > 0$ and $\delta_2 > 0$). If families in TBVC areas are less able to respond to drought, then δ_2 will be larger than δ_1. Of course, if drought is severe enough to affect fertility, child mortality or sex composition, then selection effects may dominate and alter the signs of δ_1 and δ_2. I explore these possibilities in the final section of the paper and find no evidence that composition effects drive the main results.

For the migration analysis, I specify a similar regression to (1) at the district-year level. I use Census data on each person's district of residence between 1948 and 1986 (data are described in detail in the next section) to construct district-year (dt) level measures of outmigration. I estimate the following equation for the percent of adults in district d in year t migrating away in year t ($PERCENTMOVE_{dt}$), controlling for district λ_d and year δ_t fixed effects[22]:

$$PERCENTMOVE_{dt} = \beta_0 + \beta_1 DROUGHT_{dt} + \beta_2 DROUGHT_{dt} * TBVC_d + \lambda_d + \delta_t + \varepsilon_{dt} \qquad (2)$$

where $DROUGHT_{dt}$ and $TBVC_d$ are defined as in equation (1). Since migration may respond in anticipation of or in response to a drought, I also estimate an extended specification adding indicators for drought in the year after t and drought in the year before t and the interaction of each measure with $TBVC_d$.[23]

Year fixed effects included in equation (2) control for year-specific common shocks to outmigration, for example, a national drought, or the nationwide intensification of pass law enforcement (see Lemon 1984 for trends in pass law enforcement). District fixed effects control for persistent level differences in unobservable characteristics affecting outmigration across

---

[22] Rather than collapse data to the magisterial district level (an administrative unit demarcated in the 1996 Census), I use a higher level of aggregation (the district council) to ensure there are enough observations in each geographic unit. More details on all datasets, data aggregation and variable construction appear in Data Appendix 1.

[23] Migration may respond in advance of a drought if individuals have some idea of when the next drought might occur (e.g. if some part of drought is cyclical) and if external limits on mobility mean they cannot migrate when they want to but rather must take opportunities for working outside homelands as they arise.

districts. For example, pre-existing migrant networks may be stronger in some areas. The key parameters of interest are β_1, the effect of drought on outmigration from non-TBVC areas and β_2, the differential effect of drought on outmigration from TBVC relative to non-TBVC areas. Given the discussion in Section 1, I expect $\beta_1 > 0$ and $\beta_2 < 0$ for the *apartheid* period. Such a pattern of signs would confirm that TBVC residents faced real differences in spatial mobility restrictions.

Throughout, I correct standard errors for heteroscedasticity and arbitrary forms of serial correlation at the district level by clustering on the latitude and longitude of the district (Cameron, Gelbach and Miller 2006). This follows one of Bertrand, Duflo and Mullanaithan's (2004) suggestions for the issue of serial correlation in errors in a difference-in-differences setting.

The key identification assumption in equations (1) and (2) is that there are no contemporaneous shocks to outmigration or health during drought years. This assumption rules out (for example) labor demand shocks emanating from the predominantly white economy that fall disproportionately on either TBVC or non-TBVC districts in a drought year.[24] The no contemporaneous shocks assumption is reasonably defensible since many separate natural experiments (multiple drought events in different districts across many different years) identify the effects of drought.

Causal estimates of the differential effects of drought on disability and outmigration across TBVC and non-TBVC areas provide reduced form evidence for the impact of spatial mobility on long-term health. For outmigration to be the key channel through which families mitigate the impact of drought on infant health, I additionally assume that (after controlling for district fixed effects) TBVC and non-TBVC residents differ only in the limits they face on free labor mobility. The historical context discussed in Section1 provides initial motivation for this assumption.

---

[24] Mariotti (2011) analyzes one such shock: she looks at the impact of an acute shock to labor demand from the Transkei generated by increased demand for domestic labor on the gold mines of the Witwatersrand. However, this positive labor demand shock affected some of the TBVC areas rather than the more spatially integrated non-TBVC areas.

Table 1 goes further to provide evidence that while TBVC and non-TBVC areas appear to have different access to external labor markets, they are not obviously different in ways that have consequences for child health. The adult sample is balanced on age (23 years), gender, education (7 years of schooling) and overall rates of serious disability (5.5%) across TBVC and non-TBVC areas. Mental disability is the only indicator that is slightly higher on average in TBVC areas (0.008 versus 0.007) and statistically significant at the 5% level. Female cohorts who have completed childbearing by 1996 have similar fertility rates across areas (4.7 children per woman) and child mortality rates (1 child per woman). If TBVC and non-TBVC areas differ in access to health care facilities, attitudes towards child health investments, or incomes, we would have seen larger and more significant mean differences between areas in this table.

At the individual level, similar fractions of people report ever moving away from a prior district (6.3% versus 6.1%) but people move out of TBVC areas significantly later (on average). At the district-year level, the percent of adults moving out of a TBVC area in any given year is significantly lower (0.21%) compared with non-TBVC area outmigration (0.25%), reflecting the higher restrictions on mobility faced by TBVC residents. Finally, drought prevalence is somewhat higher in TBVC than non-TBVC areas, although this difference is only significant for district-year comparisons (drought occurs in 6.8% of years for TBVC areas and 4.1% of years in non-TBVC areas, p-value = 0.080). This would be one of the differences between areas accounted for by including district fixed effects in equations (1) and (2).

The final piece of my empirical strategy addresses the question of whether remittances plausibly link adult outmigration from homelands with improved household resources for those left behind during drought. There is no data on remittance flows during the *apartheid*. However, long-standing migrant networks in the former homelands allow me to estimate the remittance response to drought using data from more recent years. I look within households situated in former rural homelands in 1996 and investigate whether remittances are more likely to flow into households with a migrant worker than into households without a migrant worker after a drought in 1995. I control for differences in remittance receipts between households in districts that did not experience drought in 1995. Since I can only exploit cross-sectional variation in drought prevalence in this specification, I also control for historical drought prevalence to soak up district-level unobservables affecting remittances. Under the assumption that historical drought

12

prevalence adequately controls for unobservable differences between districts that affect remittances, I can treat drought in 1995 as uncorrelated with remaining unobservable factors. I estimate this difference-in-difference regression for all households in the former homeland areas as well as separately for TBVC and non-TBVC areas. The results provide evidence that remittances link outmigrants with the households they leave behind, and especially so during drought.

3. Data and key variables

i.   *Measuring drought*

Figure 2 shows the distribution of droughts during *apartheid*. This is the main source of variation identifying the immediate impacts of local shocks on outmigration and the long-term effects on disability rates across TBVC and non-TBVC areas. Each bar represents the fraction of TBVC and non-TBVC districts experiencing a local drought in a given year. There is substantial variation over time: some years are entirely drought-free (e.g. 1975) while in other years (the early 1980s) over 30% of districts experience drought. In most years, a smaller, positive fraction of districts experience drought.

I use rainfall data from over 1,000 weather station locations to construct a district-year specific drought measure using the Standardized Precipitation Index (SPI) (McKee, Doesken and Kleist 1993).[25] The SPI measures the probability of observing a recent rainfall event based on the distribution of all rainfall events for a given time scale and place. It characterizes South African droughts well (Roualt and Richard 2003). Following the climatological literature, I define $DROUGHT_{dt}$ in each district d and year t to be 1 for values of the SPI below -1.5 and 0 otherwise (McKee et al 1993).

Much work in economics uses rainfall shocks to proxy for short-run income shocks. I focus on drought rather than rainfall shocks because it is Africa's most prevalent natural disaster (Bensen and Clay 1993). Furthermore, South Africa's staple crop (maize) is rain-fed rather than irrigated.

---

[25] There is no consensus on how the onset, duration or completion of a drought should be marked (Wilhite, 2001; World Meteorological Organization 2006), however, the climatological literature has shown the robustness of the SPI in capturing precipitation deficiencies that extend over time (Roualt and Richard 2003).

Limited water availability reduces maize output by interrupting growth at several points in the growing season (Le Roux 2009). Insufficient rainfall over an extended period has particularly negative consequences. As I show in Data Appendix 1, maize yields appear more sensitive to rainfall deficiencies than to rainfall excesses. Drought is therefore a relevant measure of an important local economic shock in South Africa.

ii.   *Measuring disability, fertility and population composition using Census data*

The health analysis sample comes from the 10% individual record data from the 1996 South African Census. The sample consists of African individuals born between 1948 and 1986 (age 10-48) whose current district (for never movers) or prior district (for movers) is in a rural TBVC or non-TBVC homeland. Individuals report whether they have any serious disability and the type of disability: vision, hearing or speech, mental or physical disability (e.g. paralysis). I construct an indicator for "Any serious disability?" and a "Number of serious disabilities" variable for the main analysis.

Table 1 Panel A shows disability prevalence in the sample. 5.2% of individuals report a serious disability and the largest disability is vision-related (2.3%). 1.1% of the sample has a hearing/speech disability and 1.4% have a physical disability; mental disability is reported at lower rates (0.7%). Across TBVC and non-TBVC rural areas, differences in disability rates are relatively small and not statistically significant, except for mental disabilities. Recall that the sample includes individuals aged 10 to 48 in 1996, meaning that these disability rates do not merely reflect diseases of old age.

Disruptions to the nutritional and disease environment of gestating women and young children can affect fertility timing, fertility outcomes, child mortality and sex composition of the population (Almond and Mazumdar 2011, Martorell 1994). Changes in any of these factors could affect the composition of my sample and potentially drive results. I look for evidence of these composition effects by estimating the impacts of drought in TBVC and non-TBVC areas on the number of children ever born and the number who have died. For this exercise, my sample consists of women who have likely completed childbearing by 1996 (women ages 40 to 60). For

each woman, I compute the fraction of her childbearing years (1951-1996) exposed to drought.[26] Conditional on district and year of birth fixed effects, I estimate whether this drought exposure measure predicts differences in total fertility and total child mortality for women from TBVC and non-TBVC areas. Table 1 Panel A shows that drought exposure in this sample of women is about the same across TBVC and non-TBVC areas. Average total fertility rates are high at 4.7 children per woman. Child mortality is also high: the average number of children who have died (among women with any children) is 1. These mortality rates are slightly higher in TBVC areas relative to non-TBVC areas, but the difference is not statistically significant.

I check whether drought could affect total cohort size or cohort sex composition by estimating regressions similar to (2) for the outcomes of log total males and log total females born in a given district in a given year between 1948 and 1986. Even though differential outmigration across TBVC and non-TBVC areas in response to drought could plausibly change the composition of mothers left behind (and possibly the quality of births), I find no evidence for drought affecting the composition of surviving cohorts.

iii.   *Measuring outmigration using Census data*

Migration is an inherently difficult variable to measure well, especially using household survey data.[27] Demographers instead characterize migration rates using Census data. The benefit of Census data is that it provides comprehensive coverage of migrant groups across the country, unlike general household surveys that draw from a subset of districts. The coverage of the Census also allows aggregation of migration data to broader geographic units (for our purposes, the district level). But, can cross-sectional Census data be leveraged to understand the dynamic process of historical migration?

I model how outmigration responds to local shocks using a novel way of combining migration information in South Africa's 1996 Census with assumptions about the extent of repeated migration during *apartheid*. I construct a pseudo-panel dataset of individual-year observations

---

[26] I use the prior district designation as the relevant location for drought exposure.
[27] No South African household surveys or Census data capture comprehensive migration histories (Casale and Posel 2006).

capturing where each person lived in each year using the Census questions: "Where do you live now? Where did you live before this? What year did you move here?" This dataset indicates whether a person moved out of a given district in any given year, based on the last move data. Within this pseudo-panel dataset, I observe the total number of adults in each district in each year between 1948 and 1986 and the number who move away from each district in each year. I use this to describe historical outmigration from each district in each year.[28] That is, I collapse the data to district-year level and generate the percent of adults living in each district that migrated away in each year (details are described in the Data Appendix). This *PERCENTMOVE* variable is the main outcome in equation (2).

While the "last move" data contains rich information on duration of residence information, it possesses some unusual properties. Earlier migrations are rendered invisible by migrations later on (Schmertmann 1999). Because of the design of these Census questions, the data do not allow us to observe multiple moves. I deal with this important limitation in three ways.

First, to make progress with the empirical analysis I assume that no individual ever moves more than once. This strong assumption implies that Census outmigration histories are complete (Schmertmann 1999). As long as the identification assumptions discussed in Section 2 are satisfied, OLS regressions of equations (1) and (2) provide consistent estimates of the parameters of interest. All of the main results are presented under this assumption.

Second, I consider what happens when this assumption of "no multiple moves" fails. I characterize the problem as one of measurement error in binary variables.[29] Misclassification of migration occurs when a person moves more than once. As long as this error is uncorrelated with

measured drought exposure and drought itself is measured without error, estimates from the migration analysis in (2) will be downwards biased (Bound, Brown and Mathiowetz 2001) while estimates from the health analysis in (1) are unaffected.[30]

However, in some years, drought misclassifications will be correlated with measurement error in migration status since drought is measured at the district-year level. With measurement error in drought, equation (1) underestimates the impact of drought on disabilities. The bias in the outmigration regressions is more complex because of the correlation between measurement error in the dependent and independent variables in equation (2).

I develop intuition for the form of this measurement error bias in a simple difference-in-differences regression of outmigration on drought (i.e. ignoring the TBVC interaction term) in the Measurement Error Appendix. In this simple setting, the fraction of people who move more than once (multiple movers) and the observed fraction of drought exposures drive the measurement error bias in the drought coefficient. As long as there are few multiple movers and a small fraction of individuals exposed to drought, this simple difference-in-differences comparison always underestimates the impact of drought on migration. Recent data from two different household surveys indicates that multiple permanent movements are quite rare even in 2008 (between 1% and 13% of adults are multiple movers). Multiple movements during *apartheid* were likely even more rare, given the system of legal restrictions on mobility.

Unfortunately, the analytical approach to deriving the measurement error bias does not extend easily to the triple difference specification in equation (2). As a third and final strategy, I implement a specific robustness check for the outmigration results motivated by the intuition that last move data always underestimates actual migration (Schmertmann 1999) and that last move data is more accurate for recent moves. I restrict the sample to a later period so that last move outmigration data should be more accurate. Reassuringly, the main migration results stand up to this robustness check. It appears that measurement error in migration is unlikely to be large enough to overestimate the effect of drought on migration.

---

[30] If drought exposure itself is mis-measured, the bias caused by misclassification of a binary independent variable is still downwards as long as the total misclassification rate is less than one (Aigner 1973).

Figure 3 uses the district-year level panel data to show the percent of adult's permanently leaving TBVC and non-TBVC areas between 1948 and 1986. The maximum value for outmigration is small in level terms. Only 0.6% of the district-level adult population migrates away in later years. There is also a strong upward trend in outmigration rates over time. Importantly, Figure 3 shows that outmigration from non-TBVC areas is higher than outmigration from TBVC areas in almost all years; consistent with the idea that spatial mobility was more restricted in TBVC districts under *apartheid*.

4. Main Results

i. Effects of early childhood drought exposure on disability

Table 2 presents the main results from estimating the triple difference regression in equation (1). For each of the outcomes "Any serious disability" and "Number of serious disabilities", I present the full sample results in columns (1) and (2) and results for male and female samples in columns (3) to (6).

Results from the full sample indicate that a person exposed to drought in his birth year has a 0.1 percentage point higher chance of a serious disability and 0.016 more disabilities. The effects of drought exposure on disability are three to four times larger for individuals born into TBVC areas. While neither the main nor the interaction effects are precisely estimated in the first two columns, the drought parameters are jointly significant at the 10% level. The effect of drought on disability rates in TBVC areas (the sum of the β_1 and β_2) is significantly different from zero at the 5% level.[31]

Columns (3) through (6) show that the male sample drives these disability results. Drought in the birth year raises the rate of male disability by about 1 percentage point; it raises the number of serious disabilities by almost 0.01. These male results are precisely estimated and statistically significant at the 1% level. They are also economically meaningful magnitudes: relative to the

---

[31] These disability results are largely robust to aggregation to the district and year level. They are also not driven by increased disability rates among current or former mine workers. Results are robust, and sometimes more precisely estimated, when estimated for the sample of men excluding men who report current or former occupation is in mine work (about 10% of the male sample).

average level of disability in the male sample, drought in TBVC areas raises male disability rates by almost 20%. We can reject the hypothesis that drought has no effect on the chances of disability and the number of disabilities for males later in life at the 1% level (*p*-value of the joint test for each outcome is 0.000 and 0.001).

In contrast, drought in the birth year does not have similar impacts on disability rates among females from any homeland area. I cannot reject the joint hypotheses that drought has no impact on female disability rates (*p*-value of the joint null is 0.33 and 0.39 for any disability and number of disabilities respectively). I estimate the pooled regression with a full set of gender interactions (results not shown) and strongly reject the equality of male and female *Drought*TBVC* coefficients (*p*-values = 0.000 and 0.001 for each outcome). These differences between male and female disability effects of drought are consistent with theories that male infants are more fragile than female infants (Barker, 1995; Kraemer 2000). For the rest of the disability analysis, I focus on understanding these results for the male sample.

Table 3 explores the effects of drought exposure at various critical ages in early childhood. I present estimates using the extended disability specification controlling for drought exposure at each age from *in utero* up to age four and all interactions of exposure variables with $TBVC_d$. Every drought coefficient and drought interaction term is positive for both outcomes. This means that drought exposure under age five generally undermines male health in all areas (*p*-value for the joint test is 0.000). In particular, drought exposure *in utero*, at age one and age two increases male disability rates in both TBVC and non-TBVC areas. *In utero* exposure alone raises the likelihood of disability by a statistically significant 0.75 percentage points. Yet, males born into TBVC areas are still significantly worse off when drought occurs in their birth year: they are 0.8 percentage points more likely to have any disability and have 0.0079 more disabilities.

Table 4 explores which component of male disability is most sensitive to drought exposure around birth. Going back to the basic specification, I present results for each of the component disabilities (vision, hearing/speech, physical and mental) in columns (1) to (4) of the table. All interaction terms are positive, indicating that drought exposure around birth raises disability rates even more in TBVC areas. The strongest results come from vision and physical disability outcomes. For each of these disabilities, the total effect of drought on TBVC cohorts (i.e. the

sum of the main effect and the interaction effect) is statistically different from zero (*p*-value is 0.021 for physical disability and 0.002 for vision disability). These component effects are also large relative to mean levels of sight or physical disability: drought exposure at birth raises both types of disabilities by about 20% in TBVC areas.

Consistent with prior evidence on the impact of childhood exposure to shocks on later life health, I find that early-life drought exposure has significant negative impacts on the prevalence of disabilities for African males from former homeland areas. These negative effects are concentrated in cohorts exposed to drought *in utero* and at ages one and two. However, this drought exposure accounts for an even higher fraction of total disability for African males born into TBVC areas, particularly for drought exposure in the year of birth. My results likely underestimate these disability effects because the Census does not record birth district, introducing measurement error into the drought variable. These differences in health responses to drought across TBVC and non-TBVC areas imply far-reaching consequences of limits to spatial mobility. The next section provides direct evidence on the migration mechanism at work in non-TBVC areas.

ii. Evidence on mechanism: Effect of local drought on adult outmigration

Section 1 argued that the main difference between TBVC and non-TBVC rural homelands was the intensity and duration of external mobility restrictions imposed in each group. Table 5 confirms this interpretation. I show the results from estimating the basic specification in equation (2) for the full sample period (1948-1986) in column (1) and the extended specification in column (2). The dependent variable is the percent of adults moving out of a district in a given year.

The table clearly shows that drought induces more outmigration. In all areas, outmigration is higher in drought years and in years following and (to a lesser extent) preceding a drought. In column (1), the estimate of β_1 indicates that drought induces 0.088% more outmigration. Relative to mean outmigration at the district-year level (0.24%), this is a 36% increase in adults moving out. Outmigration in the year following a drought (in column (2)) is also high at 0.099% while outmigration in anticipation of a drought is about half this size (0.047%) and marginally

significant. These positive impacts of drought on permanent outmigration from non-TBVC areas are reasonably large and precisely estimated even after controlling for a full set of year and district fixed effects.

In contrast, drought has a significantly smaller impact on outmigration from TBVC areas. The estimates of β_2 in columns (1) and (2) are large, negative and statistically significant in a drought year. The total effect of drought on outmigration from TBVC areas is the sum of the main effect and interaction term: 0.021% (0.088 - 0.067) or a 10% increase in outmigration relative to the mean. A muted migration response to drought from TBVC areas is also evident in the year following a drought (column (2)).

Next, I test for differences in migration responses using a more narrow definition of TBVC areas. In columns (3) and (4), TBVC takes a value of one in the years during which the relevant homeland is legally independent from South Africa and otherwise it is zero. This narrow version of TBVC status captures an even starker difference in external constraints on mobility, since TBVC residents required a passport to enter South Africa during the years of independence. The basic pattern of outmigration response to drought persists when switching to this alternative definition. Outmigration from non-TBVC areas is higher by between 0.061% and 0.082% in drought years and in years following a drought while outmigration from TBVC areas exposed to drought is significantly lower.

Finally, I test the sensitivity of the results to measurement error concerns. Columns (5) and (6) present results for the same regressions, restricting the sample to the later period 1960-1986.[32] Even though some statistical power is lost due to a shrinking sample, the messages from the more recent sample are overwhelmingly the same. The percent of adults out-migrating from a homeland district is significantly higher during a drought but this effect is almost entirely confined to non-TBVC areas. For example, in column (5), drought raises outmigration from a non-TBVC area by 0.066%, while in TBVC areas the effect is significantly smaller, at only 0.026%. The robustness of the outmigration response to drought in these final two columns of

---

[32] It is not sensible to test for differential outmigration responses in the year before the Census, since the legal context governing spatial mobility under *apartheid* had changed by 1995.

Table 5 reassures us that measurement error bias arising from misclassified migration status (for multiple movers) is likely small.[33]

Overall, Table 5 shows that non-TBVC homeland residents coped with drought by exploiting their relatively lower external limits on migration. An important question is whether remittances from these labor migrants could have supplemented the resources of households left behind, thereby insuring against negative health impacts of drought. In the next section, I explore this question using available data from more recent times.

iii.   *Evidence on mechanisms: Effect of drought on remittances*

Table 6 shows that drought induces remittance flows towards households in rural ex-homeland areas and that migrant workers facilitate these flows. The table presents results from household-level regressions of remittance receipts ("Any remittances") in 1995 on district-level drought in 1995, on an indicator for whether the household has a migrant worker attached, and on the interaction of these two variables. Because households with and without migrant workers could differ on a number of unobservables, it is important to control for the presence of a migrant worker attached to the household. I also control directly for the fraction of drought years experienced in each district between 1948 and 1995 to account for differences in remittance behavior induced by the "drought-proneness" of an area. Column (1) presents results for the entire sample, columns (2) and (3) break this out for non-TBVC and TBVC districts, and column (4) estimates the pooled regression where the TBVC indicator is interacted with drought, with migrant worker, and with the *Drought*Any migrant worker* variable.

---

[33] The coefficient on Drought in the following year*TBVC is positive and significant in columns (4) and (6), although small relative to the main impact of drought on outmigration from non-TBVC areas. One reason for this sign could be because TBVC workers do not have the freedom to move when they need to (i.e. at the time of drought) but rather when external labor demand conditions dictate. They may be more likely to take these arising opportunities if drought is imminent (and there are long cycles in drought that make it feasible for people to anticipate drought). It is difficult to directly test this hypothesis since there are multiple droughts in different districts. To see whether this positive interaction term changes when I exclude periods during which there are known shocks to external labor demand within some of the TBVC areas, I exclude the years 1975-1979 and re-estimate equation (2) for the later period (see Mariotti 2011 for details of this labor demand shock). The positive interaction term becomes much smaller and insignificant, while the other results remain the same. This suggests that in the later periods, workers from TBVC areas may have moved in anticipation of drought, when the opportunity presented itself.

Even though this evidence on remittances comes from the post-*apartheid* period when legal barriers to labor migration are no longer in place, the legacy of historical mobility restrictions persists. In 1996, just over 30% of all households in rural TBVC and non-TBVC districts reported at least one migrant worker (this fraction is slightly higher in non-TBVC households) and one in five received some type of remittance the year before the Census. Conditional on having any migrant worker attached to the household, the chances of a household receiving any remittance were one in two. About 7% of households without migrant workers also received some remittances (results not shown).

Looking at the estimates in Table 6, we see that in households without a migrant worker, drought exposure in 1995 has no significant impact on remittance receipts. In contrast, migrant worker households in drought-exposed districts were a significant 2.68 percentage points more likely to receive any remittances in 1995 (column (1)). This is a 13% increase relative to the mean rate of remittances.

Splitting the sample into TBVC and non-TBVC areas, we see that having a migrant worker in the household facilitates remittances after drought in both areas. This relationship is more precisely estimate and somewhat larger in non-TBVC areas, reflecting the stronger historical migrant links between these areas and the rest of the economy. For non-TBVC areas, drought exposure raises the chances of receiving remittances by over 16% (*p*-value of the *F*-test for the sum of *Drought* and *Drought*Migrant worker* is 0.000). Combined with the prior results on outmigration, Table 6 supports the notion that remittances link outmigrants with rural households remaining behind in the former rural homelands and that these remittance flows do response positively to local economic shocks. Having a migrant worker attached to a household can help families to protect incomes against the negative effects of local drought.

iv. Ruling out composition effects: Fertility, mortality and sex selection

In the final section, I explore whether drought-induced changes in cohort composition account for the differential effects of drought on male disability in TBVC areas. Drought may affect

composition through changes in total fertility, child mortality or sex selection.[34] For example, if fewer children are born during drought due to higher fetal death rates, lower conception rates or planned pregnancy timing, the survivors may be stronger ("selection" dominates "scarring" as in Bozzoli, Deaton and Quintana-Domeque 2009). Positive survivor bias would lead us to underestimate the impact of drought on disability in all areas. If positive selection were stronger in TBVC areas, this would underestimate the impact of drought in TBVC areas even more. Alternatively, if positive selection was stronger in non-TBVC areas, then survivor bias could explain the male disability differentials across TBVC and non-TBVC areas.

To check for differential drought impact on fertility and child mortality across TBVC and non-TBVC areas, I estimate regressions for the number of children ever born and the number of children who have died for women ages 40 to 60 in the 1996 Census. The sample is restricted to women whose first district is a rural homeland area. I regress each outcome on the fraction of childbearing years (ages 15-40) in drought and an interaction with of this exposure with the TBVC status of the district. I include controls for first district fixed effects and birth year controls to capture age effects. Importantly, I assume each woman bears her children in the same district reported as her first district. This assumption generates a noisy assignment of first district for the (likely small) fraction of women who moved multiple times during *apartheid*.

Table 7 first presents results for the full sample of women (columns (1) and (2)) and then for subsamples of women with different levels of education. Although the point estimates on drought are large, they are imprecisely estimated for all groups except those women with high levels of education. For this subsample of women with at least secondary education, a mean level of drought exposure during childbearing years (0.07) reduces total births by 0.11 (2.6% relative to mean) but this is not different across TBVC and non-TBVC areas. For the remaining groups,

---

[34] There is also a direct effect of drought on population composition through outmigration. However, since I use year of birth and district of birth to assign drought exposure measures to individuals, the disability results are not driven by any migration-related composition changes.

there is no strong evidence that drought affects total fertility or total child mortality differentially by TBVC status.[35]

Finally, I use the entire sample of individuals born 1948 to 1986 to check whether drought in a given district and year affects total cohort size at the district-year level. I estimate equation (1) for the outcomes of (log of) district-year total males and total females. There is no evidence from the basic specifications in columns (1) and (3) that drought exposure in birth year affects cohort size for men or women in either TBVC or non-TBVC areas. The extended specifications indicate a small negative effect of drought exposure *in utero* and age one for males and females in TBVC areas. These differences are only significant at the 10% level and do not differ by gender. Cohort size effects are unlikely to explain why males from TBVC areas experience larger disability effects of drought exposure at birth. Any survivor bias in the male sample from TBVC drought cohorts would lead us to underestimate the differential impact of drought on disability rates in TBVC areas.

5. Conclusions

My paper presents new evidence on the long-term consequences of spatial barriers to mobility in a developing country where incomes are connected to variations in local weather conditions. Using externally imposed differences in migration barriers stemming from South Africa's *apartheid* policies, I estimate that early childhood drought exposure raises disability rates by 20% for African males born into areas facing the harshest mobility restrictions. This is over and above the average impact of drought on disability for exposed cohorts across all areas. I also use a new method to construct migration histories using a cross-section of Census data. Using these constructed migration histories I estimate that adult outmigration from less mobility-restricted areas is over 30% higher during a drought year but only 10% higher in areas facing tougher restrictions. I document a link between outflows of adults and inflows of money by showing

---

[35] The lack of mortality differentials in drought years and across TBVC and non-TBVC areas exposed to drought also reassures us that the disability results are unlikely to be driven by HIV-related mortality. While HIV-related illness can lead to disabilities of the sort measured in the Census, it also raises death rates. If mortality rates among children of women ages 40 to 60 in 1996 are not different by drought exposure or by TBVC status, it is unlikely that of our disability differences are driven by the HIV/AIDS epidemic.

recent evidence on remittance receipts responding to drought in households with migrant workers.

While the policies of *apartheid* are unique to South Africa, restrictions on mobility within a country exist in different forms throughout the world. Some barriers are institutional, as in China and the former Soviet Union. Others, particularly in Africa, are geographic in nature or related to inadequate transportation infrastructure. The development literature has noted the direct effects of limited spatial mobility on income volatility and has viewed spatial mobility as a form of insurance against economic shocks. My results identify a specific implication of migration barriers for health human capital accumulation over the long run. In environments prone to frequent local environmental shocks like drought, enhancing spatial mobility and labor market integration could generate large welfare gains through the health-protective effects of labor migration and remittance behaviors.

Highlighting how spatial mobility acts as insurance also provides new insights into the long-term consequences of economic shocks experienced in early life. There is a wealth of credible empirical evidence that negative shocks to the nutritional and disease environment in early life have severe short-run effects on child health and significant long-term effects on health in adulthood. Less is known about how families mitigate the impact of these negative shocks on the health of their children (Almond and Currie 2011a, Almond and Currie 2011b,). The South African results suggest that where they are able to, families use labor mobility over space to weather the effects of drought, with long-term gains in health human capital accumulation.

References

Aigner, Dennis J. 1973. "Regression with a binary independent variable subject to errors of observation" *Journal of Econometrics* Vol. 1 pp. 49-60

Alderman, Harold, John Hoddinott, and Bill Kinsey. 2006 "Long term consequences of early childhood malnutrition", *Oxford Economics Papers* No. 58: 450-474

Alderman, Harold, Hans Hoogeven, and Mariacristina Rossi. 2009 "Preschool Nutrition and Subsequent School Attainment: Longitudinal Evidence from Tanzania." *Economic Development and Cultural Change* Vol 57

Almond, Douglas. 2006 "Is the 1918 Influenza pandemic over? Long-term effects of *in utero* influenza exposure in the post-1940s US population", *Journal of Political Economy* Vol. 114: 4

Almond, Douglas and Janet Currie. 2011a. "Human capital development before age 5", in Handbook of Labor Economics, Vol. 4b. pp. 1315-1486. Elsevier

Almond, Douglas and Janet Currie. 2011b. "Killing me softly: The fetal origins hypothesis", *Journal of Economic Perspectives*, Vol 25. No. 3, Summer pp. 153-172

Almond, Douglas and Bhashkar Mazumdar. 2011. "Health capital and the prenatal environment: The effect of Ramadan observance during pregnancy", *American Economic Journal: Applied Economics*, Vol. 3. No. 4. pp.56-85.

Aguilar, Arturo and Marta Vicarelli. 2012. "El Niño and Mexican children: Medium-term effects of early-life weather shocks on cognitive and health outcomes", Working paper.

Amaral, Ernesto F. L. 2008 "Improvements of techniques to estimate migration rates: An application with Brazilian Census Data" *Population Review*, Vol 47: 2

Banerjee, Abhijit, Esther Duflo, Gilles Postel-Vinay and Tim Watts. 2010. "Long-run health impacts of income shocks: Wine and Phylloxera in Nineteenth-Century France", *Review of Economics and Statistics*, November. Vol. 94(2): 714-728

Barker, D. J. P. 1995. "Fetal Origins of Coronary Heart Disease", *British Medical Journal* Vol. 311: 171-174

Bell, Martin. 2005. "Towards a rigorous cross-national comparison of internal migration: Who collects what?" Paper presented at the International Union for the Scientific Study of PopulationTwenty-Fifth Conference, 18-23 July, 2005, Tours, France

Benson, Charlotte and Edward Clay. 2000. Drought: A Global Assessment, Vol. I, Routledge

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan, 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* Vol. 119: 1 pp. 249-275

Bound, John, Charles Brown and Nancy Mathieowitz. 2001. "Measurement error in survey data" in Handbook of Econometrics. (ed) E.E. Learner and J.J. Heckman. Pp. 3705-3843, North Holland

Bozzoli, Carlos, Angus Deaton and Climent Quintana-Domeque. 2009. "Adult height and childhood disease" *Demography* Vol. 46: 4, November, pp. 647-669

Brown, Barbara. 1987. "Facing the 'Black Peril': The politics of population control in South Africa", *Journal of Southern African Studies*, Vol 13: 3

Bryan, Gharad, Shyamal Chowdhury and A. Mushfiq Mobarak, 2012 "Risk aversion and seasonal migration", Mimeo, Yale University

Burgess, Robin, and Dave Donaldson. 2010. "Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era." *American Economic Review*, 100(2): 449–53

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2006. "Robust inference with multi-way clustering", *NBER Technical Working Paper* No. 327

Casale, Daniela and Dorrit Posel. 2006 "Migration and remittances in South Africa", Mimeo, University of KwaZulu-Natal

Chen, Yuyu and Li-An Zhou. 2007. "The long-term health and economic consequences of the 1959-1961 famine in China", *Journal of Health Economics*, Vol. 26 No. 4. pp. 659-681.

Clemens, Michael A. 2011. "Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?" *Journal of Economic Perspectives* Vol. 25: 3 pp. 83–106

Coovadia, H., R. Jewkes, P. Barron, D. Sanders and D. McIntyre. 2009 "The health and health system of South Africa: historical roots of public health challenges", *The Lancet*, September 5

De Brauw, Alan and John Giles. 2008. "Migrant labor markets and the welfare of rural households in the developing world: Evidence from China", World Bank Policy Research Paper No. 4585

Hoddinott, John, and Bill Kinsey. 2001. "Child Growth in the Time of Drought." 2001. *Oxford Bulletin of Economics and Statistics* 63(4): 409-36

Jayachandran, Seema. 2006 "Selling labor low: Wage responses to productivity shocks in developing countries" *Journal of Political Economy* Vol 114: 3. Pp. 538-575

Jensen, Robert. 2000. "Agricultural Volatility and Investments in Children." *American Economic Review, Papers and Proceedings*, 90(2): 399-404

Kaplan, Greg. 2012. "Moving back home: Insurance against labor market risk", *Journal of Political Economy*, Vol. 120: 3 pp. 446-512

Kraemer, Sebastian. 2000. "The fragile male", *British Medical Journal* Vol. 321: 7276. Pp. 1609–12

Le Roux, Noelien. 2009. "Seasonal maize yield simulations for South Africa using a multi-model ensemble system", Unpublished thesis manuscript, University of Pretoria, April

Lemon, Anthony. 1984 "State control over the labor market in South Africa" *International Political Science Review*, Vol. 5, No. 2: 189-208

Lindeboom, Maarten, France Portrait and Gerard J. van den Berg. 2010 "Long-Run Longevity Effects of a Nutritional Shock Early in Life: The Dutch Potato Famine of 1846–1847" *Journal of Health Economics*, Vol 29, No. 5 pp 617-629

Lucas, Robert E. B. and Oded Stark. 1988 "Motivations to Remit: Evidence from Botswana" *Journal of Political Economy* Vol. 93, No. 5 pp 901-918

Maccini, Sharon and Dean Yang. 2009 "Under the weather: health, schooling and socio-economic consequences of early-life rainfall", *American Economic Review* 99(3): 1006–26

Mariotti, Martine. 2011 "Living standards in South Africa's former homelands", Working Paper, Department of Economics, Australian National University.

Martorell, R., L. Kettel Khan, and D.G. Schroeder. 1994 "Reversibility of stunting: Epidemiological findings in children from developing countries", *European Journal of Clinical Nutrition* 48, S45{S57}

Martorell, Reynaldo. 1999. "The Nature of Child Malnutrition and its Long-Term Implications." *Food and Nutrition Bulletin* 20(3): 288-92.

McKee, T. B., N. Doesken and J. Kleist 1993 "The relationship of drought frequency and duration to time scales", Preprints, 8th Conference on Applied Climatology, pp. 179–184. January 17–22, Anaheim, California

Meng, X. and Nancy Qian. 2009 "The long run health and economic consequences of famine on survivors: Evidence from China's great famine", *IZA Discussion Paper* No. 2471.

Paulson, Anna L. 2000. "Insurance motives for migration: Evidence from Thailand", Kellogg School of Management, Northwestern University, Mimeo

Posel, Dorrit and Daniela Casale. 2006. "Internal migration and household poverty in post-apartheid South Africa" in Poverty and Policy in Post-Apartheid South Africa, edited by Ravi Kanbur and Haroon Bhorat. Pretoria: Human Sciences Research Council Press pp. 351-365

Posel, Dorrit. 2011. "Households and labour migration in post-apartheid South Africa", *Studies in Economics and Econometrics*, Vol. 34: 3

Redding, Stephen J., and Daniel M. Sturm. 2008. "The Costs of Remoteness: Evidence from German Division and Reunification", *American Economic Review*, Vol. 98: 5 pp. 1766–97

Rosenzweig, Mark and Oded Stark. 1989. "Consumption Smoothing, migration and marriage: Evidence from Rural India" *Journal of Political Economy*, Vol. 97: 4.

Rosenzweig, M. 1988. "Labor markets in low income countries" in Hollis Chenery and T.N. Srinivasan (ed.), Handbook of Development Economics, edition 1, volume 1, chapter 15, pages 713-762, Elsevier.

Rouault, M. and Y. Richard. 2005. "Spatial extent and intensity of droughts in Southern Africa" *Geophysical Research Letters*, Vol. 32: L15702, 4

Secretary for Bantu Administration and Development. 1967 "General Circular No. 25 – Settling of non-productive Bantu resident in European areas, in the homelands" in Maré, G. 1980 Population Relocation in South Africa, Appendix 5. South African Institute of Race Relations: Johannesburg

Savage, Michael. 1986 "The imposition of pass laws on the African population in South Africa 1916-1984", *African Affairs*, Vol. 85: 339

Schmertmann, Carl P. 1999. "Estimating multistate transition hazards from last-move data", *Journal of the American Statistical Association* Vol. 94: 445 pp. 53-63

Simkins, Charles. 1983. Four essays on the past, present & possible future of the distribution of the black population of South Africa, South African Labour and Development Research Unit, University of Cape Town

Sjaastad, Larry A. "The Costs and Returns of Human Migration." *The Journal of Political Economy*, 1962, *70*(5, Part 2: Investment in Human Beings), pp. 80-93.

Van den Berg, Gerard, Pia Pinger and Johannes Schoch. 2011 "Instrumental variable estimation of the causal effect of hunger early in life on health later in life", *IZA Discussion Paper* No. 6110

Wilhite, D. 2001 Drought: A global assessment, vol. I, ch. "Drought as a natural hazard: Concepts and definitions", Routledge

Wilson, Francis. 1972. Labour in the South African Gold Mines 1911-1969, Cambridge University Press

Wilson, Francis. 1996. "Challenges for the post-apartheid economy" *American Economic Review Papers and Proceedings*, Vol. 86: 2 pp. 322-325

Wolpe, Harold. 1988 Race, class and the apartheid state. James Currey Press, London

Yang, Dean, and HwaJung Choi. 2007. "Are Remittances Insurance? Evidence from Rainfall Shocks in the Philippines." *World Bank Economic Review*, 21(2): 219–48

Yang, Dean. 2011. "Migrant Remittances", *Journal of Economic Perspectives*, Vol. 25: 3 pp. 129–52

FIGURE 1
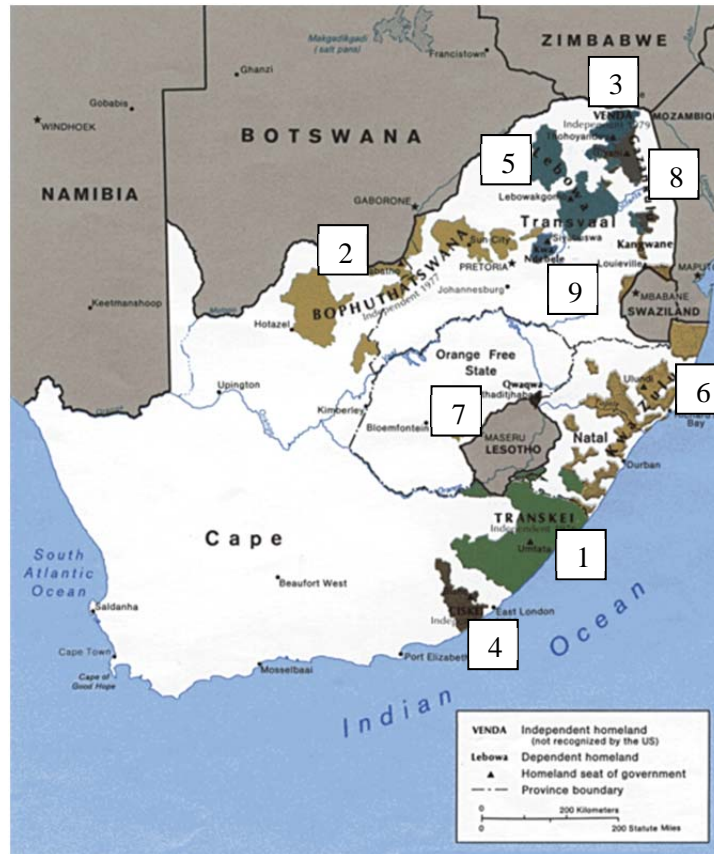
Source: Political map of South Africa 1986, Perry-Castañeda Library Map Collection, http://www.lib.utexas.edu/maps/south_africa.html, accessed July 2011. Homelands are (in order of establishment dates): TBVC areas: Transkei (1), Boputhatswana (2), Venda (3), Ciskei (4) and non-TBVC areas: Lebowa (5), KwaZulu (6), Qwaqwa (7), Gazankulu (8), Kangwane (9), KwaNdebele (10)

FIGURE 2

Graphs by TBVC status of district
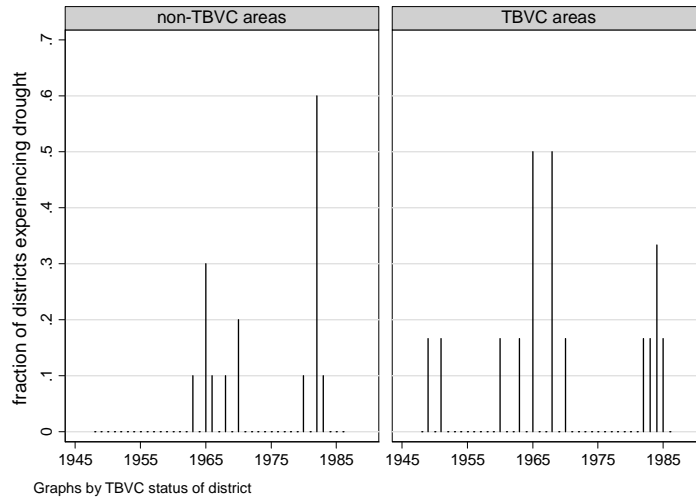
Figure 2 shows the fraction of South African homeland districts experiencing a drought annually between 1948 and 1986. The left hand panel shows drought in non-TBVC (less restricted) areas, the right hand panel for TBVC (more restricted) areas. The drought indicator is based on the Standardized Precipitation Index (SPI) as described in Data Appendix 1.

FIGURE 3
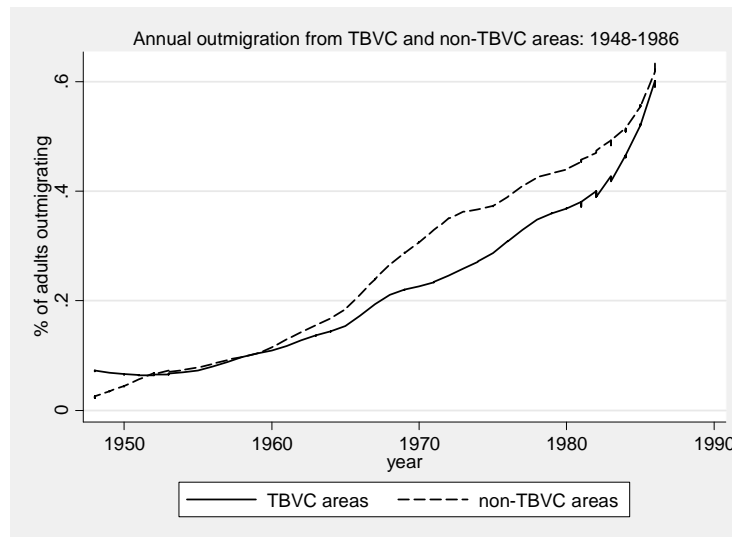
Figure 3 presents lowess-smoothed outmigration rates from TBVC and non-TBVC areas over time using Census data from 1948-1986. The y axis shows the average percent of adults who outmigrate from a district in a given year, across all districts. Kernel is Epanechnikov, bandwidth is 0.3.

Table 1: Summary Statistics for South African 1996 Census data, 10% individual sample

|	Means			P-value of difference in means	N. observations	Min.	Max.
	Full sample	TBVC sample	Non-TBVC sample				
Panel A: Individual-level data							
1. Cohorts 1948-1986							
Age	23.742	23.775	23.714	0.90	655,532	10	48
Female	0.545	0.543	0.546	0.68	655,532	0	1
Years of education (for adults age 20-40)	7.362	7.603	7.154	0.34	297,868	0	15
Any serious disability	0.052	0.054	0.051	0.55	655,532	0	1
Number of serious disabilities	0.057	0.059	0.054	0.26	655,532	0	4
Sight disability	0.023	0.024	0.022	0.64	655,532	0	1
Speech/hearing disability	0.011	0.012	0.011	0.26	655,532	0	1
Physical disability	0.014	0.015	0.014	0.83	655,532	0	1
Mental disability	0.007	0.008	0.007	0.04	655,532	0	1
Ever moved from prior district as an adult	0.062	0.063	0.061	0.86	655,532	0	1
Year moved from prior district	1986	1986.9	1985.7	0.01	67,927	1950	1993
Current district is TBVC district	0.440	0.939	0.004	0.00	655,532	0	1
Drought in year of birth	0.067	0.073	0.062	0.65	655,532	0	1
2. Female cohorts ages 40-60 in 1996							
Number of children ever born	4.793	4.909	4.688	0.527	79,532	0	23
Number of children died	1.002	1.096	0.912	0.616	70,836	0	19
Fraction of child-bearing years in drought	0.071	0.076	0.067	0.622	79,532	0	0.4
Panel B: District-year level data 1948-1986							
Percent of adults who migrate out	0.240	0.213	0.254	0.07	624	0	1
Fraction of districts experiencing drought	0.051	0.068	0.041	0.00	624	0	1
Number of districts (total=16)		6	10				

Individual and district-level means for African respondents in the 1996 Census, 10% individual records sample. TBVC stands for Transkei, Boputhatswana, Venda and Ciskei, the four earliest homeland states. Non-TBVC indicates one of the remaining six homeland areas. Individual-level data includes people born 1948-1986 (age 10-48 in 1996) who are currently or previously living in any districts formerly part of a homeland. Fertility data is reported for females who have completed childbearing in 1996, i.e. cohorts born 1936-1956. District-year level data is restricted to individuals who are adults during 1948-1986.

Table 2: Effects of drought exposure in birth year on disability later in life

Dependent variable	*Full sample [N=655,532]*		*Male sample [N=298,475]*		*Female sample [N=357,057]*	
	Any disability	Number of disabilities	Any disability	Number of disabilities	Any disability	Number of disabilities
	(1)	(2)	(3)	(4)	(5)	(6)
Drought in birth year	0.0012	0.0016	-0.0008	0.0004	0.0029	0.0027
	(0.0018)	(0.0022)	(0.0031)	(0.0038)	(0.0020)	(0.0023)
Drought in birth year*TBVC	0.0038	0.0045	0.0100***	0.0097***	-0.0013	0.0002
	(0.0023)	(0.0030)	(0.0028)	(0.0033)	(0.0030)	(0.0036)
District fixed effects?	Y	Y	Y	Y	Y	Y
Year of birth fixed effects?	Y	Y	Y	Y	Y	Y
Mean of outcome	0.052	0.056	0.051	0.055	0.053	0.057
p-values for F-tests						
All drought parameters jointly =0	0.064	0.085	0.000	0.001	0.330	0.391
Sum of Drought and Drought*TBVC=0	0.019	0.027	0.000	0.002	0.535	0.368
Drought in birth year parameters are equal for males and females[1]			0.490	0.795		
Drought*TBVC parameters are equal for males and females[1]			0.000	0.001		

Robust standard errors clustered on the latitude and longitude of the prior district. Levels of significance: p<0.001***, p<0.05**, p<0.01*. All regressions include full set of birth year and prior district fixed effects; TBVC indicator is absorbed by district fixed effects. Drought exposure is a binary variable constructed using values of the Spatial Precipitation Index; TBVC indicates whether an individual reports a prior district is TBVC or not. Sample restricted to 1996 Census data on individuals born between 1948 and 1986.

[1]P-values for tests of equality between male and female coefficients are estimated using results from fully-interacted pooled regressions.

Table 3: Effects of drought exposure in early childhood on disability later in life: Males

Dependent variable	*Male sample [N=298,475]*	
	Any disability	Number of disabilities
	(1)	(2)
Drought *in utero*	0.0075**	0.0085***
	(0.0031)	(0.0032)
Drought in birth year	0.0011	0.0025
	(0.0034)	(0.0040)
Drought at age 1	0.0071**	0.0072*
	(0.0034)	(0.0041)
Drought at age 2	0.0095***	0.0091***
	(0.0025)	(0.0031)
Drought at age 3	0.0009	0.0016
	(0.0029)	(0.0034)
Drought at age 4	0.0021	0.0025
	(0.0036)	(0.0040)
Drought *in utero* *TBVC	0.0026	0.0017
	(0.0034)	(0.0037)
Drought at birth*TBVC	0.0081**	0.0079*
	(0.0034)	(0.0041)
Drought at age 1*TBVC	0.0032	0.0027
	(0.0048)	(0.0065)
Drought at age 2*TBVC	0.0007	0.0039
	(0.0042)	(0.0056)
Drought at age 3*TBVC	0.0048	0.0019
	(0.0047)	(0.0056)
Drought at age 4*TBVC	0.0072	0.0084*
	(0.0044)	(0.0051)
District fixed effects?	Y	Y
Year of birth fixed effects?	Y	Y
Mean of outcome	0.051	0.055

p-values for F-tests

All Drought, Drought*TBVC parameters jointly=0	0.000	0.018

Robust standard errors clustered on the latitude and longitude of the prior district. Significance levels: p<0.001***, p<0.05**, p<0.1*. All regressions include full set of birth year and prior district fixed effects; TBVC indicator is absorbed by district fixed effects. Drought exposure at each age is a binary variable constructed using values of the Spatial Precipitation Index; TBVC indicates whether an individual reports a prior district is TBVC or not. Sample restricted to 1996 Census data on individuals born between 1948 and 1986.

Table 4: Effects of drought exposure in birth year on component disabilities: Males

Dependent variable:	Male sample [N=298,475]			
	Sight disability	Hearing/speech disability	Physical disability	Mental disability
	(1)	(2)	(3)	(4)
Drought in birth year	0.0009	0.0002	-0.0009	0.0000
	(0.0024)	(0.0009)	(0.0014)	(0.0008)
Drought in birth year*TBVC	0.0042	0.0012	0.0032**	0.0013
	(0.0027)	(0.0014)	(0.0013)	(0.0009)
District fixed effects?	Y	Y	Y	Y
Year of birth fixed effects?	Y	Y	Y	Y
Mean of outcomes	0.020	0.011	0.015	0.010
p-values for F-tests				
All drought parameters jointly =0	0.007	0.591	0.014	0.371
Sum of Drought and Drought*TBVC=0	0.002	0.307	0.021	0.215

Robust standard errors clustered on the latitude and longitude of the prior district. Levels of significance: $p<0.001$***, $p<0.05$**, $p<0.01$*. All regressions include full set of birth year and prior district fixed effects; TBVC indicator is absorbed by district fixed effects. Drought exposure at each age is a binary variable constructed using values of the Spatial Precipitation Index; TBVC indicates whether an individual reports a prior district is TBVC or not. Sample restricted to 1996 Census data on individuals born between 1948 and 1986.

Table 5: Effect of drought on percent of adults who outmigrate from TBVC and non-TBVC homelands

Sample	Full sample: 1948-1996				Later period: 1960-1986	
	Broad measure of TBVC status [1]		Narrow measure of TBVC status [2]		Broad measure of TBVC status [1]	
Dependent variable	Percent of outmigrants from district d in year t					
	(1)	(2)	(3)	(4)	(5)	(6)
Drought year	0.088**	0.078**	0.061***	0.059***	0.066***	0.060***
	(0.039)	(0.036)	(0.001)	(0.001)	(0.001)	(0.001)
Drought year*TBVC	-0.067**	-0.061**	-0.046***	-0.051***	-0.040***	-0.039***
	(0.031)	(0.028)	(0.002)	(0.001)	(0.002)	(0.002)
Drought last year		0.099*		0.082***		0.078***
		(0.055)		(0.001)		(0.002)
Drought last year*TBVC		-0.067		-0.107***		-0.039***
		(0.060)		(0.003)		(0.002)
Drought next year		0.047*		0.041***		0.025***
		(0.024)		(0.000)		(0.001)
Drought next year*TBVC		-0.007		0.025***		0.014***
		(0.023)		(0.002)		(0.002)
District fixed effects?	Y	Y	Y	Y	Y	Y
Year of birth fixed effects?	Y	Y	Y	Y	Y	Y
P-values for F-tests						
All drought parameters=0	0.079	0.347	0.000	0.000	0.000	0.000
Sum of Drought and Drought*TBVC=0 for:						
Drought this year	0.180	0.265	0.000	0.000	0.000	0.000
Drought last year		0.227		0.000		0.000
Drought next year		0.052		0.000		0.000
N observations	624	624	624	624	432	432
R2	0.63	0.64	0.63	0.64	0.62	0.63
Mean of outcome	0.24%	0.24%	0.24%	0.24%	0.32%	0.32%

Robust standard errors clustered on the latitude and longitude of the district; p<0.001***, p<0.05**, p<0.1*. Sample is restricted to 1996 Census data on African men and women who report their prior district was in a former homeland area of South Africa and who are 18 or older in 1996. Unit of observation is the district-year. Outcome is the percent of adults who move away from a prior district in a given year. Drought is an indicator for whether there was a drought in the district in a given (or prior or following) year. All regressions control for a full set of year and district fixed effects.

[1] Broad TBVC measure is an indicator for whether the prior district falls into one of the independent homeland areas or not and is constant through 1946-1986.

[2] Narrow TBVC measure is the broad TBVC measure refined to turn on only during the years in which TBVC states were independent.

Table 6: Effect of drought in 1995 on remittances to former homelands in 1995

Sample	All households [N=222,355]	Non-TBVC areas [N=115,126]	TBVC areas [N=107,229]	All households [N=222,355]
Dependent variable		Any remittances received in 1995?		
	(1)	(2)	(3)	(4)
Drought last year	0.0023	0.0089	-0.0058	0.0086
	(0.0045)	(0.0062)	(0.0038)	(0.0062)
Any migrant worker in HH?	0.4127***	0.3913***	0.4362***	0.3911***
	(0.0192)	(0.0214)	(0.0296)	(0.0215)
Drought last year*Any migrant worker in HH?	0.0268***	0.0327***	0.021	0.0328***
	(0.0092)	(0.0108)	(0.0146)	(0.0108)
TBVC district				-0.005
				(0.0132)
Drought last year*TBVC				-0.0130*
				(0.0066)
TBVC*Any migrant worker in HH?				0.045
				(0.0364)
Drought last year*TBVC*Any migrant worker in HH?				-0.013
				(0.0179)
Fraction of 1948-1996 period in drought	0.5399***	0.4360	0.7203**	0.5289**
	(0.1904)	(0.2751)	(0.3190)	(0.2107)
Mean of outcome	0.20	0.20	0.20	0.20
p-value for F-test				
Sum of Drought and Drought*Migrant=0	0.00	0.00	0.26	0.00
Sum of all drought terms=0[1]				0.20

Robust standard errors clustered on the latitude and longitude of the current district. $p<0.001$***, $p<0.05$**, $p<0.1$*. Sample restricted to 1996 Census data on households residing in one of the TBVC or non-TBVC districts in 1996. Unit of observation is the household. Fraction of sample with a migrant worker attached to household in 1996 is 0.318 in non-TBVC areas and 0.314 in TBVC areas.

[1.] Sum of all drought terms is Drought + Drought*Migrant + Drought*TBVC + Drought*TBVC*Migrant

Table 7: Effects of drought exposure during childbearing years on fertility outcomes among African women who have completed childbearing by 1996

Sample	Female sample		Female sample by level of education					
			None		Some primary		Some secondary	
Dependent variable	N. kids ever born	N. kids died	N. kids ever born	N. kids died	N. kids ever born	N. kids died	N. kids ever born	N. kids died
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fraction years in drought (ages 15-40)	0.105	0.122	0.2621	0.134	-0.0849	0.058	-0.116*	0.044
	(0.1360)	(0.0851)	(0.1702)	(0.0901)	(0.0832)	(0.0656)	(0.0557)	(0.0468)
Fraction years in drought (ages 15-40)*TBVC	-0.2440	0.0116	-0.2605	0.0553	-0.1028	-0.0495	-0.0756	0.0638
	(0.2436)	(0.0597)	(0.2987)	(0.0639)	(0.1818)	(0.0683)	(0.1039)	(0.0564)
District fixed effects?	Y	Y	Y	Y	Y	Y	Y	Y
Year of birth fixed effects?	Y	Y	Y	Y	Y	Y	Y	Y
N	79,532	70,836	37,989	33,312	17,407	15,852	21,882	19,952
Mean of outcome	4.79	1.00	4.96	1.15	5.11	1.13	4.35	0.68

Robust standard errors clustered on the latitude and longitude of the first district. p<0.001***, p<0.05**, p<0.1*. All regressions control for first district fixed effects and year of birth; TBVC indicator absorbed by district fixed effects. Sample restricted to women who have completed childbearing in 1996 (ages 40-60). The fraction of years exposed to drought is constructed using drought prevalence in the first district during the years in which the woman is between age 15 and 40 inclusive (1951-1996). Sample is restricted to women who have completed childbearing in the 1996 Census: that is, women aged 40-60 in 1996.

Table 8: Effects of drought exposure in early childhood on adult population composition, Census 1996

Dependent variable	Ln males		Ln females	
Specification	Basic	Extended	Basic	Extended
	(1)	(2)	(3)	(4)
Drought *in utero*		-0.044		-0.045
		(0.040)		(0.028)
Drought in birth year	0.026	0.022	0.025	0.028
	(0.045)	(0.047)	(0.034)	(0.037)
Drought at age 1		0.009		0.050
		(0.062)		(0.043)
Drought at age 2		-0.016		0.025
		(0.050)		(0.035)
Drought at age 3		0.028		0.030
		(0.043)		(0.043)
Drought at age 4		-0.016		-0.013
		(0.045)		(0.030)
Drought *in utero* *TBVC		-0.1497*		-0.1271*
		(0.087)		(0.067)
Drought in birth year*TBVC	-0.109	-0.104	-0.077	-0.075
	(0.102)	(0.103)	(0.073)	(0.073)
Drought at age 1*TBVC		-0.112		-0.1422*
		(0.116)		(0.083)
Drought at age 2*TBVC		-0.085		-0.069
		(0.098)		(0.058)
Drought at age 3*TBVC		-0.069		-0.051
		(0.109)		(0.091)
Drought at age 4*TBVC		-0.113		-0.096
		(0.088)		(0.074)
District fixed effects?	Y	Y	Y	Y
Year of birth fixed effects?	Y	Y	Y	Y
p-values for F-tests				
All Drought, Drought*TBVC parameters jointly=0	0.317	0.000	0.420	0.000
All Drought*TBVC parameters jointly=0	0.280	0.517	0.295	0.000
N	624	624	624	624

Robust standard errors clustered on the latitude and longitude of the prior district. p<0.001***, p<0.05**, p<0.1*. All regressions control for birth year and first district fixed effects. Regressions in columns 1-6 are estimated using observations aggregated to the year of birth-first district level; regressions in columns 7 and 8 are estimated using individual level data. Sample is restricted to cohorts born 1948-1986 (ages 10-48) in 1996.
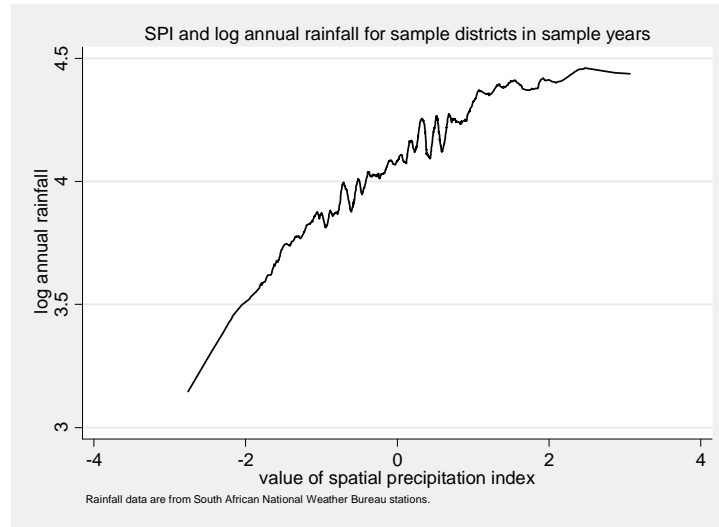
For Online Publication

Appendix 1: Data

1. Rainfall Data and Drought

The South African Weather Service http://www.weathersa.co.za/web/ provided the raw historical rainfall data These data contain monthly rainfall measures at the weather station level for over 1,600 weather stations across South Africa from 1920 to 2009. I spatially match the GIS locations of rainfall stations to corresponding district boundaries and aggregate rainfall totals to the district-year level.
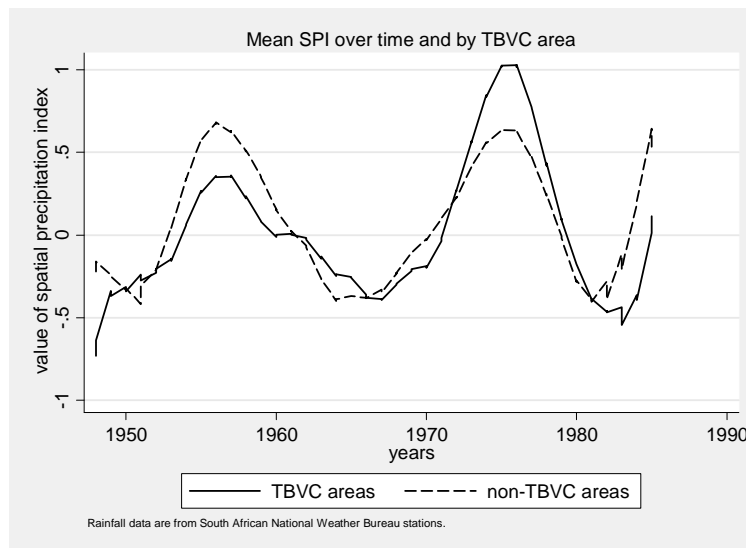
To create a measure of drought, I construct the Standardized Precipitation Index (SPI) at the district and year level (McKee, Doesken and Kleist 1993). The SPI measures the probability of observing a recent rainfall event based on the distribution of all rainfall events for a given time scale and place. Since rainfall is not normally distributed, the SPI procedure calls for a gamma distribution to be fit to the empirical data distributions. I fit a gamma distribution to the annual total rainfall of each district. and generate estimates of the scale and location parameters for district-specific rainfall patterns. For each year in the data, and the district-specific gamma distribution, I compute the probability of observing the total rainfall that was measured in each year and translate this into a normally distributed random variable using the normal CDF. This number is the district-year-specific SPI, where positive numbers reflect above-average rainfall and negative values reflect below-average rainfall. The positive relationship between log rainfall and the SPI measure across all districts is shown in Appendix Figure 1.

Following the climatological literature (e.g. McKee et al 1993) I define an indicator $DROUGHT_{dt}$ for each district (*d*) and year (*t*) that takes a value of 1 when the SPI is less than -1.5, and 0 otherwise. The spatial specificity of this measure is helpful because the same quantitative rainfall deficit may indicate inadequate precipitation in historically wetter districts but not in historically drier districts. Appendix Figure 2 shows lowess-smoothed graphs of the district level mean SPI values across TBVC and non-TBVC areas for the years 1948 to 1986. The pattern of serial correlation in SPI values is common to the two areas and implies some predictability of drought.

Appendix Figure 1

SPI and log annual rainfall for sample districts in sample years

Rainfall data are from South African National Weather Bureau stations.

Appendix Figure 2

Mean SPI over time and by TBVC area

Rainfall data are from South African National Weather Bureau stations.

There is a tight link between the SPI measure and South African maize production. Using province-level data from the South African Maize Board for the period 1964 to 1984 and for the commercial maize-growing provinces (Transvaal and the Orange Free State), I estimate the relationship between the SPI measure and maize yields. Appendix Figure 3 shows the lowess-smoothed relationship between the log of South Africa's annual maize output (in tons) against the Spatial Precipitation Index using an Epanechnikov kernel with a 0.5 bandwidth. This positive relationship is asymmetric. Output appears more sensitive to low values of the SPI than it is to

higher, positive values of the SPI. Figure 3 suggests that drought in particular captures an important negative productivity shock in agriculture.

Appendix Figure 3

Maize production and SPI in South Africa's maize-growing regions

bandwidth = .5
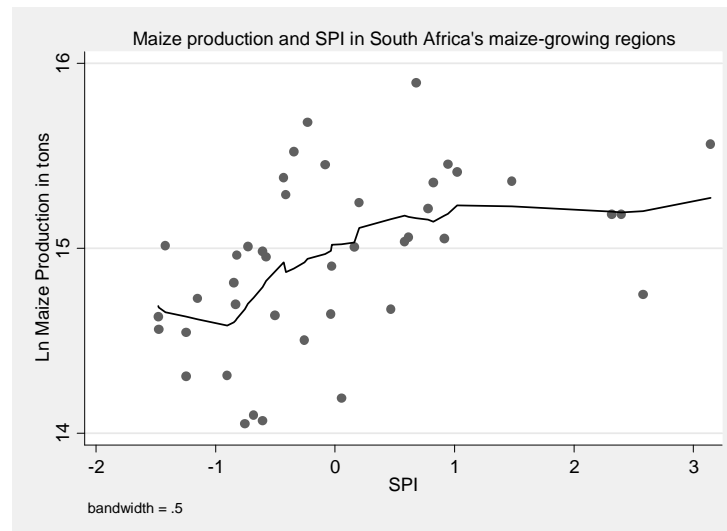
2. Homeland Boundary Data and the TBVC assignment

GIS data on sub-national boundaries for the 1996 and 2001 Census were obtained from Statistics South Africa (www.statssa.gov.za). I use the 2001 district council Census boundaries as the main geographic unit of observation since these areas are large enough to treat as distinct local labor markets and contain sufficient population in each year to make aggregation feasible.[1]

To define which of these districts belong to former homeland areas, I obtained online maps of the ten homelands with the predominant map dated 1986 (see Figure 1 in the main text). I overlaid these homeland maps onto Census boundaries and, where there was overlap, assigned districts to homelands. I created an indicator $TBVC_d$ that takes a value of 1 if a district overlapped with any of the TBVC homelands, and is 0 for those districts overlapping the remaining six homelands.[2] Of the 53 district councils in South Africa, 16 of them (30%) fall

---

[1] Magisterial districts are too small to contain sufficient population and rainfall measurements for my analysis.
[2] TBVC stands for Transkei, Boputhatswana, Venda and the Ciskei.

into prior rural homeland areas. Of these 16 areas, six fall in the former TBVC areas and the remaining 10 fall in the non-TBVC areas.

3. 1996 Census Data

The 10% sample of individual records from the 1996 South African Census was obtained from Statistics South Africa (www.statssa.gov.za).

i. Migration variables

This Census asked all individuals about their current district of residence, their former district of residence and the year in which they moved to their current district. I use this information to define several variables relevant to migration:

- The district of current residence
- The district of former residence: this is the same as district of current residence for individuals who report never having moved, for individuals who have moved since the end of apartheid (1994-1996), for individuals who reported moving during childhood to their current residence, and for those who report moving within a district
- An indicator for whether a person moved before 1994 (the end of apartheid) and during their adult lives. This indicator is 1 if a person's former residence differs from their current residence and if they report the year they moved to their current residence.

For the analysis in the paper, I eliminate all individuals who report a prior residence in a district covering one of South Africa's largest cities (36% of the sample). I also eliminate those who report living in (for never movers) or formerly living in (for movers) districts outside of South Africa. Less than 1% of the sample has a usual residence outside South Africa and less than 5% have a prior residence outside of South Africa. Of the remaining sample of adults who report a former residence (for movers) or current residence (for never movers) located in rural South Africa, 97% have complete information on current and former district of residence and the year of moving to current residence. For the 3% who report a current residence and no information about year of moving, I assign them to be non-movers.

For the migration analysis, I further restrict the sample to African adults aged 18 and older in 1996 who report a current (for never movers) or prior residence (for movers) in South Africa that is predominantly rural and located in one of the former homeland areas.

For each year in which a respondent is 18 or older, I identify what district they lived in under the strong assumption that each person moves only once. That is, I create a pseudo-panel dataset describing the place of residence by year of adulthood. I match this panel to drought at the district-year level. Finally, I collapse the resulting individual-year-districts dataset to district-year level for the migration analysis.

ii.   *Health, fertility and population outcomes*

For the disability analysis, and for the analysis of cohort size and sex composition, I use the sample of African adults who lived in any of the former homeland areas between 1948 and 1986. I match the cross-sectional data on outcomes at the individual level to the drought data on year of birth and prior district.

Note that the Census does not capture place of birth information, so I assume that a person's prior residence is their birth district. This means that birth district is potentially misclassified for people who move multiple times. Appendix 2 discusses the implications of this measurement error.

For the analysis of fertility and child mortality outcomes, I restrict the sample to African women aged 40 to 60 in 1996 and create a variable that represents the fraction of their childbearing years (ages 15-40) that they experienced drought. I assign drought exposure at the district level using the prior district reported by these women.

References

McKee, T. B., N. Doesken and J. Kleist 1993 "The relationship of drought frequency and duration to time scales", Preprints, 8th Conference on Applied Climatology, pp. 179–184. January 17–22, Anaheim, California

For Online Publication:

Appendix 2: Measurement error in Census migration data

The nature of "last move" Census data induces measurement error in migration and, potentially, in drought exposure. This note characterizes the resulting measurement error bias that arises when modeling how outmigration responds to drought exposure. Intuitively, the bias is related to the size of the population that actually moves more than once, to the fraction of observed drought events, and to the fraction of multiple movers who have misclassified drought exposure due to invisible prior migrations.

To fix ideas, note the Census contains three types of individuals: people who have never moved from their district of residence ("never movers"), people who have only ever moved districts once ("single movers") and those who have moved multiple times ("multiple movers"). Complete migration histories are known for the never movers and single movers. Misclassification only occurs for multiple movers, in moves before their last. For these multiple movers, earlier moves are made invisible by later moves (Schmertmann 1999). In addition, since a drought exposure variable for each year of life is assigned to an individual based on their district of residence in each year, misclassified migration could induce misclassification in drought exposure. This has implications for regressions of migration on drought.

To illustrate, consider the following difference-in-differences model:

$$(1)\ y_{ijt}^* = \alpha + \beta D_{ijt}^* + \lambda_j + \delta_t + \varepsilon_{ijt}$$

where y_{ijt}^* indicates whether a person moves away from district j in year t (y_{ijt}^*=1) or not (y_{ijt}^*=0), D_{ijt}^* indicates whether a person was exposed to drought in district j in year t (D_{ijt}^*=1) or not (D_{ijt}^*=0), ε_{ijt} is an error term and λ_j and δ_t are district and year fixed effects respectively. Throughout this appendix, starred outcome and independent variables denote true values of these variables; unstarred variables denote observed outcomes.

A person's observed migration status in each year (y_{ijt}) can be related to their true migration status (y_{ijt}^*) as follows:

(2) $y_{ijt} = y_{ijt}^* + v_{ijt}$

    i) $v_{ijt} = 0$ if $y_{ijt} = y_{ijt}^* = 0$ or 1

    ii) $v_{ijt} = -1$ if $y_{ijt} = 0$, $y_{ijt}^* = 1$

Condition (2i) describes never movers, single movers and multiple movers who are on their last move. For these cases, there is no measurement error in migration status, so v_{ijt} is always zero. Condition (2ii) describes the case of a misclassified non-move for multiple movers, when a real move is unseen because it was prior to the last move. Since every reported move is a true move, $v_{ijt} = 1$ ($y_{ijt} = 1$, $y_{ijt}^* = 0$) is ruled out.

Observed drought exposure for each person in each district of each year (D_{ijt}) can be related to true drought exposure (D_{ijt}^*) in a similar way:

(3) $D_{ijt} = D_{ijt}^* + w_{ijt}$

    i) $w_{ijt} = 0$ if $D_{ijt} = D_{ijt}^* = 0$ or 1

    ii) $w_{ijt} = 1$ if $D_{ijt} = 1$, $D_{ijt}^* = 0$

    iii) $w_{ijt} = -1$ if $D_{ijt} = 0$, $D_{ijt}^* = 1$

Condition (3i) describes never movers, single movers and multiple movers on their last move who have no measurement error in drought exposure. Condition (3ii) describes the misclassification of non-drought exposure for a multiple mover whose prior move is unobserved (i.e. if $v_{ijt} = -1$). Condition (3iii) describes misclassification of drought exposure for a multiple mover whose prior move is unobserved (i.e. if $v_{ijt} = -1$).

Using (2) and (3) to substitute out true unobserved values of drought and migration in (1), we can estimate the following using the Census pseudo-panel data on last moves:

(4) $y_{ijt} = \alpha + \beta D_{ijt} + \lambda_j + \delta_t - \beta w_{ijt} + v_{ijt} + \varepsilon_{ijt}$

If we assume $cov(D_{ijt}, \lambda_t + \delta_j + \varepsilon_{ijt}) = 0$ (essentially, observed drought is randomly assigned) we can make progress describing the measurement error bias in β_{OLS}:

$$\text{plim } \beta_{OLS} \quad = cov(D_{ijt}, y_{ijt})/var(D_{ijt})$$

$$= cov(D_{ijt}, \alpha + \beta D_{ijt} - \beta w_{ijt} + v_{ijt} + \lambda_j + \delta_t + \varepsilon_{ijt})/var(D_{ijt})$$

$$= \beta + cov(D_{ijt}, -\beta w_{ijt} + v_{ijt})/var(D_{ijt})$$

$$= \beta(1 - cov(D_{ijt}, w_{ijt})/var(D_{ijt})) + cov(D_{ijt}, v_{ijt})/var(D_{ijt}) \qquad (5)$$

Hence, $\text{plim}(\beta_{OLS}) \neq \beta$. The first term in this expression represents measurement error bias coming from misclassified drought exposure. The second term represents additional bias generated by the relationship between misclassified migration and misclassified drought exposure. If it is the case that measurement errors in drought and migration are uncorrelated ($\text{cov}(D_{ijt}, v_{ijt})=0$), we would be left with the standard downwards bias from measurement error generated by misclassification of a binary independent variable (Aigner 1973, Bound, Brown and Mathiowetz 2001 p. 3725-3726).[1]

Because of how drought exposure is assigned to an individual, it is unlikely that migration and drought errors are uncorrelated, so we must evaluate $\text{cov}(D_{ijt}, w_{ijt})$ and $\text{cov}(D_{ijt}, v_{ijt})$ to understand the net effects of the two biases in (5). The joint probability distribution of migration and drought variables is useful for this exercise:

Appendix 2 Table 1: Joint PDF for migration and drought variables

Observed migration status (y_{ijt})	True migration status (y_{ijt}^*)	Measurement error in migration (v_{ijt})	Observed drought exposure (D_{ijt})	True drought exposure (D_{ijt}^*)	Measurement error in drought (w_{ijt})	Drought*error in drought ($D_{ijt} * w_{ijt}$)	Drought*error in migration ($D_{ijt} * v_{ijt}$)	Probabilities
Migration status and drought exposure for never movers, single movers, and some multiple movers: No error								
0	0	0	1	1	0	0	0	P1
0	0	0	0	0	0	0	0	P2
1	1	0	1	1	0	0	0	P3
1	1	0	0	0	0	0	0	P4
Migration and drought for multiple movers only: Misclassification errors								
0	1	-1	1	1	0	0	-1	P5
0	1	-1	0	0	0	0	0	P6
0	1	-1	1	0	1	1	-1	P7
0	1	-1	0	1	-1	0	0	P8*

*P8=1-P1-P2-P3-P4-P5-P6-P7

From this, the following useful moments are available:

$E[D_{ijt}]=P1+P3+P5+P7$ which we will estimate by $\overline{D}$, the fraction of observed drought exposures

$E[w_{ijt}]=P7 - P8$

$E[D_{ijt}w_{ijt}]= P7$

[1] As discussed in Bound et al (2001), this result requires that the misclassification rate of drought exposure (sum of false negatives and false positives) does not exceed 1. This does not seem to be an unreasonably strong assumption to make in our context.

$E[D_{ijt}v_{ijt}] = -P5 - P7$

$E[v_{ijt}] = -P5-P6-P7-P8$

$\quad = -(1-P1-P2-P3-P4)$ which we will estimate by $-(\bar{M})$ and where $\bar{M}$ is the fraction of incorrectly classified multiple movers.

$V[D_{ijt}] = E[D_{ijt}^2] - E[D_{ijt}]^2$

$\quad = (E[D_{ijt}])(1 - (E[D_{ijt}]))$ which we estimate by $\bar{D}(1 - \bar{D})$

$V[v_{ijt}] = E[v_{ijt}^2] - E[v_{ijt}]^2$

$\quad = (P5+P6+P7+P8) - (P5+P6+P7+P8)^2$

$\quad = (1-P1-P2-P3-P4)(1-(1-P1-P2-P3-P4))$ which we estimate using $(\bar{M})$ $(1-\bar{M})$

With these, we can compute expressions for each bias term in (5). For the first term:

$cov(D_{ijt}, w_{ijt})/var(D_{ijt}) = (E[D_{ijt}w_{ijt}] - E[D_{ijt}]E[w_{ijt}])/[(E[D_{ijt}])(1 - (E[D_{ijt}]))]$

$\quad = (P7 - (P7-P8)*E[D_{ijt}])/((E[D_{ijt}])(1 - (E[D_{ijt}]))])$

$\quad = (P7*(1- E[D_{ijt}]) + P8*(E[D_{ijt}]))/((E[D_{ijt}])(1 - (E[D_{ijt}]))])$

$\quad = P7/(E[D_{ijt}]) + P8/(1- E[D_{ijt}]) > 0$

We can estimate this expression as:

$\quad P7/\bar{D} + P8/(1- \bar{D})$ $\qquad\qquad\qquad$ (6)

To address the second term in (5):

$cov(D_{ijt},v_{ijt}) /var(v_{ijt}) = (E[D_{ijt}v_{ijt}] - E[D_{ijt}]E[v_{ijt}])/var(v_{ijt})$

$\quad = (-(P5+P7) - E[D_{ijt}]E[v_{ijt}])/([(E[v_{ijt}])(1 - (E[v_{ijt}]))])$

We can estimate this as:

$\quad \bar{D}/(1 - \bar{M}) - (P5+P7)/(\bar{M}) (1-\bar{M})$ $\qquad\qquad$ (7)

Putting (6) and (7) together, we can rewrite (5) as:

$\text{plim } \beta_{OLS} = \beta(1-cov(D_{ijt}, w_{ijt})/var(D_{ijt})) + cov(D_{ijt},v_{ijt})/var(D_{ijt})$

$\quad = \beta*[1 - (P7/\bar{D}) - P8/(1 - \bar{D})] + \bar{D}/(1 - \bar{M}) - (P5+P7)/(\bar{M})(1-\bar{M})$

The sign of this bias is ambiguous. Under the assumption that the true β is positive, we can

consider what values of actual drought exposure $(\bar{D})$ and fraction of mismeasured movers $(\bar{M})$

would create a net downwards bias. Then, we can use information from external datasets to learn whether these values are plausible in this South African case.

The net bias in (5) will be downwards when

a) $1 > (P7/\bar{D}) + P8/(1 - \bar{D})$ and

b) $\bar{D}/(1 - \bar{M}) - (P5+P7)/(\bar{M})(1-\bar{M}) < 0$ and small

If we further make the reasonable assumption that P7=P8 (the misclassification of drought exposure for misclassified multiple movers is symmetric), and rearrange a) and b), these conditions become:

a)' $P7 < \bar{D}*(1 - \bar{D})$

b)' $\bar{M}\,\bar{D} - P5 - P7 < 0$ and small

In words, these conditions imply that the bias in (5) is downwards in contexts with smaller fractions of multiple movers (small $\bar{M}$) and smaller observed fractions of drought-exposed cohorts (small $\bar{D}$).

- In my data, $\bar{D} = 0.051$ implying $\bar{D}*(1 - \bar{D}) = 0.048$
- I use an upper bound of 0.13 for $\bar{M}$; hence, $\bar{D}\bar{M}=0.006$ or smaller.[2]

Using the values in the South African data, we as long as P7< 0.048 and 0.006 – P5 – P7<0 and small, then β_{OLS} is likely downwards biased. Put another way, the fraction of misclassified multiple movers mistakenly assigned to drought exposure (P7) would have to be larger than 0.048; and the fraction of misclassified multiple movers with any drought exposure (P5+P7) would have to be smaller than 0.006 in order for (5) to overestimate the impact of drought on outmigration. This seems unlikely to be the case in the South African setting.

---

[2]Data from the 2007 South African Community Survey and the 2007 Cape Area Panel Study indicate that the fraction of Africans who move more than once in the past five years is between 0.01 and 0.13 respectively. The 2007 Community Survey collects data on more than 300,000 African adults including their province of current and prior residences. The 2007 Cape Area Panel Survey is a sample of young adults (ages 24-33) drawn from a province with a highly mobile population, hence the higher rates of misclassification.

Implications for estimating the triple difference model for migration

The focus of the paper is on results from triple-difference specifications of the relationship between drought and migration outcomes:

$$y^*_{ijt} = \beta_0 + \beta_1 D^*_{ijt} + \beta_2 D^*_{ijt}*TBVC_{ijt} + \lambda_j + \delta_t + \varepsilon_{ijt} \qquad (8)$$

where true migration is mismeasured for multiple movers and true drought exposure and TBVC status of the prior district could both be mismeasured. While the analytical framework for measurement error bias described above is not helpful in signing the bias in this specification, it inspires a specific robustness check. Restricting the sample over which (8) is estimated to later periods should provide a less biased measures of β_1 and β_2, since misclassification of migration (and hence drought) only occurs for moves prior to the last one. Hence, I test the robustness of the migration results estimating (8) over more recent subsamples of the data.

Implications for the health regressions

In the analysis of health outcomes (disability rates) using (8), the only way that measurement error can creep in is through misclassification of drought exposure and TBVC status of birth district. Neither of these errors is likely to be related to measurement error in disability, so we expect estimates of β_1 and β_2 to be biased downwards (Bound, Brown and Mathiowetz 2001).

References

Aigner, Dennis J. 1973. "Regression with a binary independent variable subject to errors of observation", *Journal of Econometrics*, Vol. 1 pp. 49-60.

Bound, John, Charles Brown and Nancy Mathieowitz. 2001. "Measurement error in survey data" in Handbook of Econometrics. (ed) E.E. Learner and J.J. Heckman. Pp. 3705-3843, North Holland

Schmertmann, Carl P. 1999. "Estimating multistate transition hazards from last-move data", *Journal of the American Statistical Association* Vol. 94: 445 pp. 53-63