Published in final edited form as:

Popul Res Policy Rev. 2018 December ; 37(6): 1079–1108. doi:10.1007/s11113-018-9493-1.

Does It Take a Village? Migration among Rural South African Youth

Tyler W. Myroniuk,

(1) George Mason University, Department of Sociology and Anthropology

Michael J. White,

(1) Brown University, Population Studies and Training Center, Department of Sociology

Mark Gross,

(1) Cabrini University, Sociology and Criminology Department

Rebecca Wang,

(1) Brown University, Population Studies and Training Center, Department of Sociology

Carren Ginsburg,

(1) MRC/ Wits Rural Public Health and Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, South Africa

(2) INDEPTH Network, Ghana

Mark Collinson

(1) MRC/ Wits Rural Public Health and Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, South Africa

(2) DST/MRC South African Population Research Infrastructure Network (SAPRIN)
Johannesburg, South Africa

(3) INDEPTH Network, Ghana

Abstract

In a rural African context, the saying, “it takes a village to raise a child,” suggests that community characteristics are substantially important in children’s lives as they transition to adulthood. Are these contextual factors also related to youth migration? Demographers are uncertain about how community characteristics improve our understanding of an individual’s propensity to migrate, beyond individual and household factors. In many low and middle-income country settings, youth become migrants for the first time in their lives to provide access to resources that their families need. We employ discrete-time event history models from the 2003/2011 Agincourt Health and socio-Demographic Surveillance System in rural South Africa to test whether markers of development in a village are associated with the likelihood of youth and young adults migrating, distinguishing between becoming temporary and permanent migrants during this critical life cycle phase. We find that village characteristics indeed differentially predict migration, but not nearly as substantially as might be expected.

Introduction

In the past few decades, scholars have increasingly called attention to the contextually embedded nature of migration in empirical work (Beauchemin and Schoumaker 2016; Bilsborrow et al. 1987; Bilsborrow 2016; Massey, Goldring, and Durand 1994; White and Lindstrom 2005), in response to a well-established literature—theories from both neoclassical economic and the new economics of labor migration perspectives—on individual and household economic drivers of migration. As countries and global regions continue to urbanize and migrants are drawn towards cities for economic reasons, it is important to assess if conditions in one's origin community are associated with out-migration at the community-level, beyond individual and household factors. In low and middle-income country (LMIC) settings, community-level indicators of economic development have been found to be both push and pull factors, or at other times inconsequential, depending on the circumstances and operationalization of “development” (Beauchemin and Schoumaker 2005; Bohra and Massey 2009; Lindstrom 1996; Lindstrom and Lauster 2001; Massey and Espinosa 1997). While these characteristics have clearly been shown to matter in existing research, the question remains: *how much* do they matter?

We seek to deepen demographic knowledge by examining the extent to which village characteristics predict the likelihood of an individual migrating *within one's country*. In this sense, we are talking about understanding the degree to which meso-level factors—those that are intermediate between society-wide changes and individual determinants—influence these phenomena in places undergoing important socioeconomic transitions. With enough variation across local communities (such as the village-level data we assemble), one can better discern the process in which meso-level contexts affect outcomes. To be sure, population re-distribution goes hand-in-hand with economic development. In turn, migration is the key demographic mechanism facilitating this process in most societies. While it is widely acknowledged that migration takes a multitude of forms—ranging from repeated circulation to a straightforward permanent relocation—there is little systematic evidence of exactly how these various forms are understood empirically with respect to meso-level factors.

High demands on data have been noted as a barrier to evaluating these relationships because detailed longitudinal information on individuals, households, and communities is required (Beauchemin and Schoumaker 2016; Bilsborrow 2016; Black and Skeldon 2009). Demographic surveillance systems offer valuable sources of relevant, rich longitudinal data, and we use one such source here: the Agincourt Health and socio-Demographic Surveillance System (Agincourt HDSS). Health and socio-Demographic Surveillance System data, in general,¹ are able to offer key insights into the timing and type of migration behavior undertaken in LMIC settings. Furthermore, and of particular relevance to our subject matter here, the Agincourt HDSS offers a sub-district-wide census database, updated annually since 1992, that connects migrants to their origin households and villages. Agincourt is conducive to such efforts as it is a rural area in northeastern South Africa, where migration—often

¹) As of 2018 there are 48 HDSSs in 19 countries, in Africa, Asia, and Oceania. They are part of the INDEPTH Network and more information can be found at <http://www.indepth-network.org>.

labor migration—is highly prevalent and necessary for families to survive (Collinson et al. 2006; Collinson 2010).² Because there are 26 villages nested within one sub-district, we are able to tease out community-level effects that are not merely attributable to sociodemographic and ecological characteristics, which differ across wider, more scattered locales, as found in most survey research. In our study, any “village effects” we uncover are attributable to highly localized variation within a relatively homogenous geographic setting.

We are keenly interested in how time-varying village characteristics—after accounting for household and individual factors—are associated with the chances of first migrations, as well as temporary and permanent migrations undertaken by youths over the course of eight years, from 2003–2011. Once rural South African youths turn 15 years old, they are unlikely to be migrants following their parents but rather independent, prospective migrants for the first time in their lives. Individuals are first classified as part of the working-age population at age 15 (Statistics South Africa 2012). Prior empirical work in Agincourt shows that the rates of internal migration rise to the levels of prevalence among 15- to 24-year-olds higher than any other age groups (Ginsburg et al. 2016). These migrations overwhelmingly involve moves to urban and other rural areas outside of the Agincourt study site. To our knowledge, researchers have not tested whether village and other community characteristics differentially predict temporary *and* permanent migration in a longitudinal framework, in addition to a first migration. The distinction between a first and repeated migration is important, because it may have varying economic and livelihood consequences for household members, including youth. We acknowledge that research on this relationship is scant and, at times, inconsistent. We are fortunate to be able to work within a sub-district that has appreciable variation across its 26 villages and exploit available time-varying indicators of socioeconomic status of the village: a village-level asset index (derived from households) and village-level educational attainment (derived from individuals aged 24 years and above). The coefficients of variation for both indicate there is adequate variation over villages and time (see Appendix Table 1). Substantively, assets, as an indicator of overall change in village wealth, functionally amplify or mitigate relatively stable inequalities between villages such as soil quality, rainfall, and distance to nearest tarred road. The Agincourt study site also has a well-developed education system where younger cohorts have had better access to and quality of education than older cohorts. Assets and education effectively capture important village-level changes in this context.

By examining the behavior of this population, within a cluster of villages encompassing only about 420 square kilometers, and employing detailed longitudinal data, we have a unique opportunity to isolate the extent to which village characteristics matter for migration. Few data sources in LMICs have this capacity. Migration is prevalent throughout Africa, so our investigation holds promise both for its intrinsic importance to South Africa and for the understanding it can provide for demographic dynamics elsewhere. If it takes a village to raise a child, should it not also take a village to raise a migrant?

² In this context, “rural” primarily refers to limited access to public sector services and infrastructure like tarred roads, electricity, sanitation, and water.

The Presumed Impact of Development on Migration in Rural Areas

Demographers have had little theoretical guidance with respect to how much community and/or village effects ought to contribute when predicting the likelihood of migration. As countries, regions, and other locales develop their economies and infrastructures, more development likely leads to more out-migration from rural areas. Nevertheless, this relationship between development and migration can dramatically change over time—a process in which macro-level societal changes are translated into individual migratory behavior and thus net migration at the community or regional level (Rees et. al 2016). For example, in China, higher rates of investment in some provinces have led to large-scale changes in migration flows—thus drawing individuals in at the expense of other provinces (Poston and Zhang 2008). Such processes lend credence to the argument that such contextual factors play an important role in individuals' and households' decisions to migrate (see Martin and Taylor 1996; Zelinsky 1971). Macro-level economic development theories, although not perfectly aligned with these demographic research questions, indicate that conditions of development at a greater geographic scale ought to significantly influence migration. Yet the temporal scales depicted in this body of research are unclear.

However, there is often no consistent expectation with respect to whether any particular contextual characteristic ought to promote or discourage out-migration. For instance, in Nepal, rural development seems to hinder international migration, and the mere presence of electricity in villages lowers the likelihood of individuals migrating internally (Bohra and Massey 2009). In Burkina Faso, measures of rural development, such as improvements to quality of life measures, have been found to retain individuals, while the development of roads push individuals to migrate to urban areas (Beauchemin and Schoumaker 2005). Within Mexico and the U.S., nonagricultural wage opportunities in rural areas are associated with lower chances of migrating elsewhere (Lindstrom and Lauster 2001), while Mexicans in communities more heavily reliant on agriculture than others have higher chances of migrating to the U.S., except when the presence of a vocational school offers more opportunities to secure higher wages and decreases an individual's chances of moving to the U.S. (Massey and Espinosa 1997).

Aside from wages and other village development characteristics, migration is also influenced by the environmental and political context. In Burkina Faso, more rainfall predicts higher odds of longer-term out-migration (Henry, Schoumaker, and Beauchemin 2004); across dry states in Mexico, declines in rainfall are associated with higher odds of migration to the U.S. (Nawrotzki, Riosmena, and Hunter 2013); and in Indonesia, greater police presence, geographic isolation, and local conflict increases the prevalence of out-migration (Czaika and Kis-Katos 2009). Even the overall prevalence of out-migration from a particular village will influence the chances of individuals migrating away (Massey, Goldring, and Durand 1994). Individuals in these areas might be encouraged to migrate because many others around them have, which, in turn, leads to more established networks in destinations (Massey and Espinosa 1997).

Despite these contributions to migration theory, the macro-level unit of analysis—the geographic scope of development—is often ambiguous (de Haas 2008; de Haas 2010). Most

of the studies listed above work at larger geographic scales than is our concern for this investigation. As is well documented in the geography literature, measurement of a demographic phenomenon of interest is sensitive to geographic scale, potentially leading to substantial bias in estimates (Fotheringham and Wong 1991; Leyk et al. 2012). Over larger geographic units, contextual characteristics often become more heterogeneous, and, thus, associations between contextual measures and outcomes tend to be stronger at smaller geographic units. Given the difficulty in ensuring migration processes operate at, or correspond to, the geographic scale employed in a study (often an administrative unit), contextual effects may easily be over- or underestimated (Fussell 2014; Montello 2001). Additionally, the exact mechanisms—and their explanatory power—linking community characteristics and migration are difficult to ascertain.

Youth and Young Adult Migration in sub-Saharan Africa

Researchers have defined a host of reasons as to why individuals might migrate, but in LMICs, such as those in sub-Saharan Africa (SSA), this research tends to focus on individual and household determinants. Because large shares of households in LMICs have limited economic means, migration can serve as a mechanism for both upward mobility and managing risks. The household members who migrate occupy an alternative economic niche, such as an urban job, counterbalancing the household's rural livelihood. Household members often make a collective, strategic decision as to whom in the household might yield the best economic returns through migration (Stark and Taylor 1991). Often it is a household's younger members who are valuable economic resources. Youths can send remittances to their families via labor or through marriage into a wealthier family. Of course, this is predicated on a family's abilities or desires to send their youth away. In many LMIC settings, including those in SSA, the poorest often cannot afford to do so while the wealthiest do not need to do so (Dustmann and Okatenko 2014; Masson 2001; McKenzie and Rapoport 2007).

In recent years, other factors such as co-residential context have been tested as potentially important determinants of youth migration in African contexts, even if the direction of effects or causal order are unclear. For example, children in South Africa whose mothers are temporary migrants are more likely to move to another household (although often nearby as opposed to long distances) rather than remain at home with another adult, unless that adult is a "maternal substitute" (Madhavan et al. 2012, p. 713). Such transitions are facilitated by the presence of extended kin in their destinations (Madhavan, Mee, and Collinson 2014). While child migration is often a response to circumstances concerning parents or caregivers, youth migration is much more likely to be related to employment opportunities. In some resource-deprived contexts, such as rural Malawi, labor migration is an expected journey that a young individual should take as part of the process of moving into adulthood (Englund 2002). Perhaps this parallels another setting: the suggested "rite of passage" undertaken by Mexican young adult males in efforts to become independent of family, and potentially prepare for marriage (Kandel and Massey 2002). Still, some aspects of migration are out of the control of young individuals, such as becoming orphaned due to HIV/AIDS, which places children in dire circumstances that might require several migrations in order to find the best match of foster parents or guardians (Ansell and Young 2004). In the end, much of this research

emphasizes that these migratory decisions are strongly influenced by individuals, immediate families, and extended families. However, analyses are often restricted by data that cannot simultaneously track changes in additional inputs into a young person's transition into adulthood via migration (Juárez et al. 2013), such as household or meso-level contextual changes.

Background: Agincourt, South Africa

In this paper, we study a population from villages located within a geographically defined area: the Agincourt HDSS, a rural sub-district in the northeastern part of South Africa near the border of Mozambique. This area is part of a former homeland where Black South Africans were forcibly resettled between 1920 and 1970 under apartheid regulations (Madhavan et al. 2008). Located next to the Kruger National Park conservation area, this rural community is more aptly described as *remote* rather than *agricultural*. The land is arid, unproductive, and better suited for game farming rather than large-scale agriculture (Collinson et al. 2009). Infrastructure developments since the democratic transition in 1994 have occurred more slowly than expected, and to date, household sanitation remains rudimentary as most villagers lack piped water to their dwellings (Kahn et al. 2012).

The overwhelming majority of youth in the Agincourt HDSS grow up in impoverished communities and face poor employment prospects. As is typical for much of the rural interior in South Africa, local unemployment rates remain high at 25% for men and 48% for women (Blalock 2014). Additionally, informal sector activities are pervasive. It is common to find vendors selling food and other items along the numerous unpaved roads. Many working-age individuals thus engage in labor migration and opt to take the full-day bus journey to Gauteng Province, the economic engine of the country (and the continent) where greater Johannesburg and Pretoria are located.

Historically, the Agincourt HDSS has long served as a major migrant-sending community. During the apartheid era, the region containing the Agincourt study site was a source of labor—under restricted and cyclical movement—for the industries of Gauteng, including mining. At the end of apartheid in 1994, Black South Africans in rural areas migrated freely — sometimes permanently and sometimes in a temporary and cyclical fashion—to urban parts of South Africa (and elsewhere in the country). This newer migration occurred at much higher rates than in the apartheid era due to the abolishment of travel and residency restrictions based on race (also known as the “Pass Laws”) in 1986 (Reed 2013). As such, both out-migration and return-migration are highly prevalent in the Agincourt HDSS (Collinson et al. 2006).

In the contemporary Agincourt HDSS setting, as in the broader LMIC settings described above, households draw on migration to maximize their economic achievement. Poor households tend to rely on men to temporarily migrate for work, with the poorest households needing to rely additionally on female labor migration (Collinson 2010). Remittances from household members who have migrated, on either a temporary or a permanent basis, offer access to important resources for many origin households.

This study site offers a unique opportunity to investigate the relationship between mesolevel community factors and internal migration. While this region shares many features typical of major migration sending areas in southern Africa, discernible differences in socioeconomic conditions and infrastructure developments exist across Agincourt HDSS villages. This is partially due to geographic positioning. Variation in rainfall patterns within the site (which follows an increasing East-West gradient, meaning that villages further east, and therefore also closer to Mozambique and containing more Mozambican residents, receive less rainfall) indicate there are significant differences in the spatial distribution of natural resources (Leyk et al. 2012). In line with the “environmental capital” hypothesis, Hunter et al. (2013) demonstrate villagelevel availability of natural resources in the Agincourt HDSS is associated with more economically secure households that can support temporary migration. Ease of transportation is another a major source of variation between villages. The most developed villages lie along—or within several kilometers of—the major tarred road running through the sub-district. In contrast, a large fraction of other residents must travel anywhere between 10 km and 23 km to reach this major thoroughfare. Thus, residents of these villages, which account for over half of all villages, have greater difficulty accessing surrounding towns, public service provision, and shops that sell basic household items (Hargreaves 2000).

Both quantitative survey data and qualitative interview data suggest the village-level socioeconomic variation is *broadly* summed up in the following four categories (Hargreaves 2000): 1) Central villages: Well-established South African villages, centrally placed in the Agincourt HDSS sub-district; 2) Established: Fairly well-developed villages—on or near the main tarred road; 3) Underdeveloped: More remote villages that are fairly isolated; 4) Refugee villages: Mainly (former) Mozambican refugee settlements. These categorizations have remained quite stable over the past two decades (unpublished, Agincourt HDSS documentation). While we do not need to use these for our analyses, nor are the categorizations immutable, they offer more insight into this research context.

In addition to geographic variation, demographic variation is found among Agincourt HDSS villages. In 2000, just prior to the beginning of our analytic observation period, the Agincourt HDSS was marked by great variation in fertility and mortality between villages. Total fertility ranged from 2.0–4.5 children and age standardized mortality rates ranged from roughly 43 per 1,000 to 150 per 1,000 (Hargreaves 2000).

Hypotheses

Despite acknowledging the great potential for the influence of contextual effects on migration, researchers have given limited direct attention to exactly how much explanatory power meso-level characteristics—those of villages in our case—add to our understanding of the determinants of migration, regardless of whether they are push or pull factors. We hypothesize that when predicting youth migration:

H1: Models incorporating time-varying village effects will have *substantively* more explanatory power than those with only individual and household characteristics.

We test this hypothesis precisely because the literature suggests that this is both likely and substantively important. We recognize that additional regressors will increase variance explained; however, our concern is whether the inclusion of village effects is truly meaningful after accounting for the added parameters.

Our key village-level measures—wealth and educational attainment—could be indicative of higher quality of life and reflect village residents' economic advantages, which have been linked to retaining individuals (Beauchemin and Schoumaker 2005; Massey and Espinosa 1997). Despite a lack of clear guidance from recent empirical work in other contexts, migration from rural to urban areas in South Africa—and here in the Agincourt study area—is, on average, undertaken by individuals and households that are seeking upward economic mobility rather than undertaken by only those who are well-off. Although the *direction* of village effects and their associations with migration is of secondary importance in this paper, we nonetheless hypothesize that:

H2: When village resource-laden characteristics, such as wealth, or educational attainment, increase, the likelihood of all types of migration ought to decrease.

Analytic Methods—Since the Agincourt HDSS is an annual census, it contains in-depth demographic and other household data on all individuals living in 26 villages over an area of approximately 420 square kilometers. This study site is ideal for our purposes as geographic mobility is high in this population. Additionally, the Agincourt HDSS has the benefit of providing a target population from a circumscribed geographic area that, crucially, remains sensitive to local variations in contextual characteristics. The Agincourt HDSS has over twenty waves of harmonized data, reflecting the experiences of more than 90,000 people who have lived in the sub-district since 1992 (Kahn et al. 2012). Individuals can enter the sample via birth or in-migration, and leave via death or out-migration. What surveillance systems lack in generalizability, they make-up for in detailed, dynamic, and consistent measurements of entire populations and origin households over time—particularly with respect to migration. But, more importantly, the small area (relative to an entire country or region) allows us to eliminate oft-unobserved and confounding traits that limit the extent to which the effects of villages are accurately assessed in most other demographic research. With that said, at least in the Agincourt HDSS, trends in fertility, migration, and mortality are similar to those at the national level in South Africa (Kahn et al. 2007).

As is common with surveillance systems, individuals who permanently leave Agincourt are not tracked. Thus, the most useful analyses are done prospectively up through the end year, the event, or censoring. However, those who leave their household on a temporary basis—most often outside the surveillance area boundary—continue to be followed in the Agincourt HDSS. Still, there are some years—accounting for a small fraction of observations—where an individual's residence status is not resolved, and residence status is treated as missing. In our present analysis with a sample of younger individuals, attrition prior to the event of interest (migration) is uncommon.

To address our research questions on the relationship between village characteristics and differences in migration propensities, we employ two types of event history analyses: 1) discrete-time event history analyses to examine the chances of experiencing a first migration; and 2) discrete-time, competing risks event history analyses to examine the chances of temporary and permanent migration, simultaneously, relative to non-migration in a given year (see Gray and Mueller [2012] for an example in Ethiopia employing these regression techniques). We restrict our analyses to the years 2003 to 2011 (inclusive), since data collection on household assets (a key variable, which we lag by two years to capture temporal ordering of predictors of migration, described below) began in 2001. As we alluded to earlier, males and females who turned 15 years old, in each year from 2003 to 2011, were selected to represent the group of youth most at risk of migrating. In essence, we devise a person-year file from 2003 to 2011 with the addition of a new cohort of 15-year-olds each year. Thus, we observe some of our sample of youths, as they become adults, reaching up to age 23 (in a few instances age 24, depending on the timing of interviews) over the observation period. Although item non-response exists in our analytic sample, we do not consider it extensive enough to invoke multiple imputation techniques; listwise deletion is employed.

Duration to First Migration.: In the models estimating the chances of a first migration, 31,293 records from 7,381 males and 29,138 records from 7,373 females are used. Person-years are dichotomized into “Did Not Migrate” (=0) and “Migrated” (=1) for each year. Our observation window begins the year the individual turns 15 and continues until the person migrates, is censored through death (89 individuals or 0.6% of the entire sample), or reaches the end of the window in 2011. In this basic model, we consider duration to any migration without regard to type of move. We employ logistic regression models and account for the clustering of village effects (see Cameron and Miller 2015; Abadie et al. 2017 for examples justifying the value of clustering at higher levels of aggregation, and not the individual). Lastly, village-level fixed effects parameters are not explicitly estimated, although such modelling can be found in the appendix and the consequences of this choice are discussed below. Equation 1 depicts the estimation procedure for our full models:

$$\ln\left(\frac{p_{ijt}}{1 - p_{ijt}}\right) = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + \beta_3 V_{jt} \quad 1$$

The outcome is the log odds of first migration for an individual i , in village j , in a year t . The model intercept is β_0 . $\beta_1 X_{it}$ represents a set of time-varying, t , covariates (with their associated slopes) for an individual, i , while $\beta_2 Z_i$ represents a set of time-invariant covariates for an individual. Time-varying village level characteristics j are captured in $\beta_3 V_{jt}$. All of these variables are described in detail below, after outlining estimating procedures. We draw our observations from the 21 villages that have been part of the Agincourt HDSS since its inception. Another 5 villages added in 2007 are excluded, since this change occurred after our observation window opened.

Competing Risks of Migration by Type.: In models estimating the competing risks of migration, a total of 7,885 males, contributing upwards of 35,840 person-years, and 7,893 females, contributing upwards of 33,405 person-years, are employed. We categorize person-years into three potential outcomes for each year between 2003 and 2011: never moved (=0), temporary migration (=1), and permanent migration (=2). Temporary migration is a repeatable event and a competing risk with permanent migration since one cannot be both at the same time; time lapses between temporary migrations are allowed in this modelling strategy, and temporary migration episodes often lead to permanent migrations for individuals in Agincourt.³ We take advantage of annual information on an individual's residence status within the household and the event that "terminated" their membership in that household. Temporary migrations are defined by residing in the origin household for less than six months of the year, while still being considered a household member. Permanent migration is defined by migration—change of residence to a location *outside of the surveillance sub-district*—and a change in status to no longer being considered a household member. We also assign a permanent migration for year t to a circumstance in which a temporary migration (in year t) is followed by a permanent migration (in year $t+1$), since often the nature of the move (in the surveillance database) is only fully resolved in the year $t+1$ household visit. Empirically, permanent "migrations" *within the surveillance sub-district* are a small, albeit problematic subset of migrations for the purposes of these analyses, and not identified here.⁴ Conceptually, this type of permanent "migration" is more appropriately a form of local mobility, which is not the focus of this research. Death is not modelled as a competing risk due to the aforementioned very small number of cases. However, individuals' records are included up until death, when they are censored from the analyses. Since the Agincourt HDSS observation is discontinued once a person becomes a permanent migrant, such individuals are effectively censored. The end of the observation period, 2011, also censors our observations.

For the purposes of these analyses, we model these cause-specific hazards simultaneously by employing multinomial logistic regression models and account for the clustering at the village level. As in our first set of regressions, village fixed effects are excluded, but sensitivity analyses and the consequences of this choice are discussed below. Equation 2 depicts the estimation procedure for our full models:

$$\ln\left(\frac{p_{sijt}}{p_{rijt}}\right) = \beta_{0st} + \beta_{1s}X_{sit} + \beta_{2s}Z_{si} + \beta_{3s}V_{sjt} \quad 2$$

The outcome here is the log odds of temporary or permanent migration s , treated as separate events, occurring relative to the reference outcome r of not having moved for an individual i ,

³) This is analogous to a hypothetical biomedical model where an individual acquires a disease, is cured, relapses at a later point, and then dies.

⁴) In the entire Agincourt HDSS, we estimate that there is less than 3% attrition due to those leaving for another village within Agincourt—most of which is due to marriage migration (which we address in detail below). Although it is possible to reconcile these individuals' migration spells with the Agincourt HDSS, complete reconciliation of individual IDs is compromised because of the timing of spell exposures recorded (between moving out and recapturing individual at another census wave). The second spell in this type of move will not have information on the migrant's history and, if it did, would very likely include an unidentifiable time gap between residence spells.

in village j , in a year t . The intercept in this model is β_{0st} . The term $\beta_{1s}X_{sit}$ is included to capture the effect of a set of time-varying t covariates for an individual, i , in migration state, s , while $\beta_{2s}Z_{sj}$ represents a set of time-invariant covariates for an individual. Time-varying village level characteristics, j , are captured by $\beta_{3s}V_{sjt}$. As in Equation 1, only 21 villages are used for these analyses.

In Equations 1 and 2, we estimate the effects of two push and/or pull migration factors at the village-level. The first time-varying independent variable of interest (in each year from 2003–2011) is a village-level composite, 11-item durable goods asset score (*Village Assets*). It is aggregated, as an additive index, from a household inventory (items include the presence of a toilet, a tap in the house or yard, electricity, a fridge, a television, a radio, a cellphone, a car, a video/DVD player, a stove, and a living room) and lagged two years (the 2001 information informed the 2003 person-year). We exclude the values from a youth's own household, to reduce the extent to which collinearity exists on the left- and right-hand sides of our equations. The second time-varying independent variable is the average number of years of education in a village for individuals 24 years and older (*Village Years Educ.*), and thus older than any youth in our risk set (aggregated from individual measures, also lagged two years). The information strongly suggests that villages are distinct from one another along these measures, despite their relatively close physical proximity. Due to resource limitations for the Agincourt HDSS fieldwork, education and household assets are collected every other year. For instance, if there is information on education in 2001 and 2003, but not 2002, the 2001 value will be assigned to 2002. Still, in creating village-level indicators, each value for each year is unique since some individuals' records from 2001 might not count towards a village-level measure in 2002 due to an individual's absence or attrition in the Agincourt HDSS. This feature limits the extent to which we can match the timing of migration episodes with these key time-varying co-variates. However, this is a known methodological sacrifice that comes with answering such migration questions when using surveillance site data (see Bocquier 2016). We also control for the changing prevalence of any type of village out-migration (*Village % Mig.*), which is lagged two years; values from our sample youth are excluded in these calculations as well.

Our time-varying controls include continuous age; a dichotomous measure of whether a youth matriculated (completed) secondary school (which is more reliable than a household indicator of highest education—given prevalent fluctuations in household composition for Agincourt HDSS residents—or education of the household head since Black South African youth have differing access to education than their parents had during apartheid; also age at matriculation is typically 18 and most youth in the sample reach this age prior to censoring; the asset score of the observed individual's household; and the number of migrations for respondents, once they turned 15 years old (thus, assumed independent migrations and not including migrations prior to age 15). We do not include a time-varying marital status variable—despite marriage being a prominent predictor of migration, generally in SSA—for empirical reasons: 1) the Agincourt HDSS only asks respondents about union status and few youths answer the questionnaire during fieldwork (often the head of the household, or other adult household member); and 2) even if we utilize a relatively unreliable and highly gendered “reasons for migration” variable to identify and exclude all individuals who listed marriage as a reason for migration (less than 5% of our sample, and 99% female), our results

do not substantively differ from our main analyses (sensitivity analyses, not presented). Unlike other SSA settings where marriage is likely to be a much more definitive factor, youth migration due to marriage is uncommon in Agincourt. With the exception of a negligible number of cases where education was described as a reason for migrating, almost by definition, every other migration would likely be for economic circumstances.

Our time-invariant controls include cohort (year the individual turned 15 and entered the risk set); gender; whether an individual is a Mozambican refugee (due to Agincourt's proximity to Mozambique's southwest border, and this sizeable population likely has distinct migration behaviors compared to other Agincourt residents); and each individual's familial position in the household (child, head, or other).

Our modelling strategy assesses the extent to which time-varying village characteristics and/or time-invariant village characteristics better explain individuals' propensities to migrate. Our models estimate: 1) individual/household effects only; and 2) individual/household effects and village time-varying effects. These different models are marked in our regression tables below. We also examine three measures of goodness of model fit in our analyses: Pseudo R^2 , Akaike information criterion (AIC), and Bayesian information criterion (BIC).⁵

Descriptive Statistics

At the individual level, some 35.7% of males and 42.5% of females in our sample experienced at least one migration between 2003 and 2011. Table 1, however, describes our sample in terms of person-years. We see the greater representation of males among temporary migrants and females among permanent migrants. There are very modest, if any, differences by gender for most covariates. Notably, differences are quite modest for school matriculation and "other" status (in relationship to the household head). We also observe the structural condition that those who are in earlier cohorts contribute a larger share of the person-years we analyze.

To get a better sense of males' and females' propensities to migrate and the extent to which they are different, adjusting also for exposure, we present Kaplan-Meier survival estimates. Figures 1, 2, and 3 respectively depict the duration to any migration, duration to first temporary migration, and duration to first permanent migration. By age 22, about half of all males will have experienced their first migration, whereas half of females will have done so by age 21. Gender differences are modest for the onset of temporary migration; between the ages 20 and 21, roughly a quarter of both males and females have undertaken temporary migration. For permanent migration, the gender differences are more pronounced; by age 19 roughly a quarter of females will have permanently migrated, while less than 25% of males permanently migrate before age 24. The differences by gender in both figures are likely related to permanent migration due to marriage for young women. Nonetheless, these figures confirm that migration is highly prevalent among Agincourt's youth.

⁵) Akaike information criterion (AIC) and Bayesian information criterion (BIC) are interpreted, as measures of goodness of model fit, per the suggestions of Raftery (1996) and Burnham and Anderson (2004).

First Migration

Table 2 provides estimates of youth males' and females' chances of a first migration, regardless of whether it is temporary or permanent. These models, as structured, give us insight into the relative explanatory power of individual/household characteristics and time-varying village effects. When comparing models 1–3 and 4–6, we find that village-level characteristics contribute possibly 1) minimal explanatory power (Pseudo R^2); 2) a substantial amount of explanatory power for males and a modest amount for females (AIC); and, 3) no additional explanatory power compared to individual/household characteristics (BIC). In these modelling specifications, the extent to which village characteristics add explanatory power is unclear.

In sensitivity analyses (Appendix Table 2), fixed village characteristics (in the form of village dummy variables) were included in addition to each model above in Table 2, respectively. In each model specification, it appears that the inclusion of fixed village characteristics—such as geographic location and soil type—add substantially more explanatory power than those modeled in Table 2, above, based on AIC and BIC goodness of fit statistics; Pseudo R^2 indicates only marginal improvements. Yet, Hausman Tests indicate that the inclusion of fixed village characteristics does not change coefficients in a systematic way, which suggests that models without these characteristics are appropriate specifications.⁶

With respect to specific predictors, as expected, age, cohort, schooling, position in the household, and whether one is a Mozambican refugee (only in the case of females, though) predict migration out of the origin household. We see that higher time-varying village prevalence of migration (*Village % Mig.*), from year to year, is associated with higher chances of male migration (models 2 and 3, $p < .001$), while we only see a similar, marginal effect for females (model 6, $p < .10$). The time-varying village assets and education indicators are not significantly associated with male or female migration.

We consider, however, that the pooling of all types of migration in these models may obscure variation in the associations of individual or village effects with specific type of moves. For that reason, we now move to analyses of the competing risks of migration by type.

Competing Risks of Migration

Table 3 estimates models (under the multinomial logit specification with repeated measures) of male and female youth's chances of temporary (Temp) and permanent (Perm) migration, relative to non-migration in any given year (No). As in Table 2, these models test whether including only individual/household effects, or adding in time-varying village effects contribute most efficiently to explanatory power in these migration choices. Like the results

⁶ In several instances, the Hausman Test is undefined because of issues with degrees of freedom from clustering on villages and including fixed village characteristics. To our knowledge, there are no empirical guidelines in remedying the results of this test even if employing “wild bootstrapping” procedures that could potentially more effectively estimate models with village clustering and dummies.

found in Table 2, when adjudicating the explanatory power of village characteristics with Pseudo R^2 , AIC, and BIC—as model goodness of fit guides—no clear pattern emerges. For both males and females, the Pseudo R^2 statistics indicate effectively no additional village explanatory power in estimating males' and females' temporary and permanent migration compared to assessing only individual/household effects when considering time-varying village characteristics. AIC indicates that models including time-varying village characteristics might be a substantially better fit for explaining the competing risks of migration. Nevertheless, BIC suggests that these village effects add a small amount of explanatory power for males (models 2 and 3) and no additional explanatory power—beyond individual/household characteristics—for females (models 5 and 6).

In sensitivity analyses (Appendix Table 3), fixed village characteristics (in the form of village dummy variables) were included in addition to each model above in Table 3, respectively. Based on AIC and BIC, in each model specification, it appears that the inclusion of fixed village characteristics add substantially more explanatory power than those modeled in Table 3, even if there are differences between the estimates within Appendix Table 3. Again, Pseudo R^2 indicates only marginal improvements in explanatory power and model fit in Appendix Table 3, compared to Table 3. As in regressions predicting youths' first migration, Hausman Tests indicate that the inclusion of fixed village characteristics does not change coefficients in a systematic way, which suggests that models without these characteristics are appropriate specifications.⁷

We find that one's number of prior migrations and household asset score (as an indicator of household wealth) strongly and consistently predict both temporary and permanent migration for both genders. More prior migrations are linked to greater chances of temporary and permanent migration. More household assets are linked to greater chances of temporary migration, but lower chances of permanent migration; individuals from wealthier households might be taking advantage of having the ability to move and work away from home to improve the household's economic circumstances, while not desperately needing to permanently move away from Agincourt to improve one's livelihood as they come from relatively privileged households in the first place. These results indicate the value of distinguishing migration type. Other individual/household covariates, such as age, household role, cohort, and education predict temporary and permanent migration similarly for males and females, while one's Mozambican refugee status is only associated with the chances of migration for females. We also include the measure of village migratory prevalence, which is positively linked to temporary migration for males, and only marginally so for females (model 5).

Discussion

Demographers have long-called for the measurement of the contextually embedded nature of migration, and thus, whether community-level characteristics predict migration

⁷) As with Appendix Table 2, there are degrees of freedom issues with several Hausman Tests in Appendix Table 3, with no clear empirical remedy. The coefficients for time-varying village characteristics of assets and education are positively associated with permanent migration, relative to not migrating in a given year. Only for males, is village level of education positively—albeit marginally—associated with temporary migration ($p < .10$).

(Beauchemin and Schoumaker 2016; Bilsborrow et al. 1987; Bilsborrow 2016; Massey, Goldring, and Durand 1994; White and Lindstrom 2005), as theorized in macro-level economic development literature (Martin and Taylor 1996; Zelinsky 1971). This line of inquiry has advanced migration theory and shown that the context in which individuals reside—beyond the household—is important in predicting migration behavior. In the limited available empirical research, scholars have found a wide range of associations between characteristics of community development—such as electricity, the availability of non-agricultural wages, paved roads, or a combination of a number of other factors—and the likelihood of out-migration. In fact, such proxies for development have been found to be sometimes positively associated with these markers, sometimes negatively associated, and sometimes not at all associated (Beauchemin and Schoumaker 2005; Bohra and Massey 2009; Lindstrom 1996; Lindstrom and Lauster 2001; Massey and Espinosa 1997).

Although the mixed direction of meso-level effects might be expected when different types of community characteristics are employed to estimate the chances of migration in different settings, an unanswered “first principles” question is simply how much explanatory power these contextual factors contribute, especially in an estimation structure where meso-level effects cannot be driven by variation at a higher-level of geography. In our approach, we are able to assess local community effects—a village within a sub-district—and allow these to complement the conventional concerns for the influence of individual and household factors in the Agincourt HDSS, South Africa, where internal, out-migration, is highly prevalent.

Our results are indeed more complex, as they allow for influences at multiple levels—individual, household, and village—and also take into account different forms of migratory behavior. Overall, our results align with conventional models of the individual and household determinants of micro-level migration in LMICs. Yet our results challenge expectations about meso-level contextual effects that we drew from the prevailing literature. This reflects in a more nuanced way, we argue, the livelihood experiences of many migratory (and potentially migratory) persons in LMIC settings.

Hypothesis 1 tested the notion that variation in local village-level conditions (within subdistrict) would add appreciably to predicting migration. By assessing three related, although distinct, measures of goodness of fit (in addition to supplemental sensitivity analyses) for the relationship between individual, household, and community characteristics and migration among female and male youth migrants, we find that time-varying village effects could add substantial explanatory power beyond typical individual and household characteristics based on AIC, but Pseudo R^2 and BIC measures suggest that these village effects only marginally add to the overall explanatory power of these models. These three goodness of fit indicators weigh factors that go into the regression modelling process differently. Most evidently, BIC is less sensitive to sample size than AIC and more robust to Type II errors (see Dziak et al. 2012). Thus, BIC suggests that individual and household characteristics explain the most about youth migration, while AIC suggests village characteristics add appreciably (Burnham and Anderson 2004; Raftery 1996). In general, the Pseudo R^2 measure also tends to increase—implying better model fit—when additional parameters are added in regression models. Since Pseudo R^2 values improve only marginally with the addition of village effects, in line with BIC, it seems that the time-varying

indicators—average village assets and education—as well as the control for village migration prevalence do not add much explanatory power in understanding why Agincourt youths migrate.

Our second hypothesis homed in on the type of migration experience and tested the notion that village-level characteristics would add appreciably to the ability to predict temporary versus permanent (versus no) migration. Different threads of the extant literature offer alternative expectations on this point. On the one hand, village-level resources (assets and educational experience) might promote migration, as these features translate into resources that can help the new migrant or provide a flow of information. On the other hand, such resources point to smaller differentials with potential destinations and less incentive to move.

We postulated that, across villages, higher average household assets and average years of education (means across villages and across years) would be associated with males and females having lower risks of becoming migrants based on prior research and the fact that in a rural, somewhat remote, South African district, relative deprivation plays an important role in individual and household decision making. One would expect those who are comparatively disadvantaged, via residing in comparatively disadvantaged villages, to try to improve their economic standing with more extreme tactics, such as migration. Our empirical analysis indicates very little evidence for such effects, and the opposite in several instances for female temporary migration.⁸

Our results bring forth questions about the most suitable geography, time frame, and type of village effects to answer questions about meso-level effects. In these data, we see changes in the overall levels of household assets and average years of education, although not dramatically (in an 8-year window), even in a country experiencing rapid social and economic change. Some theorists (de Haas 2008; Martin and Taylor 1996; Zelinsky 1971) have described, seemingly longer—yet unspecified—periods of economic and developmental change alongside shifting patterns of migration. Yet, meso-level theory—unlike macro-level theories posited by likes of de Haas (2008); Martin and Taylor (1996); or Zelinsky (1971)—has not explicitly expressed why variation across or within villages might occur, nor how much variation is necessary to distinguish villages and locations within villages from one another. One of the downsides of a geographically homogenous area, albeit one with between-village variation in key characteristics, as is an important aspect of our empirical identification strategy, people live, work, and interact in close proximity. Information thus travels more quickly, and these geographic subareas have somewhat interconnected labor markets. We might not be able to observe the right type of village variation important in influencing migration. So, we cannot discount economic theory about the importance of higher levels of geography and its influence on migration decisions.

To expand upon this line of research, multiple sites and a wider array of markers of meso-level characteristics over a longer-span of time would be necessary in order to flesh out the

⁸) Analyses found in Appendix Table 3 found consistently positive effects of village assets and education on the chances of migration for males and females.

broader relationship between development and migration. Although data are not well standardized across INDEPTH surveillance sites, and different social and economic contexts show vastly different migration determinants and outcomes, there are efforts underway within South Africa itself to harmonize the HDSS data across two other sites in Limpopo and KwaZuluNatal provinces. We also limit our analyses to individuals who enter the observation window as 15-year-olds. In doing so, we also are considering cohorts of potential migrants who have shared many similar experiences throughout their lives, in addition to the expanded migratory opportunities afforded to Black South Africans after the end of apartheid. Even though these youths have parents who experienced the settlement restrictions of apartheid, the migration choices observed in this study were independent of legal restrictions and therefore an important window into understanding the future of rural out-migration in South Africa. Macro-level theorists have considered entire populations and their prospective migrants, which includes all age groups. Although including all individuals who could move would be an ideal scope in which to test how much village characteristics predict migration in the Agincourt HDSS, it was important to minimize endogeneity concerns associated with age heterogeneity and potentially extensive prior migration histories of adult residents. Future analyses ought to consider a greater diversity of birth cohorts, while at the same time employing a strategy to account for the inevitable heterogeneity in past livelihood and migratory experiences among these cohorts, to avoid serious bias in any results.

The positive association between education and temporary migration (in our main regression tables, females) and permanent migration (in Appendix tables) suggests Agincourt HDSS communities with higher levels of education experience human capital flight whereby presumably more highly educated young people tend to look farther afield for employment and economic returns on their education. While not definitive, these results are consistent with data from the National Income Dynamics Study (NIDS) showing that rural areas have been undergoing a process of de-agrarianization over the past 20 years (Daniels et al. 2013). Manufacturing and other non-agricultural employment sectors commonly found in rural areas have also seen declines in the share of rural residents they employ. One obvious policy response to any “brain drain” effect—should this truly be the case (migration for unskilled employment is also prevalent in the Agincourt HDSS, see Ginsburg et al. 2016)—would be to promote employment and livelihood opportunities for higher skilled or more educated workers in these rural areas that could help to insulate against this human capital loss. Alternatively, strategies to direct resources to individuals and households adversely affected by the departure of human capital could also prove beneficial.

Policies promoting (or at least acknowledging) temporary, rather than permanent, migration might also act as a buffer against human capital loss in rural areas. Investments aimed at promoting infrastructure and better transportation between rural villages and major migrant destination areas like Gauteng Province, could facilitate stronger linkages between migrants and their origin villages and households. For instance, railways connecting Johannesburg and other destinations in Gauteng already exist but are unused for mass transit purposes. Developing mass transit infrastructure would additionally provide employment opportunities along those transportation corridors.

Furthermore, policies that promote ease of remitting might be of particular benefit. Remittances are often a major motivating factor in the decision of an individual or household's migration decisions and play a major economic role for many South Africans; therefore, making the process of remitting easier and less costly would assist in the accumulation of origin household assets irrespective of whether this promotes or retards migration itself. The recent failure of Vodacom's M-Pesa in South Africa, a widely successful mobile banking and remittance sending platform in Kenya, suggests that there is still a great deal of room to develop innovative financial products, many of which could offer opportunities to more easily send and receive remittances (Mbele 2016; Mcleod 2016).

The research presented here builds towards a meso-level theory of the extent to which village characteristics matter in migration. We confirm that several conventional individual and household traits are quite important in this setting characterized by dramatic socioeconomic transition. Our results also point to the critical importance of understanding the relative impact of context at its various geographic levels with respect to migration. The micro-to-meso mechanisms driving this relationship are not fully understood, which makes this an area of migration research that is primed for expansion.

Acknowledgements:

We would especially like to thank Andy Foster, Margot Jackson, and David Lindstrom for their technical advice, in addition to Steve McGarvey, Leah VanWey, Zhenchao Qian, Taryn Dinkelman, and Jo Fisher for their constructive comments on earlier versions of this paper. This paper was presented at the 2016 Population Association of America annual meeting in Washington, DC. We would also like to thank Yashas Vaidya for consolidating Agincourt Health and socio-Demographic Surveillance System (Agincourt HDSS) data into a harmonized panel file used in this research. We are grateful for support of this project from the Providence/Boston Center for AIDS Research (CFAR) under NIH grants P30AI042853, the Eunice Shriver National Institute of Child Health and Human Development 1R01HD083374 (Brown University); NIH R24HD041020, P2C HD041020 (Brown University Population Studies and Training Center); and additional support for the maintenance of the Agincourt HDSS from Wits School of Public Health, Wellcome Trust (grants 058893/Z/99/A, 069683/Z/02/Z, 069683/Z/08/Z), Medical Research Council of South Africa, and Brown University's Population Studies and Training Center. We gratefully acknowledge the South African Medical Research Council for funding Carren Ginsburg's Career Development Award.

References

- Abadie A, Athey S, Imbens GW, & Wooldridge J (2017). When should you adjust standard errors for clustering? (No. w24003). National Bureau of Economic Research.
- Ansell N, & Young L (2004). Enabling households to support successful migration of AIDS orphans in southern Africa. *AIDS care*, 16 (1), 3–10. [PubMed: 14660139]
- Bilsborrow RE, McDevitt TM, Kossoudji S, & Fuller R (1987). The impact of origin community characteristics on rural-urban out-migration in a developing country. *Demography*, 24 (2), 191–210. [PubMed: 3609405]
- Bilsborrow RE (2016). Concepts, definitions and data collection approaches In White MJ (Ed.), *International handbook of migration and population distribution* (pp. 109–156). New York and London: Springer.
- Black R, & Skeldon R (2009). Strengthening data and research tools on migration and development. *International Migration*, 47 (5), 3–22.
- Beauchemin C, & Schoumaker B (2005). Migration to cities in Burkina Faso: Does the level of development in sending areas matter? *World Development*, 33 (7), 1129–1152.
- Beauchemin C, & Schoumaker B (2016). Micro Methods: Longitudinal Surveys and Analyses In White MJ (Ed.), *International handbook of migration and population distribution* (pp. 175204). New York and London: Springer.

- Blalock CL (2014). Labor migration and employment in post-apartheid rural South Africa Unpublished dissertation. University of Colorado-Boulder.
- Bocquier P (2016). Migration analysis using Demographic Surveys and Surveillance Systems In White MJ (Ed.), *International handbook of migration and population distribution* (pp. 205-223). New York and London: Springer.
- Bohra P, & Massey DS (2009). Processes of internal and international migration from Chitwan, Nepal. *International Migration Review*, 43 (3), 621–651. [PubMed: 21423821]
- Burnham KP, & Anderson DR (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304.
- Cameron AC, & Miller DL (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372.
- Collinson M, Tollman SM, Kahn K, Clark SJ, & Garenne M (2006). Highly prevalent circular migration: households, mobility and economic status in rural South Africa In Tienda M et al. (Eds.), *Africa on the move: African migration and urbanisation in comparative perspective* (pp. 194–216). Johannesburg: Wits University Press.
- Collinson MA, Clark S, Gerritsen A, Byass P, Kahn K, & Tollman S (2009). The dynamics of poverty and migration in a rural South African community, 2001–2005. University of Colorado-Boulder Institute of Behavioral Science Working Paper.
- Collinson MA (2010). Striving against adversity: the dynamics of migration, health and poverty in rural South Africa. *Global Health Action*, 3, 5080.
- Czaika M, & Kis-Katos K (2009). Civil conflict and displacement: village-level determinants of forced migration in Aceh. *Journal of Peace Research*, 46 (3), 399–418.
- Daniels, R.C., Partridge, A., Kekana, D., and Musundwa, S. (2013). *Rural Livelihoods in South Africa*. Cape Town: SALDRU, University of Cape Town. SALDRU Working Paper Number 122/ NIDS Discussion Paper 2013/4.
- De Haas H (2008). Migration and development: A theoretical perspective. *International Migration Institute, Oxford University*, Working Paper 9, 1–56.
- De Haas H (2010). The internal dynamics of migration processes: A theoretical inquiry. *Journal of Ethnic and migration Studies*, 36(10), 1587–1617.
- Dustmann C, & Okatenko A (2014). Out-migration, wealth constraints, and the quality of local amenities. *Journal of Development Economics*, 110, 52–63.
- Dziak, J.J., Coffman, D.L., Lanza, S.T., & Li, R. (2012). Sensitivity and specificity of information criteria. Technical Report Series #12–119, Pennsylvania State University, 1–30.
- Englund H (2002). The village in the city, the city in the village: migrants in Lilongwe. *Journal of Southern African Studies*, 28 (1), 137–154.
- Fussell E, Hunter LM, & Gray CL (2014). Measuring the Environmental Dimensions of Human Migration: The Demographer's Toolkit. *Global Environmental Change: Human and Policy Dimensions*, 28, 182–91. [PubMed: 25177108]
- Fotheringham AS and Wong D (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A*, 23(7), 1025–1044.
- Ginsburg C, Bocquier P, Béguy D, Afolabi S, Augusto O, Derra K, Odhiambo F, Otiende M, Soura A, Zabré P, White MJ, & Collinson MA (2016). Human capital on the move: Education as a determinant of internal migration in selected INDEPTH surveillance populations in Africa. *Demographic Research*, 34 (30), 845–884.
- Gray C, & Mueller V (2012). Drought and population mobility in rural Ethiopia. *World Development*, 40(1), 134–145. [PubMed: 22523447]
- Henry S, Schoumaker B, & Beauchemin C (2004). The impact of rainfall on the first outmigration: A multi-level event-history analysis in Burkina Faso. *Population and Environment*, 25 (5), 423–460.
- Juárez F, LeGrand T, Lloyd CB, Singh S, & Hertrich V (2013). Youth migration and transitions to adulthood in developing countries. *The ANNALS of the American Academy of Political and Social Science*, 648 (1), 6–15.
- Kahn K, Tollman SM, Collinson MA, Clark SJ, Twine R, Clark BD, Shabangu M, Gómez-Olivé FX, Mokoena O, & Garenne ML (2007). Research into health, population and social transitions in rural

- South Africa: Data and methods of the Agincourt Health and Demographic Surveillance System. *Scandinavian Journal of Public Health*, 35(69 suppl), 8–20.
- Hargreaves, J. (2000). Village typology: The Agincourt field site descriptive results and preliminary discussion. Unpublished internal document, MRC/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt).
- Hunter LM, Nawrotzki R, Leyk S, Maclaurin GJ, Twine W, Collinson M, & Erasmus B (2014). Rural outmigration, natural capital, and livelihoods in South Africa. *Population, Space and Place*, 20(5), 402–420.
- Kahn K, Collinson MA, Gomez-Olive FX, Mokoena O, Twine R, Mee P, Afolabi SA, Clark BD, Kabudula CW, Khosa A, Khoza S, Shabangu MG, Silaule B, Tibane JB, Wagner RG, Garenne ML, Clark SJ, and Tollman SM (2012). Profile: Agincourt Health and socio-Demographic Surveillance System. *International Journal of Epidemiology*, 41, 988–1001. DOI: 10.1093/ije/dys115 [PubMed: 22933647]
- Kandel W, & Massey DS (2002). The culture of Mexican migration: A theoretical and empirical analysis. *Social Forces*, 80(3), 981–1004.
- Lindstrom DP (1996). Economic opportunity in Mexico and return migration from the United States. *Demography*, 33(3), 357–374. [PubMed: 8875068]
- Lindstrom DP, & Lauster N (2001). Local Economic Opportunity and the Competing Risks of Internal and U.S. Migration in Zacatecas, Mexico. *International Migration Review*, 35 (4), 12321256.
- Leyk S, Maclaurin GJ, Hunter LM, Noawrotzki R, Twine W, Collinson M, and Erasmus B (2012). Spatially and temporally varying associations between temporary outmigration and natural resource availability in resource-dependent rural communities in South Africa: A modeling framework. *Applied Geography*, 34, 559–68. [PubMed: 23008525]
- Madhavan S, Townsend NW, & Garey AI (2008). ‘Absent Breadwinners’: Father–child connections and paternal support in rural South Africa. *Journal of Southern African Studies*, 34(3), 647–63. [PubMed: 20396616]
- Madhavan S, Schatz E, Clark S, & Collinson M (2012). Child mobility, maternal status, and household composition in rural South Africa. *Demography*, 49 (2), 699–718. [PubMed: 22271363]
- Madhavan S, Mee P, & Collinson M (2014). Kinship in Practice: Spatial Distribution of Children’s Kin Networks. *Journal of Southern African Studies*, 40 (2), 401–418.
- Martin PL, & Taylor JE (1996). The anatomy of a migration hump In Taylor JE (Ed.), *Development strategy, employment, and migration: Insights from models* (pp. 43–62). Paris: Organization for Economic Cooperation and Development (OECD).
- Massey DS, Goldring L, & Durand J (1994). Continuities in transnational migration: An analysis of nineteen Mexican communities. *American Journal of Sociology*, 99 (6), 1492–1533.
- Massey DS, & Espinosa KE (1997). What’s driving Mexico–U.S. migration? A theoretical, empirical, and policy analysis. *American Journal of Sociology*, 102 (4), 939–999.
- Masson MPR (2001). *Migration, Human Capital, and Poverty in a Dual-Economy of a Developing Country* (No. 1–128). International Monetary Fund.
- Mbele, L. (2016, May 11). Why M-Pesa failed in South Africa. *BBC Africa Business Report*. <http://www.bbc.com/news/world-africa-36260348>
- McKenzie D, & Rapoport H (2007). Network effects and the dynamics of migration and inequality: theory and evidence from Mexico. *Journal of development Economics*, 84(1), 1–24.
- McLeod D (2016, June 05). Why M-Pesa ‘will never work in South Africa.’ *Sunday Times*. <https://www.timeslive.co.za/sunday-times/business/2016-06-05-why-m-pesa-will-neverwork-in-south-africa/>
- Montello DR (2001). Scale in geography In J Smelser N and Baltes PB (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 13501–13504). Oxford: Pergamon.
- Nawrotzki RJ, Riosmena F, & Hunter LM (2013). Do rainfall deficits predict U.S.-bound migration from rural Mexico? Evidence from the Mexican census. *Population research and policy review*, 32(1), 129–158. [PubMed: 23913999]
- Poston DL, & Zhang L (2008). Ecological analyses of permanent and temporary migration streams in China in the 1990s. *Population Research and Policy Review*, 27(6), 689–712.

- Reed HE (2013). Moving across boundaries: migration in South Africa, 1950– 2000. *Demography*, 50(1), 71–95. [PubMed: 22956415]
- Raftery AE (1996). Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika*, 83(2), 251–266.
- Rees P, Bell M, Kupiszewski M, Kupiszewska D, Ueffing P, Bernard A, Charles- Edwards E & Stillwell J (2016). The impact of internal migration on population redistribution: An international comparison. *Population, Space and Place*. DOI: 10.1002/psp.2036
- Stark O, & Taylor JE (1991). Migration incentives, migration types: The role of relative deprivation. *The Economic Journal*, 101 (48), 1163–1178.
- White MJ, & Lindstrom DP (2005). Internal migration In Poston DL and Micklin M (Eds.), *Handbook of Population* (pp. 311–346). New York: Kluwer Academic Publishers.
- Zelinsky W (1971). The hypothesis of the mobility transition. *Geographical Review* 61 (2), 219–249.

Figure 1:
Unconditional Kaplan-Meier Estimates—Age at First Migration (any type) Agincourt 2003–2011

Figure 2:
Unconditional Kaplan-Meier Estimates—Age at First Temporary Migration Agincourt,
2003–2011

Figure 3:
Unconditional Kaplan- Age at First Permanent Migration Agincourt 2003–2011

Table 1:

Descriptive Statistics (Person-Years) 2003–2011 (Percentages and means, with standard deviations in parentheses where appropriate)

<i>Dependent Variables</i>	<i>Males</i>	<i>Females</i>
Competing Risks of Migration		
Did Not Move	85.0%	83.2%
Temporary Migrant	11.6%	10.7%
Permanent Migrant	3.4%	6.1%
<i>Village-Level Independent Variables</i>		
Average Household Asset Index (11 Item Score) in Village	5.9 (1.3)	5.9 (1.2)
Average Number of Years of Education in Village (Age 24+)	7.0 (1.1)	7.0 (1.1)
Average Percent Out-Migration in Village	28.0 (3.6)	28.0 (3.7)
<i>Individual-Level Controls</i>		
Asset Index (11 Item Score) in Household	6.7 (2.5)	6.7 (2.5)
Cohort		
1 (2003 start)	19.1%	18.6%
2 (2004 start)	18.7%	17.9%
3 (2005 start)	16.1%	14.8%
4 (2006 start)	13.2%	13.2%
5 (2007 start)	10.9%	11.6%
6 (2008 start)	8.1%	9.1%
7 (2009 start)	7.3%	7.7%
8 (2010 start)	4.5%	4.7%
9 (2011 start)	2.3%	2.5%
Matriculated School	7.0%	7.8%
Age	18.0 (2.2)	17.8 (2.1)
Person is Child in Household	68.1%	67.5%
Person is Head of Household	1.1%	0.7%
Person is Other in Household	30.8%	31.8%
Person is Mozambique Refugee	26.0%	24.4%
<i>N (Person-Years)</i>	35840	33406
<i>(N [Distinct Individuals])</i>	(7381)	(7373)

Table 2:
Event History Models, 2003–2011, 15-Year Old Cohorts, Males and Females, First Migration

	Males			Females		
	Individual/HH Only	Individual/HH+ Village Time-Varying (Assets)	Individual/HH+ Village Time-Varying (Education)	Individual/HH Only	Individual/HH+ Village Time-Varying (Assets)	Individual/HH+ Village Time-Varying (Education)
	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.12 ***	0.11 ***	0.12 ***	0.07 ***	0.06 **	0.07 ***
Matriculated (No) Yes	1.40 ***	1.42 ***	1.42 ***	0.89 ***	0.90 ***	0.90 ***
Cohort—Linear	−0.10 ***	−0.10 ***	−0.09 ***	−0.10 ***	−0.11 ***	−0.10 ***
HH Role (Child) Head	0.18	0.18	0.18	−0.07	−0.08	−0.08
Other	0.56 ***	0.56 ***	0.56 ***	0.55 ***	0.55 ***	0.55 ***
Mz. Refugee (No) Yes	0.01	0.05	0.03	0.11 +	0.15 +	0.16 *
HH Assets	0.01	0.00	0.01	−0.00	−0.01	−0.01
Village %Mig.		0.02 ***	0.03 ***		0.01	0.01 +
Village Assets		0.01			0.03	
Village Educ.			−0.02			0.03
Constant	−4.66 ***	−5.36 ***	−5.37 ***	−3.45 ***	−3.76 ***	−3.87 ***
Observations	29919	29919	29919	27419	27419	27419
Pseudo R^2	0.062	0.063	0.063	0.035	0.036	0.036
<i>AIC</i>	15495.1	15484.3	15484.0	17228.7	17225.7	17225.7
<i>BIC</i>	15561.5	15567.3	15567.0	17294.4	17307.9	17307.9

+ p<.10,

* p<.05,

** p<.01,

*** p<.001. Models include cluster-robust standard errors for villages.

Table 3:

Event History Models Repeated Events, 2003–2011, 15-Year Old Cohorts, Males and Females Competing Risks of Temporary versus Permanent Migration

	Males				Females			
	(1)		(2)		(3)		(4)	
	Temp. Vs. No	Perm. Vs. No	Individual/HH+ Village Time-Varying (Assets)	Temp. Vs. No	Perm. Vs. No	Individual/HH+ Village Time-Varying (Education)	Temp. Vs. No	Perm. Vs. No
Age	0.21 ***	-0.14 ***	0.21 ***	-0.13 ***	0.21 ***	-0.14 ***	0.12 ***	-0.03 *
Matriculated (No)	1.50 ***	0.26 +	1.52 ***	0.27 +	1.52 ***	0.27 +	1.28 ***	-0.08
Cohort—Linear	-0.12 ***	-0.12 ***	-0.13 ***	-0.11 ***	-0.12 ***	-0.12 ***	-0.09 **	-0.13 ***
HH Role (Child) Head	-0.04	0.01	-0.04	0.01	-0.04	0.01	-0.20	-0.01
Other	0.08	1.17 ***	0.08	1.17 ***	0.08	1.17 ***	0.11	0.81 ***
Mz. Refugee (No)	0.04	0.04	0.10	0.02	0.10	0.00	-0.22 *	0.34 ***
# Prior Mig.	1.25 ***	0.49 ***	1.24 ***	0.49 ***	1.24 ***	0.49 ***	1.18 ***	0.37 ***
HH Assets	0.07 ***	-0.07 ***	0.07 ***	-0.07 ***	0.07 ***	-0.07 ***	0.11 ***	-0.07 ***
Village %Mig.			0.05 ***	0.01	0.03 ***	0.01	0.02 +	0.01
Village Assets			0.02	-0.05			0.09 +	0.01
Village Educ.					0.01	-0.05		0.11 +
Constant	-6.91 ***	-0.50 *	-7.78 ***	-0.80 *	-7.85 ***	-0.65	-5.58 ***	-1.78 ***
Observations	34259	34259	34259	34259	34259	34259	31458	31458
Pseudo R²	0.296	0.297	0.297	0.297	0.297	0.297	0.243	0.244
AIC	24281.2	24262.7	24263.0	24263.0	24263.0	24263.0	26803.3	26787.0
BIC	24433.1	24431.5	24431.8	24431.8	24431.8	24431.8	26953.7	26954.1

p<.001. Models include cluster-robust standard errors for villages.

p<.01,

**

p<.05,

*

p<.10,

+

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Appendix Table 1:

Variation in Village-Level Predictors Over Time

Year	Asset Index (11 Item Score) in Village		Years of Education in Village (Age 24+)	
	Mean	Coefficient of Variation	Mean	Coefficient of Variation
	(Std. Dev.)		(Std. Dev.)	
2003	4.1 (1.4)	34.3	5.4 (1.4)	27.1
2005	4.2 (1.4)	33.9	6.1 (1.4)	22.8
2007	5.0 (1.3)	25.8	6.3 (1.3)	20.9
2009	5.9 (1.4)	24.0	6.8 (1.3)	19.9
2011	6.9 (0.8)	11.8	7.4 (1.2)	16.9
N (Villages)		21	21	

Appendix Table 2:
Event History Models, 2003–2011, 15-Year Old Cohorts, Males and Females, First Migration

	Males			Females		
	Individual/HH (1)	Individual/HH ⁺ Village Time-Varying (Assets) (2)	Individual/HH ⁺ Village Time-Varying (Education) (3)	Individual/HH (4)	Individual/HH ⁺ Village Time-Varying (Assets) (5)	Individual/HH ⁺ Village Time-Varying (Education) (6)
Age	0.12 ***	0.03	0.05	0.07 ***	-0.02	0.01
Matriculated (No) Yes	1.39 ***	1.41 ***	1.42 ***	0.90 ***	0.90 ***	0.91 ***
Cohort—Linear	-0.10 ***	-0.19 ***	-0.17 ***	-0.10 ***	-0.20 ***	-0.16 ***
HH Role (Child) Head	0.13	0.13	0.13	-0.02	-0.03	-0.02
Other	0.56 ***	0.56 ***	0.56 ***	0.56 ***	0.56 ***	0.56 ***
Mz. Refugee (No) Yes	-0.00	-0.01	0.00	0.15 *	0.15 *	0.15 *
HH Assets	0.01	0.00	0.01	-0.00	-0.01	-0.00
Village %Mig.		0.03 *	0.02		0.01	-0.00
Village Assets		0.25 **			0.28 ***	
Village Educ.			0.39 **			0.34 **
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.61 ***	-5.38 ***	-6.75 ***	-3.52 ***	-3.58 ***	-4.74 ***
Observations	29919	29919	29919	27419	27419	27419
Pseudo <i>R</i> ²	0.065	0.066	0.066	0.039	0.040	0.039
<i>AIC</i>	15446.5	15429.7	15433.5	17165.6	17143.1	17157.7
<i>BIC</i>	15512.9	15504.5	15508.3	17223.1	17217.0	17231.7

⁺ p<.10,

* p<.05,

** p<.01,

p<.001. Models include cluster-robust standard errors for villages.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Appendix Table 3:

Event History Models Repeated Events, 2003–2011, 15-Year Old Cohorts, Males and Females Competing Risks of Temporary versus Permanent Migration

	Males				Females			
	Individual/HH Only		Individual/HH+ Village Time-Varying (Assets)		Individual/HH+ Village Time-Varying (Education)		Individual/HH+ Village Time-Varying (Assets)	
	Temp. Vs. No	Perm. Vs. No	Temp. Vs. No	Perm. Vs. No	Temp. Vs. No	Perm. Vs. No	Temp. Vs. No	Perm. Vs. No
	(1)	(2)	(3)	(4)	(5)	(6)		
Age	0.22 ***	0.19 ***	0.19 ***	0.13 ***	0.13 *	0.11 **	0.13 *	-0.14 ***
Matriculated (No) Yes	1.49 ***	0.26 ⁺	1.48 ***	0.28 *	0.31 *	0.27 ***	1.27 ***	-0.05
Cohort—Linear	-0.12 ***	-0.13 ***	-0.14 ***	-0.08 ***	-0.09 ⁺	-0.11 **	-0.07 ***	-
HH Role (Child) Head	-0.03	-0.06	-0.03	-0.20	0.06	0.13	0.05	0.25 ***
Other	0.09	1.18 ***	0.09	0.10	0.84 ***	0.10	0.10	0.84 ***
Mz. Refugee (No) Yes	0.09	-0.03	0.09 ⁺	-0.04	-0.03	-0.13 ***	0.34 ***	0.36 ***
# Prior Mig.	1.24 ***	0.49 ***	1.24 ***	0.49 ***	1.16 ***	1.17 ***	0.37 ***	0.37 ***
HH Assets	0.07 ***	-0.06 ***	0.07 ***	-0.06 ***	0.11 ***	0.10 ***	0.08 ***	-
Village %Mig.		-0.01	-0.01	0.04 *	0.06	-0.05 **	0.02 *	0.01
Village Assets		0.06	0.06	0.26 **	0.02	0.38 ***		
Village Educ.			0.15 ⁺	0.59 **		0.15		0.61 ***
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-6.70 ***	-0.67 **	-6.30 ***	-	-5.49 ***	-4.00 ***	-2.34 ***	-4.61 ***
		1.71 **	4.00 ***	2.04 ***				4.58 ***

	Males				Females			
	Individual/HH Only		Individual/HH+ Village Time-Varying (Assets)		Individual/HH+ Village Time-Varying (Education)		Individual/HH+ Village Time-Varying (Assets)	
	Temp. Vs. No	Perm. Vs. No	Temp. Vs. No	Perm. Vs. No	Temp. Vs. No	Perm. Vs. No	Temp. Vs. No	Perm. Vs. No
	(1)	(2)	(3)	(4)	(5)	(6)		
Observations	34259	34259	34259	31458	31458	31458		
Pseudo R^2	0.300	0.301	0.301	0.248	0.249	0.249		
<i>AIC</i>	24130.6	24121.9	24116.5	26641.7	26596.5	26605.1		
<i>BIC</i>	24265.7	24290.7	24285.3	26783.8	26763.7	26772.3		

⁺ p<.10,

* p<.05,

** p<.01,

*** p<.001. Models include cluster-robust standard errors for villages.