INDIVIDUAL MIGRATION AND HOUSEHOLD INCOMES

Julia Garlick
Murray Leibbrandt
James Levinsohn

Individual Migration and Household Incomes
Julia Garlick, Murray Leibbrandt, and James Levinsohn
NBER Working Paper No. 22326
June 2016
JEL No. O1,O12,O15

ABSTRACT

We estimate the returns to internal migration in South Africa. These appear to be the first nationally representative estimates of the return to migration for any African country-- a somewhat surprising claim for a literature that's over 60 years old. We develop a framework to analyze individual migration in the context of income pooling within endogenously formed households. We apply this framework to estimate the return to migration from the perspective of the migrant (as is typically done) as well as from the perspectives of the sending and receiving households.

Julia Garlick
Department of Economics
Yale University
julia.garlick@yale.edu

Murray Leibbrandt
Saldru
School of Economics
University of Cape Town
Private Bag X03, Rondebosch, 7701
South Africa
Murray.Leibbrandt@uct.ac.za

James Levinsohn
Yale School of Management
PO Box 208200
New Haven, CT 06520
and NBER
James.Levinsohn@yale.edu

Individual Migration and Household Incomes

Julia Garlick
Yale University

Murray Leibbrandt
University of Cape Town

James Levinsohn
Yale University

June 1, 2016

Abstract

We estimate the returns to internal migration in South Africa. These appear to be the first nationally representative estimates of the return to migration for any African country– a somewhat surprising claim for a literature that's over 60 years old. We develop a framework to analyze individual migration in the context of income pooling within endogenously formed households. We apply this framework to estimate the return to migration from the perspective of the migrant (as is typically done) as well as from the perspectives of the sending and receiving households.

1 Introduction

Is migration a way of getting ahead? This sounds like a simple question. And, for the individual living alone who leaves one location to set up residence in a new location, again alone, it is. In many African countries, though, the very question is ill-posed, for migration frequently involves an individual leaving one household in which income is pooled and joining another in which income is also pooled. In this context asking "Is migration a way of getting ahead?" begs the follow-on "For whom?"

We investigate this question using recently available South African panel data. In so doing, we provide the first nationally representative estimates of the impact of migration on incomes in South Africa and, perhaps more surprisingly, the first for any African country. This lack of national-level evidence is less surprising when one notes that such a study, by design, requires nationally representative panel data and in Africa there have been, until recently, none. With the advent of South Africa's National Income Dynamics Survey (NIDS), the data are now available.

Some of the contributions of our analysis are specific to South Africa– a country with a long and infamous history of migration. For example, one of the most surprising things about internal migration in South Africa is its sheer prevalence. We find that about half of all South Africans live in a household impacted by migration over only a four year span. When we restrict our analysis to Black[1] South Africans, who comprise about 80 percent of the population and on whom we focus our analysis, the figure is even higher. In South Africa, migration matters. The exact magnitude of the causal impact of migration on incomes is of course also South Africa-specific.

Other contributions of the paper extend beyond South Africa and are far more general. We highlight two.

First, we provide a framework with which to analyze the economic impact of migration when *individuals* migrate and *households* pool income. When we ask whether migration is a way of getting ahead, we examine this from the perspective of the migrant, from the perspective of the sending household (the household the migrant left) and from the perspective of the receiving household (the household which the migrant joins). In the presence of income pooling, examining only the first, which is the norm in the literature, provides part of the story and viewed alone this may give a distorted view of the economic impact of migration. In the presence of income pooling, it's possible for the migrant to be better off but for the sending and receiving households to each be worse off, or for the migrant to be better off without any change in his or her individual income.

Second, even when we restrict the focus to the actual migrants (as opposed to members of sending and receiving households), we highlight the importance of analyzing the returns to migration for all migrants. The economics literature has typically focused on migration by labor market participants. That's important, of course, but in our national sample, less than half of all migrants are working age. The returns to migration, in the presence of income sharing within the household, is more nuanced than just looking at a worker's income.

Our focus is squarely on measuring the returns to migration for all those impacted – the migrants,

---

[1]In South African parlance, this is the population group referred to as "African." Hereafter we use the term "Black."

the household left behind, as well as the household joined– and to do so at a nationally representative level.

The literature on measuring the returns to migration in Africa is surprisingly sparse given the prominent role assigned to migration in seminal models of economic development (e.g. Lewis (1954), Harris and Todaro (1970), and more recently Young (2013).) The empirical side of migration research is better represented in South Asia. Recent studies, each focused on different aspects of temporary migration, include Morten (2013) (India) and Bryan, Chowdhury and Mobarak (2014) (Bangladesh.) While we know of no nationally representative studies examining the impact of internal migration on incomes in any African country, Beegle, Weerdt and Dercon (2011) examine the issue with a baseline sample of 912 households from the Kagera region of Tanzania.

In our data, migration leads to the formation of new households since households moving *en masse* are not the norm. Empirical research on endogenous household formation dates back at least to Rosenzweig and Stark (1989), which examined the issue in the context of marriage in India– a context massively different than the South African. In the South African context, research has focused on correlates of migration and endogenous household formation. For example, Hamoudi and Thomas (2014) examine the role the State Old Age Pension plays in household formation while Ardington, Case and Hosegood (2009) analyze the role that cash transfers play in enabling migration by essentially staking the migrant. Neither of these papers estimates the return to migration for any party involved.[2] Finally, we are unaware of any study, anywhere, that examines the returns to migration from the perspectives of the migrant, the sending household, *and* the receiving household.

Every careful analysis of the returns to migration must address the issue of selection. There are multiple approaches to addressing the selection issue in the context of migration and they are nicely reviewed in McKenzie, Gibson and Stillman (2010). The authors employ an ingenious strategy and examine how various econometric approaches to estimating the returns to migration compare to

---

[2]Prior empirical studies of migration in South Africa have, by necessity, relied on repeated cross-sectional data. See for example Posel and Casale (2006), Posel, Fairburn and Lund (2006), and Budlender and Lund (2011). These papers also highlight various aspects of the fluid nature of South African households.

the results from a randomized controlled trial (RCT). They use data from a New Zealand lottery that allowed a quota of Tongans to migrate hence mimicking a randomized control trial (RCT). The authors run a horse race to see which estimator most closely replicates the RCT results.[3] The paper is convincing but leaves open the question of whether the results are in some way special to 250 migrants from Tonga to New Zealand who have already secured a job - a question that in its generic form pervades the RCT approach.[4]

Our econometric strategy for tackling selection is most similar to that of Ham, Li and Reagan (2011). The authors estimate the returns to migration in the U.S.. They demonstrate that matching methods, appropriately specified and applied to nationally representative panel data, can provide well-identified estimates of the returns to migration for the migrants. This is exactly our approach.

In the next section, we introduce the data upon which we rely. Descriptive statistics that provide context and background are given in Section 3. Section 4 presents a framework for thinking about individual migration and household incomes. Econometric strategies are discussed there. The causal impacts of migration are presented in Section 5, and Section 6 concludes.

2   Data

We use data collected in the first three waves of South Africa's National Income Dynamics Study (NIDS). These waves were collected in 2008, 2010, and 2012 and they comprise a panel in which individuals were tracked across waves and re-interviewed. The three waves of our data span a broad-based macroeconomic contraction (2008-2010) and then a modest recovery (2010-2012). The data are publicly available and can be downloaded from the NIDS website.[5] The data, with supplied

---

[3]They show that the best non-experimental estimator is an instrumental variable (IV) approach, followed by matching methods.

[4]Another important way in which this paper's results are special is that the IV estimator (which performs quite well) uses instruments that don't have clear analogues in nationally representative non-experimental data.

[5]See: www.nids.uct.ac.za. The exact data used are *National Income Dynamics Study 2012, Wave 3 [dataset]* (2014), *National Income Dynamics Study 2010-11, Wave 2 [dataset]* (2014), and *National Income Dynamics Study 2008, Wave 1 [dataset]* (2014).

weights, comprise a nationally representative sample. All household members (or their proxy) were surveyed and household residency is defined by whether the individual slept in the house for at least four nights during the week. Household members are coded as having moved if they changed residences between waves and this is verified using (non-public access) GPS data. A detailed description of the data collection protocols, sampling methodology, attrition, and other technical matters is found in DeVilliers, Woolard, Daniels and Leibbrandt (2013).

The panel is constructed as follows. The residents of each household interviewed in Wave 1 are designated continuing sample members (CSMs). These original CSMs are augmented only by children born to or adopted by female CSMs in subsequent waves. New individuals enter each subsequent wave as new co-residents of CSMs. These new individuals remain in the survey only for so long as they live with a CSM. Hence, there are three ways a new individual enters the sample: as a new child of an existing female CSM; as a child or adult who joins a household containing a CSM (new temporary sample member– TSM); and as a child or adult whose household acquires a CSM (a new TSM). There are three ways an individual exits the sample. A CSM exits if they die or migrate internationally. TSMs also exit the sample if they either move out of a CSM household or if the CSM moves out of their household. In each of the three waves of data collection, CSMs who were not found or were found but refused an interview were searched for again in the next round and, if found, were asked for an interview. In this way, Wave 3 recovered a number of CSMs who were not included in Wave 2.

Table 1 gives the sample size for each wave by race.[6] In Wave 1, there were 28,278 individuals in the sample, 22,255 of whom were Black. These are the data to which sampling weights are later applied. Given the sampling design, wave 2 could have fewer respondents only if CSMs died, moved out of the country, or otherwise attritted from the survey. In fact, Wave 2 had more participants, a total of 34,978, and this increase results from new TSMs who were co-resident with a CSM at the time of the interview. Put another way, but-for attrition, the 34,978 figure would have been

---

[6]The South African population is divided into several officially recognized racial groups, following the categories formalized by the Apartheid government. In official parlance, "Coloured" people are members of a long-standing and culturally distinct mixed-race population.

even larger. The yet larger sample size of 38,191 in Wave 3 reflects the same phenomena. The last column of Table 1 gives the total number of unique individuals surveyed across the entire three waves - 41,306.

Attrition across the waves averaged about 16 percent. The modal reason for the attrition was not that the individual was not tracked but rather that the individual or the household refused to participate when re-contacted for the next wave. Attrition rates varied markedly by racial groups with Whites having the highest (about 50 percent) and Blacks the lowest (about 13 percent.)

While all four population groups are included in NIDS, our analysis of migration is restricted to Blacks. We do not include Whites for several reasons. The household dynamics underlying the migration decision are massively different for this group. The multi-generation households that are such an integral part of the demographic landscape of South Africa are not very present. Much of the observed migration is due to "working age" adults going to or from university, and finally the data are subject to high attrition rates. We also do not include the Asian/Indian population group. In addition to many of the issues that pertain to Whites, there is also a small-numbers problem such that drawing any inferences about nation-wide patterns for this group would be problematic.[7] Finally, we also do not include the Coloured population group, although this is more of a close call. The Coloured population is somewhat geographically concentrated and is less likely to live in multi-generational households or in "skip generation" households, in which grandparents can care for grandchildren. Each of these attributes make the migration decision different than it is for the Black population.[8] Finally, the Black population comprises about 81 percent of the (unweighted) sample so we retain the vast majority of our data with our focus on Blacks.

Table 2 provides an (unweighted) count of Blacks in NIDS by migration status. To fix ideas, during the four years spanned by the three waves of NIDS, 5,860 of the 21,590 individuals in the sample reported migrating at least once between 2008 and 2012. Hence, over a quarter of the Blacks in the sample changed residence over only a four year period. The bottom panel of Table 2 provides

---

[7]DeVilliers et al. (2013) warns that there are too few Asian/Indians to support statistical analysis of this group.
[8]When observed, migration by Coloured respondents involves shorter distances and fewer city changes, so it seems to be a different sort of choice than migration by Black respondents.

the age breakdown of migrants. Although not reported in the table, women are slightly more likely to have migrated than are men. A key message from the table is that the pattern of men moving to the mines for work while the women and children stayed home and received remittances is not visible. That historic pattern is no more. Indeed, almost half of the migrants are minors (not miners).

The household dynamics of migration are complicated. Less than a third of the migration events are a move in which all household members moved.[9] Combining that fact with the large number of children moving indicates that most moves involve some but not all of the sending household, often including children, moving into another existing household. The prevalence of children moving also highlights the importance of looking at all migrants since focusing only on labor market participants will miss a substantial fraction of movers.

3 Descriptive Analyses

We begin with descriptive analyses of internal migration. These set the stage for the analysis of the causal impact of migration that follows. We focus our descriptive analysis around three questions. First, how many South Africans live in households impacted by migration? Second, when individuals migrate, do they move across communities or within and, when moving across communities how much migration is of the traditional rural to urban sort and how much is within urban, within rural or even urban to rural? Third, what happens to the incomes of movers versus non-movers?

3.1 The prevalence of households impacted by moving

Table 2 provided a simple count of migrants. This count, though, vastly understates the number of South Africans impacted by migration. Because migration typically involves moves other than

---

[9]About a third of individuals moving between 2008 and 2012 were part of a household that moved *en masse*. If we simply count the number of migration events, instead of the number of migrants, then the fraction of migration events that involve an *en masse* move is much less than one third.

an *en masse* move of the entire household, many household members other than the mover are impacted. Consider an unemployed 27 year old woman living in a 5 person household. If she moves and joins a 2 person household, she leaves 4 members of the sending household. Assuming someone in that sending household has an income, the per capita income of the 4 sending household members rises with the migration *ceteris paribus.* And again assuming the migrant remains unemployed, she drags down the household per capita income of the two members of the receiving household. In this example, there are 7 people impacted by the move. Table 3 counts the number of individuals in our sample who are and who are not impacted by a move, and when someone is affected by a move, we report how many are in sending households and how many are in receiving households. We perform this exercise for moves between 2008 and 2010, for moves between 2010 and 2012, and finally for any move during the span of the sample.

We begin with a discussion of the top panel of Table 3. From 2008 to 2010, 51.98 percent of Blacks in our sample were affected by a move into or out of their household. For the 2010-2012 period, the figure jumps to 56.91 percent, and over the four years spanning our data, over 80% of all Blacks lived in a household that either sent or received a migrant. This strikes us as a stunningly high number for a period as short as four years. The fact that migration was more prevalent during the relative macroeconomic upswing than during 2008-2010 is suggestive of moving to opportunity as opposed to a push out of the nest, but we reserve this for more careful analysis below.

The percentages reported for sending and receiving households are given in rows three and four. Of households that were affected by a move in the 2008-2010 period, 41.5 percent had someone migrate out of the household while 85 percent had someone move into the household. This highlights an important and somewhat surprising phenomenon. Many households that send a migrant also receive a migrant during the same period. The pattern of more people being affected by receiving a migrant than by sending one holds throughout the sample period.

Because of how the sample is constructed with the inclusion of Temporary Sample Members (TSMs), we were concerned that the figures in the top panel of Table 4 may not be representative. To illustrate this concern, consider who joins the sample in Wave 3. These new sample members, by

design, are either members of a household that was joined or formed by a migrant or the TSM is him/herself a migrant into a household of CSMs. In order to better understand the extent of this possible bias, we restricted our sample to the 2008 Wave 1 sample. Using only these individuals and following them through time, we repeat the analysis of the top panel and report results in the bottom panel of Table 3. We find that the large proportion of the sample impacted by migration is not an artifact of the sample design. Using the entire sample over the four year span, 80.53 percent were affected and this figure falls only modestly to 73.99 percent when only Wave 1 sample members were included. The vast majority of Black South Africans lived in a household affected by migration.

3.2   Type of move

Table 4 reports moves as categorized by the urban/rural distinction so often used in the literature. A half century of models of migration have focused on the role played by rural to urban migration in economic development, and that is South Africa's history too. Migration in contemporary South Africa is more nuanced (as one would expect vis-a-vis a model) and, perhaps unexpectedly, just plain different.

There are three key messages in Table 4. First, most moves are on the diagonal. That is, most moves are within rural or within urban. Only about one fifth of all moves are from a rural origin to an urban destination. Second, moves are more common among rural households and two thirds of these moves are to a rural destination. Third, the urban sector is the modal destination, but just barely (2331 moves versus 2014 moves.)[10]

---

[10]In results not reported in the table, we find that 60 percent of all moves are within District Council. There are 53 District Councils in South Africa so this is a more granular geographic unit than Province. In our results below, we make more detailed use of information on the distance moved. The point here is simply that most moves result in relocation not that distant from what had been home.

3.3 Income changes and migration

Table 5 reports changes in log per-capita household income by whether or not the respondent moved and by gender of the respondent. The sample includes all Blacks.[11] Focusing first on the 2008 to 2010 period, male non-movers saw household per capita income rise by 4.3 percent while the figure for female non-movers is 4.1 percent. (All data are in real terms.) Over the same time span, male movers had an increase in per-capita household income of 25.6 percent and females 11.5 percent. For the 2010-2012 period, a period during which the economy was picking up, non-movers had an increase in per-capita household income of about 18-19 percent. Female movers saw incomes rise by 51 percent while male movers experienced a 43 percent increase. Over the entire sample period, non-movers saw an increase in per-capita household income of about 21.7 percent while the figure for movers was about 43 percent (female) and 44 percent (male.) These strike us as large differences in real income over a relatively short time span.

4 Framework

4.1 Framing the Question

When measuring the impact of migration on incomes, the first issue that must be addressed is which measure of income. If incomes are pooled and shared equally within the household, then household per capita income is the appropriate measure. If, on the other hand, incomes are not pooled, then individual income is the appropriate measure. Neither of these extremes are going to be exactly right for all households. While incomes do tend to be pooled within the household in South Africa, it is less clear that every household member receives the same share of household income.

Were individual income the appropriate measure, the entire issue of the impacts of migration on sending and receiving households vanishes. This case also begs the question of how the roughly 40% of the sample under the age of 16 as well as all unemployed adults survive.

---

[11]Figures in this table reflect sample weights.

The ideal outcome measure would be household income appropriately weighted by the individual household member's (or his/her proxy's) bargaining power. Alas, this ideal is not observed in our data. When measuring returns to migration for the migrant (as opposed to the sending or receiving household), we report results using both the household per capita income and individual income as outcome measures, although we are confident that the former is the more appropriate measure. By providing results with both perfect pooling (household per capita income) and no pooling (individual income), we estimate the impact of migration under drastically different assumptions.

When measuring household per capita income, remittances deserve careful attention since they can potentially appear twice in the data. If both the remittance-sending and remittance-receiving households are surveyed, remittance income accrues first to the sender (through earnings) and then to the receiver (through remittances.) Because many remittance networks arise as a result of migration, the returns to migration may be distorted if remittances are not handled with care. We assign remittances to the recipient household, not the sending household.

In the presence of income pooling within the household, the economic impact of migration is nuanced. In order to be clear just who comprises the household at a given point in time, it is helpful to establish some notation.

Consider a given household. Denote the set of individuals who migrate between period t and $t+1$ by M. The sending household members are denoted T (for "trailing".) This is the set of individuals who co-reside with M in period t but not in $t+1$. The members of the receiving household are denoted R. This is the set of individuals who co-reside with M in period $t+1$ but not in period t. Table 6 illustrates some examples of household composition with migration and helps fix ideas.

Return now to the question posed at the outset, "Is migration a way of getting ahead?"

The first line in Table 6, Example 1, gives the example of a household that moves *en masse*. That is, the members of the household in period t are the same as the members in $t+1$. Because household composition does not change, it's straightforward to determine if the household members are better

off with the move. Individual incomes within households are observed in NIDS in period t and $t+1$, so one simply computes per capita household income before and after the move to measure whether migration left the household with a higher or lower per capita household income. This situation is not typical in the data – with the exception of one person households that migrate, the migrating household usually loses or gains members.

Example 2 looks at migration from the perspective of the migrants when only some members of the period t household migrate. In this example, the only household members in common across the two periods are the migrants (or migrant, since it may be an individual rather than a group of migrants). If the question being asked is "Is migration a way of getting ahead for the migrants?," example 2 is the appropriate comparison. Note that in the presence of income pooling, it's entirely possible for the migrant's individual income to fall with migration but for her per capita household income to rise (and vice versa.) Because $M_t + T_t$, the migrants' original household, and $M_{t+1} + R_{t+1}$, the migrants' new household, are each observed in NIDS, it is straightforward to measure whether on average migrants' per capita household income increases or decreases with migration.

Next consider Example 3 in Table 6. A comparison of the per capita household incomes of $M_t + T_t$ to that of T_{t+1} is answering yet a different and still well-defined question. Example 3 asks "What happens to the sending household in the presence of migration?" Still using per capita household income as the appropriate measure of income, this framing analyzes whether migration is good for those household members who live in the household that the migrants left. If for example, it's the unemployed, school-going, or non-participants in the labor market who leave the nest, we measure how this migration has benefited the sending household members. Or if it's the employed and more productive members who leave, how badly are sending household members harmed by migration? In our data, $M_t + T_t$ and T_{t+1} are each observed. Hence it's straightforward to measure whether on average sending household members experienced higher or lower per capita household incomes from migration.

Finally, consider Example 4 in Table 6. This comparison asks whether migration benefits the receiving household. This too is a well-defined variant of the core question "Is migration a way of getting

ahead?" In this case, our focus is on the receiving household, so R_t consists of households in year t that received a migrant, typically a TSM, in year $t+1$. Hence in Example 4, we observe R_t, R_{t+1}, and M_{t+1} but only by happenstance would we observe M_t. By focusing our lens on the receiving household, we are able to track household incomes both before and after receiving the migrant(s) and hence address the impact of receiving a migrant on household per capita income.

Each row in Table 6, then, frames the question "Is migration a way of getting ahead?" from a different perspective. The first row asks the question from the perspective of the household that moves *en masse.* The second asks from the perspective of migrants who left one household to join another. The third asks the question from the perspective of the household members left behind while the last asks from the perspective of the household that received the migrants. We answer each.

4.2 Econometric Strategies

To address whether migration is a way of getting ahead, we need to measure the causal impact of migration. To do so, we need a way to infer how a migrant (or a migration-affected household) would have fared absent the migration event. This, of course, is not observed so any measurement of the returns to migration must be obtained from comparing migrants to non-migrants. This immediately raises the problems of identifying comparable non-migrants, and controlling for the role of selection. Selection appears in two forms for migration - selection into migration, and selection of destination. We do not make any attempt to address the latter, so our estimates of the returns to migration include destination effects.

Given that selection is inherent in the migration decision, the cleanest way of addressing it is to run an RCT. This is the approach favored by McKenzie et al. (2010) and Brian, Chowdhury and Mobarak (2011). Those RCTs involved migration from Tonga to New Zealand for applicants to a lottery who had proof of employment in New Zealand, and financially incentivizing temporary migration in Bangladesh, respectively. Our goal is more expansive. We wish to understand whether

13

migration is a way of getting ahead for the millions of Black South Africans who elect to move and the tens of millions in households impacted by a move. While one can imagine an RCT that spoke to this question, the practicalities of actually implementing such an RCT across a country as ethnically and geographically diverse as South Africa are daunting. Rather, we rely on non-experimental data. Given this, the next question is the choice of estimator.

An often preferred approach to the endogeneity induced by selection bias is an instrumental variables (IV) estimator. The advantage of the IV estimator is that it addresses selection based both on observable and unobservable characteristics. The feasibility of this approach hinges on whether there are good instruments. In our context, instruments need to be correlated with the migration decision and orthogonal to income shocks. This is a tall order to fill. While there are special cases when a clever instrument exists, we have come up short.[12] This is in part due to the scope of the question we address - migration on a national level.

Truly random migration is very seldom observed outside an experimental setting (and history has not looked kindly on those examples that do exist.) Instead of looking for an estimation strategy that recovers the impact of migration were migration status randomly assigned across the entire population, we instead estimate the effect of migration for those who moved (the average treatment effect on the treated, in the language of program evaluation), and we do so using matching estimators. This approach simply does not speak to the economic impact of a policy that reduced the costs of migration for the entire population. But for the question posed at the outset, "Is migration a way of getting ahead?", our matching estimators are on point.

Matching estimators in the context of migration were discussed by McKenzie et al. (2010). Ham et al. (2011) used matching to estimate the returns to migration for young men in the US. Matching estimators are generally considered inferior to experimental estimators because they can control for selection only on observables. In formal terms, matching assumes that the distribution of potential incomes of migrants and non-migrants are independent of migration conditional on the set of covariates, X. Let D denote migration status, with $D = 1$ for migrants (migrant-households)

---

[12]McKenzie et. al. use distance to the Department of Labor office since it turned out that simply knowing about how the lottery worked was an important determinant to whether one applied for the lottery.

and $D = 0$ for non-migrants. Similarly, Y_1 is income after migration and Y_0 is income for non-migrants in the corresponding period. Then the assumption underlying matching is that

$$(Y_1, Y_0) \perp\!\!\!\perp D | X \qquad (1)$$

If this is true, then conditional on covariates X, non-migrants have the same income distribution that migrants would have experienced without migration, and migrants have the same income distribution that non-migrants would have experienced had they migrated. Matching estimators can then calculate the return to migration by creating a weighted sample of non-migrants such that the distribution of observable characteristics in each group is the same. However, assuming that the returns to migration do not affect the migration decision, even with a large selection of control variables, is probably wrong.

Heckman, Ichimura, Smith and Todd (1998b) and Rosenbaum and Ruben (1983) demonstrate that a weaker condition is sufficient for a valid matching estimator, namely

$$E\left(Y_0 | P(X), D = 1\right) = E\left(Y_0 | P(X), D = 0\right) \qquad (2)$$

where $P(X) = Pr\left(D = 1 | X\right)$. The use of the index $P(X)$ avoids the dimensionality problem that arises with using a large number of covariates, and only mean-independence of the non-migration income is assumed. This amounts to allowing the returns to migration to differ across migrants and non-migrants, while requiring that the non-migration incomes of each group have the same mean. Individuals can self-select based on their expected post-migration income, provided their incomes without migration do not differ. This is the result that Ham et al. (2011) use to justify their matching estimator. Because it does not claim mean equality for Y_1, this estimator cannot be used to measure the average return to migration for the population, or even for a sub-sample of likely migrants. It can only measure the returns for those who migrated, because only Y_0 is assumed equivalent for migrants and non-migrants. It does not speak to the income that non-migrants would experience if they migrated, but only to the income that migrants would experience had they not

15

migrated.

Even in this less restrictive case, matching estimators may still be biased compared to experimental estimators. The extent and sources of this bias were studied in detail by Heckman, Ichimura and Todd (1997) in their evaluation of non-experimental relative to experimental methods using a US job-training program. They identify three contributors: nonoverlapping support between treatment and control populations; different distributions of covariates X within the two populations; and genuine selection bias due to selection on unobservables. In the cases they examine, the larger share of measured bias was due to the first two contributors, not to true selection bias. If matching methods are correctly applied, these first two sources of bias can be eliminated and the remaining bias in measurements, due to selection on non-observables, will be small. We perform bounds tests to examine this concern.

The two additional sources of bias that commonly arise in nonexperimental evaluations are due to geographic mismatch between treatment and control groups, and the use of different survey instruments (Heckman et al. (1997)). For our purposes, the latter is not of concern. Information on both migrants and non-migrants was collected in the same nationally representative survey. We additionally have access to sufficiently detailed geographic information to place migrants and non-migrants into the same (pre-migration) labor markets, which increases the plausibility that Y_0 is truly equivalent for both groups. The limiting factor in practice is sample size, which prevents matching within District Council. We instead match within the same type of labor market, explained further in Section 4.2.1.

Many migration papers (and many papers studying income effects more broadly) advocate for the use of differenced data. We experiment with differenced data, but in the case of migration, first-differencing turns out to be problematic. In general, differencing is an effective strategy for dealing with unobservable individual attributes that do not change over time and for which the impact on income is time invariant. In the case of migration, the impact of unobservables on income may well not be invariant to the change in location inherent in migration. Indeed, the notion that an unobserved component of ability will be more highly rewarded in a new location may be the

reason for migration. Another problem with differenced data arises when when migration involves a change in household composition and when income is measured by household per-capita income (as in Examples 2, 3, and 4 in Table 6.)

An example illustrates the issue. When the household is comprised of multiple individuals, correlates of household per capita income for individual i include information about other members of i's household. Some of these correlates will be unobserved. For example, if individual i's household includes a cousin, William, who is lazy and stupid, this would exert a negative influence on the residual in a regression of i's per capita household income on a set of observables. If William is in the household both periods, differencing the data will sweep out this unobservable influence on household per capita income. For Example 1 in Table 6, DD estimators work as expected. When migration involves a change in the household composition, though (as in Examples 2, 3, and 4 in Table 6), DD estimators run into problems. This is because the unobservable that captures cousin William's negative impact on household per capita income in period t may no longer be present in period $t + 1$. Hence, when household composition changes, in the presence of income pooling first differencing the data no longer sweeps out all the time invariant unobservables that might impact household per capita income, and which might be correlated with migration. Because of this issue, we rely principally on matching estimators.

Independent of exactly which matching estimator is used and on which variables we base the match, a logically prior question is just which match identifies the causal impact of migration in each of the four examples in Table 6. That is, on what should one match to identify the causal impact of migration on the individual migrant (Example 1), the migrant who switches households (Example 2), the sending household (Example 3), and the receiving household (Example 4)?

In each case, we start by noting the change in income that is observed in the data. We then ask, "What is the unobserved counterfactual change that, when compared to the actual change, identifies the causal impact of migration?" Answering this requires pinpointing just what part of the counterfactual change is unobserved and then selecting the appropriate match to "proxy" for this unobserved.

In Example 1, we observe the change in per capita household income for the migrant whose household moved *en masse.* We want to know how that migrant's household per capita income would have changed had they not moved. Since we observe the migrant's income in period t prior to the move, the missing piece of information is the migrant's per capita household income in period $t+1$ had she not moved. The match, then, looks for someone who is like the migrant in period t but who did not move. This non-mover's income in period $t+1$ is our estimate of what the migrant's income would have been and so allows us to estimate the counterfactual income change against which to compare the actual income change. The difference is the causal impact of migration.

Example 2 is similar. We observe the actual change in per capita household income for the migrant who, in this case, changes households. We want to know what the migrant's per capita household income would have been if she had stayed in her original household. The match, then, looks for an individual like the migrant, in a household that is like the migrant's in period t but did not experience a migration event and asks what their period $t+1$ per capita household income is. The difference between the migrant's actual change in per capita household income and the matched estimate of what it would have been absent leaving their original household is the causal impact of migration. It might seem that a good proxy for the migrant's per capita household income in $t+1$ but for the move is the observed per capita income of the sending household members (those who did not move.) This would be appropriate if the household were atomistic and did not somehow re-optimize after the departure of the migrant. This is probably not defensible.

In Example 3, we observe the actual change in per capita household income for the sending household. The unobserved counterfactual is what this change would have been had the migrant not departed. We observe the sending household's actual per capita income in period t so the unobserved is the sending household's period $t+1$ per capita household income but for the departure of the migrant. This is identical to the unobserved in Example 2. The only difference is that in Example 3, we compare the counterfactual change in income to the sending household's actual change while in Example 2 we compare the counterfactual change to the migrant's actual change. Our matching algorithm, then, again looks for a household that is like the migrant's in period t but

did not lose a household member (or members) due to migration and then asks what their period $t + 1$ per capita household income is.

Example 4 highlights the causal impact of migration on the receiving household. We observe the actual change in per capita household income for the receiving household.[13] The counterfactual is how the receiving household's per capita household income would have changed if it had not taken in the migrant(s). We observe the receiving household's period t income so the unobserved is the receiving household's income in $t + 1$ had they not taken in the migrant(s). The match, then, finds a household that is like the receiving household in period t but which did not receive a migrant and asks what that household's period $t + 1$ income is.

Two remaining implementation issues are which observables are used to conduct the match and which particular matching estimators are used. Each is discussed in turn.

4.2.1 Conditioning variables

The same set of conditioning variables is used for all our specifications. Ham et al. (2011) demonstrates, specifically in the context of measuring the return to migration using matching estimators, that a comprehensive approach is best. All variables that affect the income or wage should be included, as well as all available variables that are correlated with the underlying variable driving the migration decision. Thus, traditional income determinants such as age and education will be included as well as variables that potentially influence the migration decision - household structure, location and prior earnings. We do not condition on full labor market histories to avoid excluding individuals with limited participation or wage information and individuals who are migrating for non-labor market reasons. The exact variables used are quartics in age and education, as well as an interaction between age and education, gender, marital status[14], province, community type, inter-

---

[13]This point tends to generate some confusion. We observe a CSM household in year t. If a migrant joins tis household by year $t + 1$, we again aboserve the newly expanded household inclusive of the migrant (a TSM) in the next wave. What we do *not* observe is the migrant in period t but in Example 4 our focus is on the receiving household.

[14]Due to the ubiquity of long-standing cohabitation of non-married couples in South Africa, individuals who cohabit with a partner are counted as 'married'.

actions between all the above and gender, and income in the survey wave prior to migration.

The inclusion of income merits special mention. Our matching estimator finds a non-migrant who "looks like" a migrant based on pre-migration characteristics. By including income, the implicit assumption is that had the migrant not moved, his/her income would have continued on a similar trajectory to that of the non-migrant. The migrant may well have particular unobservables that impact income– this is not ruled out. Our approach simply assumes that these particular unobservables come into play due to migration. Had the migrant not moved, these unobservables would have continued to impact income as they had prior to migration as reflected in the migrant's pre-migration income. This is why a non-migrant with similar income is an appropriate control for what the migrant could have expected had they not moved.

Household-level variables included are the mean age and education of the household, household size, whether it is rural or urban, whether it contains a pensioner, a female pensioner, a child under seven or a child under three, and the fraction of the household that is female, employed, prime-aged (18-65), and the fraction that are under eighteen, under sixteen[15], under seven[16] or under three. We match within the type of community in which the household resides so as to capture the potential importance of local or regional labor markets[17], so that migrants must be matched to non-migrants who reside in the same type of labor market as they did initially.

4.2.2   Matching Estimators

Propensity score matching has been the traditional solution to the dimensionality problem created by having many covariates on which to match. A probit (or logit) model is used to calculate the probability that any one individual moves given their covariate values. Movers are then matched to non-movers with similar probabilities of moving. This requires defining "similar." The simplest

---

[15]These children cannot legally work and their guardians are eligible for Child Support Grants.

[16]Children under seven do not have to be enrolled in school.

[17]The four potential types are urban formal, urban informal, rural formal, and former Tribal Authority. The latter two differentiate between rural areas with reasonably well-functioning local labor markets and infrastructure, and rural areas in the formerly Black areas of South Africa, which have a long history of low government provision of services and infrastructure, very low formal employment and very high poverty rates

approach is to use nearest neighbor matching - the mover is matched to the non-mover with the closest propensity score value. However, this is inefficient - it uses only one of many potential matches and thus discards much useful information. A partial solution is to use an average of the K nearest neighbors (K=2,3,etc) instead of the single nearest neighbor. This reduces the standard errors of the estimates, as more information is available, but is problematic because the 'nearest' neighbors for a particular individual will be of varying closeness depending on the density of the data around that individual.

Heckman et al. (1997), Heckman, Ichimura and Todd (1998a) and Heckman et al. (1998b) incorporate local regression into matching. Instead of choosing one (or K) control individuals to match to each treated individual, everyone with a propensity score within a window around each treated individual's propensity score is used to create a weighted average counterfactual income for the migrant, with weights decreasing in their distance from the migrant. This has the advantage of increasing the information used (and thus decreasing the variance of the estimates) while limiting the increase in bias through the weighting procedure. Fan and Gijbels (1996) recommend the use of a local linear, or at times a local cubic, regression.

Frolich (2004) argues that kernel regression (essentially a local regression of degree zero) is more robust to specification errors than linear regression. However, this problem can be partially addressed through the use of a variable bandwidth and is most problematic when the control group is not substantially larger than the treatment group (less than five to one). The ratio of non-migrants to migrants among Blacks is closer to four to one, and the ratio of those affected by migration to those unaffected is almost one to one, so this is a concern for our analysis. However, local linear regression matching is more robust to asymmetric distributions of control individuals around treated individuals, which is a feature present in our data (Caliendo and Kopeinig 2008).

We employ multiple estimators to determine the robustness of our estimates. Our preferred estimator uses a local linear regression with a normal kernel, to make use of more information than a nearest neighbor match while limiting the bias from decreased match sensitivity.

5   The Causal Impacts of Moving

We organize our base case results by first examining the causal impact of migration from the migrant's perspective. We then conduct our analyses from the perspective of the sending household and finally from that of the receiving household.

5.1   Returns to Migration for the Migrant

We begin by examining the returns to migration using the change in individual (as opposed to household per capita) income as the outcome measure. Because many migrants have zero indvidual incomes, we report these results in levels, not logs. This focus on the change in the migrants' individual incomes mimics the typical focus in the literature as in McKenzie et al. (2010) and Ham et al. (2011).

Table 7 presents estimated returns for moves that occurred during 2008-2010, moves during 2010-2012, and again for any move between 2008 and 2012. Within each time period, we report results using the propensity score matching (P-score) estimator as well as the difference-in-difference (DD) estimator. Finally, we report returns for all migrants and then separately for migrants by sex and by whether or not the migrant holds a Matric certificate (the rough analogue to an American high school diploma.)

The results in Table 7 are fuzzy and often conflicting. While returns tend to be higher in the 2010-2012 period than in 2008-2010, the estimated returns are often imprecise and frequently one cannot reject that the return is zero. To put the magnitude of these estimates in perspective, mean individual income is about R1450 when everyone is included and about R2500 if one conditions on non-zero income.[18] Focusing on the top row which reports returns for the entire sample, the P-score estimator gives a negative return in 2008-2010 while the DD estimate is about zero. In 2010-2012, the P-score estimate is insignificantly different from zero while the DD estimate is positive. For the

---

[18]Prices are deflated to a base period of December 2012 at which point the exchange rate was about R8.7 to the US Dollar.

entire period, the matching estimator finds no return to migration for the entire sample nor for any sub-group, and the DD estimator does not yield a return different from zero for men although it does for women. Table 7 raises an important question: Why are over a quarter of all Blacks moving when the economic return to doing so is often not statistically different from zero? Our answer is in the next table.

Table 8 also presents returns to migration for the migrant, but this time the outcome variable is the change in household per capita income (HHPCI) instead of individual income. Because HHPCI is never zero, we estimate returns using the change in log income resulting in more easily interpreted coefficients. Also, all estimates are obtained using the P-score matching estimator since first differences in HHPCI no longer make sense with changing household composition. The columns headed "HHPCI" use changes in log HHPCI. We include results for changes in log individual income for purposes of comparison and these are reported under the column header "Ind. Income."[19] The results are striking. There are three main messages from Table 8.

First, a very different picture of the returns to migration emerges when HHPCI is the metric. The returns to migration are much higher when the outcome variable incorporates the pooled nature of household income. The return to migration is typically at least twice as high with HHPCI as the metric. Clearly, much of the gain from migration comes not just from a (possibly) new labor market outcome but also from the shared income of the endogenously formed household. We view these results as validating the focus on household rather than individual income. The rather stunning magnitude of migration observed in the data as noted in Section 2 begins to make economic sense when viewed through the lens of the household.

Second, these are large returns to migration. Focusing on the top row of the table, during the recessionary period of 2008-2010, the return to migration was 28.4%. In the more expansionary period of 2010-2012, the return jumped to almost 39% and across the entire period, the return was almost 31%. When we ask whether migration is a way of getting ahead, the answer for the migrant

---

[19]The individual income results exclude all individuals who had incomes of zero. Hence, these results are obtained from a (non-random) subsample relative to Table 7.

is clearly "yes." All of these returns using HHPCI are also quite precisely estimated.[20]

Third, the returns to migration are heterogeneous and the next four rows of the table illustrate this. From 2008-2010, the gains to migration accrued to those without a Matric– returns of 43% for males and 20% for females. The returns to migration for those with a Matric were around 20% but those returns are not precisely estimated. During the 2010-2012 period, this pattern reversed and returns to migration were higher for those with the Matric. The returns during this expansionary period were quite large. Female migrants saw a 77% return to migration and males a 56% return. Returns for those without the Matric were more modest but still substantial (38% for males and 20% for females.)

The broad pattern is one in which migration is a way of getting ahead for the migrant and that the returns are higher during the macroeconomic upswing. Even during the downswing, though, migration enhanced migrants' incomes. The results support a "moving to opportunity" view of migration rather than a "push out of the nest" view. Another interpretation is that migrants either choose their destination households well or that they are invited into destination households that are better off than the household left behind. What is clear is that migration causally increases the migrants' household per capita income.

5.2 Returns to Migration for Associated Households

We turn next to the returns to migration for the sending and receiving households. As noted in section 3, over 80% of Blacks lived in a household impacted by migration over the 2008-2012 period. Table 9 summarizes how members of these households fared. Households impacted by a migration

---

[20]Unlike some of the recent literature on matching, we do not bootstrap the standard errors. Imbens (2014) and Abadie and Imbens (2008) demonstrate that bootstrapping is not in general a valid way to calculate standard errors as matching estimators are not asymptotically linear. Instead, we use the standard error procedure proposed in Abadie and Imbens (2006) and available in the Stata program *psmatch2*, written by Leuven and Sianesi (2003). The overall idea of this procedure is that, to estimate consistent asymptotic variances for the sample average treatment for the treated (as opposed to for the population ATT), one does not need consistent estimates of the conditional outcome variances for treated and control groups at *all* values of the covariates. It is sufficient to have the average of these variances over the distribution of outcomes, which, when weighted with the inverse probability of group assignment (treatment and control, respectively), can be used to construct a consistent estimator of the variance of sample ATT. Alternative standard error treatments are reported in Section A.

event might send a migrant, might receive a migrant, or both. The last of these turns out to be quite common.

The first row reports the returns to households that sent a migrant (or migrants) irrespective of whether the household might have also received a migrant. Sending a migrant increased the sending household's HHPCI by 21-22% and this substantial return was constant across the sample period. Some of these sending households might also have received a migrant. When we exclude these households and look at households that only sent a migrant, the returns are given in the second row of Table 9. Although the return falls to 11.3% in the 2010-2012 period, the overall 2008-2012 return for "sending only" households is about the same as in the first row. Sending households clearly benefited from the departure of the migrant(s). All of the returns in the first two rows of Table 9 are precisely estimated.

The next two rows report the returns to households that received a migrant (but might also have sent one) and households that only received a migrant, respectively. Taken together, these rows present a cogent story. If a household only received a migrant, the household suffered an average decline in HHPCI of about 21% over the 2008-2012 period.[21] Across the two-year periods, the declines were 16.5% and 11.5% respectively, and the overall decline from 2008-2012 of 21% means that many households received migrants in both periods. When we look at any household that received a migrant, many of which also sent a migrant, the overall change in income is a decline of about 2% and is insignificantly different from zero. This suggests that the cost of receiving a migrant was often offset by the benefit of sending one.

If we restrict the sample to households that both received and sent a migrant, we find that these households clearly benefited from the reshuffled household roster. These results are reported in the

---

[21]Observant readers will note that for some groups in Tables 9 and 8, income changes for the full period (2008-2012) are higher than those for either single period. This is a result of two mechanisms. First, most households' income changes over three periods will be greater than the change between two periods. Regressions and matching remove the linear component that's common to mover and non-mover households. However, when log-changes are calculated, neither approach compensates for the fact that the base income is different in 2010 and 2008. The total percentage change between 2008 and 2012 will not necessarily be the same as the sum of the percentage change between 2008 and 2010, and the percentage change between 2010 and 2012. Second, not all households (and migrants) are observed in all periods. For instance, some households are not surveyed in 2010, but return to the sample in 2012. Such households will be picked up in the 2008-2012 calculation, but not in either single period calculation.

last line of Table 9. In the first period, the causal impact of migration was a 14% increase in HHPCI. This figure rose to 25.3% in the second period and was 19% for the entire sample period.

The overall picture is one in which migration causes an increase in HHPCI for the sending household and causes a decrease in HHPCI for the receiving household. The fact that migration into an existing household is so common and yet comes at a cost to members of the receiving household speaks to the strong social fabric of Black South African households.

5.3 Robustness of Results

We have examined the robustness of our results, and details are included in the Appendix. Here, we summarize that analysis.

Our robustness analysis addressed five concerns. First, are our results sensitive to the choice of estimator, the method of computing standard errors, and/or the choice of conditioning variables? This comprises a standard and extensive sensitivity analysis. Second, are we confident selection on unobservables might not be such that the estimates we obtain are an artifact of the selection on observable approach we adopt. Third, we investigate the plausibility of the idea that migrants' income paths, had they not migrated, would have evolved as the non-migrants' income paths did. Fourth, we experiment with an alternative definition of migration and examine returns when we restrict our definition of migration to include only long distance moves. Fifth, we examine the heterogeneity of returns to migration across different subsamples.

We find that the estimated returns to migration are quite robust with respect to the choice of estimator and the choice of conditioning variables. Standard errors, too, are robust to alternative estimation approaches. We find that the degree of selection on unobservables necessary to make our estimated returns disappear is implausibly large. Using two methodologies, we find support for the selection on observables approach we adopt. We also find support in the data for our assumption that the migrants' income paths, had they not migrated, would have evolved like the non-migrants' income paths. This too supports our empirical strategy.

When we restricted the definition of migration to only include long distance moves, the returns to migration for the migrants increased, sometimes substantially, and this leads us to conclude the returns we report in the main text are lower bounds. When we examined returns for various sub-samples of the data, we found returns were highest for urban to rural migrants. Dividing the sample by age of migrant also revealed substantial heterogeneity of returns.

5.4 Summing Up

When we consider the entire 2008-2012 sample period and review returns for the migrant, for households that only sent a migrant, and households that only received, we find that migrants experienced a 31% increase in HHPCI, members of sending only households experienced a 21% increase in HHPCI and the members of receiving only households saw HHPCI fall by 21%. With HHPCI as the metric, migration clearly benefited the migrant. The sending household was also better off with the migrant's departure but this gain was almost exactly offset by the decline suffered by the receiving household.

These results are consistent with a pattern in which the migrant was a drain on household resources and so moves to opportunity, leaving the sending household and the migrant better off. Some of the "opportunity" to which the migrant moves, though, may be the pooled resources of the receiving household since these receiving households suffer declines in HHPCI.

The macroeconomy also matters. The returns to the migrant are highest when the economy was expanding and the costs imposed on the receiving household are lower during expansionary times. The sending household benefits less from "unloading" a less productive household member when times are better. All of these patterns are reasonably intuitive and highlight the relevance of *when* migration happens and not just where it occurs.

Given the remarkable prevalence of migration in South Africa, it is perhaps not surprising that the migrants themselves benefit from changing households. In results reported in the Appendix, in Table 15, we find that moves that are more likely to involve a change in the labor market, proxied

either by moving more than 50km or by changing District Councils, resulted in even larger gains for the migrant. Also, moves from rural to urban areas exhibited yet larger gains for the migrant. Over the 2008-2012 period, migrants moving from rural areas to urban areas experienced a return of almost 80% (See Table 16).

It is common in the presence of large returns to migration to question why yet more individuals don't migrate. First, it's important to realize that the estimated returns are the average treatment effect on the treated. It would be incorrect to extrapolate these returns to the entire population. Second, most migrants move into an existing household and while the migrant is clearly better off, the members of the receiving household are not. While the social fabric that leads to accepting a migrant into a household appears strong, that welcome may not be unlimited. This may place a limit on migration. Third, rather than wonder why if the returns are so high more South Africans do not migrate, perhaps it is more appropriate to note that because returns are so high, a huge number of South Africans do in fact migrate.

6   Conclusions

Migration in South Africa, as in many African countries, is very much a story about endogenous household formation. Some members of the original household migrate and join another already existing household. In this context, examining the economic returns to migration is a multi-layered exercise.

We find that migration among Black South Africans is the norm. Over only a four year period, about 80% of Blacks are impacted by a migration event. For the migrant, the economic returns to moving are substantial– HHPCI increased about 28% with migration during the economic downswing and by about 39% during the subsequent upswing. These are big returns. Migration in South Africa typically involves a reshuffling of the household roster, and this, combined with pooled household incomes, means that migration impacts both the household the migrant left and the one that was joined. We find that on average the sending household members enjoyed about a 21% increase in

HHPCI with the departure of the migrant(s), while members of receiving households suffered an approximately offsetting decline. These results highlight the importance of better understanding the economics of household formation.

We conclude with a cautionary note. While it is critical that reliable estimates of the returns to migration account for selection, doing so has meant choosing our econometric strategy carefully but it has also meant being circumspect about the question we answer. For example, it's often argued that lowering the costs to migration would enhance economic growth. That's probably true, but our estimates answer a more narrow question. Our estimates of the returns to migration only apply to those who opted to migrate (or to send or receive a migrant) and so it would be inappropriate to infer that our returns would apply to the population on the whole.

References

Abadie, Alberto and Guido Imbens, "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects," *NBER Technical Working Paper*, 2002, *283*, 1–51.

____ and ____ , "On the Failure of the Bootstrap for Matching Estimators," 2004.

____ and ____ , "Large Sample Properties of Matching Estimators For Average Treatment Effects," *Econometrica*, 2006, *74*, 431–497.

____ and ____ , "On the Failure of the Bootstrap for Matching Estimators," *Econometrica*, 2008, *76(6)*, 1537–1557.

____ and ____ , "Matching on the Estimated Propensity Score," August 2009. NBER Working Paper 15301.

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, March 1999, *67* (2), 251–333.

Altonji, Joseph G., Christopher R. Taber, and Todd E. Elder, "Selection on Observed and Unobserved VVariable: Assessing the EffectiEffect of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

Ardington, Cally, Anne Case, and Victoria Hosegood, "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa," *American Economic Journal: Applied Economics*, 2009, *1*, 22–48.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon, "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey," *The Review of Economics and Statistics*, 2011, *93* (3), 1010–1033.

Brian, G., S. Chowdhury, and A.M. Mobarak, "Seasonal Migration and Risk Aversion: Experimental Evidence from Bangladesh," 2011. Yale University School of Management Working Paper.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Under-investment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 2014, *September*.

Budlender, D. and F. Lund, "South Africa: A legacy of family disruption," *Development and Change*, 2011, *42* (4), 925–46.

Caliendo, Marco and Sabine Kopeinig, "Some Practical Guidance For the Implementation of Propensity Score Matching," *Journal of Economic Surveys*, February 2008, *22* (1), 31–72.

DeVilliers, L., I. Woolard, R. Daniels, and M. Leibbrandt, "National Income Dynamics Study Wave 3 Users Manual," Technical Report, Southern Africa Labour and Development Research Unit, Cape Town 2013.

Fan, Jiaqiang and Irene Gijbels, *Local Polynomial ModModel and Its Applications: MonogrMono on Statistics and Applied Probability 66*, Chapman & Hall, 1996.

Frolich, Markus, "Finite Sample Properties of Propensity-Score Matching and Weighting Estimators," *Review of Economics and Statistics*, 2004, *86*, 77–90.

Ham, John, Xianghong Li, and Patricia Reagan, "Matching and Semi-Parametric IV Estimation, A Distance-Based Measure of Migration, and the Wages of Young Men," *Journal of Econometrics*, April 2011, *161* (2), 208–227.

Hamoudi, Amar and Duncan Thomas, "Endogenous Coresidence and Program Incidence: Re-examining South Africa's old age pension," *Journal of Development Economics*, 2014, *109*, 30–37.

Harris, John and Michael Todaro, "Migration, unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 1970, *60*, 126–142.

Heckman, J., H. Ichimura, and P. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 1997, *64*, 605–654.

___ , ___ , and ___ , "Matching as an Econometric Evaluation Estimator.," *Review of Economic Studies*, 1998, *65*, 261–294.

___ , ___ , J. Smith, and P. Todd, "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 1998, *66*, 1017–1098.

Imbens, Guido, "Matching Methods in Practice: Three Examples," March 2014. NBER Working Paper 19959.

Lechner, M., *Identification and estimation of causal effects of multiple treatments under the conditional independence assumption.*, Heidelberg: Physica,

___ , "Some practical issues in the evaluation of heterogeneous labour market proprogram by matching methods.," *Journal of the Royal Statistical Society*, 2002, *165*, 59–82.

Leuven, Edwin and Barbara Sianesi, "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.," 2003. This version 4.0.7.

Lewis, W. Arthur, "Economic Development with Unlimited Supplies of Labor," *Manchester School*, 1954, *22*, 139–191.

McKenzie, David, John Gibson, and Steven Stillman, "How Important is Selection: Experimental Vs Non-experimental Measures of the Income Gains from Migration," *Journal of the European Economic Association*, 2010, *8* (4), 913–945.

Morten, Melanie, "Temporary Migration and Endogenous Risk Sharing in Village India," 2013. Stanford University and Federal Reserve Bank of Minneapolis.

National Income Dynamics Study 2008, Wave 1 [dataset], Technical Report, Southern Africa Labour and Development Research Unit 2014. Version 5.2. Cape Town: Southern Afria Labour and Development Research Unit [producer], Cape Town: DataFirst [distributor].

National Income Dynamics Study 2010-11, Wave 2 [dataset], Technical Report, Southern Africa Labour and Development Research Unit 2014. Version 2.2. Cape Town: Southern Afria Labour and Development Research Unit [producer], Cape Town: DataFirst [distributor].

National Income Dynamics Study 2012, Wave 3 [dataset], Technical Report, Southern Africa Labour and Development Research Unit 2014. Version 1.2. Cape Town: Southern Afria Labour and Development Research Unit [producer], Cape Town: DataFirst [distributor].

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Validation," January 2015. Oster website, Brown University.

Posel, D. and D. Casale, "Internal labour migration and household poverty in post-apartheid South Africa," in H. Bhorat and R. Kanbur, eds., *Povery and Policy in Post-Apartheid South Africa*, Vol. Poverty and Policy in Post-Apartheid South Africa, HSRC Press- Cape Town, 2006, pp. 351–365.

Posel, Dorit, James Fairburn, and Frances Lund, "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa," *Economic Modelling*, 2006, *23* (5), 836–853.

Rosenzweig, Mark and Oded Stark, "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *Journal of Political Economy*, 1989, *97*, 905–926.

Young, Alwyn, "Inequality, the Urban-Rural Gap, and Migration," *Quarterly Journal of Economics*, November 2013, *128*, 1727–1785.

Table 1: Sample size by race and sample year

Race	2008	2010	2012	Total
African	22,255	28,193	30,881	33,604
	78.70	*80.60*	*80.86*	*81.35*
Coloured	4,161	4,823	5,280	5,606
	14.71	*13.79*	*13.83*	*13.57*
Asian/Indian	429	470	480	503
	1.52	*1.34*	*1.26*	*1.22*
White	1433	1492	1550	1593
	5.07	*4.27*	*4.06*	*3.86*
Total	28,278	34,978	38,191	41,306
	100	*100*	*100*	*100*

Population percentages shown in italics. Unweighted.

Table 2: Migration status by age and sample years

	2008-2010	2010-2012	2008-2012
No Move	18,341	20,657	21,590
Move	3,122	4,185	5,860
Age of Migrant			
0-5	510	562	874
6-10	281	407	529
11-15	330	413	631
16-20	446	607	896
21-25	475	668	800
26-30	352	497	542
31-35	214	292	364
36-40	135	221	242
41-45	124	142	207
46-50	73	125	127
51-55	62	77	99
56-60	39	58	67
61-65	15	27	28
65+	49	80	97

All Blacks, unweighted. Moves by age do not exactly equal total number of reported moves due to some missing observations for age.

Table 3: Individuals living in households affected by migration

		2008-2010		2010-2012		2008-2012	
All	Unaffected	16,136	*48.02*	14,481	*43.09*	6,543	*19.47*
	Affected	17,468	*51.98*	19,123	*56.91*	27,061	*80.53*
	Trailing	7,251	*41.51*	9,647	*50.45*	13,714	*50.68*
	Receiving	14,986	*85.79*	15,341	*80.22*	24,434	*90.29*
Wave 1 Sample	Unaffected	10,308	*46.32*	10,768	*48.38*	5,789	*26.01*
	Affected	11,947	*53.68*	11,487	*51.62*	16,466	*73.99*
	Trailing	7,251	*60.69*	7,906	*68.83*	11,973	*72.71*
	Receiving	9,465	*79.22*	8,864	*77.17*	13,903	*84.43*

Black sample only, all ages, unweighted. Population percentages shown in italics. Sending and Receiving percentages refer to the percentage of individuals affected by migration living in sending and receiving households respectively.

Table 4: Type of move (2008-2012)

	To Urban	To Rural	Total
From Urban	1447	328	1775
From Rural	884	1686	2570
Total	2331	2014	4345

Black sample only, all ages, unweighted.

Table 5: Changes in log per capita household income

		2008-2010	2010-2012	2008-2012
No Move	Male	0.043	0.182	0.217
	Female	0.041	0.191	0.217
Move	Male	0.256	0.432	0.445
	Female	0.115	0.510	0.430

Black sample only, all ages. Sample weights used.

Table 6: Household Composition with Migration

	Period t	Period $t+1$
Example 1	M_t	M_{t+1}
Example 2	$M_t + T_t$	$M_{t+1} + R_{t+1}$
Example 3	$M_t + T_t$	T_{t+1}
Example 4	R_t	$R_{t+1} + M_{t+1}$

34

Table 7: Returns to migration for the migrant using individual income

	2008-2010		2010-2012		2008-2012	
	P.kernel	DD	P.kernel	DD	P.kernel	DD
All	-612.42	-147.43	556.61	896.93	128.44	470.32
	189.10	*203.08*	*293.07*	*173.92*	*186.47*	*149.24*
Male, matric	-2475.60	-1111.92	1019.63	1075.03	14.68	763.50
	963.73	*998.74*	*582.38*	*620.84*	*793.54*	*682.69*
Male, no matric	-833.23	-95.54	1058.25	1056.68	226.80	430.91
	509.72	*491.48*	*951.71*	*468.12*	*531.73*	*425.89*
Female, matric	-1189.64	435.73	272.18	1847.10	29.75	537.59
	678.64	*557.12*	*502.12*	*538.59*	*316.28*	*266.95*
Female, no matric	-113.89	439.33	120.29	328.25	28.92	408.45
	90.71	*106.88*	*70.54*	*74.25*	*84.21*	*80.40*

Dependent variable is change in individual income (not logged). Columns labeled "P. kernel" show returns using local linear regression propensity score matching on individual and household characteristics with a Gaussian kernel. Columns labeled "DD" calculate display returns from difference-in-difference regressions, weighted by population sampling weights.

Table 8: Returns to migration for the migrant using log of income

	2008-2010		2010-2012		2008-2012	
	HHPCI	Ind. income	HHPCI	Ind. income	HHPCI	Ind. income
All	0.2840	0.0240	0.3881	0.1980	0.3089	0.1452
	0.0493	*0.0640*	*0.0378*	*0.0508*	*0.0321*	*0.0459*
Male, matric	0.1974	-0.2496	0.5623	0.3462	0.5290	0.0479
	0.1727	*0.2725*	*0.1220*	*0.1402*	*0.1074*	*0.1724*
Male, no matric	0.4254	0.0218	0.3757	0.2293	0.3577	-0.0187
	0.0998	*0.1435*	*0.0694*	*0.1527*	*0.0564*	*0.0916*
Female, matric	0.2079	-0.1051	0.7747	0.3311	0.3042	0.1615
	0.1586	*0.2126*	*0.1352*	*0.2014*	*0.0871*	*0.1179*
Female, no matric	0.2039	0.0869	0.1991	0.0465	0.2239	0.1347
	0.0689	*0.0952*	*0.0729*	*0.0782*	*0.0460*	*0.0551*

All returns are calculated using local linear regression propensity score matching on individual and household characteristics with a Gaussian kernel. For columns labeled "HHPCI", the dependent variable is the change in log of household per capita income. For columns labeled "Ind. income", the dependent variable is the change in log of individual income.

Table 9: Returns to migration for the associated households

	2008-2010	2010-2012	2008-2012
Sending	0.2128	0.2259	0.2273
	0.0238	*0.0205*	*0.0219*
Sending only	0.2233	0.1130	0.2046
	0.0312	*0.0261*	*0.0348*
Receiving	-0.0954	0.0325	-0.0227
	0.0205	*0.0188*	*0.0207*
Receiving only	-0.1652	-0.1153	-0.2096
	0.0233	*0.0220*	*0.0281*
Both	0.1413	0.2535	0.1898
	0.0327	*0.0281*	*0.0279*

Dependent variable is log of household per capita income. All specifications use household characteristics as controls for a propensity score matching approach based on local linear regressions with a Gaussian kernel.

Appendices

A   Appendix: Robustness of Results

In this section, we examine the robustness of our results. A common concern with matching estimators is their sensitivity to mis-specification and choices of conditioning variables. To address this concern, we calculate the returns to migration using several different matching algorithms and estimate standard errors for our preferred specifications using three methods common in the literature. These results are discussed in Section A.1, and can be summarized easily: the results are not very sensitive to specification, and alternative standard error calculations do not change the significance of the results. The next two sections assess two possible theoretical problems with our econometric strategy: selection on unobservables; and violations of the parallel trends assumption underlying difference-in-difference estimates. In Section A.2, we calculate the degree of selection on unobservables, relative to selection on observables, that is necessary to invalidate our results using approaches developed by Altonji, Taber and Elder (2005) and Oster (2015). Both methods indicate that implausibly high ratios are needed, which confirms that the returns to migration are robustly non-zero. In Section A.3, the relationship between our methods and a standard difference-in-difference approach is discussed, and evidence to support the existence of parallel trends is presented. In Section A.4, we explore the impact of restricting our definition of migration to only include long distance moves. Section A.5 explores the heterogeneity of the returns to migration for the migrant by various subsamples of the data.

A.1   Alternative matching algorithms and standard error calculations

A common concern with matching estimators is their sensitivity to mis-specification and choice of conditioning variables. To address this concern, we compute the effects of migration using twelve different estimators for each of the four questions discussed above - individual and household per capita income for the migrant, and household per capita income for the sending and receiving households.[22] For each case, we use four nearest neighbor matching estimators (with the number of matches set to one, two, five and ten, respectively), a local linear regression matching estimator and a kernel matching estimator, using both propensity score matching and matching on multiple covariates (Mahalanobis matching). The set of matching variables ranges from the full set of polynomials and interactions used in the results reported in Section 5 to a parsimonious set of demographic and location controls. The results in all cases are similar, typically lying within one to five percentage points of the preferred estimate. The more variable estimates are for the individual income measure and for the receiving households. As noted in the text above, many individuals do not have personal income in one or other year, so the individual income measures may suffer from both selection bias (into the labor market), from small sample concerns, and from high variability within individual. The variability for receiving households seems to be primarily due to smaller samples.

Another point of concern for matching estimators is the calculation of standard errors. There are limited results on the asymptotic properties of matching estimators, though recent work by

---

[22]These results are available on request from the authors.

Abadie and Imbens[23] suggests that the most common approach in the literature - bootstrapped standard errors - does not yield correct standard errors and is in fact systematically biased. They instead propose standard error corrections that provide consistent estimates of variances of sample estimators by focusing on distribution-average variance. This is the approach used for our main results. However, our results are robust to alternative standard error treatments. Table 10 gives results for individual migrants with household per-capita income as the metric, by gender and education level and for each of the three migration periods. In the first, fourth and seventh column, the main paper results are reproduced, with standard errors calculated using the Abadie-Imbens correction. In the second, fifth and eighth columns standard errors are bootstrapped. In the third, sixth and ninth columns, standard errors are calculated using another correction commonly found in the literature, proposed by Lechner (2001). This estimator explicitly accounts for the fact that individuals in the control group are used as matches repeatedly, and performs equivalently well to bootstrap in simulated results (Lechner 2002). This method can only be used with K nearest neighbor matching, because kernel or regression matching uses most control observations to calculate each counterfactual match. Thus, the coefficients in these columns differ slightly from the others due to the use of nearest neighbor (K=10) matching instead of local linear regression matching. The choice of standard error calculation does not affect the significance of our results. All coefficients remain significant, except those for high school graduates between 2008 and 2010, which were insignificant in our main results as well.

Overall, we are confident that our results are not sensitive to our choice of matching algorithm or standard error treatment. While the coefficients vary slightly in response to different matching algorithms, the relative sizes and significance are not affected, and the variation is small for almost all measures.

A.2 Selection on observables and unobservables

Fundamentally, matching estimators attempt to calculate the return to migration controlling for selection on observables. If we believe that a large part of the return to migration is generated by selection on an unobserved characteristic, then these calculations will be inaccurate. A common way to avoid concerns about the income effects of unobserved individual characteristics is to use differences in income over time. Provided the unobservables are time-invariant and affect income in time- and space-invariant ways, differencing income removes the unobservable individual effects. Thus, matching with differenced individual income as the outcome addresses both unobservable and observable selection, *provided* the effects do not change when individuals change location. Because we only try to estimate the effects of migration for the migrant (ie. the treatment on the treated), location-varying returns are not problematic - moving to a location in which one's attributes are better-rewarded is simply part of the returns to migration. Matching is valid provided that, absent migration, the migrant's income would have continued on the same trend as that seen for non-migrants in the origin.

We do not get the same benefits when we use household per capita income. By definition, the household of the migrant changes and thus the household unobservables change, so we do not difference out the unobservable household characteristics that may affect income. However, by

---

[23]Abadie and Imbens (2002), Abadie and Imbens (2004), Abadie and Imbens (2006), Abadie and Imbens (2008), Abadie and Imbens (2009)

matching on observed household characteristics, we can still estimate the treatment on the treated effects of sending or receiving a migrant.

Because we are not trying to estimate average population effects of migration, we are not ruling out selection on unobservable characteristics. However, it is still instructive to think about the share of the effects that is attributable to moving to a new location (location-returns to characteristics) and that which is attributable to individual "get-up-and-go". We can do this by considering selection on observables relative to selection on unobservables, as first proposed in Altonji et al. (2005) and expanded on by Oster (2015). The goal of these analyses is to determine the conditions for which the entire measured return to migration is due to selection on unobservables ("get-up-and-go") and there is no return to relocation per se. These conditions are expressed by calculating the ratio of selection on unobservables relative to selection on observables necessary to make the coefficient on migration zero. If the ratio is large, then there must be far more selection on unobservables than on observables.

Neither of these measures have a direct analogue in a matching framework, so for this section the estimates in question are those from difference-in-difference regressions of changes in log household per capita income on the matched sample. The coefficients from these regressions can be found in Table 13.

A.2.1 Selection ratios - if migration is akin to Catholic schools

Altonji et al. (2005) sets forth a thought experiment - what is the likely relationship between selection on observables and unobservables - and presents a formal econometric framework for using observed selection to assess the role of unobservables. Additionally, they propose an informal test that calculates the degree of selection on unobservables relative to selection on observables that would be required to eliminate the entire estimated effect. This method yields a ratio that, in this context, is the amount of unobservable selection relative to observable selection required for none of the income gain to be due directly to migration. Essentially, to what extent can that "get-up-and-go" explain the migrant's income change, independently of migration? The higher the ratio, the less likely that the entire measured effect is due to selection rather than migration itself. One of the key underlying assumptions of the measure is that the control variables available to the researcher are, at worst, a random sample of the relevant variables, and the unobserved variables do not contain *all* the important determinants of the outcome.

A.2.2 Selection and coefficient stability

Oster (2015) proposes a framework in which to think about selection, which incorporates changes in response to additional controls in both the coefficient of interest and the regressions' adjusted R^2. The intuition behind her approach is that coefficient stability should be considered in combination with the amount of information that additional controls add to the regression. When more controls are added to the baseline regression, a small change in the key coefficient and a large change in the R^2 is more reassuring than a small change in the key coefficient coupled with a small change in the R^2. If the additional controls contain a lot of information (and thus have a large effect on the R^2), then the response of the coefficient is more informative than if the additional controls contain little information. Thus, the change in the coefficient should be 'scaled' by the change

in the R^2. The change in the R^2 should also be considered relative to the maximum obtainable R^2. In a mechanical sense the maximum R^2 is always one, but in a world with measurement error and random fluctuations across time and individuals the maximum is likely lower. Assuming that $R^2 = 1$ provides the most conservative estimates. Her framework allows various statistics to be calculated, including a confidence interval for the coefficient of interest, and the ratio of selection on unobservables relative to observables that would be necessary to make the true effect zero, for a given maximum potential R^2.

A.2.3 Results

The results for the household per capita income measure are shown in Table 11 [24]. Each set of columns corresponds to a different observation period, indicated by the years. Each row corresponds to a different population sample. Columns one, four and seven contain the ratio of selection on unobservables relative to selection on observables that is necessary to reduce the migration coefficient to zero, based on Altonji et al. (2005). The other columns contain the equivalent statistic based on Oster (2014)'s method. The ratio in columns two, five and eight is calculated assuming that the maximum R^2 is one, and the ratio in columns three, six and nine is calculated assuming that the maximum potential R^2 is 0.8364[25].

There are two key take-aways from this table. First, the majority of ratios are substantially larger in absolute value than one, which means that selection on unobservables would need to be much greater than selection on observables to reduce the effect of migration to zero. The only exceptions using Altonji *et al*'s technique are for the period 2008-2010, for which the returns to migration are small however they are calculated. More of the Oster ratios are below or close to one, but the majority are still large.

The second notable point is that many of the ratios are negative. This indicates that selection on unobservables must be negative, relative to observable selection, in order to eliminate the returns to migration. For the Oster ratio, additional covariates increase both the coefficient on migration and the R^2. Similarly, for the Altonji *et al* ratio, including migration increases the coefficients on the observed variables. In both cases, this suggests that the added variables help to unravel heterogeneity in the data and allow for more precise estimation of the original variables' coefficients.

Overall, both these approaches indicate that migration itself, as well as individual unobservables such as "get-up-and-go", has a significant positive effect on income in the following period.

A.3   Plausibility of parallel trends assumption

Throughout this analysis, one of the underlying assumptions has been that non-migrants who are matched to migrants are comparable in characteristics, and specifically that migrants' income path, absent migration, would evolve equivalently to the non-migrants' income path. This is an

---

[24]Results using individual income are available in Table 14.

[25]This maximum 'realistic' R^2 is taken from Abowd, Kramarz and Margolis (1999)'s study of wages using a matched employer-employee French administrative dataset. It is the maximum R^2 they obtain for regressions that do not contain firm and industry characteristics, and is the highest R^2 in the literature known to us for income regressions.

assumption found in all difference-in-difference analyses and is often untestable. In an experimental setting, the assumption is plausible due to random assignment. In this non-experimental setting, the assumption cannot be tested for the entire sample. However, it is possible to test it for a sub-sample of the data - permanent survey members who were successfully interviewed in all three waves, and who never moved or moved only between 2010 and 2012.

For this group, there is information from two pre-migration periods available. If both periods are used to construct the match, then we are mechanically requiring that the matched non-migrant sample has a similar income trend (as well as trends in other variables) to the migrant sample. The returns to migration from this procedure, for which the parallel trends assumption holds by construction, can then be compared to the returns to migration obtained within this sub-sample when matching is performed using only the immediate pre-migration information. If these figures are similar, it provides evidence that the parallel trends assumption may hold in the analyses for the entire sample.

These results are presented in Table 12. The first two columns show the returns to migration calculated within the sub-sample, matching on 2010 information alone. The second two columns show the results when matching is performed on 2008 and 2010 information. Results from the full sample for the equivalent period are included in columns five and six for reference. For all cases using household per capita income, the returns to migration obtained using more information are higher than those from less migration. In all but one case (female high school graduates), both sub-sample estimates exceed the main sample results. As is the case with most of the analyses, the results using individual income are more variable due to small samples. The estimates for the entire sample, regardless of sex or education status, do match the above comment: more information is associated with a higher estimate. This provides support for the parallel trends assumption's accuracy. When migrants are matched to non-migrants with the same pre-migration income trend, the measured return to migration is higher than when they are matched to non-migrants with the same pre-migration income level. There is no obvious reason for why the results for this particular sub-sample should differ from the general sample, so this provides some evidence that our estimates of the returns to migration are not biased upwards due to violation of the parallel trends assumption.

A.4 Alternative definitions of migration

Our main results use a broad definition of migration as any switch between households. This may include many local, short-distance moves that almost certainly do not involve changing labor markets. We believe that this is a sound decision in a context of extended household networks and intra-household income sharing, because household composition is an endogenous response to the economic environment and family shocks. However, the more traditional definition of migration involves long distance moves in which migrants change labor markets and travel significant distances away from their families. As another robustness check, we examine how our results change when alternative definitions of migration are used. In Table 15, we present the returns to migration for the migrant, in household per-capita income, by gender and education level, for three additional migration definitions. In columns one, four and seven, only individuals who have moved more than the median distance of a move are counted as migrants. This still amounts to a relatively low number of kilometers moved - most of these migrants will have relocated within the same province. In the

41

second, fifth and eighth columns, only individuals who have moved at least 50km are defined to be migrants. In the third, sixth and ninth columns, only individuals who have changed districts are considered migrants. In US terms, this is somewhat equivalent to defining migration as switching MSAs or counties. These three definitions are all stricter in the sense that many individuals who have changed households, and are thus considered migrants in our main results, do not qualify as migrants under them.

In almost every instance (the exception is female high school graduates between 2010 and 2012), the returns to migration under these stricter definitions are greater than those estimated under the household change definition. We are thus confident that our main results are a lower bound on the returns to migration.

A.5   Exploring returns to migration within sub-samples

Another relevant critique of any effort to measure the returns to migration is the likelihood that the effects vary by group. As we saw above, our estimates differed by education category and gender. We present estimates for two other segmentations of society that plausibly have different migration patterns or motives: direction of migration; and age of migrant.

In development economics, migration is typically thought to take place from rural to urban areas. Table 4 showed that this is not the most common form of migration in South Africa, as the majority of the moves we observe are within urban or rural areas. Rural to urban moves are disproportionately likely to be long-distance moves, however.[26] It's natural to think that migration within urban areas may be a very different phenomenon than migration from rural to urban areas, so we next examine the effects of migration by the direction of migration. Four categories exist: urban to urban; urban to rural; rural to rural; and rural to urban. The migration effect for each of these types of move for each time period are displayed in Table 16.

There are substantial and significant differences in the income effects of migration depending on the direction. As with all other calculations, the effects differ greatly depending on the time period examined. There are two particularly noteworthy points shown in Table 16. First, moving from a rural to an urban area has income effects that dwarf those from any other type of move, ranging from 61% to 80%. This suggests that we are not wrong to think of this type of move as being an economic game-changer for migrants. Second, these effects are all precisely estimated and significantly different from each other, other than urban to rural migration, which is imprecise and indistinguishable from zero. This type of migration is less frequently observed than the others and is also more likely to be associated with income losses. This is evidence that some part of urban to rural migration may consist of 'failed' migrants returning to poorer households, or perhaps adults returning home to retire. Overall, however, this cut of the data supports our result that migration is associated with large and significant increases in income.

The same story is borne out when the results are split by age of migrant (shown in Table 17). The returns to migrant vary by age, but are large and significant at 1% for all age groups. The largest returns are for young adults between 18 and 25, who experienced income gains of 48% and 53% above the mean between 2008 and 2010, and 2010 and 2012 respectively. Even children who

---

[26]As they are over-represented among moves that resulted in someone changing district council, a local government area designation.

moved had above-average increases in income. [27] The only exception is the pension-age category, for which returns to migration are small and statistically insignificant. Very few pension-age adults migrate, resulting in imprecise estimates. We can postulate that migration for older adults occurs rarely and in response to a serious income shock, bringing the average effect down.

These results support the standard positive story of migration - of young adults moving from rural to urban areas in response to expected income gains - but do not contradict alternative stories, as moves within rural or urban areas also generate large and significant gains. Both older and younger migrants also experience gains, and these results are based on household per-capita income, so they do not provide evidence on whether the gain is due to increases in personal income or to moving to a wealthier household.

---

[27]For children, the income gain between 2008 and 2012 exceeds that between 2008 and 2010 and 2010 and 2012. This is due to variable construction: age groups are created based on age in the wave prior to the move, so children in the first and third columns are the same group (and similarly for the other age categories), while the second column refers to a different group of children (those aged 18 or less in 2010, instead of in 2008). Thus, the final column measures the two-period return to migration for individuals in particular categories in 2008.

43

Table 10: Returns to migration for the migrant, with different standard error treatments

| | 2008-2010 | | | 2010-2012 | | | 2008-2012 | | |
	Abadie-Imbens	Bootstrap	Lechner	Abadie-Imbens	Bootstrap	Lechner	Abadie-Imbens	Bootstrap	Lechner
All	0.2840	0.2840	0.2884	0.3881	0.3881	0.3750	0.3089	0.3089	0.3178
	0.0493	*0.0583*	*0.0591*	*0.0378*	*0.0358*	*0.0461*	*0.0321*	*0.0314*	*0.0369*
Male, matric	0.1974	0.1974	0.1974	0.5623	0.5623	0.5623	0.5290	0.5290	0.5290
	0.1727	*0.1955*	*0.2034*	*0.1220*	*0.1159*	*0.1485*	*0.1074*	*0.1156*	*0.1292*
Male, no matric	0.4254	0.4254	0.4254	0.3757	0.3757	0.3757	0.3577	0.3577	0.3577
	0.0998	*0.1223*	*0.1178*	*0.0694*	*0.0645*	*0.0836*	*0.0564*	*0.0589*	*0.0677*
Female, matric	0.2079	0.2079	0.2079	0.7747	0.7747	0.7747	0.3042	0.3042	0.3042
	0.1586	*0.1574*	*0.1911*	*0.1352*	*0.1362*	*0.1653*	*0.0871*	*0.0900*	*0.1026*
Female,	0.2039	0.2039	0.2039	0.1991	0.1991	0.1991	0.2239	0.2239	0.2239
no matric	*0.0689*	*0.0854*	*0.0845*	*0.0729*	*0.1093*	*0.0824*	*0.0460*	*0.0484*	*0.0547*

Dependent variable is log of household per capita income. All specifications use individual and household characteristics as controls.
Matching algorithm is a local linear regression propensity score matching estimator, using a normal kernel.

Table 11: Selection ratios

| | 2008-2010 | | | 2010-2012 | | | 2008-2012 | | |
	AET	Oster $R^2=1$	Oster $R^2=0.836$	AET	Oster $R^2=1$	Oster $R^2=0.836$	AET	Oster $R^2=1$	Oster $R^2=0.836$
All	2.911	-5.076	-6.602	-43.669	-2.580	-3.255	2.492	-3.970	-5.213
Male, matric	0.476	0.235	0.318	-7.013	-0.584	-0.713	2.115	-1.661	-2.206
Male, no matric	-3.426	-1.780	-2.327	-10.634	1.898	2.499	4.343	26.084	34.465
Female, matric	0.518	-1.291	-1.657	-124.561	-0.535	-0.649	3.113	-1.380	-1.780
Female, no matric	-1.243	-1.774	-2.328	-14.433	6.980	9.033	4.987	7.005	9.426

Dependent variable is log of household per capita income. Variable of interest is migration dummy. All specifications use individual and household characteristics as controls. For the Oster (2014) results, required controls are a quartic in age, sex, marital status and location. In all cases, the difference-in-difference regressions are run on a matched sample created by matching on the entire black sample with non-missing control information for the relevant time period. Columns labeled "AET" contain the informal selection ratio proposed in Altonji, Elder and Taber (2005). Columns labeled "Oster" contain the selection ratio proposed in Oster (2014), for different assumed maximum R^2.

Table 12: Returns to migration: testing the parallel trends assumption

	Only 2010 controls		2008 and 2010 controls		Main results sample	
	HHPCI	Ind. income	HHPCI	Ind. income	HHPCI	Ind. income
All	0.4724	0.2252	0.4964	0.2342	0.3881	0.1980
	0.0413	*0.0588*	*0.0531*	*0.0799*	*0.0378*	*0.0508*
Male, matric	0.5912	0.2352	0.8038	1.3608	0.5623	0.3462
	0.1523	*0.1638*	*0.6195*	*1.4128*	*0.1220*	*0.1402*
Male, no matric	0.4792	0.2697	0.4350	0.1008	0.3757	0.2293
	0.0741	*0.1924*	*0.1732*	*0.2359*	*0.0694*	*0.1527*
Female, matric	0.6245	0.3565	0.8954	0.4488	0.7747	0.3311
	0.1161	*0.1564*	*0.1632*	*0.2296*	*0.1352*	*0.2014*
Female,	0.3524	0.1828	0.4842	0.1370	0.1991	0.0465
no matric	*0.0672*	*0.0782*	*0.1027*	*0.1476*	*0.0729*	*0.0782*

All returns are calculated using local linear regression propensity score matching on individual and household characteristics with a Gaussian kernel. For columns labeled "HHPCI", the dependent variable is the change in log of household per capita income. For columns labeled "Ind. income", the dependent variable is the change in log of individual income. The first two panels contain results from the restricted sample of individuals seen in all three waves, who migrated between 2010 and 2012 or never migrated.

Table 13: Difference-in-difference estimates of the return to migration

	2008-2010	2008-2010	2008-2012
All	0.1255	0.3699	0.2398
	0.0279	0.0237	0.0203
Male, matric	0.0336	0.5899	0.3211
	0.0860	0.0720	0.0676
Male, no matric	0.1536	0.2384	0.2058
	0.0547	0.0448	0.0387
Female, matric	0.1598	0.5650	0.3034
	0.0799	0.0631	0.0570
Female,	0.1558	0.2890	0.1935
no matric	0.0411	0.0356	0.0294

Dependent variable is log of household per capita income. Variable of interest is migration dummy. All specifications use individual and household characteristics as controls. The regressions are run on a matched sample created by matching on the entire black sample with non-missing control information for the relevant time period.

Table 14: Selection ratios - individual income

	2008-2010			2010-2012			2008-2012		
	AET	Oster R2=1	Oster R2=0.836	AET	Oster R2=1	Oster R2=0.836	AET	Oster R2=1	Oster R2=0.836
All	-0.12	0.21	0.26	0.38	-4.94	-6.01	0.16	-3.80	-4.75
Male, matric	127.28	-0.47	-0.61	1.71	-0.79	-1.01	-0.54	-0.33	-0.45
Male, no matric	-0.56	-0.57	-0.71	-2.37	-0.03	-0.03	0.02	0.03	0.04
Female, matric	-1.02	-0.09	-0.11	21.44	-0.41	-0.50	3.49	-0.27	-0.34
Female, no matric	0.38	-0.07	-0.08	0.27	-0.50	-0.60	0.43	-0.20	-0.24

Dependent variable is log of individual income. Variable of interest is migration dummy. All specifications use individual and household characteristics as controls. For the Oster (2014) results, required controls are a quartic in age, sex, marital status and location. In all cases, the difference-in-difference regressions are run on a matched sample created by matching on the entire black sample with non-missing control information for the relevant time period. Columns labeled "AET" contain the informal selection ratio proposed in Altonji, Elder and Taber (2005). Columns labeled "Oster" contain the selection ratio proposed in Oster (2014), for different assumed maximum R^2.

Table 15: Returns to migration for the migrant, with differing migration definitions

	2008-2010			2010-2012			2008-2012		
	>median	50km+	Changed DC	>median	50km+	Changed DC	>median	50km+	Changed DC
All	0.4578	0.5227	0.4157	0.5083	0.5439	0.5013	0.5046	0.5735	0.3247
	0.0671	*0.0870*	*0.0764*	*0.0523*	*0.0589*	*0.0587*	*0.0429*	*0.0508*	*0.0464*
Male, matric	0.4718	0.6776	0.4520	0.5856	0.6124	0.6145	0.7083	0.8168	0.4313
	0.1731	*0.2197*	*0.2194*	*0.1315*	*0.1367*	*0.1413*	*0.1173*	*0.1276*	*0.1406*
Male, no matric	0.4070	0.7027	0.5694	0.4807	0.5350	0.4710	0.5128	0.5834	0.3671
	0.0988	*0.1698*	*0.1453*	*0.0966*	*0.1091*	*0.1091*	*0.0776*	*0.0920*	*0.0813*
Female, matric	0.4299	0.5242	0.4584	0.6517	0.5350	0.4771	0.4215	0.5303	0.2953
	0.1625	*0.2233*	*0.1880*	*0.1361*	*0.1556*	*0.1450*	*0.1128*	*0.1304*	*0.1258*
Female, no matric	0.3330	0.4052	0.3137	0.4195	0.4927	0.4563	0.4410	0.5207	0.3284
	0.1043	*0.1376*	*0.1190*	*0.0834*	*0.0966*	*0.0966*	*0.0692*	*0.0850*	*0.0716*

Dependent variable is log of household per capita income. All specifications use individual and household characteristics as controls.
Matching algorithm is a local linear regression propensity score matching estimator, using a normal kernel.

Table 16: Returns to migration for the migrant, by direction of migration

	2008-2010	2010-2012	2008-2012
All	0.2840	0.3881	0.3089
	0.0493	*0.0378*	*0.0321*
Urban-Urban	0.1354	0.3589	0.2478
	0.0734	*0.0664*	*0.0510*
Urban-Rural	0.0808	-0.1000	-0.1493
	0.1206	*0.1063*	*0.0882*
Rural-Rural	0.3938	0.2279	0.1794
	0.0810	*0.0613*	*0.0452*
Rural-Urban	0.6078	0.7995	0.7915
	0.0908	*0.0728*	*0.0623*

Dependent variable is log of household per capita income. All specifications use individual and household characteristics as controls. Matching algorithm is a local linear regression propensity score matching estimator, using a normal kernel.

Table 17: Returns to migration for the migrant, by age of migrant

	2008-2010	2010-2012	2008-2012
All	0.2840	0.3881	0.3089
	0.0493	*0.0378*	*0.0321*
Child(¡=18)	0.3684	0.3612	0.4346
	0.1110	*0.0890*	*0.0697*
Young adult(19-25)	0.4780	0.5308	0.4813
	0.0889	*0.0700*	*0.0574*
Youth(19-35)	0.3816	0.4589	0.3812
	0.0650	*0.0502*	*0.0421*
Prime age (25-50)	0.1989	0.3308	0.1612
	0.0715	*0.0535*	*0.0448*
Working age(19-65)	0.2953	0.3993	0.2907
	0.0549	*0.0424*	*0.0345*
Pension age(65+)	0.2122	0.0580	0.0833
	0.2779	*0.1465*	*0.1130*

Dependent variable is log of household per capita income. All specifications use individual and household characteristics as controls. Matching algorithm is a local linear regression propensity score matching estimator, using a normal kernel.