Migration and gender in South Africa: following bright lights and the fortunes of others?

DIETER VON FINTEL
ELDRIDGE MOSES

Stellenbosch Economic Working Papers: WP09/2017

www.ekon.sun.ac.za/wpapers/2017/wp092017

September 2017

KEYWORDS: Regional migration, gravity model, feminization of migration, income mobility, economics of gender, South Africa
JEL: C31, J16, J61, O15, O18, R23

ReSEP (Research on Socio-Economic Policy)
http://resep.sun.ac.za

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH

SOUTH AFRICA

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

www.ekon.sun.ac.za/wpapers

Migration and gender in South Africa: following bright lights and the fortunes of others?[1]

By Dieter von Fintel[2] and Eldridge Moses[3]

ABSTRACT

Internal migration in South Africa has a strong gender dimension. Historically, the apartheid-era migrant labour system meant that predominantly black African men moved to urban areas without their families. After the abolition of influx controls in 1986, many women relocated, presumably to join their male partners. The period of migration feminization was also coupled with labour market feminization. However, existing research shows that increased female labour supply was poorly matched by labour market absorption, leading to rising unemployment among black African women. This paper studies incentives for female migration in this context, by building a gravity model of male and female inter-municipal migration. We find that neither men nor women move primarily for family reasons. Instead, they follow the traditional male migrant route to well-lit economic centres. Women also do not migrate primarily for increases in their own labour market opportunities, but tend to flock to regions where other fortunate groups have higher earnings potential. While this might signal that migrants base relocation decisions on incorrect information (and could in turn explain why many migrants have unfulfilled expectations), our results also show that women not only move for work, but for public services. The implications are twofold if migration is to alleviate poverty in the long run.: firstly, in the short run, management of public resources must improve, as poor (women) place large emphasis on their effect; and secondly, labour market barriers – especially into the informal sector – should be better understood.

[1] We gratefully acknowledge funding from the Growth and Economic Opportunities for Women (GrOW) initiative, a multi-funder partnership between the UK's Department for International Development, the Hewlett Foundation and the International Development Research Centre. We thank Stephan Klasen, Servaas van der Berg, Rulof Burger, Cobus Burger, and participants at the GrOW and StatsSA/isiBalo conferences for helpful comments in improving this paper. Opinions and errors remain those of the authors.

[2] Department of Economics, Stellenbosch University and Institute for Labor Economics (IZA), Bonn; E-mail: dieter2@sun.ac.za

[3] Department of Economics, Stellenbosch University

1 Introduction

Gender, as a focal point in the understanding of internal migration decisions, processes and outcomes, has a relatively short history dating to the early 1990s, when migration studies increasingly directed attention to the relative paucity of studies documenting female migration (Tienda and Booth, 1991; Chant and Radcliffe, 1992). Much of the literature in the early years of internal migration research in sub-Saharan Africa implicitly assumed that the majority of migrants were men. By omission, many studies resigned women either to the roles of caregivers in the sending region, or companions to men through marriage or dependence. The gap in the literature occurred despite womens' growing presence in migration flows. This restrictive lens through which the migration of women was viewed, meant that even as migration research evolved, it remained largely insensitive to the specific drivers of *autonomous* female migration.

The focus on male mobility in the early years of sub-Saharan migration analyses was also partly a function of a dearth of nationally representative data available for researchers to analyze the role of gender in migration decisions, processes and outcomes (Camlin *et al*., 2014). At least part of the reason for the 'missing women' in the South African migration literature (in particular) is also a methodological one: analyses of permanent migration generally occurred for relocation over long distances (see for instance van der Berg *et al*., 2002; Kok *et al*, 2006; Moses and Yu, 2009), with little attention being paid to the bulk of (female) population flows, which occur *within* regions (Moses, 2017). Smaller studies find that rural women tended to migrate to smaller towns, semi-rural employment hubs and the informal settlements on the peripheries of small cities (Camlin *et al*., 2014), meaning that migration analysis at the inter-regional level understated the mobility of women.

This pattern has changed, however, since the relaxation of influx controls. The new extent of longer distance female mobility in South Africa is evidenced by Wentzel *et al*.'s (2006) finding that 42 percent of black individuals had crossed municipal[1] boundaries in the five years prior to 2001, and that 51 percent of those inter-municipal migrants were female. Analysis of the 1996 Census reveals that women accounted for approximately 47 percent of migration across district council boundaries, while the comparable figure for 2011 is 46 percent.

While the gender composition of internal labour migration has remained relatively stable since 1996, there is some evidence to suggest that the *nature* of female migration in South Africa has changed substantially in recent times. The initial increase women's migration was connected to movements with family members: the family migration rate for women was 31 moves per person-year in the 1994 to 2000 period, up from 10 in the 1986 to 1993 period (Reed, 2013). However, in the two decades of democracy, profound changes in marital arrangements (possibly driven in part by the disruptive effects of migration), fertility reductions and increases in educational attainment levels may also have contributed to more *independent* decision-making and migration by women.

Gendered analysis of migration also produces seemingly confounding results – large rural-urban migration flows occur despite the fact that females from rural origins have extremely low employment probabilities

[1] South Africa has 9 provinces, composed of 234 municipalities in total. The most recent municipal boundaries are shown in figures 1,3 and 4.

in urban areas. Men perform slightly better but also face relatively poor labour market prospects (Van der Berg *et al*., 2002). Migration patterns and motivations therefore appear to be more complex than simple disequilibrium models would suggest and therefore warrant a closer investigation of the role of gender in migration motivations and outcomes.

This paper therefore documents the gender-specific incentives to relocate within the borders of South Africa. We build zero-inflated negative binomial gravity models of migration numbers across municipalities. To account for geographic spillovers we introduce spatial filters into the specifications. We distinguish whether black African men and women move in greater numbers to areas where their own group experiences a migration income premium, or whether other factors are at play. In particular (given that long distance migration was historically dominated by single men who departed from former apartheid homelands), we study whether female migration has followed the same routes for the purposes of maintaining family ties, or whether these movements are economically profitable for women who have an incentive to move independently. Further, we explore whether the incomes of other privileged groups and *overall* (not group-specific) economic conditions in destination regions are pull factors. These results disentangle whether women move based on gains that accrue to women migrants, or whether they relocate based on the decisions of others. We posit multiple channels through which such behaviour could arise.

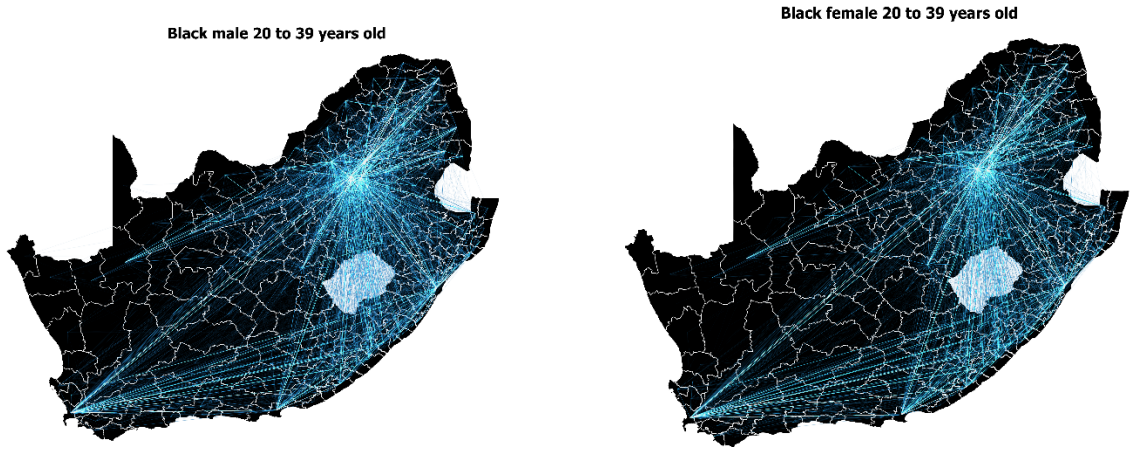
2 Literature

Historically migration and settlement in South Africa were constrained institutionally. Between assuming power in 1948 and 1991, the apartheid government devised and implemented 317 laws that governed nearly every dimension of black life, including black population movement and settlement (Chloe and Chrite, 2014: 83). The Bantu Self-Government Act (1950) relegated black individuals to homelands[2], which were created by the apartheid national government to function as independent states. Movement anywhere outside of the homelands was strictly regulated by the Pass Laws Act (1950). This act required all black individuals over the age of 16 years to carry a pass book which provided proof that the holder was employed and allowed to be in an urban area for more than 72 hours (Wilson, 2001; Gelderblom and Kok, 2006). Black Africans' migration was therefore determined by the labour needs of the (semi-)urban (white) economy, an arrangement referred to as the migrant labour system.

Migration flow maps in Figure 1 show the dominant migration corridors in South Africa. The centrality of metropolitan areas (in the Gauteng and Western Cape provinces) to inter-municipal migration flows emphasizes that the migrant labour system shaped spatial movement patterns substantially. Gauteng attracted 41 percent of inter-municipal migrants in 2011, testimony to its role as the regional centre of South African economic activity and its central location relative to the big sending regions of the former homelands.

[2] Homeland boundaries are shown in Figures 2, 4 and 5.

Figure 1 All inter-municipal migration flows of black individuals 2001 to 2011



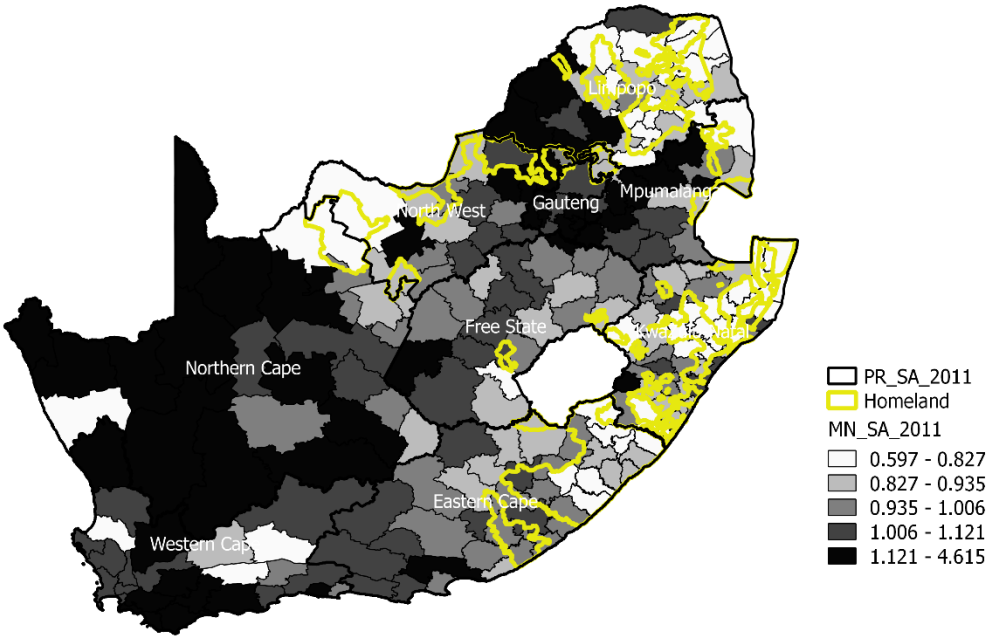
Source: Map constructed using Census 2011 data (Statistics South Africa, 2011a).

The Pass Laws also prevented the spouses or children of pass book holders from accompanying them to the urban areas they were employed in, thus firmly entrenching the historical patterns of black internal migration, settlement and progressive family disintegration previously driven by mining industry needs since the end of the 19th century. The Pass Laws, along with a number of other oppressive laws governing movement and settlement, effectively ensured that black African temporary migrants were disadvantaged relative to their permanent resident counterparts in terms of labour market opportunities, services and housing (Hindson, 1987). It also ensured that internal migration of black individuals was heavily skewed in favour of males, creating stark and persistent disparities in gender ratios between regions. These are still apparent more than two decades after the repeal of the last discriminatory laws affecting black population movement and settlement in 1991. Despite this historical pattern, Figure 1 shows that migration *patterns* (as with migration numbers) are no longer distinguishable across gender.

Figure 2, however, shows the impact of previous migration flows on regional male to female sex ratios[3]. Most municipalities in the former homelands (shown by the yellow borders) have substantially more prime-aged black women than men. In contrast, the dominant receiving municipalities in the Western Cape and Gauteng have more prime-aged men than women, indicative of heavily gendered cumulative impact of migration and settlement in South Africa's economic centres. Family migration rates for women have, however, increased in tandem with that of men (Reed, 2013). Factors constraining long-distance migration of women along with partners in the past, have seemingly abated somewhat. Figure 1 shows that migration routes of women follow those of men, pointing to possible continuance of co-migration patterns.

[3] These ratios are calculated by placing migrants in their sending municipality; in other words, they reflect demographic circumstances before the migration event (only during the one year before the census was enumerated).

Figure 2 Municipal sex ratios of black adults aged 20 to 39 years (12 months prior to Census night 2011)



Source: Map constructed using Census 2011 data (Statistics South Africa, 2011a).

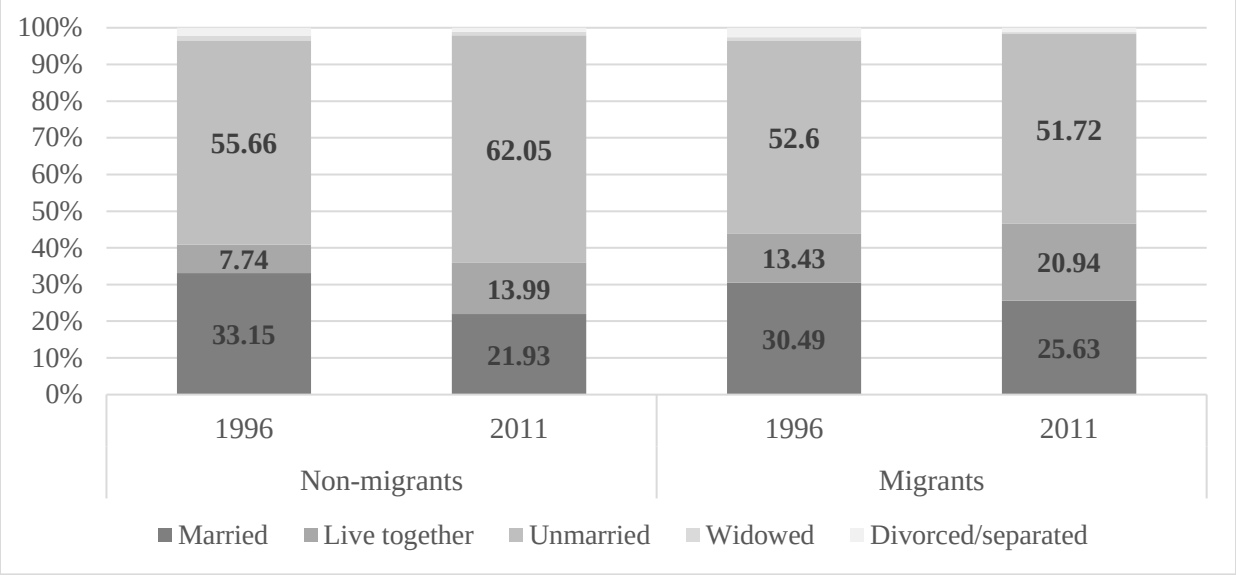
Three other possible reasons for increasing migration probabilities amongst women include the softening of attitudes towards female migration over time, changes in male-female partnership arrangements (Posel, 2004), and reductions in fertility amongst prime-aged women. Economic hardship appeared to force families to reconsider their stances on the migration of female household members, with many households not preventing and in some cases even supporting their daughters' mobility (Todes, 1998). In recent years, gendered labour allocation regimes which were previously sustained by women's economic dependence on men, household gender attitudes and social pressure, have become less relevant as women have become more educated and less likely to marry. These changes in gendered labour allocation norms may have been precipitated by the male migrant labour system, which may have contributed substantially to the breakdown of traditional family structures and gendered power relations (Spiegel, 2007; Hunter, 2001).

Figure 3 shows that the percentage of black non-migrant and migrant women between the ages of 20 and 39 years[4] who reported that they were formally married had decreased by 11 and 5 percentage points respectively. Migrant women are now more likely to be married than non-migrant women; at face value this would be consistent with increased co-migration of married couples, despite a lower prevalence of marriage over time. However, declines in formal marriage rates across all groups may have affected female migration rates positively in two other ways: (1) women not living with men in restrictive relationships would have been free to pursue personal income-maximising strategies; and (2) women not living with men might have needed to migrate to distant labour markets to compensate for the absence of male incomes.

[4] For the purpose of this paper, prime-aged women are those women aged between 20 and 39 years, who are assumed to be most likely to be entering the labour force and possible marriage markets.

Particularly large increases in co-habitation among migrants could be the result of relocation to balance the spatial divide in sex ratios; however, the less contractual nature of this living arrangement suggests that women have nevertheless migrated with a greater sense of independence than in the past.

Figure 3 Marital status of black women aged 20 to 39 years, by migrant status (1996 and 2011)



Source: Own calculations based on Censuses 1996 and 2011.

One of the more important changes affecting female internal migration patterns in the last two decades has been the increase in labour market participation rates, and by extension their increased representation in internal migration streams that were previously male-dominated. Educational attainment increases[5] and reductions in fertility were particularly large since 1996. Fertility reductions between 1996 and 2011 were particularly large amongst women between the ages of 20 and 29 years (Statistics South Africa, 2011), allowing young women to participate in labour markets more fully than was the case in the mid-1990s.

The sustained presence of women in South African migration flows, in spite of traditional norms that favour male migration probabilities, can also be explained from the "new economics of migration" perspective (Massey, 2006). Economically stretched households may engage in risk diversification strategies by sending or facilitating the migration of working-age household members to other regions. Working-age women are more likely to migrate from households where there are old-age pension-eligible women present (Ardington *et al*., 2009). The presence of older women not only provides much-needed childcare services for those women in their reproductive years (particularly when adult males are absent) but the stability of the additional pension income may also allow households the opportunity for risk diversification in the form of migration of working-age household members (Posel, 2001; Kok *et al*., 2006).

In the next section, we continue this line of enquiry. While we do not show that there have been sustained *increases* in female migration, we show that men and women have different incentives to move. Gravity models place particular emphasis in distinguishing between own (autonomous) benefits of relocation vis-

[5] The mean years of education completed by black women between the ages of 20 and 39 years (no longer attending school) increased from 7.28 years in 1996 to 10.17 years in 2011.

à-vis the role that overall economic activity and other groups (races and genders) have in determining migration flows.

3 Data and methods

3.1 Data

We study total migration flows of prime-aged (20 to 39 years old) black Africans across 234 South African municipalities, recorded for the year prior to the 2011 Census night (Statistics South Africa, 2011a)[6]. This specific population of interest is driven by two considerations: firstly, young individuals are more likely to move in search of jobs, and are in the age group where marriage is most likely[7]; secondly black African influx controls to urban areas have been rescinded, so that we are able to test whether historical migration patterns matter more than current socio-economic or familial considerations. Our analysis is also separated by gender to reflect differences introduced by the apartheid (male) migrant labour system and the rapid feminisation of the labour market more recently (Casale & Posel, 2002; Burger & von Fintel, 2015).

Each municipality is regarded as being a potential sending and destination region for migrants. However, we purposefully limit the sample to avoid capturing return migration. In doing so, we emphasise pull factors into urban areas. These are difficult to distinguish empirically from potential push factors away from rural areas (such as when migrants do not find jobs in *relatively* low - but absolutely high - unemployment receiving regions, and then return home to relatively high unemployment regions). All flows away from metropoles and towards rural municipalities are therefore discarded. We do, however, consider bi-directional flows in the case of semi-urban municipalities, as pull factors may operate across these boundaries. We implement the regional classifications of municipalities as set out in the Municipal Infrastructure Investment Framework (Palm Development Group, 2010) and adopted by a number of South African government departments. Movements of less than 100km are also discarded, as we emphasize substantive moves into other local labour markets[8].

[6] We do not study migration flows in prior years, as the data only record the *last* move of respondents. Migration in earlier years is therefore understated. Furthermore, the construction of flows from sending to receiving regions depends on knowing each migrant's previous municipality of residence. About 18% of our sample of interest does not report this information. Non-response is non-random, and is concentrated among older individuals, the less educated, those with lower incomes and the unemployed (details are available on request). Our sample is therefore biased towards those for whom moving has the highest potential economic benefit, so that coefficients on economic variables will be overstated. Many of our fuller models yield statistically insignificant coefficient estimates, so that we are confident that these effects are true zeroes.

[7] Hosegood *et. al.* (2009) report that the interquartile range of age at first marriage is 21 to 30 years for women and 27 to 35 years for men in rural KwaZulu-Natal. They also report comparable national figures. Posel & Casale (2013) assume marriageable age to be between 20 and 30 for women. We choose a broader age band in response to later marriage among men.

[8] Von Fintel (2016) shows that South Africa's 55 district councils form more appropriate labour market boundaries than the smaller 354 magisterial districts. Hence, functional labour markets are fairly large, and may cover more than one magisterial district. We place a distance restriction to ensure that we study individuals who move outside their current local labour market or social networks. The reason is to study moves based on *regional* disparities and not individual-specific preferences, which cannot be modelled by gravity analysis.

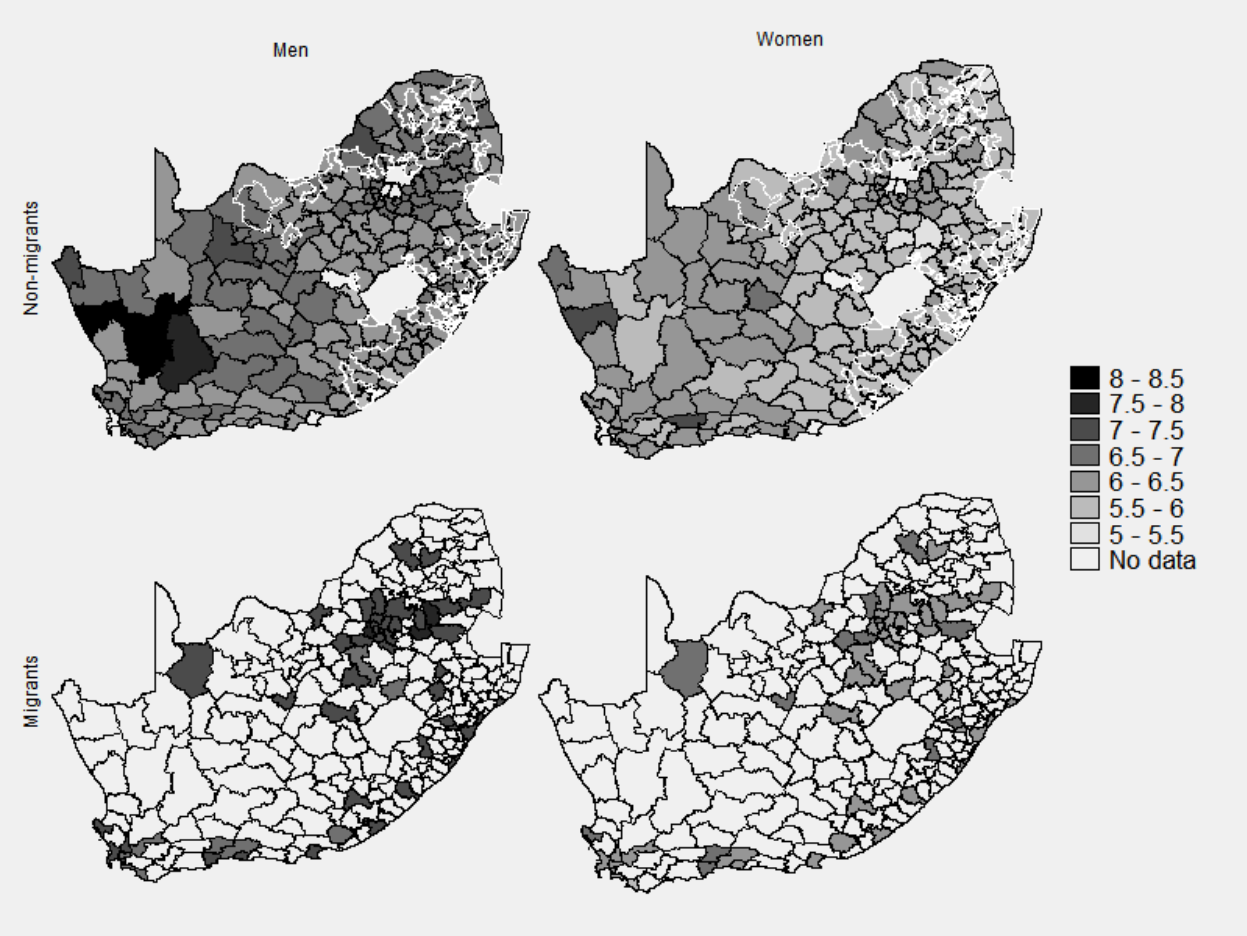
Individual income aggregates are estimated by region, using interval regressions (von Fintel, 2007)[9]. Census income data is reported in brackets, and is therefore challenging to aggregate. Interval regressions rely on the assumption of normality of *log(income)* and tobit estimation for censored data. Coefficients on region fixed effects represent the average income levels of black African individuals in each municipality, after controlling for a standard set of covariates. We construct a number of income quantities of interest using this approach: firstly, the logged ratio of the income of migrants (who moved to destination regions in the year before the migrants that we study[10]), to the income of non-migrants in sending regions; secondly, the average income of whites in destination regions. The former represents the income premium that arises between movers and stayers for each region pair, and is our primary indicator for measuring the direct benefit of relocation. Other studies do not typically make the distinction between the labour market experiences of migrants and the entire regional population in gravity models (Chun, 2008). Our use of micro data in constructing these aggregates allow us to estimate group-specific statistics. This is essential, since migrants are a self-selected group. Furthermore, our use of statistics based on past migration status partially bridges concerns of endogeneity. In other words, migration follows the incomes of previous movers in a time sequence. The opposite time progression - in that incomes of regions may rise if a select group of migrants moves there – is therefore eliminated from the analysis. Since we do not have panel data, we do not use more elaborate models to account for potential endogeneity.

In light of labour market discrimination against black Africans (Burger *et. al.*, 2016), white incomes represent one of two concepts. Firstly, it could be the aspirational income benchmark - in the absence of discrimination - of black Africans in economic centres. If black Africans have a positive probability of obtaining this benchmark, they may decide to move even if the average person from their own group does not achieve the same. Secondly, it may be a proxy for overall economic activity (and potential downstream economic activity) that is represented by the population group that has the highest incomes. Figure 4 maps the *log(income)* of non-migrants and that of previous migrants (who last moved in the year prior to our 2011 sample of migrants) in receiving regions. Men have higher incomes than women, regardless of migration status. On average, differences in incomes between migrants and non-migrants (across region pairs) are higher for men than for women. As a result, men experience a greater migration premium than women.

[9] Individuals report income data in brackets. They include incomes from all sources, including labour market earnings and cash transfers. As documented for previous censuses (Ardington *et. al.*, 2006), a large number of seemingly implausible zeroes are reported. These observations are automatically excluded from the estimation procedure after a log transformation. Our interest in regional aggregates is, however, not influenced by this sample selection problem: for the migrant sample, zero reporting rates do not differ by municipality. A regression to show this is available from the authors on request.

[10] Retrospective data allows us to classify the year of the last move to as far back as 2001. However, individuals classified as earlier migrants may not be representative of the migrant population in that year, as these individuals may have moved again and would be classified as later migrants. For this reason we only use a one year lag in constructing labour market statistics.

Figure 4 Distribution of *log(income)* of black Africans; by gender and migrant status



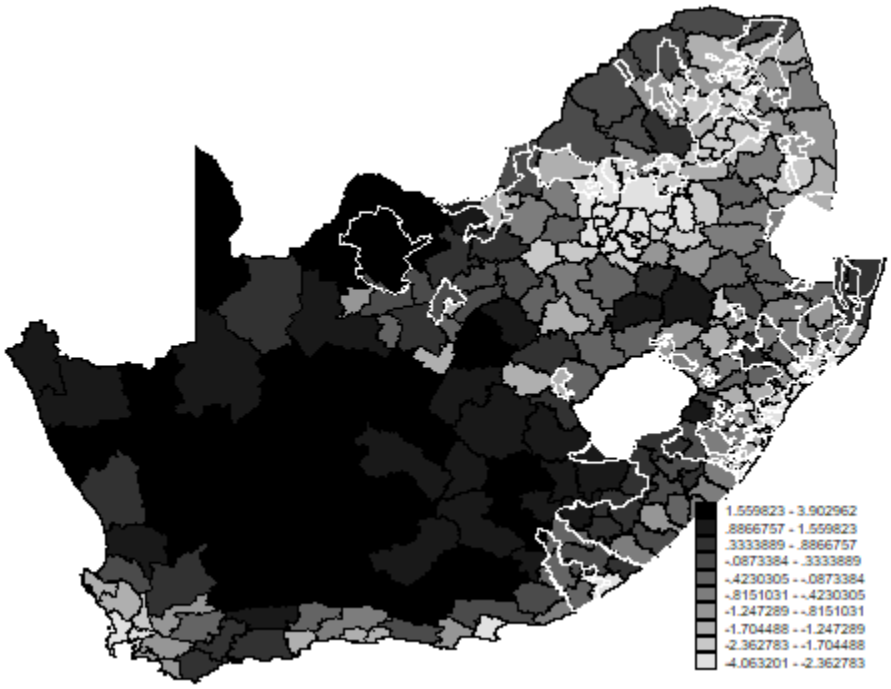
Source: Own calculations from Statistics South Africa (2011a); estimates represent the regional fixed effects from interval regressions. No data is shown for sending regions that are metropolitan. Previous migrants are located in their receiving regions. No data is shown for receiving regions that are rural.

A number of standard gravity model covariates are introduced. Distances between municipal centroids are calculated from shape files. The inclusion of distance in gravity models was originally meant to serve as a proxy for the pecuniary costs of relocation. These include transport and accommodation costs, and the cost of obtaining information about the destination region, which enjoy a positive relationship with distance. However, distance can also represent non-pecuniary costs of migration that presumably increase as proximity decreases. These include the psychological cost of separation from family, friends and familiar surroundings (Bouare, 2000: 25).

To control for regional differences in economic activity, we implement night lights data from the NOAA (2013). Henderson *et. al*. (2012) illustrate that this construct is highly correlated with national economic activity. While Jean *et. al.* (2016) show that some countries are too poorly illuminated to make useful sub-national comparisons, this source is considered sufficient to register meaningful differences in municipal economic activity in South Africa. Figure 5 shows high concentrations of night lights in metropolitan

regions and predominantly receiving migrant regions. Total logged populations (including other race and age groups) in sending and receiving regions, and unemployment ratios are estimated from Census data. As with incomes, the latter represents labour market opportunity costs, and is measured as the ratio of unemployment rates of prior migrants in destination regions to the unemployment rates of stayers in respective sending regions. Population composition is investigated in a more granular fashion also: we calculate sex ratios of non-migrant prime-aged adults in sending and receiving regions to test whether women are moving to receiving regions for the purposes of marriage or co-habitation. Finally, we construct sex ratios of past migrants: we distinguish whether both sexes follow routes that men traditionally took (due to the migrant labour system).

Figure 5 Distribution of *log(night lights luminosity)*



Source: Own calculations from NOAA (2013)

We control for other regional covariates from the Census data: regional education differences indicate whether large gaps in skills requirements of jobs pose a structural obstacle to entering a new local labour market; local marriage rates may determine whether individuals are bound by local nuclear family connections; numbers of elderly and youths in sending regions is highly correlated with receipt of social transfers in South Africa. We also consider the mediating role of social service targeting, by measuring the extent of child support grant receipt among poor households in each municipality. Data for the latter is sourced from the General Household Survey[12] (Statistics South Africa, 2011b).

[12] The poverty line is set at the food poverty line of R3864 per capita per annum (or R322 per month) in year 2000 prices, which translates to R522 per capita per month in 2010 prices. While the General Household Survey is not

3.2 Methodology

We build a standard gravity model of migration (Chan, 2008; Ramos & Surinach, 2016), where the number of migrants that relocate from sending region (municipality) *s* to receiving region *r* in the year 2011 is estimated by:

$$migration_{s;r} = \beta_0 + \beta_1 \log\left(distance_{s;r}\right) + \beta_2 \log(population_s) + \beta_3 \log(population_r) + \alpha_s' x_s + \alpha_r' x_r + \mu_s + \mu_r + \varepsilon_{s;r}$$

Standardly we expect that $\beta_1 < 0$, as longer distances are accompanied by higher relocation costs. In contrast β_2 and β_3 are positive, as the gravity theory predicts that more populated regions send and attract migrants. Our primary interest is, however, not in estimating these coefficients. We are more interested in specific sub-components of the vectors α_s and α_r, which represent sending region push factors and receiving region pull factors respectively. At times we take ratios of covariates across sending and receiving regions to find net differences. In particular, we distinguish between expected income premia that are specific to previous migrants, and regional expected income differences that accrue to the overall population (including non-migrants and individuals outside of migrants' typical social and economic networks). Our goal is to establish whether the economic benefits that actually go to migrants motivate relocation, and not overall economic activity, which may benefit others.

The specification should include both sending and receiving region fixed effects ($\mu_s; \mu_r$) to account for unobservables specific to the route. However, many characteristics in the specification do not vary over *s* and *r*, so that it is not possible to account for these without panel data. Our coefficients are therefore potentially biased and inconsistent. Even if panel data were available, exhaustive regional fixed effects may not control for omitted migration costs and benefits. For instance, individuals may choose to move to municipalities with few jobs, but cheap housing; while living in a new destination municipality, they have better access to neighbouring labour markets. Hence, higher income generating potential in *surrounding* regions is the determining pull factor, and not the (low) income generating potential in the residential receiving region. This one scenario is an example of negative spatial autocorrelation. However, other forms of spatial autocorrelation may be present in our data. We therefore estimate a spatial error model (SEM), in order to account for spatial spillovers in unobservables (Arbia, 2016):

$$migration_{s;r} = \beta_0 + \beta_1 \log\left(distance_{s;r}\right) + \beta_2 \log(population_s) + \beta_3 \log(population_r) + \alpha_s' x_s + \alpha_r' x_r + \gamma_1 W\varepsilon_s + \gamma_2 W\varepsilon_r + \varepsilon_{s;r}$$

where *W* is a spatial weighting matrix with off-diagonal elements containing the inverse distance between municipalities. This approach accounts for spatial autocorrelation, and is estimable without panel data. While maximum likelihood estimators can recover γ_i, they are computationally expensive, since the dimension of *W* is $(r^2 - r)\ x\ (s^2 - s)$. Spatial filtering presents a simpler variant that is also adaptable in multiple estimation settings (Griffith, 2003; Chun, 2008). We calculate the eigenvectors of $\left(I - \frac{1'1}{n}\right) W \left(I - \frac{1'1}{n}\right)$, where I is an identity matrix and 1 is a unit vector. Eigenvectors represent the

stratified at the municipal level, we are confident that our estimates reflect the local populations. We do not show the map of this quantity, but it is smooth across space.

structure of autocorrelation, while the sum of associated eigenvalues ($\sum_i \lambda_i$) represent the degree of autocorrelation. The gravity model simply includes eigenvectors that represent "large" degrees of autocorrelation as control variables. Our inclusion rule is $\left|\frac{\lambda_i}{\lambda_1}\right| > 0.25$, where λ_1 is the largest eigenvalue.

Our dependent variable is most appropriately modelled with count data models (Chun, 2008; Ramos & Surinach, 2016). Two issues arise. Firstly, the basic Poisson link function is inappropriate in the case of over-dispersion (where $Var(migration) > E(migration)$), so that we consider the negative binomial link function. Secondly, many of the region pairs have no migration flows between them, so that it is necessary to explore the viability of zero-inflated count models. Two stages are estimated in one maximum-likelihood routine. The first models the probability that there are zero flows, with a logit-style likelihood component. The second models the number of migrants as a count data model. The first stage corrects estimates in the second stage. Diagnostic tests can easily distinguish between the alternatives. Vuong Z-statistics allow us to detect over-dispersion; likelihood ratio tests allow us to detect the necessity of the first stage. Our use of spatial filtering in this context circumvents the need to derive a more complex likelihood function to account for spatial autocorrelation.

Table 1 Gravity model estimates of the number of male inter-municipal migrants

Dependent: Number of male migrants	1	2	3	4	5	6
Men: income ratio (receive:send)	0.147***	0.115**	0.067	0.037	0.085	0.067
	(0.052)	(0.053)	(0.054)	(0.055)	(0.061)	(0.064)
Women: income ratio (receive:send)	-0.074	-0.080	-0.015	0.000	0.056	0.134
	(0.075)	(0.075)	(0.077)	(0.079)	(0.084)	(0.087)
Men: white income (receiving)	0.275	0.163	0.579*	1.103***	1.017***	1.178***
	(0.284)	(0.285)	(0.311)	(0.345)	(0.361)	(0.432)
Women: white income (receiving)	-0.358	-0.291	-0.585**	-1.101***	-1.078***	-1.109***
	(0.247)	(0.247)	(0.256)	(0.285)	(0.294)	(0.323)
Men: unemployment ratio (receive:send)	0.521*	0.449*	0.156	0.504*	0.390	0.283
	(0.268)	(0.268)	(0.273)	(0.274)	(0.281)	(0.287)
Women: unemployment ratio (receive:send)	0.348	0.410	0.368	-0.180	-0.137	-0.156
	(0.317)	(0.316)	(0.306)	(0.316)	(0.332)	(0.346)
ln(lights ratio) - receive:send		0.095***	0.128***	0.126***	0.115***	0.125***
		(0.034)	(0.035)	(0.038)	(0.039)	(0.041)
Non-migrant sex ratio (men:women - sending)			0.086	-0.064	-0.077	-0.149
			(0.244)	(0.258)	(0.289)	(0.296)
Non-migrant sex ratio (men:women - receiving)			1.398***	1.136**	1.317**	1.290**
			(0.391)	(0.502)	(0.605)	(0.637)
Previous migrants sex ratio (men:women - sending)			0.042	0.041	0.034	0.032
			(0.054)	(0.051)	(0.052)	(0.053)
Previous migrants sex ratio (men:women - receiving)			0.128**	0.152**	0.130**	0.132*
			(0.058)	(0.065)	(0.065)	(0.068)
% poor children receive CSG (receiving)						0.935***
						(0.348)
% poor children receive CSG (sending)						-0.197
						(0.249)
Constant	-9.658***	-9.949***	-10.605***	-8.228***	-10.137***	-9.390***
	(1.197)	(1.192)	(1.212)	(1.889)	(1.986)	(2.107)
Spatial Filters, Distance and Population	Y	Y	Y	Y	Y	Y
Education				Y	Y	Y
Marriage, Children and Elderly					Y	Y
N	2180	2180	2178	2178	2172	2071
Log-likelihood	-7866.333	-7862.388	-7848.510	-7795.825	-7776.549	-7455.733
Vuong Z: ZINB vs NB	31.935	32.426	31.957	33.789	33.166	31.922
Likelihood ratio χ^2: ZINB vs ZIP	58097.994	57638.410	55889.405	47516.602	45969.750	44404.377

NOTES: Own calculations from Statistics South Africa (2011a) and other sources. *p<0.1; **p<0.05; ***p<0.01. Dependent variable is the number of male migrants between municipalities. Only rural and semi-urban to semi-urban and metropolitan flows of more than 100km are considered. Estimation follows a spatially filtered zero-inflated negative binomial regression.

Table 2 Gravity model estimates of the number of female inter-municipal migrants

Women	7	8	9	10	12	13
Men: income ratio (receive:send)	0.119**	0.072	0.049	0.006	0.111*	0.112*
	(0.054)	(0.054)	(0.057)	(0.056)	(0.061)	(0.062)
Women: income ratio (receive:send)	0.013	0.004	0.044	0.090	0.137	0.190**
	(0.076)	(0.076)	(0.079)	(0.079)	(0.084)	(0.086)
Men: white income (receiving)	0.731***	0.659***	0.581***	0.872***	0.708**	0.716*
	(0.178)	(0.175)	(0.209)	(0.231)	(0.335)	(0.372)
Women: white income (receiving)	-0.566**	-0.488**	-0.510**	-0.802***	-0.793***	-0.811***
	(0.229)	(0.228)	(0.235)	(0.246)	(0.280)	(0.298)
Men: unemployment ratio (receive:send)	-0.168	-0.278	-0.493*	-0.221	-0.412	-0.468
	(0.282)	(0.279)	(0.290)	(0.280)	(0.290)	(0.292)
Women: unemployment ratio (receive:send)	1.054***	1.176***	1.176***	0.397	0.452	0.376
	(0.332)	(0.328)	(0.324)	(0.328)	(0.339)	(0.348)
ln(lights ratio) - receive:send		0.145***	0.166***	0.137***	0.148***	0.173***
		(0.034)	(0.034)	(0.038)	(0.039)	(0.040)
Non-migrant sex ratio (men:women - sending)			0.342	-0.285	-0.478*	-0.481
			(0.258)	(0.256)	(0.290)	(0.293)
Non-migrant sex ratio (men:women - receiving)			0.579	0.714*	0.263	0.314
			(0.407)	(0.433)	(0.526)	(0.556)
Previous migrants sex ratio (men:women - sending)			-0.073	-0.028	-0.051	-0.056
			(0.055)	(0.054)	(0.055)	(0.055)
Previous migrants sex ratio (men:women - receiving)			0.087	0.134**	0.114*	0.102*
			(0.058)	(0.062)	(0.059)	(0.060)
% poor children receive CSG (receiving)						0.978***
						(0.344)
% poor children receive CSG (sending)						-0.502**
						(0.253)
Constant	-8.641***	-9.341***	-9.150***	-6.757***	-6.204**	-3.791
	(1.224)	(1.220)	(1.244)	(2.592)	(3.107)	(3.218)
Spatial Filters, Distance and Population	Y	Y	Y	Y	Y	Y
Education				Y	Y	Y
Marriage, Children and Elderly					Y	Y
N	2180	2180	2178	2178	2172	2071
Log-likelihood	-7391.091	-7381.837	-7374.953	-7304.026	-7259.454	-6965.538
Vuong Z: ZINB vs NB	36.513	36.069	36.079	34.233	34.042	32.884
Likelihood ratio χ^2: ZINB vs ZIP	49688.131	49204.242	48211.015	38247.623	36412.232	34871.027

NOTES: Own calculations from Statistics South Africa (2011a) and other sources. *p<0.1; **p<0.05; ***p<0.01. Dependent variable is the number of male migrants between municipalities. Only rural and semi-urban to semi-urban and metropolitan flows of more than 100km are considered. Estimation follows a spatially filtered zero-inflated negative binomial regression.

4 Results

4.1 Model choice

Before turning to the main results, we briefly refer to the diagnostics validating the chosen modelling strategy. Firstly, Table A1 in the online appendix shows the differences between models that include and exclude spatial filters. Parsimonious and more elaborate models are compared. At first glance, the technique does not have substantial influence on most coefficient estimates. It suggests that spatial spillovers are somewhat irrelevant, because we analyse larger movements of more than 100km. However, in certain instances, some covariates grow in magnitude and statistical significance, while they fall in others. We therefore find it necessary to correct for spatial autocorrelation; all other results in this paper therefore also include spatial filters.

In all models estimated in this paper, likelihood ratio tests conclude in favour of using negative binomial as opposed to Poisson link functions. Similarly, Vuong tests highlight the need to account for excess zeroes in all instances. Table A2 in the online appendix show the first stages, which model the probability that there was no movement along a particular migration route; results are similar across all specifications in this paper. Significant exclusion restrictions differ by gender: for men, departing from a region in the former homelands lowers the probability of zero flows. Historically, mainly male migrant workers moved from former homeland regions, so that this pattern is sensible. Regions with higher gross value added from mining are more likely to attract zero women migrants. Typically men migrated towards mining regions.

4.2 Gravity models

Our modelling strategy distinguishes between the push and pull effects that affect migrants only, and those that are relevant for the broader local origin and destination labour markets. Our goal is to establish whether relocation decisions are based on information about the broader economy, or whether migrants also take into account that their group may have a different experience to the rest of the labour market which they are entering. These concepts may be disconnected, with aggregate economic activity not necessarily translating to household income potential of migrants. This is especially true if migrants experience barriers to entering the destination labour market successfully.

Starting with basic specifications in columns 1 in Table 1 (for men) and 7 in Table 2 (for women), we first focus on economic benefits that are specific to migrants. Income growth[13] that results from *male* migration is associated with larger number of relocations of *both* genders. To the contrary, there is no statistically relevant relationship between migration flows and women's migration income premia. Women also move in greater numbers to regions where unemployment rates of *male* migrants are relatively lower (compared to stayers in their sending regions), though this effect is not statistically significant. At first glance, it appears that both black African men and women move to areas where *men* have greater labour market fortunes. We explore two potential explanations for this finding. Firstly, higher male migration premia may only be a placeholder for better overall economic conditions, and the result may not be a causal effect. Secondly, it is possible that women move for the economic benefits that accrue to male partners, and not to them directly.

[13] We do not refer to growth as within-person increases in incomes after relocation. We do not have longitudinal data to do so. Instead, we consider the difference between previous migrants' incomes and those of individuals who stayed behind in sending regions for each region pair.

They therefore potentially move for family and marital reasons, and not own economic benefits. Feminisation of the labour market, together with growth in women-headed households (Casale & Posel, 2002) discredit the latter explanation in favour of the first. Additional control variables in later specifications will disentangle this effect empirically.

Black African women tend to move to regions where the incomes of white men are high. In later specifications, this finding also appears for black African men. Again, this variable could serve as a proxy for *overall* economic activity, as white men have the highest incomes and are the largest employment creators in South Africa. However, an alternative explanation emerges, in the form of a misplaced reference group. Migrants appear to have high expectations for social mobility; they base their relocation decisions on the labour market fortunes of privileged groups, and not primarily on their own. Migrants' reference groups therefore potentially cross racial boundaries; this contrasts with the general population, whose reference group is people who live close by and are a member of the same race (Kingdon & Knight, 2007). It is possible that black African migrants move based on information of income generating potential which generally does not accrue to them, so that this cross-racial reference group is over-optimistic. Such a scenario would also explain why migrants generally face higher unfulfilled expectations and lower subjective well-being (Mulcahy & Kollamparambil, 2016). Alternatively, some black African migrants believe that they have some positive probability of attaining to average white incomes, even if the average black African does not achieve this.

Up to this point, most results had multiple, ambiguous explanations. We build more elaborate models to distinguish between the influence of individuals' income potential and overall economic activity. In columns 2 and 8 we introduce regional differences in night lights luminosity as a control. Notably, if the receiving regions' night lights luminosity exceeds that of the sending region, more individuals are likely to migrate between those areas. The relationship is intuitive, and true in most specifications for both genders.

However, we are more interested in how this control affects other coefficients. For women, the coefficient on the own race male migration income premium reduces to insignificance, while it becomes smaller for men. At this point it is possible to distinguish between the two hypotheses for womens' migration behaviour: we find affirmative evidence that black African mens' incomes proxy for overall economic conditions in destination regions, and not that women simply follow male partners.

Nevertheless, women do not base their decisions on *realized* migration premia of their own. These premia are not a robust determinant of the number of female migrants. Additionally, despite controlling for general economic activity, white male incomes remain a motivating factor. For black African women, therefore, this asymmetry could represent a misaligned income reference group and provides one explanation for why internal migrants in South Africa tend to have unfulfilled expectations (Mulcahy & Kollamparambil, 2016). Black African women therefore follow "bright lights" and the incomes of other groups: while the same income benefits do not accrue to this group, it is nevertheless possible that women move for better local amenities that are associated with the *income* fortunes of other groups. From a *labour market* perspective, movements away from poor regions can – in many instances – be connected to "failed migration". Burger & von Fintel (2014) show that labour force participation increased among younger black African women, but without accompanying absorption into jobs. Our results complement this evidence: women who move to more prosperous regions do not - in many instances – have higher incomes than those who remained in their sending regions. By inference, they do not achieve income mobility through spatial mobility. They

may, however, experience mobility in terms of public service provision. We investigate this possibility below.

Columns 3 and 9 introduce various regional sex ratios to understand whether the gender spatial imbalance (see Figure 2) influences relocation. In other words, migration is potentially not only a balancing force to equilibrate regional economic differences, but also to facilitate marriage. We have partially addressed this hypothesis for women above, and conclude that they are not primarily migrating in response to partner incentives. Not surprisingly, therefore, none of the sex ratios are significant in the female equations. For black African men, however, two patterns do emerge. They are significantly more likely to migrate to regions where recent migrants are predominantly male, and also where non-migrants are predominantly male. Migration does not equilibrate a spatial gender disparity; instead, men move to local labour markets where many men already live (and, in some cases, work). While our findings supports other research - which show that sex ratios are less influential for marriage than actual economic potential (Posel & Casale, 2013) - they are also consistent with the notion that male migrant routes have a strong historical persistence. Men still move along the same routes that were entrenched by the apartheid-era migrant labour system.

The coefficient on own income premium for black African men becomes insignificant when sex ratios are introduced in column 3. Additionally, the incomes of white men now have a similar reference group effect as it does for black African women. It is not primarily *new* income generation opportunities that induce movement; instead, men follow old migration patterns and move to areas where privileged groups prosper, but where benefits do not necessary accrue to black in-migrants. Hence, men move with the expectation of some positive probability of sharing in the gains of others, even if the average person in their group does not attain to these benefits.

Successive specifications introduce further controls to test the robustness of our results. In columns 4 and 10 we control for differences in average education levels across regions; this is to capture potential structural hindrances that could prevent the unskilled from moving to regions where jobs have a high skill content requirement. Structural unemployment exists across sectors in South Africa (Bhorat & Hodge, 1999), but also has a spatial dimension. Nevertheless, our core results remain unchanged: even if such obstacles existed and prevent labour market access, both men and women remain attracted to well-lit areas and follow historical migration routes. In columns 5 and 11 we control for sending region marriage rates and logged number of dependents in sending and destination regions. Again, our results remain robust. Full specifications in Table A1 of the appendix additionally control for concurrent migration of the opposite sex. Conclusions remain stable. It adds further evidence to suggest that migration is not primarily driven by marriage and familial considerations.

Up to now our one robust result is that migrants move towards well-lit regions, but without decent own-income benefits. However, this finding does not accommodate the possibility that people do not only relocate for incomes and jobs, but may do so for improvements in publicly provided benefits. In columns 6 and 12 we include indicators for the effectiveness of cash transfer targeting. The variable measures the regional proportion of children in poor households who receive the widespread Child Support Grant (CSG). Findings show that regions with good targeting attract more migrants, while regions with poor targeting experience out-migration. Fuller specifications in Table A1 in the appendix show that this result is only

robust for women. While individuals – especially women – move for the purposes of public services, it does not change that their dominant motivating factor is the extent of *general* economic activity[14].

Our findings for men and (more so) for women appear to be consistent with the equilibrium approach to migration modeling, popularised by Graves (1980, 1983), which proposes that real wage differentials between regions are also partly reflective of spatial differences in non-economic factors. Graves (1980) initial focus on regional differences in natural amenities such as temperature and climate was supplemented by later work that emphasized the roles of regional differences in public services, and social and cultural amenities in explaining migration flows (Gyourko & Tracey, 1991; Glaeser et al., 2001; Shapiro, 2006). The empirical evidence in favour of the equilibrium approach is particularly strong in the United States where natural amenities and urban amenities have both been found to be significant predictors of migration patterns and the perpetuation of labour market and other economic disparities between rural and urban regions (Partridge, 2010), rather than the eventual equalization of wages predicted by the disequilibrium model that dominated migration analysis until the late 1980s.

5 Conclusions

Following the abolition of apartheid-era influx controls in 1986, both migration flows and labour market participation underwent a feminization process. Previously, movement to urban centres was dominated by black African men. Our results show that these male migration routes persist, and that women have subsequently followed the same paths. While it is possible that this pattern occurred to re-unify existing nuclear families, our models show that by 2011 women were not primarily moving in response to spatially imbalanced sex ratios. Neither did they move to regions in response to increases in women's incomes. Instead, women relocate to regions with brighter lights and higher incomes of cross-racial references groups. These patterns mirror previous results, which suggest that women from younger generations rapidly entered the labour market (presumably in well-lit urban regions), but did not experience matching absorption into jobs (Burger & von Fintel, 2014). If women no longer primarily move for family reasons, and also do not move in response to labour market benefits, the question is what motivates long distance relocation? Such a pattern is potentially consistent with migration based on information from a misplaced, more fortunate reference group, and that is associated with unfulfilled expectations of migrants (Mulcahy & Kollamparambil, 2016). However, women do not only move for labour market opportunities, but relocate to areas with better public benefits. Therefore, women in poor regions do not only seek livelihoods from the job market – which does not serve them as well as men – but they also follow service provision in areas with brighter night lights.

Patterns of urbanization in Africa will therefore not only depend on the growth of formal and informal sector job opportunities, but also on the extent to which service provision in rural areas continue to lag those in urban areas. In the context of frequent service delivery protests (Bedasso & Obikili, 2016), South Africa's

[14]This final modification to the model does, however, revive the significance of own incomes for women. Nevertheless, this finding is not robust, as the fullest specification in Table A1 of the appendix highlights. Controlling for past migration of both genders renders these coefficients insignificant.

challenge is two-fold: increasing emphasis by the poor on public provision necessitates improved governance and management of public resources, in order to provide services where the labour market does not offer sufficient opportunity; more pertinently, however, the barriers that prevent urban labour market access in the long run must be adequately addressed. Entry into the informal sector is uncharacteristically difficult in South Africa (Kingdon & Knight, 2004). Fewer restrictions on this part of the economy could improve labour market attachment, and – in the long run – reduce the reliance on public provision. Such changes would, however, likely have the effect of mass urbanization, given the enduring socio-economic spatial imbalances in South Africa. Whether cities are prepared for such an influx is doubtful.

6 References

Anjomani, A. 2002. Regional growth and interstate migration. *Socio-Economic Planning Science*, 36 (4): 239–265.

Arbia, G. 2014. *A primer for Spatial Econometrics*. Basingstoke: Palgrave Macmillan.

Ardington, C., Lam, D., Leibbrandt, M. & Welch, M. 2006. The sensitivity to key data imputations of recent estimates of income poverty and inequality in South Africa. *Economic Modelling*, 23: 822-835.

Ardington C., Case, A. & Hosegood V. 2009. Labor supply responses to large social transfers: longitudinal evidence from South Africa. *American Economic Journal: Applied Economics*,1(1): 22–48.

Argent, M., Jones, R. and Holmes, J. 2014. The Amenity Principle, Internal Migration, and Rural Development in Australia. *Annals of the Association of American Geographers*, 104 (2): 305 – 318.

Bedasso, B. and Obikili, N. 2016. A Dream Deferred: The Microfoundations of Direct Political Action in Pre- and Post-democratisation South Africa. *The Journal of Development Studies*, 52(1):130-146

Bouare, O. 2002. Determinants of internal migration in South Africa. *South African Journal of Demography*, 8 (1): 23–28.

Bozzoli B. 1991. *Women of Phokeng: Consciousness, Life Strategy, and Migrancy in South Africa, 1900–1983*. Heinemann: Portsmouth, NH.

Bunea, D. 2012. Modern Gravity Models of Internal Migration: The Case of Romania. *Theoretical and Applied Economics* Volume XIX, 4(569): 127–144.

Burger, R.P. and von Fintel, D.P. 2014. Rising unemployment in a growing economy: A business cycle, generational and life cycle perspective of post-transition South Africa's labour market. *Studies in Economics and Econometrics*, 38(1):35-64.

Burger, R.P., Jafta, R. and von Fintel, D.P. 2016. *Affirmative action policies and the evolution of post-apartheid South Africa's racial wage gap*. 2016/66. Helsinki: UNU-WIDER.

Camlin, C., Snow, R. and Hosegood, V. 2014. Gendered Patterns of Migration in Rural South Africa. *Population, Space and Place* **20***:* 528–551.

Casale D. & Posel D. 2002. The continued feminisation of the labour force in South Africa: an analysis of recent data and trends. *South African Journal of Economics*,70 (1): 156–184.

Chant, S. and Radcliffe, S. 1992. Migration and development: the importance of gender. In *Gender and Migration in Developing Countries.* Chant, S. (ed). Belhaven Press: New York.

Choe, C. & Chrite, E. L. 2014. Internal Migration of Blacks in South Africa: An Application of the Roy Model. *South African Journal of Economics*, 82: 81–98.

Chun, Y. 2008. Modeling network autocorrelation within migration flows by eigenvector spatial filtering. *Journal of Geographic Systems*, 10:317-344.

Glaeser, E.., Kolko, J. & Saiz, A. 2001. Consumer city. *Journal of Economic Geography*, 1: 27–50.

Graves, P. 1980. Migration and climate. *Journal of Regional Science*, 20: 227–237.

Graves, P. 1983. Migration with a composite amenity: the role of rents. *Journal of Regional Science*, 23: 541–546.

Greenwood M.J. 1985. Human migration: Theory, models and empirical studies. *Journal of Regional Science*, 25: 521–544.

Griffith D.A. 2003. *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization*. Springer: Berlin.

Gyourko, J. and Tracy, J. 1991. The structure of local public finance and the quality of life, *Journal of Political Economy*, 99: 774 – 806).

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. Measuring Economic Growth from Outer Space. *American Economic Review,* 102(2): 994-1028.

Hindson, D. 1987. Orderly Urbanisation and Influx Control: from Territorial Apartheid to Regional Spatial Ordering in South Africa, in Tomilson, R., and Addleson, M. (eds). Regional Restructuring Under Apartheid: Urban and Regional Policies in Contemporary South Africa, Ravan Press, Johannesburg.

Hosegood, V., McGrath, N. and Moultrie, T. 2009. Dispensing with marriage: Marital and partnership trends in rural KwaZulu-Natal, South Africa 2000-2006. *Demographic Research*, 20: 279-312.hose

Hunter M. 2007. The changing political economy of sex in South Africa: the significance of unemployment and inequalities to the scale of the AIDS pandemic. *Social Science & Medicine*, 64 (3): 689–700.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., and Ermon, S. 2016. Combining satellite imagery and machine learning to predict poverty. *Science*, 353 (6301): 790–794.

Kingdon, G. and Knight, J., 2004. Unemployment in South Africa: The Nature of the Beast. *World Development*, 32(3): 391-408.

Kingdon, G. and Knight, J. 2007. Community, comparisons and subjective well-being in a divided society. *Journal of Economic Behavior and Organization*, 64(1): 29–90.

Kok, P., Gelderblom, D., Oucho, J. and van Zyl J. 2006. *Migration in South and Southern Africa: Dynamics and Determinants*. Human Sciences Research Council: Cape Town.

Massey, DS. 2006. Patterns and processes of international migration in the twenty-first century: Lessons for South Africa. In *Africa on the Move: African Migration and Urbanisation in Comparative Perspective.* Tienda, M., Findley, S., Tollman, S. & Preston-Whyte, E. (eds). Johannesburg: University of Witwatersrand Press.

Moses, E. and Yu, D. 2009. *Migration from the Northern Cape.* A Southern Africa Labour and Development Research Unit Working Paper Number 32. Cape Town: SALDRU, University of Cape Town.

Mulcahy, K. and Kollamparambil, U. 2016. The Impact of Rural-Urban Migration on Subjective Well-Being in South Africa. *The Journal of Development Studies*, 52(9):1357-1371.

National Oceanic and Atmospheric Administration (NOAA). 2013. *Version 4 DMSP-OLS Nighttime Lights Time Series*. Available [Online]: https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

Palmer Development Group. 2010. *The Municipal Infrastructure Investment Framework (MIIF): An Assessment of Investment Requirements for Municipal Infrastructure and Implications of such Investment, Round 7 Review – 2009/10.* Unpublished report prepared for the Department of Cooperative Governance and Traditional Affairs and Development Bank of Southern Africa, May 2010.

Partridge, M. D. (2010) The dueling models: NEG vs. amenity migration in explaining US engines of growth. *Papers in Regional Science*, 89, 513_536.

Posel, D. 2001. Intra-family transfers and income pooling: a study of remittances in Kwazulu-Natal. *South African Journal of Economics*, 69 (3): 501 – 528.

Posel, D. 2004. Have Migration Patterns in post-Apartheid South Africa Changed? *Journal of Interdisciplinary Economics Special Issue*, 15 (3–4): 277 – 292.

Posel, D. & Casale D. 2003. What has been happening to internal labour migration in South Africa, 1993–1999? *South African Journal of Economics*, 71 (3): 455– 479.

Posel, D. and Casale, D. 2013. The relationship between sex ratios and marriage rates in South Africa, *Applied Economics*, 45 (5): 663-676.

Ramos, R. and Surinach, J. 2016. A gravity model of migration between the END and the EU. *Tijdschrift Voor Economische en Sociale Geografie*, 108(1): 21-35.

Reed, H. 2013. Moving across Boundaries: Migration in South Africa, 1950 – 2000. *Demography*,50 (1): 71 – 95.

Shapiro, J. 2006. Smart cities: quality of life, productivity, and the growth effects of human capital. *The Review of Economics and Statistics*, 88: 324–335.

Spiegel A. 1991. Polygyny as myth: towards understanding extramarital relations in Lesotho. *African Studies*, 50 (1): 145 – 166.

Stark O. 1991. *The Migration of Labour*. Basil Blackwell: Cambridge.

Statistics South Africa 2011a. Census 2011. Available [Online]: http://interactive.Statistics South Africa.gov.za:8282/webview/

Statistics South Africa 2011b. General Household Survey. Available [Online]: http://interactive.Statistics South Africa.gov.za:8282/webview/

Statistics South Africa. 2015. *Census 2011: Fertility in South Africa.* Pretoria: Statistics South Africa.

Tienda, M. and Booth, K. 1991. Gender, migration and social change. *International Sociology* **6** (1): 51–72.

Todes. A. 1998. Gender, place, migration and regional policy in South Africa. In *Changing Gender Relations in Southern Africa: Issues of Urban Life*, Larsson, A., Mapetla, M., and Schlyter, A. (eds). Lesotho: The Institute of Southern African Studies and the National University of Lesotho.

Van der Berg, S., Burger, R., Leibbrandt, M. & Mlatsheni C. 2002. *Migration and the changing rural–urban interface in South Africa: what can we learn from census and survey data?* Presented at Labour Markets and Poverty in South Africa, 22–24 October, Johannesburg, South Africa.

Von Fintel, D.P. 2007. Dealing with earnings bracket responses in household surveys – how sharp are midpoint imputations? *South African Journal of Economics*, 75 (2): 293-312.

Von Fintel, D.P. 2016. *Wage flexibility in a high unemployment regime: spatial heterogeneity and the size of local labour markets*, Stellenbosch Economic Working Paper 09/16.

Wadycki, W.J. 1974. Alternative opportunities and interstate migration: Some additional results. *Review of Economics and Statistics*, 56: 254 – 257. Cited in Bouare, O. 2000. Determinants of internal migration in South Africa. *South African Journal of Demography*, 8 (1): 23 – 28.

Wentzel M., Viljoen, J., and Kok P. 2006. Contemporary South African migration patterns and intentions. *In Migration in South and Southern Africa: Dynamics and Determinants*, Kok, P., Gelderblom, D., Oucho, J., and van Zyl, J. (eds). Human Sciences Research Council: Cape Town.

1 Appendix

Sample	Men				Women			
ln(Distance)	-0.547	-0.852	-0.153	-0.192	-0.554	-0.954	-0.183	-0.271
	(0.038)***	(0.051)***	(0.043)***	(0.049)***	(0.039)***	(0.051)***	(0.042)***	(0.047)***
Men: income ratio (receive:send)	0.046	0.067	-0.039	-0.023	0.04	0.049	0.025	0.006
	(0.052)	(0.054)	(0.058)	(0.064)	(0.053)	(0.057)	(0.052)	(0.059)
Women: income ratio (receive:send)	0.063	-0.015	0.098	0.058	0.13	0.044	0.074	0.174
	(0.072)	(0.077)	(0.076)	(0.090)	(0.074)*	(0.079)	(0.073)	(0.082)**
Men: white income (receiving)	-0.235	0.156	0.233	0.475	-0.797	-0.493	-0.182	0.036
	(0.221)	(0.273)	(0.236)	(0.300)	(0.239)***	(0.290)*	(0.227)	(0.274)
Women: white income (receiving)	0.135	0.368	-0.253	-0.445	0.79	1.176	0.122	-0.302
	(0.275)	(0.306)	(0.344)	(0.394)	(0.289)***	(0.324)***	(0.298)	(0.340)
ln(population receiving)	0.539	0.491	0.127	0.134	0.519	0.472	0.152	0.175
	(0.044)***	(0.045)***	(0.046)***	(0.052)***	(0.044)***	(0.047)***	(0.048)***	(0.056)***
ln(population sending)	0.629	0.631	0.167	0.102	0.594	0.6	0.382	0.382
	(0.031)***	(0.035)***	(0.116)	(0.171)	(0.033)***	(0.036)***	(0.116)***	(0.157)**
ln(lights ratio) - receive:send	0.186	0.128	0.016	0.024	0.216	0.166	0.036	0.015
	(0.028)***	(0.035)***	(0.029)	(0.038)	(0.028)***	(0.034)***	(0.026)	(0.035)
Men: unemploy ratio (receive:send)	0.503	0.579	-0.068	-0.115	0.674	0.581	-0.028	-0.09
	(0.302)*	(0.311)*	(0.360)	(0.384)	(0.195)***	(0.209)***	(0.317)	(0.330)
Women: unempl. ratio (receive:send)	-0.456	-0.585	-0.099	-0.242	-0.510	-0.510	-0.181	-0.057
	(0.240)*	(0.256)**	(0.271)	(0.318)	(0.208)**	(0.235)**	(0.251)	(0.289)
Constant	-8.478	-10.605	-3.666	-4.392	-7.137	-9.150	1.060	-1.240
	(0.927)***	(1.212)***	(1.595)**	(2.059)**	(0.952)***	(1.244)***	(2.312)	(3.011)
Spatial Filters	N	Y	N	Y	N	Y	N	Y
Other controls	N	N	Y	Y	N	N	Y	Y
N	2178	2178	1708	1708	2178	2178	1708	1708
Log-likelihood	-7966.595	-7848.51	-5926.669	-5887.444	-7492.063	-7374.953	-5488.889	-5444.577
Vuong Z: ZINB vs NB	32.386	31.957	27.752	28.084	36.46	36.079	30.797	31.851
Likelihood ratio χ^2: ZINB vs ZIP	69363.712	55889.405	14836.403	14263.634	60361.707	48211.015	9961.842	9391.529

NOTES: * p<0.1, ** p<0.05, *** p<0.01. Own calculations. Robust standard errors in parentheses. Other controls include regional average education, regional dates of marriage, regional numbers of elderly and youth, regional cash transfer targeting and past migration patterns. The first stage is shown in table A2.

Table A1 Second stage zero-inflated binomial regression gravity model estimates: with and without spatial filters

Sample	Men				Women			
ln(Distance)	0.757	1.069	0.877	1.122	0.765	0.96	0.905	1.078
	(0.088)***	(0.114)***	(0.101)***	(0.129)***	(0.086)***	(0.109)***	(0.099)***	(0.125)***
Men: income ratio (receive:send)	-0.093	-0.055	-0.052	0.065	-0.308	-0.217	-0.299	-0.123
	(0.109)	(0.126)	(0.125)	(0.149)	(0.106)***	(0.121)*	(0.122)**	(0.144)
Women: income ratio (receive:send)	-0.125	-0.014	-0.156	-0.158	-0.032	-0.145	0.043	-0.282
	(0.149)	(0.178)	(0.174)	(0.219)	(0.145)	(0.172)	(0.168)	(0.211)
Men: white income (receiving)	0.021	-0.515	0.398	-0.251	0.396	0.258	0.553	0.743
	(0.474)	(0.653)	(0.523)	(0.729)	(0.455)	(0.621)	(0.506)	(0.699)
Women: white income (receiving)	-0.874	-0.083	-1.378	-0.498	-0.577	-0.084	-0.716	-0.094
	(0.536)	(0.674)	(0.597)**	(0.755)	(0.518)	(0.642)	(0.574)	(0.720)
ln(population receiving)	-0.578	-0.451	-0.641	-0.453	-0.637	-0.579	-0.695	-0.623
	(0.115)***	(0.127)***	(0.131)***	(0.145)***	(0.111)***	(0.123)***	(0.128)***	(0.142)***
ln(population sending)	-0.628	-0.825	-0.639	-0.854	-0.59	-0.775	-0.62	-0.811
	(0.072)***	(0.087)***	(0.085)***	(0.103)***	(0.071)***	(0.084)***	(0.083)***	(0.099)***
ln(lights ration) - receive:send	-0.216	-0.131	-0.188	-0.042	-0.211	-0.04	-0.228	-0.04
	(0.054)***	(0.074)*	(0.064)***	(0.088)	(0.053)***	(0.071)	(0.062)***	(0.086)
Men: unemployment ratio (receive:send)	0.654	0.084	1.038	0.446	-0.455	-0.863	0.195	-0.016
	(0.474)	(0.529)	(0.548)*	(0.627)	(0.418)	(0.455)*	(0.541)	(0.606)
Women: unemployment ratio (receive:send)	-0.662	0.030	-0.826	-0.244	0.168	0.588	0.018	0.093
	(0.480)	(0.560)	(0.527)	(0.624)	(0.463)	(0.539)	(0.521)	(0.605)
Proportion of sending in former homeland	-0.508	-0.481	-0.651	-0.625	-0.187	0.193	-0.261	0.187
	(0.177)***	(0.237)**	(0.207)***	(0.299)**	(0.167)	(0.222)	(0.196)	(0.279)
Metropolitan area (receiving)	-0.073	-0.452	0.074	-0.444	-0.164	-0.263	-0.023	-0.286
	(0.297)	(0.359)	(0.331)	(0.415)	(0.289)	(0.349)	(0.324)	(0.397)
Mining GVA (receiving)	0.027	-0.025	-0.005	-0.079	0.075	0.052	0.082	0.028
	(0.038)	(0.058)	(0.043)	(0.069)	(0.037)**	(0.056)	(0.042)*	(0.066)
Constant	15.69	13.347	16.38	14.254	12.21	10.595	11.937	7.985
	(2.491)***	(3.235)***	(2.806)***	(3.662)***	(2.415)***	(3.112)***	(2.704)***	(3.492)**
Spatial Filters	N	Y	N	Y	N	Y	N	Y
Other controls (second stage)	N	N	Y	Y	N	N	Y	Y
N	2178	2178	1708	1708	2178	2178	1708	1708
Log-likelihood	-7966.595	-7848.51	-5926.669	-5887.444	-7492.063	-7374.953	-5488.889	-5444.577
Vuong Z: ZINB vs NB	32.386	31.957	27.752	28.084	36.46	36.079	30.797	31.851
Likelihood ratio χ^2: ZINB vs ZIP	69363.712	55889.405	14836.403	14263.634	60361.707	48211.015	9961.842	9391.529

NOTES: * p<0.1, ** p<0.05, *** p<0.01. Own calculations. Robust standard errors in parentheses. The second stage is shown in table A1.

Table A2 First stage zero-inflated negative binomial regressions: with and without spatial filters