

## ***Social Networks, Employment and Worker Discouragement: Evidence from South Africa***

Justine Burns, Susan Godlonton and Malcolm Keswell

### **Abstract**

Social networks are increasingly being recognized as having an important influence on labour market outcomes, since they facilitate the exchange of job related information. Access to information about job opportunities as well as perceptions about the buoyancy of the labour market depend critically on the social structures and the social networks to which labour market participants belong. In this paper, we examine the impact of information externalities generated through network membership on labour market status. Using Census data from South Africa, a country characterized by high levels of unemployment and worker discouragement, we adopt an econometric approach that aims to minimise the problems of omitted variable bias that have plagued many previous studies of the impact of social networks. Our results suggest that social networks may enhance employment probabilities by an additional 3 to 12%, and that failure to adequately control for omitted variables would lead to substantial over-estimates of the network co-efficient. In contrast, the impact of social networks on reducing worker discouragement is much smaller, at between 1-2%.

Keywords: Unemployment, Job-search, Discouraged Worker, South Africa, Social Networks  
JEL Classification: J64, Z13.

## 1. Introduction

Networks matter for economic and social outcomes, since they allow for complex social interactions to take place between individuals, thereby facilitating information spillovers and learning between network members, as well as the transmission of norms and values (Banerjee, 1992; Bikhchandani, Hirshleifer and Welch, 1992). While information spillovers have been widely examined by economists, as affecting fertility decisions, education decisions (Coleman et al, 1966), participation in criminal activities, (Besley and Coate, 1992; Borjas, 1992, 1994, 1995, Case and Katz, 1991), and consumption (Abel, 1990), in this paper we focus primarily on information externalities generated through network membership as they pertain to employment prospects.

Arguably, it is now a stylized fact that individuals rely on friends, family and acquaintances (or “weak ties” in the parlance of Granovetter, 1974<sup>1</sup>) to find jobs, and while reliance on these social networks may vary by location and demographic characteristics<sup>2</sup>, these channels are usually productive (Calvo-Armegnol and Jackson, 2004; Blau and Robbins, 1990; Montgomery, 1991). Moreover, reliance on networks to find employment typically results in better matching between the job candidate and the available job, reduces employer uncertainty about worker productivity, and may even enhance job satisfaction and employee loyalty (Datcher, 1983; Devine and Kiefer, 1991; Marsden and Gorman, 2001). Obviously, this is conditional on the “quality” of one’s network. When an individual’s contacts are unemployed, the likelihood of getting information about jobs through contacts is reduced, thereby impeding such network benefits, reducing active search behaviour (Devine and Kiefer, 1991; Kingdon and Knight, 2001) and increasing the duration of unemployment. Thus, heterogeneity in network effects is important since it can explain changes in wages and employment inequality over time. Arrow and Borzekowski (2004) show that networks explain 15% of the unexplained variation in wages and a substantial portion of the disparity between black and white income distributions. This is exacerbated if one allows for the fact that social networks are largely endogenous or self-selected.

Within the context of very high levels of unemployment and worker discouragement in South Africa (see Table 1) combined with relatively low absorption of the unemployed into the informal sector (Kingdon and Knight, 2001), understanding the effect of networks on employment prospects is arguably critical, especially since job search methods in South Africa are predominantly passive, with most jobs being obtained through word-of-mouth and other informal recruitment methods (Kingdon and Knight, 2001). Using the Special Retrospective Survey of Employment and Unemployment (SRS) dataset, Statistics South Africa (1998a) reports that just under 20% of individuals cite high costs as the reason they do not engage in active job search. This decision not to search owing to high costs may be exacerbated for individuals living in areas characterised by high levels of worker discouragement. Such individuals may believe that there is little chance of finding work, even

---

<sup>1</sup> The extent to which networks serve as efficient conduits of information concerning job opportunities depends on the strength of ties amongst individuals in the network – Granovetter (1974:1995) argues that strong ties are those that link close friends while weak ties are those that link acquaintances, and that weak ties are relatively important in helping individuals find access to employment, since these weak ties bring information about opportunities available outside the individuals immediate circle and locale.

<sup>2</sup> For example, the empirical evidence to date suggests that network effects are particularly important in poorer communities (Elliot, 1999) and that reliance on networks varies demographically, with, for example, women being less reliant on networks than men. (Ports, 1993; Bradshaw, 1973).

if they spend the necessary money to reach the job market, since this is the experience of others in their community, and so, may rationally decide to not actively search as the opportunity cost of the transport is too high. At the same time, large numbers of firms rely on informal recruitment methods. Standing et al (1996) use firm level data from South Africa to show that 41.4% of all firms studied relied on friends and relatives of their existing employees to recruit new workers. This reliance on informal methods may be attributable to the poor signaling content of the school-leaving exam qualification, given the continuing discrepancies in educational quality, as well as to the high costs associated with formal recruitment procedures, especially for relatively low skilled jobs.

Yet, until relatively recently, the South African literature on job search and employment has focused on individual job search and work choices in isolation, ignoring the potential influence that social networks, constituted by families, peers and acquaintances, might have on these individual decisions. Networks facilitate the exchange of job related information (Ioannides and Loury, 2004), be it about actual job openings or about the paucity of such opportunities. Thus, access to information about job opportunities as well as one's perceptions about the buoyancy of the labour market may depend critically on social structures and the social networks to which individuals belong. The work that has been done examining the impact of social networks on employment status in South Africa has mostly defined the network as the number of other household members who are employed. (see Wittenberg and Pearce, 2000; Mlatsheni and Rospabe, 1999; Schoer and Leibbrandt, 2006). Wittenberg and Pearce (2000) find that in South Africa, these networks positively influence access to jobs, while Mlatsheni and Rospabe (1999) find that they significantly increase the probability of youth being in wage employment. Wittenberg's (1999) non-parametric analysis is also suggestive of a household network effect as it indicates clustering of employed and unemployed individuals<sup>3</sup>. Using a non-representative survey for a specific magisterial district located near Cape Town, Schoer and Leibbrandt (2006) demonstrate that unemployed individuals who report that they have contacts in the labour market are significantly more likely to rely on social networks in searching for a job, and that the difference between relying exclusively on social networks as opposed to more active search methods depends crucially on contact availability. Moreover, they also show that unemployed individuals who report that they rely on social networks to search for jobs are more likely to live in households where at least one household member is employed compared to unemployed individuals who rely on more active search methods, such as place-to-place searches or relying on newspapers. However, all of these studies are plagued by omitted variable bias (which we discuss in more detail below). Specifically, it is difficult to disentangle whether unemployed individuals relying on social networks to find jobs are simply more likely to live with employed individuals who support them while they engage in search behaviour, or whether the presence of an employed individual in the household lowers job search costs for the unemployed and improves access to job market information. Schoer and Leibbrandt (2006) argue that their data suggest that household employment provides access to financial resources which unemployed household members use to pursue *active* search methods, but that when unemployed individuals are constrained in their ability to engage in active search due to domestic duties, other employed individuals in the household then act as transmitters of information.

---

<sup>3</sup> Some correlations that are noted in the study based on the non-parametric techniques include: (1) If a spouse is employed then the partner will more likely be employed too, (2) Search patterns of parents are transmitted to their children.

Recognising that most studies of network effects suffer from omitted variable bias, Bertrand et al (2000) devise an estimation strategy to try to minimise these biases in their work examining the impact of social networks on the take up of welfare grants in the USA. They use language groups to proxy for social networks, arguing that information flows more easily between those who speak the same language, and include language fixed effects as well as geographic area fixed effects to minimise problems arising from omitted variable bias. They find that networks may increase the responsiveness of welfare use to policy shocks by an additional 15 to 27%. Our empirical work draws directly on their approach, but we differ in that we use age-language<sup>4</sup> cohorts to define the social network, and we focus on the impact of social networks on labour market status. We include age-language cohort fixed effects, as well as geographic area fixed effects in order to minimise potential biases that might arise from omitted variable bias. Our network estimates, while small are not insubstantial, and suggest that social networks alone may enhance employment probabilities by an additional 3 to 12%. In contrast, the impact of social networks on reducing worker discouragement is much smaller, at between 1-2%.

## Methodology

Assume that the true model governing the probability that an individual is employed is given by the following:

$$\Pr(EMPL_{ijk}) = Netw_{ijk}\alpha^* + X_i\beta^* + Y_j\gamma^* + Z_k\delta^* + \varepsilon_{ijk}$$

where  $i$  indexes individuals,  $j$  indexes geographic areas, and  $k$  indexes age-language cohort.  $EMPL$  is a dummy variable that takes a value of one if the individual is employed, and zero otherwise.  $Netw$  measures the effect of information and social pressure from individuals in one's social network on one's employment status,  $X$  are observed and unobserved personal characteristics,  $Y$  are observed and unobserved area characteristics, and  $Z$  are observed and unobserved age-language cohort characteristics.

Estimating this model is made difficult by the fact that data on actual contacts and the extent of one's social network are rarely collected. Consequently, mean neighbourhood characteristics are often used to proxy for social networks, making the implicit assumption that contacts are randomly distributed within the neighbourhood. While there is now a substantial body of literature to suggest strong correlations between mean neighbourhood outcomes and individual outcomes (see Jencks and Meyer, 1990), causal statements about the impact of social networks on individual outcomes typically cannot be made due to problems of omitted variable bias (Manski, 1993). More specifically, omitted personal characteristics may be correlated with mean group behaviour (for example, individuals living in areas with a high incidence of unemployed workers may themselves be less ambitious), or alternatively, omitted neighbourhood characteristics may be correlated with the mean incidence of unemployed workers in the area (for example, deep rural areas may have fewer job opportunities, thus making it less likely that individuals are able to find jobs, or engage in

---

<sup>4</sup> Bertrand et al (2000) define the network on the basis of language only. We combine age and language, so that individuals who speak language X (e.g. English) and are in a common age cohort Y (e.g. age 15-24 years) are considered to be part of the same social network.

active job search). Moreover, even if data on actual contacts were available, the problem remains that individuals select their contacts or friends. Since individuals with many contacts may be qualitatively different from individuals with only a few contacts, any estimates obtained in this manner may also suffer from omitted variable bias.

We use age-language cohorts within magisterial districts<sup>5</sup> to proxy for social networks. In other words, all individuals between the age of 15-24 who speak Zulu and reside in magisterial district X are considered part of the same social network. This network would be distinct from that for individuals aged 15-24 who speak English but who also reside in the same magisterial district X<sup>6</sup>. Importantly, this social network proxy should be interpreted as a measure of potential contacts rather than actual contacts, since we still do not have explicit information about peer networks. We then include fixed effects for magisterial districts as well as age-language cohorts to minimise potential bias arising from omitted variable bias at the geographic area level as well as at the age-language cohort level.

While there are arguably many ways one could define a social network (along religious lines, for example), we adopt the age-language cohort measure since we think it plausible that speaking a common language is important in accessing information about job opportunities, and that individuals are more likely to spend time with, and thus acquire information from, individuals who are roughly in the same age grouping as themselves and live in the same locality.<sup>7</sup> International evidence certainly suggests that individuals are more likely to associate with others in similar social and occupational categories, with the result that social networks tend to sort along dimensions such as race, ethnicity, age and religion. (Nan Lin 2001; Waldinger, 1996). Bakalian (1993) shows that friendship groups tend to be loosely aligned along ethno-linguistic lines, while Alba (1990) argues that the use of one's mother tongue is an important determinant of ethnic identity. These ethno-linguistic group niches hold important implications for the value of job contacts and connections for members of certain groups entering the labour market for the first time or who are changing jobs. While

---

<sup>5</sup> A lower level of disaggregation was restricted by the availability of data at this lower level which is not publicly released. In the sample there are 354 magisterial districts. The magisterial districts vary considerably in size ranging from 85 to 35,334 individuals. For language groupings, we use the eleven official languages in South Africa, and we define our age brackets as follows: Bracket 1 (15-24 years); Bracket 2 (25-34 years); Bracket 3 (35-44 years); Bracket 4 (45-54); Bracket 5 (55-65 years). Our age-language cohorts reflect all the possible age-language combinations that arise by utilizing these two variables.

<sup>6</sup> Loury (2002) and Montgomery (1991) both make the case that the social ties that may matter most may be to individuals who are slightly older than the potential employee. To the extent that this is true, and to the extent that the relevant social ties exist across the age brackets as we have defined them, our network co-efficients will be an under-estimate. However, since our age brackets span 9 years, we think the bias arising from this is likely to be quite small.

<sup>7</sup> Geographical proximity on its own is an insufficient proxy. Firstly, if one observes a correlation between outcomes for individuals who are spatially close, one needs to be able to disentangle whether this arises because these individuals have access to the same information sources (correlated effect); because they share some similar personal characteristics (contextual effect) or because they learn from one another's behaviour (endogenous social effect) (Manski, 2000). Moreover, Conley and Udry (2002) utilize a detailed data set for Ghana that contains questions about actual farming contacts to demonstrate that simply relying on geographical proximity as a network indicator may produce significant network effects in a regression context when, in fact, they do not exist. When relying solely on geographic proximity as a measure of social networks, they find evidence of significant social learning in relation to the farming of maize-cassava. In stark contrast, however, when they use an index of contacts based on actual contact information collected from a comprehensive set of questions including "Who do you get advice on farming from?", there is no evidence of significant social learning in this context.

such networks may facilitate access into employment, they may facilitate access to very specific sectors of the labour market, thereby restricting social mobility to some degree<sup>8</sup>.

The network measure is defined as:  $Netw_{jk} = C_{jk} * \overline{EMPL_k}$  where  $C_{jk}$  is the density of group  $k$  in area  $j$ , where area is defined by the magisterial district<sup>9</sup> and group  $k$  is defined by a common age-language group. This provides a measure of the number of potential contacts one has in one's area.  $\overline{EMPL_k}$  is the mean incidence of employed individuals from age-language group  $k$  in the population, and provides a measure of the employment status of one's network contacts<sup>10</sup>. Thus, the equation being estimated is:

$$EMPL_{ijk} = (C_{jk} * \overline{EMPL_k})\alpha + X_i b + g_j + d_k + C_{jk} f + e_{ijk}$$

where  $g_j, d_k$  are fixed effects for magisterial districts and age-language groups. We include  $C_{jk}$  on its own as a control, but do not include  $\overline{EMPL_k}$  as control on its own as this is accounted for through the age-language cohort fixed effects. The interaction of these two terms provides a measure of the influence of social networks on individual behaviour. A positive  $\alpha$  suggests evidence of networks effects in operation. Although our dependent variable is binary, we choose to estimate a linear probability model as opposed to a logit or probit, since the latter are computationally difficult given the large number of fixed effects in our model.<sup>11</sup>

Including numerous fixed effects allows us to control for many of the omitted variable biases that might arise in estimating such a model. Unobserved differences between areas such as variation in job availability, or differences in the extent to which employers rely on informal versus formal recruitment methods which may also be geographically correlated are

---

<sup>8</sup> For example, Waldinger (1996) cites evidence of Chinese workers dominating restaurant, laundry and garment industries in the US, while Italian workers tend to dominate construction jobs.

<sup>9</sup> Following Bertrand et al (2000),  $C_{jk}$  is the proportion of individuals in area  $j$  that belong to group  $k$  (where the group refers to a common age-language cohort) divided by the number of people in South Africa from that group i.e.  $\ln \frac{C_{jk} / A_j}{L_k / T}$  where  $C_{jk}$  is the total number of people in area  $j$  who belong to group  $k$ ;  $A_j$  is the

number of people who live in area  $j$ ;  $L_k$  is the total number of individuals in the population who are in the same age-language group, and  $T$  is the total population. This normalisation ensures that smaller magisterial districts are not under weighted which would otherwise be the case.

<sup>10</sup> We do not use the incidence of employed individuals from age-language group  $k$  who live in area  $j$ , excluding individual  $i$  as our measure of  $\overline{EMPL_k}$  since this specification introduces possible omitted variable bias in that it may reflect unobserved characteristics that the individual has in common with others from the same language group living in the same area. To minimize this problem, we use the mean incidence of employed individuals of the language group in the population. More specifically, this is measured as the mean deviation of the group's employment status relative to the mean employment status of the entire sample used in the analysis (see Bertrand et al (2000) for details on method).

<sup>11</sup> We use STATA to estimate our regressions. Given the computational difficulty associated with estimating a probit model in the presence of a large number of fixed effects, we ran a probit regression on a subsample of our data and compared these results to those obtained from a linear probability model run using the same subsample. The subsample represented 30% of our full sample and the associated observations reflected linguistic diversity across the magisterial districts. The signs and significance on the network coefficient are the same in both models, and the magnitude of the coefficient estimates are larger in the probit regressions than in LPM. This suggests that the estimates of network effects presented in this paper may be conservative.

accounted for through the inclusion of magisterial district fixed effects<sup>12</sup>. To the extent that different age-language cohorts experience levels of discrimination in the job market or hold differing levels of job experience, these are accounted for through the inclusion of the age-language cohort fixed effects. The direct inclusion of  $C_{jk}$  in the regression deals with any omitted personal characteristics that are correlated with  $C_{jk}$ , for example, unobserved characteristics such as ambition or trustworthiness may increase an individual's employment probability, and may also reduce the probability that the individual lives among their own language group. This would result in a positive estimate on  $f$  but would not affect the estimate on  $\alpha$ . However, it is important to point out that including  $C_{jk}$  in the regression controls for fixed differences between people who choose to live among their own age-language group and those who do not, but these differences could themselves vary by age-language group. For example, living away from your age-language cohort may signal success if you are from a group that has a high incidence of discouraged unemployed individuals (and vice versa). This differential self-selection might produce a network effect when in fact, none exists. Given that we measure  $C_{jk}$  at the magisterial district level (i.e. a high level of aggregation already) we do not think this will be a significant source of bias in our results. In addition, we include controls for whether or not the individual has relocated in the intervening period between the 1996 and 2001 Census, as an attempt to control for the propensity of individuals to move away from their group<sup>13</sup>.

## Data and Sample Characteristics

The data used in the analysis is the 10 % sample of the 2001 Census survey conducted by Statistics South Africa. Since we are interested in the economically active population, we limit the sample to individuals aged 15-65 years and exclude all non-housing units<sup>14</sup> and non participants (with the exception of discouraged workers) in the labour force.<sup>15</sup> As expected, the unemployment figures differ depending on whether one adopts the broad or narrow definition. As Table 1 demonstrates, 47% of the sample is unemployed according to the broad definition of unemployment (i.e. including the discouraged unemployed) compared with 42% when the narrow definition is used. Notably, in our sample, almost a fifth of unemployed individuals had not actively searched for a job in the four weeks prior to the Census<sup>16</sup>. A

<sup>12</sup> Since neighbourhood effects are eliminated through the inclusion of these controls, it may be that our actual estimates underestimate the true extent of network effects.

<sup>13</sup> Note that in our empirical specifications, we also estimate a set of regressions focusing on the probability that an unemployed individual is discouraged (that is, non-searching unemployed) as opposed to engaging in active job search. In other words,  $\Pr(DU_{ijk}) = Netw_{ijk}\alpha^* + X_i\beta^* + Y_j\gamma^* + Z_k\delta^* + \varepsilon_{ijk}$  where we limit the sample to unemployed individuals only, and where  $DU$  is a dummy variable that takes a value of one if the individual is a discouraged unemployed individual, and zero otherwise (i.e. unemployed searching). Everything else stays the same. As such, we might expect  $\alpha$  to have a negative sign, since if a large fraction of one's age-language cohort in one's area is employed, this should improve the transmission of information about job possibilities and the buoyancy of the labour market, thereby making individuals more inclined to bear job search costs.

<sup>14</sup> That is, all prisons, residential hostels, student residences, old age homes, and worker hostels are all excluded.

<sup>15</sup> The reason that these distinct labour force states are excluded is that the analysis seeks to answer the question whether the decision to search (which can only be based on those unemployed – either broad or narrow) is a function of the experience of other individuals of one's own network.

<sup>16</sup> The narrow unemployment rate of 42% based on the 2001 Census is substantially higher than the narrow rate of unemployment of 29.5% estimated by the 2001 September Labour Force Survey (LFS). However, the broad unemployment rates are less divergent, with the 2001 Census estimating a broad unemployment rate of 47%

report released by Statistics South Africa based on the Special Retrospective Survey of Employment and Unemployment (SRS) found that 33 % of respondents stopped looking for work due to lost hope in finding employment, while another 25 % mentioned that it was due to no jobs in the area. (StatsSA, 1998a). This accords with work by Kingdon and Knight (2001) who provide convincing evidence that the lack of job search by unemployed individuals is largely attributable to high search costs and low returns to search, as opposed to a decision to voluntarily remain unemployed. Moreover, they show that unemployed individuals who do not engage in active job search are discouraged, are significantly worse off in material terms than the searching unemployed, and that they are taken into account in local wage determinations as general labour force participants. This evidence suggests to us that for analytical purposes, adopting the broad definition of unemployment may be more appropriate, even though the narrow definition is used for official purposes in South Africa. Moreover, we are interested in exploring the impact of social networks on job search behaviour, particularly worker discouragement.

Table 2 presents summary statistics for the sample by employment<sup>17</sup> status, revealing some interesting differences between the employed and unemployed. Black individuals are under-represented amongst the employed and over-represented amongst the unemployed relative to their share of the population. This is especially true for those speaking Zulu and Xhosa, and to a lesser extent, Sepedi . The converse holds true for all three of the other race groups. Unemployment and the incidence of discouraged workers is higher amongst women than men, and worker discouragement is particularly high for individuals living in rural areas. Over a fifth of employed individuals had relocated in the intervening period between the last census and that conducted in 2001 compared with fewer than 15% of unemployed individuals.

A striking difference between employed and unemployed individuals lies in their potential access to job information via other employed household family members. On average, an employed individual lives in household where over half of all other adult members have a job. This stands in stark contrast to the unemployed, where on average, only 10% of all other household adult members are employed.

Just over a third of unemployed individuals are married compared with over 60% of their employed counterparts. To some extent, this difference may reflect an age cohort difference. Consider that just under a third of all unemployed and discouraged workers are between the ages of 15 and 24. In fact, to put it more starkly, over 60% of unemployed workers are between the ages of 15 and 35, with unemployment figures declining for those in older age cohorts (Table 2). This accords with work by Wittenberg (1999) who reports that approximately half of the strictly unemployed at age 30 have been unemployed for 3 years or

---

compared to a broad unemployment rate of 41.5%. These differences may be attributed to the fact that the LFS is a far more comprehensive labour market survey instrument than the Census, designed to capture a broad range of employment activities in the informal and subsistence agriculture sectors, particularly among those working only a few hours per week. The labour force survey questionnaire includes many more prompts to identify such people and then counts them as employed. This is not the case with the Census. This may account for the large differences in the narrow rates of unemployment between the two datasets.

<sup>17</sup> An individual is recorded as being employed in the Census if they responded “Yes” to the question “In the seven days before 10 October, did (the person) do any work for pay (in cash or kind), profit or family gain for one hour or more?”. The responses include (a) Yes: formal registered (non-farming) (b) Yes: informal unregistered (non-farming) (c) Yes (farming) (d) Yes, has work but temporarily absent (e) No, did not have work. An individual is recorded as employed if they responded Yes to (a) through (d).

more. Clearly, it is difficult for new young labour market entrants to enter the market and find employment.

Since the network measure used in this paper defines the network in terms of age-language cohorts, Tables 3 and 4 present unemployment figures by language grouping and age cohort. These statistics tell much the same story. Over half of all individuals who speak an indigenous African language are unemployed, with unemployment and the incidence of discouraged workers being particularly high amongst those who speak Xhosa, Zulu, Sepedi and Tshivenda (Table 3). Moreover, this unemployment is largely concentrated amongst the younger cohorts in these groups. For example, over 80% of all individuals between the ages of 15-24 who speak Xhosa, Zulu or Sepedi are unemployed, while almost 60% of individuals in the 25-34 year age cohort in these language groups find themselves unemployed (Table 4). Interestingly, despite these higher rates of youth unemployment, the incidence of discouraged workers is not necessarily higher amongst younger cohorts within language groupings (Tables 3 and 4).

## **Regression Results**

In Table 5, we present the network coefficient estimates from our baseline regressions, demonstrating the very large decline in the co-efficient estimates as the various fixed effects are included in the regression specification. For the employment regressions, the network co-efficient declines from 0.5 when no additional controls are included to 0.05 once age-language cohort fixed effects and magisterial district fixed effects are included. Despite this decline, the network coefficient remains highly significant. The decline is even more dramatic in the regressions estimating the probability that an unemployed worker is discouraged as opposed to actively seeking work, where the co-efficient declines effectively to zero and becomes insignificant. Clearly, failure to adequately control for omitted variable bias through the inclusion of fixed effects would lead to substantial over-estimates of the network co-efficient.

Table 6 presents regression estimates that include additional controls for the personal characteristics of the individual and their household. Once these additional controls are included, the network co-efficient in the regressions estimating employment probabilities declines from 0.047 to 0.042 and remains statistically significant (compare the network co-efficient in Column 3 of Table 5 to the network co-efficient in Column 1 of Table 6). In Column 2 of Table 6, we include an additional control for the number of other household members who are employed. As the number of other adults in the household who are employed increases, this has a positive and significant impact on an individual's employment prospects, and this is associated with a small decline in the estimated network co-efficient from 0.042 to 0.037. In the parlance of Granovetter (1974), one might designate these other employed household members as strong ties, since the individual is in close and frequent contact with these individuals. Thus, the fact that the network co-efficient remains positive and statistically significant even after controlling for this source of information about employment opportunities, may suggest that relatively "weaker ties" constitute an important source of job-related information.

Including additional controls for personal characteristics in the regressions concerned with job search status results in the network co-efficient becoming statistically significant, and increasingly marginally in magnitude (compare the network co-efficient in Column 6 of Table

5 to that in Column 3 of Table 6). However, the inclusion of a control for the number of other household members who are employed, has no impact on explaining worker discouragement and its inclusion does not alter the network co-efficient at all which remains small, negative but significant (Table 6, Column 4).

It is difficult to interpret the actual magnitude of the network coefficient in these regressions given the way it is derived, but Bertrand et al (2000) argue that it is possible to arrive at a measure of the magnitude of the network effect, by asking to what extent social networks would magnify a policy shock affecting employment probabilities<sup>18</sup>. They assume that the impact of the policy shock is linear, and argue that the marginal change induced solely by social networks is given by  $\frac{1}{(1-\alpha C_k)} - 1$ . Table 7 presents the indirect network impact on employment probabilities and employment search status by age-language cohort<sup>19</sup>. The network effects, while small, are not insubstantial. With respect to employment prospects, social networks alone may enhance employment probabilities by between 3 and 12% (Panel A of Table 7). Recall that these effects are in addition to any direct changes induced by employment policies. In contrast, the influence of social networks on worker discouragement are much smaller, reducing discouragement by 1-2% at most (Panel B of Table 7)<sup>20</sup>.

Table 8 presents the indirect network impact on employment probabilities and worker discouragement for specified groups. We focus first on the regressions that estimate employment status. In order to ensure that our results are not being driven by the largest language groups dominating the results, we re-run the specification excluding Xhosa and Zulu speakers (since these are the largest language groups in the sample). Excluding these groups increases the magnitude of the network coefficient, and the results suggest that social networks would improve employment prospects by an additional 8% if these two groups were

<sup>18</sup> Applying the Bertrand et al (2000) thought experiment to our work, consider a policy  $\psi$  that has a linear impact on employment outcomes. This variable can be added to the model, and be scaled to ensure that a one percentage point increase in  $\psi$  results in a one percentage point increase in employment probabilities in the absence of any network effects, to give

$EMPL_{ijk} = \psi + (C_{jk} * \overline{EMPL}_k)\alpha + X_i b + g_j + d_k + C_{jk} f + e_{ijk}$ . In equilibrium, the increase in employment probabilities will be greater than the increase in  $\psi$ , owing to the multiplier effect generated by social networks on employment probabilities. Averaging both sides of the equation for each age-language cohort and differentiating with respect to  $\psi$ , gives  $\frac{d\overline{EMPL}_k}{d\psi} = 1 + \overline{C}_k * \frac{d\overline{EMPL}_k}{d\psi} \alpha$  where  $\overline{C}_k$  is the mean of

$C_{jk}$  within each age-language cohort. Solving this equation provides a measure of the responsiveness of each age-language cohort's employment prospects in response to a policy change. However, since this includes the direct effect of the policy change, one is subtracted from this formula to arrive at the marginal change induced solely by social networks, namely  $\frac{1}{(1-\alpha \overline{C}_k)} - 1$ .

<sup>19</sup> The indirect network effect is calculated using  $\frac{1}{(1-\alpha C_k)} - 1$  where  $\alpha$  is the relevant co-efficient from Row 3 of Table 6 (when all fixed effects are included as well as additional controls for personal and household characteristics).

<sup>20</sup> Given the approach we adopt here, we are not able to say anything about the precision of these calculated effects.

excluded from the analysis. The network co-efficient on employment status for women is double that for men, with the indirect effect suggesting that social networks increase employment prospects for women by three times that for men. The network coefficient for those living in rural areas is half that of those living in non-rural areas, increasing employment prospects for those in rural areas by 6% compared to 7% for those in non-rural locations<sup>21</sup>. Interestingly, networks appear to have a larger positive impact on employment probabilities for those who have not relocated than compared with those individuals who had relocated in the period between the 1996 Census and the 2001 Census. Finally, the network effect on employment probabilities is also larger (and more significant) for individuals with access to a telephone, increasing employment probabilities by an additional 5% compared to only 2% for individuals without access to this means of communication.

In contrast, the results for the effect of social networks on employment search status are much weaker, being both smaller in magnitude and in statistical significance. Excluding Xhosa and Zulu speakers from the analysis does not alter the network co-efficient at all, and in all the additional specifications we run, the results suggest that worker discouragement is only significantly reduced through the operation of social networks for those without disabilities, those who have relocated, and those with access to a telephone. In each of these cases, worker discouragement is reduced by 1% at most.

## Discussion

In this paper, we have presented estimates of the magnitude of social networks on employment probabilities and worker discouragement, controlling for as many sources of omitted variable bias as possible. Our results illustrate the importance of properly controlling for omitted variable bias in this type of work, since failure to do so clearly results in an overstatement of the magnitude of the network effects. Even after controlling for various fixed effects and controls for personal and household characteristics, our network coefficients remain significant, although they have a substantially smaller impact on reducing the incidence of worker discouragement amongst the unemployed than they do on affecting employment probabilities. Indeed, our results suggest to us that there is a lot about worker discouragement that remains unexplained through the usual socio-economic analysis, and may point to the need for a more psychologically grounded behavioural analysis in order to explain this phenomenon more fully.

Secondly, while it is clear that informational externalities provided through other employed adult household members are important for enhanced job prospects, information and job opportunities acquired through this means accounts for only a small proportion of the measured social network effect, suggesting an equally important role for “weaker ties” within the social network. However, at the same time, our results also suggest that informational externalities provided through other employed adult household members do not affect the

---

<sup>21</sup> Note that despite the difference in the network coefficient, the derivation of the indirect network effect also takes the value of  $C_k$  into account. The relatively small difference in the final network effects is thus accounted for through corresponding differences in these values for those in rural versus non-rural locations.

incidence of discouragement among unemployed individuals. This is a phenomenon worthy of further investigation.

Our results confirm that the importance of social networks in terms of employment probabilities and in affecting worker discouragement varies demographically. Social networks significantly enhance employment prospects for women, for individuals who did not relocate in the 1996-2001 period, and for those with access to telephones. In contrast, social networks appear to significantly reduce worker discouragement amongst individuals who did relocate in the 1996-2001 period, and those with access to telephones. However, it is important to note that the absence of a significant network effect for specific groups of individuals in our results may reflect lower usage of networks by the specified group to find jobs or to gather information about jobs, but it could also reflect differences in the likelihood of the individual in the specific group being made a job offer for jobs that one hears of through word of mouth.

It is also the case that employer characteristics will determine to some extent the context in which different types of job search methods operate. In some cases, worker characteristics may be easily observable, while in other cases, recommendations from existing employees may be especially important or valuable. The strength of network effects will vary in these circumstances. In addition, job seekers may vary in terms of their contact availability and in terms of the quality of their networks. In other words, the impact of the social network on individual outcomes depends not only on the size and quality of the network, but also on whether the individual is able to access it effectively. For example, migrants may have greater difficulty accessing the social network since they are relative newcomers to the area, and it is perhaps unsurprising that social networks have smaller effects on the employment probabilities of those individuals who had relocated. However, social networks significantly reduce worker discouragement for those who did relocate. Since it is plausible that these individuals may have relocated to look for work, it is unsurprising that they are significantly more likely to be engaged in active job search. This suggests to us that individuals who move away from their own age-language cohorts may well be qualitatively different from those who stay behind, and again underscores the importance of trying to control for this source of bias in network studies.

In this paper, we do not identify the specific ways in which the effectiveness of social networks depend on differences in the characteristics of job seekers, in the characteristics of the contacts they use, or the relationship between job seekers and their contacts and features of the job environment. We simply document the existence of network effects for various groups. Thus, a large number of questions remain unanswered, and in the absence of more detailed survey information of job search and recruitment methods, may remain so. Providing convincing explanations for the difference in magnitude of network effects for different groups is an example of one such question. For example, is it that groups that rely more heavily on social networks to access employment opportunities or gather job information do so because they have skills which are difficult to observe, with the result that employers rely more heavily on social networks in their hiring decisions, or is it that these groups face explicit labour market discrimination in formal hiring processes and thus, have to rely on networks to access employment. Or is it that some groups, for example women, gain more from their social networks than men because their contacts have better information about job opportunities and do a better job of passing this information on.

Moreover, while our statistics describing the sample characteristics suggest that some age-language cohorts may find it more difficult to access opportunities in the labour market than others, an unanswered question is whether ethno-linguistic job niches exist in the South African labour market, in much the same way as documented by Waldinger (1991) for the USA. To the extent that such ethno-linguistic niches exist, this will reinforce the importance of social ties in job search and employment prospects, but may also serve to inhibit social mobility. Arguably, understanding these kinds of dynamics, which will hinge critically on collecting better data on social networks, will contribute substantially to a better understanding of the complexities inherent in the South African labour market.

## References:

- Abel, A (1990) "Asset Prices Under Habit Formation and Catching up with the Joneses", *American Economic Review*, LXXX, 38-42.
- Alba, R.D. ((1990) "Ethnic Identity: The Transformation of White America; Yale University Press; New Haven, CT.
- Arrow, K.J. and R. Borzekowski (2004) "Limited Network Connections and the Distribution of Wages", mimeo.
- Bakalian, A (1993) "American-Americans: From Being to Feeling American, Transaction Publishers, New Brunswick, NJ.
- Banerjee, A. (1992) "A Simple Model of Herd Behaviour", *Quarterly Journal of Economics*, CVII, 797-818
- Bertrand, M., Luttmer, E., and Mullainathan, S. (2000): "Network Effects and Welfare Cultures", *Quarterly Journal of Economics*, 116 (1), 351 – 377.
- Bikhchandani, S; D. Hirshleifer and I.Welch (1992) "A Theory of Fads, Fashion, Custom and Cultural Change in Informational Cascades", *Journal of Political Economy*, C, 992-1026.
- Blau, D. and P. Robbins (1992) "Job Search Outcomes for the Employed and Unemployed", *Journal of Political Economy*, 98:3, 637-55.
- Bradshaw, T (1973) "Jobseeking Methods Used by Unemployed Workers" *Monthly Labor Review*, 96:2, 35-40.
- Calvo-Armengol, A. and M.O. Jackson (2004) "The Effects of Social Networks on Employment and Inequality", *American Economic Review*, 94:3, 426-454.
- Case, A.C. and L.F. Katz (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths", *National Bureau for Economic Research Working Paper No. 3705*, May.
- Coleman, J.S. et al (1996) "Equality of Educational Opportunity", Government Printing Office, Washington DC.
- Conley, T. and Udry, C. (2002): "Learning about a New Technology: Pineapple in Ghana", mimeo.
- Datcher, L. (1983) "The Impact of Informal Networks on Quit Behaviour", *Review of Economic Statistics*, 65, 3, 491-95.
- Devine, T and N. Kiefer (1991) "Empirical Labor Economics: The Search Approach", Oxford University Press, NY.
- Elliot, J (1999) "Social Isolation and Labor Market Isolation: Network and Neighbourhood Effects on Less-Educated Urban Workers" *Sociology Quarterly*, 40:2; 199-216.
- Ellwood, D.T. (1982): "The Spatial Mismatch Hypothesis: Are there Teenage Jobs Missing in the Ghetto? in Freeman, R.B. and Holzer, H. (eds): *The Black Youth Employment Crisis*. University of Chicago Press, Chicago.
- Granovetter, M. S. (1974) "Getting a Job: A Study of Contacts and Careers", first edition, Harvard University Press, Cambridge, MA.

Ioannides, Y.M. and L.D. Loury (2004) "Job Information Networks, Neighborhood Effects and Inequality", *Journal of Economic Literature*, XLII, December, 1056-1093.

Jencks, C and S. Mayer (1990) "The Social Consequences of Growing Up in a Poor Neighbourhood" in Laurence E. Lynn and Michael G.H. McGeary (eds) *Inner-City Poverty in the United States*, National Academy Press, Washington DC.

Kingdon, G. and Knight, J. (2001): "Are searching and Non-searching Unemployment distinct states when unemployment is high? The Case of South Africa", Centre for the Study of African Studies, Oxford University.

Kingdon, G. and Knight, J. (2006): "The Measurement of Unemployment when Unemployment is High", *Labour Economics*, Vol 13. No.3, 291-313.

Klasen, S. and Woolard, I. (2000): "Surviving Unemployment without State Support: Unemployment and Household Formation in South Africa. Unpublished mimeo.

Lin, N (2001) "Social Capital: A Theory of Social Structure and Action". Cambridge University Press, Cambridge and New York.

Manski, C. (1993): "Identification of Endogenous Social Effects: the Reflection Problem", *Review of Economic Studies*, Vol. 60, No. 3.

Marsden, P and E.H. Gorman (2001) "Social Networks, Job Changes, and Recruitment" in *Sourcebook of Labor Markets: Evolving Structure and Processes*, Ivar Berg and Arne Kalleberg (eds); Kluwer Academic/Plenum Publishers, NY.

Mlatsheni, C. and Rospabe, S. "Why is youth unemployment so high and unequally spread in South Africa" Unpublished mimeo.

Moll, P. (1996): "The Collapse of Primary Schooling Returns in South Africa, 1960 – 1990". *Oxford Bulletin of Economics and Statistics*, Vol. 58, pp. 185 – 209.

Montgomery, J.D. (1991) "Social Networks and Labor-Market Outcomes: Towards an Economic Analysis; *American Economic Review*, Vol. 81, No. 5 (December).

Ports, M. (1993) Trends in Job Search Methods: 1970-1992", *Monthly Labour Review*, October, 63-67.

Rees, A. (1986): "An Essay on Youth Joblessness", *Journal of Economic Literature*, Vol. 24, pp. 613 – 628.

Schoer, V., Leibbrandt, M. (2006). 'Determinants of job search strategies: Evidence from the Khayelitsha/Mitchell's Plain survey' *South African Journal of Economics*, 74 (4): 702-724.

Standing, G., Sender, J. and Weeks, J. (1996): "Restructuring the Labour Market: The South African Challenge: An ILO Country Review, International Labour Organization, Geneva, Switzerland.

StatsSA. (1998a): "Unemployment and Employment in South Africa", Statistics South Africa, Pretoria.

StatsSA (1998b): "Unemployment and Employment in South Africa: October Household Survey: 1994 – 1997", Statistics South Africa, Pretoria.

Waldinger, R.D. (1996) "Still the Promised City?", Harvard University Press, Cambridge, MA.

Wittenberg, M. (1999): "Job search and household structure in an era of mass unemployment: A semi-parametric analysis of the South African labour market" Unpublished mimeo, University of Witwatersrand.

Wittenberg, M and C. Pearce (1996) "Youth Unemployment: Some Perspectives from the South African Living Standards and Development Survey" in L. Chisholm et al *Out-of-School Youth Report: Policy and Provision for Out-of-School and out-of-Work Youth*, Education Policy Unit, University of Witwatersrand, 1996.

Table 1: Narrow and Broad Unemployment rates

		Census 2001
A.	Total employed	768,701
B.	Total unemployed (official definition)	561,432
C.	Total economically active = A+B	1,330,133
D.	Narrow unemployment rate=B/C	0.422
E.	Total unemployed (broad definition)	689,414
F.	Total economically active (broad definition) = E+A	1,458,115
G.	Broad unemployment rate = E/F	0.472
H.	Not searching/discouraged as fraction of unemployed	0.186

*Notes:*

- Data source: 10 % sample of the 2001 South African Census
- Sample is limited to the economically active population, that is, individuals aged 15-65 years; all non-housing units and non-participants (with the exception of discouraged workers) in the labour force are excluded.
- An individual is recorded as being employed in the Census if they responded “Yes” to the question “In the seven days before 10 October, did (the person) do any work for pay (in cash or kind), profit or family gain for one hour or more?”. The responses include (a) Yes: formal registered (non-farming) (b) Yes: informal unregistered (non-farming) (c) Yes (farming) (d) Yes, has work but temporarily absent (e) No, did not have work. An individual is recorded as employed if they responded Yes to (a) through (d).
- The narrow unemployment rate of 42% based on the 2001 Census is substantially higher than the narrow rate of unemployment of 29.5% estimated by the 2001 September Labour Force Survey (LFS). However, the broad unemployment rates are less divergent, with the 2001 Census estimating a broad unemployment rate of 47% compared to a broad unemployment rate of 41.5%. These differences may be attributed to the fact that the LFS is a far more sophisticated labour market survey instrument than the Census, designed to capture a whole range of employment activities in the informal and subsistence agriculture sectors, particularly among those working only a few hours per week. The labour force survey questionnaire includes many more prompts to identify such people, which is not possible during Census enumeration. This may account for the large differences in the narrow rates of unemployment between the two datasets.

Table 2: Mean statistics for sample by employment status

Variable	All		Employed individuals		Unemployed individuals		Non-searching unemployed	
Individual is Black	0.77	(0.42)	0.62	(0.48)	0.91	(0.29)	0.92	(0.27)
Individual is Coloured	0.10	(0.30)	0.14	(0.34)	0.07	(0.25)	0.05	(0.22)
Individual is Asian/Indian	0.03	(0.17)	0.04	(0.21)	0.01	(0.10)	0.01	(0.09)
Individual is White	0.10	(0.30)	0.19	(0.39)	0.02	(0.13)	0.02	(0.13)
Male	0.46	(0.50)	0.56	(0.50)	0.44	(0.50)	0.41	(0.49)
Years of education	8.24	(4.55)	9.65	(4.80)	7.71	(4.36)	6.85	(4.57)
Age	33.35	(13.46)	37.38	(10.62)	32.06	(10.70)	32.20	(11.39)
Fraction of other household adults (age 16-65) employed	0.25	(0.29)	0.53	(0.29)	0.10	(0.15)	0.08	(0.14)
Has access to a telephone	0.91	(0.28)	0.96	(0.21)	0.90	(0.30)	0.86	(0.35)
Has a disability	0.05	(0.22)	0.03	(0.16)	0.04	(0.19)	0.05	(0.21)
Lives in rural location	0.38	(0.49)	0.23	(0.42)	0.41	(0.49)	0.58	(0.49)
Married	0.42	(0.49)	0.61	(0.49)	0.36	(0.48)	0.36	(0.48)
Married Male	0.19	(0.39)	0.38	(0.49)	0.14	(0.35)	0.13	(0.33)
Individual relocated between 1996 and 2001 Census	0.14	(0.35)	0.21	(0.41)	0.13	(0.34)	0.11	(0.31)
Age bracket 1: 15-24 years	0.33	(0.47)	0.11	(0.32)	0.29	(0.45)	0.31	(0.46)
Age bracket 2: 25-34 years	0.25	(0.43)	0.32	(0.47)	0.35	(0.48)	0.32	(0.47)
Age bracket 3: 35-44 years	0.20	(0.40)	0.31	(0.46)	0.21	(0.41)	0.19	(0.40)
Age bracket 4: 45-54 years	0.13	(0.34)	0.19	(0.39)	0.11	(0.32)	0.12	(0.33)
Age bracket 5: 55-65 years	0.08	(0.28)	0.07	(0.25)	0.04	(0.19)	0.05	(0.21)
Afrikaans	0.14	(0.35)	0.22	(0.42)	0.07	(0.26)	0.06	(0.24)
English	0.09	(0.29)	0.16	(0.37)	0.03	(0.16)	0.02	(0.14)
isiNdebele	0.02	(0.13)	0.01	(0.12)	0.02	(0.13)	0.02	(0.14)
isiXhosa	0.17	(0.38)	0.12	(0.32)	0.21	(0.41)	0.21	(0.41)
isiZulu	0.22	(0.42)	0.17	(0.38)	0.28	(0.45)	0.25	(0.43)
Sepedi	0.09	(0.29)	0.07	(0.26)	0.10	(0.30)	0.13	(0.34)
Sesotho	0.08	(0.28)	0.08	(0.27)	0.10	(0.30)	0.08	(0.28)
Setswana	0.09	(0.28)	0.08	(0.27)	0.09	(0.29)	0.10	(0.30)
Siswati	0.03	(0.16)	0.02	(0.15)	0.03	(0.16)	0.03	(0.16)
Tshivenda	0.02	(0.15)	0.02	(0.13)	0.03	(0.16)	0.03	(0.18)
Xitsonga	0.04	(0.20)	0.04	(0.19)	0.05	(0.22)	0.06	(0.24)

*Notes:*

Data source: 10 % sample of the 2001 South African Census

Category Unemployed” include non-searching unemployed.

Figures in brackets are standard deviations.

Table 3: Mean statistics for Language and Age Cohorts by Employment Status

Language group/Age group	Employed	Unemployed	Non-searching unemployed
Afrikaans	76.81	23.19	15.18
English	86.69	13.31	14.06
isiNdebele	46.31	53.69	20.93
isiXhosa	37.99	62.01	18.64
isiZulu	39.90	60.10	16.45
Sepedi	43.02	56.98	24.13
Sesotho	45.81	54.19	15.72
Setswana	47.95	52.05	20.01
Siswati	46.72	53.28	17.97
Tshivenda	42.35	57.65	24.99
Xitsonga	42.86	57.14	22.15
15-24 years	29.23	70.77	19.74
25-34 years	49.61	50.39	16.98
35-44 years	61.18	38.82	17.24
45-54 years	64.38	35.62	19.83
55-65 years	66.86	33.14	23.86

*Notes:*

Data source: 10 % sample of the 2001 South African Census

Category “Unemployed” include non-searching unemployed.

Table 4: Mean statistics for Age-Language Cohorts by Employment Status

Language	Age cohort	Employed	Unemployed	Non-searching unemployed
Afrikaans	15-24 years	57.17	42.83	15.12
	25-34 years	78.51	21.49	14.80
	35-44 years	82.64	17.36	14.46
	45-54 years	84.78	15.22	16.32
	55-65 years	87.24	12.76	19.61
English	15-24 years	70.37	29.63	12.62
	25-34 years	88.28	11.72	14.35
	35-44 years	90.37	9.63	12.91
	45-54 years	90.38	9.62	16.05
	55-65 years	91.34	8.66	20.34
isiNdebele	15-24 years	20.71	79.29	22.28
	25-34 years	42.10	57.90	20.30
	35-44 years	58.78	41.22	18.57
	45-54 years	62.08	37.92	22.28
	55-65 years	57.18	42.82	24.37
isiXhosa	15-24 years	17.75	82.25	20.45
	25-34 years	36.66	63.34	16.56
	35-44 years	47.08	52.92	17.60
	45-54 years	49.54	50.46	20.16
	55-65 years	51.40	48.60	24.73
isiZulu	15-24 years	18.30	81.70	17.79
	25-34 years	39.38	60.62	14.83
	35-44 years	50.66	49.34	15.58
	45-54 years	52.30	47.70	18.11
	55-65 years	52.86	47.14	21.50
Sepedi	15-24 years	17.56	82.44	26.22
	25-34 years	39.52	60.48	22.14
	35-44 years	55.47	44.53	22.70
	45-54 years	57.73	42.27	25.83
	55-65 years	57.71	42.29	31.14
Sesotho	15-24 years	20.38	79.62	17.48
	25-34 years	42.76	57.24	14.65
	35-44 years	55.64	44.36	14.32
	45-54 years	58.99	41.01	16.50
	55-65 years	60.32	39.68	19.83
Setswana	15-24 years	22.58	77.42	22.74
	25-34 years	44.09	55.91	18.41
	35-44 years	58.74	41.26	18.05
	45-54 years	61.36	38.64	20.44
	55-65 years	62.82	37.18	25.91
Siswati	15-24 years	23.38	76.62	18.89
	25-34 years	45.56	54.44	16.23
	35-44 years	58.46	41.54	18.42
	45-54 years	59.82	40.18	19.44
	55-65 years	58.63	41.37	21.96
Tshivenda	15-24 years	17.45	82.55	27.15
	25-34 years	39.05	60.95	22.57
	35-44 years	53.93	46.07	23.73
	45-54 years	57.19	42.81	28.71
	55-65 years	57.51	42.49	30.91
Xitsonga	15-24 years	23.61	76.39	23.45
	25-34 years	41.63	58.37	21.12
	35-44 years	53.72	46.28	20.25
	45-54 years	55.52	44.48	24.18
	55-65 years	56.62	43.38	27.59

Table 5: Regression estimates of network co-efficient as additional fixed effects are included

	Probability individual is employed			Probability that unemployed individual is discouraged/not searching		
	(1)	(2)	(3)	(4)	(5)	(6)
Contact availability	-0.060** (0.004)	-0.045** (0.005)	-0.022** (0.003)	0.023** (0.003)	0.020** (0.003)	0.002 (0.001)
Network effect	0.499** (0.020)	0.118** (0.011)	0.047** (0.013)	-0.063** (0.006)	-0.050** (0.005)	-0.004 (0.003)
Constant	0.553** (0.011)	0.782** (0.016)	0.729** (0.011)	0.155** (0.006)	0.140** (0.006)	0.197** (0.005)
Observations	1254829	1254829	1254829	608846	608846	608846
R-squared	0.09	0.18	0.21	0.01	0.01	0.05
Age-language cohort fixed effects (55 groups)	No	Yes	Yes	No	Yes	Yes
Magisterial district fixed effects (354 groups)	No	No	Yes	No	No	Yes

*Notes:*

- Data source: 10 % sample of the 2001 South African Census
- Significance: \*\*=1%; \*=5%
- Robust standard errors in parentheses; standard errors control for clustering at the magisterial district level.
- Variable “Contact availability” refers to  $C_{jk}$  as specified in the text.
- Variable “Network effect” refers to  $Netw_{jk} = C_{jk} * \overline{EMPL_k}$  as specified in the text.

Table 6: Regression estimates of network co-efficient including additional controls for personal characteristics

	Probability that individual is employed		Probability that unemployed individual is discouraged	
	(1)	(2)	(3)	(4)
Contact availability	-0.018** (0.002)	0.002 (0.001)	0.001 (0.001)	0.001 (0.001)
Network effect	0.042** (0.010)	0.037** (0.008)	-0.006* (0.003)	-0.006* (0.003)
Individual is Black	-0.148** (0.009)	-0.020** (0.006)	-0.062** (0.009)	-0.062** (0.009)
Individual is Coloured	-0.135** (0.010)	0.021** (0.006)	-0.042** (0.008)	-0.042** (0.008)
Individual is Asian/Indian	-0.069** (0.010)	0.053** (0.005)	-0.032** (0.009)	-0.032** (0.009)
Individual is male	0.004 (0.004)	-0.040** (0.002)	-0.021** (0.002)	-0.021** (0.002)
Years of education	0.024** (0.000)	0.012** (0.000)	-0.005** (0.000)	-0.005** (0.000)
Years of education squared	-0.000** (0.000)	-0.000** (0.000)	0.000** (0.000)	0.000** (0.000)
Individual relocated between 1996 and 2001	0.055** (0.004)	-0.035** (0.002)	-0.000 (0.003)	-0.000 (0.003)
Individual has access to a telephone	0.034** (0.005)	0.024** (0.002)	-0.026** (0.005)	-0.026** (0.004)
Individual has a disability	-0.048** (0.003)	-0.034** (0.002)	0.025** (0.004)	0.025** (0.004)
Individual is married	0.021** (0.004)	-0.017** (0.004)	0.001 (0.002)	0.001 (0.002)
Individual is married male	0.178** (0.006)	0.219** (0.006)	-0.015** (0.002)	-0.015** (0.002)
Individual lives in rural area	0.020 (0.013)	0.011** (0.004)	0.073** (0.008)	0.073** (0.008)
Fraction of other adults in household who are employed		0.958** (0.015)		0.002 (0.006)
Constant	0.437** (0.010)	0.114** (0.008)	0.281** (0.009)	0.280** (0.009)
Observations	1254829	1254829	608846	608846
R-squared	0.26	0.54	0.06	0.06

- Data source: 10 % sample of the 2001 South African Census; Significance: \*\*=1%; \*=5%
- Robust standard errors in parentheses; standard errors control for clustering at the magisterial district level.
- Variable “Contact availability” refers to  $C_{jk}$  as specified in the text.
- Variable “Network effect” refers to  $Netw_{jk} = C_{jk} * \overline{EMPL_k}$  as specified in the text.
- All regressions include fixed effects for age-language cohorts and magisterial districts.

Table 7: Indirect network impact on employment probabilities and worker discouragement by age-language cohort

Panel A: Indirect network effects on employment probabilities					
All	0.046 (0.008)				
	15-24 years	25-34 years	35-44 years	45-54 years	55-65 years
Afrikaans	0.036 (0.008)	0.036 (0.007)	0.035 (0.007)	0.033 (0.007)	0.033 (0.007)
English	0.039 (0.008)	0.039 (0.008)	0.038 (0.008)	0.038 (0.008)	0.039 (0.008)
isiNdebele	0.081 (0.011)	0.070 (0.010)	0.070 (0.010)	0.072 (0.010)	0.077 (0.010)
isiXhosa	0.040 (0.008)	0.032 (0.007)	0.035 (0.007)	0.038 (0.008)	0.047 (0.008)
isiZulu	0.038 (0.008)	0.033 (0.007)	0.033 (0.007)	0.035 (0.007)	0.039 (0.008)
Sepedi	0.062 (0.009)	0.051 (0.009)	0.052 (0.009)	0.055 (0.009)	0.063 (0.010)
Sesotho	0.054 (0.009)	0.048 (0.009)	0.050 (0.009)	0.051 (0.009)	0.053 (0.009)
Setswana	0.063 (0.010)	0.058 (0.009)	0.059 (0.009)	0.059 (0.009)	0.064 (0.010)
Siswati	0.103 (0.012)	0.091 (0.011)	0.091 (0.011)	0.092 (0.011)	0.096 (0.011)
Tshivenda	0.120 (0.012)	0.096 (0.011)	0.105 (0.012)	0.110 (0.012)	0.120 (0.012)
Xitsonga	0.068 (0.010)	0.054 (0.009)	0.056 (0.009)	0.060 (0.009)	0.069 (0.010)

  

Panel B: Indirect effects on search status					
All	-0.007 (0.003)				
	15-24 years	25-34 years	35-44 years	45-54 years	55-65 years
Afrikaans	-0.006 (0.003)	-0.006 (0.003)	-0.006 (0.003)	-0.005 (0.003)	-0.005 (0.003)
English	-0.006 (0.003)	-0.006 (0.003)	-0.006 (0.003)	-0.006 (0.003)	-0.006 (0.003)
isiNdebele	-0.012 (0.004)	-0.011 (0.004)	-0.011 (0.004)	-0.011 (0.004)	-0.011 (0.004)
isiXhosa	-0.006 (0.003)	-0.005 (0.003)	-0.005 (0.003)	-0.006 (0.003)	-0.007 (0.003)
isiZulu	-0.006 (0.003)	-0.005 (0.003)	-0.005 (0.003)	-0.005 (0.003)	-0.006 (0.003)
Sepedi	-0.009 (0.004)	-0.008 (0.003)	-0.008 (0.003)	-0.008 (0.004)	-0.010 (0.004)
Sesotho	-0.008 (0.004)	-0.007 (0.003)	-0.008 (0.003)	-0.008 (0.003)	-0.008 (0.003)
Setswana	-0.009 (0.004)	-0.009 (0.004)	-0.009 (0.004)	-0.009 (0.004)	-0.010 (0.004)
Siswati	-0.015 (0.005)	-0.013 (0.004)	-0.013 (0.004)	-0.013 (0.004)	-0.014 (0.005)
Tshivenda	-0.017 (0.005)	-0.014 (0.005)	-0.015 (0.005)	-0.016 (0.005)	-0.017 (0.005)
Xitsonga	-0.010 (0.004)	-0.008 (0.003)	-0.009 (0.004)	-0.009 (0.004)	-0.010 (0.004)

Notes: The marginal change induced solely by social networks is calculated as  $\frac{1}{(1 - \alpha C_k)} - 1$ .

For indirect effects on employment probability,  $\alpha = 0.037$  (from Table 6, Column 2). For indirect effects on employment search status,  $\alpha = -0.006$  (from Table 6, Column 4). Standard errors in brackets, calculated using delta method.

Table 8: Network co-efficients and indirect network impacts on employment and worker discouragement for specified groups

Specification/Group	Probability that individual is employed	Indirect network effect on employment status	Probability that unemployed individual is discouraged	Indirect network effect on reducing worker discouragement
	$\alpha$		$\alpha$	
Exclude Xhosa and Zulu speakers	0.056*** (0.009)	0.082 (0.011)	-0.007* (0.004)	-0.009 (0.004)
Has access to telephone	0.043*** (0.008)	0.053 (0.009)	-0.005* (.003)	-0.006 (0.000)
Does not have access to telephone	0.015* (0.008)	0.022 (0.010)	-0.000 (0.000)	0.000 (0.000)
Individual is male	0.022*** (0.005)	0.026 (0.006)	-0.004 (0.005)	-0.005 (0.005)
Individual is female	0.050*** (0.011)	0.065 (0.013)	-0.006 (0.004)	-0.007 (0.005)
Individual lives in rural area	0.032*** (0.006)	0.055 (0.008)	0.000 (0.000)	0.000 (0.000)
Individual does not live in rural area	0.067*** (0.010)	0.066 (0.010)	0.003 (0.003)	0.003 (0.003)
Individual has disability	0.033*** (0.011)	0.045 (0.014)	-0.025 (0.018)	-0.031 (0.020)
Individual does not have disability	0.036*** (0.008)	0.045 (0.009)	-0.005 (0.003)	-0.006 (0.003)
Individual relocated	0.025*** (0.007)	0.022 (0.006)	-0.012 (0.007)	-0.010 (0.006)
Individual did not relocate	0.044*** (0.009)	0.058 (0.010)	-0.004 (0.003)	-0.005 (0.004)

Significance: \*\*\*=1%; \*\*=5%; \*=10%

Std errors (in brackets) in regressions control for clustering at magisterial district level, and include controls for personal and household characteristics, as well as age-language cohort fixed effects, and magisterial district fixed effects. Std errors for indirect effects calculated using delta method.