Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa

Jeremy Magruder

January 3, 2007

Job Market Paper

Abstract

This paper examines the role of parents in helping children secure employment in South Africa. I use longitudinal data on young South Africans to examine the covariance of children's employment with parent's usefulness in job search. I find that fathers serve as useful network connections to their sons (and not daughters), but that mothers do not seem to be useful network connections for any of their children. The possibilities of specific human capital, correlated networks, and multiple non-labor force choices are considered and do not alter the estimated father-son effect. The results suggest that part of the decline in observed economic mobility in South Africa is attributable to the increase in importance of small-scale networks along with high unemployment.

1 Introduction

When jobs are scarce, the system by which they are allocated becomes a prime determinant of inequality and economic mobility. Under perfect competition, the most able are allocated jobs. Informational problems, however, may compromise this insight, and a recent literature documents the importance of networks in mitigating these problems and allocating jobs. Since full networks are rarely observed in data, this literature has looked for variation in the size or well-being of groups which seem likely to be correlated with an individual's personal contacts, such as people who live near the individual or are of the same ethnicity or both. Since this literature only identifies groups correlated with networks rather than actual network members, it cannot speak to the scale of networks which are relevant to individuals. However, there is reason to suspect that the scale of the relevant network plays an important role in the contribution of the network regime to economic mobility. If small scale networks – for example, immediate families – are quite important, then networks may work to concentrate access to employment as only people with employed family members learn of job openings. The end result of these small scale networks could be a much more persistent state of unemployment, as individuals lacking this network-based social capital have less access to jobs.

Economists studying mobility have long been aware of this fact. From early papers, economists mention the possibility of "connections" (e.g. Becker and Tomes 1979) as a potential source for the intergenerational correlations they describe. However, few papers have been able to explicitly test for the importance of parents as a source of job information or reference. In part, this neglect is practical: these studies have primarily been concerned with lifetime outcomes, like permanent income, for which networks can not be separated from correlated individual fixed effects, such as preferences, genetic and wealth endowments, and educational investment. However, another reason may be the contexts of these papers – nearly all of the literature is set in the US or Europe, where unemployment is low enough, and labor markets are competitive enough, that having a parent who represents a good employment connection seems far less likely to be important than the human capital which he invests in you.

South Africa represents a different context, which is more relevant for many developing countries. Unemployment is severe, especially for the majority black population. Depending on the definition used, between 16 and 28% of prime age males are unemployed (between 19% and 33% for blacks). Unlike in other countries with similarly high official statistics, there is no substantial informal sector or subsistence agriculture to engage unemployed adults, so these numbers represent "true" unemployment. Further, while unemployment has remained high in South Africa since at least the late 1970s, the distribution of the unemployment has changed substantially since the fall of apartheid. In particular, the depth of unemployment has become more severe, so that increasingly large percentages of the unemployed have never held a job or have not worked in many years. This seems especially surprising given the increase in *de jure* economic opportunity available to non-whites after apartheid ended. At the same time, education has become a weaker predictor of employment, suggesting that changes in the distribution of human capital have not caused this change in persistence. Legal changes associated with the end of apartheid created a shift in the regime allocating jobs, and anecdotal evidence suggests that networks are now governing job allocation. Quantifying not only the importance of networks to job allocation but also the relevant scale of networks is critical to understanding the distribution of wealth and economic opportunity in South Africa.

This paper sits at the nexus of the network and intergenerational correlations literatures. Using a panel data set of young adults in Cape Town, South Africa, I ask if parents have become important network connections for their children. Unlike past network studies, this approach looks for fluctuations in the ability of parents as individual network members to provide job information and references, observing that if parents are individually important in securing jobs for their children, then small scale networks must be important. The longitudinal aspect of this panel allows the comparison of intertemporal variance in employment numbers in parents' industries with changes in children's labor force behavior, suggesting that networks may be important. As parents' wealth may be correlated with labor demand in their industries, I take advantage of gender-segregation and geographic specificity in jobs to create two control groups who should experience wealth effects but cannot take advantage of network aid. Estimation reveals that when fathers' industries are hiring, sons are more likely to work if their fathers are in the province, but that sons with absent fathers and all daughters are less likely to work. This negative relationship for individuals whose fathers are not network connections is consistent with the hypothesis that wealth effects are diminishing the labor supply of these young adults. The point estimate is large; when a father's industry grows by 10%, his son is 4% more likely to work if the father is present. In contrast, mothers do not appear to represent effective network connections for either sons or daughters, in part due to differences in the utilization of networks in the different industries in which men and women work in South Africa. These results are robust to the consideration of specific human capital, correlated household and neighborhood networks, and expanding young adults' choice sets to allow multiple non-labor force options such as schooling and leisure.

I open this paper by discussing the empirical literatures on networks and intergenerational correlations, and provide simple examples of the implications of the scale of relevant networks for the persistence of unemployment and how networks can exacerbate intergenerational correlations from other sources described in the literature. I then describe in greater detail the institutional structure and the unemployment situation in South Africa, and describe the evidence for how networks have supplanted bureaucratic job allocation with the liberalization of labor markets following apartheid. Next, using a new panel dataset of young adults in the Cape Town area, I argue that fathers are providing jobs to their sons, but not to their daughters, when they live close enough to be able. I then verify that this effect does not appear to be correlated with specific human capital or correlated networks. I close by discussing the extent to which these paternal connections are decreasing economic mobility in South Africa.

2 Networks, Intergenerational Correlations, and Job Allocation

A growing literature in economics examines the role of networks in spreading job information and changing employment rates. Calvo-Armengol and Jackson (2004) outline a theoretical model which captures the important points. In their framework, individuals are born into a certain exogenous network, with links to a graph of other individuals, some of whom are employed and some are not. Employed individuals hear about jobs with some probability; if they learn of a new job, they tell a network member at random. If that network member is unemployed, he accepts the job; if he is employed, he passes the information on to another network member. Unemployed members of your network are thus your competitors for job information while employed members are suppliers of jobs. Calvo-Armengol and Jackson find that if labor force participation is costly, then initial conditions create large variations in overall employment rates, as individuals in low-employment networks find the benefits of job search to be lower than those in high-employment networks relative to the costs of labor market participation.

Recent empirical support for the importance of networks is strong, and is surveyed in Ioannides and Loury (2004) . Within the economics literature, empirical work has taken two approaches. In the first, researchers look for random variation in the size of an individual network. Munshi (2003) finds that village-level rainfall shocks in Mexico are correlated with migration to the United States, which, combined with past migration patterns gives him an instrument for network size of Mexican immigrants in different cities. He finds that having more migrants from your village a few years earlier boosts individual chances of employment. Beaman (2006) directly tests the Calvo-Armengol and Jackson theory using refugees in the US. Taking advantage of a randomized refugee assignment program, Beaman finds that refugees have a competitive relationship with other refugees settled at the same time and one year beforehand, but learn about job openings from older settled refugees in their ethnic group and city.

The second approach looks for correlations based on geographical distance. For example, Conley and Topa (2002) examine spatial correlations in unemployment with respect to several distance metrics, including miles, travel time, occupational distance, and ethnic similarity measures across census tracts. They find that ethnic similarity is the most powerful predictor of correlations in employment outcomes, with physical distance being the only other measure that can explain spatial correlation. A different approach is taken by Bayer, Ross, and Topa (2005), who investigate whether individuals who live in the same census block are more likely to work together than they are to work with individuals from nearby blocks. Excluding individuals who work and live in the same block, they find extremely large effects: the odds of a pair working together are boosted by 50% if individuals live in the same block as opposed to a nearby block.

Due to data limitations, the literature outlined above has been focused on finding networks correlated with the "true" effective networks relevant to individuals, either through geographic or ethnic proximity, rather than identifying the effect of actual network members. While this approach is successful in identifying whether or not network effects exist, it has limitations as it does not identify the scale of the actual network. In particular, the point estimates convey little meaning, as they are average effects of some fraction of actual networks and of a group of individuals not in the network. Similarly, policy predictions are hard to come by from this approach, as the effective unit of networks is unidentified – if larger networks help individuals get jobs, we want to increase the size of the true network, not the ineffectual group who is correlated and identified due to available data. More relevant to this study, implications of networks for inequality and mobility can not be explored via a strategy which observes only correlated networks, as these implications differ strongly if the true network is a massive, diffuse group which quickly spreads out information or if it is simply small groups of family members passing jobs amongst themselves.

A simple example illustrates the importance of network scale: networks a and b both consist of N individuals, subdivided into families of K. In network a, the family is irrelevant; individuals pass job information at random to other network members. In network b, only family members are network members; individuals pass job information on only to other family members. In period t, fraction p of adults are working, and working adults hear about jobs with probability α. For simplicity, jobs are never destroyed. In network a, everyone is equally mobile: each period, $Np\alpha$ new jobs are learned of and divided among $N(1-p)$ individuals, so every unemployed person has probability $p\alpha/(1-p)$ of learning of a job. In the long run, everyone becomes employed in network a. In network b, each unemployed person's probability of learning of a job depends on the fraction of people in his family who are working; that is, if F people in his family are working, than he has probability $F\alpha/(K-F)$ of learning of a job. Clearly, unemployed people in fully unemployed families have no chance of finding employment, while unemployed individuals in a highly employed family are likely to find employment before long. Even in the long run, some families remain completely unemployed, while others become fully employed. In this example, the implications for network scale are immediate: not only are long run opportunities better in network a than in network b, unemployed individuals face more equal chances of employment. That is, unemployed individuals in high-employment families in network b have very good chances of being employed in the next period, while unemployed individuals in low or zero employment families in network b have very poor or zero chances of becoming employed. In contrast, everyone in network a faces the same chance of finding a job, which rests between these two extremes. For some individuals in network b, unemployment is extremely persistent, whereas for others it is a minor inconvenience. Though the assumptions in this example are extreme, the intuition carries through if individuals only treat their families with preference, that is, if they are more likely to tell family members about jobs than unrelated network members.

In fact, there is some indication that closely related individuals may be especially important network nodes, particularly in high-unemployment settings. Munshi and Rosenzweig (2003) are aware of this possibility in their study of social pressures created by sub-caste level networks in India, and condition on father's occupation to be sure that the sub-caste (and not family) is the right unit of analysis in their context. While the father's occupation does not drive the sub-caste effect, it remains a significant contributor to the difference in schooling between males and females, suggesting that family-level networks may remain important in India. Granovetter (1983) surveys empirical evidence from sociology which finds that disadvantaged groups in the US (like racial minorities) are more likely to use "strong" ties to find jobs, i.e. closer relationships are more important for employment in their networks. Loury (2006), using the NLSY in the US finds that 10 percent of men found jobs through prior generation male relatives, which represents more than 1/5 of all network help, suggesting that a few family nodes matter. Moreover, these referrals from older male relatives are the only type of network help which results in higher wages for young men. Further, she finds that young women almost never receive job offers from male contacts, and that jobs received from female contacts offer no wage boost for either gender. In a job-rich environment like the US, wages are perhaps a more relevant indicator of a valuable network than labor force status, and her results are qualitatively similar to mine. Finally, Kramarz and Skans (2006) have access to a database on all 16-65 year old workers earning income in Sweden from 1985-2002. They investigate whether children are more likely to get a job at the same plant as their parents than their classmates are. They find a very strong effect, particularly in high-unemployment and low-skill settings. However, since their interest is in a one-time event (the first stable job received by these young adults), they are unable to perfectly control for factors which are doubtless correlated between parents and children, such as geographical location, specific human capital, preferences, and abilities. In order to investigate the causes of economic mobility, it is necessary to separate these events out, and a literature which has largely ignored networks has striven to do so.

In fact, economists have examined intergenerational correlations in outcomes since Galton (1869) looked at correlations in height and in wealth. In a seminal paper, Becker and Tomes (1979) discuss the implications of dynastic utility where each generation faces a trade-off between own consumption and the consumption of the next generation, purchased through either investing in human capital or transferring physical capital. Children receive their parents' gifts as well as a "luck" endowment which will be correlated with their parents' endowments. With concave utility, richer parents invest more in their children. Between investment choices which are a function of economic success, gifts, and correlated "luck," several avenues for intergenerational correlation are discussed. Noteworthy to this investigation, Becker and Tomes acknowledge the importance of parental "connections" in intergenerational correlations, but do not explicitly consider them. Theoretical investigations since have followed their example and focused on different mechanisms for intergenerational correlation. For example, Banerjee and Neuman (1993) find that credit constraints can imply an unequal and immobile long-run distribution of wealth if investment opportunities are non-convex, while Mookherjee and Ray (2002) find that credit constraints imply that equality of outcomes is unstable if occupations are diverse simply due to differential optimal investment in children. If families are important as network connections, then the occupation-based effects that these authors highlight will become stronger.

Since Becker and Tomes work, many empirical investigations have sough to disentangle the relative contribution of each of these inputs and to estimate the elasticity of childrens' permanent income with respect to parents' income. While a literature summary is beyond the scope of this paper, Solon (1999) surveys the literature, and creates a stylized model which makes transparent several potential avenues for correlation. Specifically, parents face a trade off between consuming themselves and investing in their childrens' consumption. In addition to the wealth effects and correlated luck discussed above, Solon adds neighborhood sorting, by which children of the wealthy find themselves in neighborhoods with better schools, peers, etc., which may also contribute to intergenerational correlation. Nearly all of these empirical investigations take place in the US and Europe, and they often find an elasticity of one generation's permanent income on the next generation's on the order of .2-.4. Because this literature concerns itself with long term outcomes, they are largely unable to highlight the roles of specific agents who remain constant in a network, and leave as an open question the role of inherited social capital and network referrals. However, this is a significant oversight: network based intergenerational correlations require creative governmental policy. The policy prescriptions to counteract other sources of correlation are better known. For example, underinvestment in education can be remedied through government policy and preference-based correlations are not important to a utilitarian social planner. In contrast, network connections are much harder for policy makers to create and maintain the potential to inhibit economic opportunity. If these are important, spreading job information and preventing nepotism in job hires may be valuable contributions of public policy. This paper will try to correct this oversight, taking both the intergenerational literature and the network literature into account, and, using time-variation within individuals' behavior, disentangle network-based intergenerational correlations from the other sources described above.

A simple adaptation of Becker and Tomes (1979) model (adapted via Solon 1999) suggests immediately the importance of networks for intergenerational mobility. Parents in generation $g-1$ face a choice over their consumption and investment in their children in generation g. Working individuals receive a wage equal to their human capital; for simplicity human capital is unrelated to experience. Parents have Cobb-Douglass utility over their own consumption and the consumption of their child,

$$U\left(C_g, C_{g-1}\right) = \alpha \ln\left(C_g\right) + (1-\alpha) \ln\left(C_{g-1}\right) \tag{1}$$

. The parents budget constraint is

$$y_{g-1} = H_{g-1}\sum_t W_{g-1,t} = C_{g-1} + I_g$$

$W_{g-1,t}$ is an indicator indicating whether parents are working in time t, and they receive a wage of their human capital H_{g-1} if working. They spend that wage on their own consumption and investment in their children, C_{g-1} and I_g.

In turn, the young adult finds himself with human capital $rI_g + E_g$, where E_g is his luck endowment and r the annualized return on invested human capital. Over his lifetime, he therefore earns

$$y_g = (rI_g + E_g)\sum_t W_{gt} \tag{2}$$

Where r represents the rate of return on human capital investment and E_g the error in human capital for generation g. Naturally, both of these only result in income when the young adult is working, $W_{g-1,t} = 1$. The first order conditions quickly lead to the optimal investment choice, $I_g = (1-\alpha)\, y_{g-1} - \frac{\alpha}{r}E_g$. With Cobb-Douglass utility, investment choices are not related to the future labor force behavior of the child (as long as $\sum_t W_{gt} > 0$); in a more general utility model the parent may consider his sons future work opportunities in making investment choices.

Hence

$$y_g = (1-\alpha)\,(ry_{g-1} + E_g)\sum_t W_{gt}$$

Next, I allow for network effects, by modeling generation g's labor force status.

$$W_{gt} = \gamma_1 W_{g-1,t} + u_{gt} \tag{3}$$

where γ_1 is the offer rate for a working parent in generation $g-1$ discussed in Calvo-Armengol and Jackson (2004). If parents are important network connections, $\gamma_1 > 0$. If not, $\gamma_1 = 0$. We can then write

$$y_g = (1-\alpha)\,(ry_{g-1} + E_g)\sum_t (\gamma_1 W_{g-1,t} + u_{tg}) \tag{4}$$

Solon's description of this model focuses on the term $(1-\alpha)\,(ry_{g-1} + E_g)$, and the motivations for intergenerational correlation that he considers are contained within it. In contrast, the network based correlations are shown to be multiplying this effect, which is logical: if parents' income resembles their childrens' due to any correlation in human capital, this capital can only earn a return when it is being exercised, i.e. the individual in question is working. That is, all of the constraints to economic mobility reviewed in the literature are exacerbated in the presence of intergenerational networks.[1]

3 Unemployment and Labor Markets in South Africa

The context for this study is South Africa, where unemployment is severe. Table 1 reveals that between 1993 and 2004, the narrow measure which requires individuals to be actively searching ranged from 14-25%, while the broader measure which requires only a reported desire to work hovered between 30 and 40% of Adults. This unemployment is concentrated among blacks, rural people, and the young, although rates remain high for coloureds and urban individuals. Kingdon and Knight (2006) advocate using the broader definition in this context, on the grounds that local wages are more sensitive to broad unemployment than narrow unemployment.

The cause of this unemployment is hotly debated. This paper does not contribute to that debate, but a few institutional suspects emerge and are relevant to the discussion of job allocation. First, formal sector wages in South Africa are very high for a country of that level of development, for several reasons. Capital-intensive production methods were strongly encouraged under apartheid, in an effort to boost wages for white workers (at the expense of unemployed blacks) (e.g. Seekings and Nattrass 2005). These historically high wages have been maintained in the years since through South Africa's union structure and high minimum wages in uncovered industries. Industrial Bargaining Councils extend union arbitration decisions to all firms in an industry; Moll (1996) discusses how this system provides incentives for wages to grow above standard union-arbitration levels, as large firms will agree to excessively high wages so as to reduce the profitability and competition of small and medium size enterprises. High minimum wage rates have been

[1] In this example, network capabilities are unrelated to income. However, introspection and existing literature suggests that there may be some relationship. A more general model would make this explicit and $\gamma_1 W_{g-1,t}$ would be replaced by $\Gamma(y_{g-1,t}, W_{g-1,t})$. One might in principle imagine $\Gamma(\cdot)$ to be positively or negatively correlated with income, as high income managers have more control over hiring decisions while traditional wisdom suggests that close relations are more important in (generally lower paid) blue collar jobs. While a complicated $\Gamma(\cdot)$ could in principle cause networks to either increase or decrease intergenerational correlations, the basic intuition that networks change the structure of intergenerational correlations goes through regardless of the wealth derivatives of Γ.

established in all industries not covered by union arbitration in recent years. These high formal sector wages increase incentives to queue for a scarce highly paid job, and reduce the costs of unemployment. Looking at union effects, Schultz and Mwabu (1998) find that regions with high union wages relative to non-union wages are associated with higher youth unemployment, and that halving this wage differential could increase employment rates of young African men by 15 percent (3.3 percentage points). Secondly, unlike other countries with similar levels of official employment, there is no appreciable informal sector or subsistence agriculture. Agriculture was consolidated and mechanized under apartheid, and the efficient large farms remain major players in the South African economy, removing subsistence agriculture as an option for rural people.

The absense of an informal sector is perhaps the most intriguing mystery about South Africa. Forbidden under apartheid, the informal sector has grown somewhat in years since, with the fraction of workers who are either self-employed or domestic workers rising from 19% in 1993 to 25% in 2004 , but it remains anemic compared to other countries with similar unemployment and poverty rates (e.g. Kingdon and Knight 2006). The South African Labour Force surveys make a special effort to capture informal employment, where respondents are given a battery of questions about different types of work that they may have done "even for only one hour," including unpaid work in a household business, farming for subsistence, begging, etc. Despite this effort, very few people report any behavior resembling the common place informal economic activity found in other developing countries. Kingdon and Knight (2004) find that the unemployed are worse off in most measurable dimensions than informally-employed individuals, and hypothesize that barriers to entering the informal sector must exist. Another possibility is that waiting for a formal sector job is preferable to informal employment, particularly as wages are high enough for those who do have jobs that, combined with a generous governmental pension program, most individuals have a support network in place. Whether by choice or by restriction, unemployment appears to be the default outcome of most individuals who do not find formal sector jobs.

In this context, understanding the system by which formal sector jobs are allocated is critical for understanding the distribution of wealth, opportunity, and mobility. In fact, while unemployment has remained at high and similar levels, this system has changed dramatically. A variety of legal restrictions challenged economic opportunity under apartheid. For rural blacks, job allocation was purely bureaucratic under the old system. In order to be allocated a job, individuals had first to apply at the local labor bureau, who would indicate if they could seek work, where they could seek work, and what sectors they could apply to. Migration was also restricted without proof of gainful employment, so that all employment opportunities were forced to go through one's local bureaucracy (Seekings and Nattrass 2005). Moreover, coloureds and urban blacks faced laws of job reservation (by which particular categories of jobs were reserved for different racial groups), which similarly restricted the opportunity sets available to these groups. Starting in the 1980s, these bureaucratic regulations began to break down, and increasingly firms began hiring friends and relatives of employees. The importance of networks in the new South African economy is attested to by a 1996 survey of manufactures in which 41% of firms report that when they need to fill vacancies, they hire friends and relatives of employees (Standing, Sender, and Weeks 1996).

Another important characteristic of a job allocation regime for mobility is the permanence of a system. One possibility is that whatever criteria the labor market is now rewarding with jobs is more equal, or, equivalently, that individual employment probabilities are similar across individuals, even if they are small. On the contrary, a different allocation regime could now have limited mobility, so that some individuals find work relatively easily and experience frequent employment, whereas others find work only with great difficulty and spend most or all of their lives in perpetual unemployment. The latter appears to be a closer approximation for South Africa. Since 1993, average durations of unemployment have risen sharply, for both men and women. The fraction of unemployed adults who have never worked has increased substantially since 1993, particularly for young adults, accompanying higher unemployment rates. Table 3 reports these unemployment rates as well as the fraction of unemployed who have never worked and the time since the end of their last job for those who have worked, split into adults aged 20-35 and 35-60 in 1993 and 2003. For all young blacks and young coloured men, the rise in percentage who have never worked is truly striking, growing by 20 percentage points for young black males (and 15 percentage points for young coloured men). For older adults, the relevant indicator of persistence of unemployment is duration, and the rise in the fraction who have been unemployed for all groups for at least three years is similar in scale. Further, while unemployment was more persistent for women than for men in 1993, the reverse appears to be true now, especially among the coloured population. Further, a closer examination of table 2 reveals that unemployment has actually become shorter in duration for young adults who have ever worked, which suggests that a new exclusionary criterion has emerged to regulate employment – those who satisfy this criterion work frequently, with short unemployment spells in between, while those who do not have little chance of finding employment. Similar results exist for the white sample, although the fraction unemployed is very small and these estimates are imprecise. Table 3 displays probit marginal effects of a year of primary or secondary schooling, all conditional on urbanization and a quadratic in age, and reveals that this criterion is not education, as observable human capital has become a weaker predictor of employment since 1993. By 2003, the marginal effect of an extra year of secondary schooling on the likelihood of getting a formal sector waged job had fallen by 60-75% for coloured men and for all blacks. Since unemployment did not increase by a similar margin in this period, this indicates a shift in the criteria used to hire workers, away from observable human capital. While education is no longer a variable which determines economic opportunity in South Africa, it has been replaced by something far more restrictive, especially for men.

If networks are now solving the information problems endemic to liberalized labor markets, these may be compromising the ability of different groups to take full advantage of the opportunities generated by a free labor market. Seekings and Nattrass (2005, p. 282) discuss the distributional implications of using network channels for job allocation in greater detail: "It is surely the case that the number of discouraged unemployed in South Africa is large in part because vacancies are so often filled using [network] channels. Thus it is likely that, among the unemployed, there are some with good prospects for employment and others with poor prospects and that the former are more likely to be members of households with working members." This argument is one of network scale: if household members are inportant network connections, then network scale must be small indeed and the sorts of implications for persistent unemployment may resemble the severe results from network b in the simple numerical example given earlier. Indeed, if networks are now regulating employment in South Africa, we may expect that these networks would be used first and foremost to generate intergenerational correlations, as parents have both selfish and altruistic reasons to be concerned first about the employment of their children. However, there has been extremely little work on the interaction between networks and intergenerational correlations, and integrating these two literatures is central to analyzing the South African labor market.

4 Data

The Cape Area Panel Study is a random sample of 4758 young adults aged 14-22 in 2002 who live in the Cape Town Metropolitan Area. Located in the South-Western corner of South Africa, Cape Town is the second-largest city in the country. These young adults were interviewed first in 2002. A subset of 1360 young adults were reinterviewed in 2003, with the remainder reinterviewed in 2004, and all were reinterviewed in 2005. At each interview after 2002, a monthly calendar of past behavior was collected, allowing the creation of a full panel with all behavioral variables (e.g. working, schooling enrollment) taken to be behavior in September of that year (chosen to coincide with the labor force surveys described below). Attrition is a problem in this study, due to the migratory nature of these young adults; 11% of these young adults disappear before data can be recorded for September 2003 and 26% before data can be recorded for September 2004. Attrition was especially problematic for the whites in this sample, where 22% were lost by the end of the 2003 and 44% by 2005 (many of these moved out of country, most of the rest out of the Western Cape province). Due to the substantial attrition in the white sample, I restrict my analysis to blacks and coloureds; 10% of these people are lost to attrition by the end of 2003 and 23% by 2005. I examine the effects of attrition in robustness checks below.

Table 4 reports summary statistics for this sample. Similar to the national employment tables presented above, few of these young adults are working, with only 25% of males and 18% of females working. This difference largely reflects the difference in the percentage of men and women who found jobs through networks; 6 or 7% of both men and women found jobs through their own means (e.g. applying at factories, sending out CVs, etc.), whereas 14% of men compared to 8% of women found a job after a friend or family member told them about it or referred them for it. Also worth noting is that many of these young adults have fathers who are either deceased or do not live in the province, and that many of those with fathers in the province do not cohabit with the father at baseline. The geographical heterogeneity within families is due in part to the migration restrictions which were lifted only at the end of apartheid, so that many in this sample are recent migrants. As a large urban area, Cape Town is a destination for migration, so absent fathers are often either those who have not migrated yet or who have returned to their more rural homeland area. Fewer young adults have mothers who live away, though this still represents a large fraction of my sample. In the Appendix, I describe the construction of provincial employment data.

4.1 Intergenerational Correlations in Cape Town

Few studies have investigated intergenerational correlations in developing countries, and the fact that education seems to be of little help in finding a job may make us doubt our prior beliefs about what could matter in securing employment. The CAPS data allows a descriptive and reassuring test, similar to Altonji and Dunn (1991). If intergenerational correlations are important in a network sense, we would hope to see that childrens' industries are correlated with their parents. Tables 5 and 6 report coefficient estimates from systems of seemingly unrelated regressions, where each dependent variable is a dummy for working at baseline in a one-digit industry and the right hand side is dummies for the father or mother being in that industry. Of the ten one digit industries, Industry 2 (mining) is excluded from all regressions and industry 4 (Utility Provision) is excluded from regressions involving females as there are very few observations of young adults working in these industries. Sons industries appear highly correlated with their fathers' while daughters' industries are highly correlated with their mothers' , but there appears to be little cross-gender intergenerational correlation once education, age, and race are controlled for. Men work in very different industries and occupations than women in this context, so this absence of cross-gender correlation is reassuring. Moreover, this seems similar to Loury (2006), who finds in the US that men often utilize older male relatives to get jobs, and that women often use older female relatives, but few use relatives of the opposite gender.

A natural extension of the above analysis is to consider occupational correlations as well as industries. While industries seem more likely to be the relevant correlation for networks (as network members can pass on job information at the plant regardless of the occupation), other intergenerational correlations are more likely to result in occupational similarity. In particular, specific human capital and preference-based effects seem likely to create an intergenerational correlation in occupation. Tables 7 and 8 report coefficient estimates for occupations from seemingly unrelated regressions, following the methodology used for industries. The difference in the occupation-industry correlations are striking. While industries are correlated along gender lines, occupational correlations are completely absent for fathers with children of both genders and strong for mothers with all children. Below, when I consider whether the transmission of specific human capital or preferences could be driving my network results, I will take advantage of the fact that occupations are not correlated between fathers and sons.

5 Empirical Strategy

Two approaches to estimating this problem are taken. In the first, I consider only the choice to work or not work. In the second, the choice set is expanded to include the possibility of schooling. The advantage of the second approach is completion; the first approach, however, has the advantages of more power and better data – enrollment data in this survey are suspect, as many individuals remain in school despite matriculating to the next grade infrequently. The linear probability model used in the binary choice problem also has the advantage of easily interpretable coefficients, while the multinomial conditional logit used in the second approach can not deliver consistent marginal effects without strong additional assumptions due to the fact that individual fixed effects are unidentified.

To estimate probabilities of working, I begin with a linear probability model. Inserting individual heterogeneity into Equation 3, young adult g in family f faces

$$W_{fgt} = \gamma_{1ft} W_{fg-1,t} + u_{fgt}$$

I allow for potential interactions of parental wealth and child endowments in the error term by specifying

$$u_{gt} = \gamma_2 y_{fg-1,t} + \beta X_{fgt} + \xi_{fg} + \delta_t + \nu_{fgt} \tag{5}$$

where $y_{fg-1,t}$ is the parent's income in family f at time t, X_{fgt} is a vector of covariates including age and gender, ξ_{fg} is a fixed effect for the young adult in generation g (which is doubtless correlated with his luck endowment), δ_t is a time-fixed effect describing the strength of the labor market in year t and ν_{fgt} is a person-time-specific error.

Combining these equations

$$W_{fgt} = \gamma_{1ft} W_{fg-1,t} + \gamma_2 y_{fg-1,t} + \beta X_{gt} + \xi_{fg} + \delta_t + \nu_{fgt} \tag{6}$$

That is, for family f, the young adult's probability of working is related to his parent's offer rate at time t if his parent is working, his parent's income, a vector of covariates, a time fixed effect, an individual fixed effect and the error term.

Let ϕ_{it} be the offer rate in industry i at time t. If $i = 0$ denotes unemployment, $\phi_{0t} = 0\ \forall t$. Then, I specify that, for parent f who works in industry i,

$$\gamma_{1ft} W_{fg-1,t} = \gamma_1 \phi_{it} + \xi_{fg-1} + \nu_{fg-1t} \tag{7}$$

that is, a parent's offer rate is equal to the time t offer rate in the industry in which he works at time 0, a parent fixed effect, and a parent-time-specific error. Hence,

$$W_{fgt} = \gamma_1 \phi_{it} + \gamma_2 y_{fg-1,t} + \beta X_{fgt} + \xi_{fg} + \xi_{fg-1} + \delta_t + v_{fgt} + \nu_{fg-1,t} \tag{8}$$

Taking fixed effects at the individual level eliminates the parent and son time-invariant effects, ξ_{fg} and ξ_{fg-1}. Using fixed effects on this equation generates the baseline estimation

However, two immediate concerns need to be addressed. First, the offer rate in an individual industry is unobserved. In practice, I use log provincial employment in the two digit industry that the parent is working in to proxy the offer rate. In even the simplest model, employment is a function both of labor supply and labor demand. Hence if industry-specific labor supply increases, we would expect employment in that industry to increase. Simultaneity is not a concern as individual labor supply decisions are too small to impact provincial employment statistics. Trends in overall labor supply are picked up by the time trends utilized, and time-constant individual labor supply components are captured by the fixed effects. Nonetheless, I will be picking up labor supply effects if innovations in industry-specific labor supply are correlated with the industry in which one's parent works (at baseline). That is, if sons of construction workers suddenly tend to want work in construction but not manufacturing more in 2003 relative to 2004 (and sons of manufacturing workers do not), then labor supply is contaminating my estimates. I assume that this is not the case, that is, that trends in industry-specific labor supply are uncorrelated with the industry in which parents are employed.

Secondly, if permanent income is uncertain, then even if I could observe the offer rate perfectly

I would face the problem that parents' welfare will be correlated with this offer rate – if an industry shifts its labor demand outwards, its employees may receive higher wages. At the least, they are more likely to continue working. This is reflected in the coefficient on parent's wealth, γ_2, in the above model. To address this issue, I use two "control groups" of people who can not use individual parents as useful connections. I established above that young women's industry of employment is uncorrelated with their fathers'. Hence, daughters should not realize the impact of paternal connections, but should still be sensitive to permanent wealth effects. Secondly, job information is geographically concentrated – one cannot accept a job at a plant at a location too distant from where one lives. Therefore, both sons and daughters whose fathers live in a different province should not be affected by the job information, and only affected by the correlated implications for the fathers. This approach generates an overidentifying restriction to test: daughters whose fathers live nearby should look no different from those whose fathers live away in terms of sensitivity to trends in fathers' industries. Symmetric tests can be considered for mothers and daughters versus mothers and sons, although data limitations discussed below eliminate the possibility of maternal proximity as a test. All regressions presented in this paper are conditional on individual and year fixed effects.

6 Estimation Results

Table 9 reports the results of this estimation for fathers. Sons appear to be strongly impacted by trends in employment in their fathers industries. The point estimate is quite large: if employment increases by 10% in one's father's industry, then the son is 4% more likely to work. The median industry in this context grows by 8%, so that sons of median workers are about 3% more likely to work each year than those with unemployed fathers. Recalling that 25% of male-years are spent working in this sample, this effect is large. Consistent with the overidentifying test, daughters whose fathers live away look the same as those whose fathers live in the province, in dramatic contrast to sons. Sons whose fathers live away and all daughters seem to be negatively effected by increases in fathers employment, consistent with the labor supply hypothesis of a negative wealth effect for employment. Also, the father-here coefficient reaches conventional significance levels, and the joint test for a male coefficient is very precisely estimated (as is the joint test for a female coefficient). Comparing males to females, the same story is told, and more precisely: sons' employment responds strongly and positively to trends in their fathers' industries, while daughters are less likely to work when their fathers' industries are doing well, suggesting that wealth effects make young adults less likely to work in this age group. Moreover, these effects are not diluted by allowing fixed effects for years of completed education and age or separate time trends by race, as shown in columns five and six.

A further test of the theory is to examine the difference between jobs received through network connections and those found through young adults' own effort. Though unfortunately I cannot observe exactly who got the job for the young adult, the CAPS data does allow me to differentiate between young adults who received jobs through network connections (they report either that a household member or friend or relative outside of the household told them about the job or got them the job at their workplace) and those who found jobs through sending out CVs or inquiring at factories. In table 10, I create two variables: Net job is equal to a 1 if and only if the young adult is both working and reports a network member's aid in finding the job and is a 0 otherwise. Self job is defined similarly for jobs found through non-network-based search means. Columns one and 2 present regressions similar to the baseline with these variables. Columns 3 and 4 observe that the problem is in fact more complex than one of binary choice: individuals who have pre-existing self jobs are less likely to accept offered network jobs, for they would only accept offers better than the one they currently have. In fact, the true effect of interest is whether an individual has access to jobs or not, and this network access is unobserved if the individual has already secured a job through other means. As a result, columns 3 and 4 only include individuals who either have the type of job in question or are not working as a first-pass solution to this problem. We may be skeptical about the quality of self-reports on how jobs were acquired; yet, though the estimates become less precise, the coefficient pattern appears the same for network jobs as it does for employment status, and there appears to be no relationship between self-sought jobs and trends in employment in the fathers' industry, precisely as we would expect if trends in employment in the father's industry are capturing network effects.

Mothers' industries are correlated with daughters' industries just as fathers' are with sons', so it is natural to wonder if similar network effects are seen between mothers and daughters as between fathers and sons. Caution must be taken with this analysis, as very few mothers both live in a different province and are actually working. In fact, there are only 30 young men and 24 young women whose mothers live in a different province and who have time-variance both in their own labor force status and in employment in their mother's industry, which makes the here-away comparison highly suspect for mothers and highly sensitive to individual outliers. Moreover, the potential for attrition bias appears grave here: of the black and coloured sample whose mothers live in a different province, 40% attrit by 2005, as opposed to 18% of young adults whose mothers live in the same province. As a result of these data limitations, I restrict the analysis of mothers to average son and daughter effects.

Table 11 reports coefficients from similar linear probability models to the father-son baseline. As is immediately obvious, there is no discernible effect on daughters of changes in mothers industries. In fact, the coefficients on sons appears larger than daughters, though it is indistinguishable from zero and dwarfed by the father-son effect. Looking for likelihood of network employment, we find even smaller point estimates of mothers' effects, still all indistinguishable from zero. The finding from this analysis is that South African daughters do not benefit from their mothers as network connections, which recalls Loury's (2006) absence of a wage boost for women who get jobs with help from their older female relatives, and is consistent with the higher persistent employment return to education for South African women. Several possible motivations could be given in this context for the lack of effect. First, women work in different industries than men, and it may simply be that the industries in which women work rely less heavily on networks. Mechanically, women work in fewer industries, so that there is less variance in mothers' employment than in fathers'. Three of the four main industries that South African women work in are highly informal (domestic work, clothing and textile manufacturing, and self-owned business, the fourth industry is retail), and it may be true that employment statistics reflect labor demand more poorly in these industries and that employment status is a less relevant statistic in these industries, due to the greater possibility of underemployment. Similarly, employment may be less scarce in these industries, so that a daughter could work at her mother's store without the need of a job opening, again disrupting the link between labor demand fluctuations and employment. Unionization in South Africa is dominantly male, and these jobs have the highest wage premia(Schultz and Mwabu 1998), which suggests that job rationing may be strongest in male industries. These results recall Loury's (2006) conclusion that, as women work in fewer high-paid jobs, they are less frequently privy to wage offers which are substantially better than other sources of job information available to young women.

We can test the hypothesis taht networking may simply be less important in the industries in which females work in South Africa. Using the CAPS data, a suggestive decomposition can be constructed. Baseline data allows the construction of shares of young adults who got jobs with networks in each gender-industry group. That is, for each gender g and two-digit industry i, we can construct the fraction of young adults of gender g who are working in industry i and who had network help in finding the job. This gives a measure of how important networks are to securing jobs for people of each gender in each industry. Comparing the difference in the importance of networks in fathers' industries for boys to the importance of networks in mothers' industries for girls gives the total difference in importance of networks for finding work. We can decompose this difference into two parts: first, we can examine the difference in importance of networks in fathers' industries from the importance of networks in mothers' industries for both genders. This gives an estimate of the difference in networkability for mens and womens' industries. In turn, the difference in network importance for sons and daughters in the same parents' industry gives a measure of the unexplained part of the smaller network effects for females. As industries are measured imperfectly (at just the two-digit level), we might reasonably think of this as an overestimate of the unexplained portion. Table 12 reports this decomposition, and reveals that daughters who work in industries that mothers work in are about 10 percentage points less likely to report network help than sons who work in industries that fathers do. Networks seem less important in mothers' industries than in fathers' for both sons and daughters, and this difference accounts for over half of the difference between mother-daughter and father-son pairs. In turn, daughters are across the board less likely to receive network help, which accounts for the remainder of the difference. It appears that at least part of the reason that no mother-daughter network effect is estimated is due to networks being less utilized in women's industries.

6.1 Specific Human Capital and Correlated Networks

The strong results for fathers and sons require further investigation to eliminate the possibility of type one error, and indeed, two sources of potential endogeneity may confound the above analysis. First, if sons supply labor to only their fathers' industries for non-network based reasons, then the above result could be obtained. That is, if young adults have industry-specific human capital from a lifetime of learning from their fathers which makes them only qualified or only desiring to work in the same industry as their fathers, then their employment will react only to labor demand in that industry without network effects (and the identifying assumption on labor supply may be incorrect). Fortunately, the CAPS dataset allows a direct test of this hypothesis. Specifically, respondents are queried about the industry that their father worked in "most of the time [while] you were growing up" in addition to the question of what industry your father is working in now. Unsurprisingly, these variables are correlated, but there is substantial variation; many fathers moved in and out of employment, and among those who were employed in both periods, 35% switched industries. If industry-specific preference or human capital is driving these results, we would expect that to be correlated with employment trends in fathers' historical industries as well as their current ones. That is, whatever learning process leads to this human capital, it seems likely to have occured at earlier ages as well and so young adults should also be sensitive to employment trends in industries that the father used to work in. Column 2 of Table 13 reports this test on a male only sample and we find that, in fact, conditioning on log employment in the fathers' historical industry does not impact the parameters of interest, nor is it itself significant. This suggests that industry-specific human capital is not driving the network effects observed in the baseline estimates.

However, specific capital may not be industry-specific. In particular, the very reasons that we find fathers working in different industries now relative to the past may be that a new industry is now better rewarding the capital which the father and son share. If the old industry still rewards this capital to some extent, then unemployment must be voluntary, which doesn't seem especially appealing given that only 25% of these young adults are working; nonetheless, a check against this hypothesis is possible. A natural classification of skills which are rewarded differentially in different industries is occupation, so the above argument would suggest that the family-specific human capital is now being especially rewarded in the occupation-industry cell that fathers are currently working in. In contrast, networks need not be occupation-specific; a father can learn of or give reference for any opening at the plant he works, not simply the ones in his occupation. Table 13 explores this possibility by aggregating occuptions in two ways. In Column 3, I use a very coarse characterization of occupation, where occupations are divided into skilled white collar, unskilled white collar, and blue collar, while column 4 allows employment at the one-digit occupation within an industry level. Columns three and four reveal that the relevant effect is at the industry, not the occupation-industry level, suggesting that specific human capital is not behind this effect.

This paper was motivated by a discussion of the necessity of capturing the right network scale, and, indeed, a potential problem could exist if fathers' industries are highly correlated with other industries in a young adults' network. That is, your father is someone who often lives in your neighborhood, and certainly someone who often lives in your household. These networks may have nothing to do with paternity and rather be attributable to randomly selecting adults who are close to these young adults. In truth, some of these concerns are more worrying than others for the interpretation of this analysis – economic mobility is impacted by networks in a very similar way if household heads are giving an advantage to their younger generation household members as if fathers do so, although the flexibility of household structure in this region suggest that household-based mobility restrictions may be less binding than paternity-based restrictions. Nonetheless, table 14 examines several possible correlated networks, again restricting attention to males. First, in column 3, I look at trends in the log employment in the industry of the male head of household. In households where the reported head was not an employed male, I define the male head as the spouse of the head, if he is male and working, or the first employed male on the household roster at baseline who is too old to be considered in this study of young adults (older than 22 at baseline). The strong network effect remains on fathers who live in the province, and the point estimate remains about the same (and, in all specifications, remains significant at about the 5% level). In fact, these household effects both seem very small in absolute value and never attain significance, suggesting that the father is of particular importance within the household. Column 4 considers neighborhood effects. The CAPS data includes identifiers for which enumeration area the survey was conducted in, which amount to neighborhoods of an approximately three block square. In column 4, I condition on the fraction of black and coloured males in other households in the same neighborhood who are working as well as the log employment in the modal industry among adult males in the neighborhood. This neighborhood youth employment appears very important for predicting young males' employment status, suggesting that neighborhood trends are quite important, and that this employment rate is capturing some signal about neighborhood effects. Of course, we can not firmly attribute this effect to neighborhood networks since this neighborhood employment rate is likely correlated with many other neighborhood trends. Nonetheless, including this neighborhood effect does not impact the coefficient on fathers or its significance, as a young adult's neighborhood status turns out to be uncorrelated with trends in the father's industry. This suggests that the family network effect is unrelated to neighborhood trends. In contrast, the log modal industry variable appears insignificant. Finally, since the rest of these potential network nodes are by definition located close to the young adult, their proximity was not included in interaction terms. But we may be concerned that the fathers' proximity is correlated in some way with the efficacy of these network nodes, and that I am attributing that correlation to being a fathers' network effect. I consider this in columns 6 and 7. After controlling for potential interactions between father's proximity and network variables, the coefficients on the father effect become somewhat more precise without changing the point estimates substantially. Modal industry effects appear negative when the father is absent and zero when he is present.

Table 15 repeats the basic estimates of table 14, but utilizes the information on how a job was received. Looking at network jobs versus self-found jobs, we find that the trends in fathers employment remain strong for network jobs and absent for self-reported jobs in the presence of the above controls for correlated networks. The significant modal industry results from Table 13 do not appear to be impacting employment through network channels. In contrast, the strong relationship between individuals' working status with local youth employment rates in their neighborhood appears to be largely confined to jobs found with network assistance, suggesting that neighborhood level networks may be important as well.

6.2 Attrition

Attrition is important in this survey, and it is possible that attrition is correlated with network effects as I have estimated them, though the individual fixed effects used in the analysis would eliminate any time-invariant differences. In fact, some of the attrition may even be causal and part of what I attempt to estimate– while fathers who live away are useless sources of job information for sons who remain in Cape Town, they may be helping out sons who disappear from my sample, because they might move to be with their fathers. Indeed, as table 16 reports, attritors are more likely to have fathers living in other provinces than non-attritors, although their fathers tend to be much less employed as well. Therefore, the above analysis would predict they are less likely to be working than their non-attriting counterparts, on average, and indeed they were working less at baseline. To test if attrition could be responsible for my results, I make a variety of extreme assumptions about the attritors' behavior so as to include them in my estimation sample. In column 1 of Table 17, I assume that they are all working. In column 2, I assume that none are working, while in column 3, those whose fathers live in other provinces are assumed to work and those whose fathers living nearby are assumed to not work. Column 4 assumes the opposite. Even under these extreme assumptions, the pattern of coefficients remain the same throughout, and the network effects retain significance. Since attritors work less at baseline than none-attritors, we may think that column 2 represents the most likely case, where none of them are working. In this case, all coefficients remain similar to the baseline.

6.3 Instrumental Variables

Finally, concerned about the identification assumption above, that innovations in industry-specific labor supply are uncorrelated with industries in which fathers work, we can attempt an instrumental variables strategy suggested by Bartik (1991) and Blanchard and Katz (1992). These authors argue that the local share of a national industry is more or less fixed over time, and that industry supply trends are more local while labor demand trends are more national in nature. If so, then taking the average provincial share of an industry and multiplying that by national employment rates gives an estimate of labor demand in the industry. In the case of these two studies, the authors argue that local units are small enough that local trends will not substantially impact national employment numbers. However, with only nine provinces in South Africa, this assumption seems less appealing. Instead, I use the average ratio of local employment to employment in the other eight provinces over my study period, and multiply employment in the other provinces by that rate. The results presented in table 18 reveal that this approach, unsurprisingly, gets less power, but the male-female difference remains identified. Moreover, a regression-based Hausman test (Wooldridge 2002) does not give an ex ante reason to prefer the instrumented estimates over the OLS, lending credence to the identification assumption. An additional advantage of this estimation is that it is immune to the sampling error concerns discussed in the appendix (as sampling error is independent across provinces), and grants reassurance that this effect is not generated by sampling error.

6.4 Multinomial Choice

In alternate model suggests that individuals choose between three options: unemployment, employment, and school enrollment. That is, using V_{fjt} to denote the value to individual f of choice j in time t, individual f will choose option $j \in J$ if $V_{fjt} > V_{fkt}$ $\forall k \in J \backslash \{j\}$. Within the intergenerational correlations model outlined above, we can model the time t valuation of decision j to parent f who works in industry i.

$$V_{fjt} = \gamma_{1j}\phi_{it} + \gamma_{2j}y_{fg-1,t} + \beta_j X_{fjgt} + \xi_{jfg} + \zeta_{jfg-1} + \delta_{jt} + v_{jfgt} + \nu_{jfg-1,t}$$

Unfortunately, standard multinomial choice techniques are biased in the presence of fixed effects due to the incidental parameters problem. Since individual fixed effects are critical to my identification strategy, and I have relatively few time periods, the bias is large and complicates matters tremendously. Fortunately, Chamberlain (1980) describes a multinomial conditional logit which allows fixed effects. This approach utilizes the convenient form of the logistic error term, which allows an analytic marginal density which is conditional on the observed choices in the data. In this model, rather than choosing from their original choice set, individuals choose from an idiosyncratic set of permutations of the choice set, where each potential permutation has the same end distribution of choices as the observed data. That is, if I observe a young adult who goes to school in the first two periods and works in the third, the multinomial conditional logit looks at the probability that he chooses that outcome versus other permutations where he is in school twice and working once. This conditional choice probability does not depend on the fixed effect, as time-invariant fixed effects determine the end distribution of choices but not the timing of those choices. However, this model does have limitations. Most important for my case is that it has less power than the linear probability model considered above, and that it is impossible to estimate marginal effects without ad hoc and arbitrary assumptions, as marginal effects are functions of the unobserved and unidentified individual fixed effect. As with any model with a logistic error term, the multinomial conditional logit suffers from the Independence of Irrelevant Alternatives property. Nonetheless, this approach gives us some idea of whether schooling decisions are impacting the network effects estimated in the previous section. As few (3%) of young adults are both enrolled in school and working, this category is ignored in the analysis and these individuals are categorized as working, so that we have 3 time periods and 3 choices (school, working, and unemployment).

Formally, this approach yields

$$V_{fjt} = e^{\gamma_{1j}\phi_i\left(W_{fg-1,t},t\right)+\gamma_{2j}y_{fg-1}+\gamma_{3j}h(I_{fgt})+\xi_{ijg}+\delta_t+\nu_{fgt}}$$

where ν_{ftg} is distributed according the logistic distribution and I_{fgt} is the parental investment in a child. Since schooling is a choice in this model, years of completed schooling seem a natural control variable. However, past choices of schooling may be correlated with unobserved time-variant heterogeneity: if some secular trend makes some individuals desire schooling more in one year, it may well differentially impact those with different degrees of schooling. In order to avoid any concerns over bias resulting from this endogeneity, I present multinomial conditional logits with and without years of completed schooling. Estimations which include schooling breaks education into having completed at least grade seven (as nearly all members in my sample have completed at least grade six), a linear effect for grades eight through eleven (i.e. some secondary school) and a second dummy for having completed at least secondary school (grade twelve). .

Defining $m_{ftj} = 1$ if the young adult from family f chooses option j in time t and $m_{ftj} = 0$ otherwise, the multinomial conditional logit conditions on $s_{fj} \equiv \sum_t m_{fth}$, which leads to the conditional log likelihood contribution of

$$L = \sum_f \ln \frac{\exp\left(\sum_t \sum_j \left(\gamma_{1j}\phi_i\left(W_{fg-1,t}\right) + \gamma_{2j}y_{fg-1} + \gamma_{3j}h\left(I_g\right) + \delta_{tj}\right) m_{ftj}\right)}{\sum_{d\in B_f} \exp\left(\left(\sum_t \sum_j \gamma_{1j}\phi_i\left(W_{fg-1,t}\right) + \gamma_{2j}y_{fg-1} + \gamma_{3j}h\left(I_g\right) + \delta_{tj}\right) d_{tj}\right)}$$

where $B_f \equiv \left\{d = (d_{11}, d_{12,\ldots,} d_{33})\,|d_{tj} = 0 \text{ } or \text{ } 1,\ \sum_j d_{tj} = 1,\ \sum_t d_{tj} = s_{fj},\ j \in \{1,2,3\}\right\}.$

Table 19 reports the result of the multinomial conditional logit. While this specification does not have enough power to separate the effects of distant fathers from proximate ones, the signs look the same as in the linear probability model. Moreover, for working, the male-female difference in working remains precisely estimated. If we believe that sons and daughters are identically impacted by wealth effects, then this model is overidentified (as girls are picking up the wealth effect), and the significance on the male effect is sufficient to accept the hypothesis that family networks matter, although the point estimate is not correct as this is the average effect of sons with present fathers and sons with absent ones. The average effect of employment in the father's industry on schooling for both males and females appears to be close to zero, and the joint tests reveals no overall male effect. Thus it appears that young adults are not systematically more likely to prefer schooling to unemployment when their fathers' industries are doing well. Individual coefficients are, however, extremely noisy, so firm conclusions are hard to draw. When years of completed schooling are conditioned upon (column 6), the joint male effect appears present, although point estimates appear very similar to the regressions which don't condition on schooling. The education coefficients are also worth discussion in this regression. While individuals with an extra year of secondary school are much more likely to attend one more year of schooling, they simultaneously look no different in terms of their labor force status. Moreover, completion of secondary school only makes young adults more likely to be unemployed rather than working. Of course, labor force participation decisions are likely correlated with education; yet it remains true that while network effects appear huge, education does not appear to be positively associated with employment, consistent with the tiny effects seen in the national statistics presented earlier.

7 Conclusions

Fathers appear to be extremely important network connections for sons in South Africa. A simple example illustrates the implications of this result: a son with a father in construction, the 75th percentile industry in terms of growth between 2002 and 2004, would be an additional 5.3 percent more likely to be working per year if only his father lived nearby. In this high unemployment setting, having a father whose industry is doing only as well as the median industry in that year gives his son a 3.2% larger likelihood of working each year, which is substantial considering that only 25% of the boys are working in this sample. Since 55% of black and coloured young men in this region have absent, unemployed, or deceased fathers, this suggests that the majority of these young adults are at a large disadvantage. Moreover, those with present fathers are already advantaged: as table 20 indicates, households with present, working fathers have, on average, 80% more income than households without present, working fathers. These families also tend to be the ones advantaged in terms of physical human capital, as children are more frequently enrolled in school and perform slightly better on an IQ test administered by the CAPS team. At the same time, young women do not appear to benefit from these network connections. This result resembles Loury's (2006) findings for the US, and the majority of the male-female difference appears to be attributable to the differences in industries in which men and women work.

The importance of networks forces the immediate (and not uncontroversial) conclusion that unemployment in South Africa is not entirely voluntary – unless jobs are truly scarce, variations in labor demand should not be affecting young adults differentially. The long term prognosis is severe: a characteristic which the majority of young adults do not have and can not obtain is very helpful for employment. Indeed, a limitation of this study is that this problem is so severe that there are no easy policy recommendations, and government may be forced to adopt some creative policy. For example the South African government could encourage the spread of job information and try to prevent discrimination against unconnected, qualified individuals through anti-nepotism laws. More certainly, the policy implication of this work is that the luxury of using government policy to sustain high wages at the cost of high unemployment is very costly. Rather than simply creating unemployment, these policies create long run poverty traps which are inherited by children.

Finally, this study emphasizes the limitations of utilizing correlated groups as proxies for networks. It appears that fathers are individually important connections for sons. If trends in correlated groups affect fathers as well, then a correlated groups strategy may lead to the incorrect inference that a large network is important for job procurement when in fact only a few closely related individuals matter very much. Without understanding the relevant network scale, it is impossible to understand the role of networks in the propagation of inequality, and difficult to derive welfare conclusions from the statement that networks are important. This study has an immediate implication for survey design: if investigators want to delve into the roles of networks in their context, the development of survey instruments to capture individual network members is an important priority.

8 Appendix: Employment Data Construction

Provincial employment data used in the empirical specification (and summary statistics for the economy as a whole given above) are constructed by the author using two-digit industries from the September Labour Force Surveys in 2002-2004 using the sampling weights calculated by Statistics

South Africa. These are nationally representative samples of a rotating panel survey, which each surveys 67836 to 73797 adults. To alleviate sample size concerns, occupation-industry employment numbers are constructed by estimating the fraction of a two-digit industry in each occupation in each year using national data and multiplying the overall employment numbers in that industry by that fraction. The employment data are quite noisy at the two digit levels, especially for relatively small industries. Because the analysis will focus on log changes, the possibility of a small industry not being found in one period due to sampling error could create large outliers. As a result, I observe that my naive employment estimates for industry i at time t, $\widehat{Emp}_{it} = w_{it} \left(Emp_{it} + u_{it} \right)$, where u_{it} is sampling error, w_{it} represent sampling weights, and Emp_{it} represents the number of employed individuals observed in the national survey. Fortunately, we know the distribution of this sampling error: drawing randomly from the population, means that if share p_i of the population works in industry i and I sample n people, then each sampled individual faces a multinomial distribution with probability p_i of being in industry i. Employment is a sum of binomials, which means that $u_{it}\tilde{} N\left(0, p_i \left(1 - p_i\right) n\right)$. However, employment estimates face a lower bound of zero, and hence in expectation I underestimate true employment for small industries as the sampling error is truncated. As a result, I correct my estimates so that u_{it} is mean zero, by estimating the mean sampling bias in small industries, using mean sampling weights for each year and average fraction of the population found employed in industry i from 2001-2004 as an estimate of p_i. The analysis presented is not sensitive to small changes in p_i estimates.

However, the presence of sampling error suggests another concern. Because there is some error in my estimates of employment, the coefficient may be biased downward due to attenuation bias, and the standard errors may be biased downwards as well (e.g. Murphy and Topel 1985). Indeed, this problem is present in any study which uses macro statistics constructed from surveys. Fortunately, the variance of this sampling error is very small: using the delta method derives the distribution for the error in log employment to be $w_{it}^2 n p_{it} \left(1 - p_{it}\right) / Emp_{it}^2$, which is small whenever employment in the industry is larger than the industry-specific sampling weight. As argued above, this error term is not normal in finite samples due to truncation for small industries, however, we can approximate it by assuming normality and estimating p_{it} from my employment estimates. This allows a Murphy-Topel (1985) correction term to the standard errors as a robustness check. Table 21 reveals that this correction is very small, and the results presented above carry through.

References

Altonji, J. G., and T. Dunn (1991): "Relationships Among the Family Incomes and Labor Market Outcomes of Relatives," *Research in Labor Economics*, 12, 269–310.

Banerjee, A. V., and A. F. Newman (1993): "Occupational Choice and the Process of Development," *The Journal of Political Economy*, pp. 274–298.

Bartik, T. J. (1991): *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

Bayer, P., S. Ross, and G. Topa (2005): "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," NBER Working Paper No. 11019.

Beaman, L. (2006): "Social Networks and the Dynamics of Labor Market Outcomes: Evidence from Refugee Resettlement," mimeo, Yale University.

Becker, G. S., and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *The Journal of Political Economy*, 87(6), 1153–1189.

Blanchard, O. J., and L. F. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 1992(1), 1–75.

Calvo-Armengol, A., and M. O. Jackson (2004): "The Effects of Social Networks on Employment and Inequality," *The American Economic Review*, 94(3), 426–454.

Chamberlain, G. (1980): "Analysis of Covariance with Qualitative Data," *The Review of Economic Studies*, 47(1), 225–238.

Conley, T., and G. Topa (2002): "Socio-Economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, 17(4), 303–327.

Galton, F. (1869): *Hereditary Genius: An Inquiry Into Its Laws and Consequences*. Macmillan, London.

Granovetter, M. (1983): "The Strength of Weak Ties: A Network Theory Revisited," *Sociological Theory*, 1, 201–233.

Ioannides, Y. M., and L. D. Loury (2004): "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42(4), 1056–1093.

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

Kingdon, G. G., and J. Knight (2006): "The Measure of Unemployment When Unemployment is High," *Labour Economics*, 13(3), 291–315.

Kramarz, F., and O. N. Skans (2006): "Nepotism at Work? Family Networks and Youth Labor Market Entry," Working Paper.

Loury, L. D. (2006): "Some Contacts are More Equal Than Others: Informal Networks, Job Tenure, and Wages," *Journal of Labor Economics*, 24(2), 299–318.

Moll, P. (1996): "Compulsory Centralization of Collective Bargaining in South Africa," *The American Economic Review*, 86(2), 326–329.

Mookherjee, D., and D. Ray (2002): "Is Equality Stable?," *The American Economic Review*, 92(2), 253–259, Papers and Proceedings of the One Hundred Fourteenth Annual Meeting of the American Economic Association.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *The Quarterly Journal of Economics*, 118(2), 549–599.

Munshi, K., and M. Rosenzweig (2003): "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy," BREAD Working Paper 038.

Murphy, K. M., and R. H. Topel (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business and Economic Statistics*, 3(4), 370–379.

Schultz, T. P., and G. Mwabu (1998): "Labor Unions and the Distribution of Wages and Employment in South Africa," *Industrial and Labor Relations Review*, 51(4), 680–703.

Seekings, J., and N. Nattrass (2005): *Class, Race, and Inequality in South Africa*. Yale University Press, New Haven and London.

Solon, G. (1999): "Intergenerational Mobility in the Labor Market," in *The Handbook of Labor Economics, Vol. 3A*, ed. by O. Ashenfelter, and D. Card, pp. 631–660. Elselvier-North Holland, New York.

Standing, G., J. Sender, and J. Weeks (1996): *The South African Challenge: Restructuring the South African Labor Market.* International Labour Organisation, Geneva.

Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data.* MIT Press, Cambridge, Massachusetts, and London.

LF Status		Working	Searching	Unemployed Wants Work	Not in LF
Black Males, 20-34	1993	49.08	14.83	12.6	23.49
	2003	38.58	26.55	17.93	16.94
change, 1993-2003		-10.5	11.72	5.33	-6.55
Black Males, 35-60	1993	76.09	7.95	7.61	8.35
	2003	63.61	13.6	8.95	13.83
change, 1993-2003		-12.48	5.65	1.34	5.48
Black Females,20-34	1993	33.21	13.88	16.55	36.36
	2003	24.9	27.8	27.67	19.63
change, 1993-2003		-8.31	13.92	11.12	-16.73
Black Female,35-60	1993	47.71	8.24	10.44	33.61
	2003	47.08	11.94	14.45	26.53
change, 1993-2003		-0.63	3.7	4.01	-7.08
Coloured Males,20-34	1993	70.61	18.03	2.62	8.74
	2003	63.96	20.56	7.19	8.29
change, 1993-2003		-6.65	2.53	4.57	-0.45
Coloured Males,36-60	1993	76	8.48	2.11	13.41
	2003	68.98	8.53	4.4	18.09
change, 1993-2003		-7.02	0.05	2.29	4.68
Coloured Females,20-34	1993	54.79	16.66	5.2	23.35
	2003	50.2	20.03	13.58	16.2
change, 1993-2003		-4.59	3.37	8.38	-7.15
Coloured Females,35-60	1993	45.81	5.8	2.92	45.48
	2003	49.4	8	6.38	36.22
change, 1993-2003		3.59	2.2	3.46	-9.26
White Males, 20-34	1993	87.9	3.15	0.93	8.02
	2003	81.43	5.49	1.64	11.45
change, 1993-2003		-6.47	2.34	0.71	3.43
White Males,35-60	1993	91.15	2.33	0.45	6.08
	2003	87.99	2.71	0.75	8.55
change, 1993-2003		-3.16	0.38	0.3	2.47
White Females,20-34	1993	63.38	5.68	4.12	26.83
	2003	67.5	6.32	3.05	23.12
change, 1993-2003		4.12	0.64	-1.07	-3.71
White Females,35-60	1993	53.43	3.68	4.7	38.18
	2003	60.26	2.67	2.48	34.59
change, 1993-2003		6.83	-1.01	-2.22	-3.59

Table 1: South African Unemployment

Table reports the fraction of individuals by race-age-gender group who are in each labor force category.

LF Status		Unemployed	% Never Work	Time since last job <1 year	1-3 years	3years +
Black Males, 20-34	1993	27.43	59.09	29.21	41.16	29.62
	2003	44.48	79.75	41.6	34.35	24.05
change, 1993-2003		17.05	20.66	12.39	-6.81	-5.57
Black Males, 35-60	1993	15.56	25.68	20.83	33.14	46.04
	2003	22.55	31.35	15.93	17.31	66.76
change, 1993-2003		6.99	5.67	-4.9	-15.83	20.72
Black Females,20-34	1993	30.43	66.66	28.96	44.25	26.79
	2003	55.47	80	34.6	33.12	32.27
change, 1993-2003		25.04	13.34	5.64	-11.13	5.48
Black Female,35-60	1993	18.68	45.39	20.48	35.27	44.25
	2003	26.39	49.83	13.43	18.98	67.58
change, 1993-2003		7.71	4.44	-7.05	-16.29	23.33
Coloured Males,20-34	1993	20.65	36.57	39.71	47.94	12.35
	2003	27.75	52.08	50	31.6	18.4
change, 1993-2003		7.1	15.51	10.29	-16.34	6.05
Coloured Males,36-60	1993	10.59	11.55	41.89	41.44	16.67
	2003	12.93	20.45	22.65	18.53	58.82
change, 1993-2003		2.34	8.9	-19.24	-22.91	42.15
Coloured Females,20-34	1993	21.86	40.83	36.5	51.75	11.75
	2003	33.61	43.4	39.9	33.84	26.26
change, 1993-2003		11.75	2.57	3.4	-17.91	14.51
Coloured Females,35-60	1993	8.72	21.19	33.33	44.09	22.58
	2003	14.38	23.63	13.83	18.01	68.17
change, 1993-2003		5.66	2.44	-19.5	-26.08	45.59
White Males, 20-34	1993	4.08	23.76	53.25	45.24	4.76
	2003	7.13	68.99	52.17	39.13	8.7
change, 1993-2003		3.05	45.23	-1.08	-6.11	3.94
White Males,35-60	1993	2.78	2.86	53.92	32.35	13.73
	2003	3.46	12.79	20.57	26.24	53.19
change, 1993-2003		0.68	9.93	-33.35	-6.11	39.46
White Females,20-34	1993	9.8	18.56	26.51	34.42	39.07
	2003	9.37	55.59	20.57	26.24	53.19
change, 1993-2003		-0.43	37.03	-5.94	-8.18	14.12
White Females,35-60	1993	8.38	6.85	24.41	27.09	48.49
	2003	5.15	41.38	30.89	36.59	32.52
change, 1993-2003		-3.23	34.53	6.48	9.5	-15.97

Table 2: Duration of Unemployment

Fraction who have never worked and who belong to each duration category is the fraction of currently unemployed in that grouping.

probit, wage work	Primary Ed, 1993	Primary, 2003	Secondary Ed, 1993	Secondary, 2003	% change, Secondary
Black Males, 20-34	-0.0296 ***	-0.0014	-0.0021	-0.0080 **	2.8095
	(0.003)	(0.005)	(0.004)	(0.003)	
Black Males, 35-60	-0.0023	0.0038	0.0486 ***	0.0126 ***	-0.7407
	(0.003)	(0.003)	(0.005)	(0.004)	
Black Females, 20-34	-0.0030	0.0007	0.0319 ***	0.0000	-1.0000
	(0.002)	(0.004)	(0.003)	(0.003)	
Black Females, 35-60	0.0082 ***	0.0072 ***	0.0830 ***	0.0285 ***	-0.6566
	(0.002)	(0.003)	(0.004)	(0.003)	
Coloured Males, 20-34	-0.0051	0.0259	0.0459 ***	0.0141	-0.6928
	(0.007)	(0.017)	(0.007)	(0.011)	
Coloured Males, 35-60	-0.0001	-0.0065	0.0626 ***	0.0243 ***	-0.6118
	(0.005)	(0.009)	(0.008)	(0.009)	
Coloured Females, 20-34	0.0163 ***	0.0133	0.0603 ***	0.0353 ***	-0.4146
	(0.006)	(0.016)	(0.007)	(0.010)	
Coloured Females, 35-60	0.0218 ***	0.0163 *	0.0876 ***	0.0473 ***	-0.4600
	(0.005)	(0.009)	(0.008)	(0.008)	
White Males, 20-34	0.0146		-0.0190 **	0.0246	-2.2947
	(0.027)		(0.008)	(0.040)	
White Males, 35-60	0.0247		0.0034	-0.0022	-1.6471
	(0.021)		(0.006)	(0.032)	
White Females, 20-34	0.0269	0.0665	0.0607 ***	0.0387	-0.3624
	(0.025)	(0.107)	(0.008)	(0.045)	
White Females, 35-60	0.0121		0.0467 ***	0.0100	-0.7859
	(0.024)		(0.006)	(0.028)	

Table 3: Employment returns to Educaton

Estimated marginal effects of eduction from a probit on waged work. Domestic work and self-employment are excluded. All estimates are conditional on urbanization and a quadratic in age

	Males				Females			
	mean	sd	min	max	mean	sd	min	max
Black	0.490	0.500	0	1	0.514	0.500	0	1
Age	18.693	2.608	14	24	18.749	2.549	14	24
Father in Prov	0.651	0.477	0	1	0.630	0.483	0	1
Father in HH	0.443	0.497	0	1	0.398	0.490	0	1
Father Works	0.449	0.497	0	1	0.435	0.496	0	1
works	0.248	0.432	0	1	0.181	0.385	0	1
in school	0.451	0.498	0	1	0.458	0.498	0	1
grade	9.568	3.243	0	25	10.154	3.297	0	25
How job was found								
network job	0.137	0.344	0	1	0.088	0.283	0	1
found job by self	0.069	0.254	0	1	0.061	0.239	0	1

Table 4: CAPS Summary Statistics

Summary statistics in my estimation sample of black and coloured young adults. Found job through network is equal to 1 if and only if the respondent is both working and found a job through the means described, and a zero otherwise, and found job by self is defined analogously

	Father's Industry					
Child's Industry	Sons	Sons	Sons	Daughters	Daughters	Daughters
Informal Services	0.0001	0.0853	0.0948	0.0010	0.0240	0.0055
	(0.020)	(0.080)	(0.074)	(0.019)	(0.037)	(0.029)
Agriculture	0.0487**	0.1526***	0.1509***	0.0688***	0.0738***	0.0231
	(0.023)	(0.022)	(0.022)	(0.022)	(0.020)	(0.017)
Manufacturing	0.0100	0.1129***	0.1167***	0.0135	0.0353***	0.0179
	(0.011)	(0.020)	(0.019)	(0.011)	(0.013)	(0.011)
Utility Provision	0.2110***	0.2156***	0.2152***			
	(0.016)	(0.022)	(0.022)			
Construction	0.0152	0.1216***	0.1231***	0.0117	0.0290**	0.0155
	(0.011)	(0.020)	(0.019)	(0.010)	(0.012)	(0.011)
Retail	0.0124	0.1185***	0.1177***	0.0155	0.0364***	0.0192
	(0.011)	(0.020)	(0.019)	(0.012)	(0.014)	(0.012)
Transportation	0.0136	0.1097***	0.1121***	0.0173	0.0325**	0.0192
	(0.014)	(0.021)	(0.020)	(0.013)	(0.015)	(0.012)
Business	0.0011	0.0934***	0.0982***	0.0176	0.0399**	0.0230
	(0.020)	(0.026)	(0.025)	(0.018)	(0.066)	(0.014)
Services	0.0212*	0.1169***	0.1120***	0.0107	0.0316**	0.0168
	(0.012)	(0.019)	(0.019)	(0.012)	(0.014)	(0.012)
N	605	292	292	570	291	291
Excluding unemployed Parents	No	Yes	Yes	No	Yes	Yes
Race, Age, Ed Fixed Effects	No	No	Yes	No	No	Yes
F-test: all Industries	178.16	139.8	140.34	13.77	16.95	4.21
p-value	0	0	0	0.088	0.0306	0.8373
F-test: All but agriculture	174.67	120.64	123.47	5.37	9.74	3.9
p-value	0	0	0	0.6155	0.2036	0.7909

Table 5: Correlations between Fathers' Industries and Childrens' Industries

Reports results from seemingly unrelated regressions where the dependent variable is a dummy for the child working in a one-digit industry and the independent is a dummy for the father working in that industry (and, in some specifications, fixed effects for each age and education level and racial group).

	Mother's Industry					
Child's Industry	Sons	Sons	Sons	Daughters	Daughters	Daughters
Informal Services	0.0052	0.0286	0.0092	0.0312*	0.1495***	0.1267***
	(0.010)	(0.019)	(0.021)	(0.017)	(0.025)	(0.026)
Agriculture	-0.0006	0.0196	-0.0006	0.5175***	0.6495***	0.5891***
	(0.025)	(0.031)	(0.035)	(0.046)	(0.049)	(0.075)
Manufacturing	0.0031	0.0314	0.0085	0.0335**	0.1858***	0.1707
	(0.007)	(0.019)	(0.021)	(0.015)	(0.028)	(0.028)
Resources	-0.0010					
	(0.024)					
Construction	0.0111	0.0473	0.0244	0.0005	0.0746	0.0426
	(0.035)	(0.045)	(0.056)	(0.059)	(0.074)	(0.071)
Retail	0.0034	0.0303	0.0138	0.0250	0.1884***	0.1747***
	(0.008)	(0.020)	(0.021)	(0.020)	(0.031)	(0.030)
Transportation	0.0003	0.0236	0.0084	0.0003	0.0553	0.0349
	(0.058)	(0.072)	(0.072)	(0.049)	(0.052)	(0.048)
Business	-0.0014	0.0216	0.0082	0.0641	0.2037***	0.1858***
	(0.019)	(0.029)	(0.031)	(0.042)	(0.051)	(0.047)
Services	0.0005	0.3974	0.0045	0.0186	0.1617***	0.1385***
	(0.007)	(0.019)	(0.020)	(0.016)	(0.027)	(0.026)
N	605	282	282	568	260	260
Excluding unemployed Parents	No	Yes	Yes	No	Yes	Yes
Race, Age, Ed controls	No	No	Yes	No	No	Yes
F-test: all Industries	0.61	3.29	0.85	136.14	201.88	93.64
p-value	0.9999	0.9148	0.999	0	0	0
F-test: All but agriculture	0.6	3.29	0.8	9.94	57.67	45.36
p-value	0.9997	0.857	0.9975	0.1922	0	0

Table 6: Correlation between Mothers' and Childrens' Industries

Reports results from seemingly unrelated regressions where the dependent variable is a dummy for the child working in a one-digit industry and the independent is a dummy for the mother working in that industry (and, in some specifications, fixed effects for each age and education level and racial group)

	Fathers' Occupation					
	Sons	Sons	Sons	Daughters	Daughters	Daughters
Legislators, Senior Officials, and Managers	-0.0094	-0.0065	-0.0080			
	(0.014)	(0.015)	(0.015)			
Professionals	-0.0017	-0.0015	-0.0052	0.0006	-0.0020	0.0071
	(0.016)	(0.010)	(0.021)	(0.021)	(0.017)	(0.016)
Technicians and Associate Professionals	0.0272	0.0378	0.0147	0.0373	0.0240	0.0217
	(0.029)	(0.034)	(0.036)	(0.027)	(0.026)	(0.025)
Clerks	0.0367	0.0494	0.0597	0.0195	0.0067	-0.0013
	(0.047)	(0.052)	(0.050)	(0.039)	(0.037)	(0.035)
Service and Market Sales workers	0.0138	0.0275	0.0215	0.0322	0.0174	0.0184
	(0.029)	(0.034)	(0.034)	(0.024)	(0.023)	(0.022)
Agriculture and Fisheries Workers	0.0505	0.0581	0.0568	-0.0008	-0.0047	-0.0033
	(0.035)	(0.039)	(0.042)	(0.041)	(0.038)	(0.030)
Craft and Related Trades Workers	-0.0053	0.0128	0.0056	0.0055	0.0021	0.0034
	(0.015)	(0.022)	(0.022)	(0.013)	(0.013)	(0.013)
Plant and Machine Operators	-0.0118	-0.0042	-0.0112	0.0004	-0.0095	-0.0100
	(0.019)	(0.024)	(0.024)	(0.016)	(0.017)	(0.017)
Elementary Occupations	0.0288	0.0482	0.0514	0.0242	0.0238	0.0189
	(0.021)	(0.028)	(0.028)	(0.019)	(0.020)	(0.020)
N	605	292	292	570	291	291
Excluding unemployed Parents	No	Yes	Yes	No	Yes	Yes
Race, Age, Ed Fixed Effects	No	No	Yes	No	No	Yes
F-test: all occupations	6.8	7.51	8.57	5	3.64	3.11
p-value	0.6575	0.5841	0.478	0.07573	0.8883	0.9274

Table 7: Occupational Correlations: Fathers

Reports results from seemingly unrelated regressions where the dependent variable is a dummy for the child working in a one-digit occupation and the independent is a dummy for the father working in that occupation (and, in some specifications, fixed effects for each age and education level and racial group).

	Mothers Occupations					
	Sons	Sons	Sons	Daughters	Daughters	Daughters
Legislators, Senior Officials, and Managers	0.0007	0.0060	-0.0033	0.0016	0.0179	0.0138
	(0.021)	(0.017)	(0.018)	(0.016)	(0.023)	(0.018)
Professionals	0.0316**	0.0569***	0.0582	0.0512***	0.0658***	0.0407**
	(0.013)	(0.015)	(0.015)	(0.015)	(0.019)	(0.016)
Technicians and Associate Professionals	0.0110	0.0515	0.0529	0.0203	0.0472*	0.0369
	(0.026)	(0.037)	(0.036)	(0.025)	(0.029)	(0.023)
Clerks	0.0165	0.0545*	0.0475*	0.0236	0.0584**	0.0392*
	(0.021)	(0.029)	(0.028)	(0.021)	(0.026)	(0.020)
Service and Market Sales workers	0.0173	0.0711**	0.0777***	0.0228	0.0564**	0.0360*
	(0.021)	(0.030)	(0.029)	(0.019)	(0.023)	(0.190)
Agriculture and Fisheries Workers	0.0023	0.0227	0.0486	0.0024	0.0177	0.0118
	(0.070)	(0.079)	(0.080)	(0.065)	(0.073)	(0.062)
Craft and Related Trades Workers	-0.0630**	-0.0193	-0.0362	0.0225	0.0501***	0.0362**
	(0.022)	(0.033)	(0.033)	(0.016)	(0.019)	(0.016)
Plant and Machine Operators	0.0151	0.0619**	0.0633**	0.0235	0.0540**	0.0359**
	(0.018)	(0.026)	(0.025)	(0.019)	(0.021)	(0.018)
Elementary Occupations	0.0203	0.0809***	0.0788***	0.0275**	0.0721***	0.0437***
	(0.015)	(0.025)	(0.026)	(0.013)	(0.019)	(0.017)
N	603	281	281	570	263	263
Excluding unemployed Parents	No	Yes	Yes	No	Yes	Yes
Race, Age, Ed Fixed Effects	No	No	Yes	No	No	Yes
F-test: all occupations	18.59	29.17	31.67	17.25	20.16	10.34
p-value	0.0172	0.0003	0.0002	0.045	0.0169	0.3238

Table 8: Intergenerational Correlations: Mothers Occupation

Reports results from seemingly unrelated regressions where the dependent variable is a dummy for the child working in a one-digit occupation and the independent is a dummy for the mother working in that occupation (and, in some specifications, fixed effects for each age and education level and racial group).

Work	1	2	3	4	5	6
Log Employment, Father's Industry (FEmp)	- 0.0105	- 0.092**	- 0.1103	0.0013	0.0073	0.0174
	(0.033)	(0.042)	(0.103)	(0.086)	(0.087)	(0.091)
Male*FEmp		0.166***		- 0.2878	- 0.2963	- 0.2827
		(0.064)		(0.202)	(0.204)	(0.206)
Father in Province*FEmp			0.1061	- 0.1010	- 0.1033	- 0.1171
			(0.109)	(0.096)	(0.097)	(0.101)
Male*Father in Province*FEmp				0.4789**	0.4854**	0.4692**
				(0.215)	(0.217)	(0.218)
observations	11014	11014	11014	11014	11014	11014
F-test: Male effect				8.35	4.89	4.76
p-value				0.0002	0.0076	0.0086
F-test: Female effect				4.17	2.29	2.49
p-value				0.0155	0.1016	0.0831
Education Fixed Effects	No	No	No	No	Yes	Yes
Age and Race-Year fixed Effects	No	No	No	No	No	Yes

Table 9: Baseline results – Fathers

In all regressions, an indicator for working is the dependent variable. Standard errors are clustered at the household level.

Dependent variable	Net Job	Self Job	Net Job	Self Job
Log Emp., Father's Indus (Femp)	- 0.0946	0.0703	- 0.0774	0.0379
	(0.106)	(0.084)	(0.093)	(0.065)
Male*FEmp	- 0.1853	- 0.0167	- 0.1770	0.0028
	(0.196)	(0.113)	(0.210)	(0.109)
Father in Prov*FEmp	0.0738	- 0.0679	0.0320	- 0.0415
	(0.111)	(0.088)	(0.100)	(0.073)
Male*Father in Prov*FEmp	0.2732	0.0255	0.2926	0.0299
	(0.204)	(0.119)	(0.221)	(0.117)
Excluding Other jobs?	No	No	Yes	Yes
F: Male	1.6	0.04	1.87	0.31
	0.2016	0.9583	0.1536	0.7313
F: Female	0.55	0.35	0.94	0.17
	0.5742	0.7016	0.3895	0.8401
no obs	11014	11014	9877	9379

Table 10: Network Jobs and Self-Sought Jobs

Net Job is equal to one if and only if the young adult is both working and reports the assistance of friends or relatives in finding a job, and a zero otherwise. Self job is similarly defined for job search methods which did not rely on networks. FEmp is log employment in the father's baseline industry. Standard errors are clustered at the household level.

Dependent Variable	Work	Work	Work	Net Job	Self Job
Log Employment, Mother's industry (MEmp)	0.0047	0.0101	0.0161	0.0163	0.0176
	(0.050)	(0.050)	(0.050)	(0.041)	(0.032)
Male*MEmp	0.0526	0.0380	0.0380	0.0166	0.0009
	(0.074)	(0.075)	(0.075)	(0.071)	(0.053)
Log Employment, Father's Industry (FEmp)		- 0.0924**	- 0.0915**	- 0.0266	0.0065
		(0.042)	(0.042)	(0.034)	(0.026)
Male*FEmp		0.1615***	0.1565***	0.0743	0.0051
		(0.065)	(0.064)	(0.057)	(0.033)
N	11014	11014	11014	11014	11014

Table 11: Impacts of Mothers as Connections

In columns 1-3, an indicator for working is the dependent variable, where in column 4 it is an indicator for both working and using network help in finding a job and column 5 is an indicator for both working and using your own means to find this job. Standard errors are clustered at the household level.

	% Jobs through Network		
	Mother's Industry	Father's Industry	Difference: Fathers-Mothers
Daughters	0.4679	0.5175	0.0496
	(0.125)	(0.233)	(0.009)
N	987	938	
Sons	0.5004	0.5649	0.0645
	(0.118)	(0.153)	(0.007)
N	833	802	
Difference: Sons-Daughters	0.0325	0.0474	
	(0.006)	(0.009)	
Difference: Father-Son Effect minus Mother-Daughter effect:			0.0970
			(0.004)

Table 12: Comparisons: fraction of networked jobs

Fraction of employees who received network help is calculated for each gender-industry cell, and is applied to the industries in which mothers and fathers work. The total difference between sons and daughters is the difference between sons in fathers industies and daughters in mothers industries, which is decomposed into the differences in the industries in which men and women work and the differences between sons and daughters within those industries.

	1	2	3	4
Log Emp, Fathers' Industry (FEmp)	-0.288	-0.304	-0.324	-0.319
	(0.193)	(0.197)	(0.205)	(0.217)
Father in Prov*FEmp	0.377*	0.423**	0.450**	0.457**
	(0.201)	(0.205)	(0.214)	(0.225)
Log Emp., historical indus		0.028	0.028	0.028
		(0.083)	(0.083)	(0.083)
Father in Prov*Hist Emp		-0.093	-0.093	-0.092
		(0.100)	(0.100)	(0.100)
Coarse Occupation Emp			0.026	0.006
			(0.074)	(0.109)
Father in Prov*Coarse Emp			-0.037	-0.047
			(0.077)	(0.114)
Fine Occupation Emp				0.039
				(0.113)
Father in Prov*Fine Emp				0.02
				(0.119)
F-test: Father Network	2.63	3.41	3.48	3.76
p-value	(0.072)	(0.033)	(0.031)	(0.023)
F-test: Specific Capital		0.76	0.48	0.81
p-value		(0.470)	(0.749)	(0.563)
Observations	4985	4985	4985	4985

Table 13: Specific Capital examination

In all regressions, a dummy for working is the dependent variable. FEmp is log employment in the fathers industry at baseline. Log employment in the historical industry is the current employment in the industry which your father worked in "while you were growing up". Coarse occupations are divided into high-skilled white collar, low-skilled white collar, and blue collar, while fine occupations are occupations at the one digit level. Occupational employment is log employment in the occupation-industry cell in which the father works. Standard errors are clustered at the household level.

	1	2	3	4	5	6
	work	work	work	work	work	work
Log Emp., Father's Indus (FEmp)	0.071	-0.288	-0.287	-0.286	-0.289	-0.276
	(0.049)	(0.193)	(0.194)	(0.192)	(0.193)	(0.191)
Father in Province*FEmp		0.377*	0.401*	0.401*	0.432**	0.416**
		(0.201)	(0.208)	(0.207)	(0.213)	(0.211)
log Employment, Male Head Indus			-0.036	-0.032	0.024	0.042
			(0.069)	(0.068)	(0.113)	(0.109)
Father in Prov*Male Head Emp					-0.101	-0.121
					(0.140)	(0.137)
neighborhood youth employment rate				0.172**		0.200**
				(0.063)		(0.100)
Father in Prov*Neighbor Emp Rate						-0.038
						(0.123)
Log Employment, Modal Industry				-0.049		-0.122*
				(0.042)		(0.069)
Father in Prov* Modal Indus Emp						0.113
						(0.083)
F-test: Father-network Effect		2.63	2.29	2.31	2.59	2.4
p-value		(0.072)	(0.101)	(0.100)	(0.076)	(0.091)
F-test: all network variables			1.8	2.93	1.46	2.27
p-value			(0.145)	(0.012)	(0.213)	(0.020)
Observations	4985	4985	4985	4985	4985	4985

Table 14: Correlated Networks

In all regressions, an indicator for working is the dependent variable. FEmp is log employment in the father's industry at baseline. The Male Head is the male head of household or the first working male on the household roster. The neighborhood youth employment rate is the fraction of males in other households in the same sampling cluster as the respondent who are working, while the modal industry is the modal industry among adult males in that sampling cluster. Standard Errors are clustered at the household level.

	1	2	3	4	5	6
Dependent Variable	Net Job	Net Job	Net Job	Self Job	Self Job	Self Job
Log Emp, Father's Industry (FEmp)	- 0.2852	- 0.2871	- 0.2866	0.0531	0.0523	0.0548
	(0.177)	(0.177)	(0.177)	(0.076)	(0.076)	(0.077)
Father in Prov*FEmp	0.3484*	0.3356*	0.3713*	- 0.0420	- 0.0257	- 0.0404
	(0.182)	(0.188)	(0.191)	(0.079)	(0.082)	(0.088)
Log Employment, Male Head Indus		0.0200	0.1025		- 0.0249	- 0.0470
		(0.066)	(0.101)		(0.038)	(0.056)
Father in Prov*Male Head Emp			- 0.1364			0.0374
			(0.129)			(0.075)
Neighborhood Youth Employment Rate		0.1365**	0.1025		0.0249	0.0911
		(0.054)	(0.091)		(0.042)	(0.071)
Father in Prov*Neighborhood Emp Rate			0.0461			- 0.0898
			(0.110)			(0.086)
Log Employment, Modal Industry		- 0.0156	- 0.0311		0.0100	- 0.0070
		(0.037)	(0.054)		(0.028)	(0.039)
Father in Prov*Modal Indus Emp			0.0233			0.0264
			(0.067)			(0.050)
Obs	4985	4985	4985	4985	4985	4985
F-test: Father effect	2.27	1.61	1.98	0.37	0.54	0.32
p-value	0.1032	0.2008	0.1386	0.6914	0.584	0.7243
F-test: All networks		2.41	1.65		0.3	0.35
p-value		0.0341	0.1046		0.9114	0.9478

Table 15: Robustness Checks with Job Search Channels

In Columns 1-3, the dependent variable is an indicator for both working and using network help to find the job. In columns 3-6, the dependent variable is both working and having found the job without using networks. FEmp is log employment in the father's industry at baseline. The Male Head is the male head of household or the first working male on the household roster. The neighborhood youth employment rate is the fraction of males in other households in the same sampling cluster as the respondent who are working, while the modal industry is the modal industry among adult males in that sampling cluster. Standard errors are clustered at the household level.

	Non-Attritors in 2002		Attritors in 2002	
	Mean	SD	Mean	SD
Black	0.477	0.500	0.669	0.471
Female	0.542	0.498	0.590	0.492
Age	17.720	2.466	18.485	2.403
Works	0.196	0.397	0.166	0.373
Father in Province	0.660	0.474	0.498	0.500
Father Working	0.444	0.497	0.380	0.486
Mother in Province	0.834	0.372	0.625	0.484
Mother Working	0.464	0.499	0.382	0.486
Job was found:				
through network	0.102	0.303	0.067	0.251
by self	0.043	0.202	0.030	0.169
N	3178		949	

Table 16: Attrition Summary Statistics

	1	2	3	4
Log Emp., Father's Indus. (FEmp)	0.0559	- 0.0131	0.0867	- 0.0177
	(0.131)	(0.061)	(0.140)	(0.063)
Male*FEmp	- 0.2232	- 0.2304	- 0.1990	- 0.2545
	(0.245)	(0.161)	(0.253)	(0.169)
Father in Province*FEmp	- 0.2187**	- 0.0872	- 0.2072	- 0.0986
	(0.139)	(0.073)	(0.146)	(0.077)
Male*Father in Prov*FEmp	0.4478*	0.4000**	0.3714	0.4763**
	(0.256)	(0.173)	(0.261)	(0.183)
Num Obs	12372	12372	12372	12372
F-test: Male Effect	5.85	4.81	4.09	6.54
p-value	0.0029	0.0082	0.0168	0.0015
F-test: Female Effect	6.23	1.63	4.01	3.42
p-value	0.002	0.1956	0.0181	0.0329

Table 17: Attrition Robustness

The dependent variable in all regressions is an indicator for working. In Column 1, all attritors are assumed to be working, while in column 2, none are. Column 3 presumes that attritors whose fathers are absent are working while those with present fathers are not, while column 4 assumes the opposite. Standard errors are clustered at the household level.

	work	Instrument F-stat	work	Instrument F-stat
Log Employment, Fathers Industry (FEmp)	- 0.1370	292.41	- 0.1736	135.32
	(0.122)		(0.437)	
Male*FEmp	0.3161*	304.88	- 0.0199	155.44
	(0.170)		(0.750)	
Father in Prov*FEmp			0.0399	148.65
			0.4555	
Male*Father in Prov*FEmp			(0.336)	158.44
			0.7704	
Regression Based Hausman test p-value	0.6082		0.7911	
num obs	11028		11028	
num individuals	4127		4127	

Table 18: Instrumental Variables Results

In both regressions, the dependent variable is an indicator for working. The instruments used are the average ratio of local employment to employment in other provinces in the father's two digit industry multiplied by the employment numbers in other provinces in that two digit industry.

Working Coefficients	1	2	3	4	5	6
Log Employment, Father's Indus (FEmp)	- 0.3292	- 0.9927	- 1.1561***	0.8980	- 1.1614**	0.2488
	(0.274)	(2.592)	(0.404)	(1.917)	(0.415)	(1.804)
Male*FEmp			1.6235***	- 3.4982	1.6032**	- 1.1697
			(0.563)	(3.618)	(0.574)	(3.032)
Father in Province*FEmp		0.6520		- 2.1760		- 1.5171
		(2.380)		(1.962)		(1.851)
Male*Father in Prov*FEmp				5.3368		2.9775
				(3.665)		(3.088)
Education <8 years					- 0.3837	- 0.2888
					(0.520)	(0.517)
Years of Education-8, Ed>7 and Ed<12					- 0.0684	0.0187
					(0.172)	(0.169)
At least 12 years of Education					- 1.7166**	- 0.9222
					(0.808)	(0.777)
Wald test: Male Effect				10.926		9.716
p-value				0.0042		0.0078
Wald test: Female Effect				9.5077		9.1188
p-value				0.0086		0.0105
In School Coefficients						
FEmp	- 0.0867	1.398	0.0046	0.6673	0.0341	0.5480
	0.2585	3.0087	0.3545	0.9829	0.371	0.991
Male*FEmp			- 0.1201	2.6747	- 0.0657	2.1580
			0.5145	2.4514	0.540	2.349
Father in Province*FEmp		- 1.6251		- 0.7765		- 0.5227
		2.2245		1.0516		1.062
Male*Father in Prov*FEmp				- 2.8489		- 2.3740
				2.5112		2.415
Education <8 years					- 0.3910	- 0.3681
					0.234	0.233
Years of Education-8, Ed>7 and Ed<12					0.7239***	0.7216
					0.109	0.106
At least 12 years of Education					2.5692***	2.5674****
					0.569	0.557
Wald test: Male Effect				1.2935		0.9939
p-value				0.5237		0.6084
Wald test: Female Effect				0.5459		0.3096
p-value				0.7611		0.8566

Table 19: Multinomial Conditional Logit results

FEmp is log employment in the fathers industry. The multinomial choice set is working, schooling, or neither, and the coefficient estimates presented are the differential value for working or schooling over choosing neither.

	Fathers Here and Working	Other Households
Household Per Capita Income	899.99	526.29
	(965.164)	(676.986)
Years of Education	9.01	8.90
	(1.956)	(2.093)
Test Score	26.68	24.21
	(7.962)	(8.251)
Household Members	5.64	5.71
	(2.224)	(2.752)
Enrolled in School	0.67	0.58
	(0.471)	(0.494)

Table 20: Comparison of Households with Network advantage to those without

Column 1 presents sample means for young adults with fathers who are here and working while column 2 presents the same statistics for young adults with fathers who are either absent, deceased, or unemployed.

Baseline Results	
Log Emp, Father's Indus (FEmp)	-0.0139
Unadjusted Std. Dev.	(0.120)
Adjusted Std. Dev	(0.125)
Male*FEmp	-0.1984
Unadjusted Std. Dev.	(0.189)
Adjusted Std. Dev.	(0.192)
Father in Prov*FEmp	-0.0846
Unadjusted Std. Dev	(0.124)
Adjusted Std. Dev	(0.125)
Male*Father in Prov*FEmp	0.3897
Unadjusted Std. Dev	(0.195)
Adjusted Std. Dev	(0.203)
num obs	11028
num individuals	4127
Wald: Males, Unadjusted Variance	16.7240
p-value	0.0002
Wald: Males, Adjusted Variance	9.9406
p-value	0.0069
Wald: Females, Unadj. Variance	8.3470
p-value	0.0154
Wald: Females, Adjusted Variance	6.3832
p-value	0.0362

Table 21: Baseline Estimates with Unadjusted and Murphy-Topel (1985) Standard Errors

Coefficients are not yet adjusted for attenuation bias induced by the sampling error. The Dependent variable is an indicator for working. FEmp is log employment in the father's industry at baseline.