Essays on Labour Market in Developing Countries

Peng Zhang

Faculty of Economics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Newnham College

June 2018

Abstract

This PhD thesis focuses on determinants of labour market outcomes in development economics with a special interest in South Africa and China.

After an introduction in chapter 1, the key chapter 2, *Ethnic Diversity and Labour Market Outcomes: Evidence from Post-Apartheid South Africa* joint with Sara Tonini, investigates how ethnic diversity amongst black South Africans affects their employment opportunities in the post-Apartheid era. We find that ethnic diversity has a positive impact on the employment rate of the black South Africans, and it only affects ethnic groups with relatively large population size. To address the endogeneity of ethnic composition, we explore the location of historical "black homelands" and argue that districts more equally distant to multiple homelands are more ethnically diverse. In our instrumental variable regressions, a one standard deviation increase in ethnic diversity index increases employment rate by 3 (5) percentage point in 1996 (2001), which is around 8% (13%) of the average employment rate. We then propose a model of a coordination game to explain these findings. A more ethnically diverse place requires a higher rate of inter-ethnic communication to maintain social connection. As inter-ethnic communication requires more skills than intra-ethnic connection, people in ethnically diverse districts are motivated to invest more in social skills to be able to communicate with those outside their own group. The acquisition of these social skills makes them better equipped for the labour market.

The remaining two chapters look into the intergenerational transmission of socio-economic status in South Africa and China. Chapter 3, *Returns to Education, Marital Sorting and Family Background in South Africa* joint with Patrizio Piraino, applies the model of Lam (1993, JPE) which combines intergenerational transmission of ability and assortative mating to investigate the relative explanatory power of father-in-law's and father's background for male wages. In the empirical analysis, after correcting for potential measurement errors in earnings and education, we find that father-in-law's schooling is more correlated with male workers' labour market earnings, employment rate and labour force participation than own father's schooling in contemporary South Africa. This difference is more obvious when parental educational levels are higher.

Chapter 4, *Higher Education Expansion and Intergenerational Mobility in Contemporary China*, studies how higher education affects the upward mobility of people from relatively disadvantaged families. Intergenerational occupational mobility is stimulated when children from different social classes end up in similar occupations. Whether or not they have similar occupational status depends not only on their level of education but also the occupational returns to education. Given there is already a convergence in educational achievements between children from different social classes in contemporary China, in this paper, I focus on their occupational returns to education. Occupational status is measured by the widely-accepted ISEI scaling system ranging from 16 to 90 points with large number indicating higher occupational status. I

take advantage of an exogenous college expansion policy in 1999 as a natural experiment and find that one additional year of education increases the occupational status of their first job by 2.243 (2.774) points on average along the ISEI scale in OLS (IV) regressions. And children from upper-class families do not necessarily have higher returns to education than children from other social classes. The average occupational returns to education are higher for the most recent job than the first job, but the difference among social classes is still not significant.

To my parents, my grandfather and other family members for all their support.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Introduction and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution. It does not exceed the prescribed word limit of 60,000 words.

Peng Zhang
June 2018

Acknowledgements

I am deeply indebted to my supervisors Pramila Krishnan and Toke Aidt for their continuous support and valuable suggestions throughout my PhD study. I am also very grateful to Kaivan Munshi, Hamish Low, Gabriella Santangelo, Meredith Crowley, Kai Liu, Jane Cooley Fruehwirth, Sriya Iyer and Julia Shvets whose comments significantly improved the thesis. I would also thank Murray Leibbrandt and Ingrid Woolard who offered me the opportunity to visit the University of Cape Town and work with great people there. I learnt especially a lot from my talented coauthor Patrizio Piraino. Tim Brophy and Lynn Woolfrey provided super useful guidance in helping me get familiar with data in South Africa.

I also thank my outstanding friends and colleagues in Cambridge Muhammad Farid Ahmed, Jake Bradley, Yujiang Chen, Jeroen Dalderop, Sihua Ding, Axel Gottfries, Felix Grey, Bowei Guo, Lu Han, Jiaqi Li, Simon Lloyd, Rafe Martyn, David Minarsch, Ekaterina Smetanina, John Spray, Chen Wang, Daniel Wales, Ruochen Wu and Jasmine Xiao for insightful discussions. Moreover, I would like to thank my friends, colleagues as well as roommates Sara Tonini and Alessandro Tondini when we were PODER fellows at the University of Cape Town. It was such a great time both personally and academically to be with amazing people like you.

In particular I would like to thank my parents and my grandfather whose continuous support is of fundamental importance throughout my PhD. You always mean a lot to me.

My PhD study also got the financial support from Cambridge Overseas Trust, PODER program (Policy Design and Evaluation Research in Developing Countries) as part of the Marie-Curie Fellowship and Luca D' Agliano Scholarship.

Peng Zhang
June 2018

Contents

Chapter 1

Introduction

This thesis focuses on development economics with a special interest in identifying constraints which prevent workers from being well equipped for labour market opportunities in South Africa and China. These constrains, such as complexity in social and ethnic interactions, not only restrict contemporaneous labour market opportunities but also can have long-lasting effects over generations due to the intergenerational persistence in socio-economic outcomes. Furthermore, they can affect both the quantity (e.g. employment rate) and quality (e.g. occupational choice and wage) of jobs accessible to workers.

I investigate two broad questions. Firstly, which features in developing countries affect contemporaneous labour market outcomes? Secondly, what is the role of intergenerational transmission in determining children's labour market opportunities? Chapter 2 answers the first question by investigating a factor especially highlighted in African countries - ethnic relationships, and identifying its impact on the post-Apartheid labour market in South Africa. Although it has been shown that ethnic diversity, or ethnic fragmentation, is an important reason why African countries fall behind in economic growth, this chapter suggests that ethnic diversity has a surprisingly positive effect on the employment rate among the black South Africans, which can be explained by a model of a coordination game.

The other two chapters answer the second question by providing theory and evidence on intergenerational transmission of socio-economic status in China and South Africa. Chapter 3 compares parents and parents-in-law to see which of the two can be more correlated with children's labour market outcomes in post-Apartheid South Africa. Consistent with previous evidence from Brazil, we also find the seemingly counterintuitive result that father-in-law's educational background has a stronger explanatory power than father's schooling for male wages, employment rate and labour force participation rate in South Africa, after correcting for potential measurement errors in both earnings and education. It is because it is possible that more information on children's unobserved characteristics is captured by father-in-law's education than father's education.

In Chapter 4, I study intergenerational mobility in China's labour market in the recent decades after the economic reform in 1978. Instead of taking the common approach by focusing on the difference in educational levels among social classes, I investigate if occupational returns to education are different among children from different social classes. I find that upper-class children do not necessarily have higher returns to education than children from other groups. Therefore, equalisation in educational opportunities is an effective way to stimulate the inter-

generational mobility of children from more disadvantaged classes.

The methodology of the thesis lies on exogenous historical events, geographical features and policy change to identify causal relationships, and/or use theoretical models to explain the mechanism. Data comes from administrative censuses, large-scale household surveys and geographical and archival sources.

Due to the word limit, some robustness checks in these chapters are not included in the thesis but can be found in the online version of the papers, which will be specified in the corresponding paragraphs.

Chapter 2

Ethnic Diversity and Labour Market Outcomes: Evidence from Post-Apartheid South Africa

Sara Tonini and Peng Zhang

Abstract. This paper investigates how ethnic diversity amongst black South Africans affects their labour market outcomes in the post-Apartheid era. We find that ethnic diversity has a positive impact on the employment rate of the black South Africans, and it only affects ethnic groups with relatively large population size. To address the endogeneity of ethnic composition, we explore the location of historical "black homelands" and argue that districts equally distant to multiple homelands are ethnically diverse. In our instrumental variable regressions, a one standard deviation increase in ethnic diversity index increases employment rate by 3 (5) percentage point in 1996 (2001), which is around 8% (13%) of the average employment rate. We also disentangle the two components in the ethnic diversity index and show that the variation in our diversity index comes from the dispersion of group size. We then propose a model of a coordination game to explain these findings. A more ethnically diverse place has less dispersion of group size, which implies a higher rate of inter-ethnic communication needed to maintain the overall level of social connection. As inter-ethnic communication requires more skills than intra-ethnic connection, people in ethnically diverse districts are motivated to invest more in social skills to be able to communicate with those outside their own group. The acquisition of these social skills makes them better equipped for the labour market. The key mechanism of the model is verified by both numerical simulation and empirical evidence.

Keywords: Ethnic diversity; Employment; Social skill; South Africa; Homeland.

JEL classification: O12, J40, Z13, N37.

2.1 Introduction

Many developing countries are characterised by a diverse composition of ethnic groups. A growing body of literature studies the link between ethnic diversity and economic performance, which is summarised in Alesina and La Ferrara (2005). The majority of the literature finds a negative association between ethnic diversity and many socio-economic indicators. More ethnically fractionalised communities can experience slower economic development as measured by GDP per

capita (Easterly and Levine, 1997). They may also have higher social costs which are reflected in lower levels of trust and participation in social activities (Alesina and La Ferrara, 2000, 2002), inefficient public goods provision (Alesina et al., 1997; Gomes et al., 2016) and higher inequality (Alesina et al., 2016b). The ethnic cleavage may also be detrimental to the establishment of a culture of inclusiveness and tolerance which is favorable to economic growth.

Much less is known on how ethnic diversity affects individual outcomes, especially labour market performance which is of great importance in driving economic development (Anand et al., 2016). This paper adds to this micro-level discussion by investigating how ethnic diversity amongst black South Africans (i.e. within-black ethnic diversity) affects their labour market outcomes in post-Apartheid era. We focus on the black people as the black population represents an overwhelming part of the whole South African labour force.[1] Moreover, the inter-ethnic relationship amongst the black can be more important than the black-white (or black-coloured) division in social interaction as the long-term Apartheid regime separated different racial groups and confined their choice of residence, which persists in the post-Apartheid years. It is therefore not common that the black and the white (or the black and the coloured) reside in the same community and have close interactions. Therefore, the coexistence of the black and white (or black and coloured) in the same district may not necessarily imply social interactions between them in reality.

We focus on how the employment rate of the black South Africans responds to the composition of black ethnic groups in the district of their residence.[2] Post-Apartheid South Africa provides a unique and interesting setting for the study of the diversity-labour market nexus. On the one hand, ethnic identity remains distinct even after generations of integration. This is because ethnicity became a salient concept during Apartheid (from 1948 to 1994) when the Apartheid government deteriorated inter-ethnic relationships by reinforcing the ethnic solidarity to prevent black ethnic groups from forming a coalition to fight against the white government (Gradin, 2014). On the other hand, the Apartheid regime has largely destroyed both the regional path dependence in demand of black labour and the intergenerational occupational persistence in labour market outcomes by compressing the educational and job opportunities of the black South Africans universally. The Apartheid government imposes strict labour regulations to prevent the black South Africans from performing semi-skilled and skilled jobs or running their own business in "white" areas. Therefore the post-Apartheid era is the first time since the early 20th century when the majority of the black South Africans could freely make decisions on occupations or set up their own business. Thus contemporaneous labour market outcomes of the black might convey less information on the persistence in regional labour demand and inherited

[1] For example, according to the census data, the proportion of the white over the whole South African population decreases from 18.2% in 1980 to 10.9% in 1996 with the coloured population staying small and stable (13.7% in 1980 and 11.7% in 1996). Among the working age (15-64) population, the black South Africans make up 74.25% of the whole labour force participants. The majority of workers (i.e. those who have a job) are also black (58.81% black, 17.7% coloured and 22.7% white).

[2] There is literature about ethnic diversity at the workplace level, which shows the complementarities between workers from different cultural backgrounds as a rationale for the existence of a global firm (Lazear, 1999b). We argue that it is better to focus on the ethnic diversity in places of residence than places of work in our setting. Firstly, the model we propose in the paper links ethnic diversity to interactions an individual has ever had in his daily life, which is better captured at places of residence. Secondly, we later on show that in the data around 60% of the whole black population do not have a job, which means the information on their place of work is not available. Thirdly the overwhelming majority of black South Africans live and work in the same magisterial district (i.e. the geographical unit in our analysis). For example, in 1996 census data, the correlation between district of work and district of residence among the whole black population who are employed is 0.98. Therefore it does not make too much sense to distinguish the two concepts.

abilities which are the confounders in our analysis.

Baseline results, based on 1996 and 2001 census data, show that black individuals are more likely to be employed in a more ethnically diverse district. Especially they are more likely to work as an employee, as opposed to setting up their own business.

One challenge in interpreting this as a causal relationship is that the formation of ethnic diversity in a district may not be random. For example, if a district has more job opportunities or higher levels of development, it will attract people from diverse ethnic backgrounds. These people will be more likely to be employed simply due to the higher labour demand in those districts. Or if people with some specific characteristics (i.e. higher ability) are attracted by more ethnically diverse districts, they might also perform better than their counterparts wherever they go. A simple OLS regression will generate biased estimates of the effect of ethnic diversity on employment.

We therefore turn to an instrumental variable strategy, which relies on the location of historical black settlements (known as "homelands"). Following the standard assumption in the literature about migration (Alesina et al., 2015), we assume that the magnitude of migration decreases with the distance between the original homelands and the destination districts. In particular, our instrument exploits the fact that a district tends to host a more diverse population if it is equally distant to multiple homelands. On the contrary, a district becomes more homogeneous if it is relatively close to one homeland but far away from the rest. Importantly, the equidistance to multiple homelands remains a strong predictor of ethnic diversity even after controlling for the proximity of the district to the closest homeland. This further confirms that what can be captured by this instrument is not purely the absolute distance to these homelands but the equidistance to multiple homelands.

In our main IV regressions, a one standard deviation increase in ethnic diversity index in 1996 (2001) increases employment rate by 2.98 (4.56) percentage points, which is 8.12% (13.04%) of the average employment rate in 1996 (2001). This positive effect only holds for the black ethnic groups with relatively large population size.

We further decompose ethnic diversity index into the inverse of the number of different ethnic groups and the dispersion of group size. A clear investigation of the mechanism through which ethnic diversity works on labour market outcomes requires disentangling these two components. This can also be solved with our instrumental variable approach. By construction, the number of ethnic groups is fixed in our instrumental variable, which is exactly the number of historical homelands. Therefore the only variation in the instrumental variable comes from the difference in the distance between the destination district and different homelands, which captures the difference in the population size among different ethnic groups in the destination. Both OLS and IV regressions based on this decomposition shows that when the number of ethnic groups is fixed, a more even distribution of group size (which leads to a higher degree of ethnic diversity) in the district of our interest can increase the employment rate of the black South Africans.

We propose a model of a coordination game in the spirit of the literature on social interaction to explain these findings. Utility comes from both intra- and inter-group communication. We assume inter-ethnic communication is more costly than intra-ethnic connection (because one needs to overcome barriers such as language). Given the number of ethnic groups, a more ethnically diverse place has less dispersion of group size, which implies a higher rate of inter-ethnic communication needed to maintain the overall level of social connection. Therefore it is

more necessary for people in ethnically diverse districts to invest in social skills to be able to communicate with those outside their own groups. Their labour market outcomes will improve accordingly as these additional social skills can help them in finding jobs, either by reducing search cost or by improving their productivity.

Our key mechanism can also explain why only groups with large population size respond to ethnic diversity. Starting from the situation where everyone in the district invests in social skills in order to participate in inter-ethnic communication, groups with larger size are more likely to deviate from this coordination because they can get enough social connection by intra-ethnic communications. This is especially the case in an ethnically homogeneous place where they are the dominant groups, but is less likely to be the case if the district is more diverse as their population share becomes smaller. For groups with smaller size who heavily rely on inter-ethnic connection, they do not have the incentive to deviate and will always participate in inter-ethnic interaction and invest in social skills regardless of the diversity level.

We conduct both numerical simulation and analysis based on real data to verify the key mechanism of the model. We fix the number of different ethnic groups and explore the dispersion of group size. The results consistently show that holding other parameters constant, less dispersion of group size (i.e. larger diversity) incentivises people to invest in social skills. Our numerical simulation also shows that our results can be reconciled with papers finding the negative correlation between ethnic diversity and economic development, as the level of investment in social skills can potentially decrease with diversity when per unit cost of investment is too high. IV regressions similar to our main analysis based on 1996 census data also find that our proxy of social skills increases with ethnic diversity and this effect only exists among groups with large size.

Contributions This paper contributes to the literature in four ways. Firstly, we find an innovative way to capture the exogenous variation in ethnic diversity. Our instrumental variable has advantages over instruments exploring simple geographical features. For example, distance to certain places is commonly used as an instrument for migration but whether this is orthogonal to economic conditions has been challenged.[3] By construction we control for the distance to the closest homeland and explore the remaining variation in equidistance to multiple homelands, which could be less problematic than the simple distance measures. Alternatively, one can use the historical ethnic diversity directly as an instrument for contemporary diversity, as explored in Miguel and Gugerty (2005) who use the historical distribution of ethnic residence in two districts in Kenya as an instrumental variable to study ethnic diversity and public goods provision. Such a historical distribution of ethnic settlements might also be correlated with other factors. For example, they find that places where several settlements intersect are in lack of sufficient public goods provision. This might however just be because public policies are less effective at the border between different districts in general, whether or not these districts represent a diverse composition of ethnic groups. Our instrument mitigates this violation of exclusion restriction by focusing on districts outside these settlements instead of the settlements themselves. More importantly, by construction we can have places relatively far from all homelands but still with reasonably high ethnic diversity level as long as they are equidistant to all homelands. These places are less likely to be affected by the initial conditions of original homelands.

[3]For example, a place close to an economic centre might get the positive spillover from the centre, or a place close to the road might perform better than others simply because the demand for road is higher in a better place.

Secondly, our instrumental variable also manages to disentangle the two components in diversity index: number of groups and dispersion of group size. In our instrument the number of different groups is fixed (i.e. the number of homelands is fixed), and the variation only comes from the dispersion of group size. Therefore ethnic diversity has a clear interpretation in our story: a more diverse place means the distribution of group size is more even. Accordingly the employment opportunity is driven by the degree of dispersion of group size, which is directly related to our theoretical model.

Thirdly, we contribute theoretically to the mechanism through which ethnic diversity affects economic performance. Traditional network literature emphasises the importance of group size. In particular, social connection increases with the size of own group, which means the network effect decreases with the degree of ethnic diversity. This indicates a negative association between ethnic diversity and socio-economic outcomes and contradicts our empirical findings. We propose that what drives our whole story is not the absolute amount of social connection but the composition of social interaction. A more ethnically diverse place does not necessarily have more total amount of social interaction, but it has more skill investment because a larger proportion of communication takes place across ethnic lines, which is more challenging than intra-ethnic communication and therefore motivates people to invest more in skills. Furthermore, traditional explanations on why diversity improves labour market performance, such as knowledge spillover, skill complementarity and discrimination, are not completely compatible with our empirical evidence.[4] Our model of coordination game provides a new perspective on how ethnic diversity positively affects labour market outcomes.

Moreover, our mechanism expands the literature on the importance of skill composition in labour market by linking skill mix to ethnic relations. Labour economists have highlighted the importance of skill mix in the labour force (Acemoglu and Autor, 2011). In particular, higher social skills in the workplace can facilitate people's trading of tasks based on each other's comparative advantage, therefore increasing overall productivity (Deming, 2017). Taking a step back, we provide some insight on how to motivate the acquisition of these social skills in preparation for the labour market. Our mechanism shows that this could potentially be achieved by encouraging ethnic diversity of their communities and stimulating inter-ethnic communication.

Fourthly, we contribute to the literature on South African labour market by emphasising another dimension of inter-group relations in addition to black-white divisions, and showing this also has important implications on labour market outcomes of the black. Studies on South Africa have been focusing on the segregation between black and white while each group within the black population is implicitly seen as being homogeneous. What we show in this paper is that each black ethnic group has distinct features and the inter-ethnic relationship amongst the black population is important in their economic opportunities.

Focusing on the within-black ethnic diversity can also deal with the major obstacles to contemporary unemployment amongst the black South Africans. Banerjee et al. (2008) propose that the stagnancy of the high unemployment rate among the black in post-Apartheid South Africa might be mainly due to high search cost in job hunting and little growth in the informal sectors. On the one hand, social skill acquisition in an ethnically diverse district can reduce this high search cost. On the other hand, as the informal sector is not powerful enough to provide more employment opportunities, black South Africans still rely heavily on jobs in formal sectors

[4]Detailed discussion is in the theoretical section of the paper.

where skill complexity is required and social skills can be very important.

Related Literature This paper mainly relates to two strands of literature. The first one is the empirical analysis on the relationship between ethnic diversity and economic development. A general perspective is that ethnic diversity is negatively associated with economic opportunities at the regional level. It is the case especially in African countries characterised by high ethnic fragmentation (Michalopoulos, 2012; Michalopoulos and Papaioannou, 2013).[5] Ethnic fragmentation harms the economic performance in these countries as it is associated with under-investment in public goods (Michalopoulos and Papaioannou, 2013), conflict (Amodio and Chiovelli, forthcoming) and collective action failures resulting from difficulties in imposing social sanctions in diverse places (Miguel and Gugerty, 2005).

Discussions at the micro level are relatively scarce. There is some firm-level microeconometric evidence on the direct effect of ethnic divisions on workers' productivity in Kenya which documents that upstream workers undersupply downstream workers at the sacrifice of total output if these people come from different ethnic groups (Hjort, 2014). Another strand of literature looks at how entrepreneurs from a specific ethnic group make use of their ethnic networks to develop social capital and mobilise resources (Iyer and Shapiro, 1999), but this is not directly linked to ethnic diversity. Thus, how the level of ethnic fractionalisation affects labour market outcomes remains unclear.

Some papers established a causal relationship between ethnic diversity and economic outcomes. The first approach relies on the exogenous change of ethnic diversity in the time dimension, for example due to the implementation of new jurisdictions (Alesina et al., 2016a). The second approach is based on natural or quasi-experiments which directly affect the level of ethnic diversity. For example, Algan et al. (2016) explore an exogenous allocation of public housing in France at the apartment block level and Dahlberg et al. (2012) make use of a policy on the compulsory allocation of refugees in Sweden. In South Africa, however, ethnic diversity does not change dramatically over time, which means there is not enough time variation to identify changing levels of diversity. It is also hard to find proper natural or quasi-experiments due to the political sensitivity of ethnic topics in this country. Therefore, the above two commonly established identification strategies in the current literature are not feasible in our setting.

The second strand of literature concerns the theoretical models on social interaction. There are two key differences between our model and several models documenting social interactions in response to diversity in current literature. On the one hand, unlike models relying on the intrinsic ethnic-specific parameters of taste, preference or discrimination (for example, Morgan and Vardy (2009) shows minority candidates produce noisier signals of their ability), we show that ethnic diversity still affects people's decision in investments in social skills without documenting those assumptions. This is in line with the recent finding that ethnic diversity can be independent of cultural diversity (Desmet et al., 2017). On the other hand, unlike Glaeser et al. (1992) which requires that communication is more extensive or the amount of social connection is larger in more diverse places (Alesina and La Ferrara, 2000), in our model the overall level of social interaction does not necessarily increase with ethnic diversity (overall social interaction is the

[5]More research in developed world finds support for the positive side of diversity (Andersson et al., 2005; Niebuhr, 2010; Ottaviano and Peri, 2006). The relationship between diversity and economic performance can also be non-linear. For example, Nikolova et al. (2013) use data from the post-soviet states and show that entrepreneurship is increasing in ethnic heterogeneity at low level of diversity, while it loses its positive impact when diversity reaches a certain threshold.

sum of intra- and inter-ethnic connections). Ethnic diversity results in more investments in social skills because inter-ethnic communication is more costly (or requires more skills) than intra-ethnic connection.

The mechanism in our paper is the closest to, yet distinct in important aspects from, two existing papers. In the story in Lazear (1999a), he finds that immigrants to the U.S. have higher English proficiency when there are smaller proportions of people from their native country in the communities in their destination. Our paper also documents that people are incentivised to learn English to have access to more potential communication partners (in our story we generalise "language" to a broader concept of social skill). The key difference is that they focus on the assimilation of the immigrants to the U.S and therefore the majority group (i.e. the U.S. native) do not respond to the diversity level in different communities. However, both the theoretical model and empirical findings in our paper show the opposite - only groups with large size (analogue to the U.S. native in his paper) are affected by ethnic diversity whereas smaller groups (analogue to the minority group of immigrants in the U.S.) behave indifferently between ethnically diverse and homogeneous places.[6] What generates this difference is that his model is featured by unilateral assimilation of the immigrants to the U.S. while in our model social interaction and skill investments are bilateral. This makes more sense especially in ethnically diverse places where no ethnic group has overwhelming group size. Also due to strong ethnic identities, groups with smaller size will invest in a common or official language rather than the language of the large group. In our modelling part, we show further that unilateral assimilation is not consistent with our empirical results.

In another model on social interactions between different groups, Alesina and La Ferrara (2000) assume that individuals prefer to communicate with people with similar income, race or ethnicity and conclude that homogeneous communities have higher levels of social capital. Instead of making the direct assumption of group-based preference, we treat this as an implicit implication of the model and argue that people have preference towards groups similar to them because the cost of intra-ethnic communication is lower.

The paper unfolds as follows. In Section 2, we provide a historical overview of the pattern and formation of ethnic diversity as well as summary statistics on labour market in South African context. In Section 3, we describe the data sources and how we construct the variables of our interest. Section 4 details the empirical methodology, focusing on the instrumental variable and its validity. In Section 5, we comment on the results about how ethnic diversity affects labour market outcomes in post-Apartheid South Africa and how this impact differs across sub-groups. Section 6 proposes a theoretical model with numerical simulation and empirical evidence to explain the main empirical results and rule out some alternative explanations. Finally we draw some conclusions and policy implications in Section 7.

2.2 Institutional Setting

2.2.1 Ethnic groups in South Africa and the formation of ethnic diversity

None of the black ethnic groups are indigenous in South Africa. All of them migrated from eastern and central Africa to southern Africa starting from centuries ago, as part of the so-

[6] We control for the proportion of the black over the whole population in our analysis and focus on within-black communication.

called "Bantu migration".

Before explaining the narratives, two concepts should be made clear. The first is "homeland" which refers to the original settlements of those ethnic groups when they first moved to South Africa. The second is "white areas" or "white South Africa"[7] which refers to places in South Africa outside those homelands. Many years after arrival in South Africa, those black people moved out of their original homelands and ended up in these "white areas" due to different reasons, mainly the pressure of conflicts with the British and Dutch colonisers as well as other ethnic groups. Therefore, "white areas" are not areas where only white people reside, but places outside original black homelands (the proportion of the black over the whole population can still be large in those "white areas").

Based on Mwakikagile (2010) and Gradin (2014), we provide historical narratives on the mass migration of ethnic groups from central Africa towards South Africa, the original settlements of these ethnic groups and the migration of these people out of their homelands to "white areas" in South Africa. The timeline about the history of the settlements and migration of the black ethnic groups outside their own settlements up to the time of South Africa's independence can be found in the upper panel of Figure 2.1.

The indigenous groups in South Africa are San and Khoikhoi (both are "coloured" groups) residing in the southwestern and southeastern coast about 2000 years ago. Around 700s A.D., black Africans had settled in the northern part of what is South Africa today.[8] They were members of different Bantu ethnic groups who had moved southward from East-Central Africa (the Great Lake district around Congo) and spoke related languages.

Ethnicity-specific information on the Bantu migration from eastern and central Africa towards South Africa and the formation of ethnic diversity in "white areas" are summarised in Appendix A.1. The table contains information on the timing of their migration into South Africa, geographical location of original homelands, timing of migration outside homelands and the Bantustans assigned to them during Apartheid (which will be explained in constructing our instrumental variable). For example, Zulu are believed to be descended from a leader named Zulu born in the Congo Basin area. In the 16th century, they migrated to the south and eventually settled in the eastern part of South Africa, an area now known as Kwazulu-Natal. The Zulu empire in the 1800s witnessed their vast migration and expansion of territory.

One indication from the narratives is that the black had settled in the country long before Europeans arrived. For example, the diaries of shipwrecked Portuguese sailors attest to a large Bantu-speaking population in present-day Kwazulu-Natal by 1552. In 1652 Jan van Riebeeck and about 90 other people set up a permanent European settlement as a provisioning station for the Dutch East India Company at Table Bay on the Cape of Good Hope, beginning the era of European colonisation.

Due to the pressure from the potential conflicts with white colonisers and the other ethnic groups, the nine black ethnic groups began to move out of their homelands or change their territories. By the early 1700s, there were already some African groups migrating into the interior of the country to shield themselves from European domination. By 1750 some white farmers, known as Boers, expanded to the region where they encountered the Xhosa and Zulu. Starting from 1789, a series of wars and conflicts over land and cattle ownership broke out

[7] It became an official terminology during the Apartheid regime.

[8] Some argue it is as early as the third century (Gradin, 2014).

between the Boers and the black ethnic groups. In early 1800s the British replaced the Dutch at the Cape as the dominant force. The Boers, defeated by the British, migrated eastwards into today's Kwazulu-Natal and Free State where the conflicts between the Boers and Zulu people continued. Many other ethnic groups have encountered similar conflicts.

The destination of their migration is not well-documented. This information, however, can be reflected from today's distribution of ethnic groups across South Africa. This pattern of migration will also affect today's distribution of ethnic diversity. For example, a place would be more diverse potentially if more ethnic groups moved in. Details will be shown in the next section. One thing which needs to be emphasised here is that in most of the cases the key driving force of emigration from ethnic homelands is the conflict either with the white or with other ethnic groups rather than the economic benefits in the destination.

Importantly, further evidence shows that the mass migration both from central to southern Africa and from homelands to "white areas" within South Africa took place mainly before the spur of industrialisation and modern economy. The discovery of mineral resources is a milestone in the economic development and transformation towards modern South Africa. Diamonds were first discovered in 1867 along Vaal and Orange rivers, and in Kimberley in 1871. In 1886, gold was first discovered in Witwatersrand, around today's Johannesburg, which stimulated trade and construction in large dimensions. All this took place after the Bantu migration. This means the migration from homelands to "white" areas, although not completely random, may not be purely driven by the economic prospects in the destination.

In 1910 the Union of South Africa was established, which declared the superior socio-economic status of the white politically and created a white-dominated society. Since then racial discrimination has been a prominent feature of South African society even before the official institution of Apartheid, and the mobility of the black was largely restricted.

Summary demographic statistics about the nine ethnic groups are reported in Table 1.1 for 1996 data and Table 1.2 for 2001 data. The distribution of population share among these nine groups and their labour market outcomes are similar in these two years. In both 1996 and 2001 there are three out of nine ethnic groups (Xhosa, Zulu and South Sotho) who have relatively large population size (i.e. their share of the whole population is over 20%). We define them as *large* groups. Another two ethnic groups have smaller size (Tswana and North Sotho), and are therefore defined as *medium* groups. The remaining four ethnic groups have much smaller population share (less than 5%) and are defined as *small* groups.

2.2.2 The role of Apartheid in shaping inter-ethnic relations and labour market outcomes

Since mid-1900s, inter-ethnic relationships and labour market outcomes have been significantly shaped by the Apartheid regime and related regulations. The regime reinforced the ethnic identity and destroyed much of the path dependence in the opportunities for education and labour market for the black. The timeline of the Apartheid regime can be found in the lower panel of Figure 2.1.

Starting in 1948, the ruling Afrikaner National Party (NP) implemented a program of *apartness* and formalized a racial classification system, which transformed into official *Apartheid* by the 1951 *Bantu Authorities Act* and 1953 *Bantu Self-Govern Act*. Each individual living in South Africa belonged to one of the four races (White, Indian, Colored, Black), which essentially de-

fined an individual's social and political rights. In addition, the government over-emphasised the differences among the various ethnic groups, in the spirit of the *"divide et impera"* principle. The ethnic segregation, on top of the racial separation, was to guarantee the political and economic supremacy of the white minority. This exacerbated division of ethnic groups served as a tool for the white to control the black in an easier way (Gradin, 2014).

With the introduction of the *Promotion of Black Self-Government Act* in 1959, the government delimited a number of scattered rural areas as "native reserves" for blacks (called "Bantustans"), one for each ethnic group. The designated areas for the reserves amounted to 13 percent of the total South African territory, while the blacks accounted for more than 75 percent of the total population. Blacks' land ownership was restricted, as well as their ability to freely move and settle in the white South Africa. Internal migration was severely regulated until the repeal of the *Pass Laws Act* in 1986. With the forced removal of the blacks from the "white areas" of South Africa, the Bantustans became over-densely populated territories, where land was overgrazed and afflicted with serious soil erosion. The economic development of these reserves never materialized, leaving their inhabitants in acute poverty (Christopher, 2001). In 1970, the regime promulgated the *National States Citizenship Act*, which provided citizenship to blacks in their homelands. The ultimate aim was to create a number of ethnicity-based independent states.

In conclusion, the Apartheid regime used separation along racial lines and ethnic lines as a fundamental device for the demarcation of physical and social boundaries for all interactions.

One thing which needs to be pointed out is that Apartheid did not shift the big picture of the magnitude and distribution of ethnic diversity in these "white areas", despite the campaign of forced-removal during this time. During the Apartheid period, 3.5 million (equivalent to $\frac{1}{5}$ of the black South African population in 1980) were forcibly removed from their homes and dumped in areas designed for the black by the Apartheid government. However, our data shows that this forced removal did not lead to large changes in the pattern of the distribution of black South Africans across "white" districts. In 1996 census data, 79.61% of the black population in the "white areas" of our interest never moved in their life. 11.82% moved within their birth district and only 6.63% migrated across districts. These inter-district migrants did not dramatically change the ethnic diversity of "white" districts, as we find the high correlation of district-level ethnic diversity between 1996 and 1985 (the correlation is 0.88, calculated from 1985 and 1996 census by the authors). Therefore it is still reasonable to link contemporaneous distribution of ethnic diversity to the location of historical homelands, despite the large campaign of black migration during the Apartheid era.

The Apartheid regime also severely limited the job opportunities and resources among the black (Posel, 2001). The *Bantu Education Act* of 1953 ensured that non-whites received a substandard quality of education, while access to occupation was regulated by the 1956 *Industrial Conciliation Act*. Whites were authorized to determine the racial allocation of jobs (Mariotti, 2012) and to reserve certain professions for themselves, especially in the manufacturing sector. In particular, the black were banned from semi-skilled and skilled occupations. Similarly, blacks were not allowed to run their own businesses in white areas. In fact, only with the advent of the democracy, in 1993, non-whites were able to make their free occupational choices. This, together with the reallocation of industries, changed the industrial and occupational structures in white areas, which partly weakened the path-dependence in regional demand of black labour.

Moreover, the intergenerational occupational persistence, which has been shown to be particularly relevant for employment (Sørensen, 2007; Pasquier-Doumer, 2012; Magruder, 2010), does not represent a very important issue in the early post-Apartheid era. In other words, blacks may rely more on resources outside their families in overcoming the entry barriers to jobs (barriers such as information about trade partners and market opportunities, informal credit and insurance arrangement).

2.2.3 Labour market in post-Apartheid South Africa

High unemployment rates and large proportion of discouraged workers remain important issues in the South African labour market in the post-Apartheid era (Bhorat and Oosthuizen, 2005; Leibbrandt et al., 2009). Based on 1996 census data, over 60 percent of the working-age black population are either unemployed or out of labour force. A large share of the unemployed in 2005 have never worked in their life. To make things worse, skill-biased technological changes lead to an increase in capital-labour ratio in late 1980s and the whole 1990s, further reducing demand for unskilled labour. At the same time, real wage has been stable or decreasing between 1995 and 2005 (Banerjee et al., 2008). The increase in the supply of unskilled labour, together with the shrinkage in labour demand due to skill-biased technical change as well as the exodus of the white (who are the owners of capital and factories) largely leads to this persistent unemployment issues in the contemporary South African labour market (Banerjee et al., 2008). Furthermore, there is a very low informal employment rate in South Africa, which is only 7.7% - 9.7% based on various measures of informality in September 2004 Labour Force Survey (Heintz and Posel, 2007), possibly because there are also entry barriers in those informal sectors (Kingdon and Knight, 2004). This means the formal wage-employed sector is still the main force in absorbing increased labour supply.

Summary statistics on labour market outcomes based on 1996 and 2001 census data confirm this pattern. In Table 1.1 and Table 1.2, in the overall sample, less than 40% are employed over the whole working-age black population, among which self-employment rate is particularly low (3.2% in 1996 and 2.3% in 2001). The slight rise in unemployment rate from 1996 to 2001 is consistent with the current finding that unemployment rate peaked between 2001 and 2003 in South Africa (Banerjee et al., 2008).

There is, however, large heterogeneity among different ethnic groups. In general, groups with medium and small sizes are more active in the labour market and more likely to be employed, both in self- and wage-employed jobs. This indicates that groups with smaller size are in general more active in the labour market and more competitive in job search, which can be explained by the theoretical model later on in the paper.

2.3 Data

For our empirical analysis, we make use of different data sources. We rely on census data for main analysis. There are three years of census data in the post-Apartheid era: 1996, 2001 and 2011, all of which are the 10% sample from the original national sample in publicly available sources. We do not use 2011 census as both the classification and boundary of magisterial districts have changed dramatically after 2001, making it less reliable to match the new system of magisterial districts in 2011 to the older ones. More importantly, in publicly available 2011 census data,

there is no information on which magisterial district each individual resides in. As respondents in 1996 and 2001 census cannot be matched, we use them as two separate cross-sectional data-sets.

The unit of analysis is the Magisterial District (MD).[9] There are 354 magisterial districts in South Africa, with an average territory size of 3447.5 km^2 and average population size of 0.1 million in 1996. It is particularly convenient to use the MD as a small-scale geographical unit for comparative analysis, given that all other administrative divisions have been revised and re-demarcated repeatedly since the first democratic election in 1994. It also provides a reasonably large geographical unit to define labour market. Our final sample consists of 210 districts in 2001 census (205 in 1996 census), which are the "white" areas outside the historical homelands. Take 2001 census as an example. The excluded districts are either part of the homelands and thus had distinct political status and partially different laws and labour market regulations (124 districts)[10], or districts where the black population in 2001 accounted for less than 1% of the overall population (11 districts[11]), or they cannot be matched with 1985 census data that is explored in the instrumental variable approach (9 districts).[12]

Status in employment. In both 1996 and 2001 census data, we construct an individual-level binary variable for unemployment. The dummy takes value 1 if one is unemployed or economically inactive and 0 if one is employed (either self-employed or an employee). Among workers who are employed, we also consider the allocation of them between self-employment and wage-employment jobs. More in details, an individual is considered to be self-employed if s/he declares to be either self-employed, an employer or work in the family business. To do this, we create another dummy variable only for employed people. It equals 1 if one is self-employed and takes value 0 if s/he declares to be an employee. We only consider working-age black population (15-64 years old).

Ethnicity. Following Amodio and Chiovelli (forthcoming), the ethnolinguistic group each individual belongs to is identified using the information on the first language they speak in the 1996 and 2001 census. There are nine black ethnic groups in the country: Xhosa, Zulu, Swazi, Ndebele, North Sotho, South Sotho, Tswana, Tsonga, and Venda. Following Desmet et al. (2012), we rely on Lewis' *Ethnologue* tree of ethnolinguistic groups (Lewis et al., 2009) to build our measures of ethnic diversity.[13] For each magisterial district and census year, we calculate the relative shares of each ethnic group within the black population and combine them into ethnic diversity index: the *fractionalisation index*.[14] Universally used in the empirical

[9]We calculate the ethnic diversity of the magisterial districts where individuals reside in. There are three reasons why we do not use district of work for the main analysis. Firstly, the mechanism we provide in this paper regarding how ethnic diversity affects labour market outcomes is more related to the districts where one resides (i.e. places where one has social interaction even before entering the labour force) than where one works, which we will explain in the theoretical model. Secondly, the correlation between district of work and district of residence is very high so that they provide similar information. Thirdly, more than half of the black population are unemployed or out of labour force. Therefore the information on their district of work is unavailable and has to be replaced by the information on district of residence, making the district-level information among this group and that among the employed people less comparable.

[10]The boundary of the homelands does not coincide with the boundary of contemporary MD. Taking a conservative method, we define district with less than 10 % overlap with homelands as "white" districts.

[11]This figure is 16 in 1996 census data, which is why the total number of districts of our interest is 205 in 1996.

[12]OLS regression results remain unchanged if we include the nine districts which cannot be matched with 1985 census data.

[13]The nine black ethnolinguistic groups of South Africa belong to the Niger-Congo language family and correspond to level 11 in the tree of ethnolinguistic groups.

[14]We consider another index: polarization index in the robustness check (results are in the online version). It has been proved that fractionalisation index performs better in explaining economic outcomes than polarisation index (Alesina et al., 2003).

literature on ethnic diversity (Desmet et al., 2017; Easterly and Levine, 1997; Alesina et al., 2003; Alesina and La Ferrara, 2005), the ethno-linguistic fractionalisation index (ELF) is a decreasing transformation of the Hirschmann-Herfindahl concentration index and is defined as

$$ELF = 1 - \sum_{k=1}^{m} s_k^2$$

where s_k is the population share of ethnolinguistic group k and m is the overall number of groups. Intuitively, the index measures the probability that two individuals who are randomly drawn from the population belong to different ethnic groups. A larger value of the fractionalisation index indicates higher level of diversity in the magisterial district.

Figure 2.2 shows how ethnic diversity, measured by the ELF index, is distributed in the districts of our interest in 1980, 1985, 1996 and 2001. Districts in darker colours are those with higher ethnic diversity. There is large variation in ethnic diversity levels across South Africa. In general, districts in the northeastern part of the country are more ethnically diverse than those in the southwestern part. In addition, some districts in the middle part of the country are the most ethnically diverse ones. These patterns will be explained when we construct instrumental variables. Districts coloured in white are those inside original homelands, with less than 1% of the black population or that cannot be matched to 1985 census data. A cross-year comparison shows that the degree of ethnic diversity in these districts is very stable. The patterns are extremely similar between year 1996 and 2001. The spatial distribution of ethnic diversity during Apartheid (1980 and 1985) is slightly different but places with higher (lower) degree of diversity remain ethnically diverse (homogenous) over time. This reveals that the formation of ethnic diversity is a historical event and not largely driven by contemporary migration. A comparison between 1980 and 1996 (or 2001) confirms that the Apartheid regime did not drastically shift the spatial distribution of ethnic diversity.

A more detailed investigation of the distribution of ethnic groups in districts with different degrees of ethnic diversity is in the last column of Table 1.1 and Table 1.2 for year 1996 and 2001 respectively. Ethnic groups with relatively larger population size (e.g. Xhosa) are distributed in more homogeneous places while small groups (e.g. Venda) are in more diverse districts. This implies that homogenous places are dominated by groups with large size over the national population while the distribution of population size over different groups is more even in ethnically diverse districts.

Demographic, socio-economic and geographical controls. From the censuses, we also derive a number of controls, which we introduce in our regressions either at the individual level or as aggregated information at the district level. Individual characteristics include gender, age, educational attainment, marital status and whether one's father is alive. Among the district-level controls, we consider population density, proportion of the blacks, proportion of people working in manufacturing and service sectors, whether the district is mainly rural or urban, and whether there is a river and road crossing the district. Additionally, we introduce other geographical factors, which can potentially shape the economic activities of a region. Starting from the Mineral Resources Data System[15], we compute the density of mine for each district. Our geographical unit here, magisterial district, is large enough to capture activities related to

[15]Mineral Resources Data System, MRDS, is a collection of reports describing metallic and nonmetallic mineral resources throughout the world. Spatial data is available at: https://mrdata.usgs.gov/mrds/.

the mining sector. Furthermore, the density of mine has two advantages over a simple dummy for the presence of mining activities. Firstly, it takes into account the number of mineral resources in each district as the magnitude of the effect of mines can increase with the number of mines available at the district level. Secondly, it captures the fact that mineral resources have larger economic effects in more condensed districts either due to higher population density or lower travel cost to the mines. In order to account for the agricultural suitability of land, we use the measure of terrain ruggedness from Nunn and Puga (2012).[16] We also include the measure of soil quality as another proxy for agricultural suitability. Data comes from the Harmonized World Soil Database from the Food and Agricultural Organization of the United Nations. It is a discrete index ranging from 1 to 7, with a descending order of soil quality.[17] As a proxy for the economic development at the local level, we use the National Oceanic and Atmospheric Administration night-time light satellite images data for 1996 and 2001 (Michalopoulos and Papaioannou, 2013).[18] We also include the number of conflicts in each district as it has been proved to be correlated with ethnic diversity (Amodio and Chiovelli, forthcoming) and potentially affects economic prosperities. "Conflicts" here incorporate violence outside the context of a civil war, including violence against civilians, militia interactions, communal conflict, and rioting. A detailed discussion of conflicts in post-Apartheid South Africa can be found in Amodio and Chiovelli (forthcoming).

The rationale of taking into account these control variables is to control for the main drivers of economic development especially employment which are correlated with ethnic diversity. A detailed discussion is in the section about empirical model specification. Details on the sources of data and methods in constructing district-level control variables are presented in the Appendix C.1.

Before looking into the data, it is worthwhile to point out some differences in information collected in 1996 and 2001 census. Firstly, 1996 census distinguishes between those who are unemployed and out of labour force (i.e. economically inactive) while 2001 census combines these two categories. We thus conduct analysis separately as well as jointly for these two groups in 1996 data, and compare the results based on the joint group with the corresponding results using 2001 census.

Secondly, we also explore labour market outcomes other than employment status to enrich our analysis on South African labour market, including wage, income and working hours. information on working hours is only available in 2001 census data. We thus focus on 2001 census in calculating hourly income. In addition, a drawback of the income information in the census data in both years is that it calculates income from all possible income sources, including labour market income, social grant and other sources like bonus, rent or interest. As a result, another dataset (i.e. Labour Force Survey) is required for a more precise measurement of wage, which will be discussed in the empirical results.

Thirdly, 1996 census data asks information on both first and second language spoken whereas 2001 census only asks people about the first language they speak. Therefore, we only look at

[16] We also tried the measure of slope from the same data source. The results are very similar. We do not include ruggedness and slope at the same time as they are highly correlated (the correlation is larger than 0.9), which potentially leads to multicollinearity issues in regressions.

[17] In the soil quality index, 1 = No or slight limitations; 2 = Moderate limitations; 3 = Severe limitations; 4 = Very severe limitations; 5 = Mainly non-soil; 6 = Permafrost area; 7 = Water bodies.

[18] Night-light data is at 30-second grid level. Here we take the average night-time light density within each magisterial district by summing up the night-light measure over these grids and dividing it by area of the district.

1996 census to test our channel of social skill acquisition using proficiency of a second language as a proxy for social skills.

Fourthly, in the robustness check, we reinforce our analysis by looking at natives and migrants separately to see if our results are purely driven by the selection of migrants in each district. For migrants in each district, we have full information on the exact year of their migration to the current magisterial districts only in 1996 census. In 2001 census only migration between 1996 and 2001 is recorded. Therefore, in 1996 census data non-migrants are defined as those who either never moved or moved within magisterial districts and migrants are defined based on cross-district migration. In 2001 census non-migrants are those who did not migrate between 1996 and 2001 or migrated within magisterial districts while migrants are people who moved across districts between 1996 and 2001.

Table 2.1 and 2.2 compare districts whose ethnic diversity is above and below the medium level of ethnic fragmentation in 1996 and 2001, respectively. The last column shows the p-value corresponding to the t-statistics on the difference between districts with high and low ethnic diversity. In both years more diverse places perform significantly better in all indicators of employment, including employment rate, proportion of self-employed people and employees over the whole working-age black population. Among those people who are employed, there is some difference among sectors and occupations. In 1996 census places with higher level of diversity have a larger proportion of people in the manufacturing sector and less in the service sector and this pattern will change once we include our control variables in regressions. Districts with larger ethnic diversity also have less proportion of people in the unskilled occupations among all workers. The similar pattern holds in 2001 census.

The negative correlation between unemployment and ethnic diversity at the district level is further confirmed in Figure 2.3 where we plot the proportion of unemployed (including economically inactive) people over the whole working-age black population against ethnic diversity in each district. The downward-sloping line between these two variables is observed in both 1996 and 2001.

2.4 Empirical Methodology and Specification

2.4.1 Baseline model specification and potential bias

We study the relationship between ethnic diversity among the black population living in "white areas" in South Africa and their labour market outcomes. In particular, we examine whether the within-black ethnic diversity affects blacks' employment opportunities. We start by examining the cross-sectional evidence and investigating the relationship separately for year 1996 and 2001. For both of the years we specify our linear probability model as follows:

$$Empl_{ikdp} = \alpha + \beta ELF_{dp} + \gamma \mathbf{X}_{ikdp} + \delta \mathbf{Z}_{dp} + v_{ikdp} \tag{2.1}$$

where $Empl_{ikdp}$ is a dummy variable for the labour market outcome for individual i of ethnicity k in district d in province p, taking value 1 if one is unemployed or economically inactive, and 0 if employed. We also report the results for wage-employment, self-employment (including self-employed, employer and working in the family business) and the substitution

between wage-employment and self-employment within the subsample of the employed people. ELF_{dp} takes the value of the within-black ethnic diversity index (i.e. fractionalisation index computed in Section 2.3[19]) in district d in province p. X_{ikdp} is a vector of individual-level characteristics (age, gender, educational attainment, marital status, whether one's father is alive which is a proxy for family financial and non-financial support). Z_{dp} is a set of both time-varying demographic and economic controls and time-invariant geographical characteristics at the district level, which will be explained in more detail below.

Unobservables which potentially affect employment rate are included in the term v_{ikdp}. v_{ikdp} can therefore be decomposed into the following items:

$$v_{ikdp} = \theta_p + \lambda_k + \epsilon_{ikdp} \tag{2.2}$$

ϵ_{ikdp} is the random error term. θ_p is province fixed effect which mainly controls for historical path dependence in job opportunities in each province, as well as province-level fiscal variables including social grant provision and policies on taxation and redistribution. There is also evidence that there is inequality between ethnic groups (Alesina et al., 2016b) and that the gaps between different ethnic groups lie in their demographic structure, location, education and labour market outcomes (Gradin, 2014). Therefore we introduce λ_k, ethnic group fixed effects, which allows us to control for mechanical compositional effect and ensures we are comparing individuals from the same ethnic group across districts exposed to different levels of diversity.

Cross-sectional estimates suffer from omitted variable bias originating from ϵ_{ikdp}. For example, the existence of a local economic centre in the district could both create the demand for labour and encourage diversity, in that job opportunities attract individuals from other districts with different ethnic backgrounds. Or more energetic individuals with higher work spirits, who are intrinsically more likely to be employed than the average population, may sort to more diverse districts which have more active atmosphere. In these cases, our results will suffer from upward bias as both ethnic diversity and employment rate are positively correlated with the unobserved district and individual characteristics.

To address the concern that the results are driven by these confounding factors, we first include a rich set of district controls Z_{dp} to limit the information in unobserved items. To account for market size effects, we introduce the population density and urban/rural status of the district. As proxies for local economic development, we use the average night-time light density across 30-second grid areas within each district, and the share of blacks in the district population. For the industrial structure of the district which potentially leads to differences in labour intensity of firms, we control for the proportion of people employed in manufacturing and service sectors. Furthermore, to control for the direct spillover from homelands, we include the distance to homelands. To control for the potential cost of ethnic diversity like conflicts, we add the number of violence in each district in the corresponding years, as conflict has been proved to be associated with ethnic diversity (Amodio and Chiovelli, forthcoming) and potentially job opportunities for the blacks (for example, there might be more closure of factories in more turbulent districts). Finally, to control for agricultural suitability and other geographic factors relevant to the local economic activities we use the terrain ruggedness, the existence of a river

[19]We use the results about polarization index as a robustness check. Results are in the online version.

and a road crossing the district and the density of mineral resources.

The remaining district-level omitted variables are included in ϵ_{ikdp}. Our results will be biased if they are correlated with employment rates. All this will be dealt with using the instrumental variable discussed later on.

Unobserved information at the individual level in ϵ_{ikdp} might also bias the OLS result. We therefore cluster standard errors at the district level to allow for correlation of the error term across individuals in the same district. Furthermore, as a robustness check, we conduct regressions only on people who are born and remain in the districts (i.e. native people) as well as those who only migrated within districts. If the main results still hold among the native, the potential selection of people moving into places with different levels of diversity based on individual-level criteria will not largely drive the whole story. This will be discussed in more detail in the empirical results.

The relationship between ethnic diversity and labour market outcomes can also be investigated at the district level. The results are in the online version of the paper.

2.4.2 Instrumental variable approach

Our instrument for ethnic diversity exploits the historical origins of the location of blacks' homelands. As explained in the institutional setting, the nine black ethnic groups moved long ago from the northern territories of the African continent and settled in different regions of today's South Africa, with one ethnic group occupying one settlements (i.e. defined as "homelands"). Assume the magnitude of migration from the homelands to outside districts decreases with the distance between them and distance is the only determinant in migration. When they moved out of these homelands to the outside districts (i.e. "white" districts which we are focusing on in this paper), the territories that are equally distant to multiple homelands are more likely to be inhabited by individuals with different ethnic origins, and therefore the ethnic diversity will be the highest. On the contrary, places close to only one homeland and far away from the rest become ethnically homogeneous as they have one group dominant in population size migrating from the closest homeland. Visually, this prediction is confirmed by the distribution of ethnic diversity in South Africa in 1996 (Figure 2.2). As shown before, places with relatively higher level of diversity are not necessarily places at the border or close to economic centres of the country, but are those in the middle and northeastern part of the territory surrounded by multiple homelands. Furthermore, districts on the far western part of the country present reasonably high level of ethnic diversity although being far away from all homelands. This is because these districts are still equally equidistant from all the homelands.

We therefore need an instrument that captures the equidistance of each district to all the original homelands. Our instrumental variable strategy proceeds in two stages. First, similar to Alesina et al. (2015), we estimate a parsimonious gravity model of migration based on 1985 census data (i.e. pre-1994 distribution of ethnic groups). We aim at predicting the level of within-black ethnic diversity in each white district d, solely as a function of a factor that is plausibly exogenous to labour market outcomes of the blacks: the distance of the district to the homelands. Second, we start from the predicted stocks to construct a diversity index. Specifically, we estimate:

$$N_{dk85} = \alpha + \beta_1 Dis_{dk} + \gamma_k + \epsilon_{dk85} \tag{2.3}$$

where N_{dk85} is the actual stock of individuals belonging to ethnic group k in (white) district d in 1985; Dis_{dk} is the bilateral Euclidian distance between the centroid of district d and the closest border of homeland for ethnic group k[20]; and γ_k is the homeland fixed effect. The determinants in our model are the ones traditionally used in the related literature (Mayda, 2010; Beine et al., 2013; Ortega and Peri, 2014; Dumont et al., 2010). In particular, the physical distance between two districts (the homelands and the white areas) accounts for the migration costs, while the homeland fixed effects take into account common shocks in living conditions in the original settlement and the stock of population of each ethnic group in homelands, which can also influence migration decision. Following Santos Silva and Tenreyro (2006), we estimate the model by using the pseudo poisson maximum likelihood (PPML) estimator, which better suits the count data in the dependent variable.[21]

By imposing a universal β_1 to all ethnic groups, we assume that the per-unit migration cost is the same for everyone, regardless of their ability and ethnicity. In addition, by ignoring any characteristics of the destination (e.g. population size, economic development and job opportunities) in the gravity model, we impose the condition that the benefit of migration is the same for everyone. Therefore by construction our predicted number of migrants from each homeland is only determined by the distance between homeland and destination.

In principle, the migration stocks could be predicted by 1996 and 2001 data. Nevertheless, we prefer to use the 1985 census data to rule out the selection of migration resulting from the movements of the black population after 1994 (this happened even as early as the repeal of the Pass Law in 1986). In fact, as previously documented (Section 2.2), the blacks had very limited freedom in choosing their own residential location and were strictly regulated in inter-district migration before 1986. After 1986 these constraints were loosened and the blacks had some freedom to decide where to resettle. Therefore, the distribution of ethnic groups in 1985 is less affected by the simultaneous change of labour market conditions and blacks' selection into "white areas" in the post-Apartheid era. Another reason why we use the 1985 distribution of the black population is that the equidistance to different homelands is a feature which stays stable over time. By sticking to 1985 data we can construct an instrumental variable whose value stays the same between 1996 and 2001 to make the IV regression results in these two years more comparable.[22]

Using the predicted stocks $\widehat{N_{dk}} = \widehat{\alpha} + \widehat{\beta_1} Dis_{dk} + \widehat{\gamma_k}$, we calculate the predicted share of ethnic group k in the black population of district d and construct the instrument for the fractionalisation index ELF:

[20]The reason why we use the centroid of the districts instead of capital city is that capital cities are not well-defined at the magisterial district level. We use the border instead of the centroid of the homeland because the shape of the homeland is highly irregular and scattered. Furthermore, the distribution of population within homeland is highly uneven, making the centroid of homeland a less reliable measure in capturing the distance between the destination and the location of potential migrants from homeland.

[21]We do not control for the population size in the destination in the gravity model as it might be endogenously determined by the level of economic development in the destination which potentially affects the flow of migrants into the destination. Here our aim is not to get the most precise estimate of bilateral migration but to construct the counterfactual number of migrants in each district under a hypothetical setting where bilateral migration is only determined by distance between the original homeland and destination.

[22]We do not find much variation in fragmentation index between 1996 and 2001, which means ethnic diversity stays relatively stable over time.

$$\widehat{ELF} = 1 - \sum_{k=1}^{m} \widehat{s_k}^2 \quad \text{with } \widehat{s_k} = \frac{\widehat{N_{dk}}}{\sum_{k=1}^{m} \widehat{N_{dk}}} \tag{2.4}$$

The same instrumental variable approach with the same model specification at the first stage can be applied to district level regressions.

The remaining challenge is to find a proper measure of the original homelands for each ethnic groups. As there is no document about the exact location and boundary of these homelands, we use the territories of Bantustans during Apartheid as proxies for these original homelands. As is discussed in the institutional setting, with the ascent of the apartheid regime, the white-dominated government of South Africa designated specific territories as pseudo-national homelands (i.e. "native reserves", called "Bantustans" in the official documents) for the country's black African population. The Bantustans were organized on the basis of ethnic and linguistic groupings and were a major administrative device for the exclusion of blacks from the "white areas" of South Africa. The location of the Bantustans is based on the government's knowledge and documents about the historical location of homelands of each ethnic group. Ten Bantustans were created for these nine ethnic groups (There are two Bantustans for Xhosa people - Transkei and Ciskei. Other groups each occupies one Bantustan).[23]

To verify that the location and territory of Bantustans can be treated as proxies for the original homelands for the black people, we compare the distribution of these Bantustans and the "Murdock map". This map, drawn by an anthropologist George Murdock in 1953[24], provides the information on what the dominant ethnic group is in each geographical unit on the map of the whole African continent at the end of the 19th century. As reflected in the Murdok's map (panel (a) in Figure 2.4) (each colour represents a certain group dominating the corresponding place in terms of population size), up to the end of the 19th century, each of the nine groups has occupied some specific areas of the country. The Murdock map reveals the distribution of dominant ethnic group in each geographical unit rather than the exact location of original homelands. And the boundary of the geographical units on this map does not coincide with the border of magisterial districts in South Africa. Therefore, the Murdock map can only roughly implies the spatial distribution of each ethnic group in South Africa, which is a result of both the distribution of original homelands and the migration of ethnic groups from these original settlements to other places.

Comparing Murdock's map in panel (a) and the distribution of Bantustans under the Apartheid system in panel (b) in Figure 2.4, we can find large overlaps between the Bantustan and the region where the same group have dominated historically in Murdock's map. For example, places around the Bantustan designed for Tswana people (the dark green part in panel (b)) are also the places dominated by Tswana people (labeled with the same dark green colour) at the end of the 19th century in Murdock's map in panel (a). Therefore, it is reasonable to use the distribution of Bantustans as proxies for the location of original ethnic homelands.

The map in Figure 2.5 presents the value of predicted diversity index together with the

[23]Therefore we treat Transkei and Ciskei as one homeland in the gravity model. When we calculate the distance between each district and the original homeland of Xhosa people, we measure the distance between each district and Transkei and Ciskei respectively and choose the smaller one.

[24]The map has been digitized by Nathan Nunn, starting from "Tribal Map of Africa" which is a fold out map from the book "Africa: Its peoples and Their Culture History" by George Murdock, 1959.

distribution of Bantustans across the country. The white places with slashes are either places which cannot be plausibly considered as "white" South Africa of our interest as they have more than 10% overlap with Bantustans, or places which cannot be matched with 1985 census data. The spatial pattern of predicted value of ethnic diversity in this figure is similar to the distribution of ethnic diversity in Figure 2.2 based on the real data. Again, places with the highest predicted ethnic diversity are those amid multiple homelands (mainly in the middle and northeastern part of the country). A more important feature is that the distance to the closest homeland (proxied by Bantustans) does not completely determine the level of predicted ethnic diversity. That is to say, places close to a specific Bantustan (and far from the other ones) may not be highly diverse. It is particularly the case for the districts around the Bantustans of Transkei, Ciskei, Kwazulu and Bophuthatswana.

Test of validity of the instrumental variable

Identification requires the instrument to capture the ethnic diversity pattern observed in 1996 and 2001 and to be uncorrelated with any other determinants of the blacks' labour market outcomes. The first condition is satisfied provided that: 1) The historical distribution of ethnic groups within the country varies with and is closely related to the distance of the destination region ("white" district) from multiple Bantustans, and 2) Apartheid did not overturn the historical pattern. As for the second condition, the non-randomness of blacks' homelands could cast doubts on its fulfillment. The proximity to the Bantustans might well be correlated with unobserved factors other than diversity, affecting the blacks' labour market outcomes.

However, the instrument exploits the distance to *multiple* ethnic homelands as a predictor for diversity. As mentioned above, the map in Figure 2.5 shows that districts with higher predicted diversity are the ones that are "equally" distant to multiple homelands, and not necessarily the ones that are the closest to a specific homeland. For example, although being contiguous to one of the Bantustans - Transkei (identified with the red color in Figure 2.5), districts in the South-East are among the most ethnically homogeneous areas because they are located at the periphery of other homelands. To further ensure that the instrument only captures the relative distance to multiple homelands and not the proximity to a single Bantustan, in the regressions we control for the distance to the closest homeland. As all the homelands are located at the eastern part of the country, controlling for distance to the closest homeland can also deal with the problem that the instrumental variable might purely capture the west-east division of the country.

We argue that, conditional on proximity to a single homeland, the distance to multiple homelands is as good as random. The most direct narrative evidence is that according to the timeline in the institutional setting, the mass migration of the black largely occurred before the discovery of mines, rise of industrial sectors and modern development. This means the migration from homelands to "white" areas is not purely driven by the higher economic prosperity in the destination.

For a more rigorous test of the validity of our instrumental variable, we run regressions to show that the predicted ethnic diversity index is not correlated with potential confounders which determine ethnic diversity and employment simultaneously, conditional on all the control variables in our first stage regressions. Firstly, we test the correlation between the instrumental variable and potential job opportunities. According to agglomeration economics, economic

centres, as clusters of economic activities, business and capital inflow, may act as the hub of job creation. Therefore, distance to economic centres may capture the potential job opportunities an individual is exposed to, based on the spillover of economic prosperity from the economic centres. There are five main economic centres in South Africa: Cape Town, Pretoria, Durban, Port Elisabeth and Johannesberg. In the validity test we calculate the distance from the centroid of each magisterial district to the closest economic centre and correlates it with predicted fragmentation index discussed above.

The second potential confounding factor is the economic activity of the white. On the one hand, as the Apartheid regime destroyed the self-employment opportunities, leadership and the training towards skilled occupations of the black in the "white" South Africa, the majority of the employers of wage-employed black people are the whites. Although our main regressions focus on the blacks, the population size and the employment status of the whites are also important in determining black people's employment rate, as they might be the providers of potential jobs to the black workers. On the other hand, the dominance and wealth of the white might potentially affect the migration decision of the early black migrants. Black people from different ethnic groups may move to a district where the white behave relatively better as there are more opportunities (or poorer as there is less stress/competition from the white) and thus the ethnic diversity of the black might be correlated with the behaviours of the white. We then calculate the employment rate of the white among their working-age population for each magisterial district in our sample and see if it relates to ethnic diversity of the black.

Thirdly, path dependence also matters in determining contemporary employment opportunities. As the distribution of black settlements is not completely random, the equidistance to multiple original settlements might reveal some socio-economic characteristics besides the distance itself (i.e. customs, early conflict or the distribution of ancient civilisations) which have long-term impact on contemporary development. This persistence of particular socio-economic features is usually a concern in literature which constructs instrumental variables with geographical variables. However, in our special setting, the Apartheid regime before our sample period compressed the opportunities of education, job opportunities and residential choice nationwide among the black and potentially destroyed part of such historical path dependence. If we can show that the path dependence which potentially correlates with equidistance to homelands was largely destroyed by the Apartheid regime due to the shift in residential patterns and the re-allocation of economic activities both for the black and the white, we will be safer to claim that the historical persistence is not likely to affect contemporary employment opportunities directly. As there is no reliable data to reveal the employment pattern of the black during apartheid, we use the employment pattern of the white in 1980 as a proxy for the remaining path-dependence in employment close to the end of the apartheid and see if it correlates with our instrumental variable measured with 1996 and 2001 data. For the employment status of the white in 1980, we do not consider self-employment as the definition of self-employment is not quite clear under Apartheid regime and therefore has large measurement errors.[25] We also consider the population size of the white in 1980.

The fourth potential confounding factor is the magnitude of contemporary migration. Although historical migration was not mainly driven by economic prospects, it might still be the

[25]There are four censuses during Apartheid: 1960, 1970, 1980 and 1985 census. We only consider 1980 census as the data quality is higher than that in 1960 and 1970 census. Publicly available 1985 census data has no information on employment status.

case that contemporary diversity results from contemporary migrants which are driven by economic opportunities. Therefore, we need to show that our predicted diversity does not relate to the magnitude of contemporary migration which refers to cross-district migration ever happening in one's life in 1996 census and cross-district migration between 1996 and 2001 in 2001 census.

Table 3 shows the results on the validity of the instrumental variable based on 1996 and 2001 census data. We regress a set of variables that potentially affect employment rate on predicted fractionalisation index conditional on all the control variables in the main regressions discussed above. Panel A, B, C and D present the tests on the relationship between predicted ethnic diversity and job opportunities, economic activities of the white, path dependence and contemporary migration, respectively. We obtain the coefficients of the tests by regressing the corresponding dependent variables (as reported in the table) on predicted ethnic diversity conditional on all the control variables in the main regression. These dependent variables include: distance to the closest economic centre, proportion of white people who are self-employed over the white population in 1996 and 2001, proportion of white people who are employees over the white population in 1996 and 2001, proportion of white people over the whole population in 1996, 2001 and 1980, proportion of white people who are employees over the white population in 1980 and the number of contemporary migrants in each district. We do not find systematic relationships between these potential confounders and our instrumental variable, which means that the predicted ethnic diversity can be considered as a valid instrumental variable.

Other potential threats to the instrumental variable

This section discusses some remaining potential threats to the instrumental variable which are not likely to be measured with available data.

Firstly, one may argue that the original distribution of ethnic homelands is not completely random. The fact that one place is close to multiple homelands might mean that these homelands are themselves close to each other. Similarly, one possible pre-requisite for a place to be close to only one homeland is that those homelands might be scattered and relatively far away from each other. If the whole region is equipped with better endowments (geography, climate or soil quality) than the others at the time of the Bantu migration from central Africa, this place could attract more than one ethnic groups to establish their homelands, whilst regions with only one ethnic homeland or regions where the distribution of homelands is more scattered might be less attractive in resources and endowments. Therefore, our instrumental variable - the predicted diversity index might just capture the distribution of homelands and the original endowments of the whole surrounding region.

This is not likely to be the case for the following reasons. The first reason is that our instrumental variable captures the equidistance to different homelands conditional on the distance to the closest homeland. By construction places far away from all homelands can still have reasonably high predicted diversity, as long as it is of equidistance to all these homelands. These places are less likely to be affected by the original endowments and resources of ethnic homelands. The second reason is that we have already controlled for geographical endowments (ruggedness, soil quality and river) in each district which are potentially correlated with their initial development by affecting their agricultural production. The third reason is that if our instrumental variable mainly captures the initial economic development and the endowments or resources of the region

rather than ethnic diversity, the predicted diversity index should be correlated with the labour market outcomes among both black and white population. However, as is shown in Table 3, our instrumental variable is not systematically correlated with the employment rate of white workers. Therefore, it is unlikely that the initial endowments in the regions surrounding ethnic homelands challenge the exclusion restriction of the instrumental variable.

Secondly, there is a possibility that districts close to multiple homelands might be the trading centres for people from those homelands whilst trade flows in districts close to only one homeland are less. This might also lead to the difference between these two types of places in the initial economic prosperity and the establishment of cities resulting from trade. Here we show this is unlikely to severely violate the validity of our instrumental variable. Our instrumental variable by construction allows for the case that a place far away from all homelands can be reasonably diverse if it is equidistant to different homelands. And this place is less affected by the initial trade flows among homelands. Furthermore, places with more initial trade flows might become contemporaneous economic centres due to the path dependence in city development and the accumulation of capital and labour. In our validity test we do not find a systematic pattern of the distance to the closest economic centre and predicted diversity index.

Thirdly, one may worry that certain events which attract diverse migrants might happen coincidentally in places close to multiple homelands. For example, the homeland for Tswana group (i.e. the Bantustan of Bophuthatswana) and places in Mpumalanga and Limpopo Province (in the northeastern part of the country) are rich in mineral resources. If our instrumental variable mainly captures the distribution of mineral resources, and if the discovery of mines in a district motivates people of diverse backgrounds to migrate into the district and at the same time boosts economic development, what can be reflected in the predicted ethnic diversity is mainly the effect of mineral resources. In our analysis we have controlled for the density of the mines in each district. More importantly, narrative evidence reveals that the mass migration from central Africa (which can be dated back to the 11th and 12th century) and the emigration from homelands to "white" South Africa happened well before the discovery of mineral resources (mainly starting from the 19th century). Therefore, the discovery of mines and the related events are not likely to violate the validity of our instrumental variable.

First stage results

Table A0 in the Appendix reports the estimated parameters of the gravity model. It suggests that the distance between a white district d and an ethnic group's homeland is strongly negatively correlated with the size of the same ethnic group's population living in district d. Table 4 presents the first-stage regression of the instrument at the individual level. We are interested in both working-age population (age 15-64) and a subsample which have already finished full education (age 25-64). All regressions include province fixed effects and all control variables. Columns 1 and 2 (3 and 4) report the first-stage regression results based on 1996 (2001) census data. In both years the predicted fragmentation index $\widehat{ELF}$ is positively associated with the observed index ELF. The F-statistics are very high in all regressions (i.e. much larger than 10), indicating that the instrument is a very strong predictor of ethnic diversity. Comparing column 1 and 2 reveals that the F-statistics remain stable in both the full sample and the subsample. Comparison between columns 3 and 4 confirms the same pattern in year 2001.

2.4.3 Supplementary approach: district-level fixed effects

The fact that we have two-year cross-sectional census data and that the territory of magisterial districts stay stable between 1996 and 2001 motivate us to find a way to construct panel data at district level as a supplementary approach to the instrumental variable specification, which is to control for district-level confounders directly by including district fixed effect. Therefore we construct a balanced panel by matching the magisterial districts between 1996 and 2001[26] and conduct the district-level model by adding magisterial district fixed effect. Any time-invariant variables in Z_{dp} and θ_p are dropped automatically. Instead we add time fixed effect u_t in the model.[27]

$$Empl_{dt} = \alpha + \beta ELF_{dt} + \delta \widetilde{\mathbf{Z}_{dt}} + \sigma_d + u_t + \epsilon_{dt} \tag{2.5}$$

We report the results of this district-level fixed effects model right after the main analysis.

2.5 Empirical Results

2.5.1 Ethnic diversity and labour market outcomes

Ethnic diversity on employment

Table 5 summarizes the main results on the effect of ethnic diversity (measured by fractionalisation index) on unemployment rate. The dependent variable is a dummy which equals 1 if one is unemployed or out of labour force and 0 otherwise (including people who are self-employed and employees). In 1996 census data which distinguishes people who are unemployed and out of labour force, we create dummies for unemployment and labour force participation and look at how they respond to ethnic diversity separately. Columns 1-6 report the results in year 1996 while columns 7-8 are for year 2001 when unemployed workers and people out of labour force are combined into one category in the original census data. Panel A in Table 5 reports the results based on the cross-sectional OLS regressions at the individual level. Panel B in Table 5 provides the corresponding estimates based on the instrumental variable regressions. We provide results both for the full sample (which gives a lower bound of the effect of diversity on employment) and a subgroup of people aged from 25 to 64 who have finished their education (which gives an upper bound of the effect of diversity on employment). All regressions control for the individual and district level characteristics as well as ethnicity fixed effects discussed above.

In most of the OLS and IV regressions in Table 5 the coefficients of ethnic diversity on

[26] Among 205 magisterial districts in 1996 and 210 districts in 2001, 205 of them can be matched, given that we exclude districts with less than 1% of black people over the whole population.

[27] A potential further specification is to combine the above two approaches and rely on fixed effect-IV approach. The rationale to do this is that some district-level unobservables might change over time which cannot be captured by time-invariant σ_d. In this case, we have the first difference specification:

$$\Delta Empl_{dt} = \alpha + \beta \Delta ELF_{dt} + \delta \Delta \widetilde{\mathbf{Z}_{dt}} + \Delta f_{dt} + \epsilon_{dt}$$

Ideally we can find an instrumental variable for f_{dt}. A similar case to this specification can be found in Dustmann et al. (2017). However, this first-difference specification at district level with instrumental variable is not appropriate here because there is little variation in both the real-world ethnic diversity and the predicted ethnic diversity (i.e. the equidistance to different homelands does not change in the time dimension) over time, which is not sufficient for reliable statistical inference.

unemployment (or labour force participation or these two outcomes altogether) are significantly negative, indicating that within-black diversity increases the rate of employment and labour force participation. Comparing panel A and panel B, the negative and significant coefficients of ethnic diversity remain in IV regressions in many columns. In panel B, comparing columns 2, 4 and 6 reveals that ethnic diversity increases employment mainly by decreasing the number of people who are actively looking for jobs but still unemployed, rather than bringing people into the labour force. Table 5 also shows that the coefficients of ethnic diversity are larger and more significant for the subgroup of people who have finished education than those for the full sample, which confirms that the full-sample analysis gives a lower bound of the effect of ethnic diversity.

We focus on the full sample (i.e. lower bound) to calculate the magnitude of the effects of ethnic diversity on employment based on the results in columns 5 and 7. In panel A in column 5, a one standard deviation increase in ethnic diversity index in 1996 is associated with a 2.15 percentage point decrease in unemployment (including inactivity), which is 3.51% of the average unemployment (including inactivity) rate.[28] Similarly, in panel A in column 7, a one standard deviation increase in ethnic diversity index in 2001 is associated with a 3.88 percentage point decrease in unemployment (including inactivity), which is 5.97% of the average unemployment (including inactivity) rate.[29] Correspondingly, in IV regressions, a one standard deviation increase in ethnic diversity index in 1996 (2001) decreases unemployment (including inactivity) by 2.61 (4.40) percentage point, which is 4.24% (6.91%) of the average unemployment (including inactivity) in 1996 (2001).

Comparing the magnitude of estimates in OLS and IV regressions in both years shows that the magnitude of the effects of ethnic diversity on unemployment rate increases largely between 1996 and 2001 (from 3.51% of the average unemployment rate to 5.97% in OLS and from 4.24% to 6.91% in IV) and IV estimates are slightly larger than OLS estimates. This can be explained by the fact that IV regressions capture LATE for workers at the margin of being affected by ethnic diversity, who might be the most responsive to diversity levels.

We further breaks down employment status into two categories: self-employment and wage-employment. The results (in the online version of the paper) show that within-black ethnic diversity has a positive effect on black's labour market outcomes of the blacks mainly in wage employment. There are two potential reasons to explain why the effect of ethnic diversity on self-employment is not obvious. Firstly, as described in summary statistics, self-employment rate is only 2-3% among the black South Africans, which means the variation of self-employment rate across districts might be too limited for reasonable statistical inference. In addition, measurement errors in self-employment might be large. If these measurement errors are not random, it will also bias our results. Secondly, it is reasonable that self-employment does not respond as much as wage-employment. Current literature relates self-employment, especially small-scale entrepreneurs, to trust, tolerance or cohesive networks (Barr, 1998) and argues that self-

[28] It can be calculated that the standard deviation of ethnic diversity in 1996 is 0.2659. The coefficient of diversity index in panel A in column 6 is -0.081. Therefore a one standard deviation in diversity index decreases unemployment by 0.081 * 0.2659 = 0.0215. From Table 1.1 we know that the average unemployment (including inactivity) rate among the black in "white" districts is 0.613. Therefore this point decrease is 0.0215/0.613= 3.51% of the average unemployment rate.

[29] It can be calculated that the standard deviation of ethnic diversity in 2001 is 0.2586. Therefore in 2001 one standard deviation in diversity index decreases unemployment by 0.146 * 0.2586 = 0.038. From Table 1.2 we know that the average unemployment (including inactivity) rate among the black in "white" districts is 0.636. Therefore this point decrease is 0.038/0.636= 5.97% of the average unemployment rate.

employment increases with trust. Since ethnic diversity can potentially lower the level of trust (Alesina and La Ferrara, 2000, 2002), the overall effect of ethnic diversity on self-employment can be ambiguous.

Table 6 further presents how ethnic diversity affects workers' choice between self-employment and being an employee. As self-employment rate is between 2% - 3% of the whole working-age black population, we drop self-employed people from the whole sample and investigate if ethnic diversity increases the probability of being an employee against unemployed in columns 1 and 5, as most of the effects of ethnic diversity on employment takes place in wage-employed jobs.

Columns 3, 4, 7 and 8 only include employed people and look at the allocation of these workers between self- and wage- employment. The dependent variable equals 1 if one is self-employed and 0 if wage-employed. This is to investigate the effect of ethnic diversity on the potential substitution between self- and wage-employment among employed black population. We replicate the results of the main analysis by restricting the sample to people who are either wage-employed or self-employed (i.e. excluding the unemployed and the inactive). Although the self-employment rate might be too low for enough variations to generate significant statistical inference, we find that the coefficients of ethnic diversity are consistently negative in OLS and IV regressions in both years. That is to say, ethnic diversity helps unemployed individuals get into employment; a large fraction of those newly employed people opt for working for others as an employee.

Focusing on the full sample (excluding self-employed people) in columns 1 and 5, we find that a one standard deviation increase in ethnic diversity index increases the rate of wage employment among the working-age black by 2.29 (3.76) percentage point in 1996 (2001) in OLS regressions, which is around 6.23% (10.67%) increase of the average wage-employment rate in 1996 (2001). In IV regressions, a one standard deviation increase in ethnic diversity index increases the rate of wage employment among the working-age black by 2.98 (4.56) percentage point in 1996 (2001), which is around 8.12% (13.04%) increase of the average wage-employment rate in 1996 (2001).[30]

Ethnic diversity on wage, income and working hours

In this section we investigate labour market outcomes other than employment to get a more thorough picture of how labour market responds to ethnic diversity in post-Apartheid South Africa. We replicate the above individual-level regressions (both OLS and IV) by replacing the dependent variables with other labour market outcomes, including working hours, hourly wage and monthly earnings. As information on working hours is only available in 2001 census data, we only conduct these analyses based on 2001 data. For data on working hours, if values of self-reported weekly working hour are larger than 80, we treat them as outliers and exclude them from regressions. In addition, we trim the income data by excluding values above 5 standard deviation of the mean income. Hourly wage is constructed by dividing monthly earnings by monthly working hours (i.e. four times of weekly working hours).

Data on monthly income in 2001 census includes both labour market earnings and income from other sources such as dividend, rent or social grant. We first report the results based on these rough measures of monthly earnings and replicate the regressions with more precise data on labour market earnings and working hours.

[30]The wage employment rate is 36.69% in 1996 and 34.97% in 2001 after excluding self-employed people.

Panel A in Table 7 reports the OLS and IV regression results on these labour market measures based on 2001 census data. Dependent variables include: log monthly income, log hourly earnings and weekly working hours. As self-employed workers and employees have very different determinants of working hours and earnings, and that ethnic diversity mainly increases wage-employment rate, we only focus on employees in all regressions.[31] Columns 3 and 6 indicate that ethnic diversity does not affect weekly working hours among the employees. Therefore the increase in employment in response to ethnic diversity comes from the extensive margin by increasing employability of unemployed and inactive people, rather than the intensive margin (measured by weekly working hours). And this extension of the extensive margin of labour is not achieved at the sacrifice of decreased intensive margin.

Columns 1, 2, 4 and 5 show some evidence on the increase in both monthly and hourly income among the black employees in response to ethnic diversity. As stated above, information on income in census data incorporates all potential income sources. Therefore we need another dataset which asks information on labour market earnings in particular. We turn to October Household Survey 1996 to replicate all the results in Panel A.[32] We do not choose year 2001 because starting from year 1998 there is no information on the magisterial districts each individual lives in. The results are in Panel B in Table 7. Columns 3 and 6 confirm that weekly working hours are not responsive to ethnic diversity. In columns 1, 2, 4 and 5 the effects of ethnic diversity on measures of labour market earnings are not significant, possibly because the increase in employment can come from both the supply and demand side of the labour market, or because the measures of nominal earnings are not adjusted for price levels (as there is no price or living cost data at the magisterial district level).

2.5.2 Supplementary approach: district-level fixed effects

As a supplementary approach to the instrumental variable approach, we provide estimation results on district-level fixed effects models based on the model specification (2.5) in Table 8. We construct a balanced panel between 1996 and 2001 (205 magisterial districts each). The measures of labour market outcomes (i.e. dependent variables) are: proportion of people who are unemployed or inactive among the whole working-age black population; proportion of employed workers among the whole working-age black population (excluding self-employed people); ratio of the number of self-employed workers versus employees and log monthly income among employees.

Similar to the main IV regression results, larger diversity is associated with more employment, mainly in wage-employment, while there is no significant correlation between ethnic diversity and monthly income. In particular, in district fixed effects regressions we find some evidence that larger diversity is associated with a higher ratio of wage-employment in relation to self-employment.

The magnitude of coefficients in Table 8 are larger than those in Table 5 and Table 6, which can be explained by two possible reasons. Firstly, district-level regressions do not include ethnicity fixed effect which is used to capture ethnicity-specific unobservables that affect the

[31] There are more observations in columns 3 and 6 than others because there are missing values in income and we trim the income values above 5 standard deviation from the mean.

[32] It is an annual survey staring from 1993 (which was renamed as Labour Force Survey conducted twice a year from 2000 and became a quarterly survey from 2008). In 1996 survey 72890 individuals are covered, among which 16082 have information on work status.

labour market outcomes of each ethnicity such as the attitudes towards work and leisure and ethnic-specific skills. It is however not appropriate to include this fixed effect in the district-level regressions due to the potential multicollinearity problem, as the proportion of each ethnic group in a district is already a component of the ethnic diversity index.

Secondly, the relatively larger coefficients of panel regressions might reflect some time-varying district-level unobservables. For example, people are more likely to move to ethnically diverse districts as time goes by as a result of increased benefits in the destination (i.e. the economy of the districts with larger ethnic diversity might grow more rapidly than that in more homogeneous districts). This does not affect our IV regressions as our instrumental variable is not likely to be correlated with the economic development in the destination. However, panel regressions with district-level fixed effects may lead to upward bias of the key estimators as they do not take into account these time-varying unobservables.

2.5.3 Heterogeneous effects of ethnic diversity on employment

Table 9 splits the whole sample into several sub-samples to investigate the heterogeneity in the impact of ethnic diversity on labour market outcomes with individual-level regressions. In particular, we replicate the regressions in the main specification by carrying out the same analysis on these sub-samples. By excluding workers who are self-employed, we use a dummy dependent variable which takes the value 1 if one is an employee and 0 if one is unemployed or inactive.[33] Panel A in Table 9 replicate the same regressions in columns 1 and 5 in Table 6 by splitting the working-age black population into ethnic groups with different population size. Panel B and C look at the allocation of employees among different sectors and occupations in response to ethnic diversity by regressing the probability of working in particular sectors or occupations on diversity index among employees.

Panel A splits the sample by group size. As is shown in Table 1.1 and 1.2, we have three "large" groups whose population share is above 20%, two "medium" groups whose share is between 10% and 20% and the remaining "small" groups each making up less than 5% of the whole black population. We look at these three groups separately and discuss how they are affected by ethnic diversity. The results reveal that only the group with "large" size are positively affected by diversity. None of the columns show that "small" groups respond to ethnic diversity while evidence on the "medium" group is mixed. It is not very likely that the results are purely driven by the lack of power of statistical inference due to smaller sample size. In all the regressions for "medium" and "small" groups, the t-statistics is far from being large enough to generate significant inference. Furthermore, in some regressions the coefficients of ethnic diversity are negative, especially for those in the "small" group in 1996.

In panel B and C, we look at the allocation of industries and occupations among employed workers to show that the improvement in employment rate is not accompanied by the increase in less skilled jobs or the expansion of primary sectors. Otherwise this will lead to a less favourable industrial and occupational structure.

Both 1996 and 2001 census data provides information on the industrial sectors they work, which we classify into agriculture, manufacturing and service sectors. Panel B presents the results on this allocation. There is no evidence to show that ethnic diversity affects the industrial

[33] We also conduct the analysis with a dummy on whether one is unemployed (including inactive people) or not. The results are quite similar.

structure of the districts in IV regressions. This further confirms the idea that the employment opportunities generated from ethnic diversity are not purely driven by the expansion of manufacturing sector due to the revolutionary events like the discovery of mines, nor from the expansion of the primary sector.

We study the allocation of employees further by looking into occupations to show if ethnic diversity leads to a less skilled occupational structure. In both 1996 and 2001 census for each worker there is information on the occupation classified into a detailed 3-digit code. We aggregate this 3-digit coding system into types of occupations based on their skill levels: manager, professional, clerk, service worker, craft worker, skilled worker in agricultural sector, machine operator and unskilled worker. The dependent variables in Panel C are dummies on whether one works in one of these occupations. According to the regression results, ethnic diversity decreases people's chance of becoming a machine operator and increases their probability of being a manager, professional employee and clerk. One common feature is that occupations such as manager, professional and clerk require more language and social skills while the demand for social skills is the least among machine operators. This is closely linked to our mechanism through which ethnic diversity influences labour market outcomes, which will be discussed in the modelling part.

2.5.4 Robustness check

We conduct a series of robustness checks in this section to consolidate the result that ethnic diversity increases employment rate among working-age black population.

Firstly, in the main analysis we use population density and the proportion of black people over the whole population as our control variables. They two altogether capture the information on the total population size of the black in the destination districts. As our census data is a 10% subsample of the original census data and the size of population is calculated with the post-stratification weights, these two variables may suffer from measurement errors. Our first robustness check is to replace these two control variables with the total distance from the destination district to all black homelands. The idea is that if all black people in the "white" districts come from historical homelands and the migration from those homelands to the destination decreases with distance, total distance to all homelands will be a proxy for the pool of black population in the destination. We therefore replicate the main analysis using total distance to all homelands to capture population density and proportion of black people. The results are very similar (results are in the online version).

Secondly, we provide some further evidence on the argument that our result is not purely driven by the sorting of migrants. We show that the positive correlation between ethnic diversity and labour market outcomes does not purely come from the migrants with higher abilities moving to more diverse places and therefore are performing better in job searching. We divide the whole working-age black population into three sub-samples with different levels of sorting: people who were born and stay in the district or people migrating within districts (i.e. "native" people); people moving across districts (i.e. "migrants"); immigrants moving from other countries ("immigrants").[34] In Table 10 we run the same IV regressions[35] as those in the main

[34]Note that "migrants" and "immigrants" in 2001 census data are those who move across districts or countries between 1996 and 2001, whereas in 1996 census they are the people whose last migration was across districts or countries.

[35]OLS regressions have very similar results. We only show the results about IV regressions here.

analysis separately for these three groups in 1996 and 2001. The dependent variables include a dummy on whether one is unemployed and a dummy on whether one is an employee (excluding self-employed workers).

Columns 1 and 4 show that in both years ethnic diversity positively affects the labour market outcomes of native people who are the least likely to sort to places with higher ethnic diversity, as they were born in these districts and remained there, or moved within districts. The positive effect of ethnic diversity on employment also exists among immigrants in columns 3 and 6, the mostly selected sample based on ability and preference (although the number of immigrants in South Africa belonging to one of the nine ethnic groups is very small compared with the whole black population). Interestingly, there is no effect of ethnic diversity on employment among migrants across districts. As we discussed in the validity of the instrumental variable, there are two potential mechanisms of selection among migrants. Either the selection occurs in the original place, meaning people with higher ability choose to move out; or the selection takes place at the destination, meaning people sort to places with higher economic prosperity or job opportunities or more socially active environment when they decide where to move. The result about cross-district migrants here suggest that the first selection mechanism is more important - migrants are of higher ability and therefore behave better wherever they end up, which indicates that the relationship between ethnic diversity and employment is not solely driven by the selection of destinations.

Another potential threat to the interpretation of our results as illustrating a positive impact of ethnic diversity on employment is the emigration of the white after the end of Apartheid. It has been observed that there has been a large emigration of the white out of South Africa after 1994 and that white people moved out of the country for the fear of the worsening economic conditions, weaker government capacity, or the revenge from the black after the nightmare of Apartheid. A place having larger diversity might just indicate that the power of the white is weaker in these places (so that the black community can grow and attract people with a diverse background). If this is the case, there would be more white people emigrating from South Africa in a district with larger ethnic diversity index. The mass emigration of the white may lead to many job vacancies to be filled by black workers, consequentially improving the job opportunities of the black. If this story is true, the correlation between ethnic diversity index and employment rate in a district cannot reflect the impact of ethnic diversity as ethnic diversity index here is just a proxy for the power of the white.

We therefore regress the number of the white in 1996 and 2001 respectively and the difference in the number of white residents between 1985 and 1996 (or 1985 and 2001) on ethnic diversity index for each district, using the same set of control variables. We find in Table 11 that the ethnic diversity index is associated with neither the absolute number of the white population nor the difference in the white population before and after the end of Apartheid (which captures the emigration of the white). This confirms that ethnic diversity is positively related to employment not simply because these places have more job vacancies left by the white people who emigrated from the country.

Thirdly, related literature suggests using Conley's standard errors in regressions to account for the spatial correlation in error terms (Michalopoulos and Papaioannou, 2013). Corresponding to individual-level regressions, we depend on district-level regressions and use Conley's standard errors to replicate the analysis. It is required to set up a cutoff distance above which there is

no spatial correlation. Current cross-country analysis in Africa uses 2000km as the cutoff value (Michalopoulos and Papaioannou, 2013). In our paper, we reduce the cutoff value to 1000km for within-country regressions. Spatially correlated error terms are both implemented in OLS and GMM (using the same IV as that is in the main analysis) regressions. Appendix Table A1 reports all these results. Ethnic diversity still has a positive impact on employment opportunities in all columns in both OLS and IV regressions. Comparing GMM estimates with the previous main IV analysis, we find that the magnitude of the effect of ethnic diversity is in general larger in regressions with spatially correlated errors.

Last but not least, we use non-linear econometric methods to estimate the main regressions. Given that our outcomes are measured by binary variables, we replicate our results by estimating a logit model, a probit model and a probit model with the instrumental variable in both 1996 and 2001. Results are summarized in the Appendix Table A2.1. Marginal effects at average ethnic diversity index are reported in all columns. The positive effect of ethnic diversity on both employment as a whole and wage-employment in particular (excluding self-employed people in columns 4 - 6) is robust to these specifications. The magnitude of the marginal effects is very similar to those in Table 5 and Table 6 in baseline regressions. For example in logit regressions in 2001, the coefficient of ethnic diversity on wage employment is 0.145, which is roughly the same as the corresponding coefficient in OLS regressions in Table 6 (0.144 in column 5). In IV regressions the magnitude in non-linear models is smaller than that in linear IV models but the significance remains the same. For example, in probit regressions with our instrumental variable based on 2001 census data, the coefficient of ethnic diversity on employment is 0.140 while in the corresponding IV regression it is 0.176 (column 5 in Table 6).

In Appendix Table A2.2 we also implement multinomial regressions to take into account the decisions of both self-employment and wage-employment. We construct a variable of employment status which equals 0 if one is unemployed, 1 if one is self-employed and 2 if one is wage employed. Columns in Appendix Table A2.2 captures the marginal effects of ethnic diversity on the decision of self-employment and wage employment, relative to the outcome of unemployment. Columns 1 and 2 report the cross-sectional multinomial logit regression results while columns 3 and 4 report the multinomial probit regressions with our instrumental variable. The results indicate that wage employment rate responds positively to ethnic diversity while there is no robust evidence on self-employment rate. The magnitude of the effect of ethnic diversity in multinomial logit regressions is similar to that in our main analysis while the magnitude in IV regressions is much larger in multinomial models than in linear probability models. For example, the coefficient is 0.135 in 2001 in multinomial logit regressions and 0.125 in OLS regressions with linear probability models. The corresponding coefficient is 0.981 in multinomial probit regressions with instrumental variable and 0.182 in IV regressions with linear probability models.

2.5.5 Decomposing ethnic diversity index

We can decompose ethnic diversity index into two components: the number of groups and the dispersion of group size. Suppose there are m ethnic groups in a district. Group k has a population share s_k over the whole population. Our ethnic diversity index is thus decomposed in the following way:

$$ELF = 1 - \sum_{i=1}^{m} s_k^2 = 1 - \sum_{k=1}^{m}((s_k - \frac{1}{m}) + \frac{1}{m})^2$$

It is obvious to get the following decomposition:

$$ELF = 1 - \underbrace{\frac{1}{m}}_{\text{1/No. of groups}} - \underbrace{\sum_{k=1}^{m}(s_k - \frac{1}{m})^2}_{\text{Dispersion of group size}} \tag{2.6}$$

Leaving aside the constant term (i.e. 1), the first item in the ethnic diversity index is an inverse of number of ethnic groups and the second item captures the dispersion of group size. We therefore replicate the main analysis by using these two items to replace ethnic diversity index for both the whole sample and the subsamples with different population size and see how they are correlated with wage employment rate.

The corresponding results are in columns 1-4 in Table 12.1. We find that in 1996 and 2001, both of the two terms are negatively correlated with wage employment rate, which means that employment rate increases with the number of ethnic groups in a district and decreases if the distribution of group size becomes more uneven (column 1). Again, a detailed investigation of subgroups shows that these two components only affect ethnic groups with relatively large size (column 2 - 4).

As both components are significantly correlated with wage employment rate, we need to disentangle these two factors for a further investigation of the mechanism through which ethnic diversity improves labour market opportunities. Here our instrumental variable not only provides an exogenous variation on ethnic diversity but also helps disentangle these two components by fixing one and exploring the variation in the other. By construction the instrumental variable is calculated based on the distance from each district to all the historical black homelands. That is to say, the number of groups in the predicted ethnic diversity index is fixed, which is the same as the total number of homelands. Therefore the only variation in the instrumental variable comes from the uniformity in the distance to different homelands (which refers to the distribution of population size among all these ethnic groups). By applying this instrumental variable, ethnic diversity has a clear meaning here: a more diverse place implies the distribution of group size is more even, which is independent of the number of ethnic groups.

To verify this argument, we run the IV regressions similar to our main analysis in Table 12.2. In Panel A, we control for the number of groups (i.e. the corresponding variable is an inverse of the number of groups) and use the predicted ethnic diversity index as an instrumental variable for the dispersion of group size for both 1996 and 2001. IV regressions in column 1 - 6 imply that given the number of groups, a decrease in the dispersion of group size (i.e. group size is distributed in a more even way, which is the case in a more diverse place) will significantly increase the wage employment rate. Again, this only works for ethnic groups with relatively large population size. Furthermore, the instrumental variable remains strong in all these columns as the F statistics is close to or larger than 10, especially for large groups.

In Panel B we control for the dispersion of group size and instrument the number of groups (i.e. the corresponding variable is an inverse of the number of groups). In all columns the instrumental variable is weak as F-statistics is around or below 3. This confirms that the instrumental variable does not capture the number of different groups.

Summary of empirical results. The above empirical section consolidates the following results which are the basis for the theoretical model in the next section:

1. Ethnic diversity increases employment among the working-age black population and this mainly takes place in wage-employed jobs.

2. The positive effect of ethnic diversity on employment only works for the ethnic groups with relatively large size.

3. Ethnic diversity affects employment opportunities through the change in the dispersion of population size among different ethnic groups.

2.6 How Does Ethnic Diversity Affect Employment: A Theoretical Model and Mechanism

We propose a theoretical framework consistent with our empirical findings above to explain the positive effects of ethnic diversity on employment and the heterogeneity of the effects across sub-groups. More specifically, we focus on social skill investment which increases with ethnic diversity. The model can be verified by both numerical simulation and empirical evidence from real data.

2.6.1 A theoretical framework

The story is as follows. Assume that inter-ethnic communication requires more skills than intra-ethnic interaction. In a more diverse place, the necessity to communicate with individuals from different ethnic groups motivates people to learn and practise more social skills. The acquisition of this extra skill, which is helpful in reducing coordination costs or increasing labour productivity (which we will discuss later on), can make them more competitive in the labour market and increase their chances of finding jobs.

In more detail, people obtain utility from interacting with those both inside and outside their own ethnic group. Establishing a relationship with someone from a different ethnic group requires more skills than with those from the same ethnic group (this may be due to barriers like language). In a more ethnically diverse place people have to communicate with a larger proportion of individuals outside their own ethnic group to maintain a certain level of social connection. Therefore they put in more efforts in developing social skills, as long as the benefit of interacting with a different ethnic group outweighs the cost of learning efforts. Social skills here can be of many types, including both cognitive skills like language and non-cognitive skills like pro-social traits. When people are in the labour force, these skills are beneficial to their labour market performance, in addition to their human capital investment.

What needs to be emphasised here is that more ethnically diverse places do not necessarily have more overall social interaction but the investment in social skills should be higher because a larger proportion of social interaction comes from inter-group connection which requires more skills than intra-ethnic communication.

The distinction between social connection and investment in social skills is analogue to the literature which differentiates social connectedness and network formation (Chay and Munshi, 2015). Their story implies that there exists a threshold above which social connectedness and

network-based outcomes are positively correlated. Similarly, in our story, the level of social connection can be high in both ethnically homogeneous and diverse places, but investment in social skills is only high when a large proportion of this social connection takes place between ethnic groups as intra-ethnic communication is relatively costless.

Model setup

We provide a model of a coordination game to explain the mechanism. We assume that individuals gain utility from social interaction at the cost of investing in social skills. As the cost of communicating with a different ethnic group is larger than that with the same group, we assume communication within each ethnic group is costless. The cost of investment in social skills for inter-ethnic interaction is c per unit. We also assume that the amount of investment in social skills x_{ik} equals the output of the investment (i.e. the amount of skills acquired) for individual i in ethnic group k. We have the following setup of a coordination game:

Players. Each group only differs in terms of their population size. Suppose there are m ethnic groups in total. We denote these different groups as m different sets $N_1, N_2, \ldots N_m$, each with a group size n_k and $k = 1, 2, \ldots m$. The overall population in each district is N, so that $\sum_{k=1}^{m} n_k = N$. Each ethnic group then has the share s_k and $k = 1, 2, \ldots m$ over the whole population in the district. Here $s_k = \frac{n_k}{N}$.

Strategies. Each individual i in group k invests x_{ik} in social skills. For simplicity we assume x_{ik} is a binary variable which equals 1 (0) if i invests (does not invest).[36] One can only participate in inter-ethnic social interaction if he invests in social skills. The total amount of people each individual i in ethnic group N_k with a group size n_k has access to in the inter-ethnic communication is calculated as $x_{ik} \sum_{j \neq k} \sum_{q \in N_j} x_{jq}$. There is complementarity between i's own investment in social skills and the overall investment level of people outside group k. Therefore the total number of people interacting with i (both inside and outside his own group) can be calculated as $n_k + x_{ik} \sum_{j \neq k} \sum_{q \in N_j} x_{jq}$.

One important feature of the strategy is that by construction we assume skill investment is bilateral. If $x_{ik} = 0$, i cannot benefit from social interaction even if everyone outside his group invests in social skills. As is discussed in the literature review part, papers discussing the social behaviour of ethnic minorities in the American society argues that investment in social skills is unilateral as ethnic minorities try to assimilate to the American society by learning the language spoken by the American majority while the Americans do not put in any efforts in learning additional language. However, in our case, if we assume that skill investment is unilateral such that the "small" ethnic group assimilates to the "large" group by learning their language while the "large" group does not need to make any investment, we should observe the pattern that only the "small" ethnic group responds to ethnic diversity, which directly contradicts our empirical finding (where we find only "large" ethnic group responds to ethnic diversity). Therefore, a more reliable assumption in our setting is that both groups make efforts in learning additional social skills. A reasonable example is that both groups learn a common language. This is also consistent with our proxy of social skill later on, which is the proficiency

[36] One can potentially treat x_{ik} as a continuous variable or make x_{ik} heterogeneous in communicating with different ethnic groups. For example, similar to Akerlof (1997), we can introduce the investment of x_{ikj} if individual i is interacting with group j, and x_{ikj} is a decreasing function of social distance between groups k and j. However, this binary setting of x_{ik} is already enough to explain the key empirical findings about ethnic diversity discussed above.

of English/Afrikaans as the second language. In this case one can communicate with people from another ethnic group only if both learn a second official language.

Utility. Individual i belonging to group k obtains utility from social interaction which depends on the size of his own group n_k and the number of people he can reach in other ethnic groups, the latter relying on both his own investment in social skills and the efforts from other ethnic groups. The utility from overall social interaction is written as $f(n_k + x_{ik} \sum_{j \neq k} \sum_{q \in N_j} x_{jq})$, which is assumed to be increasing at a diminishing rate. That is, $f' > 0$ and $f'' < 0$. The implication is that utility from social interaction increases as more people participate in communication, but this has a diminishing return as people get tired from social life when the number of contacts increases. We can thus write the net utility U_{ik} from overall social interaction for individual i in group k as follows:

$$U_{ik} = f(n_k + x_{ik} \cdot \sum_{j \neq k} \sum_{q \in N_j} x_{jq}) - c x_{ik}$$

We then normalise the amount of social interaction by the overall population size N in the corresponding district. By doing this we control for the whole population size and what matters in social interaction is the share of each group over the whole population rather than the absolute level of group size. There are three reasons to do this normalisation. Firstly, each group's share of population size, instead of the absolute level of group size, is directly linked to our measure of ethnic diversity index. Secondly, controlling for the magnitude of overall population size in each district is consistent with our empirical analysis where we control for the population density in each district and investigate the remaining variation in ethnic diversity. Thirdly, the interpretation of the utility function becomes more intuitive. As the total amount of time for social interaction is limited for each individual, what matters more in social connection is not the total amount of people one has access to, but the probability of establishing connection to a person one randomly meets in the district per unit time. After this normalisation, the utility function becomes:

$$U_{ik} = f(s_k + \frac{x_{ik} \cdot \sum_{j \neq k} \sum_{q \in N_j} x_{jq}}{N}) - c x_{ik} \tag{2.7}$$

Here we also assume that the per unit cost of social interaction is the same in different districts. In principle one can extend the model by allowing the cost c to vary across districts. If this is the case, how the investment in social skills responds to ethnic diversity might be ambiguous. On the one hand, it can increase with the level of ethnic diversity, which will be explained by our mechanism. On the other hand, it may decrease with ethnic diversity as more ethnically diverse districts might have more conflicts, which discourage people from social interaction. However, as we find the positive effect of ethnic diversity on employment in the empirical part, we argue that the positive side of ethnic diversity is more important than its negative side. Furthermore, how conflict responds to ethnic diversity has been discussed in other literature already and is not the central focus of this paper. Therefore, we only focus on explaining the positive effect of ethnic diversity by simplifying other potential factors at the negative side. In our numerical simulation we also set different values for c to see how this affects our results.

Equilibrium. In this paper we focus on pure strategy Nash equilibrium. In this game,

player i from group k chooses either $x_{ik} = 1$ or $x_{ik} = 0$ to maximise his total utility from social interaction. In the pure strategy Nash equilibrium, no one has the incentive to deviate from his current decision.

Clearly the coordination game has multiple equilibria. For example, $x_{ik} = 0, \forall i, k$ is a Nash equilibrium. This is because starting with this initial condition, no one has the incentive to deviate. In more detail, for an individual i in group k, his utility from social interaction is:

$$U_{ik} = \begin{cases} f(s_k) - c, & \text{if } x_{ik} = 1 \\ f(s_k), & \text{if } x_{ik} = 0 \end{cases}$$

Therefore individual i always gets higher utility by not investing in social skills. That is to say, in order for the social interaction to happen, there might be some initial efforts to stimulate communication.

Key features of the Nash equilibrium with the maximal level of skill investment

Each Nash equilibrium is characterised by its own level of skill investment so that it does not make sense to conduct comparative statics across different Nash equilibria in this setting. Moreover, the ultimate Nash equilibrium in a given district with a given distribution of group size purely depends on its initial condition (i.e. how many groups choose $x = 1$ and how many choose $x = 0$ in this district originally). As there is no particular selection criterion of the initial condition in each district, it is reasonable to assume that the initial conditions are assigned randomly to each district. In this case, each district falls in any of its own possible Nash equilibria with equal probability. Therefore, the expected level of skill investment in each district at the equilibrium is determined by the range of its possible Nash equilibria (i.e. the Nash equilibria with the maximal and minimal level of investment in social skills). That is to say, to capture the expected level of skill investment in each district, we can just focus on the range of possible Nash equilibria in each district instead of discussing each Nash equilibrium individually.

As discussed before, the Nash equilibrium with the minimal level of skill investment is the same for all districts (i.e. everyone chooses $x = 0$). In this case, the range of possible Nash equilibria in each district is only determined by the Nash equilibrium with the maximal level of skill investment. Therefore in the following discussion we only focus on the equilibrium where the number of individuals investing in social skills is as large as possible, and see how this equilibrium state changes in response to group size. By doing this, we can indirectly demonstrate how the expected level of skill investment changes with ethnic diversity in each district.

One important feature of this particular equilibrium is that to guarantee the maximum participation in inter-ethnic communication, individuals always choose to invest in social skills unless the net utility from doing so is strictly smaller than that from deviating. In other words, even if the individual is indifferent between investing and not investing, he will always choose to invest in social skills.

In addition, we derive the following two lemmas which capture the key characteristics of the Nash equilibrium in this coordination game with the maximal level of skill investment.

Lemma 2.6.1. *In each district, people from the same ethnic group choose the same amount of investment.*

Proof. Suppose player 1 and player 2 both come from ethnic group k with group size n_k. Without loss of generality we assume $x_{1k} = 1$ and $x_{2k} = 0$. We focus on the pure strategy equilibrium with the maximum number of skill investment. As both 1 and 2 maximise their utility from social interaction, we have:

$$\begin{cases} f(s_k + \frac{\sum_{j\neq k}\sum_{q\in N_j} x_{jq}}{N}) - c \geq f(s_k), & \text{for player 1} \\ f(s_k + \frac{\sum_{j\neq k}\sum_{q\in N_j} x_{jq}}{N}) - c < f(s_k), & \text{for player 2} \end{cases}$$

Clearly these two inequalities contradict each other. Therefore we must have $x_{1k} = x_{2k} = 1$ or $x_{1k} = x_{2k} = 0$. □

Based on this, we have lemma 2.6.2:

Lemma 2.6.2. *In each district, people from different groups choose the same amount of investment as long as the population size of these groups is the same.*

Proof. Suppose player i and player j come from ethnic group k and l, and $s_k = s_l$. Without loss of generality we assume $x_{ik} = 1$ and $x_{jl} = 0$. According to lemma 2.6.1, everyone from group k (l) chooses $x_{ik} = 1$ ($x_{jl} = 0$). As both i and j maximise their utility from social interaction, we have:

$$\begin{cases} f(s_k + s_l \cdot 0 + \frac{\sum_{p\neq k, p\neq l}\sum_{q\in N_p} x_{pq}}{N}) - c \geq f(s_k), & \text{for player } i \\ f(s_l + s_k \cdot 1 + \frac{\sum_{p\neq k, p\neq l}\sum_{q\in N_p} x_{pq}}{N}) - c < f(s_l), & \text{for player } j \end{cases}$$

When $s_k = s_l$, these two inequalities hold altogether if and only if $f(s_k + \frac{\sum_{p\neq k, p\neq l}\sum_{q\in N_p} x_{pq}}{N}) - c > f(s_k + s_k + \frac{\sum_{p\neq k, p\neq l}\sum_{q\in N_p} x_{pq}}{N})$ for each possible x_{pq} in group p. As $f' > 0$, $s_k \geq 0$, $c > 0$, this inequality cannot hold.

Therefore we must have $x_{ik} = x_{jl} = 1$ or $x_{ik} = x_{jl} = 0$. □

2.6.2 Social interaction, skill acquisition and distribution of group size

Analytical predictions: staring from a symmetric case

Combining lemma 2.6.1 and lemma 2.6.2, we can link the size distribution of ethnic groups to social skill investments. To guarantee the maximal level of skill investment in equilibrium, we start with the initial condition where $x_{ik} = 1, \forall i, k$ and study people's incentive to deviate from this condition.

We derive an analytical proposition based on a symmetric case where all groups in a district have the same population size (i.e. group size is distributed evenly). We later on show that it is not feasible to prove the proposition with an arbitrary distribution of group share. So we will provide a numerical simulation based on the generalised density function of group share to verify the proposition.

Consistent with the empirical strategy, we fix the total number of ethnic groups in a district and see how a more even (uneven) distribution of these groups affect social skill acquisition. Suppose the number of groups m is fixed but groups are not distributed evenly. Starting from the point where each group has the same population size and compare it with the case of asymmetric size distribution among all these groups, we have the following proposition:

Proposition 2.6.1. *Suppose the total number of different groups is given. Compared with the symmetric case where each group has the same population size in the district and everyone invests in social skills, social skill investment decreases when the dispersion of group size in a district increases (i.e. the distribution of population size among different groups becomes more uneven).*

Proof. Given the total number of different groups m, the dispersion of group size can be captured by $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$.[37] Starting from the symmetric case where every group has $s = \frac{1}{m}$, when the distribution of group size is more uneven, the gap in the population share among all these groups becomes larger. One implication is that if $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$ becomes larger, either there exists one s_k which is extremely large or there exist several s_k which are larger than the fixed mean value of group share $\frac{1}{m}$ (Otherwise the overall population size is smaller than N). As a result, to show that a higher proportion of people will deviate from the initial condition when the distribution of group size is more dispersed, we just need to show that larger groups are more likely to deviate.

Starting from $x_{ik} = 1, \forall i, k$, the utility of social interaction for individual j in group k is:

$$U_{jk} = \begin{cases} f(s_k + (1 - s_k)) - c, & \text{if } x_{jk} = 1 \\ f(s_k), & \text{if } x_{jk} = 0 \end{cases}$$

Individual j in this group will deviate if:

$$f(1) - c < f(s_k) \Rightarrow s_k > s^* \tag{2.8}$$

Suppose in the symmetric case no one deviates, which means $f(1) - c \geq f(\frac{1}{m})$. When group sizes are more unevenly distributed in a district, the population share of the largest group(s) becomes larger than $\frac{1}{m}$. Suppose in a district, s_k is the largest group in the distribution of group size ($s_k > \frac{1}{m}$), it is straightforward that it is more likely to have $s_k > s^*$ when group sizes are more unevenly distributed in the district. In this case the largest group k will deviate and choose $x_{jk} = 0$. For the remaining groups, suppose group l is the second largest group. Given the largest group deviates from investment, the same logic shows that for group l to deviate as well, we must have:

$$f(1 - s_k) - c < f(s_l) \tag{2.9}$$

Since $f' > 0$, we find that the motivation of deviating increases with group size. In particular, when the dispersion of group size is larger in a district, more groups will have relatively large sizes so that they will deviate from investment in social skills.

□

[37] In principle the dispersion of group size can also be captured by the variance or standard deviation of group share. Here $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2 = m * Var(s_k)$. We do not use the $Var(s_k)$ to measure the dispersion here, because to prove the level of investment changes with number of different groups and the dispersion of group size, we must hold one fixed and get the other to vary. In statistics $Var(s_k)$ decreases intrinsically with m. Therefore, it is very hard to find the same $Var(s_k)$ for different m. It is better to scale $Var(s_k)$ up by a scalar m. In another case, if we want to hold m constant and see the changes in the dispersion, it does not matter whether we use $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$ or $Var(s_k)$ as $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2 = m * Var(s_k)$.

Numerical simulation with a convoluted density function of group size: algorithm and results

In this section we give a more generalised verification of proposition 2.6.1 with numerical simulation by allowing for a convoluted density function of group size in a district. The logic is similar to that behind the equation 2.8.

Suppose there are m groups in total in a district. Without loss of generality we rank them by an ascending order of group size. We have:

$$s_1 \leq s_2 \leq \ldots \leq s_k \leq s_{k+1} \leq s_m$$

From the starting point where everyone invests in social skills, the largest group will deviate if:

$$f(1) - c < f(s_m)$$

After the largest group deviates, the second largest group will deviate if:

$$f(1 - s_m) - c < f(s_{m-1})$$

In general, suppose s_k is the last group which deviates from the initial condition. We have:

$$\begin{cases} f(1 - s_m - s_{m-1} - \ldots - s_{k+1} - s_k) - c = f(s^*) \\ s_{k-1} \leq s^* < s_k \end{cases} \tag{2.10}$$

Given the population share of each group over the total population in the district, the largest s^*, which gaurantees the maximal level of social interaction at the equilibrium, is unique.

The total proportion of people deviating from the initial condition is $Y = \sum_{s_k > s^*}^{m} s_k$. And the overall level of skill investment is $1 - Y = 1 - \sum_{s_k > s^*}^{m} s_k$.

It is not feasible to get an analytical analysis on how Y changes with m or $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$ with an arbitrary distribution of group size. This is mainly because Y depends on both the number of groups which deviate from the initial condition and the population share of these groups. These two are not easily captured simultaneously by a density function of group share. Furthermore, with different parameter values c, how Y reacts to the dispersion of group size is not always unambiguous. For example, in some particular distribution, we may find that Y decreases with the dispersion of group share, which we have encountered in our numerical simulation. However, if the amount of tests in the numerical simulation is large enough, we can get overwhelming results that support our propositions.

Algorithm. We need to numerically show that for a convoluted distribution of group share in a district, the proportion of people who deviate from investing in social skills (i.e. Y) increases with the dispersion of group size $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$ when the number of groups is fixed. Suppose the utility function is $f(x) = \sqrt{x}$. c in principle can take any positive values. We conduct our simulation based on three of them: $c = 0.1$, $c = 0.2$ and $c = 0.5$. For each c, the steps of simulation are as follows.

1. Draw s_k, $k = 1, 2, \ldots, m$ from a convoluted distribution of s, but make sure that $\sum_{k=1}^{m} s_k = 1$ (Appendix A.3 explains how to make the constrained draws in more detail).

2. Choose a particular m as we want to fix the number of groups.

3. Rank each s_k in an ascending order.

4. Suppose the largest group share is s_m. $Y = 0$ if $\sqrt{s_m} \leq \sqrt{1} - c$. Otherwise move to the next step.

5. Suppose the second largest share is s_{m-1}. $Y = s_m$ if $\sqrt{s_{m-1}} \leq \sqrt{1 - s_m} - c$. Otherwise move to the next step.

6. Continue until we find s^* which satisfies Equation 2.10. $Y = \sum_{s_k > s^*}^{m} s_k$.

7. $Y = 1$ if we search till the smallest s_1 but still could not find such s^*.

8. Operate another draw of s_k. Repeat the steps 3-7 to get different Y. In our simulation we conduct 100000 tests.

9. Fix m, we calculate the standard deviation of group share in each test (i.e. $SD(s)$) and finally draw a figure of mean value of Y over $SD(s)$.

The simulation results are in Figure 2.6. Here we fixed the number of groups m and see how the proportion of people who deviate from investment changes with the dispersion of group size in a district. For each value labeled in the x axis, we conduct 100000 tests to get the mean value of Y. For each test, we also do the same simulation for different c. Consistent to proposition 2.6.1, the probability of deviating increases with the dispersion of group size in a district. This is robust to different numbers of groups we set in our simulation. The intuition is that when the distribution of group size becomes more uneven, there is a larger chance that we can have groups with very large sizes and these are the groups which are the most likely to deviate.

One interesting finding is that when c is relatively large and the number of groups is small (which means each group is important), the proportion of people who deviate can decrease with the dispersion of group size (panel a in Figure 2.6). This is reasonable because for example we can have two districts, one having only one very large group and the other having several large groups. The relative magnitude of their overall population share in the corresponding district can be ambiguous. This means, if the conflict level in a district is too high (i.e. cost of investment is too large), a more diverse district (i.e. less dispersion of group size) can potentially have less investment in social skills, which might be harmful to economic outcomes.

Social interaction, skill acquisition and ethnic diversity

We prove from the above proposition that skill investment is higher when the group size is more evenly distributed. And how does these relate to ethnic diversity?

Equation 2.6 indicates that ethnic diversity decreases with the dispersion of group size. Based on proposition 2.6.1, we have the following proposition 2.6.2:

Proposition 2.6.2. *Social skill investment increases with ethnic diversity (which means a more even distribution of group size given the number of groups).*

Following proposition 2.6.2, we also have proposition 2.6.3:

Proposition 2.6.3. *Ethnic groups with relatively small group size are not affected by ethnic diversity.*

This is because starting from $x_{ik} = 1, \forall i, k$, in the Nash equilibrium with the maximal level of social interaction, the small group will not deviate as long as their group size is below a certain level (regardless of the strategies of the large group). In other words, they always choose to participate in inter-ethnic communication and invest in social skills regardless of ethnic diversity levels. Therefore the small groups will in general have more social skill investment than the large groups but their social skill investment is not affected by ethnic diversity of the district. The intuition is that as the small groups get relatively less utility from intra-group communication, they rely more on inter-group connection and therefore are less sensitive to changes in the level of ethnic diversity.

One thing to notice is that in our data, "large", "medium" and "small" groups are defined by the group size in the national population while in the model "small", "medium" and "large" groups are defined at district level. However, definitions at these two levels are compatible in our data. A detailed investigation of the population share in each district in both 1996 and 2001 shows that in general groups with large population size at the national level are also the dominant group in ethnically homogeneous districts, while groups with small population share at the national level also makes up a very small part of the population in those districts. In diverse places the population share of these groups becomes more balanced.

Social skills and labour market outcomes

The social skills acquired through inter-group interaction in a diverse place can improve workers' employment opportunities in several ways.

Less search cost in job hunting. Social skill lowers the cost of searching for potential jobs, therefore increasing labour supply. More social skills help individuals build closer and stronger intra-group contacts. For example, people with higher social skills are better at making use of networks and other methods in gaining job information or asking for referrals. Current literature shows that social network is an important factor in providing more job opportunities for low-educated labour both in South Africa (Magruder, 2010) and in other developing countries (Munshi, 2003).

Increased productivity of certain skills. Recent literature incorporates different tasks in the production function (Acemoglu and Autor, 2011) and highlights the importance of social skills (Deming, 2017). Under the framework that low and high-skilled workers have their own comparative advantages in dealing with different tasks and the range of tasks performed by low-skilled workers is determined by where their comparative advantages are, Deming (2017) explains that social skill increases the productivity of certain tasks by allowing workers with comparative advantages to trade their tasks, which leads to more efficient production. In our story, acquiring additional social skills may also potentially increase the productivity of certain tasks and increase the employment chances for low-skilled workers by allowing them to perform a wider range of tasks.

Overcoming skill deficit. A simple explanation on why social skill stimulates employment is that it works as a substitute for other skills required by employers. In particular, low-educated workers may lack skills necessary for certain occupations, which prevents them from getting the position. For example, if the candidate for the position of a salesman lacks necessary skills of communication, proficiency of additional language may compensate for this communication skills. As the substitutability between social skill and skills acquired through formal education

helps more people qualified for the positions they apply for in a more diverse place, the employment rate will increase accordingly. Skill acquisition from inter-group interaction here functions in a way similar to what is emphasised in related literature that community-based network can work as a substitute for endowments by helping individuals from disadvantaged families get out of low-skill occupational traps (Munshi, 2011).

2.6.3 Ethnic diversity, social skill acquisition and employment: empirical evidence

In this section we provide some evidence to show that social skill acquisition increases with ethnic diversity. There is no straightforward information in census data on social skills. The closest one we can approach is the information on second language at home, including whether or not one speaks a second language and which language they speak. A black person is considered to have some proficiency in a second language if he speaks either one of the nine ethnic languages or a common language (English or Afrikaans). Language is often considered as a cognitive skill which can be learnt from school. In this setting, however, controlling for educational background and investigating into the heterogeneity in the acquisition of language skills among sub-groups, we hope the proficiency of the second language can capture some information on the skills one acquires from inter-group interactions.

More importantly, whether one speaks a second language (and which language he speaks) reflects more of his investment in social skills than the inheritance of language skills from his parents. This results from a series of laws and regulations during the Apartheid regime. Firstly, inter-racial marriage was prohibited during Apartheid starting from 1949 when the Prohibition of Mixed Marriage Act came into effect. The act was repealed in 1985 by the Immorality and Prohibition of Mixed Marriages Amendment Act. In 1996 and 2001 census, parents and spouse of the working-age black people of our interest either lived through Apartheid when marriage between black and white (or black and coloured) was abandoned, or they got married before the independence of South Africa from the British colonisation when there was already informal racial segregation. Thus it is not very likely that the proficiency of English or Afrikaans among the current generation was purely obtained from their parents in the inter-racial marriage. Even among the black population, inter-ethnic marriage is also rare. As discussed at the beginning of the paper, inter-ethnic relationship was deteriorated during Apartheid so that marrying someone from another ethnic group is not a common case. Appendix Table A3 shows that in 1996 census, the contemporary inter-ethnic marriage rate is less than 4%. This phenomenon is even more rare in the parental generation as their inter-ethnic marriage rate is only 1%. Although there is sample selection as only spouse and parents cohabiting with the household head are included in the census, this statistics can still reflect the low inter-ethnic marriage rate.

Furthermore, whether one speaks a second language is not very likely to capture the language proficiency of individuals before they decided to move out of the homelands. As is discussed in the institutional setting, there were almost no indigenous black people in the "white" areas in South Africa and the contemporary population in these districts are mainly the decedents of the migrants from different homelands before the arrival of white colonisers. Therefore it is unlikely that those ancestors learnt English or Afrikaans before migration. Furthermore, in the institutional setting, we have already shown that over 90% of the black population surveyed in 1996 census either never moved or moved within districts up to the time when they were surveyed.

Even among recent migrants in 2001, intra-district migration is much larger than inter-district migration. This further shows that the distribution of ethnic diversity in our census data is largely inherited from the historical pattern rather than driven by contemporaneous migration who potentially acquired language skills before migrating.

To prove the channel in our theoretical model, we first show that ethnic diversity improves social skill acquisition (i.e. measured by second language proficiency) and then we demonstrate that higher social skill is correlated with higher employment rate conditional on ethnic diversity. As the information on second language proficiency is only reported in 1996 census data, we only show the results in 1996 census in this section.

Appendix Table A3 also reveals that the proportion of people who speak a second language is not too small. Among the whole black population, around 22.5% speaks a second language, 8.7% (13.8%) of which speaks a common language (ethnic language). In the regression analysis we focus on the common language (English or Afrikaans) instead of ethnic language as the former one is more related to labour market performance in wage-employment and less likely to reflect family inheritance as the ban on inter-racial marriage was more strict than inter-ethnic marriage within the black.[38]

We introduce a dummy variable on whether one can speak English or Afrikaans as a second language and regress it on ethnic diversity in 1996, conditional on the same set of control variables in the main analysis. Simple OLS regressions may suffer from the same problem as discussed before. For example, there are two potential types of selection of migrants related to their language proficiency. Firstly, migrants with higher ability are able to move out of the homelands and these people might have already mastered a second language prior to migration. Secondly, migrants with better language proficiency choose to move to a more diverse area where there are more job opportunities. If the first type of selection is the case, people with higher ability than their counterparts in the original homelands can potentially move to both ethnically homogenous and diverse places. Thus we should not see any correlation between ethnic diversity and proficiency of second language if language skills are purely captured by the selection of migrants at the time of moving out of homelands. The second selection of migration comes from the fact that migrants with higher ability (including language efficiency) move to more diverse places as migrants who cannot speak a common language may find it difficult to communicate with people outside their ethnic groups. To deal with this selection, we use the same instrumental variable approach as implemented in the main analysis (i.e. using predicted value of ethnic diversity in 1996 as an instrument for real ethnic diversity).

Table 13.1 shows both OLS and IV regression results about how ethnic diversity affects individuals' second language proficiency. Panel A and B investigate the results for the whole black population and the heterogeneity of the effects of ethnic diversity by group size. The coefficients in Panel A in both OLS and IV regressions are significantly positive, indicating that ethnic diversity increases the probability of learning a second language (English or Afrikaans). In Panel B, a comparison between groups with large, medium and small population size indicates that ethnic diversity has a strong and positive effect on language skills only among the ethnic groups with relatively large population size, which is consistent with proposition 2.6.3 in the model. In addition, the instrumental variable remains strong in both whole-sample and sub-

[38]But in regressions the proficiency of both common language and ethnic language can respond to ethnic diversity.

sample regressions.

One concern in interpreting the positive impacts of ethnic diversity on language proficiency as a result of social interaction is that in a more diverse place the importance of English or Afrikaans is more highlighted. For example, firms will have a more favourable environment for employees to learn an official language, as employees have to serve customers or talk to colleagues whose first language is different from these employees' own ethnic language. In this case language proficiency is developed after one has found a job rather than before, and the probability of finding a job is affected by other factors. To show that the improvement in language proficiency is driven by the need for social interaction instead of a skill purely developed in the workplace, we split the sample into subgroups with different ages: young people below the age of 15, working-age people (15-64) and people who are retired ($\geq$65) in Table 13.2. We find the positive and significant effect of ethnic diversity on language proficiency in IV regressions among all the three age groups, which indicates that the language skill can be achieved even before people enter the labour force, and therefore is not purely driven by the requirement from the workforce.

As a further check of our mechanism, we decompose the ethnic diversity index into the inverse of number of groups and the dispersion of group size and conduct the same OLS and IV regressions as described in Table 12.1 and 12.2 by replacing the dependent variables with a dummy on whether one can speak English or Afrikaans as the second language. The results are in columns 5 - 8 in Table 12.1 and 7 - 9 in Table 12.2. All the results imply that language proficiency increases when the distribution of group size becomes more even, and this only works for groups with large population size.

We then look at whether acquisition of social skills improves labour market outcomes by regressing employment probabilities on the proficiency of a second language (English or Afrikaans) conditional on ethnic diversity, which is presented in Table 14. The dependent variable in Panel A is a dummy on whether one is employed or not (including unemployed and inactive) while in Panel B the dependent variable equals 1 if one is an employee and 0 if one is unemployed or inactive. The independent variable in all these OLS regressions is a dummy on whether one can speak English or Afrikaans as a second language. Again we look at the whole sample and the difference among groups with large, medium and small population size. In all regressions learning a second common language is positively and significantly associated with a higher employment rate (both overall employment rate and wage-employment rate).

2.6.4 Summary of the theoretical model and mechanism

In summary, diversity along ethnic lines could provide individuals with social skills, which improves their employability. That is to say, even if ethnic diversity does not necessarily increase the amount of overall social interaction within a district, it may still motivate people in more diverse areas to learn and practise more skills. This is because communication with individuals from different ethnic groups requires more efforts and skills than intra-ethnic interaction. The acquisition of this extra skill, which is helpful in reducing coordination cost or increasing productivity of certain skills, could increase an individual's chances of finding a job.

The key point of the story is that it is not the overall amount of social interaction that drives the whole story, but the composition of the social interaction. That is to say, more diverse districts do not necessarily have larger amount of social interaction but a larger proportion of the interaction comes from inter-ethnic communication, which is more challenging than intra-ethnic

connection and therefore gives people more motivation to invest in social skills.

In our model, without imposing any intrinsic difference in taste, skills or attitudes between different ethnic groups, the tradeoff between the cost of and benefit from developing social skills leads to the conclusion that inter-ethnic social interaction and investment in social skills are the most likely to occur in a place where the distribution of group size is relatively even, which implies a larger ethnic diversity. It is because starting from an initial condition where everyone invests in social skills, less people deviate from this investment decision in the equilibrium state in a more ethnically diverse place. This effect occurs mainly among the ethnic groups with relatively large group size. In the labour market, the acquisition of these extra social skills is helpful in lowering the barrier to formal jobs by reducing coordination and search costs, by increasing productivity of certain skills or by substituting for some necessary skills which are otherwise not available.

2.6.5 Ruling out some alternative explanations

Ethnic diversity might positively affect the labour market outcomes of the blacks through several channels. Here we rule out some alternative explanations through which ethnic diversity improves labour market outcomes based on our data and narratives.

Labour supply: skill complementarity. There might be some skill complementarities among different ethnic groups, as each may have their own comparative advantages in skills. For example, South Sotho are believed to have special skills as shaft-sinkers on the mines (Guy and Thabane, 1988). Therefore, diversity generates creativity and innovation by combining people with different skills. In this case, we can also expect diversity to affect differently individuals with different level of education. A priori, we would expect to find a stronger effect for the higher educated whose activities would benefit more from knowledge-sharing and problem solving.

In Appendix Table A4 we replicate the main results by splitting the sample into people with high and low educational levels. According to the compulsory schooling law in the post-Apartheid South Africa, one has to go to school when reaching the age of seven and stays at school until the age of fifteen or the ninth grade. We therefore use 9 years of schooling as a cutting point between high and low-educated people. "High education" refers to people with more than 9 years of schooling (i.e. high school, college and postgraduate) while "low education" means no education, primary and junior high school education. We present both OLS and IV results in both years.[39] In 1996 the positive and significant effect of ethnic diversity on wage-employment rate only exists among low-educated working-age black population. The magnitude of the coefficients of ethnic diversity index is also larger among the low-educated group. In IV regressions in 2001 the positive effect of ethnic diversity still only holds for low-educated people. However, there is some difference in its magnitude between 1996 and 2001. From 1996 to 2001 the magnitude of the coefficient of ethnic diversity index increases largely from 0.05 to 0.12 for high-educated people while for low-educated people the increase is smaller (from 0.14 to 0.19). A more detailed split of the sample reveals that the increase in the magnitude of the effect of ethnic diversity on wage-employment rate takes place only among college graduates while for high-school graduates the coefficient is insignificant and the magnitude is still around 0.05.

[39]The results are robust to other definitions of "high" and "low" educational categories. For example, we also split the sample into people with more and less than 7 years of schooling, and people whose years of schooling are above and below the mean value in the district where they live.

All this indicates the substitutability rather than the complementarity between education and ethnic diversity.

Furthermore, if ethnic diversity generates skill complementarity, it might also give birth to new occupations as new skills can be learnt from other ethnic groups and this creates opportunities for occupations which rely on otherwise infeasible tasks. Therefore, if ethnic diversity stimulates new ideas and skills, we may observe a larger range of occupations in a more diverse place. We regress the range of occupations in each district[40] on ethnic diversity. The results from 1996 and 2001 census are in Appendix Table A5.1 and Appendix Table A5.2. We do not find any positive relationship between diversity and potential new occupations in either OLS or IV regressions.

Labour supply: social grant. Social grant, such as Old Age pension, potentially dis-incentivise labour force participation in South Africa (Banerjee et al., 2008). At the same time there is a possibility that a more ethnically homogenous place is associated with higher level of public goods provision, which might include social grants. In particular, governments in a more ethnically homogeneous place might be willing to offer more social grants due to the nepotism towards the dominant group in that place or less coordination cost among ethnic groups. If the receipt of social grants dis-incentivise working-age people to enter labour force, this could also explain the association between higher ethnic diversity and higher employment rate. However, this is not the case in our setting for two reasons. Firstly, provision of social grants is mainly designed at the national level, which does not vary across magisterial districts. Secondly, we include province fixed effects to account for potential discrepancy of social grants at province level.

Labour demand: discrimination. Discrimination in the labour market is a potential reason why homogeneous places discourage employment, as employers deliberately prevent the minority groups from gaining job opportunities and therefore the demand for minority labour is declined (Goldberg, 1982). It has been proved that the disutility from discrimination against minority groups in the production network harms the productivity of co-workers (Hjort, 2014; Borjas and Bronars, 1989).

A more diverse place can reduce the discrimination against minority groups by encouraging higher level of tolerance and openness. As the chance of interacting and communicating with other ethnic groups increases in a more ethnically diverse place, discrimination in the labour market becomes less of an issue, either because employers have access to more information about the productivity and behaviours of ethnic minorities, or because they are more open to people from different backgrounds.

If this story is the case, we would expect that ethnic groups with smaller size benefit more from increased ethnic diversity than those with relatively larger size, which contradicts our empirical evidence.

Labour demand: diversity of taste. Another potential driving force of labour demand might be the diversity of taste. As people from different ethnic groups have diversified tastes for consumption goods, the variety of consumption increases when a place becomes more ethnically diversified. This induces the diversity of production as well, resulting in higher variety of labour inputs in the production process. When different labour inputs are complementary in

[40] We measure the range of tasks by counting the total number of different occupations observed in each district. Occupations are counted in 3-digit code level.

the production function, this love for variety of labour increases the total demand for labour, therefore improving workers' chance in the labour market. However, if this is the case, we should see the positive effect of ethnic diversity among both large and small ethnic groups, which also contradicts the empirical findings. There is also related literature about how greater diversification of sectoral demands reduces unemployment (Neumann and Topel, 1991). However this works under the condition that workers are mobile enough, which is not likely to be a prevalent case in South Africa where many black people locate far away from economic centres and the transportation cost is very high to them.

2.7 Conclusion and Discussion

This paper provides empirical support for the positive role played by within-black ethnic diversity and blacks' labour market outcomes in post-Apartheid South Africa based on an instrumental variable approach. We also propose a theoretical model to explain how the need for inter-ethnic social interaction stimulates investment in social skills in more ethnically diverse places, making black workers better equipped for the labour market.

The finding reveals that ethnic identity, together with inter-ethnic relationship, is still a distinctive feature shaping people's social life and labour market in modern South African society. The distinction between ethnic groups does not fade away after years of integration, resulting from the Apartheid regime which reinforced ethnic identity. In addition, although the climate of hatred and mistrust generated by the Apartheid system had substantial repercussions on the social fabric, inter-ethnic connections still occur within the black population.

Our result is different from, yet can be reconciled with, the association between ethnic diversity and inter-ethnic cleavages or the erosion of social cohesion. Firstly, most of those literature highlights the underprovision of public goods and social capital in ethnically fragmented communities in developing countries (Alesina et al., 2016a), or the conflict between different ethnic groups (Amodio and Chiovelli, forthcoming). Our story takes a different angle by focusing on skill investment motivated by social interaction. This can just be another side of inter-personal relations which can co-exist with conflicts or coordination problems. Secondly, in our numerical simulation, we show that the level of investment in social skills can decrease with ethnic diversity (i.e. increase with the dispersion of group size) when per unit cost of investment c is large enough. This means, if the conflict level in a district is too high (i.e. cost of investment is too large), ethnic diversity can potentially decrease investment and the consequential economic growth. Thirdly, we have shown in our model that the initial condition in skill investment is important in shaping the ultimate equilibrium. If the society starts from the situation where no one actively participates in inter-ethnic communication, benefits from inter-ethnic connection will stay at the low level forever. Therefore, societies where ethnic diversity is negatively associated with socio-economic indicators might have worse initial conditions in inter-ethnic interaction.

We also find the heterogeneous effects of ethnic diversity on labour market outcomes for different sub-groups. In particular, labour market outcomes of the ethnic groups with larger size are more responsive to ethnic diversity. This indicates that our story is not likely to be the case where the minority assimilates to the majority by integrating into their culture and language, nor is it the story that diversity alleviates discrimination against minority groups (in both cases only the small group will respond to diversity level). Rather, in our story groups with both large

and small sizes participate in social interaction and invest in social skills in response to ethnic diversity.

More importantly, what drives our story is not the number of groups but the distribution of group size. Different from many other papers, we do not impose any intrinsic difference in taste or preference among different ethnic groups. We show that people respond differently in places with low and high levels of ethnic diversity not because ethnically diverse districts bring about more groups which contribute to something unique in these diverse places, but because the relative size of their group results in different motivations to invest in social skills.

Could any interventions be designed to increase employment opportunities for the black South African? As is presented in the theoretical framework, a successful intervention must encourage more inter-ethnic connection which can motivate people to invest in more social skills. It can be an efficient policy as we show that the initial investment in social skills is important to the ultimate equilibrium. Therefore, an attempt at fostering inter-ethnic communication in a more diverse society will have long-lasting effects on overall skill investments. Policies which directly improve black people's social skills may also be effective in preparing them for better employment opportunities.

These interventions to improve people's labour market performance have far-reaching implications not only in different aspects of South African society but also in dealing with ethnic issues all over the world. On the one hand, reducing unemployment can have other important consequences on South African society. For example, it has been estimated that in contemporary South Africa a 10 percentage point reduction in unemployment lowers the Gini coefficient by 3 percent (Anand et al., 2016).

On the other hand, this paper can also shed light on dealing with inter-ethnic relations in other African countries or even developed countries. In recent decades, Western societies have also become considerably more ethnically diverse due to the net immigration flows and the growing presence of ethnic communities (Putnam, 2007), which gives rise to more social problems. For example, there is some negative evidence of ethnic diversity on the support for redistribution which in particular harms low-income earners (Dahlberg et al., 2012). Furthermore, current immigration policies in the US and the European refugee crisis also require urgent modification in policy interventions to improve inter-ethnic relationships and explore the positive impact of ethnic diversity on economic outcomes, to which our mechanism about inter-ethnic interactions can be generalised. Our identification strategy can also be generalised to studies on other types of diversity or migration. For example, replacing homelands with individuals' countries of origin, one can instrument the ethnic composition of immigrants in Europe or the U.S. with a measure of equidistance to multiple home countries (Alesina et al. (2015) implements an approach similar to this).

From 700s A.D.	1652	Before mid-1800s	1867	1910
Bantu migration from central and eastern Africa	First permanent European settlement	Emigration from homelands to "white areas"	First discovery of mines	Establishing the Union of SA

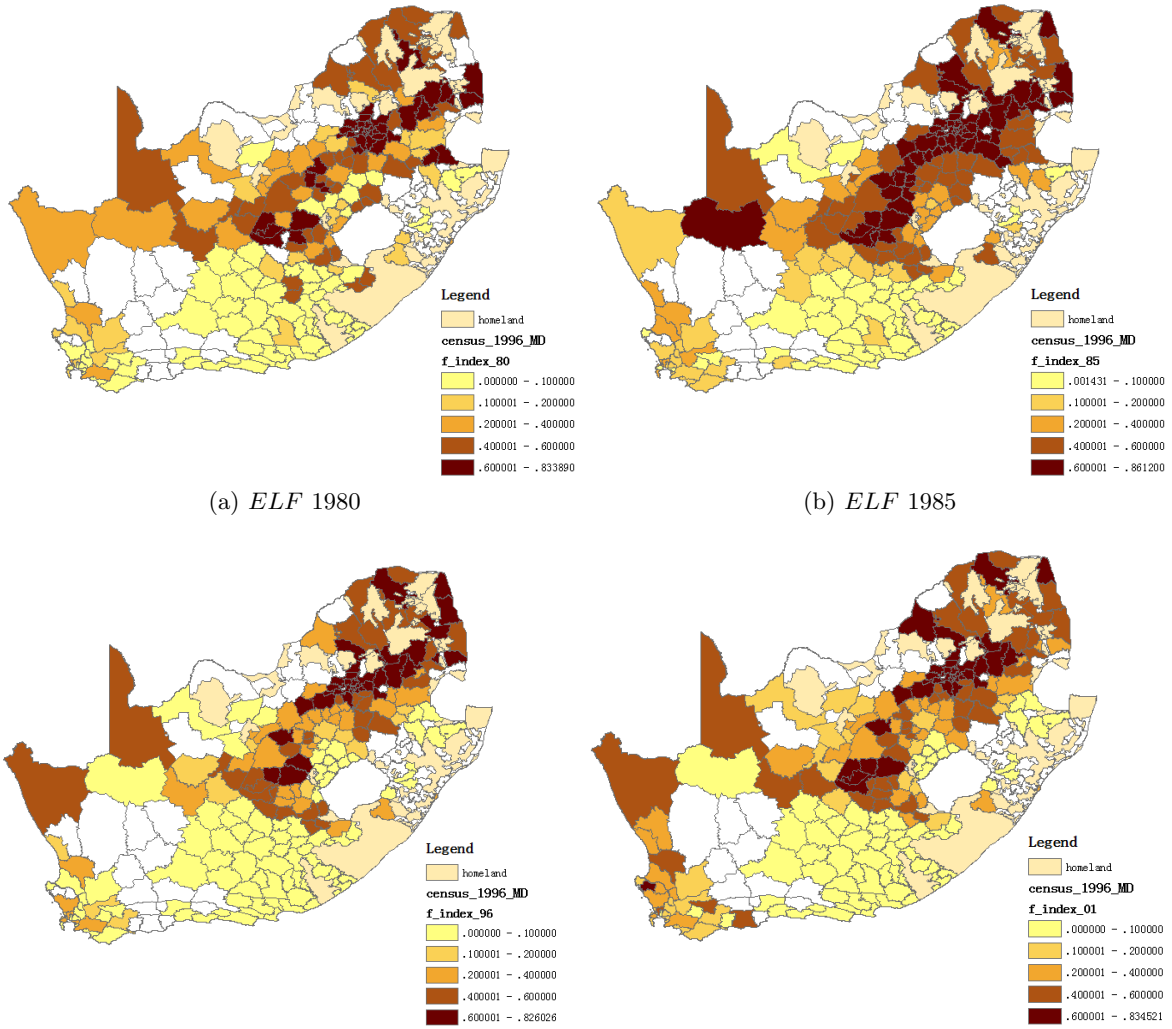
(a) Timeline of Bantu migration and early development in South Africa

1948	1951	1953	1976	1986	1994
Independence	Bantu Authorities Act	Bantu Self-Govern Act	Transkei Independence	End of Pass Law	End of Apartheid

(b) Timeline of modern South Africa starting from Apartheid

Notes: The figures presents the timeline of important nodes in South African history: Bantu migration from central and eastern Africa, emigration of ethnic groups from original homelands, the White colonisation, the discovery of mines and Apartheid regime. Sources of narratives: Mwakikagile (2010) and Gradin (2014).

Figure 2.1: Timeline of Bantu migration, historical development and Apartheid regime in South Africa

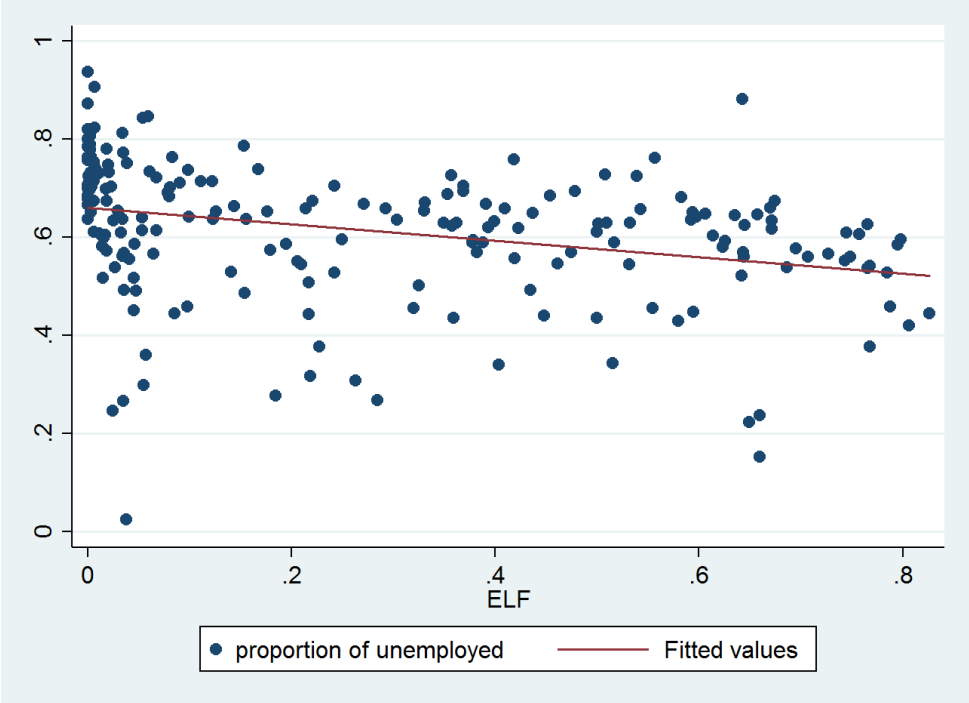


(a) *ELF* 1980 (b) *ELF* 1985

(c) *ELF* 1996 (d) *ELF* 2001

Notes: The figure presents the geographical pattern of ethnic diversity across South African districts in 1980, 1985, 1996 and 2001 for the "white areas". Within-black ethnic diversity is measured with Fractionalisation Index analogue to Herfindahl Index. The results are calculated by the authors based on the corresponding census data.

Figure 2.2: Distribution of ethnic fractionalisation index: 1980. 1985, 1996, 2001

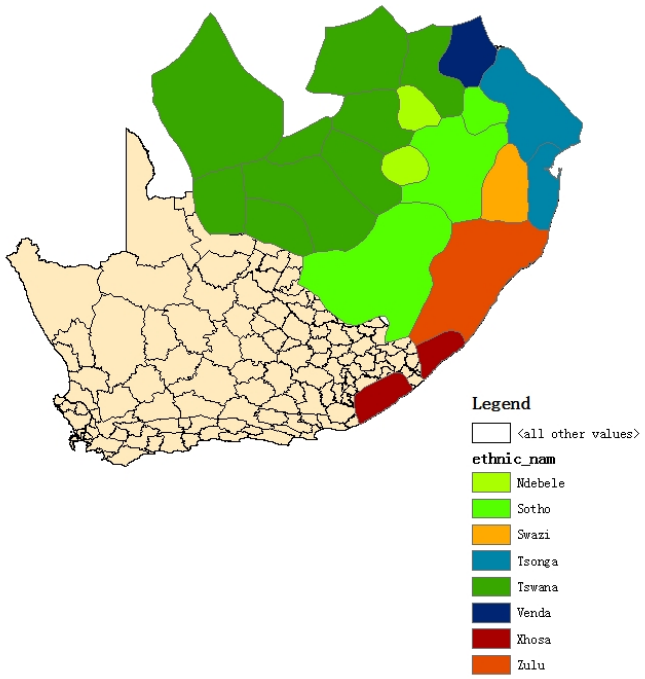


(a) Unemployment and ethnic diversity (ELF) 1996

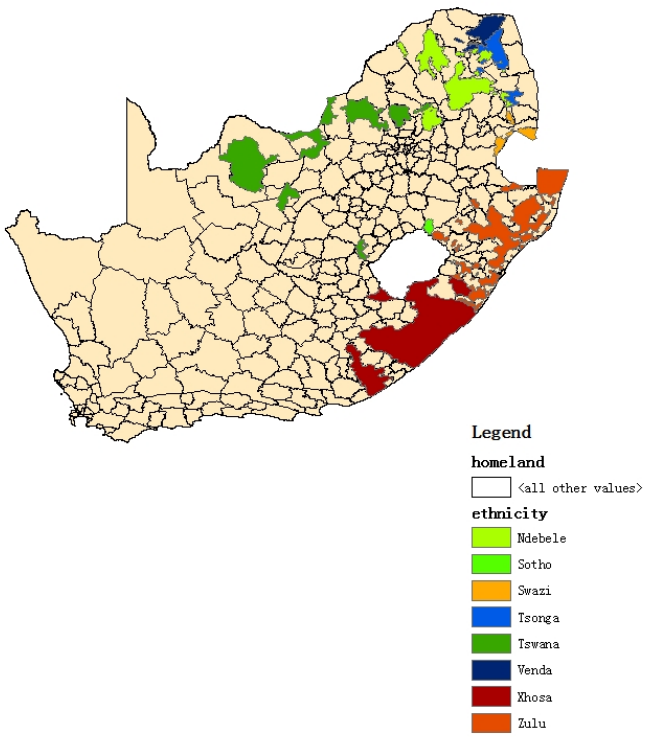
(b) Unemployment and ethnic diversity (ELF) 2001

Notes: The figures present the results on the correlation between ethnic diversity and unemployment rate. Both are measured at the magisterial district level (therefore unemployment rate is calculated as the proportion of unemployed people over the whole working-age black population in a district). The results are calculated by the authors based on 1996 and 2001 census data.

Figure 2.3: The relationship between ethnic diversity and unemployment in South Africa in 1996 and 2001



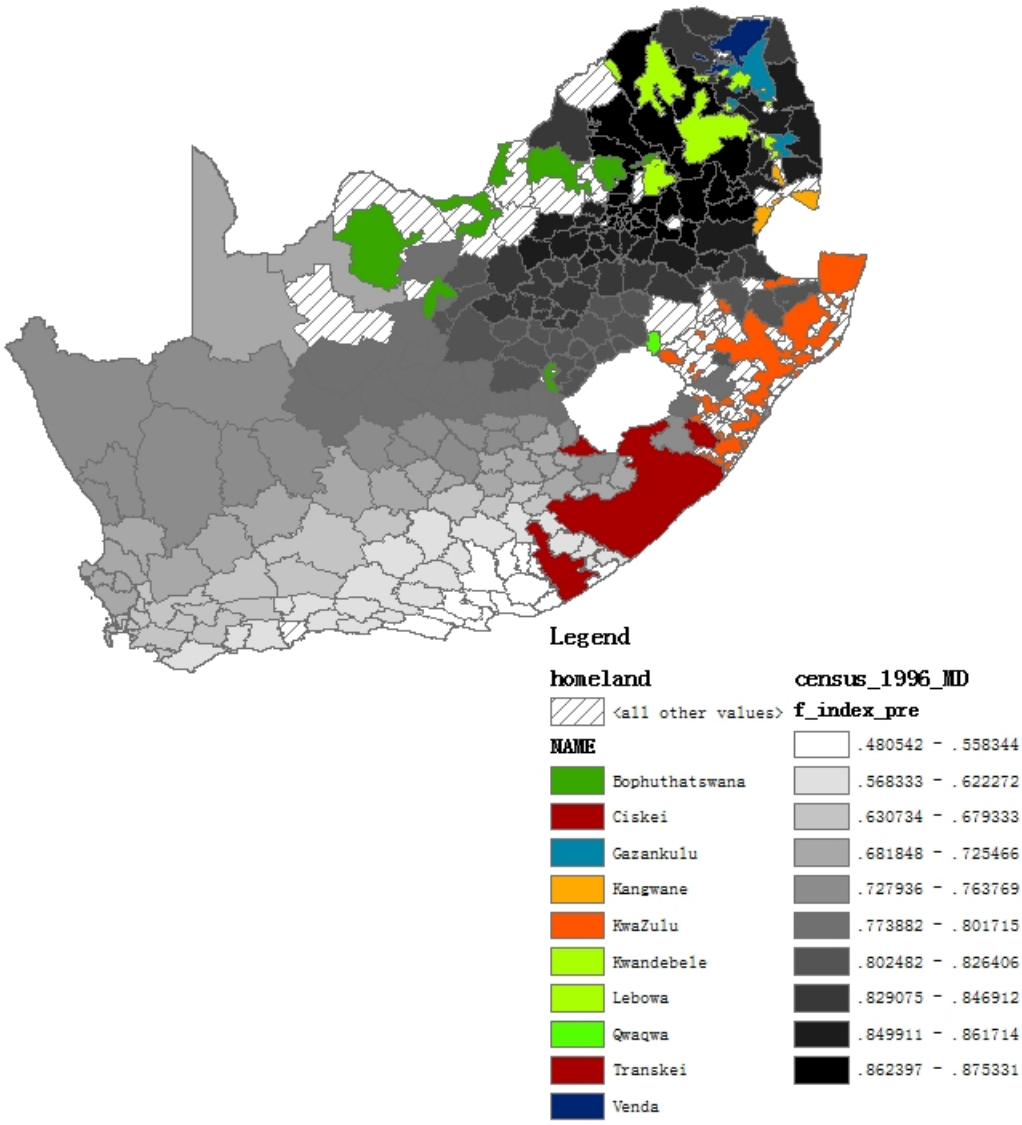
(a) Murdock's map



(b) Bantustan

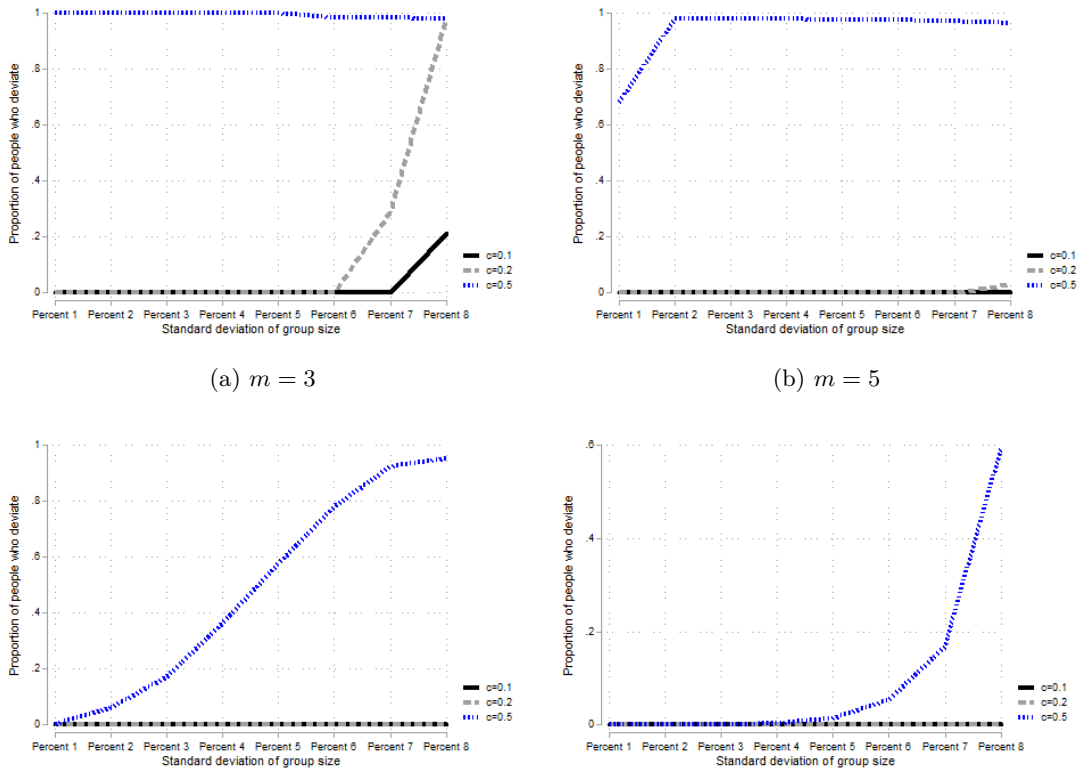
Notes: The figures compare the distribution of ethnic groups in South Africa in Murdock map and the location of Bantustans as proxies for ethnic homelands. Murdock map comes from George Murdock's 1959 work which illustrates the dominant ethnic group in each geographical unit, which is highly consistent with the Bantustans for these ethnic groups assigned by the Apartheid government. This confirms that the location of these Bantustans can well reflect the spatial distribution of original homelands for those ethnic groups.

Figure 2.4: Comparison between the historical settlements of the black ethnic groups and Bantustans



Notes: The figure shows the spatial distribution of our instrumental variable for ethnic diversity - the predicted ethnic fractionalisation index. Following the idea that districts more (less) equidistant to multiple homelands are more (less) diverse, we first calculate the stock of each ethnic group in each district based on the distance between the district and the corresponding homeland with a gravity model. The instrumental variable is a predicted fractionalisation index calculated based on the predicted stock of black migrants.

Figure 2.5: Distribution of predicted ethnic fractionalisation index



(a) $m = 3$

(b) $m = 5$

(c) $m = 7$

(d) $m = 9$

Notes: The figures show the results on the numerical simulation of the proportion of people who deviate from investing in social skills in response to the dispersion of group size in the district. Each group has population share s_i and there are m groups in total. The dispersion of group size is measured by $\sum_{i=1}^{m}(s_i - \frac{1}{m})^2$. In each graph we hold the number of groups m constant. We also consider different per unit cost of investment c.

Figure 2.6: Numerical simulation results on how the level of investments in social skills responds to dispersion of group size

	Population size	Share of the black population	Self employed	Wage employee	Unemployed	Inactive	Unemployed +inactive	ELF
Xhosa	2229452	0.252	0.027	0.312	0.273	0.388	0.661	0.230
			[0.162]	[0.463]	[0.445]	[0.487]	[0.474]	[0.301]
Zulu	2073036	0.234	0.037	0.349	0.264	0.350	0.614	0.521
			[0.189]	[0.477]	[0.441]	[0.477]	[0.487]	[0.273]
South Sotho	2009582	0.227	0.026	0.349	0.242	0.383	0.625	0.467
			[0.159]	[0.477]	[0.428]	[0.486]	[0.484]	[0.23]
Tswana	1039138	0.117	0.026	0.384	0.224	0.367	0.590	0.525
			[0.158]	[0.486]	[0.417]	[0.482]	[0.492]	[0.243]
North Sotho	770110.7	0.087	0.038	0.412	0.236	0.314	0.549	0.700
			[0.192]	[0.492]	[0.424]	[0.464]	[0.498]	[0.116]
Tsonga	295688.6	0.033	0.075	0.434	0.247	0.244	0.491	0.715
			[0.263]	[0.496]	[0.431]	[0.430]	[0.500]	[0.109]
Ndebele	185065.8	0.021	0.040	0.358	0.227	0.375	0.602	0.689
			[0.195]	[0.480]	[0.419]	[0.484]	[0.490]	[0.111]
Swazi	181467.1	0.020	0.041	0.405	0.224	0.330	0.554	0.604
			[0.198]	[0.491]	[0.417]	[0.470]	[0.497]	[0.166]
Venda	80189.34	0.009	0.048	0.497	0.202	0.252	0.455	0.714
			[0.214]	[0.500]	[0.402]	[0.434]	[0.498]	[0.113]
Overall	8863729.5	1.000	0.032	0.355	0.251	0.361	0.613	0.274
			[0.177]	[0.479]	[0.434]	[0.48]	[0.487]	[0.266]

Note: The number and proportion of each ethnic group in the whole black population are calculated in the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. Population size is not always an integer because it is weighted by each person's weight in the census data. Employment outcomes are calculated from individual-level 1996 census data among the working-age blacks. "Self-employed" refers to the proportion of self-employed people in each ethnic group over the whole working-age population of the corresponding ethnic group. Other labour market outcomes are calculated in similar ways. The mean degree of ethnic diversity index is calculated at the district level. All other statistics are calculated at the individual level.

Table 1.1: Summary statistics of demographics and employment among black ethnic groups in 1996

	Population size	Share of the black population	Self employed	Wage employee	Unemployed+inactive	ELF
Xhosa	3105625	0.249	0.017	0.299	0.684	0.251
			[0.130]	[0.458]	[0.465]	[0.298]
Zulu	2798132	0.224	0.025	0.331	0.643	0.558
			[0.156]	[0.471]	[0.479]	[0.264]
South Sotho	2531013	0.203	0.020	0.324	0.657	0.500
			[0.139]	[0.468]	[0.475]	[0.221]
Tswana	1373413	0.110	0.018	0.373	0.610	0.578
			[0.132]	[0.484]	[0.488]	[0.225]
North Sotho	1341608	0.107	0.027	0.396	0.577	0.689
			[0.163]	[0.490]	[0.494]	[0.148]
Tsonga	552403.3	0.044	0.048	0.421	0.531	0.714
			[0.214]	[0.494]	[0.50]	[0.128]
Ndebele	292188.3	0.023	0.029	0.370	0.601	0.708
			[0.168]	[0.483]	[0.490]	[0.131]
Swazi	324071.7	0.026	0.028	0.376	0.597	0.579
			[0.164]	[0.484]	[0.491]	[0.189]
Venda	172927.4	0.014	0.034	0.457	0.509	0.724
			[0.183]	[0.498]	[0.500]	[0.119]
Overall	12491382	1.000	0.023	0.341	0.636	0.302
			[0.149]	[0.474]	[0.481]	[0.259]

Note: The number and proportion of each ethnic group in the whole black population are calculated in the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. Population size is not always an integer because it is weighted by each person's weight in the census data. Employment outcomes are calculated from individual-level 2001 census data among the working-age blacks."Self-employed" refers to the proportion of self-employed people in each ethnic group over the whole working-age population of the corresponding ethnic group. Other labour market outcomes are calculated in similar ways. The 2001 census data does not distinguish unemployed and economically inactive people. The mean degree of ethnic diversity index is calculated at the district level. All other statistics are calculated at the individual level.

Table 1.2: Summary statistics of demographics and employment among black ethnic groups in 2001

	High ELF			Low ELF			
	Mean	S.d	Obs	Mean	S.d.	Obs	ttest
ELF	0.507	0.134	102	0.044	0.071	103	***
Self employment	0.028	0.044	102	0.021	0.044	103	***
Wage employee	0.400	0.110	102	0.320	0.120	103	***
Unemployed	0.570	0.110	102	0.658	0.118	103	***
Agriculture	0.466	0.138	102	0.454	0.134	103	
Manufacture	0.115	0.105	102	0.090	0.105	103	*
Service	0.419	0.114	102	0.455	0.130	103	**
Manager	0.014	0.032	102	0.012	0.045	103	
Profession	0.070	0.063	102	0.082	0.081	103	*
Clerk	0.032	0.056	102	0.020	0.049	103	***
Serve	0.073	0.059	102	0.063	0.069	103	*
Craft	0.107	0.092	102	0.125	0.108	103	
Skilld agriculture	0.121	0.079	102	0.107	0.079	103	*
Operator	0.088	0.071	102	0.062	0.063	103	***
Unskill	0.495	0.118	102	0.529	0.111	103	**

Note: This table compares labour market outcomes in districts with relatively high (i.e. above the median value) and low levels of ethnic diversity. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. Employment outcomes are calculated from individual-level 1996 census data among the working-age population. "Self-employed" refers to the proportion of self-employed people in each ethnic group over the whole working-age population of the corresponding ethnic group."Wage employee" and "unemployed" are calculated in similar ways. We only focus on people who are employed when comparing the allocation of workers across industrial sectors and occupations.

Table 2.1: Summary statistics of ethnic fragmentation and labour market outcomes in 1996

	High ELF			Low ELF			
	Mean	S.d	Obs	Mean	S.d.	Obs	ttest
ELF	0.527	0.126	105	0.077	0.084	105	***
Self employment	0.022	0.055	105	0.019	0.045	105	
Wage employee	0.396	0.114	105	0.315	0.118	105	***
Unemployed	0.582	0.118	105	0.667	0.118	105	***
Agriculture	0.338	0.155	105	0.376	0.152	105	
Manufacture	0.183	0.130	105	0.096	0.089	105	***
Service	0.478	0.138	105	0.527	0.152	105	*
Manager	0.017	0.051	105	0.017	0.068	105	
Profession	0.082	0.075	105	0.080	0.076	105	
Clerk	0.056	0.055	105	0.054	0.084	105	
Serve	0.081	0.071	105	0.076	0.069	105	
Craft	0.059	0.071	105	0.084	0.093	105	***
Skilld agriculture	0.117	0.084	105	0.074	0.071	105	***
Operator	0.108	0.075	105	0.088	0.071	105	***
Unskill	0.480	0.120	105	0.527	0.121	105	**

Note: This table compares labour market outcomes in districts with relatively high (i.e. above the median value) and low levels of ethnic diversity. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. Employment outcomes are calculated from individual-level 2001 census data among the working-age black population. "Self-employed" refers to the proportion of self-employed people in each ethnic group over the whole working-age population of the corresponding ethnic group."Wage employee" and "unemployed" are calculated in similar ways. We only focus on people who are employed when comparing the allocation of workers across industrial sectors and occupations.

Table 2.2: Summary statistics of ethnic fragmentation and labour market outcomes in 2001

	[1]	[2]
Dependent variable	1996	2001
Panel A: Job opportunities		
Distance to the closest economic centre	-274959.5	-245255.3
	(301774.502)	(260543.026)
Panel B: Economic activities of the white		
Share of white who are self employed contemporarily	0.230*	0.031
	(0.139)	(0.136)
Share of white who are wage employed contemporarily	0.095	0.185
	(0.170)	(0.158)
Proportion of white	0.335	0.149
	(0.221)	(0.140)
Panel C: Path dependence		
Share of white who are wage employed in 1980	-0.227	-0.245
	(0.217)	(0.219)
Proportion of white in 1980	-0.118	-0.707***
	(0.255)	(0.238)
Panel D: Contemporary migration		
Number of migrants	-44201.67	9845.154
	(37163.296)	(24771.553)
District controls	YES	YES
Individual controls (district average)	YES	YES
Province fixed effect	YES	YES
Obs	205	210

Note: This table conducts validity test of the instrumental variable based on 1996 and 2001 census data. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. All regressions are at the district level. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and province fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Validity of the instrumental variable

	[1]	[2]	[3]	[4]
	1996		2001	
	Age 15-64	Age 25-64	Age 15-64	Age 25-64
Predicted ELF	1.515***	1.488***	1.653***	1.623***
	(0.320)	(0.326)	(0.292)	(0.293)
Edu	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.000	0.000	0.001	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)
Age	-0.000	-0.000*	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Married	0.003*	0.002	0.004***	0.003**
	(0.002)	(0.001)	(0.001)	(0.001)
Father alive	0.000	0.001	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Pop density	0.000***	0.000***	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Urban	0.012	0.012	0.003	0.002
	(0.011)	(0.011)	(0.010)	(0.010)
River	0.084***	0.080***	0.062**	0.060**
	(0.029)	(0.029)	(0.027)	(0.028)
Density mine	0.434	0.337	0.781	0.697
	(0.709)	(0.698)	(0.717)	(0.695)
Prop black	-0.290***	-0.289***	-0.421***	-0.421***
	(0.056)	(0.056)	(0.077)	(0.077)
Distance closest	-0.000**	-0.000**	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Ruggedness	0.005	0.006	-0.005	-0.004
	(0.008)	(0.008)	(0.007)	(0.007)
Soil quality	0.051*	0.052*	0.054**	0.053*
	(0.029)	(0.029)	(0.027)	(0.027)
Per capita light	0.340	0.353	0.583	0.562
	(0.232)	(0.231)	(0.369)	(0.361)
Road	0.009	0.013	0.011	0.015
	(0.030)	(0.030)	(0.029)	(0.028)
Conflict	0.018*	0.017*	-0.004**	-0.004**
	(0.009)	(0.009)	(0.002)	(0.002)
Proportion manu	0.305**	0.313***	0.261***	0.272***
	(0.118)	(0.117)	(0.076)	(0.074)
Proportion service	0.377***	0.373***	0.198**	0.191**
	(0.129)	(0.129)	(0.086)	(0.084)
Ethnicity fixed effect	YES	YES	YES	YES
Province fixed effect	YES	YES	YES	YES
F-statistics of the instrument	22.36	20.87	32.04	30.60
R-squared	0.875	0.879	0.887	0.890
Observations	464,130	318,610	697,369	484,639

Note: This table reports the first-stage results of the instrumental variable based on 1996 and 2001 census data and report the F-statistics of the instrumental variable. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. All regressions are at the individual level. We report all the control variables, both district-level variables especially geographical features and individual-level controls for socio-economic status. We control for ethnicity and province fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: First-stage regression results: individual level regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	Unemployed		Inactive		Unemployed + inactive		Unemployed + inactive	
	1996		1996		1996		2001	
	Age 15-64	Age 25-64	Age 15-64	Age 25-64	Age 15-64	Age 25-64	Age 15-64	Age 25-64
Panel A: OLS estimates								
Ethnic fractionalisation ELF	-0.024	-0.031	-0.057**	-0.051**	-0.081**	-0.082**	-0.146***	-0.148***
	(0.018)	(0.027)	(0.023)	(0.022)	(0.032)	(0.039)	(0.037)	(0.042)
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.033	0.075	0.153	0.094	0.195	0.123	0.175	0.107
Observations	464,130	318,610	464,130	318,610	464,130	318,610	697,368	484,639
Panel B: IV estimates								
Ethnic fractionalisation ELF	-0.142***	-0.126**	0.043	-0.036	-0.098	-0.163*	-0.170*	-0.206**
	(0.046)	(0.064)	(0.061)	(0.061)	(0.076)	(0.087)	(0.088)	(0.098)
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
F statistics of the instrument	22.36	20.87	22.36	20.87	22.36	20.87	32.04	30.60
R-squared	0.032	0.074	0.153	0.094	0.195	0.123	0.175	0.106
Observations	464,130	318,610	464,130	318,610	464,130	318,610	697,369	484,639

Note: This table reports results about the effect of ethnic diversity on unemployment rate at individual-level regressions based on 1996 and 2001 census data. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. We separate unemployed and economically inactive groups only for 1996 results as these two categories are combined in 2001 census. "Unemployed + inactive" is a dummy variable which equals 1 if one is unemployed or inactive and 0 if one is employed. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Ethnic diversity, unemployment and labour force participation: individual level regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	Wage employment		Self/wage		Wage employment		Self/wage	
	1996		1996		2001		2001	
	Age 15-64	Age 25-64	Age 15-64	Age 25-64	Age 15-64	Age 25-64	Age 15-64	Age 25-64
Panel A: OLS estimates								
Ethnic fractionalisation ELF	0.086***	0.087**	-0.024	-0.020	0.144***	0.147***	0.013	0.012
	(0.033)	(0.041)	(0.017)	(0.017)	(0.037)	(0.043)	(0.013)	(0.013)
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.194	0.126	0.011	0.01	0.173	0.106	0.008	0.008
Observations	449,200	305,099	180,535	162,333	681,529	470,552	253,809	228,519
Panel B: IV estimates								
Ethnic fractionalisation ELF	0.112	0.173*	-0.055	-0.030	0.176**	0.215**	-0.043	-0.041
	(0.077)	(0.091)	(0.043)	(0.040)	(0.087)	(0.098)	(0.037)	(0.034)
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
F statistics of the instrument	22.46	21	20.24	20.31	32.33	30.88	27.93	28.61
R-squared	0.194	0.126	0.011	0.01	0.173	0.106	0.008	0.007
Observations	449,200	305,099	180,535	162,333	681,529	470,552	253,809	228,519

Note: This table reports results about the effect of ethnic diversity on employment and the allocation between self- and wage-employment at individual-level regressions based on 1996 and 2001 census data. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. In column 1, 2, 5, 6 we drop self-employed people as they are a very small proportion of the whole working-age population. Column 3, 4, 7, 8 are based only on the employed black people. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Ethnic diversity and employment status: individual level regresions

	[1] OLS	[2] OLS	[3] OLS	[4] IV	[5] IV	[6] IV
	Log monthly income	Log hourly wage	Hour	Log monthly income	Log hourly wage	Hour
Panel A: Individual level, census data						
Ethnic fractionalisation ELF	0.326***	0.362***	-1.279	0.497***	0.414	2.482
	(0.071)	(0.090)	(1.286)	(0.190)	(0.268)	(3.809)
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
F statistics of the instrument				28.28	28.28	28.12
R-squared	0.345	0.314	0.053	0.345	0.314	0.052
Observations	228,256	228,256	232,533	228,256	228,256	232,533
Panel B: Individual level, LFS data						
Ethnic fractionalisation ELF	-0.0772	-0.0286	0.107	0.439	0.0361	23.23
	(0.253)	(0.246)	(2.943)	(0.996)	(0.870)	(16.51)
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
F statistics of the instrument				5.497	5.322	5.591
R-squared	0.482	0.474	0.054	0.480	0.474	0.018
Observations	3,478	3,615	3,660	3,478	3,615	3,660

Note: This table reports results about the effect of ethnic diversity on other labour market outcomes at individual-level regressions, including working hour, hourly wage and monthly earnings. We only report the result in 2001 as there is no information on hours of working in 1996 census. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We control for province fixed effects. Ethnic diversity is measured with fractionalisation index. All the columns only focus on employees. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Ethnic diversity, intensive margin and wage: individual level regressions

Dependent variable	[1] unemploy + inactive	[2] wage employ	[3] self/wage	[4] log monthly income
Ethnic fractionalisation ELF	-0.291***	0.341***	-0.133*	-0.382
	(0.0709)	(0.072)	(0.0696)	(0.365)
Individual controls (district average)	YES	YES	YES	YES
District controls	YES	YES	YES	YES
R-squared	0.493	0.488	0.244	0.730
Observations	410	410	410	410

Note: This table reports results about the effect of ethnic diversity on employment based on the district-level balanced panel. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables which vary over time and individual-level controls aggregated at district level and province fixed effects. Ethnic diversity is measured with fractionalisation index. The dependent variable in column 1 is the proportion of unemployed over the whole working-age black population. Column 2 is defined in a similar way but we exclude those who are self-employed. Column 3 has the dependent variable which is the ratio of the number of self-employed to that of employees at district level. Column 4 only focuses on back people who are employed. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Ethnic diversity and employment: district fixed effects models

Table 9: Heterogeneous effects of ethnic diversity on wage-employment: individual level regressions

	[1] OLS	[2] IV	[3] OLS	[4] IV
Dependent variable	1996	1996	2001	2001
Panel A: By ethnicity				
Large	0.111***	0.125	0.148***	0.182**
	(0.035)	(0.081)	(0.032)	(0.086)
Obs	320,901	320,901	459,108	459,108
Medium	0.018	-0.150	0.203***	-0.029
	(0.075)	(0.210)	(0.061)	(0.207)
Obs	91,373	91,373	149,632	149,632
Small	-0.016	-0.606*	0.035	0.210
	(0.104)	(0.326)	(0.083)	(0.415)
Obs	36,926	36,926	72,789	72,789
Panel B: By industrial sector				
Agriculture	0.068**	0.034	0.016	0.067
	(0.033)	(0.070)	(0.031)	(0.073)
Obs	165,605	165,605	180,227	180,227
Manufacturing	-0.024***	-0.008	-0.008	-0.012
	(0.009)	(0.019)	(0.010)	(0.020)
Obs	165,605	165,605	180,227	180,227
Service	-0.044	-0.026	-0.008	-0.055
	(0.027)	(0.061)	(0.028)	(0.073)
Obs	165,605	165,605	180,227	180,227
Panel C: By occupation				
Manager	0.004	0.016	0.006*	0.023***
	(0.004)	(0.010)	(0.003)	(0.008)
Obs	153,294	153,294	224,942	224,942
Profession	-0.016	0.107**	-0.011	0.092*
	(0.016)	(0.046)	(0.013)	(0.052)
Obs	153,294	153,294	224,942	224,942
Clerk	0.015**	-0.004	0.022**	0.033
	(0.006)	(0.014)	(0.010)	(0.024)
Obs	153,294	153,294	224,942	224,942
Serve	-0.022*	0.037	0.015	-0.044
	(0.011)	(0.033)	(0.016)	(0.038)
Obs	153,294	153,294	224,942	224,942
Craft	-0.015	-0.033	-0.043	0.024
	(0.029)	(0.067)	(0.027)	(0.048)

Continued on next page

Table 9 – continued from previous page

	[1] OLS	[2] IV	[3] OLS	[4] IV
Dependent variable	1996	1996	2001	2001
Obs	153,294	153,294	224,942	224,942
Skilled agriculture	0.002	-0.073	0.014	-0.059*
	(0.021)	(0.046)	(0.015)	(0.033)
Obs	153,294	153,294	224,942	224,942
Operator	-0.036**	-0.143***	-0.038**	-0.058
	(0.016)	(0.047)	(0.018)	(0.038)
Obs	153,294	153,294	224,942	224,942
Unskilled	0.068*	0.094	0.035	-0.010
	(0.035)	(0.072)	(0.035)	(0.064)
Obs	153,294	153,294	224,942	224,942
Individual controls	YES	YES	YES	YES
District controls	YES	YES	YES	YES
Province FE	YES	YES	YES	YES

Note: This table reports the main results about the heterogeneous effects of ethnic diversity on the probability of being an employee at individual-level regressions by sub-groups in both 1996 and 2001 census. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. All the columns in Panel B and Panel C only focus on employees to illustrate the allocation of employed workers across different industrial sectors and occupations. *** p<0.01, ** p<0.05, * p<0.1.

	[1]	[2]	[3]	[4]	[5]	[6]
	Unemployed + inactive			Wage employment		
	Native	Migrants	Immigrants	Native	Migrants	Immigrants
Panel A: IV estimates, 1996 census						
Ethnic fractionalisation ELF	-0.161**	0.142	-0.506	0.171**	-0.131	0.530*
	(0.079)	(0.133)	(0.321)	(0.081)	(0.134)	(0.316)
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
F statistics of the instrument	24.05	15.52	8.432	24.12	15.51	8.826
R-squared	0.191	0.193	0.299	0.188	0.196	0.330
Observations	305,458	128,215	4,657	296,864	122,956	4,283
Panel B: IV estimates, 2001 census						
Ethnic fractionalisation ELF	-0.153*	-0.247	-0.991*	0.158**	1.048*	0.252
	(0.080)	(0.211)	(0.592)	(0.080)	(0.624)	(0.209)
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
F statistics of the instrument	33.73	16.65	6.618	33.90	17.14	6.666
R-squared	0.171	0.196	0.289	0.168	0.198	0.312
Observations	568,260	119,696	20,390	556,296	116,089	19,250

Note: This table reports results about the effect of ethnic diversity on employment separately for native, migrants and immigrants at individual-level regressions based on 1996 and 2001 census data. "Native" is defined as people who were born in the district and never move out or within-district migrants. "Migrants" are cross-district migrants while "immigrants" are those who migrated from another country. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. In columns 4-6 we drop self-employed people as they are a very small proportion of the whole working-age population. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Ethnic diversity and employment: separating native, migrants and immigrants

	[1]	[2]	[3]	[4]	[5]	[6]
		1996			2001	
	Number of white in 1996	Number of white in 1985	Diffrence: 96 - 85	Number of white in 2001	Number of white in 1985	Diffrence: 01 - 85
Panel A: OLS estimates						
Ethnic fractionalisation ELF	12758.975	765.232	11993.742	-9966.072	4529.285	-14495.357
	(25151.926)	(31380.028)	(13971.953)	(29215.642)	(25530.141)	(15013.818)
R-squared	0.766	0.781	0.452	0.767	0.895	0.782
Observations	205	205	205	210	210	210
Panel B: IV estimates						
Ethnic fractionalisation ELF	145935.526	226342.130	-80406.604	33390.776	15067.044	18323.731
	(185457.367)	(204618.284)	(59332.187)	(106236.562)	(84187.211)	(33223.976)
F statistics of instruments	10.19	10.19	10.19	29.85	29.85	29.85
R-squared	0.735	0.720	0.349	0.765	0.894	0.776
Observations	205	205	205	210	210	210
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES

Note: This table looks at whether ethnic diversity is correlated with the number of white population in 1996 and 2001 and the emigration of the white from the district after the end of Apartheid at district-level regressions. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features and individual-level controls aggregated at district average. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Ethnic diversity and the emigration of the white

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	Wage employee				Common language (only 1996)			
	Whole sample	Large	Medium	Small	Whole sample	Large	Medium	Small
Panel A: 1996 census								
Dispersion of size	-0.087***	-0.110***	-0.015	0.014	-0.035*	-0.028	-0.054	0.015
	(0.033)	(0.034)	(0.073)	(0.102)	(0.018)	(0.020)	(0.033)	(0.032)
1/No. of groups	-0.170***	-0.194***	-0.207	0.405	0.016	0.020	0.127	-0.097
	(0.044)	(0.045)	(0.311)	(0.555)	(0.028)	(0.031)	(0.366)	(0.177)
R-squared	0.194	0.187	0.201	0.230	0.068	0.068	0.072	0.061
Observations	449,200	320,901	91,373	36,926	654,116	469,737	131,601	52,778
Panel B: 2001 census								
Dispersion of size	-0.148***	-0.152***	-0.207***	-0.033				
	(0.037)	(0.031)	(0.057)	(0.083)				
1/No. of groups	-0.284***	-0.266***	0.419	-0.555				
	(0.065)	(0.057)	(0.777)	(1.126)				
R-squared	0.173	0.164	0.177	0.200				
Observations	681,529	459,108	149,632	72,789				
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES

Note: This table reports results based on the decomposition of ethnic diversity index into items relating to number of ethnic groups and group share, and how these two items are associated with employment rate at individual-level OLS regressions with 1996 and 2001 census data. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. We look at both the whole sample and sub-samples split by sample size. *** p<0.01, ** p<0.05, * p<0.1.

Table 12.1: Decomposing ethnic diversity into number of groups and dispersion of population share

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	Wage employee 1996			Wage employee 2001			Common language 1996		
	Large	Medium	Small	Large	Medium	Small	Large	Medium	Small
Panel A: Instrument dispersion of size									
Dispersion of size	-0.166**	0.166	0.570*	-0.213**	0.030	-0.202	-0.136*	0.068	-0.012
	(0.077)	(0.197)	(0.319)	(0.084)	(0.203)	(0.429)	(0.073)	(0.168)	(0.097)
1/No. of groups	-0.249***	-0.073	0.864	-0.335***	0.531	-0.603	-0.086	0.224	-0.116
	(0.080)	(0.374)	(0.604)	(0.107)	(0.721)	(1.059)	(0.079)	(0.434)	(0.190)
F statistics of the instrument	17.98	7.662	9.421	23.68	12.11	4.009	18.31	8.424	10.03
R-squared	0.187	0.200	0.223	0.164	0.175	0.199	0.066	0.070	0.061
Observations	320,901	91,373	36,926	459,108	149,632	72,789	469,737	131,601	52,778
Panel B: Instrument number of groups									
Dispersion of size	-0.270	0.057	0.113	-0.505	0.065	-0.044	-0.313	-0.006	0.011
	(0.248)	(0.156)	(0.135)	(0.751)	(1.802)	(0.101)	(0.322)	(0.060)	(0.036)
1/No. of groups	-0.566	1.514	4.145	-1.781	56.536	-8.462	-0.619	1.256	-0.289
	(0.573)	(2.867)	(2.830)	(3.415)	(377.818)	(26.742)	(0.716)	(1.169)	(0.630)
F statistics of the instrument	0.897	1.128	3.477	0.204	0.0212	0.331	1.112	1.393	3.392
R-squared	0.184	0.196	0.220	0.142	-0.825	0.193	0.038	0.066	0.061
Observations	320,901	91,373	36,926	459,108	149,632	72,789	469,737	131,601	52,778
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note: This table reports results based on the decomposition of ethnic diversity index into items relating to number of ethnic groups and group share, and how these two items are associated with employment rate at individual-level IV regressions with 1996 and 2001 census data. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. We look at the sub-samples split by group size. *** p<0.01, ** p<0.05, * p<0.1.

Table 12.2: Verifying that the instrument variable captures the dispersion of group size

	[1] OLS	[2] IV
Panel A: Whole sample		
Overall	0.036**	0.109*
	(0.018)	(0.065)
F statistics of the instrument		23.07
Obs	654,116	654,116
Panel B: By ethnicity		
Large	0.028	0.161**
	(0.020)	(0.078)
F statistics of the instrument		17.30
Obs	469,737	469,737
Medium	0.052	-0.078
	(0.034)	(0.188)
F statistics of the instrument		9.052
Obs	131,601	131,601
Small	-0.014	0.024
	(0.031)	(0.093)
F statistics of the instrument		10.76
Obs	52,778	52,778
Individual controls	YES	YES
District controls	YES	YES
Province FE	YES	YES

Note: This table reports results about the effect of ethnic diversity on the acquisition of social skills (proficiency of second language as a proxy) at individual-level regressions based on 1996 census data (as there is no information on the second language in 2001 census). The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. We look at both the whole sample and sub-samples split by group size. *** p<0.01, ** p<0.05, * p<0.1.

Table 13.1: Ethnic diversity and skill acquisition: second language

	[1] <15	[2] 15-64	[3] >=65
Panel A: Large group			
ELF	0.121**	0.174**	0.236**
	(0.057)	(0.088)	(0.110)
F statistics of the instrument	17.34	17.22	14.89
Obs	108,939	342,890	17,908
Panel B: Medium group			
ELF	-0.079	-0.064	-0.458
	(0.115)	(0.209)	(0.350)
F statistics of the instrument	11.10	8.097	12.96
Obs	28,848	97,754	4,999
Panel C: Small group			
ELF	0.093	-0.005	-0.033
	(0.073)	(0.106)	(0.122)
F statistics of the instrument	11.96	9.896	15.56
Obs	10,238	40,690	1,850
Individual controls	YES	YES	YES
District controls	YES	YES	YES
Province FE	YES	YES	YES

Note: This table reports results about the effect of ethnic diversity on the acquisition of social skills (proficiency of second language as a proxy) at individual-level regressions based on 1996 census data (as there is no information on the second language in 2001 census). The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. In particular, we split the sample by age groups. *** p<0.01, ** p<0.05, * p<0.1.

Table 13.2: Ethnic diversity and skill acquisition across age groups: second language

	[1] Overall	[2] Large	[3] Medium	[4] Small
Panel A: Unemployed as dependent variable, conditional on diversity				
Second official	-0.133***	-0.123***	-0.144***	-0.176***
	(0.013)	(0.010)	(0.022)	(0.036)
Individual controls	YES	YES	YES	YES
District controls	YES	YES	YES	YES
Province FE	YES	YES	YES	YES
R-squared	0.201	0.193	0.212	0.229
Obs	461,942	329,416	93,673	38,853
Panel B: Wage employ as dependent variable, conditional on diversity				
Second official	0.132***	0.121***	0.143***	0.179***
	(0.013)	(0.011)	(0.023)	(0.036)
Individual controls	YES	YES	YES	YES
District controls	YES	YES	YES	YES
Province FE	YES	YES	YES	YES
R-squared	0.199	0.191	0.209	0.236
Obs	447,103	319,580	90,817	36,706

Note: This table reports results about the relationship between social skill acquisition (proficiency of second language as a proxy) and employment at individual-level regressions based on 1996 census data. We control for ethnic diversity and investigate whether this language skill is positively correlated with employment chances. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. In Panel A we keep the whole working-age black sample while in Panel B we drop self-employed people as they are a very small proportion of the whole working-age population. *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Skill acquisition and labour market outcomes: second language

Chapter 3

Returns to Education, Marital Sorting and Family Background in South Africa

Patrizio Piraino and Peng Zhang

Abstract. This paper investigates whether father-in-law's background has a stronger explanatory power than father's background for male wages in South Africa. We also investigate the heterogeneity of the effects of father and father-in-law's background at different educational levels and between different population groups. After correcting for potential measurement errors in both earnings and education, we find that, consistent with existing studies on Brazil, father-in-law's schooling has a larger effect on male workers' labour market earnings than own father's schooling. The gap increases with parental educational levels. Father's background plays a more important role in explaining labour market income of black South African males compared to white males, while this does not hold for father-in-law's background. Furthermore, we find that family background and marital sorting are also related to other labour market outcomes, such as employment and labour force participation.

Keywords: Intergenerational mobility; Marital sorting; Education; South Africa.

JEL classification: I24, J24, J62.

3.1 Introduction

Research on intergenerational transmission of socio-economic status shows that parental education plays an important role in children's socio-economic outcomes such as income in both developing and developed countries (Solon, 1992; Piraino, 2015; Dahl and DeLeire, 2008; Dustmann, 2005; Bingley and Cappellari, 2014; Eisenhauer and Pfeiffer, 2008; Mare, 2000). An interesting phenomenon is that in developing countries like Brazil, fathers-in-law's schooling explains more of male workers' wage than father's schooling (Lam and Schoeni, 1993), which is however not the case in more developed countries such as the US (Lam and Schoeni, 1994) and UK (Blanden, 2004). This is not likely driven by factors commonly believed to be responsible for the transmission of socio-economic status across generations such as family connections. For example, if children from better educated parents have higher income due to the network provided by the family, it is reasonably to assume that this nepotism is larger between father and son than that between father-in-law and son-in-law. This means the educational background of fathers-in-law should have smaller explanatory power than that of fathers, which contradicts the above empirical finding.

In Lam and Schoeni (1993), they provide a model that links the degree of assortative mating to intergenerational inequality to explain this paradox. The mechanism is as follows. Earnings are determined by education and unobserved individual characteristics such as ability. Family background can be considered

as a signal for this unobserved ability. For male workers, own father's education serves as proxy for only the inherited components of the ability while father-in-law's education signals both the inherited and uninherited determinants of ability. As there is more information in father-in-law's schooling than father's schooling, father-in-law's background will have a relatively larger explanatory power in male workers' earnings. In Lam and Schoeni (1994), they compare the results from a series of wage equations in the United States and Brazil and further show that the explanatory power of father-in-law's education is larger in countries characterised by a higher degree of assortative mating and lower female labour force participation.

Bringing assortative mating into research on intergenerational transmission of socio-economic status continues to attract attention in more recent literature. Papers focusing on different countries may show results different from the evidence in Brazil but most of them confirm that marital sorting is unneglectable in determining intergenerationally transmitted earnings (Holmlund, 2008). For example, Handy (2015) uses the US data and finds that assortative mating explains around one quarter of the intergenerational correlation in education and income. Using data from Germany and the UK, Ermisch et al. (2006) show that marital sorting explains a high proportion of the covariance between offspring's and parents' family income in both countries. Eika et al. (2014) focus on the contribution of gender and marital status to intergenerational mobility in the United States and Norway. In their paper, the intergenerational earnings mobility of married women is influenced by a combination of assortative mating and labour supply responses.

It is therefore worthwhile to further investigate how assortative mating and intergenerational transmission jointly determine earnings by expanding the pioneering work of Lam and Schoeni (1993, 1994). In particular, three questions remain to be answered: are the effects of father and father-in-law's background on male workers' earnings nonlinear across different educational categories? Are the current results contaminated by measurement errors and inaccurate measure of lifetime earnings (i.e. using transitory earning as a proxy for permanent earning) which are usually the challenges in research on intergenerational transmission? Are the results from Brazil applicable to other developing countries with more complicated structure of labour market and more serious unemployment issues?

Based on these, in this paper we follow the framework of Lam and Schoeni (1993, 1994) and re-examine the paradox that father-in-law's education explains more of male workers' income than father's education by looking at the intergenerational transmission and assortative mating in the South African labour market. The contributions of the paper lie in the following three aspects.

Firstly, we capture the heterogeneity of the effects of father and father-in-law's background at different educational levels and document its nonlinearity. It has been noticed that returns to education can be convex in response to educational levels in developing countries. Some studies suggest it might be simply due to selection into higher education (Colclough et al., 2009). As educational levels increase, people who choose to attend higher educational institutes will have higher abilities. Therefore the increase in estimated returns to education may only capture the increased omitted variable bias due to unobserved ability as educational levels increase. In this paper, we show that whether this explanation is true can be partly revealed from parental information. We find that this convexity also exists in returns to father and father-in-law's education. Furthermore, the magnitude and convexity of the effect of own education, although declining, do not disappear after controlling for parental background. Following the intuition that father and father-in-law's education capture unobserved ability, this result indicates that the convexity in the returns to education is not totally driven by the selection of people with higher ability to enter higher educational institutes and therefore implies that real returns to high school and college education can be high in developing countries.

Secondly, we deal with measurement errors and the challenge of measuring permanent earnings often faced by research on intergenerational transmission, based on a panel survey of households with adjusted

value of education and income according to updated information in each wave. In Lam's paper linking family background and assortative mating, he pointed out that even if the magnitude of measurement errors in schooling is small, increase in measurement error bias after controlling for parental background can account for as much as 80% of the decrease in returns to schooling. Therefore, over the last decade, increased availability of longitudinal data and advances in econometric techniques have allowed economists to overcome some of the most common measurement issues and to obtain reliable estimates of intergenerational income mobility in a number of countries (Corak, 2006; Mazumder, 2005). Similarly, we use panel data to correct for errors from transitory income data by averaging income across waves. Furthermore, as the data also reports the "best" estimate of education and income which corrects for measurement errors in any single wave, we can further deal with measurement errors by using "best" estimates rather than original values.

Thirdly, we conduct analysis similar to Lam and Schoeni (1993, 1994) with the above two improvements in the setting of South Africa to see whether their results can be generalised to other developing countries besides Brazil and how their conclusions might differ when we turn to a country with more complicated structure of labour market and more serious unemployment issues. South Africa has two features that lead to its distinguished labour market structure. On the one hand, South Africa is characterised by segmented but highly linked labour market, which can be revealed by the diverse response to intergenerational transmission between black and white. The forced racial segregation of the past has a legacy in the highly segmented society of the present. Income inequality is among the highest in the world, and vast differences remain across racial groups in a variety of socioeconomic outcomes (Leibbrandt et al., 2009).

As the degree of intergenerational transmission can be different across groups within the same country (Mazumder, 2011), especially between different races (Hertz, 2008), it is worthwhile to investigate the racial difference in assortative mating and intergenerational correlation between the black and white in South Africa. Specifically, given the fragmented labour market and high inequality between black and white in South Africa, it is reasonable to assume that the degree of assortative mating and the rate of female labour force participation, the two most important factors determining the effects of father-in-law's background, can be quite different between the black and the white. It also makes sense to focus on the assortative mating within race as inter-racial marriages are still uncommon. Furthermore, unlike the case in Brazil, labour force participation rate is very low in South Africa. As a result, it would also be valuable to look at how parental background and assortative mating determine other labour market outcomes such as probability of employment and labour force participation besides earnings.

In our paper, after controlling for measurement errors in education and income, the earnings equations of South African male workers show that parental background variables have significant explanatory power. We find that father-in-law's schooling has a larger effect on labour market earnings than own father's schooling, consistent with the result from Brazil. To further confirm the relatively larger effect of father-in-law's background, we find that the drop in the coefficients on father's schooling is larger than that on father-in-law's schooling when own schooling is included in the earnings equation, indicating that the information contained in father's schooling and own education is highly overlapped so that the marginal contribution of father's education to information on unobserved ability is relatively limited. Following the same reasoning, the drop in the coefficients on own schooling is smaller when father's schooling is included than when father-in-law's schooling is included. Furthermore, parental background and martial sorting not only explain labour market earnings but also relate to other labour market outcomes such as probability of employment and labour force participation.

In terms of the heterogeneity of the effects of parental background, first, we find that there is convexity in the structure of returns to education. In particular, the patterns we observed for the average schooling effect are mostly displayed at the higher level of education categories. This convexity declines after controlling for

parental background but does not disappear. This implies that there is an endogenous selection of people with high ability into higher education but this selection does not fully capture the pattern of convex returns to education. Second, there is heterogeneity between the black and the white in response to intergenerational transmission and assortative mating. Parental background plays a more important role in labour market income for the black while this pattern does not exist for father-in-law's education. The reason why the role of father-in-law's education does not differ between the black and white might be that assortative mating is lower and the male-female ratio of the variance in income is larger among the white than that among the black (both are verified by data). The former leads to a smaller effect of father-in-law's education relative to father's education while the latter indicates a larger relatively effect of father-in-law's schooling.

The rest of the paper is structured as follows. The next section offers a theoretical discussion of the links between marital sorting and intergenerational mobility. Section 3 describes the data used for the empirical analysis. Section 4 provides empirical framework. The results are presented and discussed in Section 5. Section 6 concludes.

3.2 Theoretical Framework

3.2.1 Model

We use the notation m for men (i.e. husband) and w for women (i.e. wife). Suppose y_m, S_m and A_m are husband's labour market income, schooling and ability, respectively. We do not consider inherited wealth (like bequest) here as we only focus on labour market outcome for empirical study[1]. F_m is father's schooling, which captures inherited part of schooling and ability. S_m^u and A_m^u are uninherited parts which determine schooling and ability. Labour market income is determined by observed schooling and unobserved overall ability. Here we have:

$$y_m = \beta_{0m} + \beta_{sm} S_m + \beta_{am} A_m + u_m \tag{3.1}$$

$$S_m = \alpha_{sm} + \gamma_{sm} F_m + S_m^u \tag{3.2}$$

$$A_m = \alpha_{am} + \gamma_{am} F_m + A_m^u \tag{3.3}$$

If we regress husband's labour market income on father's education, we have:

$$y_m = \delta_{0m} + \delta_{fm} F_m + v_m \tag{3.4}$$

Here:

$$\delta_{fm} = \beta_{sm}\gamma_{sm} + \beta_{am}\gamma_{am} \tag{3.5}$$

$$\delta_{0m} = \beta_{0m} + \beta_{sm}\alpha_{sm} + \beta_{am}\alpha_{am}$$

Similarly,

$$y_w = \delta_{0w} + \delta_{fw} F_w + v_w,\ \delta_{fw} = \beta_{sw}\gamma_{sw} + \beta_{aw}\gamma_{aw} \tag{3.6}$$

Based on Lam's model, we make the following three assumptions.

[1] Following Lam and Schoeni (1993) which treats inherited wealth and labour market earnings separately.

1. The degree of assortative mating is captured by the correlation of lifetime income between husband and wife, proxied by $\rho_{y_m y_w}$.

2. There is no additional assortative mating other than income. This means that given the income of both husband and wife, the degree of assortative mating does not depend on other socio-economic factors such as education. The intuition of this assumption is that the husband with higher income prefers a spouse with higher income. But he does not care whether this high income is from her high educational level or good luck in the labour market. Mathematically, the correlation between husband's income and wife's other socio-economic factors should be zero conditional on wife's income, meaning: $\rho_{y_m s_w \cdot y_w} = 0$, $\rho_{y_w s_m \cdot y_m} = 0$, $\rho_{y_m a_w \cdot y_w} = 0$, $\rho_{y_w a_m \cdot y_m} = 0$, $\rho_{y_m f_w \cdot y_w} = 0$, $\rho_{y_w f_m \cdot y_m} = 0$.

3. Uninherited components of ability and schooling are not correlated.

According to Lam and Schoeni (1993), from the second assumption above, we can get the following equation.

$$\rho_{y_m f_w} = \rho_{y_w y_m} \cdot \rho_{y_w f_w} \tag{3.7}$$

With the above three assumptions, we consider the following regressions.

First, directly from equation 3.4 and 3.5, if we regress husband's labour market income on his father's education and define the coefficient of father's education as $\beta_{y_m f_m}$, we have:

$$plim\beta_{y_m f_m} = \rho_{y_m f_m} \cdot \frac{\sigma_{y_m}}{\sigma_{f_m}} \tag{3.8}$$

Second, we regress husband's labour market income on his father's schooling, conditional on its own schooling. Suppose the coefficient of father's education in this regression is $\beta_{y_m f_m \cdot s_m}$. We have:

$$plim\beta_{y_m f_m \cdot s_m} = \rho_{y_m f_m} \cdot \frac{\sigma_{y_m}}{\sigma_{f_m}} [\frac{1 - \frac{\rho_{s_m f_m} \cdot \rho_{s_m y_m}}{\rho_{y_m f_m}}}{1 - \rho^2_{s_m f_m}}] \tag{3.9}$$

Third, when we regress income on father-in-law's schooling, we get the coefficient of father-in-law, defined as $\beta_{y_m f_m}$. We have:

$$plim\beta_{y_m f_w} = \delta_{f_w} \cdot \rho_{y_m y_w} \cdot \frac{\sigma_{y_m}}{\sigma_{y_w}} \tag{3.10}$$

Fourth, we regress income on father-in-law's schooling conditional on own schooling, get the coefficient of father-in-law's schooling, defined as $\beta_{y_m f_w \cdot s_m}$. We have:

$$plim\beta_{y_m f_w \cdot s_m} = \delta_{f_w} \cdot \rho_{y_m y_w} \cdot \frac{\sigma_{y_m}}{\sigma_{y_w}} [\frac{1 - \rho^2_{y_m s_m}}{1 - \rho^2_{y_m s_m} \rho^2_{y_m y_w} R^2_{y_w f_w}}] \tag{3.11}$$

Here $R^2_{y_w f_w}$ is the R^2 from a regression of y_w on f_w. The proofs of equation 3.7 - 3.11 are in the Appendix.

3.2.2 Schooling, intergenerational transmission and assortative mating

From the above four formulas, we can make the following propositions which shed light on schooling, intergenerational transmission and assortative mating. The first three propositions are explain in Lam and Schoeni (1993) with intuitions and we provide rigorous proof here.

Proposition 3.2.1. *Father-in-law's schooling can have a more explanatory power on labour market earnings than own father's schooling. And the explanatory power depends on the degree of assortative mating and the*

variance in male income relative to that in female income. That is: it is possible to have $\beta_{y_m f_w} > \beta_{y_m f_m}$ when $\rho_{y_m y_w}$ is not too small and $\sigma_{y_m} > \sigma_{y_w}$ (or equivalently: $\sigma_{u_m} > \sigma_{u_w}$).

Proof. When equation 3.7 holds, we have:

$$\frac{\rho_{y_m f_w}}{\rho_{y_w f_m}} = \frac{\rho_{y_m y_w}\rho_{y_w f_w}}{\rho_{y_m y_w}\rho_{y_m f_m}} = \frac{\rho_{y_w f_w}}{\rho_{y_m f_m}}$$

Therefore:

$$\rho_{y_m f_w} - \rho_{y_m f_m} = \frac{\rho_{y_w f_w} \cdot \rho_{y_w f_m} - \rho^2_{y_m f_m}}{\rho_{y_m f_m}}$$

When $\rho_{y_w f_w} > \rho_{y_m f_m}$, it is possible to have: $\rho_{y_m f_w} > \rho_{y_m f_m}$
Since:

$$\rho_{y_m f_m} = \frac{Cov(Y_m F_m)}{\sigma_{y_m}\sigma_{f_m}},\ \rho_{y_w f_w} = \frac{Cov(Y_w F_w)}{\sigma_{y_w}\sigma_{f_w}}$$

When other parts are equal, $\rho_{y_w f_w} > \rho_{y_m f_m}$ implies $\sigma_{y_m} > \sigma_{y_w}$ (or equivalently: $\sigma_{u_m} > \sigma_{u_w}$).

Furthermore, given equation 3.10, as $\rho_{y_m y_w}$ or $\frac{\sigma_{y_m}}{\sigma_{y_w}}$ increases, the explanatory power of father-in-law will also increase.

□

Proposition 3.2.2. *The relative drop in the coefficients on father's schooling can be larger than that on father-in-law's schooling when own schooling is included in the earnings equation. That is to say: $\frac{\beta_{y_m f_m} - \beta_{y_m f_m \cdot s_m}}{\beta_{y_m f_m}} > \frac{\beta_{y_m f_w} - \beta_{y_m f_w \cdot s_m}}{\beta_{y_m f_w}}$. It is especially true when $\rho_{y_m y_w}$ or $\frac{\sigma_{y_m}}{\sigma_{y_w}}$ is large.*

Proof. Based on equations 3.8 - 3.11, we have:

$$\frac{\beta_{y_m f_m} - \beta_{y_m f_m \cdot s_m}}{\beta_{y_m f_m}} = 1 - \frac{1 - \frac{\rho_{s_m f_m} \cdot \rho_{s_m y_m}}{\rho_{y_m f_m}}}{1 - \rho^2_{s_m f_m}}$$

$$\frac{\beta_{y_m f_w} - \beta_{y_m f_w \cdot s_m}}{\beta_{y_m f_w}} = 1 - \frac{1 - \rho^2_{y_m s_m}}{1 - \rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}$$

For proposition 3.2.2 to hold, we must have:

$$1 - \frac{1 - \frac{\rho_{s_m f_m} \cdot \rho_{s_m y_m}}{\rho_{y_m f_m}}}{1 - \rho^2_{s_m f_m}} > 1 - \frac{1 - \rho^2_{y_m s_m}}{1 - \rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}$$

This is the condition which has to be satisfied. It is easy to get:

$$\frac{\partial[\frac{1-\rho^2_{y_m s_m}}{1-\rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}]}{\partial \rho_{y_m y_w}} > 0$$

And:

$$\frac{\partial[\frac{1-\rho^2_{y_m s_m}}{1-\rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}]}{\partial R^2_{y_w f_w}} > 0$$

That is, when $\rho_{y_m y_w}$ or $\frac{\sigma_{y_m}}{\sigma_{y_w}}$ is large, the righthand side of the above inequality is smaller, meaning the condition is easier to satisfy.

□

Proposition 3.2.3. *The coefficients on own schooling can decline more when father-in-law's schooling is included, than when father's schooling is included. That is,* $\beta_{y_m s_m \cdot f_m} < \beta_{y_m s_m}$, $\beta_{y_m s_m \cdot f_w} < \beta_{y_m s_m}$, *and* $\frac{\beta_{y_m s_m} - \beta_{y_m s_m \cdot f_m}}{\beta_{y_m s_m}} < \frac{\beta_{y_m s_m} - \beta_{y_m s_m \cdot f_w}}{\beta_{y_m s_m}}$.

Proof. This can be proved from the perspective of omitted variable bias. When investigating returns to own education, the true regression function is the equation 3.1. However, we rely on the following regression instead:

$$y_m = \beta_{0m} + \beta_{sm} S_m + u_m$$

The estimated value is:

$$\hat{\beta_{y_m s_m}} = \frac{Cov(\beta_{0m} + \beta_{sm} S_m + \beta_{am} A_m + u_m) S_m)}{Var(S_m)} = \beta_{sm} + \beta_{am} \frac{Cov(S_m A_m)}{Var(S_m)} \tag{3.12}$$

When adding F_m into the regression, the omitted variable bias (i.e. $\beta_{am} \frac{Cov(S_m A_m)}{Var(S_m)}$) is reduced as F_m captures some information on the unobserved ability in u_m.

Taking father's education as an example. When adding father's education, the new estimated coefficient is:

$$\hat{\beta_{y_m s_m \cdot f_m}} = \frac{Var(F_m) Cov(Y_m S_m) - Cov(F_m S_m) Cov(Y_m F_m)}{Var(S_m) Var(F_m) - Cov^2(F_m S_m)}$$

$$y_m = \tilde{\beta_{0m}} + \beta_{sm} S_m + \beta_{am} \gamma_{am} F_m + \beta_{am} A_m^u + u_m$$

With some algebra we can prove:

$$\hat{\beta_{y_m s_m \cdot f_m}} = \beta_{sm} + \beta_{am} \frac{Cov(S_m A_m^u)}{Var(S_m)} (1 - \rho^2_{A^u F_m \cdot S}) \tag{3.13}$$

$\rho^2_{A^u F_m \cdot S}$ is the partial correlation of uninherited ability and father's education with control of own education. As father-in-law's education captures both inherited and uninherited ability while father's education only captures inherited ability, this partial correlation can be larger when father-in-law's education is added than when father's education is controlled. Thus the omitted variable bias $\beta_{am} \frac{Cov(S_m, A_m^u)}{Var(S_m)} (1 - \rho^2_{A^u F_m \cdot S})$ is smaller when we add father-in-law's education rather than father's education, meaning the coefficient of own education will decreases more towards the true value.

□

Proposition 3.2.4. *Whether the returns to own education are convex simply due to endogenous selection into higher education can be partly revealed from father and father-in-law's education.*

Proof. According to the equation 3.12, $\beta_{am} \frac{Cov(S_m A_m)}{Var(S_m)}$ is the omitted variable bias. If there is more selection into education when educational levels become higher, it means $Cov(S_m A_m)$ increases with educational levels, leading to more severely overestimated returns to own education.

According to the equation 3.13, if father's education and father-in-law's education altogether captures inherited and uninherited abilities perfectly, this convex pattern will disappear once we control for father or father-in-law's background. And the convex pattern will be less obvious when father-in-law's education is included (when $\rho_{A^u F_m \cdot S} < \rho_{A^u F_w \cdot S}$).

□

Proposition 3.2.5. *If the degree of assortative mating is lower and the gender gap in labour force participation is higher among the white than the black, father's education is relatively more important in determining male workers' income for the black than the white but the racial difference in the effects of father-in-law's background is ambiguous.*

Proof. We will show in the descriptive statistics that the degree of assortative mating is larger and the male labour force participation is lower among the black than the white. Lower male labour force participation rate indicates that uninherited ability is less important in determining the black's job market outcome and $\rho_{s_m y_m}$ (and therefore $\rho_{f_m y_m}$) will be larger among the black. From equation 3.8 and 3.10, it is obvious to find that the effect of father's education on male workers' earnings is larger among the black than the white.

Furthermore, from equation 3.9, we notice that the relative decrease in omitted variable bias resulting from unobserved ability after controlling own education is $1 - \frac{1-\frac{\rho_{s_m f_m} \cdot \rho_{s_m y_m}}{\rho_{y_m f_m}}}{1-\rho^2_{s_m f_m}}$. This decrease is larger if $\rho_{s_m y_m}$ is higher. This implies the decrease in the coefficient of parental background is larger among the black than the white after controlling for own education.

Comparing between equations 3.10 and 3.11, the relative decrease in omitted variable bias resulting from unobserved ability after controlling own education is $1 - \frac{1-\rho^2_{y_m s_m}}{1-\rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}$. It can be shown that:

$$\frac{\partial \frac{1-\rho^2_{y_m s_m}}{1-\rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}}{\partial \rho_{s_m y_m}} < 0$$

$$\frac{\partial \frac{1-\rho^2_{y_m s_m}}{1-\rho^2_{y_m s_m}\rho^2_{y_m y_w} R^2_{y_w f_w}}}{\partial \rho_{y_m y_w}} > 0$$

It is obvious that if the uninherited ability is more important ($\rho_{s_m y_m}$ is smaller) and the degree of assortative mating is lower $\rho_{y_m y_w}$ among the white than the black, the effect of controlling for own education on father-in-law's coefficient will be ambiguous, as these two forces drive the change in the coefficients of father-in-law's education in the opposite directions.

□

3.2.3 Predictions to be tested in the empirical sections

There are five predictions derived from theoretical framework which will be tested in the empirical analysis.

1. Father-in-law's schooling can have a higher effect on labour market earnings than own father's schooling.

2. The drop in the coefficients on father's schooling is larger than that on father-in-law's schooling when own schooling is included in the earnings equation.

3. The drop in the coefficients on own schooling is smaller when father's schooling is included than when father-in-law's schooling is included.

4. If the returns to own education are convex due to endogenous selection into higher education, this convexity will be reduced after including father's education and father-in-law's education and more reduction takes place after father-in-law's education is included.

5. Parental background plays a more important role in labour market income among the black than the white while the racial difference in the effect of father-in-law's education is ambiguous. Moreover, the coefficient of parental education decreases more among the black than the white after controlling

for own education but the racial difference in the change of coefficients of father-in-law's education is ambiguous.

3.2.4 Dealing with measurement errors

Measurement errors and inaccurate estimates of permanent income are challenging to research on intergenerational transmission (Nybom and Stuhler, 2016). Accordingly, these measurement errors might contaminate the explanations on our results if they differ between variables of father and father-in-law. With measurement errors, the coefficients of own education or parental background are smaller than the true ones and the bias from the measurement errors in own education will be larger if additional variables such as parental background are added into regressions. Therefore, we may still observe the decline in the magnitude of own education after controlling for father or father-in-law's education but here the results are simply driven by the amplified measurement errors after controlling for more variables, not based on any relation between intergenerational transmission, assortative mating and income.

As is pointed out in Lam and Schoeni (1993), when we regress labour market income on own education, suppose education S_m is measured with bias: $S_m^* = S_m + w_m$. Define $\lambda = \frac{Var(w)}{Var(S^*)}$. The estimated returns to education is:

$$\hat{\beta_{ys}} = \beta_s \cdot \frac{\sigma_s^2}{\sigma_s^{*2}} = \beta_s(1-\lambda) \tag{3.14}$$

When we add parental background F_m or F_w into the regression, the new estimated coefficient becomes:

$$\hat{\beta_{ys\cdot f}} = \beta_s(1 - \frac{\lambda}{1 - R^2_{s^*.F}}) \tag{3.15}$$

$R^2_{s^*.F}$ is the R^2 from a regression of S_m^* on F[2].

Comparing equation 3.14 and 3.15, after controlling for parental education, as $R^2_{s^*.F} < 1$, the measurement error in own education will become larger and as a result the coefficient of own education will decrease. If $R^2_{s^*.F}$ is larger when father-in-law's education is included, which is very likely as father-in-law's background captures uninherited components in addition to inherited parts of ability, the measurement error will increase more after adding father-in-law's education than adding father's education. In this case, the coefficient of own education will also decrease more when father-in-law's education is controlled, compared with when father's education is included.

Measurement errors lead to the same pattern as the predictions in the theoretical framework. It is therefore necessary to deal with measurement errors before staring main empirical analysis. Methods to deal with measurement errors in both dependent and independent variables will be addressed in the following section.

3.3 Data

3.3.1 Data source

Our data is based upon the National Income Dynamics Study (NIDS) wave 1 to wave 4 (year 2008, 2010, 2012 and 2014), the first national longitudinal household survey in South Africa. It is a biannual nationwide household survey from 2008 which consisted of a nationally representative sample of approximately 7,300 households. The most recent wave was conducted in 2014 with 11895 households.

[2]See Lam's paper for details about measurement error.

NIDS used a combination of household and individual level questionnaires to obtain information on a wide selection of human capital variables, labour force experiences and demographic characteristics such as education, family relationships and household composition. The present study mainly uses information from the adult questionnaire, which includes the wages and other incomes of the adults in the household, as well as their level of education and employment status. Each household member is assigned a personal identification code. For those who co-reside with their parents, NIDS also provides information on their father, mother and spouse's identification code in order to link between parents and children and husband and wife in the same household. Therefore, for people co-residing with their parents, we can obtain the information on their fathers (i.e. by linking respondent with his father via identification code) and fathers-in-law (i.e. by linking respondent with his spouse via identification code and looking at spouse's father) from this linked data. In addition, all adults are asked to complete a section on parental background, where information on the educational qualification, living status and occupation of both parents is recalled by respondents even if they do not co-reside with their parents. Therefore, for respondents living separately with parents, parental information is obtained from this section on parental background. Key information regarding non-resident and deceased parents is obtained in this retrospective section.

Individual monthly income, which is used to capture labour market outcomes, is derived from variables in the Adult and Proxy datasets. We only consider labour market income, which is the sum of wage from the main job, wage from the second job, casual wage, self-employment income, 13th cheque, bonus payment, profit share, income from helping friends and extra piece-rate income[3].

Employment status follows the International Labour Organization's definitions. Each respondent is assigned to one of the following groups: employed, unemployed (strict definition), unemployed (broad definition) and not economically active. Being employed includes having primary jobs, secondary jobs, self-employment, causal work, agricultural work and helping others with business. We consider both the strict and broad definitions of unemployment when determining respondents' employment status.

We restrict our sample to working age males (i.e. 25 - 64 years old) who are either married or living with their partner[4]. From this sample of married individuals, we limit our analysis to those couples who have non-missing information on the earnings of the husbands and parental education.

For the purposes of identifying measurement errors and capturing permanent income rather than transitory income, we make use of the panel feature of the NIDS data. Detailed methods of dealing with these two problems are described in the following section. To construct our whole sample, we first pool the data in all the four waves and only keep those observations where information on labour market income (income is treated as missing in the year of unemployment or being outside labour force[5]), education and parental education is not missing in at least one of the four waves. That means, we only include people who are employed at least in one of the four waves and whose information on education and parental background is not missing in at least one wave. This results in an unbalanced panel with a subsample of 1447 different respondents.

[3]For some of the non-respondents of individual surveys, there is a proxy dataset which contains information provided by a proxy. Information on income in the Proxy dataset is shown via a series of unfolding brackets, rather than a point estimate of the income. We choose the mid-point of the interval the respondents' income falls in. For those whose income is above the value of the highest bracket we assign the value of the upper bound of the top bracket plus the gap between the lower bound of the highest bracket and the mid-point of the second highest bracket.

[4]We also conduct analysis for female to further verify our mechanism, which will be discussed in the empirical results. We do not present the results explicitly as the sample of working women is more selective due to lower female labour force participation rate. The results are provided upon request.

[5]Being outside labour force is defined as "economically inactive" in the NIDS data.

3.3.2 Dealing with measurement errors and permanent income measures

To deal with measurement errors in education, parental education and earnings, we use the "best" variables in the NIDS data. In wave 2-4, the NIDS team reports data containing "best" variables which indicate the best estimate of the variables like education and income given the updated information in each consecutive wave. For example, according to the Wave 4 User Manual, the "best" variable of education is created based on the following two principles. First, it compares information on education across all the waves and selects the value of education which appears the most frequently as the best estimate of education. Second, in wave 4 in particular, a subsample of individuals whose educational progression is not consistent across waves were asked an additional set of questions to make sure there is no impossible progression between grades (for example, downgrade educational levels). Therefore, for education, we use the wave 4 data as the best estimate and replace the corresponding values in education in the preceding waves with wave 4 data. The same process of reducing inconsistencies also applies to the parental information. In each wave, income measure is also replaced by the "best" estimate in the corresponding time period[6].

Estimating permanent income is based on the panel feature of the data. We replace the transitory income with the average of all non-missing income values across the four waves. The income during unemployed and economically inactive period is treated as missing. Therefore, if one only works in year 2010 and 2012, the measure of income is the average of his labour market income in these two years. This sample includes people who work in at least one of the waves, which allows us to correct for some bias from self-selection into employment in each period as we do not require that people included in the sample must work in all the four waves. We construct similar measures for employment rate and labour force participation rate. If one is employed (participate in the labour force) at least half of the time between 2008 and 2014, the dummy on employment or participation will be one[7].

3.3.3 Assortative mating and labour force participation: descriptive statistics

Patterns of assortative mating and labour force participation are in table 1, for black and white separately. The "all" category also includes the coloured and Indian population. Results are both unweighted and weighted with post-stratification weights.

Degree of assortative mating is measured with the correlation of education and income (after correction for measurement errors and transitory shocks) across all waves between husband and wife. Note that according to Lam, the most proper measurement of assortative mating based on economic status is the correlation of spouses' lifetime overall income (i.e. including labour market income and other sources of income such as bequest). As there is no data on lifetime overall income for either husband and wife, we have to investigate labour market income and education as proxies for overall income instead. According to Lam and Schoeni (1993), the correlation between husband and wife's labour market income or education is smaller than the real correlation based on their overall income, and education is a better proxy than labour market income. Although our measure might lead to the potential underestimate of the effect of assortative mating, they are still indicative of level of assortative mating and the difference between black and white in South Africa.

In the weighted data, the degree of assortative mating measured by the correlation of years of education between husband and wife is 0.603 for the black and 0.525 for the white. The degree of assortative mating is less in both groups if measured by average income across waves. The correlation of income between husband

[6]We choose the "best variable" for income in each wave although if measurement errors in the dependent variables are random, they should not affect the coefficients of regressors.

[7]This, together with the fact that we focus on the married sample, the employment rate will be higher in our sample than in other analysis.

and wife is 0.4854 for the black and -0.0028 for the white. According to Lam and Schoeni (1993), the most indicative measure of assortative mating based on lifetime earnings is actually the correlation of educational level, which is high in our sample. In all the measures, assortative mating remains higher among the black than the white. Results with unweighted data differ in magnitude but the differential pattern between the black and the white remains the same. All the above evidence shows that the degree of assortative mating between spouse is reasonably high in South Africa and is higher for the black than the white.

In table 1, neither employment rate nor labour force participation is very high in general but both are higher among the white than the black. In the weighted sample, among the black, labour force participation is 78.7% for male but only 63.2% for female. The corresponding figure for the white is 90.6% for male and 69.6% for female. Among those who participate in the labour force, 85.6% of black male are employed while the employment rate for black female is only 69%. The employment rate is 98.1% for white male and 84.5% for white female. The same pattern of labour force participation and employment can be found in the unweighted sample.

In total, there is some degree of assortative mating in South African in both racial groups but it is higher among the black than the white. Labour force participation in South African labour market is not very large in general but it is larger among the white than the black and larger for male than female. Therefore, based on our theoretical framework, we should observe that the education of fathers-in-law can explain more of respondents' income than that of fathers.

3.4 Empirical Framework

After correcting for measurement errors, our test of intergenerational transmission and assortative mating is based on the following basic earnings equation:

$$Y_{it} = \beta Edu_i + f(Age_{it}) + \gamma White_i + \delta_t + \alpha + \epsilon_{it} \tag{3.16}$$

where Y_{it} is the log transform of average income of individual i entering the data in wave t. $f(Age_{it})$ includes both linear and quadratic form of age. Age is measured at the first year when individual i appeared in the survey. $White_i$ equals 1 if one's population group is white and 0 if he is black, coloured or Indian. To account for the calendar effect arising from the year of survey, we also control for the survey year t when individual i first appeared in the data.

The key variable is Edu_i. It refers to a set of educational variables, including respondent's own education in the baseline regression, his father's education in the regression about intergenerational transmission and father-in-law's education in the regression about assortative mating. Then we regress the log transform of income on both own education and father's (father-in-law's) education to test how much father's (father-in-law's) background can explain earnings. All regressions are weighted with post-stratification weights in the data.

3.5 Empirical Results

3.5.1 Earnings equation: family background and assortative mating

We first estimate the earnings equation by regressing log earnings on own education, father's education and father-in-law's education. We are primarily interested in the magnitude of these three variables. Control variables include age, age square, a dummy on whether one belongs to the white population group and a variable indicating the year when the respondent first appeared in the data. Educational level is measured

by a continuous variable of years of education.

Table 2 presents the main results [8]. Column 1 reveals a standard Mincer regression to estimate the returns to own schooling when no family background variables are included. In columns 2 and 3, own father's and father-in-law's schooling are included as the explanatory variables, respectively. In columns 4 and 5 we add father's and father-in-law's schooling to the Mincer equation in column 1. We add father's education and father-in-law's education altogether in column 6. All models have log earnings as the dependent variables and include controls for age, age squared and a race dummy (i.e. white).

The results show that average returns to schooling in South Africa are significant. One additional year of schooling increases labour market income by 19.7% (column 1). Incidentally, this value is very close to the estimated returns in Lam and Schoeni (1994) for Brazil.

In column 2, the coefficient of father's schooling reveals that individuals with better educated fathers have an earnings advantage. An additional year of father's schooling is associated to a 10.7% increase in earnings. When comparing this coefficient to that from column 3, we observe that the magnitude of the coefficient on father-in-law's schooling is higher. The log earnings advantage associated to a one-year increase in father-in-law schooling is 13.2%. Note also that column 3 has a larger explanatory power as indicated by the value of R-square.

When we add father's schooling to the equation in column 1, we observe a decline in the effect of own schooling from 0.197 to 0.177. As expected, the apparent effect of own schooling decreases when a family background variable is added as a regressor. When we focus our attention on the magnitude of the coefficient on father's schooling, we see that it is less than half of the corresponding coefficient in column 1. That is, much of the parental advantage picked up by the father's schooling variable seems to operate through the channel of own education.

Moving to model 5, we can compare the changes in the estimated coefficients of interest when father-in-law's schooling is included in the equation. Again, the effect of own schooling is mitigated by the inclusion of a family background variable. More importantly, the fall in the coefficient on own schooling is larger in model 5 than model 4. Also, if we look at the table horizontally and compare the two coefficients on father-in-law's schooling (columns 3 and 5), we see that the drop in the apparent effect of father-in-law's schooling is smaller than the drop we observe in the coefficient on father's schooling from columns 2 to 4[9].

The difference between father and father-in-law's education is further confirmed in column 6 when we add both of them at the same time. Father-in-law;s education has a larger explanatory power and the coefficient of father's education is even not significant. Moreover, the decline in the coefficient of father's education after adding own education (from 0.107 to 0.011) is still larger than that of father-in-law's education (from 0.132 to 0.0718).

The above results can be interpreted in light of the theoretical discussion in Section 2 of this paper. First, earnings equations for South African men show that parental background variables have significant explanatory power. Father-in-law's schooling has larger effects on earnings compared to own father's schooling. That husband's earnings may be more highly correlated with their wives' family background than with their own family background was one of the implications of the theoretical model we discussed above. If we consider family background as a signal for unobserved individual characteristics, the results confirm that own father's education serves as proxy for just the inherited components of earnings. Father-in-law's education signals both the inherited and uninherited determinants of earnings due to high degree of assor-

[8] We also have results based on the original values of education with the one-shot observation of income, the original values of education with average income and education corrected for measurement error with the one-shot observation of income, respectively. These tables are available in the online version of the paper.

[9] The change in coefficients of father-in-law's schooling is $\frac{0.132-0.076}{0.132} = 0.424$ and the change in father's schooling is $\frac{0.107-0.0409}{0.107} = 0.62$.

tative mating and larger variance in male income relative to female income. The fact that returns to own schooling drop more after controlling for father-in-law's education than father's education further confirms that father-in-law's education captures more information on ability than father's schooling.

The second main result is the differential drop in the coefficients on father's schooling and father-in-law's schooling when own schooling is included in the earnings equation. The theoretical model presented above allows a revealing interpretation of this finding. After controlling for schooling, we expect to see a large decline in the apparent effect of own father's education on earnings when parental advantage operates mostly through the offspring's education. On the contrary, we expect a smaller decline in the apparent effect of father-in-law's education on husband's earnings as information on uninherited ability contained in father-in-law's schooling cannot be captured by own schooling[10].

3.5.2 Heterogeneity of the effects of family background and assortative mating

Convexity of returns to education and family background: categorical educational levels

The effects of parental background and assortative mating might be nonlinear as educational levels increase. Table 3 replaces the year of schooling variable with education categories. South Africa is characterised by a strong convex relationship between education and earnings (Keswell and Poswell, 2004). Much of the estimated average returns to schooling derive from superior returns at higher education levels. In particular, a spike in the estimated returns to schooling is typically observed for individuals who have at least obtained a high school degree (the matric).

As year 7 is the final year of primary school and year 12 is the last year of secondary school (i.e. completing matriculation), we define categorical educational levels as follows: educational level = 1 if no education, = 2 if 1-6 years of schooling, = 3 if 7-11 years of schooling, = 4 if finishing matriculation, = 5 if college or postgraduate. Using the first category (i.e. no education) as the base group, we find that the results in table 3 confirm the convexity of the return structure in South Africa. In columns 1 - 3, returns to all schooling measures increase sharply with the progression of education at the categories of middle school and college. For example, in column 1, receiving a junior high school degree, compared with no education, increases labour market income by 104.2%, while a college degree leads to over 250% increase. The same pattern is also shown in father's education and father-in-law's education.

Most importantly, we note that the apparent effects of the schooling of fathers and fathers-in-law are consistent with the results shown in the previous tables. Comparing column 2 and 3, we find that the magnitude of coefficients is larger in father-in-law's education than father's education in almost all educational levels, especially at the matric level. Again, this is further confirmed in column 6 where we add both of them altogether.

Furthermore, the changing patterns in returns to parental schooling after adding one's own schooling are now mostly displayed in all education categories. For example, comparing the gap in the magnitude of father's education before and after controlling for own education (column 2 and 4) with the gap in the magnitude of father-in-law's education (column 3 and 5) reveals that coefficients of father-in-law's education decrease less than father's education in almost every educational level.

[10] To further verify the mechanism, we conduct the same analysis on the female sample and all the findings are the opposite to those based on the male sample, as expected. For example, father's education has more explanatory power than father-in-law's education (the return to father's education is 0.138 and it is 0.095 for father-in-law's education). This also confirms the intergenerational transmission between father and daughter is larger than that between father and son (0.138 versus 0.107). Returns to own education decreases more when father's education is included compared with when father-in-law's education is considered (it decreases from 0.213 to 0.2 when adding father-in-law's education and from 0.213 to 0.18 when adding father's education). Returns to father-in-law's education decline more when own education is included (returns to father's education become half of the original value when adding own education but father-in-law's education drop to only one third of the original value).

Another important result is that the convexity in returns to education drops obviously after controlling for parental background especially at the higher education categories but do not disappear completely. This implies that selection into higher education based on both inherited and uninherited ability is one of the reasons explaining the convex returns to education but may not be the only determinant. In current literature it can also be explained by the falling demand for and the increased supply of low-skilled workers, as well as the weakening quality of schooling at lower educational levels (Colclough et al., 2009).

Differences in black and white population

Differences among population groups are prominent in South African economy, especially between the black and white people. Therefore, the effects of parental background and assortative mating might also differ between the black and the white. In this section, we only focus on black and white population (i.e. excluding coloured) and interact the white dummy with categorical levels of own education, father's education and father-in-law's education. All the interaction terms reveal the differential pattern of schooling, intergenerational transmission and assortative mating between the black and the white. As there is no observation of the white people who only obtain grade 1-7 in our sample and there are only 3 observations with grade 8, we aggregate educational levels into three categories: below matriculation, matriculation and college (and above). Categorisation of parental education is also adjusted according to the sample definition. The new categories are: below 10 years of schooling, 10 -11 years, matriculation and college. The results are in table 4.1.

In table 4.1, the interaction terms between the white dummy and own education (i.e. matriculation and college) are negatively significant, indicating that the returns to own schooling are larger among the black than the white, as is shown in column 1. This might indicate that the unobserved part of the earnings equation explains a larger part of the income of the white than that of the black. In column 2, the interaction terms between the white dummy and father's education are negative when year of father's schooling is 10-11 or matriculation, meaning father's education explains less of the returns to own education for the white. Comparing the gap in parental background between column 2 and 4, we find that the coefficient of father's education decreases more among the black after including own education in the regression, further confirming that parental background is less orthogonal to own education among the black than the white.

At the same time, column 3 shows that the role of fathers-in-law is not significantly different between the black and the white, as the interaction terms between the white dummy and father-in-law's education are not significant. Although the pattern gets a bit more clear in column 6, it is still consistent with the theoretical prediction that the racial difference in the effect of father-in-law's background is ambiguous because the degree of assortative mating is lower and the gender gap in labour force participation is higher among the white than the black, which drives the importance of father-in-law's education in opposite directions. The fact that the male-female ratio of the variance in income is larger among the white than that among the black is reflected in table 4.2, which presents the male-female ratio of standard deviation in labour market earnings among the overall population, black and white.

3.5.3 Robustness check

Given the possible racial difference in education, intergenerational transmission and assortative mating, as robustness checks, we first add more race dummies to see if the regression results are still robust.

As there are four population groups in total, we add dummies on whether one is coloured, Indian or white, respectively, to the main regressions. The results are in table 5.1 for continuous level of education and 5.2 for categorical eduction. In both tables, the coefficients of education are slightly smaller than the

corresponding coefficients in tables 2 and 3. However, the patterns of intergenerational transmission and assortative mating remain the same. Comparing columns 2 and 3, we find that the coefficients of fathers-in-law are larger than those of fathers, not only in continuous educational levels but also in all educational categories. Columns 4 and 5 show that after controlling for own education, the coefficients of father's education decrease more than those of father-in-law's. There is also convexity in the returns to education in these regressions.

In table 6.1 and 6.2 we also add two variables on location and replicate the regression results in table 2 and 3. We first introduce a variable "district" to capture the district council of one's current residence (53 district councils in total). "Urban" is a dummy variable on whether one lives in urban or rural area. Comparing table 6.1 and 2, we find that the returns to both own education and family background are smaller when we control for location and urban/rural division. For example, in column 1 in both tables, the returns to own education decrease from 0.197 in table 2 to 0.189 in table 6.1. However, the two key conclusions remain the same: father-in-law's education has more explanatory power than father's education and the information captured by father's education has more overlap with own education.

3.5.4 Other labour market outcomes: employment and labour market participation

South African labour market is characterised by high unemployment rate and large temporary transition into and out of employment. Therefore, parental background may affect not only labour market income but also employment rate and participation rate. Furthermore, as the estimates of the intergenerational correlation of earnings can be sensitive to sample selection rules (Couch and Lillard, 1998), it would be indicative to include unemployed or inactive people in our sample to see how parental background affects their labour market outcomes.

Table 7 replicates regressions in table 3 but expands the sample to all people who are active in the labour market, whether employed or not. In these regressions, we replace the labour market income with the probability of being employed as the dependent variable. We use Probit models instead of linear regression models and report marginal effects in all columns.

The result shows that having a college education increases the chance of being employed by 10.8% compared with being illiterate, as is presented in column 1. In consistent with the previous results, father-in-law's background has more explanatory power than father's education at the high school and college level. According to the comparison of column 2 and 3, father-in-law's education is significantly related to employment probability only when father-in-law finishes high school or has a college degree. However, in column 4 and 5, all the effects of parental background is absorbed by the inclusion of own education, indicating that father's and father-in-law's education affects employment probability almost completely through the channel of own education.

In table 8 we replace probability of employment with labour force participation and include all the male respondents in our sample. We create a dummy variable which equals one if one is economically active in the labour market, whether employed or not. The rest of the regressions and datasets remains the same. Table 8 indicates that both father's and father-in-law's education has explanatory power in the decision of labour force participation, and the comparison between column 2 and 3, and column 4 and 5 indicates that father-in-law's education plays a more important role in determining male labour force participation than father's background, especially at high school and college level.

In total, parental background and martial sorting not only explain labour market earnings but also relate to other labour market outcomes such as probability of employment and labour force participation. These two measures might be more indicative in reflecting South African labour market than earnings, because labour force participation is very low in our sample and many potential workers will be excluded from our

analysis if we only focus on people with positive labour market earnings.

3.6 Conclusion and discussion

Following Lam and Schoeni (1993), this paper further investigates the seemingly counterintuitive phenomenon that father-in-law's background has a stronger explanatory power than father's background for male wages by focusing on South African labour market. We extend Lam's paper in the following three ways. Firstly, we capture the heterogeneity of the effects of father and father-in-law's background at different educational levels and document its nonlinearity to explain where the convexity in returns to own education comes from. Secondly, we deal with measurement errors in education and the challenge of measuring permanent earnings which might lead to biased estimate of the importance of the parental education in the estimates of returns to own schooling. Thirdly, we extend Lam's conclusions to other developing countries such as South Africa, a country with more complicated structure of labour market and more serious unemployment issues.

After correcting for measurement errors in education and income, consistent with the current literature on South African labour market (Finn and Leibbrandt, 2016), we find high intergenerational transmission in socio-economic status in South Africa. More importantly, father-in-law's schooling has a higher effect on labour market earnings than own father's schooling, consistent with the result from Brazil. In addition, the drop in the coefficients on father's schooling is larger than that on father-in-law's schooling when own schooling is included in the earnings equation, and the drop in the coefficients on own schooling is smaller when father's schooling is included than when father-in-law's schooling is included. Both of these further confirm the larger importance of father-in-law's education relative to father's background.

We also find the convexity in the returns to own education declines after controlling for parental background but does not disappear. This implies that there is an endogenous selection of people with high ability into higher education but this selection does not fully account for the high returns to own education in South Africa.

We also notice the heterogeneity between the black and the white in response to intergenerational transmission and assortative mating in this segmented labour market. Our data suggests that assortative mating is lower and the male-female ratio of the variance in income is larger among the white than the black, which leads to the fact that parental background plays a more important role in labour market income among the black while this pattern does not exist for father-in-law's education.

	Assortative mating		Employment rate		Participation rate	
	Education	Average income	Male	Female	Male	Female
Unweighted						
All	0.72	0.1681	0.871	0.727	0.779	0.579
Black	0.682	0.4822	0.844	0.697	0.748	0.567
White	0.566	-0.0062	0.983	0.887	0.887	0.668
Weighted						
All	0.685	0.1113	0.899	0.748	0.828	0.649
Black	0.603	0.4854	0.856	0.69	0.787	0.632
White	0.525	-0.0028	0.981	0.845	0.906	0.696
Obs	676	676	1841	1424	2364	2461

Note: This table shows patterns of assortative mating and labour force participation. The degree of assortative mating is measured by the correlation between husband and wife in education and income across waves. The "Employed" columns show the probability of being employed given participation in the labour market. The "Participation" columns report the probability of labour force participation for all respondents between age 25 and 65. As we focus on the married sample and people who are employed or participate in the labour force in at least one of the waves, the employment and labour force participation rate might be higher than the calculation in other papers.

Table 1: Assortative mating and labour force participation in South Africa

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Edu	0.197***			0.177***	0.158***	0.155***
	(0.0116)			(0.0122)	(0.0138)	(0.0127)
Father edu		0.107***		0.0409***		0.0111
		(0.0121)		(0.0102)		(0.0157)
In-law edu			0.132***		0.0760***	0.0718***
			(0.0193)		(0.0217)	(0.0259)
Age	-0.00786	0.0317	0.0625	0.00946	0.0355	0.0379
	(0.0422)	(0.0469)	(0.0421)	(0.0428)	(0.0368)	(0.0376)
Age sq	0.000301	-0.000289	-0.000640	0.000129	-0.000168	-0.000189
	(0.000493)	(0.000548)	(0.000487)	(0.000498)	(0.000424)	(0.000432)
White	0.494***	0.326	0.212	0.241	0.0797	0.0337
	(0.185)	(0.205)	(0.275)	(0.183)	(0.269)	(0.231)
Svy_year	0.0475**	0.0903***	0.0872***	0.0494**	0.0510**	0.0513**
	(0.0220)	(0.0260)	(0.0240)	(0.0219)	(0.0219)	(0.0217)
Constant	-89.28**	-174.5***	-169.1***	-93.53**	-97.21**	-97.93**
	(43.85)	(51.97)	(48.01)	(43.63)	(43.82)	(43.36)
Observations	1,447	1,447	1,447	1,447	1,447	1,447
R-squared	0.410	0.250	0.319	0.422	0.455	0.456

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Earnings equation with continuous level of education

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Own						
Year 1-6	0.168			0.143	0.109	0.118
	(0.149)			(0.152)	(0.152)	(0.152)
Year 7-11	1.042***			0.931***	0.812***	0.809***
	(0.133)			(0.133)	(0.141)	(0.137)
Matric	1.465***			1.287***	1.160***	1.145***
	(0.209)			(0.199)	(0.244)	(0.222)
College	2.536***			2.280***	2.054***	2.030***
	(0.171)			(0.174)	(0.177)	(0.171)
Father						
Year 1-6		0.249*		-0.0346		-0.166
		(0.137)		(0.111)		(0.111)
Year 7-11		1.100***		0.514***		0.201
		(0.159)		(0.111)		(0.134)
Matric		1.128***		0.421		0.0628
		(0.275)		(0.267)		(0.345)
College		1.398***		0.532**		0.123
		(0.224)		(0.224)		(0.239)
Father-in-law						
Year 1-6			0.545***		0.348***	0.370***
			(0.157)		(0.131)	(0.139)
Year 7-11			1.053***		0.638***	0.575***
			(0.166)		(0.178)	(0.212)
Matric			1.819***		1.079***	1.002***
			(0.278)		(0.261)	(0.323)
College			1.620***		0.766**	0.686*
			(0.304)		(0.338)	(0.387)
Age	-0.00990	0.0236	0.0426	0.00474	0.0151	0.0146
	(0.0434)	(0.0469)	(0.0418)	(0.0435)	(0.0380)	(0.0393)
Age sq	0.000321	-0.000224	-0.000406	0.000166	6.30e-05	6.47e-05
	(0.000499)	(0.000545)	(0.000491)	(0.000495)	(0.000437)	(0.000449)
White	0.404**	0.392**	0.246	0.159	0.0396	0.00111
	(0.193)	(0.188)	(0.273)	(0.179)	(0.251)	(0.207)
Svy_year	0.0500**	0.0818***	0.0824***	0.0443**	0.0486**	0.0434**
	(0.0226)	(0.0260)	(0.0240)	(0.0221)	(0.0219)	(0.0216)
Constant	-93.53**	-157.1***	-158.8***	-82.57*	-91.35**	-80.86*
	(45.04)	(51.84)	(48.12)	(43.87)	(43.69)	(43.02)
Observations	1,447	1,447	1,447	1,447	1,447	1,447
R-squared	0.419	0.252	0.318	0.437	0.467	0.471

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Earnings equation with categorical levels of education

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Own						
White * matric	-1.382**			-1.419**	-1.313***	-1.321***
	(0.588)			(0.571)	(0.479)	(0.460)
White * college	-0.906***			-0.842**	-0.871***	-0.796**
	(0.342)			(0.331)	(0.317)	(0.310)
Matric	0.794***			0.765***	0.806***	0.775***
	(0.112)			(0.112)	(0.109)	(0.110)
College	1.609***			1.575***	1.562***	1.548***
	(0.110)			(0.106)	(0.110)	(0.109)
Father						
White * 10-11		-0.795**		-0.878**		-0.921*
		(0.400)		(0.401)		(0.479)
White * matric		-0.941*		-0.797		-0.575
		(0.504)		(0.494)		(0.562)
White * college		-0.546		0.154		-0.0209
		(0.392)		(0.376)		(0.490)
Year 10-11		0.750***		0.390		0.378
		(0.286)		(0.241)		(0.246)
Matric		0.799***		0.408***		0.216
		(0.237)		(0.132)		(0.160)
College		0.849***		-0.0929		-0.156
		(0.280)		(0.254)		(0.231)
Father-in-law						
White * 10-11			0.959		1.173	1.147*
			(0.799)		(0.734)	(0.686)
White * matric			0.388		0.369	0.356
			(0.803)		(0.673)	(0.682)
White * college			0.483		0.699	0.684
			(0.828)		(0.682)	(0.688)
Year 10-11			0.420**		0.0724	0.0886
			(0.205)		(0.207)	(0.200)
Matric			1.009***		0.669***	0.613**
			(0.225)		(0.230)	(0.271)
College			0.801**		-0.0286	-0.00886
			(0.323)		(0.296)	(0.273)
Age	0.00490	0.0247	0.0523	0.0182	0.0197	0.0306
	(0.0449)	(0.0535)	(0.0473)	(0.0427)	(0.0415)	(0.0419)
Age sq	0.000103	-0.000269	-0.000563	-3.88e-05	-5.19e-05	-0.000178
	(0.000521)	(0.000620)	(0.000549)	(0.000495)	(0.000478)	(0.000482)
White	1.529***	1.414***	0.316	1.818***	0.755	1.100**
	(0.285)	(0.176)	(0.746)	(0.333)	(0.533)	(0.427)
Svy_year	0.0661**	0.0961***	0.101***	0.0664***	0.0660***	0.0674***
	(0.0258)	(0.0280)	(0.0280)	(0.0250)	(0.0252)	(0.0249)
Constant	-125.6**	-185.8***	-196.7***	-126.5**	-125.8**	-128.8***
	(51.51)	(55.77)	(56.11)	(49.84)	(50.62)	(49.89)
Observations	1,156	1,156	1,156	1,156	1,156	1,156
R-squared	0.387	0.237	0.268	0.398	0.418	0.424

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.1: Earnings equation with heterogeneous effects of parental background between black and white

	Male	Female	Male/female ratio
All	28942.42	10846.5	2.6683649
Black	8605.476	6025.592	1.4281544
White	45439.42	25388.78	1.7897441

Note: This table shows standard deviation of income in black and white population. Male and female groups are estimated separately. Data comes from NIDS wave 1 - 4.

Table 4.2: Standard deviation of income in black and white population

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Edu	0.185***			0.173***	0.156***	0.154***
	(0.0103)			(0.0116)	(0.0134)	(0.0125)
Father edu		0.0890***		0.0290***		0.00635
		(0.0110)		(0.0104)		(0.0152)
In-law edu			0.119***		0.0682***	0.0662**
			(0.0214)		(0.0243)	(0.0275)
Age	-0.00211	0.0335	0.0618	0.00937	0.0341	0.0356
	(0.0422)	(0.0466)	(0.0413)	(0.0428)	(0.0365)	(0.0374)
Age sq	0.000242	-0.000305	-0.000626	0.000128	-0.000151	-0.000165
	(0.000491)	(0.000543)	(0.000478)	(0.000498)	(0.000421)	(0.000430)
Coloured	0.312**	0.114	0.0565	0.240**	0.154	0.143
	(0.121)	(0.172)	(0.161)	(0.116)	(0.118)	(0.112)
Indian	0.756***	0.954***	0.649**	0.632***	0.396**	0.380**
	(0.170)	(0.280)	(0.261)	(0.181)	(0.198)	(0.188)
White	0.634***	0.557***	0.363	0.429**	0.195	0.163
	(0.185)	(0.197)	(0.307)	(0.182)	(0.306)	(0.265)
Svy_year	0.0460**	0.0874***	0.0855***	0.0477**	0.0499**	0.0502**
	(0.0221)	(0.0263)	(0.0244)	(0.0220)	(0.0223)	(0.0220)
Constant	-86.33*	-168.6***	-165.5***	-90.05**	-94.94**	-95.50**
	(44.01)	(52.49)	(48.82)	(43.84)	(44.52)	(43.90)
Observations	1,447	1,447	1,447	1,447	1,447	1,447
R-squared	0.427	0.272	0.329	0.432	0.460	0.460

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa, with race dummies. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.1: Robustness check: earnings equation with more race

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Own						
Year 1-6	0.188			0.167	0.121	0.132
	(0.163)			(0.159)	(0.157)	(0.155)
Year 7-11	1.022***			0.954***	0.815***	0.824***
	(0.146)			(0.140)	(0.146)	(0.140)
Matric	1.381***			1.280***	1.133***	1.138***
	(0.212)			(0.201)	(0.245)	(0.222)
College	2.417***			2.267***	2.044***	2.041***
	(0.172)			(0.178)	(0.180)	(0.173)
Father						
Year 1-6		0.243*		-0.0455		-0.175
		(0.139)		(0.115)		(0.113)
Year 7-11		0.881***		0.343***		0.113
		(0.138)		(0.109)		(0.123)
Matric		0.980***		0.300		0.0208
		(0.269)		(0.264)		(0.337)
College		1.225***		0.405*		0.0838
		(0.210)		(0.217)		(0.230)
Father-in-law						
Year 1-6			0.549***		0.352***	0.386***
			(0.158)		(0.132)	(0.137)
Year 7-11			0.986***		0.579***	0.553**
			(0.181)		(0.194)	(0.219)
Matric			1.627***		0.923***	0.895**
			(0.302)		(0.297)	(0.350)
College			1.439***		0.621*	0.592
			(0.332)		(0.350)	(0.399)
Age	-0.00176	0.0275	0.0431	0.00715	0.0153	0.0145
	(0.0423)	(0.0462)	(0.0414)	(0.0430)	(0.0378)	(0.0391)
Age sq	0.000238	-0.000257	-0.000410	0.000144	5.89e-05	6.54e-05
	(0.000485)	(0.000535)	(0.000485)	(0.000489)	(0.000434)	(0.000446)
Coloured	0.298**	0.0865	0.0502	0.207*	0.133	0.119
	(0.122)	(0.171)	(0.158)	(0.115)	(0.121)	(0.110)
Indian	0.871***	0.912***	0.653***	0.675***	0.520**	0.452**
	(0.178)	(0.267)	(0.250)	(0.187)	(0.214)	(0.219)
White	0.562***	0.592***	0.407	0.345**	0.191	0.145
	(0.189)	(0.181)	(0.309)	(0.176)	(0.283)	(0.231)
Svy_year	0.0479**	0.0822***	0.0807***	0.0447**	0.0474**	0.0434**
	(0.0227)	(0.0262)	(0.0243)	(0.0221)	(0.0221)	(0.0215)
Constant	-89.64**	-158.1***	-155.5***	-83.39*	-89.11**	-80.90*
	(45.06)	(52.31)	(48.78)	(43.89)	(44.10)	(42.88)
Observations	1,447	1,447	1,447	1,447	1,447	1,447
R-squared	0.440	0.271	0.327	0.448	0.473	0.475

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa, with race dummies. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.2: Robustness check: earnings equation with categorical education and more race

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Edu	0.189***			0.172***	0.154***	0.151***
	(0.0117)			(0.0123)	(0.0133)	(0.0122)
Father edu		0.0937***		0.0384***		0.00992
		(0.0122)		(0.0105)		(0.0158)
In-law edu			0.123***		0.0751***	0.0715***
			(0.0203)		(0.0222)	(0.0263)
Age	-0.00511	0.0279	0.0534	0.0102	0.0330	0.0351
	(0.0432)	(0.0467)	(0.0418)	(0.0436)	(0.0378)	(0.0386)
Age sq	0.000264	-0.000249	-0.000536	0.000114	-0.000144	-0.000163
	(0.000502)	(0.000543)	(0.000483)	(0.000505)	(0.000436)	(0.000445)
White	0.468**	0.291	0.154	0.235	0.0640	0.0234
	(0.187)	(0.202)	(0.278)	(0.180)	(0.270)	(0.232)
District	0.000117	0.000113	-6.92e-05	8.18e-05	-4.12e-05	-4.26e-05
	(0.000155)	(0.000177)	(0.000174)	(0.000147)	(0.000141)	(0.000140)
Urban	0.182*	0.480***	0.450***	0.148	0.139	0.132
	(0.101)	(0.109)	(0.125)	(0.0954)	(0.102)	(0.101)
Svy_year	0.0481**	0.0876***	0.0847***	0.0498**	0.0512**	0.0515**
	(0.0221)	(0.0257)	(0.0238)	(0.0220)	(0.0219)	(0.0217)
Constant	-90.67**	-169.2***	-163.9***	-94.35**	-97.60**	-98.21**
	(44.00)	(51.26)	(47.55)	(43.75)	(43.70)	(43.28)
Observations	1,446	1,446	1,446	1,446	1,446	1,446
R-squared	0.414	0.273	0.337	0.424	0.457	0.458

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa, with controls on district council and urban/rural division. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.1: Robustness check: earnings equation with regional variables

VARIABLES	[1] Income	[2] Income	[3] Income	[4] Income	[5] Income	[6] Income
Own						
Year 1-6	0.0981			0.0931	0.0654	0.0791
	(0.164)			(0.164)	(0.163)	(0.163)
Year 7-11	0.908***			0.839***	0.727***	0.734***
	(0.151)			(0.150)	(0.152)	(0.150)
Matric	1.291***			1.170***	1.052***	1.051***
	(0.218)			(0.211)	(0.242)	(0.226)
College	2.352***			2.161***	1.945***	1.935***
	(0.186)			(0.187)	(0.185)	(0.180)
Father						
Year 1-6		0.144		-0.0728		-0.193*
		(0.132)		(0.108)		(0.109)
Year 7-11		0.959***		0.463***		0.169
		(0.149)		(0.105)		(0.132)
Matric		0.956***		0.357		0.0159
		(0.273)		(0.270)		(0.343)
College		1.247***		0.481**		0.0856
		(0.247)		(0.237)		(0.243)
Father-in-law						
Year 1-6			0.469***		0.321**	0.354**
			(0.156)		(0.131)	(0.138)
Year 7-11			0.960***		0.610***	0.561***
			(0.176)		(0.182)	(0.214)
Matric			1.715***		1.048***	0.989***
			(0.286)		(0.266)	(0.325)
College			1.474***		0.713**	0.651*
			(0.318)		(0.339)	(0.387)
Age	-0.00719	0.0201	0.0339	0.00434	0.0119	0.0106
	(0.0432)	(0.0470)	(0.0417)	(0.0437)	(0.0384)	(0.0397)
Age sq	0.000281	-0.000187	-0.000308	0.000161	9.07e-05	0.000101
	(0.000494)	(0.000543)	(0.000488)	(0.000495)	(0.000442)	(0.000454)
White	0.360*	0.346*	0.180	0.145	0.0168	-0.0109
	(0.195)	(0.192)	(0.274)	(0.184)	(0.250)	(0.208)
District	0.000124	0.000106	-3.72e-05	7.65e-05	-1.30e-06	-9.56e-06
	(0.000181)	(0.000180)	(0.000169)	(0.000168)	(0.000151)	(0.000151)
Urban	0.290***	0.485***	0.469***	0.248***	0.237***	0.236***
	(0.0932)	(0.115)	(0.118)	(0.0884)	(0.0902)	(0.0858)
Svy_year	0.0509**	0.0791***	0.0800***	0.0449**	0.0490**	0.0436**
	(0.0228)	(0.0257)	(0.0236)	(0.0222)	(0.0218)	(0.0216)
Constant	-95.64**	-151.9***	-154.1***	-83.75*	-92.28**	-81.26*
	(45.44)	(51.30)	(47.36)	(44.16)	(43.66)	(43.01)
Observations	1,446	1,446	1,446	1,446	1,446	1,446
R-squared	0.428	0.275	0.337	0.443	0.472	0.476

Note: This table shows regression results of average income across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa, with controls on district council and urban/rural division. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

VARIABLES	[1] Employ	[2] Employ	[3] Employ	[4] Employ	[5] Employ	[6] Employ
Own						
Year 1-6	-0.0101			-0.00677	-0.00799	-0.00553
	(0.0383)			(0.0382)	(0.0378)	(0.0378)
Year 7-11	-0.0364			-0.0367	-0.0334	-0.0339
	(0.0380)			(0.0376)	(0.0377)	(0.0375)
Matric	0.0282			0.0313	0.0362	0.0392
	(0.0347)			(0.0356)	(0.0353)	(0.0359)
College	0.108***			0.110***	0.108**	0.113***
	(0.0388)			(0.0425)	(0.0421)	(0.0435)
Father						
Year 1-6		-0.0204		-0.0363		-0.0320
		(0.0335)		(0.0345)		(0.0343)
Year 7-11		0.0429		0.0202		0.0205
		(0.0293)		(0.0296)		(0.0300)
Matric		0.0328		-0.0153		-0.0199
		(0.0367)		(0.0342)		(0.0344)
College		0.0466		-0.000568		-0.0187
		(0.0369)		(0.0420)		(0.0377)
Father-in-law						
Year 1-6			-0.0108		-0.0151	-0.0103
			(0.0294)		(0.0288)	(0.0297)
Year 7-11			-0.0139		-0.0431	-0.0429
			(0.0266)		(0.0269)	(0.0265)
Matric			0.0789*		0.0451	0.0460
			(0.0412)		(0.0397)	(0.0384)
College			0.0884**		0.0516	0.0554
			(0.0383)		(0.0420)	(0.0373)
Age	0.00886	0.00828	0.00935	0.00896	0.0103	0.0103
	(0.00875)	(0.00927)	(0.00926)	(0.00882)	(0.00880)	(0.00885)
Age sq	-0.000106	-0.000115	-0.000128	-0.000107	-0.000120	-0.000120
	(0.000102)	(0.000103)	(0.000104)	(0.000102)	(0.000102)	(0.000102)
White	0.123*	0.142***	0.115**	0.125**	0.0884*	0.0992**
	(0.0632)	(0.0545)	(0.0519)	(0.0529)	(0.0515)	(0.0482)
Svy_year	0.00362	0.00606	0.00570	0.00357	0.00307	0.00316
	(0.00503)	(0.00547)	(0.00515)	(0.00502)	(0.00487)	(0.00489)
Observations	1,739	1,739	1,739	1,739	1,739	1,739

Note: This table shows regression results of the probability of being employed across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa. Marginal effects in Probit models are reported. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Probability of employment with categorical level of education

VARIABLES	[1] Patricipate	[2] Patricipate	[3] Patricipate	[4] Patricipate	[5] Patricipate	[6] Patricipate
Own						
Year 1-6	0.0217			0.0160	0.0163	0.0142
	(0.0311)			(0.0309)	(0.0310)	(0.0307)
Year 7-11	0.0108			0.00126	0.000141	-0.00339
	(0.0295)			(0.0294)	(0.0293)	(0.0293)
Matric	0.0594**			0.0442	0.0428	0.0374
	(0.0290)			(0.0298)	(0.0294)	(0.0299)
College	0.139***			0.102***	0.0958***	0.0794**
	(0.0353)			(0.0362)	(0.0348)	(0.0350)
Father						
Year 1-6		0.0966***		0.0811**		0.0716*
		(0.0357)		(0.0357)		(0.0367)
Year 7-11		0.0436		0.0216		-0.00574
		(0.0291)		(0.0305)		(0.0306)
Matric		0.113*		0.0779		0.0579
		(0.0637)		(0.0647)		(0.0634)
College		0.174***		0.130**		0.0908
		(0.0492)		(0.0519)		(0.0599)
Father-in-law						
Year 1-6			0.0352		0.0168	-0.000838
			(0.0413)		(0.0425)	(0.0421)
Year 7-11			0.0763***		0.0550*	0.0536**
			(0.0272)		(0.0284)	(0.0266)
Matric			0.123***		0.0903**	0.0783*
			(0.0418)		(0.0417)	(0.0436)
College			0.190***		0.154***	0.133***
			(0.0444)		(0.0445)	(0.0498)
Age	0.0422***	0.0441***	0.0440***	0.0437***	0.0440***	0.0442***
	(0.00725)	(0.00697)	(0.00722)	(0.00703)	(0.00729)	(0.00704)
Age sq	-0.000594***	-0.000622***	-0.000622***	-0.000605***	-0.000609***	-0.000608***
	(7.87e-05)	(7.49e-05)	(7.77e-05)	(7.64e-05)	(7.89e-05)	(7.65e-05)
White	0.0377	0.00619	-0.00117	-0.0235	-0.0332	-0.0692
	(0.0350)	(0.0446)	(0.0363)	(0.0461)	(0.0364)	(0.0441)
Svy_year	0.0133***	0.0156***	0.0144***	0.0135***	0.0126***	0.0128***
	(0.00437)	(0.00428)	(0.00411)	(0.00440)	(0.00421)	(0.00419)
Observations	2,364	2,364	2,364	2,364	2,364	2,364

Note: This table shows regression results of the labour force participation across waves on own education, father's education and father-in-law's education for all respondents aged 25 to 65 in South Africa. Marginal effects in Probit models are reported. Data comes from National Income Dynamics Study wave 1 - 4 (year 2008, 2010, 2012 and 2014). Variables of education are corrected for measurement errors. Observations for regressions are weighted using post-stratification weights. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Labour market participation with categorical level of education

Chapter 4

Higher Education Expansion and Intergenerational Mobility in Contemporary China

Peng Zhang

Abstract. Intergenerational occupational mobility is stimulated when children from different social classes end up in similar occupations. Whether or not they have similar occupational status depends not only on their level of education but also the occupational returns to education among these groups. Given there is already a convergence in educational achievements between children from different social classes in contemporary China, in this paper, I focus on their occupational returns to education. Occupational status is measured by the widely-accepted ISEI scaling system ranging from 16 to 90 points with large number indicating higher occupational status. I take advantage of an exogenous college expansion policy in 1999 as a natural experiment and find that one additional year of education increases the occupational status of their first job by 2.243 (2.774) points on average along the ISEI scale in OLS (IV) regressions. And children from upper-class families do not necessarily have higher returns to education than children from other social classes. The average occupational returns to education are higher for the most recent job than the first job, but the difference among social classes is still not significant.

Keywords: Intergenerational mobility; Occupational choice; Education; Contemporary China.

JEL classification: I24, J24, J62.

4.1 Introduction

Occupation is a robust and thorough indicator of individuals' socio-economic status. It has two advantages compared with income as a measurement. First, occupation is regarded as the most important indicator of overall socio-economic status in most sociology literature, which better captures not only economic achievements but also social status than pure economic measurements such as income (Ganzeboom, 1996). Second, measurement error in income data is a big problem and may result in severe bias. Furthermore, many income measures capture transitory income data which is highly affected by temporary fluctuations (Solon, 1992). The problem of measurement error is more severe in developing countries where income data from cross-sectional household surveys or short panels are not sufficient to support estimates of permanent income.

Research on intergenerational mobility has come a long way since the review in the Handbook of Labor Economics by Solon (1999), but as Black and Devereux (2011) state, there is little literature on occupational

mobility. Intergenerational occupational mobility has important implications. A high intergenerational persistence rate indicates that upper-class families can maintain their privilege by increasing children's chances of succeeding in the job market even if children from middle-class and working-class families have higher abilities (Banerjee and Newman, 1993), which leads to the mismatches in the labour market (Magruder, 2010) and the persistence in inequality over time. But evidence on intergenerational occupational mobility in developing countries is especially scarce. Current studies show that intergenerational occupational persistence is as high as 43% in India, due to the caste system (Hnatkovska et al., 2013). Similar persistence rates are also found in Indonesia, Nepal and Vietnam (Pakpahan et al., 2009).

China has a higher rate of intergenerational occupational mobility than most developing countries (Takenoshita, 2007; Chen, 2012). This has been partly attributable to its merit-based educational system. Particularly, the college entrance exam system provided equality of opportunity in tertiary education. China's job market has also become more merit-based since the 1978 reform as the role of human capital became more crucial in recent years. For example, a college education is a necessary criterion for all administrative positions (Bian, 2002).

Whether or not this educational system can be viewed as playing a key role in improving intergenerational mobility depends on two aspects: accessibility to education and returns to education. Following Kreidl et al. (2014), I use the term "occupational returns to education" to study the effects of education on occupational attainments. Occupational returns to education indicate the average additional occupational status one can achieve in response to an additional year of schooling. Therefore, if both accessibility to education and occupational returns to given education levels do not differ among children from different family backgrounds, this educational system can provide a meritocratic route for best children to become occupational elites.

Most of the existing studies pay attention to the first aspect to investigate whether access to education is related to family background. In this paper I focus on the second aspect and compare occupational returns to education across different socio-economic groups. My empirical analysis is based on the theoretical model developed by Becker (1981) and individual-level repeated cross-sectional data (waves 2003, 2005, 2008, 2010, 2012, 2013 and 2015) from the China General Social Survey (CGSS).

There are a number of difficulties in establishing causal links between education and occupational status. In this paper I follow the natural experiment approach. In particular, I take advantage of an unexpected policy change in China's higher education in 1999 as a natural experiment, which increased the college enrolment rate from 34% in 1998 to 48% in 1999. This empirical strategy has some similarities to several studies on intergenerational transmission using exogenous changes in educational policies which affected particular birth cohorts (Black et al., 2005; Carneiro et al., 2013; Oreopoulos et al., 2006; Currie and Moretti, 2003; Maurin and McNally, 2008; Chevalier, 2004). Unlike most of these papers which focus on compulsory schooling laws, the natural experiment in this paper is an exogenous change in higher education and is an intervention occurring at a later stage in education, which can play a more important role than primary education in determining people's chance of obtaining a privileged occupation.

I find that occupational returns to education are high in contemporary China. Occupational status is measured by the widely-accepted ISEI scaling system ranging from 16 to 90 points. Occupations of higher statuses receive higher scores. One additional year of education increases the occupational status of the first job by 2.243 (2.774) points on average along the ISEI scale in OLS (IV) regressions. However, children from upper-class families do not necessarily have higher returns to education than children from other social classes. The average occupational returns to education are even higher for the most recent job compared with the first job, but the difference among social classes is still not significant. Given that upper-class children have slightly higher educational levels than middle-class children and occupational returns to education are

concave, my result indicates that increasing educational levels can potentially stimulate the intergenerational mobility of middle-class children.

This paper mainly contributes to literature on intergenerational mobility in developing countries by providing an indirect way to capture intergenerational mobility and focusing on an aspect less extensively studied: occupational returns to education which can have indirect implications on intergenerational mobility. If occupational returns to education are the same regardless of family background, policies aiming at equalising educational attainments among different social classes (such as lowering the financial burden of disadvantaged families) can stimulate intergenerational mobility.

My second contribution lies in the identification of cross-partial effects of children's education and family background to see whether educational attainments and family resources are complements or substitutes in determining children's socio-economic status. This is particularly important given China's merit-based educational system because children who reach the minimum requirements in entrance exams can get promoted to higher education regardless of their family background. As a result, the comparison of children from different family backgrounds with similar levels of education helps us better understand the relative roles of family resources and own human capital.

Related literature The paper follows research on intergenerational occupational mobility and is theoretically inspired by Gary Becker's framework which explains the intergenerational transmission through education (Becker and Tomes, 1979; Becker, 1981). In his model, parents maximize their utility by choosing between their own consumption and investments in their offspring. He predicted that intergenerational mobility is affected by the propensity to invest in children, the degree of endowment transmission (family's caste, religion, race, culture, genes or reputation) and the ability to finance this investment (Becker, 1981).

Empirically, this paper relates to the literature studying how much educational attainments can explain intergenerational mobility (Bowles and Gintis, 2002; Botticini and Eckstein, 2006; Di Pietro and Urwin, 2003). Testing the causal impact of education on intergenerational transmission is the focus of recent literature[1]. The first approach takes advantage of unique datasets to compare biological and adoptive parents or identical and fraternal twins. It is shown that both pre- and post-birth factors contribute to intergenerational education transmission but post-birth factors are more important for sons' income (Bjorklund et al., 2006). This approach has the advantage that it disentangles the effects between genetic inheritance and parental socio-economic statuses. However, the observed correlation between twins should still be driven by parental characteristics which are not genetically transmitted and are potentially correlated to both education and job market performance.

Another approach in causal identification, which is more widely used and is followed by this paper, is based on exogenous shocks in the educational level as natural experiments (Black et al., 2005). Shocks or instrument variables include: changes in compulsory schooling laws (Carneiro et al., 2013; Oreopoulos et al., 2006), variation in schooling costs (Carneiro et al., 2013) and the availability of college in ones's community (Currie and Moretti, 2003).

The specific identification strategy in this paper is closest to those used by Maurin and McNally (2008) and Chevalier (2004). Maurin and McNally (2008) look at an exogenous change in college admission rate in France to identify parents' years of schooling. France was thrown into a state of turmoil in 1968 when normal examination procedures were abandoned and the pass rate for different qualifications increased enormously. As a result, a larger proportion of students were able to pursue more years of higher education than would otherwise have been possible. The instrumental variable in my paper is also based on an exogenous change in college admission rate. The major difference is that college expansion was a temporary shock which only happened in 1968 in France while the college expansion process has been continuously taking place in China

[1] A detailed discussion is in Black and Devereux (2011).

since 1999.

This expansion of higher education in China has brought about additional difficulties in identification strategies because it is hard to disentangle college expansion effects from cohort effects as the expansion has continued since 1999. Moreover, treatment effects of this natural experiment may be heterogenous in different cohorts. As a result, I follow Chevalier (2004) in constructing a group of instrumental variables by interacting the policy dummy with cohort trend.

The next section presents a conceptual framework explaining the role of education in intergenerational mobility. Section 3 discusses data and descriptive statistics of China's intergenerational occupational mobility and education expansion. Details of the empirical strategy and results are reported in sections 4 and 5. I will also discuss robustness checks, further checks which control for household fixed effects and alternative explanations in sections 6 before conclusions.

4.2 Conceptual Framework

Research on intergenerational transmission places education as the central mechanism through which advantages in human capital are passed from one generation to the next. The conceptual framework of this study follows this approach and modifies it by focusing on the cross-partial effects between education and parents' social classes. The framework is based on a simplified version of Becker and Tomes (1979) and Becker (1981).

4.2.1 Determinants of occupational status

Parents are classified into three social classes: upper class (h), middle class (m) and working class (l). The classification is based on their occupational status. Y_i^k indicates children's occupational status if their parents are from social class k, $k \in \{h, m, l\}$.

Following Becker and Tomes (1979) and Becker (1981), parents allocate resources between their own consumption and children's human capital investment E_i based on their social class k. Children get human capital investments (i.e. education) and gain their own occupational statuses. L_i refers to other determinants of children's occupational statuses including child's ability and family endowments. Becker (1981) stated that family endowments are "determined by the reputation and the connections of families, the contribution to the ability, race and other characteristics of children, or the learning, skills and other commodities acquired through belonging to a certain family". For example, parents can provide networks for job market candidates or impose their preference of certain jobs. Furthermore, children's labour market outcomes also depend on random factors. Based on the theory and the corresponding model specification in (Solon, 1999), a child i from social class k gains occupational status Y_i^k written as:

$$Y_i^k = \alpha_i + \phi^k E_i + L_i = \alpha_i + \phi^k E_i + \delta_k + \beta X_i + u_i, \ with \ k \in \{h, m, l\}. \tag{4.1}$$

The key parameter here is ϕ^k, the occupational returns to education. The relative magnitude of ϕ^k among different social classes determine whether education and family background are substitutable, complementary or separable. u_i is a random factor such as luck during job applications. Determinants of occupational status besides education (L_i) are decomposed into a social class fixed effect δ_k and some proxies for the strength of family endowments and individual characteristics which are included in X_i [2]. Variables in L_i serve as

[2] β is assumed to be constant among children from different social classes for simplicity. I also run regressions where β differs among different social groups but the magnitude and significance of core coefficients are quite similar.

control variables. A more rigorous control for L_i will be performed in robustness check where L_i is treated as a household fixed effect.

4.2.2 Social class, education and intergenerational mobility

Based on equation 4.1, a comprehensive framework which links social class, education and intergenerational mobility is provided in this section. Similar to the potential outcome framework, it estimates and compares ϕ^k among individuals from three discretised social classes. The dependent variable Y_i refers to the occupational status of individual i, which can be written as:

$$Y_i = Y_i^h D_{1i} + Y_i^l D_{2i} + Y_i^m (1 - D_{1i} - D_{2i}) \tag{4.2}$$

D_{1i} is a dummy variable which indicates whether an individual is from an upper-class family. D_{2i} is a dummy variable which indicates a working-class background. Middle class is the reference group. Y_i^k is represented explicitly by Y_i^h, Y_i^m and Y_i^l. Their corresponding functions are as follows:

$$Y_i^h = \alpha_i + (\phi_1 + \xi)E_i + \delta_1 + \beta X_i + u_i. \tag{4.3}$$

$$Y_i^m = \alpha_i + \xi E_i + \beta X_i + \epsilon_i. \tag{4.4}$$

$$Y_i^l = \alpha_i + (\phi_2 + \xi)E_i + \delta_2 + \beta X_i + v_i. \tag{4.5}$$

ϕ_1 (ϕ_2) is the difference in occupational returns to education between individuals from upper-class (working-class) and middle-class families. Class fixed effects δ_k equal to δ_1 and δ_2 for individuals with upper-class and working-class origins, respectively. As the reference group, middle class has $\delta_m = 0$.

We can link social class and education to occupational status by combining equations 4.2, 4.3, 4.4 and 4.5. The cross-partial effects of parents' social class and children's education on children's occupational status can be modeled as:

$$Y_i = \alpha_i + \phi_1 E_i \times D_{1i} + \phi_2 E_i \times D_{2i} + \xi E_i + \delta_1 D_{1i} + \delta_2 D_{2i} + \beta X_i + u_i. \tag{4.6}$$

This model specification has direct implications on intergenerational occupational mobility. Especially when educational achievements of children from middle-class and working-class families have increased, differences in occupational returns to education among individuals from different social class origins matter more than the level of education in determining intergenerational occupational mobility. If ϕ is higher among upper-class children, the reduction in educational inequality will make a less difference in intergenerational mobility.

Both ϕ_1 and ϕ_2 matter in intergenerational mobility but I focus on ϕ_1. The reason is that children from upper-class and middle-class families have had similar levels of education in recent years. However, there is still a large gap in educational achievement between them and working-class children, making ϕ_2 less indicative for intergenerational mobility. We therefore have the following predictions:

- If $\phi_1 \leq 0$ then $\phi^h \leq \phi^m$: increases in educational achievements among students from middle-class family could stimulate their intergenerational upward mobility.
- If $\phi_1 > 0$ then $\phi^h > \phi^m$; increases in educational achievements makes less difference in their intergenerational upward mobility.

4.3 Intergenerational Occupational Mobility in Contemporary China

4.3.1 Data

The primary source of data in this paper is individual-level data from the nationwide China General Social Survey (CGSS). This is a biannual (annual in 2005 and 2006) repeated cross-sectional database compiled by the Survey Research Center of the Hong Kong University of Science and Technology. The CGSS project targets civilian adults aged 18 and older. In accordance with the sampling process in China's fifth census in 2000, a national sample of 5,900 urban households was interviewed in the 2003-2006 phase, with modifications in the 2008-2015 wave due to changes in community development (Bian and Li, 2012). We use data in 2003, 2005, 2008, 2010, 2012, 2013 and 2015. Data in 2006 is omitted as the coding for occupational status is not consistent with the others and we do not use data in 2011 as the year of entrance to the labour market is missing. To rule out the confounders from the systematic rural-urban division in occupational structure and educational achievement, I focus only on urban samples in the following analysis. To further control for the influences of social chaos during the Cultural Revolution, I restrict the sample to adults born after 1969. Their educational achievements were not likely to be affected by the Cultural Revolution during 1966-1976. Furthermore, they all entered the job market after 1985 (at the age of 16) after the reform towards a marketised economy was launched. To make sure individuals in our sample finish their whole education, we restrict the sample to people aged above 25.

CGSS data has three advantages in capturing individuals' and parental occupational status. First, it partly deals with potential bias resulting from labour force participation because it collects information on the last job for currently unemployed respondents. As a result, I can approximate the occupational status of currently unemployed respondents by coding their last existing jobs, which makes my results less sensitive to selection into the labour market. Second, it provides retrospective information on the overall employment history of each individual, which makes it possible to separate their first jobs and current jobs as individuals may experience job mobility in their life. Third, the CGSS survey also contains detailed information on parents' occupation when each individual was 18 years old [3], which is just the time when parents and children make decisions on whether to pursue higher education or enter the job market.

This study focuses on the influence of fathers' occupational status on children's occupational choice in the empirical analysis. As is mentioned by Lin and Bian (1991), fathers' occupational status, compared with mothers' resources, is more crucial in determining children's job market performance in urban China where male-superior social norm still dominates. I ran similar regressions by including both parents' occupational status and found that the coefficients of mothers' background are not significant. I also use the highest occupational standing of either the father or the mother to impute the occupational status of the family to replace father's occupational status and found that the magnitude and significance of core variables do not change too much.

I use a pooled sample of both men and women to study children's occupational status. Although in some studies the samples were restricted to men because the status attainment process may be different across genders, it has also been proved in some studies that the analysis might be biased by restricting the sample to men because changes in the supply of and demand for female labor over time will also affect men's choice in the job market (Kreidl et al., 2014). This distortion is larger in China's context because male and female labor forces can be more competing in China than in other countries. For example, the manual sector is highly sex-segregated in the US, but female participation in this sector is not highly restricted in China partly because of the state protection in the pre-reform era where women even got involved in military production and manual work in heavy industries (Chen, 2012). The proportion of females who stay at home

[3] The 2005 wave asked questions on parents' occupation when the respondents were 14 years old.

for housework instead of participating in the labour force is less than 2% among Chinese women under 30 years old. Also I ran regressions by splitting sample into men and women and found that although the magnitude of core variables differ between men and women, the sign of these coefficients remain the same across genders.

Summary statistics of the national sample are presented in table A1. The gender ratio is more or less balanced. The average age is 35.33 years in 2015 wave, which is the age when respondents finish their education and are at an early stage of their occupation. In general there is a decline in the proportion of employments in state sectors across the four waves, which reflects the privatisation taking place from 1992. Years of education, especially the proportion of college graduates is increasing with time. Variations of these socio-economic variables across wave indicate that the province fixed effect and survey year fixed effects should be controlled in the empirical analysis.

Occupations in CGSS are classified based on ISCO-88 (International Standard Classification of Occupations 1988) from the year 2008 and CSCO (Chinese Standard Classification of Occupations) in the 2003 and 2005 waves. I convert CSCO to ISCO-88 classifications to make the measurements in each wave comparable [4]. ISCO-88 is a hierarchical four-digit system of nested classification of occupations based on skill requirements (Ganzeboom, 1996). For example, 1120 stands for "Senior government officials", 2110 for "Physicist, chemist and related professionals" and 8111 for "Mining plant operators" [5].

I then convert ISCO-88 codes to the Socio-Economic Index of Occupational Status (ISEI) to map each occupation into its occupational standing [6]. ISEI is an optimal scaling of occupations which ranks occupations according to their skill levels and income status. More precisely, it is a ranking of attributes of different occupations based on their potential of converting individuals' educational attainment to expected earnings (Ganzeboom, 1996). ISEI scores are continuous measurements of occupational status under occupational titles.

One advantage of ISEI scaling is that ISEI was developed without interference from any criterion which is external to the process of stratification itself. As a result, although ISEI was first created to study occupational stratification in the US, it can also be applied to China's context even if occupational prestige might be different between these two countries. Thus it is arguably the best available international standard ranking of occupations.

As is pointed out in current literature (Kreidl et al., 2014), another advantage of ISEI ranking system is that it has a relatively stable distribution across contexts and is robust to changes in the distribution of occupations as long as the underlying stratification principles remain the same. As a result, differences in occupational returns to education are less likely to be associated with differences in the distribution of occupations among different social classes. Thus the potential changes in occupational structure since 1978

[4]The code is provided by China Family Panel Studies. The codebook is available at http://www.isss.edu.cn/cfps/sj/data2010/2013-07-11/180.html. One potential problem with ISCO-88 is that there might have been changes in occupations since 1978. For example, new occupations have emerged after the process of marketisation and globalisation. To prove that this does not severely affect my results, I suppose that the emerging occupations which are not included or cannot be integrated into the existing categories are classified as the category "hard to classify" in ISCO-88. In the online version of the paper I show that the proportion of respondents whose occupations are "hard to classify" in each year is not increasing over time. I also control for year of entrance in the labour market in regressions to compare cohorts looking for jobs at the same year.

[5]According to the ISCO88 manual by International Labour Organisation in 1990, virtually every occupation can be defined as a self-employed or salaried position. Thus, ISCO-88 does not acknowledge self-employment, ownership, and supervising status. Self-employers and small shop owners are classified with workers managing establishments on someone else's behalf. Members of the Armed Forces are excluded from the sample as working in the army is a temporary job.

[6]There are two additional scales used to measure socio-economic status of different occupations in sociology (Xie, 2012). Treiman's Standard International Occupational Prestige Scale (SIOPS) is largely based on prestige measures and reputation. This is not applicable in China's context as prestige of different occupations has changed a lot after the economic reform. For example, manufacturing workers were once highly respected but they lost this privilege after 1978. Erikson and Goldthorpe's class categories (EGP) map the ISCO-88 occupation categories into a discontinuous 10-category classification, which loses much information on the relative status of different occupations.

become less of an issue when ISCO-88 codes are converted to hierarchical ISEI scores.

Operational procedures for coding can be found in Ganzeboom (1996) and detailed comparisons of ISEI and ISCO-88 are in its appendix. ISEI scores are created by computing a weighted sum of socioeconomic characteristics of each occupation. The code of converting ISCO-88 to ISEI with some adaption to China's context is provided by the China Family Panel Studies[7]. The resulting set of scores ranges from 16 to 90, with Judges (ISCO-88 is 2422) gaining the highest score. The lowest score is jointly held by two unit groups: Agricultural, Fishery and Related Laborers (ISCO-88 is 9200) and Domestic Helpers and Cleaners (ISCO-88 is 9130). A higher ISEI score represents a higher occupational status[8]. The ISCO-88 and ISEI were applied to occupations in China by Deng and Treiman (1997) using 1982 Census data and the results were reasonable. They also made slight changes in CSCO before matching it with ISCO and ISEI, which are also adapted here.

However, a single continuous measurement of fathers' occupational status is not enough to generate a complete pattern of the effects of family background due to two reasons. First, the effects of fathers' occupational status may not change in a monotonic way with the increase of fathers' ISEI scores. For example, children from middle-class families are proved to be the most motivative in current literature (Maurin and McNally, 2008). Thus middle-class children might have higher occupational returns to education than children from both upper-class and working-class families. If this is true, occupational returns to education do not change monotonically along the ISEI scale. Second, the effects of an increase in fathers' occupational status may be heterogenous even if the effects of fathers' occupational status change in a generally monotonic way. Sometimes a slight increase in ISEI score may lead to fundamentally different effects but sometimes not. For example, although production clerks (ISEI is 43) and paper-products machine operator (IESI is 38) differ little in ISEI scores, they represent two fundamentally different occupations (i.e. white-collar and blue-collar jobs, respectively). Thus they may influence children's occupational achievements to very different degrees. On the contrary, senior government officials (ISEI is 68) and corporate managers in large enterprises (ISEI is 70) are similar both in ISEI scores and occupation categories. As a result, an adequate study of family background requires that fathers' occupational status should also be treated as a set of discrete categories.

Although research can benefit from a discretised measure of occupational status, the discretisation process requires careful consideration. In this paper I compared EGP scheme and ISEI to determine the threshold in ISEI scale for each social class. Details are in the online version of the paper. I define people with ISEI scores larger than 60 to be the upper class. People scoring between 40 and 60 in their ISEI are the middle class while the rest belong to the working class. This definition also has economic intuitions because 40 is the mean value of ISEI scores of service workers who are considered to be on the boundary between middle-class and working-class people while 60 is the mean value of ISEI scores of upper-level and middle-level professional workers who are regarded to be the watershed between upper class and middle class. In empirical analysis I will run regressions with both continuous measurements of father's occupational status (as robustness check) and their discretised social classes (as main analysis).

Summary statistics of continuous occupational status and discrete social classes based on ISEI scores are in table 1. The sample consists of 16% upper-class, 41% middle-class and 44% working-class people. ISEI scores of first occupations of the respondents range from 16 to 90. People from upper-class families have almost double the chance of getting an upper-class job than people from middle-class background. Gaps

[7]The codebook is available at http://www.isss.edu.cn/cfps/sj/data2010/2013-07-11/180.html.

[8]The skill-level distinctions embedded in the logic of ISCO88 are also reflected in the ISEI scale. For example, associate professionals average 16 points less (5 points more) than professionals (clerical workers). The manual/nonmanual divide (between clerical and skilled-crafts occupations) is 11 points. In the manual ranks, craft workers are only 3 points higher than machine operators, which lead elementary occupations by 11 points, according to Ganzeboom (1996).

in ISEI scores exist among people with different educational achievements. College graduates (including postgraduates) achieve 13 points more than senior school graduates and 18 points higher than people with junior degree. They are also the most likely to get upper-class occupations (31%). I also list the occupational status of people entering the job market in different years. ISEI scores increase with entrance years, from 38.4 points in the 1978-1986 cohort to 45.5 points in the post-1998 cohort, indicating that the economic development since 1978 has resulted in more opportunities for occupations of higher socio-economic status. ISEI scores of current occupations are higher than those of first occupations, on average.

4.3.2 Intergenerational occupational mobility in contemporary China

The time trend of intergenerational mobility in contemporary China can be obtained from figure 4.1 which presents the occupational status of the first job among individuals coming from different family backgrounds each year. As the year of entrance to the labour market goes by, the society has become more mobile as the average ISEI scores among children from both middle-class and working-class families have increased. And the gap in occupational status among respondents from upper and middle classes has been narrowing since 1985, especially after 2010 when the average ISEI scores are almost the same among children from these two social groups. This might be because after the marketisation since 1978, social hierarchy and network resources, which largely constrained intergenerational upward mobility of the middle-class and working-class people, have given way to credentials and ability (Yang, 2006; Bian, 2002).

There is also a convergence in educational achievements among children from different social backgrounds. Figure 4.2 reports the educational achievements of each cohort who enter the labour market after 1985 from different social backgrounds in the CGSS sample. To achieve universal compulsory primary education, China launched its Compulsory Education Law on July 1, 1986, which made 9 years of education (6 years of primary school plus 3 years of junior high school) compulsory for students in the entire country. According to the law, all the children at age six (or seven in some cases) should have the right and obligation to finish at least junior schooling regardless of gender, ethnicity and family background (Fang et al., 2012). Figure 4.2 reveals that years of schooling of students from all three social classes are above 10 among all cohorts, indicating that China has achieved the national goal of extending universal compulsory education among all the school-aged population.

4.4 Empirical Strategies

The empirical model directly follows equation 4.6 in the conceptual framework. Main regressions are based on a discretised measurement of fathers' occupational status.

4.4.1 Baseline regressions

Baseline analysis is the OLS regression of children's ISEI scores on family background and its interaction with educational attainment. One concern is that the time trend in China's economy may directly affect people's opportunities of being employed in upper-class occupations, regardless of the development of higher education. China has been growing rapidly since 1978, transforming from a planned economy to a market-oriented one after the enforcement of the reform and opening-up policy in 1978. Great changes took place in this post-reform period, resulting in a continuous decline in employment in manufacturing industries and a rise of tertiary industries. This means job opportunities increased continuously in upper-class and middle-class occupations (especially professional and service jobs) after the 1978 reform. Although all birth cohorts in my sample went to the job market after 1985, this time trend in the economy can still affect

individuals' occupational choices because it is possible that cohorts born later on had better chances to get upper-class jobs. Furthermore, job opportunities may also differ across regions. Metropolitan cities such as Beijing and Shanghai and economically advanced cities in East China may have encountered larger economic development.

To address these issues, I further consider cohort and spatial effects. Respondents enter the job market at the same year should experience the same occupational opportunities. As in the current literature (Chevalier, 2004), I include a quadratic function of year of entrance of each individual in all regressions to allow for the curvilinear relationship between year of entrance and occupational status. Each measurement of year of entrance is subtracted by 1985 in regression models. I also add dummies on province of residence to control for spatial patterns in the changing economy. Additionally, I control for calendar effects and trends in reported education and occupation, by controlling for the year each wave of the survey was conducted.

Furthermore, more socio-economically advanced parents tend to have children at an older age. Some aspects of parents' socio-economic status, which cannot be completely captured by fathers' occupations, can also directly affect children's job market performance. In the existing literature, fathers' years of birth matter as fathers born in more recent years tend to have higher socio-economic status as a result of the economic development. Fathers' party membership is also important in determining children's job hunting in China's context (Bian, 1997). To account for this, I also control for father's birth year and father's party membership in some of the regression models[9].

Fathers' occupational statuses are discretised into three social classes. As is suggested in Section 3, I also present models where father's occupational statuses are represented by a continuous measurement of ISEI scores in the Appendix. The baseline regression model is as follows[10]:

$$\begin{aligned} Occupation_i = &(school_i \times downfather_i)\theta + (school_i \times upfather_i)\beta \\ &+ downfather_i\lambda + upfather_i\delta + school_i\gamma + entrance_i + entrance_i^2 + X_i\omega + \alpha + \epsilon_i. \end{aligned} \tag{4.7}$$

where $downfather_i$ and $upfather_i$ are dummies on fathers' occupations (working-class and upper-class occupations, respectively). Respondents from middle-class families are treated as the reference group. $Occupation_i$ refers to individuals' occupational status (can be both first and most recent jobs). $school_i$ is a continuous variable of years of schooling for each respondent. $entrance_i$ and $entrance_i^2$ includes entrance cohort trend and forms a quadratic form of entrance year to control for the time trend. X_i includes gender, ethnicity, father's birth year and father's party membership. ϵ_i is the random error term for unobservable individual characteristics. I also include dummies of province of residence and survey year to further control for variations across regions and years of survey. The model predicts that if β is negative or zero, increase in educational achievements of middle class children can stimulate their intergenerational upward mobility.

[9]Other potential confounders such as father's education can be captured by father's occupational status.

[10]The baseline OLS regression with continuous measurements of father's occupational status is:

$$\begin{aligned} Occupation_i = &(school_i \times father's\ ISEI_i)\beta + father's\ ISEI_i\lambda + school_i\gamma \\ &+ entrance_i + entrance_i^2 + X_i\omega + \alpha + \epsilon_i. \end{aligned}$$

where $father's\ ISEI_i$ represents father's occupational status measured by continuous ISEI scores. I include dummies of province of residence and dummies of the survey year to further control for variations across regions and years of survey. The model predicts that if β is negative or zero, then increase in educational achievements of children from relatively disadvantaged families can stimulate their intergenerational upward mobility.

4.4.2 Natural experiment approach

Three conditions must be satisfied to get consistent estimates of occupational returns to schooling among individuals from different social classes:

$$Cov(u_i, E_i) = 0;\ Cov(u_i, E_i \times D_{1i}) = 0;\ Cov(u_i, E_i \times D_{2i}) = 0. \tag{4.8}$$

These conditions, however, may not be satisfied in simple cross-sectional analysis. Confounding factors come from the unobservables in the error term which may be correlated with education and occupational status at the same time. For example, the variable of years of schooling may suffer from endogeneity when unobservable individual characteristics such as ability can affect both education and occupational choice positively, as individuals with higher abilities tend to obtain more education and are more likely to have higher occupational status. It is thus unable to disentangle the effect of education from that of underlying abilities. The above OLS estimation may overestimate the effect of schooling and its interaction terms.

Natural experiments can help identify causal relationship. Instrumental variables based on natural experiments can deal with the remaining variations in L_i resulting from the unobserved confounders which cannot be controlled by D_{1i}, D_{2i} and X_i and potentially affect children's occupation and education at the same time[11]. In this part, I take advantage of an exogenous change in college admission rate to address the issue of endogeneity by using it's functional form as instrumental variables for the endogenous variable of education and its interaction terms with 2SLS estimates. I also calculated LIML estimates which are more robust to weak instruments in instrumental variable regressions.

The natural experiment comes from an unexpected increase in the college admission rate in 1999. In June 1999, the central government and the Ministry of Education increased the number of students admitted to tertiary education by 0.55 million. As a result, the number of new college students increased by 48% compared with that in 1998, which is the largest increase since 1978. In addition, the admission rate increased from 34% in 1998 to 48% in 1999 (Yeung, 2013). This made year 1999 a milestone in the history of China's higher education. More importantly, this expansion in 1999 was unexpected for high school graduates and their families as the announcements were made less than one month before the college entrance exams which were held in early July. The expansion policy can thus be regarded as an exogenous experiment because it did not have enough time to dramatically change the behavior of high school graduates. The number of new college students continued to increase in the following years, as is reflected in figure 4.3 in the CGSS sample.

The expansion in 1999 lowered the thresholds (minimum requirement of exam scores) for college admission, which enabled a proportion of students on the margin of the higher education system to pursue more years of higher education than would otherwise have been possible. However, access to higher education is still a competitive process after 1999 because all students still need to take the National College Entrance Examination, which is described as "thousands of troops crossing a single-log bridge" in the public media (Yeung, 2013). This means there is still a large variation in years of schooling among cohorts who participated in college entrance exams after 1999, which makes the instrumental variable analysis plausible.

I estimate the effect of an increase in opportunities of tertiary education by exploiting this variation in years of schooling caused by the college expansion policy. I construct a dummy variable $post_i$ on this policy change as the instrumental variable for years of schooling. That is to say, $post_i = 1$ if individual i was affected by this policy change. Consistent with Maurin and McNally (2008), the effect of college expansion policy on any given birth cohort is important if it is mainly made up of students at relevant stages of tertiary

[11] I do not use other identification strategies such as difference-in-difference or regression discontinuity design as it is impossible to identify treatment and control group for the former one in my data and there are no enough observations for the latter one.

education[12]. According to the compulsory schooling law in 1986, students are required to go to primary schools at 6 or 7 years old. After 6 years of primary education, 3 years of junior schooling and 3 years of senior school education, they are generally 18 or 19 years old when making the decisions on whether or not to attend college (Fang et al., 2012). It is plausible to assume that the expansion policy in 1999 primarily affected students born at or after 1979 who were at an early stage of higher education in 1999 (around 19 years old in 1999) while individuals were not potentially affected by the expansion policy if they were above 20 years old in 1999.

Thus $post_i = 1$ if a respondent i was less than 20 years old on the policy's effective date (i.e. i was born at or after 1979) and equals 0 otherwise. I rely on this binary variable to determine whether each individual was affected by the 1999 reform.

This binary variable captures the difference in educational achievements for cohorts before and after the 1999 reform. However, it cannot distinguish the variation in the time trend that is driven just by differences across birth cohorts - for example, the fraction of people who can achieve college degrees is increasing each year (cohort effect). As a result, a cohort trend should also be included in the first stage regressions.

Furthermore, this simple binary variable does not take into account the fact that the effect of the 1999 policy may be heterogeneous for post-reform cohorts. As is shown in figure 4.3, college admission rate kept increasing over time after 1999. As a result, cohorts who took college entrance exams later can benefit more from the expansion policy, which suggests that the treatment effect of this natural experiment is not homogenous for all post-reform cohorts. Variations in the probability of attending college after 1999 indicate that this reform shifted the educational attainment of all post-reform cohorts in a non-uniform way even after controlling for trends in education. To account for the heterogeneous treatment effects, I follow Chevalier (2004) and include both the binary variable $post_i$ and the interactions between $post_i$ and a trend of cohort of birth as instruments for each individual's education. The predicted educational attainment of each individual, rather than the observed one, is used to estimate occupational returns to education.

Both year of schooling and its interaction terms with parental social classes are potentially endogenous so that both should be instrumented to avoid the problem of "forbidden regressions". I also include interaction terms between the above two instrumental variables (i.e. $post_i$ and its interaction with a trend of the cohort of birth) and parental social classes in the whole set of instrumental variables. This idea comes from Bun and Harrison (2014) who proved that if z is a valid instrument for x and there is a dummy w, $z \times w$ is better than z as an instrument for $x \times w$. This is mainly because using only z is likely to suffer from underidentification problems.

In first-stage regressions, the instruments for the endogenous variables ($school_i$, $school_i \times downfather_i$, $school_i \times upfather_i$) are $post_i$, $post_i \times cohort_i$, $post_i \times upfather_i$, $post_i \times cohort_i \times upfather_i$, $post_i \times downfather_i$, $post_i \times cohort_i \times downfather_i$. The exogenous variables included are $entrance_i$, $entrance_i^2$, $upfather$, $downfather$ and X_i[13]. First-stage regressions are run separately for each of these three endogenous variables.

Concerns of the exogeneity of this IV approach arise from the following aspects. First, some high school graduates who were expected to take college entrance exams in 1998 may have been able to anticipate the expansion policy in the next year and thus postponed their exams to 1999. This is not very likely though due to the fact that the expansion policy was largely unanticipated. Second, students who failed the exams

[12] Variations in enrolment rates across province and time may be another potential IV. However, as is pointed out in Currie and Moretti (2003), enrolment reflects both the supply of college places and the demand for these places, which can be endogenous as well.

[13] X_i is a vector variable including gender, father's party membership and father's year of birth. In regressions where a continuous measurement of fathers' occupational status is used, the instruments for the endogenous variables ($school_i$ and $school_i \times father's\ ISEI_i$) are $post_i$, $post_i \times cohort_i$, $post_i \times father's\ ISEI_i$, $post_i \times cohort_i \times father's\ ISEI_i$. The exogenous variables included are $entrance_i$, $entrance_i^2$, $father's\ ISEI_i$, X_i.

in 1998 may retake the exams in 1999, which lowers the average ability of candidates in 1999. However, it does not affect the results because re-examination happens every year after 1978, which cannot explain the sharp gap in educational achievements between cohorts born before and after 1980.

4.5 Empirical Results

4.5.1 Baseline results

I consider two dependent variables in regression analysis: the ISEI scores of individuals' first occupations ($first\ ISEI$) and current occupations ($ISEI$)[14]. Both dependent variables are important. On the one hand, it is not common for children to get elite occupations (such as managers, senior government officers or professors) in their first job even if they get the best education or come from families at the top of the occupational status. However, they may obtain elite positions at a later stage of their careers. Thus, looking at first occupations may underestimate their potentials in upward mobility. On the other hand, those who have the motivation to change their jobs may systematically differ from those who stick to their first occupations. Focusing only on current jobs may suffer from potential selection bias. As a result, I consider both the first occupation and the current occupation in empirical analysis.

In the existing literature where wages are used as dependent variables, it is common that the percentage changes in wages rather than the absolute levels of earnings are computed to estimate economic returns to education. It is however not permitted in studies of occupational status because ISEI is an interval scale with no naturally occurring zero point Kreidl et al. (2014). Thus I interpret the occupational returns to education as changes in average occupational status caused by an additional year of education.

Table 2 reports the baseline models. β is the estimate of interest. The base model (column 1 and 3) do not include parental characteristics. Column 2 and 4 add father's year of birth and father's party membership. The significance of coefficients between these two models is similar while the effect of education is slightly smaller in magnitude when fathers' background information is controlled for. The full model indicates that having one additional year of schooling provides an advantage of 2.243 points in ISEI scores in the first occupation and 2.292 in most recent occupation. However, the interaction terms are not significant in all regressions and even negative in regressions regarding current occupations. This indicates that there is no strong evidence to show that occupational returns to education are different among upper-class and middle-class children in either first occupation or most recent occupation. In the robustness check, I will further consider two possible policies on labour market to show that structural changes in China's economy after the economic reform does not affect this result.

4.5.2 Natural experiment approach and test of the IV strategy

The models accounting for the endogeneity of children's education are identified by the 1999 college expansion policy as a natural experiment. A quadratic form of year-of-entrance cohort trend, province fixed effect, survey year fixed effect and all other exogenous variables which appear in OLS regressions are controlled in all first-stage regressions.

Table 3 presents first-stage results based on discrete measurements of fathers' social classes. As there are two dummies on social classes in the regressions (upper class and working class), there are six instrumental variables in total (i.e. the reform dummy, its interaction with a birth cohort trend and all interactions

[14]The chance of upward mobility in people's career history, which might be determined by family background, is also an interesting topic regarding occupational status attainment. Thus I also used a binary variable on whether one has ever switched occupational status as a dependent variable. I do not present the results here because there is too little variation to add any information in the current data set. This issue will be investigated more carefully in future studies.

between these two variables and two dummies on social classes). Compared with the first three columns with the last three ones, I find the inclusion of father's party membership and years of birth again lowers the value of F statistic of IV from 23.03 to 12.51). But the F statistics is still large. P-values of Hansen J statistics further confirms that the IV passes the overidentification test.

Table 4 reports two stage least square (2SLS) and limited-information maximum likelihood (LIML) regressions (to account for possible weak IV problem). An additional year of schooling increases the average ISEI scores of first occupations by 2.774 in LIML estimation for the first occupation (2.243 for the corresponding OLS regression) and 3.338 for the current occupation (2.292 for the corresponding OLS regression). It also suggests there is no significant evidence to show that people from upper-class families have higher occupational returns than people from other social classes. The interaction term between schooling and upper-class dummy is even negative in these IV regressions for both first and current occupations.

Table A2 shows the results where fathers' occupational statuses are measured by a continuous scale. The dependent variable is the ISEI score of the first occupation in columns 1 and 2 while columns 3 and 4 report regression results of current occupations. Father's characteristics are controlled for in all regressions. The estimates of years of schooling are still significantly positive and slightly larger than those in OLS regressions (2.502 versus 2.272 for the first occupation, and 3.348 versus 2.295 for the current occupation). In accordance with the results in OLS regressions, the interaction term between years of schooling and fathers' occupational status measured by ISEI scores is not significant and even negative for current occupation. The magnitude and significance of core coefficients are very similar between 2SLS and LIML estimators, which further confirms the instrumental variables are strong.

Comparing coefficients in table 4 between OLS and IV regressions indicates that the magnitudes are generally larger in IV regressions. As is indicated by Angrist and Imbens (1995) and Card (2001), a higher value in the IV specification is a typical finding in the literature estimating the returns to education. Card (2001) provides some explanations. Firstly, upward bias due to unobserved abilities in OLS estimates might be relatively small compared with their downward bias due to measurement errors. Thus IV estimates can be larger than OLS estimates if instrumental variables can correct for the measurement errors in OLS regressions. Secondly, 2SLS estimators based on quasi-experimental comparisons rely on grouped data because it can only identify two groups: the treatment group and the compare group. This grouping may amplify any inherent bias in OLS regressions by reducing the variance in the endogenous variable by more than it reduces the covariance of the endogenous variable with the bias terms. In this study I introduce not only a policy dummy but also the interaction terms between the policy dummy and a cohort trend to reduce this bias in IV estimates resulting from quasi-experimental comparisons. The third reason is that the effect of education varies across individuals and IV and OLS estimates capture different aspects of the distribution of this effect. IV estimates here mainly identify marginal children who would not have attended college had there not been the expansion policy. These students typically have higher marginal effects of schooling than the average as they treasure the chance of going to universities and are more motivated in finding upper-class jobs. The fact that IV estimates are generally larger than those coming from OLS regressions is consistent with a LATE interpretation.

This gives rise to a potential challenge in interpreting the results based on IV regressions. If IV regressions provide an estimate of the average effect on occupational choice of changes in years of schooling encountered by affected students in the after-1979 cohorts and the composition of marginal people who are affected by the IV is different among social classes, the significance and magnitude of the interaction terms between schooling and social classes may just capture different LATE based on different marginal population who are affected by the IV, instead of real difference in occupational returns to education among different social classes. Figure A1 shows this might not be a serious problem. It is very likely that the marginal children

who are affected by the college expansion policy are those high-school graduates who would otherwise not have been able to go to college had it been no college expansion. Therefore, I compare the proportion of people who have high school degree (on the left hand side) and above (on the right hand side) among different social classes. I find that especially in the years right before the college expansion policy (year of birth from 1976 to 1978), the proportion of high school graduates is quite similar between children from upper-class and middle-class families. This indicates that the composition of marginal children affected by the policy may not be largely different among social classes.

Furthermore, as the college expansion policy was implemented in a period when the One Child Policy was introduced, it is necessary to show that the increase in educational level after the college expansion is due to this change in education policy rather than the One Child Policy. The One Child Policy took legal effect in 1979 (Qian, 2008), resulting in a sharp decline in the number of children born after 1979 who reached the college attending age in 1998. This policy may be related to education and labour market in two ways. First, the number of participants in college entrance exams and job markets may largely decline after 1998 because of the One Child Policy, which can relax the competition in college entrance exams and in job hunting when they graduate; this can increase their chances of being admitted by universities as well as finding a good job. Second, as children have less siblings in their households, family resources can be more concentrated and they can be taken better care of by their parents, which increases their academic and job market performance at the same time. As the One Child Policy may affect children's schooling together with their occupational status, the instrumental variables will also suffer from the problem of endogeneity if the dummy $post_i$ mainly captures the effect of the One Child Policy rather than the 1999 college expansion policy.

I show the instrumental variable captures college expansion policy instead of One Child Policy in two ways. Firstly, according to Qian (2008), before 1979 when the One Child Policy was officially launched, family planning policies in China had already begun with a four-year birth spacing law in the early 1970s, which, together with the One Child Policy, affected cohorts born in and after 1976. Thus, the effective date of the One Child Policy does not coincide with the birth of the 1979 cohort who reached age 19 or 20 in 1999 when the college expansion policy was implemented.

Based on this evidence, I regress years of schooling on birth year dummy variables for all birth cohorts from 1970 to see if the One Child Policy can solely explain the increase in educational achievements of later birth cohorts in the sample. The corresponding regression is as follows:

$$school_i = \sum_{l=1970}^{1990} d_l \zeta_l + X_i \omega + \alpha + \epsilon_i. \tag{4.9}$$

where the d_l variables represent cohort dummies from 1970 to 1990, which are variables of interest. A quadratic function of entrance year is still controlled and father's characteristics, children's gender and children's ethnicity are included in X_i.

The results are in table 5.1. Columns 1 and 3 have no father's characteristics while columns 2 and 4 include all control variables. Both years of schooling and the probability of getting college education are regressed. The coefficients or marginal effects of the birth cohort between 1976 and 1978 are either significantly negative or not significant, which reveals that One Child Policy does not solely explain the increase in educational levels. The significant increase in educational levels starts from the 1979 birth cohort.

Secondly, as the minority group is not affected by the One Child Policy, if the increase in educational level for people born after 1979 is only due to the One Child Policy, we would expect that the increase in schooling is smaller among the minority group than the Han group. Therefore, I interact the instrumental

variable with a dummy on minority group and focus on the interaction term. I run regressions for both years of schooling and the probability of getting a college degree. In table 5.2, both of the interaction terms are far from being significant, which indicates that there is no significant evidence to prove that the One Child Policy is mainly responsible for the change in educational levels.

4.6 Robustness Check And Alternative Explanations

4.6.1 College degree and occupational status

Robustness checks are divided into three parts. In the first robustness check the variable of year of schooling is replaced by a dummy on whether one has a college degree.

The 1999 expansion policy directly changes the admission rate in higher education and as a result mainly affects students with college degrees. Thus, I conduct a robustness check by replacing the variable of years of schooling with a dummy variable on whether one has a college degree. The results are in table 6. Columns 1 and 2 show results for OLS regressions and 3-6 are results based on the natural experiment with the same set of instrumental variables as above. I only report the results based on the discrete measurement of fathers' occupational status but the results based on the continuous ISEI scale are also robust. Here both occupational statuses in first occupations and current occupations are regressed against interaction terms of a dummy on whether one has a college degree with dummies on working-class and upper-class family backgrounds. Instrumental variables also pass the weak IV test and overidentification test. In OLS regressions in column 1, having a college degree provides an advantage of 12.37 points in the ISEI scale of the first occupation (which becomes 22.87 in IV regressions in column 4). Again, there is no significant evidence to show that upper-class children have higher occupational returns to education and the interaction term between schooling and upper-class dummy is even negative. The results are similar regarding the most recent occupation rather than the first occupation.

4.6.2 Possible confounding policies

Although individuals in the sample were born in a short span of time, they may still encounter structural changes in the economy, which will also affect their opportunities in job hunting.

As a result, I conduct some robustness checks by adding dummies on policy changes which can account for possible structural changes in the economy. When looking at each policy change, a dummy on whether each individual was influenced by this policy is added into regression models. This policy dummy will also be interacted with all key variables (education and interaction terms between education and fathers' social classes) in regression analysis. Thus the interaction among education, upper-class family background and policy dummy ($school \times upfather \times policy$) should be the core variable in this robustness check. Policy change does not significantly affect the results if the coefficient of this variable is insignificant.

I consider two policies which took place in China after 1978. The first reform in the 1990s is the accelerated process of marketisation after 1992. The year 1993 is important in China's economic development as China has sped up the process of reform and opening-up since then. In March 1992, the authorities declared an end to the 1988 "rectification program" of centrally imposed price controls. In October 1992, the Communist Party formally endorsed Deng Xiaoping's view that the market system is compatible with socialism and called for the establishment of a "socialist market economy". Exchange rates were unified, a labor law was issued and reforms of state-owned enterprises were announced for 18 cities in 1993 (Jaggi et al., 1996). All this may have created new job opportunities for children who entered the job market after 1993, affecting their occupational choice.

The second reform in the 1990s relates to the urban job allocation system. From the mid-1950s to the late 1980s, a government program of job assignments was the core of the job allocation system in urban China where the central government controlled the size, growth and allocation of urban jobs mainly for college graduates. After graduating from college, youths waited for state job assignments and were restricted from switching between places of employment once assigned. Potential candidates for state job assignments were screened by authorities from residences, schools and recruiting organizations. The screening process was based on each applicant's demographic information, political and academic performance. Youths who got the job through state assignments were notified only of their place of employment and did not know their particular occupations until they checked in at the assigned work place (Bian, 1997). As a result, cohorts who came to the job market before the termination of this policy had little freedom in choosing their occupations and can thus be systematically different from those who looked for jobs after that. Reforms of this urban job allocation system began in 1992 when the central government encouraged more freedom in job hunting for college graduates. This planning system was cancelled in 1996 and was completely terminated in all provinces in 1998.

I consider these two policies one by one in table 7. Dummy variables *market* and *urban* refer to them respectively. These dummy variables equal 1 if individuals were affected by the policy change. I only report the interaction terms between the policy dummies and key variables related to occupational returns to education. OLS and IV regressions are conducted in terms of both first and current occupations. Although interpretations of the results of IV regressions should be cautious because F statistics of IVs are not very high, the general pattern remains quite consistent and the results are similar between 2SLS and LIML estimators. The coefficients of interaction terms between policy dummies, schooling and fathers' social classes are insignificant in all the six columns. This indicates that individuals do not behave fundamentally different after the accelerated process of marketisation and the abolishment of the urban job allocation system.

4.6.3 Family fixed effects

Another way to deal with the unobservables in L_i is to treat L_i as a household fixed effect and control it directly using sibling data. Here L_i can be assumed to be family-specific because genetic makeup, parents' attitudes towards occupation and their child-rearing practices differ across families [15].

As there is no sibling data in CGSS, we use an alternative data set, RUMiC 2008 (Rural Urban Migration in China, 2008). RUMiC 2008 sibling data provides information on education and occupational choices of siblings originating from the same household. However, occupations in RUMiC are roughly classified into 7 large categories instead of being coded into ISCO-88 system. I therefore replace the continuous measure of ISEI status with a dummy on whether one has an upper-class job and regress equation 4.1 for each social class separately with household fixed effects to get direct estimates of ϕ. The logit regression thus becomes[16]:

$$\begin{aligned} Pr(up_h^i = 1) &= f(\alpha^k + \phi^k E_h^i + L_h^i) \\ &= f(\alpha^k + \phi^k E_h^i + \mu_h + X_h^i\beta + T_h^i\eta + \epsilon_h^i),\ with\ k \in \{h, m, l\} \end{aligned} \tag{4.10}$$

[15] These factors have been the focus of literature in intergenerational mobility. A detailed discussion is in Black and Devereux (2011).

[16] One problem with the sibling data is that individuals with siblings may have family characteristics which are fundamentally different from those of families with only one child after the implementation of the One Child Policy. Thus focusing only on sibling data will lead to selection bias by cropping individuals in one-child families out of the whole sample. To deal with the potential selection bias, I first match between families with only one child and those with multiple children based on individuals' years of schooling, gender, age and fathers' characteristics. Then I run both linear probability and logit models with household fixed effects on the matched sample.

up_h^i is a binary variable on whether or not individual i from household h is employed in an upper-class job, instead of a continuous variable on ISEI scores. μ_h is the time-invariant unobservable family characteristics which can be controlled by household fixed effect. T_h^i indicates a quadratic form of age (to replace year of entrance). ϵ_{ih} is an individual-specific random variable which captures the remaining variation in unmeasured determinants of occupational choice among siblings.

Table 8 presents the results of occupational returns to education in the corresponding linear probability and logit models with household fixed effects. Coefficients of years of schooling are all positive and significant at the 1% level among all three social classes. Although the magnitude of coefficients here is not comparable with that in the previous regressions where the dependent variable is a continuous measurement of occupational status, we find that the estimated occupational returns to education are even larger in magnitude among children from middle-class families compared with children whose fathers have upper-class jobs. For example, in logit models, children in middle-class families get a 10.5% advantage in obtaining an upper-class job from one additional year of schooling, which is only 6.82% for children from upper-class families. The number of observations varies with family background but is generally small in all these regressions, which means cohorts born recently (after 1970) are largely exposed to the One Child Policy so that they are less likely to have siblings[17].

I also did other robustness checks. For example, replacing year of entrance with year of birth in the regressions as well as using more detailed discretisation to take into account the discrepancy in classifying father's occupations. The results are both robust and can be found in the online version of the paper.

4.6.4 Alternative explanations

1. **Educational levels are different.**

   One possible explanation for the different returns to education among different social classes might be that if the returns to education are not linear with the increase in educational levels, different schooling among social classes leads to different returns to education directly. In particular, if returns to education are concave and children from middle-class families have lower educational levels than upper-class children, it is natural that middle-class children will have relatively higher returns to education. Figure A2 shows that this is not the case. I draw the returns to education for children from all three social classes and find that the occupational returns to education are always convex. When middle-class children have slightly lower educational levels than upper-class ones, as indicated by data, they will have lower returns to education if the curve of human capital production function is the same between these two groups. But in reality the returns to education are not necessarily higher among upper-class than middle-class children.

2. **Compositional change in the ability distribution.**

   The education expansion policy allowed more high school graduates to go to college by lowering the thresholds of scores in college entrance exams, leading to a compositional change in the distribution of ability among higher education candidates. As the college entrance exam system selects students with the top scores, a higher admission rate may lead to a lower average ability of the college graduates, which means students who attended exams after 1999 may have lower abilities in general. As a result,

[17] Unobserved individual characteristics may still affect education as well as chances of intergenerational occupational mobility, despite controlling for family fixed effects. For example, individuals with higher ability may get both higher educational achievement and better occupations than their siblings. Thus, it is still useful to instrument for unobserved individual factors with the college expansion policy. However, the coefficient of the dummy variable in first-stage regressions is not significant for middle-class families (not reported here). Hence, these are not a valid group of instruments for education among middle-class children and I do not report IV regressions here. Possible reasons may be that the sample size is not large enough for birth cohorts after 1979 and the variations in their year of birth and educational attainment may not be large enough.

the observed decrease in the chance of attaining an upper-class job for some students may only result from the fact that candidates who attended college after 1999 had lower abilities on average. Thus, they were less likely to find an upper-class job, even if they finished college education. Using RUMiC2008 data on college entrance examination scores, I find this is not very likely (results are in the online version).

3. **Increase in tuition fees.**

   Under the central planning regime, higher education was heavily subsidized. These subsidies have been lowered by the Chinese government since 1993 when higher education institutions became increasingly decentralized both financially and administratively. From 1995 to 2004, tuition fees increased from 800 RMB per person per year to 5000 RMB per person per year, on average. Expenditures on education ranked first in total household expenditures in the early 2000s. Increasing tuition fees resulted in financial constraints for some families, which might be the reason why poor families had to give up college opportunities and might have benefited less from the expansion policy (Yeung, 2013).

   However, family's ability to finance college tuition may not contribute too much to the difference in educational achievements and occupational choice for students in this paper. First, my major focus is middle-class families which might be less likely to suffer from this financial burden, compared with working-class or lower-income families. Second, a comparison of educational achievements of children from upper-class and middle-class families reveals that the proportions of students who have college degrees are quite similar across these two groups after 1980. This indicates that children from middle-class families are not severely affected by the financial burden from the increase in tuition fees. Thus, the influence of tuition fees on college admission is limited when the question of interest in the above analysis is how occupational choice responds to an additional year of schooling.

4.7 Conclusions

Education plays a key role in intergenerational transmission (Bowles and Gintis, 2002). Models by Becker and Tomes (1979) and Becker (1981) place education as the central mechanism through which advantages are passed from one generation to the next. Following this approach, most of the current studies on education and intergenerational transmission do not explicitly study the occupational returns to education among different social groups. The past decades have witnessed a convergence in educational levels between children from upper-class and middle-class families in China. However, this convergence in education may not lead to a convergence in occupational status if occupational returns to given educational levels are higher among upper-class children.

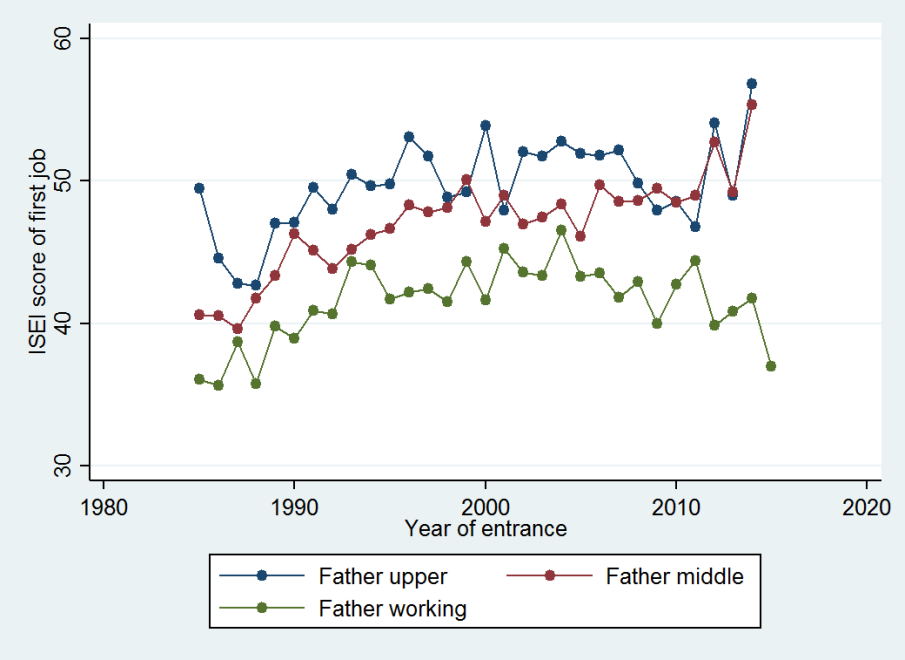
Motivated by this argument, I examine whether the occupational returns to education are different among social groups. Based on a college expansion policy in 1999, empirical results show that one additional year of education increases the occupational status of their first job by 2.243 (2.774) points on average along the ISEI scale in OLS (IV) regressions. And children from upper-class families do not necessarily have higher returns to education than children from other social classes. The average occupational returns to education are even higher for the most recent job compared with the first job, but the difference among social classes is still not significant.

Why don't children from upper-class families necessarily have higher returns to education? Normally there are higher possibilities of getting better jobs when children go to better schools (like the "Key" high school and "leading" universities in China's educational system). However, using data on school quality in CGSS, I find that the quality of education does not differ between these two groups, which might be one of

the answers to this question (results are in the online version).

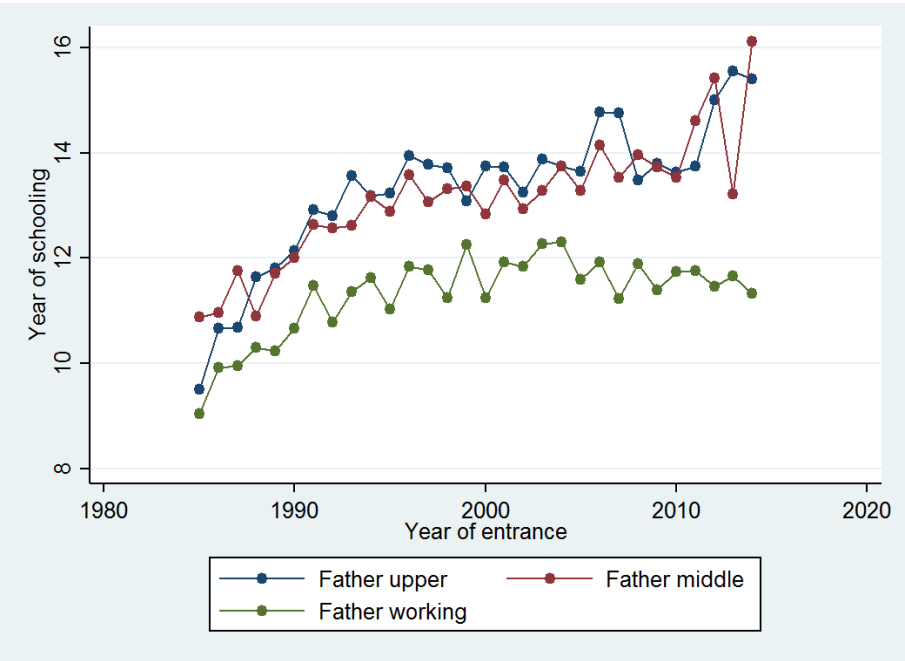
This is consistent with the most recent evidence in China which shows that education and family resources might not be complements (Emran and Sun, 2014). However, this is different from findings in other developing countries (Behrman et al., 2001), or evidence from developed countries such as Italy which points out that children's achievements still largely depend on parents' social status despite the establishment of an egalitarian education system (Di Pietro and Urwin, 2003).

This paper also sheds light on policies aiming at increasing intergenerational mobility by equalising educational attainment. It verifies the long-term belief that higher education is "a golden ticket" for people from less advantaged family background in China. As the occupational returns to education are not necessarily larger among upper-class children, policies aiming at equalising education, such as the 1999 college expansion policy, can not only increase the educational achievements among children from all different social groups, but also potentially motivate intergenerational occupational mobility.



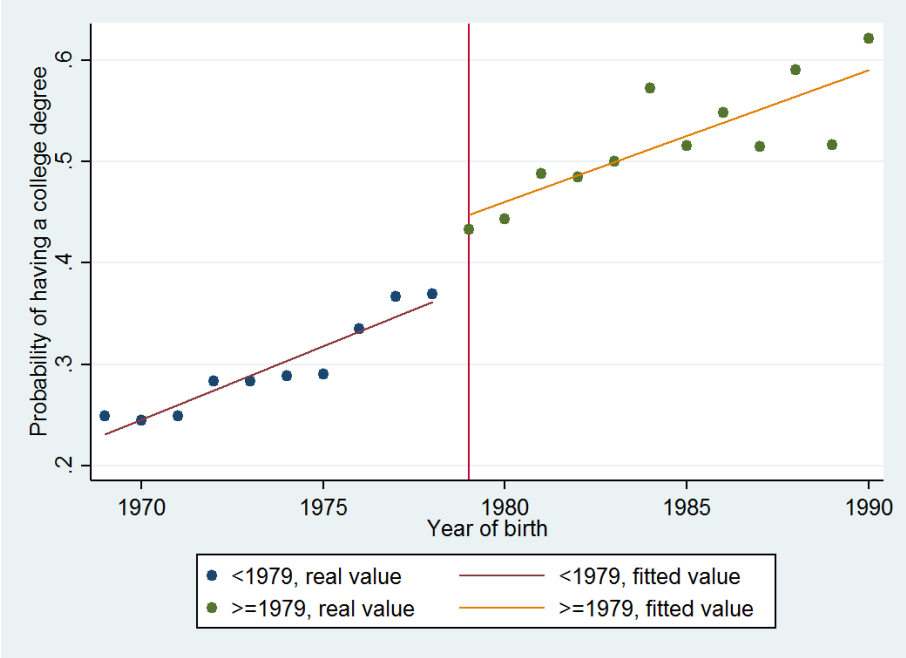
Notes: The figure plots the ISEI scores for children who enter the labour market after 1985. Children from upper-class, middle-class and working-class families are dealt with separately. Data source: CGSS 2003, 2005, 2008, 2010, 2012, 2013, 2015.

Figure 4.1: Occupational status among children from different family backgrounds over time



Notes: The picture captures years of schooling over years of entrance to the labour market. Children from upper-class, middle-class and working-class families are studied separately. Data source: CGSS 2003, 2005, 2008, 2010, 2012, 2013, 2015.

Figure 4.2: Years of schooling for each birth cohort



Notes: The picture captures the probability of getting a college degree for each birth cohort. Data source: CGSS 2003, 2005, 2008, 2010, 2012, 2013, 2015.

Figure 4.3: Probability of obtaining a college degree for each birth cohort

	Occu ISEI		First occu ISEI		First Upper		First Middle		First Down	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Overall	45.22	15.12	44.56	14.95	0.16	0.36	0.41	0.49	0.44	0.5
Father ISEI										
upper	51.21	16.07	50.24	16.34	0.27	0.44	0.45	0.5	0.29	0.45
middle	47.97	14.35	47.1	13.97	0.16	0.37	0.52	0.5	0.31	0.46
down	43	14.62	42.48	14.45	0.13	0.33	0.36	0.48	0.51	0.5
Education										
illiteratre	31.09	10.62	30.94	10.68	0.04	0.2	0.11	0.31	0.85	0.36
primary	33.47	9.92	32.63	8.57	0	0.05	0.16	0.37	0.84	0.37
junior	36.8	11.34	36.13	10.43	0.03	0.18	0.25	0.44	0.71	0.45
senior	41.83	12.86	40.79	12.24	0.06	0.25	0.41	0.49	0.53	0.5
college	54.04	14.15	53.71	14.33	0.31	0.46	0.53	0.5	0.16	0.37
Entrance year										
78-86	39.95	12.77	38.35	13.19	0.09	0.29	0.25	0.44	0.65	0.48
86-92	43.61	15.1	41.3	14.06	0.12	0.32	0.34	0.47	0.55	0.5
92-98	46.08	15.54	45	15.26	0.17	0.37	0.4	0.49	0.43	0.49
>98	45.62	14.91	45.46	14.92	0.17	0.37	0.42	0.49	0.42	0.49

Note: The table presents summary statistics of occupational status both for current occupations and first occupations in CGSS data 2003, 2005, 2008, 2010, 2012, 2013 and 2015. First occu ISEI refers to the ISEI scores for the first job. First upper, first middle, first down are dummies on whether one's first occupation is an upper-class, middle-class and working-class job.

Table 1: Summary statistics of occupational status

VARIABLES	[1] First ISEI	[2] First ISEI	[3] ISEI	[4] ISEI
school*upper class	0.142	0.152	-0.0418	-0.0385
	(0.203)	(0.204)	(0.208)	(0.209)
school*working class	0.110	0.114	0.0548	0.0569
	(0.133)	(0.133)	(0.136)	(0.137)
schoolyear	2.259***	2.243***	2.297***	2.292***
	(0.120)	(0.121)	(0.124)	(0.125)
upper class	-0.00862	-0.131	2.735	2.688
	(2.749)	(2.755)	(2.850)	(2.858)
working class	-2.745	-2.707	-1.822	-1.808
	(1.705)	(1.706)	(1.786)	(1.787)
female	1.704***	1.718***	1.723***	1.725***
	(0.323)	(0.322)	(0.328)	(0.328)
entrance year	0.236**	0.221**	-0.107	-0.109
	(0.0988)	(0.101)	(0.104)	(0.106)
entrance year2	-0.00897***	-0.00890***	0.000997	0.00101
	(0.00339)	(0.00340)	(0.00351)	(0.00352)
minority	-0.233	-0.223	-0.0839	-0.0827
	(0.709)	(0.709)	(0.714)	(0.714)
father party member		0.348		0.146
		(0.411)		(0.424)
father birthyear		0.0214		0.00410
		(0.0209)		(0.0214)
Province FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Constant	14.19***	-27.24	18.87***	10.90
	(1.840)	(40.44)	(1.994)	(41.47)
Observations	8,481	8,481	8,481	8,481
R-squared	0.303	0.303	0.289	0.289

Note: The dependent variables include individuals' ISEI scores of both first occupations (columns 1 and 2) and current occupations (columns 3 and 4). Fathers' occupational status is measure by two dummies of their social classes (upper class and working class). The base model (columns 1 and 3) only includes province fixed effects and survey year fixed effects. The second model (columns 2 and 4) adds parental characteristics as additional determinants, including father's year of birth and father's party membership. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: OLS regressions: father's social classes

	[1]	[2]	[3]	[4]	[5]	[6]
	No father's characteristics			Father's characteristics		
VARIABLES	school	school*upper class	school*working class	school	school*upper class	school*working class
Coefficients						
post	-0.874	-0.201**	-0.188	-0.554	-0.237***	0.094
	(0.796)	(0.081)	(0.244)	(0.769)	(0.086)	(0.254)
post*upper	-3.774***	-4.207***	-0.127	-3.708***	-4.197***	-0.105
	(1.152)	(0.843)	(0.304)	(1.129)	(0.842)	(0.321)
post*working	0.532	0.008	0.066	0.474	0.013	0.019
	(0.860)	(0.059)	(0.423)	(0.829)	(0.061)	(0.424)
post*cohort	0.138**	0.009	0.004	0.105*	0.015**	-0.028
	(0.055)	(0.006)	(0.018)	(0.054)	(0.007)	(0.020)
post*cohort*upper	0.266***	0.400***	0.005	0.253***	0.399***	-0.002
	(0.076)	(0.055)	(0.020)	(0.075)	(0.054)	(0.021)
post*cohort*working	-0.035	-0.001	0.096***	-0.038	-0.001	0.096***
	(0.058)	(0.004)	(0.028)	(0.055)	(0.004)	(0.028)
upper class	0.523***	12.959***	-0.001	0.527***	12.958***	0.005
	(0.145)	(0.111)	(0.036)	(0.145)	(0.111)	(0.039)
working class	-1.336***	0.004	11.092***	-1.025***	0.020	11.260***
	(0.116)	(0.011)	(0.066)	(0.117)	(0.017)	(0.071)
female	-0.443***	-0.025	-0.385***	-0.431***	-0.026	-0.377***
	(0.076)	(0.027)	(0.065)	(0.076)	(0.027)	(0.065)
entrance year	0.198***	0.055***	0.120***	0.173***	0.058***	0.096***
	(0.026)	(0.010)	(0.023)	(0.027)	(0.010)	(0.024)
entrance year2	-0.006***	-0.002***	-0.004***	-0.005***	-0.002***	-0.003***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
minority	-0.179	-0.024	-0.044	-0.166	-0.026	-0.031
	(0.182)	(0.061)	(0.153)	(0.177)	(0.061)	(0.152)
father party member				0.897***	0.039	0.497***
				(0.090)	(0.040)	(0.067)
father birthyear				0.028***	-0.004*	0.026***
				(0.006)	(0.002)	(0.005)
Province FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Constant	13.468***	-0.018	0.803***	-42.022***	7.336*	-50.398***
	(0.255)	(0.077)	(0.212)	(12.017)	(3.926)	(9.804)
Observations	8,481	8,481	8,481	8,481	8,481	8,481
R-squared	0.0262	0.0262	0.0262	0.0139	0.0139	0.0139
IV tests						
F statistics of IV	23.03	23.03	23.03	12.51	12.51	12.51
Hansen J statistics	1.303	1.303	1.303	1.326	1.326	1.326
Hansen J p-value	0.7283	0.7283	0.7283	0.7229	0.7229	0.7229

Note: The dependent variables are three endogenous variables: years of schooling and its interaction with fathers' occupational statuses which are measured by dummies of their social classes (upper class and working class). The base model (columns 1, 2 and 3) only includes province fixed effects and survey year fixed effects. The second model (columns 4, 5 and 6) adds parental characteristics as additional determinants, including father's year of birth and father's party membership. Instrumental variables include: policy dummy, its interactions with cohort trend and the interactions of the above two variables with dummies on fathers' social classes. Overidentification test is based on Hansen J statistics. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: First-stage regressions: father's social classes

VARIABLES	[1] 2SLS	[2] First ISEI LIML	[3] 2SLS	[4] LIML	[5] 2SLS	[6] ISEI LIML	[7] 2SLS	[8] LIML
schoolyear	2.772***	2.776***	2.767***	2.774***	2.985***	3.001***	3.293***	3.338***
	(0.557)	(0.560)	(0.645)	(0.654)	(0.569)	(0.578)	(0.672)	(0.695)
school*upper class	-0.560	-0.563	-0.564	-0.568	-0.661	-0.671	-0.708	-0.727
	(0.676)	(0.678)	(0.671)	(0.673)	(0.680)	(0.685)	(0.680)	(0.687)
school*working class	-0.056	-0.057	-0.058	-0.058	-0.042	-0.044	0.020	0.021
	(0.532)	(0.534)	(0.526)	(0.527)	(0.542)	(0.547)	(0.543)	(0.548)
upper class	9.212	9.249	9.266	9.313	10.746	10.863	11.220	11.447
	(8.924)	(8.948)	(8.839)	(8.868)	(8.975)	(9.040)	(8.950)	(9.036)
working class	-0.128	-0.112	-0.108	-0.096	0.228	0.276	-0.279	-0.246
	(6.808)	(6.831)	(6.659)	(6.682)	(6.961)	(7.025)	(6.899)	(6.970)
female	1.863***	1.865***	1.862***	1.865***	2.009***	2.016***	2.151***	2.172***
	(0.358)	(0.359)	(0.390)	(0.393)	(0.366)	(0.368)	(0.407)	(0.415)
entrance year	0.185	0.185	0.185	0.184	-0.201	-0.203	-0.227*	-0.233*
	(0.119)	(0.119)	(0.124)	(0.125)	(0.125)	(0.126)	(0.132)	(0.134)
entrance year2	-0.008**	-0.008**	-0.008**	-0.008**	0.003	0.003	0.003	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
minority	-0.165	-0.164	-0.165	-0.164	0.023	0.025	0.055	0.062
	(0.719)	(0.719)	(0.720)	(0.721)	(0.730)	(0.731)	(0.744)	(0.747)
father party member			0.012	0.006			-0.691	-0.732
			(0.603)	(0.610)			(0.630)	(0.648)
father birthyear			0.002	0.002			-0.047	-0.049
			(0.035)	(0.035)			(0.036)	(0.038)
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Constant	7.344	7.289	3.303	3.920	9.624	9.406	96.099	100.283
	(7.393)	(7.438)	(62.762)	(63.446)	(7.568)	(7.689)	(65.816)	(67.737)
Observations	8,481	8,481	8,481	8,481	8,481	8,481	8,481	8,481
R-squared	0.297	0.297	0.297	0.297	0.276	0.276	0.257	0.254

Note: The dependent variables include both first occupations and current occupations. Fathers' occupational statuses are measured by dummies on fathers' social classes. The base model (columns 1, 2, 5 and 6) only includes province fixed effects and survey year fixed effects. The second model (columns 3, 4, 7 and 8) adds parental characteristics as additional determinants, including father's year of birth and father's party membership. Instrumental variables include: policy dummy, its interactions with a cohort trend and the interactions of the above three variables with dummies on fathers' social classes. Both 2SLS and LIML are used. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: IV regressions: father's social classes

VARIABLES	[1] schoolyear	[2] schoolyear	[3] college	[4] college
1970	-0.431**	-0.383*	-0.0663**	-0.0620**
	(0.202)	(0.201)	(0.0310)	(0.0307)
1971	-0.616***	-0.552***	-0.0868***	-0.0799**
	(0.201)	(0.200)	(0.0322)	(0.0321)
1972	-0.167	-0.0279	-0.0196	-0.00380
	(0.209)	(0.207)	(0.0315)	(0.0313)
1973	-0.404*	-0.280	-0.0450	-0.0321
	(0.223)	(0.218)	(0.0325)	(0.0321)
1974	-0.107	0.0657	-0.0289	-0.0108
	(0.218)	(0.215)	(0.0335)	(0.0331)
1975	-0.459**	-0.277	-0.0814**	-0.0637*
	(0.220)	(0.213)	(0.0335)	(0.0327)
1976	-0.559**	-0.344	-0.0682*	-0.0466
	(0.242)	(0.236)	(0.0364)	(0.0353)
1977	-0.154	0.138	-0.0179	0.0147
	(0.250)	(0.245)	(0.0365)	(0.0354)
1978	0.0489	0.361	-0.0164	0.0171
	(0.251)	(0.242)	(0.0369)	(0.0357)
1979	0.255	0.561**	0.0197	0.0516
	(0.258)	(0.249)	(0.0387)	(0.0373)
1980	0.288	0.619**	0.0107	0.0444
	(0.261)	(0.245)	(0.0378)	(0.0356)
1981	0.649**	1.014***	0.0774*	0.114***
	(0.279)	(0.262)	(0.0407)	(0.0382)
1982	0.515*	0.888***	0.0358	0.0734**
	(0.265)	(0.248)	(0.0392)	(0.0367)
1983	0.738**	1.163***	0.0682*	0.111***
	(0.296)	(0.275)	(0.0414)	(0.0383)
1984	1.068***	1.530***	0.139***	0.186***
	(0.312)	(0.291)	(0.0441)	(0.0405)
1985	0.810***	1.336***	0.0521	0.107***
	(0.309)	(0.284)	(0.0443)	(0.0403)
1986	1.055***	1.549***	0.128**	0.178***
	(0.370)	(0.345)	(0.0515)	(0.0476)
1987	0.670*	1.231***	0.0468	0.105**
	(0.361)	(0.335)	(0.0533)	(0.0489)
1988	1.025**	1.615***	0.0699	0.131**
	(0.438)	(0.409)	(0.0578)	(0.0540)
1989	1.913***	2.527***	0.153**	0.216***
	(0.466)	(0.444)	(0.0763)	(0.0737)
1990	1.723***	2.369***	0.149**	0.214***
	(0.491)	(0.465)	(0.0727)	(0.0682)
Father control	NO	YES	NO	YES
Observations	8,481	8,481	8,478	8,478
R-squared	0.224	0.210		

Note: I regress years of schooling on birth year dummy variables for all birth cohorts from 1970 to 1990, including the cohorts between the implementations of the One Child Policy and the college expansion policy, to see if the One Child Policy can solely explain the increase in educational achievements of later birth cohorts in the sample. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.1: Test on the identification strategy: cohort effect

VARIABLES	[1] schoolyear	[2] college
college expansion * minority	-0.181	-0.0292
	(0.356)	(0.0472)
college expansion	0.810***	0.0915***
	(0.120)	(0.0167)
upper class	0.550***	0.0694***
	(0.122)	(0.0199)
working class	-1.073***	-0.131***
	(0.0956)	(0.0156)
female	-0.452***	-0.0480***
	(0.0758)	(0.0114)
entrance year	0.130***	0.0168***
	(0.0252)	(0.00356)
entrance year2	-0.00329***	-0.000353***
	(0.000869)	(0.000116)
minority	-0.0950	0.00285
	(0.210)	(0.0315)
father party member	0.910***	0.125***
	(0.0901)	(0.0136)
father birthyear	0.0358***	0.00367***
	(0.00584)	(0.000899)
Province FE	YES	YES
Year FE	YES	YES
Observations	8,481	8,478
R-squared	0.218	

Note: I interact years of schooling with minority dummy. As the minority are not affected by the One Child Policy, if the One Child Policy can solely explain the increase in educational achievements of later birth cohorts in the sample, we should expect the interaction term to be negatively significant. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.2: Test on the identification strategy: minority

	[1]	[2]	[3]	[4]	[5]	[6]
	First ISEI	ISEI	First ISEI		ISEI	
VARIABLES	OLS	OLS	2SLS	LIML	2SLS	LIML
college *upper class	1.498	-0.227	-2.168	-2.333	-2.576	-2.961
	(1.164)	(1.172)	(4.883)	(4.946)	(5.048)	(5.205)
college *working class	3.178***	2.596***	1.732	1.814	2.759	3.012
	(0.814)	(0.830)	(3.867)	(3.943)	(4.120)	(4.310)
college	12.37***	12.79***	22.183***	22.869***	26.511***	28.268***
	(0.699)	(0.714)	(4.877)	(5.210)	(5.221)	(5.884)
upper class	1.408*	2.653***	2.890	2.941	3.120	3.228
	(0.855)	(0.883)	(2.708)	(2.737)	(2.800)	(2.876)
working class	-3.322***	-2.876***	-1.479	-1.417	-1.102	-0.961
	(0.531)	(0.568)	(1.817)	(1.851)	(1.958)	(2.042)
female	1.388***	1.394***	1.828***	1.867***	2.115***	2.217***
	(0.325)	(0.331)	(0.413)	(0.428)	(0.446)	(0.480)
entrance year	0.316***	-0.0128	0.182	0.170	-0.231	-0.261*
	(0.102)	(0.107)	(0.130)	(0.134)	(0.142)	(0.152)
entrance year2	-0.0122***	-0.00236	-0.010***	-0.010***	0.001	0.001
	(0.00348)	(0.00360)	(0.004)	(0.004)	(0.004)	(0.004)
minority	-0.546	-0.393	-0.462	-0.458	-0.305	-0.295
	(0.693)	(0.697)	(0.714)	(0.720)	(0.764)	(0.787)
father party member	0.647	0.452	-0.449	-0.540	-1.262*	-1.496*
	(0.418)	(0.431)	(0.687)	(0.728)	(0.735)	(0.818)
father birthyear	0.0609***	0.0444**	0.012	0.008	-0.034	-0.045
	(0.0207)	(0.0213)	(0.034)	(0.036)	(0.036)	(0.040)
Province FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Constant	-80.56**	-43.44	8.709	16.552	102.069	122.508
	(40.14)	(41.13)	(64.054)	(67.638)	(69.206)	(76.556)
Observations	8,481	8,481	8,481	8,481	8,481	8,481
R-squared	0.288	0.272	0.224	0.213	0.109	0.061
IV tests						
F statistics of IV			9.316	9.316	9.316	9.316
Hansen J statistics			2.787	2.787	2.787	2.787
Hansen J p-value			0.4256	0.4256	0.4256	0.4256

Note: Dependent variables include both first and current occupations. Fathers' occupational statuses are measured by dummies on fathers' social classes. Both OLS and IV regression results are presented. Year of schooling is replaced by a dummy on whether one has a college degree. Instrumental variables include: policy dummy, its interactions with a cohort trend and the interactions of the above two variables with dummies on fathers' social classes. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Robustness check: using a dummy on college degree instead of years of schooling

VARIABLES	[1] First ISEI OLS	[2] ISEI OLS	[3] First ISEI 2SLS	[4] First ISEI LIML	[5] ISEI 2SLS	[6] ISEI LIML
Reform 1: accelerated marketisation						
school*upper*market	-0.0815	0.0171	0.555	0.559	1.073	1.076
	(0.105)	(0.105)	(0.736)	(0.740)	(0.805)	(0.808)
school*working*market	0.148*	0.268***	0.960	0.965	1.321	1.325
	(0.0806)	(0.0841)	(0.812)	(0.817)	(0.886)	(0.891)
school*market	0.317**	0.291**	1.372	1.373	1.637	1.637
	(0.143)	(0.143)	(1.210)	(1.215)	(1.267)	(1.271)
IV tests						
F statistics of IV			4.008	4.008	4.008	4.008
Hansen J statistics			0.841	0.841	0.841	0.841
Hansen J p-value			0.8396	0.8396	0.8396	0.8396
Reform 2: urban job allocation						
school*upper*urban	-0.0729	-0.107	0.278	0.291	0.438	0.468
	(0.0867)	(0.0853)	(0.354)	(0.367)	(0.373)	(0.399)
school*working*urban	-0.0193	0.0138	0.372	0.388	0.401	0.430
	(0.0667)	(0.0680)	(0.393)	(0.412)	(0.405)	(0.444)
school*urban	0.105	0.0789	0.259	0.253	0.613	0.618
	(0.116)	(0.117)	(0.807)	(0.829)	(0.826)	(0.870)
IV tests						
F statistics of IV			3.506	3.506	3.506	3.506
Hansen J statistics			2.681	2.681	2.681	2.681
Hansen J p-value			0.8477	0.8477	0.8477	0.8477
Control	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Note: Dependent variables include both first occupations and current occupations. Fathers' occupational statuses are measured by dummies on fathers' social classes (upper class and working class). Policy 1 refers to the accelerated marketization process after 1992. Policy 2 refers to the abolishment of urban job allocation system in 1996. Only interaction terms are reported. All the models include province fixed effects, survey year fixed effects and parental characteristics, including father's year of birth and father's party membership. Results of Weak IV and Overidentification test are also presented. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Robustness check: possible confounding policies

VARIABLES	[1] LPM Upper	[2] LPM Middle	[3] LPM Working	[4] Logit Upper	[5] Logit Middle	[6] Logit Working
schoolyear	0.0369***	0.0449***	0.0397***	0.0682***	0.105***	0.0781***
	(0.0132)	(0.0104)	(0.00606)	(0.0226)	(0.0172)	(0.00793)
age	-0.130	0.0669	0.935***	-0.305	0.0412	2.367***
	(0.145)	(0.163)	(0.0957)	(0.498)	(0.3923)	(0.117)
age_square	4.27e-05	-0.00261	-0.000259	-0.000442	-0.0101**	0.000893
	(0.00208)	(0.00218)	(0.00123)	(0.00545)	(0.00395)	(0.00196)
female	0.056	0.09*	-0.0686**	0.128	0.318***	-0.164***
	(0.0575)	(0.05)	(0.0318)	(0.0928)	(0.0702)	(0.0499)
Observations	134	131	437	134	131	437

Note: Regressions with household fixed effect are presented. The dependent variable is a dummy on upper-class job. Standard errors are in parentheses. Columns 1, 2 and 3 (or 4, 5, and 6) refer to individuals from upper-class, middle-class and working-class families. Data source: RUMiC 2008. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Regressions with household fixed effect on the matched sample

Appendix A

Appendix to Chapter 2

A.1 Appendix. Bantu migration and the formation of ethnic diversity from historical narratives

Below we provide a summary of the history of the Bantu migration from central and eastern Africa and the settlement of these groups in South Africa for each ethnic groups in details. Narrative evidence is summarised from Mwakikagile (2010) and Gradin (2014).

Ethnicity	Time of migration into SA	Homelands	Time of moving into white areas	Bantustan
Xhosa	Before 1400s	Today's Eastern Cape	After conflicts with the native Khoisan	Ciskei and Transkei
Zulu	16th century	Eastern part, today's Kwazulu-Natal	Early 1800s	KwaZulu
Swazi	15th and 16th centuries	Southern part of Tongaland in what is now Mozambique	17th and 18th centuries into the Pongola River	KaNgwane
Ndebele	Before 1835	Today's Northern Province, Mpumalanga and Gauteng	By 1835 towards Swaziland and Northern Transvaal	KwaNdebele, Lebowa
North Sotho	1500s	Today's Limpopo and Northwest	After the war with Boers and Ndebele	Qwawa
South Sotho	1500s	Today's Limpopo and Northwest	After the war with Boers and Ndebele	Qwawa
Tswana	1500s	Today's Limpopo and Northwest	After the war with Boers and Ndebele	Bophuthatswana
Tsonga	Before the early 1500s	Close to today's Mozambique	After conflicts with Zulu	Gazankulu
Venda	Before 800s A.D.	A mountainous area in the northern part close to Limpopo River	800s A.D. to Matopo Hills	Venda

A.2 Appendix. Data source and construction of district-level variables

In this section we present data sources and the construction of our district-level control variables in detail. Emphasis has been given on those geographical measures.

Variable	Data source	Construction of variable
Panel A: From census		
Area of the district	Census 1996 and 2001 district-level shape file.	Calculated from the shape file directly in ArcGIS.
Population density	Census 1996 and 2001.	Calcualte the total number of black in each district in census data and divide it by area.
Proportion of the black	Census 1996 and 2001.	Calculate the number of black over the whole population.
Proportion of manufacturing	Census 1996 and 2001.	Calculate the number of people working in manufacturing sector over the whole employed black people.
Proportion of service	Census 1996 and 2001.	Calculate the number of people working in service sector over the whole employed black people.
Urban/rural	Census 1996 and 2001.	Information on whether one lives in an urban or rural settlement is explicit in census data.

Variable	Data source	Construction of variable
Panel B: Sources on geography		
Overlap of district and homeland	A map (shape file) of homeland provided by Tim Brophy and Adrian Frith.	Intersecting the boundary of districts with that of homelands and seeing the overlap in ArcGIS.
River	Census 2001 river shape file.	Overlapping shape file of districts and river and directly calculating in ArcGIS.
Road	Census 2001 major road shape file.	Overlapping shape file of districts and road and directly calculating in ArcGIS.
Ruggedness	From Nunn and Puga (2012). We also tried the measure of slope from the same data source with similar results.	Same as Nunn and Puga (2012).
Soil quality	Harmonized World Soil Database. http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/.	Calculating average soil quality measures in a district (average of the index over grids in a district).
Density of mine	Mineral Resources Data System (MRDS) https://mrdata.usgs.gov/mrds/.	Overlapping shape file of districts and mines. Calculating number of mines in each district and dividing it by area.
Nighlight per capita	The National Oceanic and Atmospheric Administration night-time light satellite images. www.noaa.gov/stories/our-earth-at-night.	Calculating nightlight measures in a district (summation of the index over grids). Dividing it by the whole population in the district obtained from census data.
Distance from district to homeland	A map (shape file) of homeland provided by Tim Brophy and Adrian Frith.	Calculating Euclidean between centroid of districts and the border of homelands.
Distance to closest homeland	A map (shape file) of homeland provided by Tim Brophy and Adrian Frith.	Choosing the mininum value of the distance to all homelands.
Conflict	The Geo-referenced Event Dataset of the Uppsala Conflict Data Program (UCDP-GED v1.5) for 1996. The Armed Conflict Location and Event Data Project (ACLED) database for 2001.	Same as Amodio and Chiovelli (forthcoming).

A.3 Appendix. Explanation on how to draw data for simulation

In this Appendix we explain in more detail how to draw a series of $s_k, k = 1, 2, \ldots, m$ from a convoluted distribution of s under certain constraints in our simulation.

A.3.1 Hold the number of groups constant

Given a particular value of m, we just need to make sure $\sum_{k=1}^{m} s_k = 1$ when we draw these m different values of s_k.

Choose a particular value of m and hold it as a constant, we start by drawing $d_k, k = 1, 2, \ldots, m$ from a uniform distribution at any positive interval (here we use the interval [0,1]). Set $s_k = \frac{d_k}{\sum_{k=1}^{m} d_k}$. It is straightforward to prove that by this definition $\sum_{k=1}^{m} s_k = 1$. Also by this transformation s_k no longer follows uniform distribution, which satisfies our requirement that in the numerical simulation we need a convoluted density function of s. Therefore $s_k, k = 1, 2, \ldots, m$ is the series of our simulated data which represents each group's share over the whole population in the real data.

A.3.2 Hold the dispersion of group size constant

In this cases the value of $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$ has to be fixed. That is to say, we need to make sure $\sum_{k=1}^{m} s_k = 1$ and $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2 = T$ (T is a constant) when we draw these m different values of s_k.

Similarly, we start by drawing $d_k, k = 1, 2, \ldots, m$ from a uniform distribution at any positive interval (here we use the interval [0,1] again). Set $y_k = \frac{y_k}{\sum_{k=1}^{m} y_k}$.

Choose a particular value of $\sum_{k=1}^{m}(s_k - \frac{1}{m})^2$ (suppose it equals T) and hold it as a constant, we define $s_k = \frac{1}{m} + \frac{y_k - \frac{1}{m}}{\sqrt{\frac{\sum_{k=1}^{m}(y_k - \frac{1}{m})^2}{T}}}$. We can prove that $s_k, k = 1, 2, \ldots, m$ is the series of our simulated data which represents each group's share over the whole population in the real data.[1]

This is because:

$$\sum_{k=1}^{m} s_k = \sum_{k=1}^{m} \frac{1}{m} + \sum_{k=1}^{m}(y_k - \frac{1}{m}) \frac{1}{\sqrt{\frac{\sum_{k=1}^{m}(y_k - \frac{1}{m})^2}{T}}} = 1$$

$$\sum_{k=1}^{m}(s_k - \frac{1}{m})^2 = \sum_{k=1}^{m}(y_k - \frac{1}{m})^2 \frac{1}{\frac{\sum_{k=1}^{m}(y_k - \frac{1}{m})^2}{T}} = T$$

Again, although we start from the uniform distribution, after the transformation, the distribution of s_k becomes convoluted.

After making the draws, we conduct 100000 tests for each possible value of m and calculate the mean value of Y for each m (following the institutional setting, we choose $m = 2, 3, \ldots, 9$). Then we draw a figure of the mean value of Y over the corresponding m.

[1] One potential problem is that by this transformation y_k might be negative. In our simulation, with relatively large numbers of T this can occur in few occasions. As the number of tests is large enough, we just drop those test with at least one negative y_k in our simulated sample.

A.4 Appendix. Tables and figures

	Dependent variable: ethnic population N_{kd}		
	Coef.	Std. Err.	t-stat
Distance Dis_{kd}	-.0039	(.0007)	-5.17
Ethnic group fixed effects:			
Group 1	.9750	(.2139)	4.56
Group 2	.6133	(.1769)	3.47
Group 3	.1778	(.2248)	0.79
Group 4	-.4604	(.2311)	-1.99
Group 5	.2220	(.2259)	0.98
Group 6	.8940	(.1803)	4.96
Group 8	.0469	(.1833)	0.26
Group 9	-.8184	(.2776)	-2.95
Constant	9.157	(.2176)	42.08
R-squared	.092		
Observations	1989		

Note: This table reports results about the gravity model which helps estimate the stock of each ethnic group in each "white" district based on 1985 census data. The sample is for all the "white" magisterial districts which can be matched to 1996 and 2001 census. We control for homeland fixed effects and run a regression of the stock of ethnic groups on the distance between their corresponding homelands and each district using PPML models. *** p<0.01, ** p<0.05, * p<0.1.

Table A0: Gravity model predicting the stock of black population in white districts: PPML estimator

	[1]	[2]	[3]	[4]	[5]	[6]
	1996			2001		
	Uunemployed + inactive	Wage employee	Self/wage	Unemployed + inactive	Wage employee	Self/wage
Panel A: OLS estimates						
Ethnic fractionalisation ELF	-0.072***	0.073***	-0.002	-0.125***	0.127***	0.004
	(0.018)	(0.018)	(0.020)	(0.037)	(0.035)	(0.023)
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
Panel B: IV (GMM) estimates						
Ethnic fractionalisation ELF	-0.385**	0.378**	0.024	-0.193*	0.220*	-0.147
	(0.156)	(0.159)	(0.076)	(0.114)	(0.122)	(0.076)
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES

Note: This table reports results about the effect of ethnic diversity on employment and the allocation between self- and wage-employment at district-level regressions based on 1996 and 2001 census data. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features and individual-level controls. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. Dependent variables are the proportion of people in each employment status over the whole working-age black population. We use Conley's standard errors with spatial correlations for both OLS and GMM analysis. We use 1000km as the cutoff value above which there is no spatial correlation. *** p<0.01, ** p<0.05, * p<0.1.

Table A1: Ethnic diversity and employment: district level regressions using spatially correlated standard errors

	[1]	[2]	[3]	[4]	[5]	[6]
	Unemployed + inactive			Wage employment		
	Logit	Probit	IV Probit	Logit	Probit	IV Probit
Panel A: 1996 census						
Ethnic fractionalisation ELF	-0.080**	-0.078**	-0.078	0.085**	0.083**	0.082
	(0.033)	(0.033)	(0.080)	(0.034)	(0.034)	(0.081)
Observations	464,130	464,130	464,130	449,200	449,200	449, 200
Panel B: 2001 census						
Ethnic fractionalisation ELF	-0.148***	-0.145***	-0.144	0.145***	0.143***	0.140
	(0.038)	(0.038)	(0.091)	(0.039)	(0.038)	(0.089)
Observations	697,369	697,369	697,369	681,529	681,529	681,529
Individual controls	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES

Note: This table reports results about the effect of ethnic diversity on employment based on non-linear econometric models in 1996 and 2001. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. In column 4, 5 and 6 we drop self-employed people as they are a very small proportion of the whole working-age population. *** p<0.01, ** p<0.05, * p<0.1.

Table A2.1: Estimations based on non-linear econometric models

	[1]	[2]	[3]	[4]
	Mlogit		IV Mprobit	
	Self employment	Wage employee	Self employment	Wage employee
Panel A: 1996 census				
Ethnic fractionalisation ELF	-0.008	0.086**	0.106	0.586
	(0.007)	(0.034)	(0.381)	(0.418)
Observations	464,130	464,130	464,130	464,130
Panel B: 2001 census				
Ethnic fractionalisation ELF	0.012**	0.135***	0.408	0.981***
	(0.005)	(0.038)	(0.427)	(0.369)
Observations	697,369	697,369	697,369	697,369
Individual controls	YES	YES	YES	YES
District controls	YES	YES	YES	YES
Province FE	YES	YES	YES	YES

Note: This table reports results about the effect of ethnic diversity on employment based on multinomial econometric models (both with and without instrumental variables) in 1996 and 2001. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. *** p<0.01, ** p<0.05, * p<0.1.

Table A2.2: Estimations based on multinomial econometric models

	Mean	Std. Dev.	Obs
Inter-ethnic marriage			
Own generation	0.966	0.18	96,031
Parental generation	0.99	0.0966	532
Second language among married people			
Any second language	0.2356	0.424	95,580
Second English/Afrikaans	0.0888	0.284	95,580
Second ethnic language	0.147	0.354	95,580
Second language among whole sample			
Any second language	0.225	0.418	203,327
Second English/Afrikaans	0.087	0.283	203,327
Second ethnic language	0.138	0.345	203,327

Note: This table reports inter-ethnic marriage rate (i.e. marriage between different ethnic groups within the black population). Ethnicity is identified from the first language spoken by both household head and spouse for the current generation, and household head's parents for the parental generation. We also report the proportion of the black population who can speak a second language.

Table A3: Inter-ethnic marriage rate and ethnic diversity: 1996 census

	[1]	[2]	[3]	[4]
	Low edu		High edu	
	OLS	IV	OLS	IV
Panel A: 1996 census				
Ethnic fragmentation index	0.100***	0.142*	0.037	0.051
	(0.035)	(0.078)	(0.029)	(0.089)
F statistics of the instrument		23.89		14.69
R-squared	0.184	0.184	0.275	0.275
Obs	297,206	297,206	151,994	151,994
Panel B: 2001 census				
Ethnic fragmentation index	0.136***	0.191**	0.154***	0.115
	(0.041)	(0.090)	(0.029)	(0,101)
F statistics of the instrument		33.91		25.78
R-squared	0.172	0.171	0.221	0.221
Obs	390,222	390,222	291,307	291,307
Individual controls	YES	YES	YES	YES
District controls	YES	YES	YES	YES
Province FE	YES	YES	YES	YES

Note: This table reports the main results about the heterogeneous effects of ethnic diversity on the probability of being an employee at individual-level regressions by educational levels in both 1996 and 2001 census. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features, individual-level controls aggregated at district average and ethnicity fixed effects. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. "High" ("Low") education is defined as years of schooling above (below) 9. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Ethnic diversity and wage employment rate: by education level

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	Var manager	Var profession	Var clerk	Var serve	Var craft	Var skill agri	Var operator	Var unskill
Panel A: OLS estimates								
Ethnic fragmentation index	0.611	0.361	-1.323	-1.228	-0.320	-1.499	-2.040	-3.813***
	(0.965)	(3.601)	(0.911)	(1.119)	(0.589)	(1.807)	(1.704)	(1.223)
R-squared	0.793	0.865	0.825	0.762	0.538	0.818	0.827	0.803
Obs	205	205	205	205	205	205	205	205
Panel B: IV estimates								
Ethnic fragmentation index	-4.534	-10.272	-8.039***	-10.269*	-4.614*	-22.605**	-23.246***	-9.949*
	(4.595)	(15.507)	(3.034)	(5.651)	(2.405)	(9.911)	(7.629)	(5.247)
F statistics of the instrument	10.19	10.19	10.19	10.19	10.19	10.19	10.19	10.19
R-squared	0.761	0.856	0.760	0.653	0.363	0.660	0.612	0.769
Obs	205	205	205	205	205	205	205	205
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES

Note: This table reports results about the effect of ethnic diversity on the variety of occupations among employees in 1996. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features and individual-level controls aggregated at district average. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. *** p<0.01, ** p<0.05, * p<0.1.

Table A5.1: Ethnic diversity and the range of occupations: 1996

	[1] Var manager	[2] Var profession	[3] Var clerk	[4] Var serve	[5] Var craft	[6] Var skill agri	[7] Var operator	[8] Var unskill
Panel A: OLS estimates								
Ethnic fragmentation index	0.319	-0.960	-1.252**	0.459	-0.982*	-0.898	-0.713	-1.494
	(1.081)	(3.594)	(0.619)	(0.583)	(0.570)	(1.679)	(1.744)	(0.924)
R-squared	0.843	0.888	0.829	0.824	0.571	0.870	0.860	0.809
Obs	210	210	210	210	210	210	210	210
Panel B: IV estimates								
Ethnic fragmentation index	-0.589	-4.808	-5.316***	0.432	-3.628***	-11.541***	-6.318	-4.933*
	(2.093)	(10.362)	(1.856)	(1.600)	(1.349)	(3.753)	(5.270)	(2.554)
F statistics of the instrument	29.85	29.85	29.85	29.85	29.85	29.85	29.85	29.85
R-squared	0.842	0.887	0.787	0.824	0.515	0.825	0.849	0.791
Obs	210	210	210	210	210	210	210	210
Individual controls	YES	YES	YES	YES	YES	YES	YES	YES
District controls	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES

Note: This table reports results about the effect of ethnic diversity on the variety of occupations among employees in 2001. The sample is only for the "white" magisterial districts which can be matched to 1985 census and whose black population accounts for more than 1% of the overall population. We control for district-level variables especially geographical features and individual-level controls aggregated at district average. We also control for province fixed effects. Ethnic diversity is measured with fractionalisation index. *** p<0.01, ** p<0.05, * p<0.1.

Table A5.2: Ethnic diversity and the range of occupations: 2001

Appendix B

Appendix to Chapter 3

B.1 Appendix. Models

B.1.1 Proof of assortative mating and regression coefficients

We first prove equation 3.7.

Proof. From the formula of partial correlation, we have:

$$\rho_{y_m f_w \cdot y_w} = \frac{\rho_{y_m f_w} - \rho_{y_m y_w}\rho_{f_w y_w}}{\sqrt{1-\rho^2_{y_m y_w}}\sqrt{1-\rho^2_{f_w y_w}}}$$

Since we assume $\rho_{y_m f_w \cdot y_w} = 0$, we have:

$$\rho_{y_m f_w} = \rho_{y_w y_m} \cdot \rho_{y_w f_w}$$

□

Proof of equation 3.8 is as follows.

Proof.

$$plim\beta_{y_m f_m} = \frac{Cov(Y_m F_m)}{Var(F_m)} = \frac{\rho_{y_m f_m}}{\sigma^2_{f_m}} \cdot \sigma_{y_m} \cdot \sigma_{f_m} = \rho_{y_m f_m} \cdot \frac{\sigma_{y_m}}{\sigma_{f_m}}$$

□

We then prove equation 3.9.

Proof.

$$\begin{aligned}
\beta_{y_m f_m \cdot s_m} &= \frac{Var(S_m)Cov(Y_m F_m) - Cov(S_m F_m)Cov(Y_m S_m)}{Var(F_m)Var(S_m) - Cov(S_m F_m)^2} \\
&= \frac{\sigma^2_{s_m}\rho_{y_m f_m} \cdot \sigma_{y_m}\sigma_{f_m} - \rho_{s_m f_m} \cdot \sigma_{s_m} \cdot \sigma_{f_m} \cdot \rho_{y_m s_m} \cdot \sigma_{s_m} \cdot \sigma_{y_m}}{\sigma^2_{f_m}\sigma^2_{s_m} - \rho^2_{s_m f_m} \cdot \sigma^2_{s_m}\sigma^2_{f_m}} \\
&= \frac{\frac{\rho_{y_m f_m} \cdot \sigma_{y_m}}{\sigma_{f_m}} - \frac{\rho_{s_m f_m} \cdot \sigma_{y_m} \cdot \rho_{y_m s_m}}{\sigma_{f_m}}}{1-\rho^2_{s_m f_m}} \\
&= \rho_{y_m f_m} \cdot \frac{\sigma_{y_m}}{\sigma_{f_m}}[\frac{1 - \frac{\rho_{s_m f_m} \cdot \rho_{s_m y_m}}{\rho_{y_m f_m}}}{1-\rho^2_{s_m f_m}}]
\end{aligned}$$

□

For equation 3.10:

Proof.

$$\beta_{y_m f_w} = \frac{Cov(y_m F_w)}{Var(F_w)} = \frac{\rho_{y_m f_w} \cdot \sigma_{y_m} \sigma_{f_w}}{\sigma_{f_w}^2}$$

$$\rho_{y_w f_w} = \frac{Cov(y_w F_w)}{\sigma_{y_w} \sigma_{f_w}} = \frac{Cov(y_w F_w)}{\sigma_{f_w}^2} \cdot \frac{\sigma_{f_w}}{\sigma_{y_w}} = \delta_{f_w} \cdot \frac{\sigma_{f_w}}{\sigma_{y_w}}$$

Given equation 3.7, we have:

$$\beta_{y_m f_w} = \frac{\sigma_{y_m} \sigma_{f_w}}{\sigma_{f_w}^2} \cdot \rho_{y_w y_m} \cdot \rho_{y_w f_w} = \delta_{f_w} \cdot \rho_{y_m y_w} \cdot \frac{\sigma_{y_m}}{\sigma_{y_w}}$$

□

Based on this, we prove equation 3.11:

Proof.

$$\begin{aligned}\beta_{y_m f_w \cdot s_m} &= \frac{Var(S_m)Cov(Y_m F_w) - Cov(S_m F_w)Cov(Y_m S_m)}{Var(F_w)Var(S_m) - Cov(S_m F_w)^2} \\ &= \frac{\sigma_{s_m}^2 \rho_{y_m f_w} \cdot \sigma_{y_m} \sigma_{f_w} - \rho_{s_m f_w} \cdot \sigma_{s_m} \cdot \sigma_{f_w} \cdot \rho_{y_m s_m} \cdot \sigma_{s_m} \cdot \sigma_{y_m}}{\sigma_{f_w}^2 \sigma_{s_m}^2 - \rho_{s_m f_w}^2 \cdot \sigma_{s_m}^2 \sigma_{f_w}^2}\end{aligned}$$

Given the second assumption, we have:

$$\rho_{s_m f_w} = \rho_{s_m y_m} \cdot \rho_{y_m y_w} \cdot \rho_{y_w f_w}$$

Then:

$$\begin{aligned}\beta_{y_m f_w \cdot s_m} &= \frac{\frac{\rho_{y_m f_w} \cdot \sigma_{y_m}}{\sigma_{f_w}} - \frac{\rho_{y_w f_w} \cdot \rho_{y_m s_m}^2 \cdot \sigma_{y_m} \cdot \rho_{y_w y_m}}{\sigma_{f_w}}}{1 - \rho_{s_m y_m}^2 \cdot \rho_{y_m y_w}^2 \cdot R_{y_w f_w}^2} \\ &= \frac{\frac{\rho_{y_m y_w} \cdot \sigma_{y_m} \rho_{y_w f_w}}{\sigma_{f_w}} - \frac{\rho_{y_m y_w} \cdot \rho_{y_m s_m}^2 \cdot \sigma_{y_m} \cdot \rho_{y_w f_w}}{\sigma_{f_w}}}{1 - \rho_{s_m y_m}^2 \cdot \rho_{y_m y_w}^2 \cdot R_{y_w f_w}^2}\end{aligned}$$

Since $\rho_{y_w f_w} = \delta_{f_w} \cdot \frac{\sigma_{f_w}}{\sigma_{y_w}}$, we have:

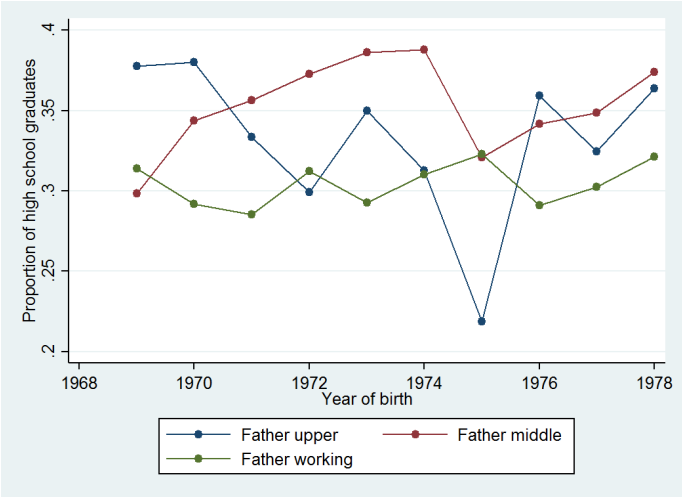
$$\beta_{y_m f_w \cdot s_m} = \frac{\delta_{f_w} \cdot \frac{\rho_{y_m y_w}}{\sigma_{y_w}} \cdot \sigma_{y_m} - \delta_{f_w} \cdot \frac{\sigma_{y_m}}{\sigma_{y_w}} \cdot \rho_{y_m y_w} \cdot \rho_{y_m s_m}^2}{1 - \rho_{s_m y_m}^2 \cdot \rho_{y_m y_w}^2 \cdot R_{y_w f_w}^2}$$

□

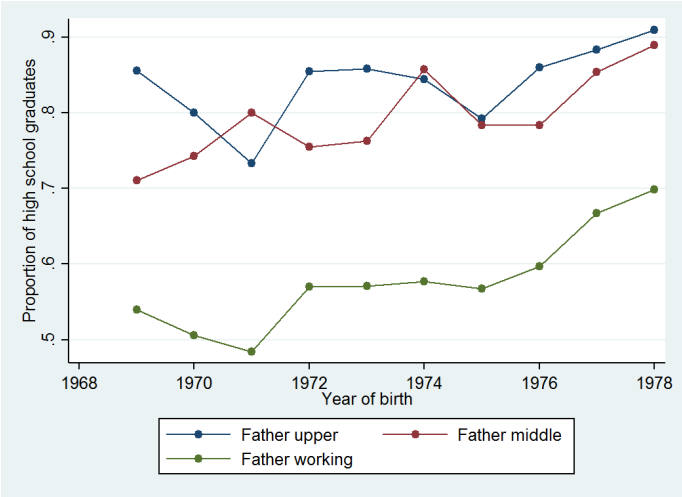
Appendix C

Appendix to Chapter 4

C.1 Appendix. Figures and tables



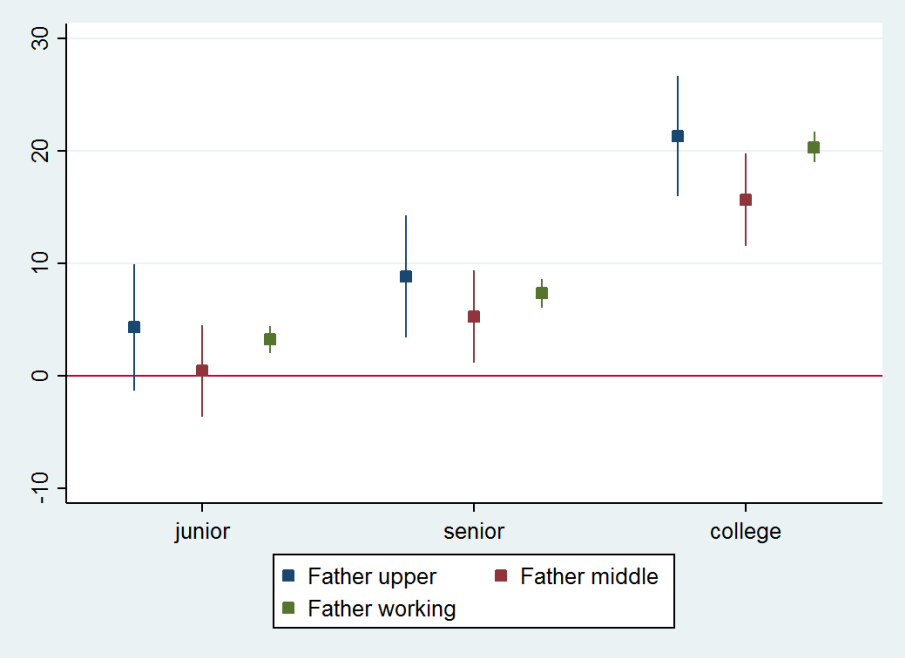
(a) Senior



(b) Senior + college

Notes: The figures capture the composition of children having high school degree (or above) among different social classes over birth cohort. The left hand side shows the proportion of people having high-school degree while the right hand side shows high-school and college degree. Data source: CGSS 2003, 2005, 2008, 2010, 2012, 2013, 2015.

Figure A1: Proportion of people having high school degree or above over birth cohort



Notes: The picture captures the occupational returns to schooling among people from different social classes. Data source: CGSS 2003, 2005, 2008, 2010, 2012, 2013, 2015.

Figure A2: Occupational returns to schooling among social classes

		2003	2005	2008	2010	2012	2013	2015
Demographics								
Age	Mean	33.56	30.07	30.89	32.2	34.43	34.64	35.33
	SD	5.39	2.82	3.53	4.45	5.46	5.6	6.31
Female	Mean	0.51	0.54	0.51	0.49	0.49	0.49	0.54
	SD	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Socio-economic status								
State	Mean	0.38	0.58	0.52	0.42	0.29	0.39	0.35
	SD	0.48	0.49	0.5	0.49	0.45	0.49	0.48
Income	Mean	35760.39	11778.6	16453.9	26479.14	38965.12	42961.41	55264.15
	SD	87024.17	12383.03	22594.41	25312.2	50993.07	48322.97	161982.66
Schoolyear	Mean	12.18	12.15	12.15	12.59	12.31	12.47	12.07
	SD	3.26	2.84	2.84	3.02	3.27	3.23	3.48
Junior	Mean	0.23	0.23	0.22	0.18	0.22	0.21	0.26
	SD	0.42	0.42	0.41	0.38	0.42	0.41	0.44
Senior	Mean	0.29	0.38	0.41	0.36	0.25	0.27	0.25
	SD	0.45	0.48	0.49	0.48	0.44	0.44	0.43
College	Mean	0.42	0.36	0.35	0.41	0.46	0.47	0.41
	SD	0.49	0.48	0.48	0.49	0.5	0.5	0.49
Party member	Mean	0.14	0.12	0.09	0.13	0.14	0.14	0.13
	SD	0.34	0.33	0.29	0.33	0.35	0.35	0.33

Note: The table presents summary statistics of CGSS data 2003, 2005, 2008, 2010, 2012, 2013 and 2015. Socio-economics status of individuals include: if they work in the state-owned sectors, income, years of schooling and communist party membership.

Table A1: Summary statistics of CGSS data

	[1]	[2]	[3]	[4]
	First ISEI		ISEI	
VARIABLES	2SLS	LIML	2SLS	LIML
schoolyear	2.498***	2.502***	3.324***	3.348***
	(0.792)	(0.800)	(0.820)	(0.833)
schoolyear *father ISEI	0.002	0.002	-0.003	-0.003
	(0.014)	(0.014)	(0.014)	(0.014)
father ISEI	0.030	0.030	0.065	0.067
	(0.176)	(0.176)	(0.179)	(0.180)
female	1.828***	1.830***	2.157***	2.165***
	(0.397)	(0.399)	(0.415)	(0.418)
entrance year	0.191	0.190	-0.233*	-0.235*
	(0.125)	(0.126)	(0.134)	(0.135)
entrance year2	-0.008**	-0.008**	0.004	0.004
	(0.004)	(0.004)	(0.004)	(0.004)
minority	-0.163	-0.163	0.066	0.069
	(0.715)	(0.715)	(0.740)	(0.741)
father party member	0.004	0.001	-0.696	-0.710
	(0.556)	(0.559)	(0.583)	(0.589)
father birthyear	0.008	0.007	-0.046	-0.047
	(0.036)	(0.036)	(0.038)	(0.038)
Province FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Constant	-6.710	-6.416	93.900	95.574
	(63.371)	(63.838)	(66.565)	(67.375)
Observations	8,481	8,481	8,481	8,481
R-squared	0.301	0.301	0.258	0.257

Note: The dependent variables include both first occupations and current occupations. Fathers' occupational status is measured by a continuous ISEI scale. Instrumental variables include: policy dummy, its interactions with a cohort trend and the interactions of the above two variables with fathers' ISEI scores. Robust standard errors are calculated. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: IV regressions: father's ISEI scores

Bibliography

Aakvik, A., J. J. Heckman, and E. J. Vytlacil (2005): "Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs," *Journal of Econometrics*, 125, 15–51.

Acemoglu, D. and D. Autor (2011): *Skills, tasks and technologies: Implications for employment and earnings*, vol. 4, Elsevier Inc.

Akerlof, G. A. (1997): "Social distance and social decisions," *Econometrica*, 65, 1005–1027.

Akguc, M., G. Giulietti, and K. F. Zimmermann (2013): "The RUMiC Longitudinal Survey : fostering research on labor markets in China. IZA working paper," .

Alesina, A., R. Baqir, and W. Easterly (1997): "Public goods and ethnic divisions," *The Quarterly Journal of Economics*, 114, 1243–1284.

Alesina, A., A. Devleeschauwer, S. Kurlat, and R. Wacziarg (2003): "Fractionalization," *Journal of Economic Growth*, 8, 155–194.

Alesina, A., C. Gennaioli, and S. Lovo (2016a): "Public goods and ethnic diversity: evidence from deforestation in Indonesia," .

Alesina, A., J. Harnoss, and H. Rapoport (2015): "Birthplace diversity and economic prosperity," .

Alesina, A. and E. La Ferrara (2000): "Participation in heterogeneous communities," *The Quarterly Journal of Economics*, 115, 847–904.

——— (2002): "Who trusts others?" *Journal of Public Economics*, 85, 207–234.

Alesina, A. and E. La Ferrara (2005): "Ethnic diversity and economic performance," *Journal of Economic Literature*, XLIII, 762–800.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016b): "Ethnic inequality," *Journal of Political Economy*, 124, 428–488.

Algan, Y., C. Hemet, and D. D. Laitin (2016): "The social effects of ethnic diversity at the local level: a natural experiment with exogenous residential allocation," *Journal of Political Economy*, 124, 696–732.

Altonji, J. G. and T. A. Dunn (2000): "An intergenerational model of wages , hours , and earnings," *The Journal of Human Resources*, 35, 221–258.

Amodio, F. and G. Chiovelli (forthcoming): "Ethnicity and violence during democratic transitions: evidence from South Africa," *Journal of the European Economic Association*.

Anand, R., S. Kothari, and N. Kumar (2016): "South Africa: labor market dynamics and inequality," .

Andersson, R., J. M. Quigley, and M. Wilhelmsson (2005): "Agglomeration and the spatial distribution of creativity," *Papers in Regional Science*, 84, 445–464.

Anger, S. and G. Heineck (2009): "Do smart parents raise smart children? The intergenerational transmission of cognitive abilities," *Journal of Population Economics*, 23, 1105–1132.

Angrist, J. D. (2001): "Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice," *Journal of Business and Economic Statistics*, 19, 2–28.

Angrist, J. D. and G. W. Imbens (1995): "Two-stage least squares estimation of average causal effects in models with variable treatment intensity," *American Statistical Association*, 90, 431–442.

Arrow, K. (1971): "The theory of discrimination," *Discrimination in Labor Markets*, 3.

Atkin, D. (2015): "Endogenous skill acquisition and export manufacturing in Mexico," *Working Paper*, 1–66.

Auten, B. G., G. Gee, and N. Turner (2013): "Income inequality, mobility, and turnover at the top in the US, 1987-2010," *The American Economic Review: Papers and proceedings*, 103, 168–172.

Awaworyi Churchill, S. (2015): "Fractionalization and entrepreneurial activities," .

Aydemir, A., W.-h. Chen, and M. Corak (2009): "Intergenerational earnings mobility among the children of Canadian immigrants," *The Review of Economics and Statistics*, 91, 377–397.

Bai, C.-E., J. Lu, and Z. Tao (2006): "The multitask theory of state enterprise reform: empirical evidence from China," *The American Economic Review*, 96, 353–357.

Bandiera, O., I. Barankay, and I. Rasul (2005): "Social preferences and the response to incentives: evidence from personnel data," *The Quarterly Journal of Economics*, 120, 917–962.

——— (2009): "Social connections and incentives in the workplace: evidence from personnel data," *Econometrica*, 77, 1047–1094.

——— (2010): "Social incentives in the workplace," *Review of Economic Studies*, 77, 417–458.

Banerjee, A., S. Galiani, J. Levinsohn, Z. McLaren, and I. Woolard (2008): "Why has unemployment risen in the New South Africa?" *Economics of Transition*, 16, 715–740.

Banerjee, A. V. and A. F. Newman (1993): "Occupational choice and the processof development," *Journal of Political Economy*, 101, 274–298.

Barr, A. (1998): "Enterprise performance and the functional diversity of social capital," Tech. rep., Centre for the Study of African Economies, University of Oxford.

Bayer, P., S. L. Ross, and G. Topa (2008): "Place of work and place of residence: informal hiring networks and labor market outcomes," *Journal of Political Economy*, 116, 1150–1196.

Beaman, L. A. (2012): "Social networks and the dynamics of labour market outcomes: evidence from refugees resettled in the US," *The Review of Economic Studies*, 79, 128–161.

Becker, G. (1981): *A treatise on the family*, Chicago: University of Chicago Press.

Becker, G. S. and N. Tomes (1979): "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Political Economy*, 87, 1153–1189.

Behrman, J. and P. Taubman (1985): "Intergenerational earnings mobility in the United States: some estimates and a test of Becker's intergenerational endowments model," *The Review of Economics and Statistics*, 67, 144–151.

Behrman, J. R., A. Gaviria, and M. Szekely (2001): "Intergenerational mobility in Latin America," *Economía*, 2, 1–31.

Beine, M., F. Docquier, and M. Schiff (2013): "International migration, transfer of norms and home country fertility," *Canadian Journal of Economics/Revue canadienne d'économique*, 46, 1406–1430.

Bennedsen, M., K. M. Nielsen, F. Perez-Gonzalez, and D. Wolfenzon (2007): "Inside the family firm: the role of families in succession decisions and performance," *The Quarterly Journal of Economics*, 122, 647–691.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How much should we trust difference-in-difference estimates?" *The Quarterly Journal of Economics*, 119, 249–275.

Bhorat, H. and M. Oosthuizen (2005): "The post-Apartheid South African labour market," .

Bian, Y. (1994): "Guanxi and the allocation of urban jobs in China," *The Quarterly Journal of Economics*, 140, 971–999.

——— (1997): "Bringing strong ties back in: indirect ties, network bridges, and job searches in China," *American Sociological Review*, 62, 366–385.

——— (2002): "Chinese social stratification and social mobility," *Annual Review of Sociology*, 28, 91–116.

Bian, Y., R. Breiger, D. Davis, and J. Galaskiewicz (2005): "Occupation, class, and social networks in urban China," *Social Forces*, 83, 1443–1468.

Bian, Y. and L. Li (2012): "The Chinese General Social Survey (2003-8): sample designs and data evaluation," *Chinese Sociological Review*, 45, 70–97.

Bingley, P. and L. Cappellari (2014): "Correlations of brothers' earnings and intergenerational transmission," .

Bjorklund, A. and M. Jantti (1997): "Intergenerational income mobility in Sweden compared to the United States," *The American Economic Review*, 87, 1009–1018.

Bjorklund, A., M. Lindahl, and E. Plug (2006): "The origins of intergenerational associations: lessons from Swedish adoption data," *The Quarterly Journal of Economics*, 121, 999–1028.

Black, S. E. and P. Devereux (2011): "Recent developments in intergenerational mobility," *Handbook of Labor Economics*, 4, 1487–1541.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005): "Why the apple doesn't fall far: understanding intergenerational transmission of human capital," *The American Economic Review*, 95, 437–449.

——— (2009): "Like father, like son? a note on the intergenerational transmission of IQ scores," *Economics Letters*, 105, 138–140.

Blanden, J. (2004): "Intergenerational mobility and assortative mating in the UK," .

Blau, P. M. and O. D. Duncan (1967): *The American occupational structure*, New York: John Wiley and Sons.

Borjas, G. J. (1992): "Ethic capital and intergenerational mobility," *The Quarterly Journal of Economics*, 107, 123–150.

Borjas, G. J. and S. G. Bronars (1989): "Consumer discrimination and self-employment," *Journal of Political Economy*, 97, 581–605.

Botticini, M. and Z. Eckstein (2006): "Path dependence and occupations," *New Palgrave Dictionary of Economics*.

Bowles, S. and H. Gintis (2002): "The inheritance of inequality," *The Journal of Economic Perspectives*, 16, 3–30.

Brugiavini, A., D. Cavapozzi, G. Pasini, and E. Trevisan (2013): "Working life histories from SHARELIFE: a retrospective panel," .

Bun, M. J. and T. D. Harrison (2014): "OLS and IV estimation of regression medoels including endogenous interactin terms," .

Burchardi, K. B. and T. A. Hassan (2013): "The economic impact of social ties: evidence from German reunification," *The Quarterly Journal of Economics*, 128, 1219–1271.

Calvó-armengol, A. and M. O. Jackson (2004): "The effects of social networks on employment and inequality," *The American Economic Review*, 94, 426–454.

Cameron, A. C. and P. K. Trivedi (2005): *Microeconometrics: methods and applications*, Cambridge, UK: Cambridge University Press.

Card, D. (2001): "Estimating the return to schooling: progress on some persistent econometric problems," *Econometrica*, 69, 1127–1160.

Carmichael, F. (2000): "Intergenerational mobility and occupational status in Britain," *Applied Economics Letters*, 7, 391–396.

Carneiro, P., C. Meghir, and M. Parey (2013): "Maternal education, home environments, and the development of children and adolescents," *Journal of the European Economic Association*, 11, 123–160.

Chadwick, L. and G. Solon (2002): "Intergenerational income mobility among daughters," *The American Economic Review*, 92, 335–344.

Chay, K. and K. Munshi (2015): "Black networks after emancipation: evidence from reconstruction and the Great Migration," .

Chen, M. (2012): "Intergenerational mobility in contemporary China : 1996-2006," Phd thesis, University of Manchester.

Cheng, Y. and J. Dai (1995): "Intergenerational mobility in modern China," *European Sociological Review*, 11, 17–35.

Chevalier, A. (2004): "Parental education and child's education: a natural experiment," .

Christopher, A. J. (2001): *The atlas of changing South Africa*, London and New York: Routledge.

Cichello, P. L., C. Almeleh, L. Mncube, and M. Oosthuizen (2011): "Perceived barriers to entry into self-employment in Khayelitsha, South Africa: crime, risk, start-up capital dominate profit concerns," .

Colclough, C., G. Kingdon, and H. A. Patrinos (2009): "The pattern of returns to education and its implications," *Research Consortium on Educational Outcomes & Poverty*, 1–5.

Conley, T. G. and G. Topa (2002): "Socio-economic distance and spatial patterns in unemployment," *Journal of Applied Econometrics*, 17, 303–327.

Corak, M. (2006): "Do poor children become poor adults? lessons from a cross-country comparison of generational earnings mobility," .

——— (2013): "Income inequality, equality of opportunity and intergenerational mobility," *The Journal of Economic Perspectives*, 27, 79–102.

Couch, K. a. and D. R. Lillard (1998): "Sample selection rules and the intergenerational correlation of earnings," *Labour Economics*, 5, 313–329.

Cui, Y., D. Nahm, and M. Tani (2013): "Earnings differentials and returns to education in China, 1995-2008. IZA working paper," .

Currie, J. and E. Moretti (2003): "Mother's education and the intergenerational transmission of human capital: evidence from college openings," *The Quarterly Journal of Economics*, 118, 1495–1532.

Dahl, M. and T. DeLeire (2008): "The association between children's earnings and fathers' lifetime earnings: estimates using administrative data," *Institute for Research on Poverty.*

Dahlberg, M., K. Edmark, and H. Lundqvist (2012): "Ethnic diversity and preferences for redistribution," *Journal of Political Economy*, 120, 41–76.

Dávila, B. A., M. T. Mora, and S. K. Stockly (2011): "Does Mestizaje matter in the US ? economic stratification of Mexican immigrants," *The American Economic Review: Papers and proceedings*, 101, 593–597.

de Parros, R. P., F. H. G. Ferreira, J. R. M. Vega, and J. S. Chanduvi (2009): *Measuring inequality of opportunities in Latin America and the Caribbean*, Washington, D.C.: The World Bank.

Deaton, A. (1985): "Panel data from time series of cross-sections," *Journal of Econometrics*, 30, 109–126.

Deming, D. (2017): "The growing importance of social skills in the labor market," *The Quarterly Journal of Economics*, 1593–1640.

Démurger, S., S. Li, and J. Yang (2012): "Earnings differentials between the public and private sectors in China: Exploring changes for urban local residents in the 2000s," *China Economic Review*, 23, 138–153.

Deng, Z. and D. J. Treiman (1997): "The impact of the Cultural Revolution on trends in educational attainment in the People's Republic of China," *American Journal of Sociology*, 103, 391–428.

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2012): "The political economy of linguistic cleavages," *Journal of Development Economics*, 97, 322–338.

Desmet, K., I. Ortuno-Ortin, and R. Wacziarg (2017): "Culture, ethnicity, and diversity," *American Economic Review*, 107, 2479–2513.

Di Pietro, G. and P. Urwin (2003): "Intergenerational mobility and occupational status in Italy," *Applied Economics Letters*, 10, 793–797.

Dumont, J.-C., G. Spielvogel, and S. Widmaier (2010): "International migrants in developed, emerging and developing countries," *OECD Social, Employment and Migration Working Paper No.114*.

Duranton, G. and D. Puga (2001): "Nursery cities: urban diversity, process innovation, and the life cycle of products," *The American Economic ReviewAmerican Economic Review*, 91, 1545–1477.

Dustmann, C. (2005): "Intergenerational mobility and return migration: comparing the sons of foreign and native born fathers," *Centre for Research and Analysis of Migration*.

Dustmann, C., T. Frattini, and I. P. Preston (2013): "The effect of immigration along the distribution of wages," *The Review of Economic Studies*, 80, 145–173.

Dustmann, C., U. Schoenberg, and J. Stuhler (2017): "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 132, 435–483.

Easterly, W. and R. Levine (1997): "Africa's growth tragedy: policies and ethnic divisions," *The Quarterly Journal of Economics*, 112, 1203–1250.

Eika, L., M. Magne, and Z. Basit (2014): "Educational assortative mating and household income inequality," *NBER Working Papers*.

Eisenhauer, P. and F. Pfeiffer (2008): "Assessing intergenerational earnings persistence among German workers," .

Emran, M. S. and Z. Hou (2013): "Access to markets and rural poverty: evidence from household consumption in China," *The Review of Economics and Statistics*, 95, 682–697.

Emran, M. S. and F. Shilpi (2010): "Intergenerational occupational mobility in rural economy: evidence from Nepal and Vietnam," *The Journal of Human Resources*, 46, 429–458.

——— (2011): "Intergenerational occupational mobility in rural economy: evidence from Nepal and Vietnam," *The Journal of Human Resources*, 46, 427–458.

Emran, M. S. and Y. Sun (2011): "Magical transition? intergenerational educational and occupational mobility in rural China: 1988-2002," .

——— (2014): "Are the children of uneducated farmers doubly doomed? farm, nonfarm and intergenerational educational mobility in rural China," .

Eraydin, A., T. Tasan-Kok, and J. Vranken (2010): "Diversity matters: immigrant entrepreneurship and contribution of different forms of social integration in economic performance of cities," *European Planning Studies*, 18, 521–543.

Erikson, R. and J. H. Goldthorpe (2002): "Intergenerational inequality: a sociological perspective," *Journal of Economic Perspectives*, 16, 31–44.

Ermisch, J. and M. Francesconi (2002): "Intergenerational mobility in Britain: new evidence from the BHPS. Institute for Social and Economic Research working paper," .

Ermisch, J., M. Francesconi, and T. Siedler (2006): "Intergenerational mobility and marital sorting," *The Economic Journal*, 116, 659–679.

Fainstein, S. S. (2005): "Cities and diversity: should we want it? Can we plan for it?" *Urban Affairs Review*, 41, 3–19.

Fairlie, R. W. (1999): "The absence of the African-American owned business: an analysis of the dynamics of self-employment," *Journal of Labor Economics*, 17, 80–108.

Fan, C. (2002): "The elite, the natives, and the outsiders: migration and labor market segmentation in urban China," *Annals of the Association of American Geographers*, 92, 103–124.

Fang, H., K. N. Eggleston, J. A. Rizzo, S. Rozelle, and R. J. Zeckhauser (2012): "The returns to education in China: evidence from the 1986 compulsory education law. NBER Working Paper," .

Fedderke, J. (2012): "The cost of rigidity: the case of the South African labor market," .

Finn, A. and M. Leibbrandt (2016): "Patterns of persistence: Intergenerational mobility and Education in South Africa," *South Africa Labour and Development Research Unit*.

Ganzeboom, H. B. (1996): "Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations," *Social Science Research*, 239, 201–239.

Giuntella, O., G. L. Mattina, and C. Quintana-Domeque (2017): "Assortative mating, intergenerational transmission and inequality: evidence from birth weight using parental grandmother fixed effects," .

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992): "Growth in cities," *Journal of Political Economy*, 100, 1126–1152.

Glaeser, E. L., J. Scheinkman, and A. Shleifer (1995): "Economic growth in a cross-section of cities," *Journal of Monetary Economics*, 36, 117–143.

Goldberg, M. S. (1982): "Discrimination, nepotism, and long-run wage differentials," *The Quarterly Journal of Economics*, 307–319.

Gomes, J., K. Desmet, and O.-O. Ignacio (2016): "The geography of linguistic diversity and the provision of public goods," .

Gong, H., A. Leigh, and X. Meng (2012): "Intergenerational income mobility in urban China," *Review of Income and Wealth*, 58, 481–503.

Gradin, C. (2014): "Poverty and ethnicity among black South Africans," .

Granovetter, M. (1983): "The strength of weak ties: a network theory revisited," *Sociological Theory*, 1, 201–233.

——— (2000): "The economic sociology of firms and entrepreneurs," in *Entrepreneurship: The Social Science View*, ed. by R. Swedberg, Oxford University Press, 244–280.

Grimm, M., J. Krüger, and J. Lay (2011): "Barriers to entry and returns to capital in informal activities: evidence from Sub-Saharan Africa," *Review of Income and Wealth*, 57.

Guy, J. and M. Thabane (1988): "Technology, ethnicity and ideology: basotho miners and shaft-sinking on the South African gold mines," *Journal of Southern African Studies*, 14, 257–278.

Haihui, Z., X. Zhaohui, S. Jiang, and G. L. Lugar (2013): *A scholarly review of Chinese studies in North America (in Chinese)*, Ann Arbor: Association for Asian Studies.

Handy, C. (2015): "Assortative mating and intergenerational persistence of schooling and earnings," *MPRA Paper*.

Hanson, G. H. and C. Mcintosh (2010): "The great Mexican emigration," *The Review of Economics and Statistics*, 92, 798–810.

Harper, D. A. (2003): *Foundations of entrepreneurship and economic development*, London, UK: Routledge.

Heintz, J. and D. Posel (2007): "Revisiting informal employment and segmentation in the South African labor market," .

Hertz, T. (2008): "A group-specific measure of intergenerational persistence," *Economics Letters*, 100, 415–417.

Hjort, J. (2014): "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 1899–1946.

Hnatkovska, V., A. Lahiri, and S. B. Paul (2013): "Breaking the caste barrier: intergenerational mobility in India," *The Journal of Human Resources*, 48, 435–473.

Holmlund, H. (2008): "Intergenerational mobility and assortative mating: effects of an educational reform," *Centre for the Economics of Education*.

Huang, Y. (2001): "Gender, hukou, and the occupational attainment of female migrants in China (1985 - 1990)," *Environment and Planning A*, 33, 257–279.

Ioannides, Y. M. and L. D. Loury (2004): "Job information networks, neighborhood effects and inequality," *Journal of Economic Literature*, 42, 1056–1093.

Iyer, G. R. and J. M. Shapiro (1999): "Ethnic entrepreneurial and marketing systems: Implications for the global economy," *Journal of International Marketing*, 83–110.

Jaggi, G., M. Rundle, D. Rosen, and Y. Takahashi (1996): "China's economic reforms: chronology and statistics," .

Jr., R. G. F. and G. C. Loury (2013): "Valuing diversity," *Journal of Political Economy*, 121, 747–774.

Kahn, C. and D. Mookherjee (1988): "A competitive efficiency wage model with Keynesian features," *The Quarterly Journal of Economics*, CIII, 609–645.

Keswell, M. and L. Poswell (2004): "Returns to education in South Africa: a retrospective sensitivity analysis of the available evidence," *South African Journal of Economics*, 72, 834–860.

Khor, N. and J. Pencavel (2006): "Income mobility of individuals in China and the United States," .

Khunou, S. (2009): "Traditional leadership and independent Bantustans of South Africa: some milestones of transformative constitutionalism beyond Apartheid," *PER: Potchefstroomse Elektroniese Regsblad*, 12, 81–122.

Kim, S. (1989): "Labor specialization and the extent of the market," *Journal of Political Economy*, 97, 692–705.

King, G. and L. Zeng (2001): "Logistic regression in rare events data," *Political analysis*, 9, 137–163.

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

Kramarz, F. and O. N. Skans (2014): "When strong ties are strong: Networks and youth labour market entry," *The Review of Economic Studies*, 0, 1–37.

Kreidl, M., H. B. Ganzeboom, and D. J. Treiman (2014): "How did occupational returns to education change over time?" .

La Ferrara, E. (2002): "Inequality and group participation: theory and evidence from rural Tanzania," *Journal of Public Economics*, 85, 235–273.

Labar, K. (2007): "Intergenerational mobility in China. CERDI working paper," .

Lam, D. and R. F. Schoeni (1993): "Effects of family background on earnings and returns to schooling : Evidence from Brazil," *Journal of Political Economy*, 101, 710–740.

——— (1994): "Family ties and labor markets in the United States and Brazil," *Journal of Human Resources*, 29, 1235–1258.

Lazear, E. P. (1999a): "Culture and language," *Journal of Political Economy*, 107, 95–124.

——— (1999b): "Globalisation and the market for team-mates," *Economic Journal*, 109, 15–40.

Lee, C.-i. and G. Solon (2009): "Trends in intergenerational income mobility," *The Review of Economics and Statistics*, 91, 766–772.

Leibbrandt, M., I. Woolard, H. McEwen, and C. Koep (2009): "Employment and inequality outcomes in South Africa," .

Lewis, M. P., G. F. Simons, and C. D. Fennig (2009): *Ethnologue: languages of the world*, vol. 16, SIL international Dallas, TX.

Li, S., J. Whalley, and C. Xing (2013): "China's higher education expansion and unemployment of college graduates," *China Economic Review*.

Li, S. and C. Xing (2010): "China's higher education expansion and its labor market consequences," .

Lin, N. and Y. Bian (1991): "Getting ahead in urban China," *American Journal of Sociology*, 97, 657–688.

Lindbeck, A. and D. Snower (2002): "The insider-outsider theory: a survey," .

Long, J. and J. Ferrie (2013): "Intergenerational occupational mobility in Great Britain and the United States since 1850," *The American Economic Review*, 103, 1109–1137.

Loury, G. C. (1981): "Intergenerational transfers and the distribution of earnings," *Econometrica*, 49, 843–867.

Loury, L. D. (2006): "Some contacts are more equal than others: informal networks, job tenure and wages," *Journal of Labor Economics*, 24, 299–318.

Luke, N. and K. Munshi (2006): "New roles for marriage in urban Africa: kinship networks and the labor market in Kenya," *Review of Economics and Statistics*, 88, 264–282.

Luke, N., K. Munshi, and M. Rosenzweig (2004): "Marriage, neworks, and jobs in third world cities," *Journal of the European Economic Association*, 2, 437–446.

Magruder, J. R. (2010): "Intergenerational networks, unemployment, and persistent inequality in South Africa," *American Economic Journal: Applied Economics*, 2, 62–85.

——— (2012): "High unemployment yet few small firms: the role of centralized bargaining in South Africa," *American Economic Journal: Applied Economics*, 4, 138–166.

Majumder, R. (2010): "Intergenerational mobility in educational and occupational attainment: a comparative study of social class in India," *The Journal of Applied Economic Research*, 4, 463–494.

Maoz, Y. D. and O. Moav (1999): "Intergenerational mobility and process of development," *The Economic Journal*, 109, 677–697.

Mare, R. D. (2000): "Assortative mating, intergenerational mobility, and educational inequality," *California Center for Population Research*.

Mariotti, M. (2012): "Labour markets during apartheid in South Africa," *The Economic History Review*, 65, 1100–1122.

Maurin, E. and S. McNally (2008): "Vive la revolution! long-term educational returns of 1968 to the angry students," *Journal of Labor Economics*, 26, 1–33.

Mayda, A. M. (2010): "International migration: a panel data analysis of the determinants of bilateral flows," *Journal of Population Economics*, 23, 1249–1274.

Mazumder, B. (2005): "Fortunate sons: new estimates of intergenerational mobility in the United States using social security earnings data," *Review of Economics and Statistics*, 87, 235–255.

——— (2011): "Black-white differences in intergenerational economic mobility in the US," *Federal Reserve Bank of Chicago*, 1–28.

McCrary, J. and H. Royer (2011): "The effect of female education on fertility and infant health: evidence from school entry policies using exact data of birth," *The American economic review*, 101, 158–195.

Mckenzie, D. and H. Rapoport (2007): "Network effects and the dynamics of migration and inequality: theory and evidence from Mexico," *Journal of Development Economics*, 84, 1–24.

McKenzie, D. J. and C. Woodruff (2006): "Do entry costs provide an empirical basis for poverty traps? Evidence from Mexican microenterprises," *Economic development and cultural change*, 55, 3–42.

Michalopoulos, S. (2012): "The origins of ethnolinguistic diversity," *American Economic Review*, 102, 1508–1539.

Michalopoulos, S. and E. Papaioannou (2013): "Pre-Colonial ethnic institutions and contemporary African development," *Econometrica*, 81, 113–152.

Miguel, E. and M. K. Gugerty (2005): "Ethnic diversity, social sanctions, and public goods in Kenya," *Journal of Public Economics*, 89, 2325–2368.

Miller, R. A. (1984): "Job matching and occupational choice," *Journal of Political Economy*, 92, 1086–1120.

Montalvo, J. G. and M. Reynal-Querol (2005): "Ethnic polarization, potential conflict, and civil wars," *The American Economic Review*, 95, 796–816.

Montgomery, B. J. D. (1991): "Social networks and labor-market outcomes: toward an economic analysis," *The American Economic Review*, 81, 1408–1418.

Montgomery, J. D. (1992): "Job search and network composition: implications of the strength-of-weak-ties hypothesis," *American Sociological Review*, 57, 586–596.

Mookherjee, D. and D. Ray (2003): "Persistent inequality," *Review of Economic Studies*, 70, 369–393.

Morgan, J. and F. Vardy (2009): "Diversity in the workplace," *American Economic Review*, 99, 472–485.

Munshi, K. (2003): "Networks in the modern economy: Mexican migrants in the U.S. labor market," *The Quarterly Journal of Economics*, 118, 549–599.

——— (2011): "Strength in numbers: networks as a solution to occupational traps," *Review of Economic Studies*, 78, 1069–1101.

Munshi, K. and M. Rosenzweig (2006): "Traditional institutions meet the modern world: caste, gender and schooling choices in a globalizing economy," *American Economic Review*, 96, 1225–1252.

Murdock, G. P. (1959): *Africa: its peoples and their culture history*, McGraw-Hill.

Mwakikagile, G. (2010): *South Africa as a multi-ethnic society*, Dar es Salaam, Tanzania: Continental Press.

Neumann, G. and R. Topel (1991): "Employment risk, diversification, and unemployment," *The Quarterly Journal of Economics*, 1341–1365.

Newey, W. K. (1987): "Efficient estimation of limited dependent variable models with endogenous explanatory variables," *Journal of Econometrics*, 36, 231–250.

Ng, I. Y., X. Shen, and K. W. Ho (2009): "Intergenerational earnings mobility in Singapore and the United States," *Journal of Asian Economics*, 20, 110–119.

Niebuhr, A. (2010): "Migration and innovation: Does cultural diversity matter for regional R&D activity?" *Papers in Regional Science*, 89, 563–585.

Nikolova, E., D. Simroth, et al. (2013): "Does cultural diversity help or hinder entrepreneurs? Evidence from eastern Europe and central Asia," Tech. rep., European Bank for Reconstruction and Development.

Nunn, N. and D. Puga (2012): "Ruggedness: the blessing of bad geography in Africa," *Review of Economics and Statistics*, 94, 20–36.

Nybom, M. and J. Stuhler (2016): "Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation," *Journal of Human Resources*, 51, 239–268.

Obukhova, E. (2012): "Motivation vs. relevance: using strong ties to find a job in Urban China," *Social science research*, 41, 570–80.

Oreopoulos, P. (2003): "The long-Run consequences of living in a poor neighborhood," *The Quarterly Journal of Economics*, 118, 1533–1575.

Oreopoulos, P., M. Page, and A. H. Stevens (2008): "The intergenerational effects of worker displacement," *Journal of Labor Economics*, 26, 455–000.

Oreopoulos, P., M. E. Page, and A. H. Stevens (2006): "The intergenerational effects of compulsory schooling," *Journal of Labor Economics*, 24, 729–760.

Ortega, F. and G. Peri (2014): "Openness and income: The roles of trade and migration," *Journal of International Economics*, 92, 231–251.

Ottaviano, G. I. and G. Peri (2006): "The economic value of cultural diversity: evidence from US cities," *Journal of Economic geography*, 6, 9–44.

Page, M. E. (2006): "Father's education and children's human capital: evidence from the World War II G. I. Bill," .

Pakpahan, Y. M., D. Suryadarma, and A. Suryahadi (2009): "Destined for destitution: intergenerational poverty persistence in Indonesia. SMERU Research Institute working paper," .

Pasquier-Doumer, L. (2012): "Intergenerational transmission of self-employed status in the informal sector: a constrained choice or better income prospects? Evidence from seven West African countries," *Journal of African Economies*, 22, 73–111.

Patel, K. and F. Vella (2013): "Immigrant networks and their implicaitons for occupational choice and wages," *The Review of Economics and Statistics*, 95, 1249–1277.

Perez-Gonzalez, F. (2006): "Inherited control and firm performance," *The American Economic Review*, 96, 1559–1588.

Peri, G. and C. Sparber (2009): "Task specialization, immigration and wages," *American Economic Journal: Applied Economics*, 1, 135–169.

Peters, H. E. (1992): "Patterns of intergenerational mobility in income and earnings," *The Review of Economics and Statistics*, 74, 456–466.

Piraino, P. (2015): "Intergenerational earnings mobility and equality of opportunity in South Africa," *World Development*, 67, 396–405.

Posel, D. (2001): "What's in a name? Racial categorisations under apartheid and their afterlife," *Transformation 47*, 50–74.

Putnam, R. D. (2007): "E pluribus unum: diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture," *Scandinavian Political Studies*, 30, 137–174.

Qian, N. (2008): "Missing women and the price of tea in China: the effect of sex-specific earnings on sex imbalance," *The Quarterly Journal of Economics*, 123, 1251–1285.

Quheng, D., B. Gustafsson, and L. Shi (2012): "Intergenerational income persistency in urban China. IZA working paper," .

Raftery, A. E. and M. Hout (1993): "Maximally maintained inequality: expansion, reform, and opportunity in Irish education, 1921-75," *Sociology of Education*, 66, 41–62.

Rees, A. (1966): "Information networks in labor markets," *The American Economic Review*, 56, 559–566.

Rosenzweig, M. R. and J. Zhang (2013): "Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of Chinese twins," *Journal of Development Economics*, 104, 245–260.

Santos Silva, J. and S. Tenreyro (2006): "The log of gravity," *The Review of Economics and statistics*, 88, 641–658.

Sicherman, N. and O. Galor (1990): "A theory of career mobility," *Journal of Political Economy*, 98, 169–192.

Simon, C. J. and J. T. Warner (1992): "Matchmaker, matchmaker: the effect of old boy networks on job match quality, earnings, and tenure," *Journal of Labor Economics*, 10, 306–330.

Sobel, R. S., N. Dutta, and S. Roy (2010): "Does cultural diversity increase the rate of entrepreneurship?" *The Review of Austrian Economics*, 23, 269–286.

Solon, G. (1992): "Intergenerational income mobility in the United States," *The American Economic Review*, 82, 393–408.

——— (1999): "Intergenerational mobility in the labor market," *Handbook of Labor Economics*, 3, 1761–1800.

——— (2002): "Cross-country differences in intergenerational earnings mobility," *Journal of Economic Perspectives*, 16, 59–66.

Sørensen, J. B. (2007): "Closure and exposure: Mechanisms in the intergenerational transmission of self-employment," in *The Sociology of Entrepreneurship*, Emerald Group Publishing Limited, 83–124.

Sun, Y. (2013): "Fate or opportunity? impact of family background on household welfare in rural China," Ph.D. thesis, The George Washington University.

Takenoshita, H. (2007): "Intergenerational mobility in East Asian countries: a comparative study of Japan, Korea and China," *International Journal of Japanese Sociology*, 16, 64–79.

Topa, G. (2001): "Social interactions, local spillovers and unemployment," *The Review of Economic Studies*, 68, 261–295.

Train, K. (2003): *Dicrete choice methods with simulation*, Cambridge: Cambridge University Press.

Treiman, D. J., M. Mckeever, and E. Fodor (1996): "Racial differences in occupational status and income in South Africa, 1980 and 1991," *Demography*, 33, 111–132.

von Fintel, D. (2016): "Institutional wage-setting, labour demand and labour supply: causal estimates from a South African pseudo-panel," .

Wahba, J. and Y. Zenou (2005): "Density, social networks and job search methods: theory and application to Egypt," *Journal of Development Economics*, 78, 443–473.

Walder, A. G. (1995): "Career mobility and the communist political order," *American Sociological Review*, 60, 309–328.

Wang, S.-Y. (2012): "Credit constraints, job mobility and entrepreneurship: evidence from a property reform in China," *The Review of Economics and Statistics*, 94, 532–551.

——— (2013): "Marriage networks, nepotism and labor market outcomes in China," *American Economic Journal: Applied Economics*, 5, 91–112.

Whyte, M. K. (1996): "The Chinese family and economic development: obstacle or engine?" *Economic Development and Cultural Change*, 45, 1–30.

Wittenberg, M. (2002): "Job search in South Africa: a nonparametric analysis," *South African Journal of Economics*, 70, 1163–1197.

Wong, R. S.-K. (1990): "Understanding cross-national variation in occupational mobility," *American Sociological Review*, 55, 560–573.

Wu, X. (2010): "Voluntary and involuntary job mobility and earnings inequality in urban China, 1993−2000," *Social Science Research*, 39, 382–395.

——— (2014): *School choice in China: a different tale?*, Abingdon, Oxon: Routledge.

Wu, X. and D. J. Treiman (2007): "Inequality and equality under Chinese socialism: the Hukou system and intergenerational occupational mobility," *American Journal of Sociology*, 113, 415–445.

Xie, Y. (2012): "China Family Panel Studies (2010) User's Mannual," .

Yang, J. (2006): *Analysis of social structure in contemporary China (in Chinese)*, Lanzhou: Gansu Renmin Press.

Yeung, W.-J. J. (2013): "China's higher education expansion and social stratification. ASIA RESEARCH Institute working paper," .

Zhang, X. and G. Li (2003): "Does guanxi matter to nonfarm employment?" *Journal of Comparative Economics*, 31, 315–331.

Zhigang, Y. and C. Lin (2013): "The trend and mechanism of intergenerational income mobility in China: an analysis from the perspective of human capital, social capital and wealth," *The World Economy*, 36, 880–898.