Essays on Education, Discrimination and Development

Rajesh Ramachandran

Supervised by Hannes Müller

Departament d'Economia i d'Historia Economica,

Universitat Autònoma de Barcelona

May 2013

Thesis Supervisor: Hannes Müller   Tutor: Caterina Calsamiglia

Author: Rajesh Ramachandran

Essays on Education, Discrimination and Development

by

Rajesh Ramachandran

Submitted to the Department of Economics

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

UNIVERSITAT AUTÒNOMA DE BARCELONA

May 2013

This is dedicated to my amma and appa, for their undying support and for teaching me the value of knowledge.

Acknowledgements

First and foremost, I would like to thank Hannes Müller my supervisor for his continued support during the making of this thesis. I will be forever grateful for his confidence in me and for continuously pushing me in the correct direction. I have immensely benefitted from the numerous discussions, feedback and criticisms received during the course of this dissertation. I have learnt not only how to approach problems from a critical perspective but also more importantly the joy of doing good research and I hope that I will continue to learn from him in the future.

A special thanks goes to Francesco Fasani, at the IAE-CSIC, from whose suggestions and comments, my first chapter has immensely benefited. I would also like to extend my gratitude to Alessandro Tarozzi, at the UPF, for his generous suggestions, advice and time. I would also like to thank the other faculty members at the department and the Institute at UAB. In particular, I would like to acknowledge the help of Ana-Rute Cardoso, Lidia Farrè and Omar Licandro.

A special thanks goes to the co-author of my second chapter, Christopher Rauh. I have, both personally and professionally, learnt enormously from him in the course of the last seven years that I have been fortunate to know him. I hope that this is the first of our many papers together. I would also like to thank Ashwini Deshpande, my co-author of the third chapter and former teacher in Delhi, who laid the foundations of my interest in research. I

have during the course of the last years learnt a lot through numerous discussions with my colleagues and other students at the department and otherwise. A special mention goes to Jack Willis.

Finally, I would also like to acknowledge the generous financial support that I have received over the course of my studies from the Universitat Autònoma de Barcelona.

The journey of writing a doctoral thesis involves many highs and lows and needless to say they could not have been traversed but for the continued support and love of Emmy; as always thank you for everything. My years in Barcelona could never have been the same if not for Chris, Gary, Giorgio, Claudia, Julia and Francois. A special thanks for making this journey memorable.

Contents

List of Figures

List of Tables

General introduction

The thesis examines three topics. The first one titled "Language use in education and primary schooling attainment: Evidence from a natural experiment in Ethiopia" looks at the effect of language policy choices on educational outcomes applying a difference in differences approach. Ethiopia introduced mother tongue instruction in primary schooling for the largest ethnic group in 1994. Using the fact that the exposure to the language policy change was jointly determined by the date of birth, language group and the region of residence, we estimate the causal effect of the provision of mother tongue instruction. Our results show that provision of mother tongue education led to an increase of 1 year of primary schooling and 1.18 years of schooling in the affected cohort, which is $\frac{1}{3}$ of the mean years of schooling in the country. Moreover the entire increase in the years of schooling can be attributed to the intensive margin of education. The language policy change increased the percentage of people completing 6 years or more of schooling by 12%. Applying our findings to a set of African countries suggests that introduction of mother tongue instruction increases the percentage of population completing primary schooling by as much as 8% points. These findings have important policy implications at a time when there is need for solutions which can increase the quality of education without requiring huge capital or infrastructural outlays.

The second chapter of the thesis, titled "Discrimination Without Taste - How Discrimination Can Spillover and Persist," co-authored with *Christopher Rauh*, introduces coordination failures driven by beliefs as a channel of persistent discrimination in activities characterized by interdependency in payoffs and interlinkages across markets. The model shows how discrimination can persist under perfect observability of individual ability, when taste for discrimination has died out, and in the absence of discriminatory social norms, i.e. under weaker conditions than normally assumed in the literature. For the case of self-employment the main result shows that individuals of the discriminated group will have lower participation rates and payoffs despite equal ability, leading to an overall welfare loss. Empirically we find that beliefs about discrimination are a significant factor in the estimation of self-employment rates of blacks in the US. The estimate suggests that reducing the share of people who believe that discrimination exists by 10% points would increase self-employment rates among blacks from 7.3% to 9.3%, an increase of 28%. The results are robust to the inclusion of year and region fixed effects and a variety of controls. The paper highlights the importance of distinguishing between the various channels of discrimination, as understanding the mechanism behind it is crucial to devising appropriate policy.

The last chapter co-authored with *Ashwini Deshpande* and titled "How Backward are the Other Backward Classes? Changing Contours of Caste Disadvantage in India" traces the socio-economic evolution of the Other Backward Classes (OBCs) since independence in India. While there is a growing literature on the political rise of the OBCs in India, where they are often seen as the new elite or the dominant castes, detailed empirical assessments of their socio-economic condition are practically non-existent. Using individual-level data from the National Sample Survey for 1999-2000 and 2009-2010, our paper is one of the first to undertake a comprehensive empirical exercise, both at the national as well as

the regional levels. We compare five age-cohorts, born between the years 1926-85, for the OBCs, SC-STs and Others' (everybody else) and examine the differences in key indicators such as educational attainment, occupation and activity status, wages and consumption expenditure through a difference-in-differences method. Our results show clear disparities in virtually all indicators of material well-being, with Others' at the top, SC-STs at the bottom and OBCs in between. We find evidence of convergence between OBCs and Others' in literacy and primary education, but divergence when higher educational categories are considered. In the realm of occupation, the younger cohorts among OBCs seem to be closing the gap vis-a-vis the Others' in terms of access to prestigious white-collar jobs. Finally comparing wage gaps for males in the labour force and estimates of labour market discrimination, we find that while average wages of Others' are higher than those for OBCs for all age cohorts, the unexplained (or the discriminatory) component is lower for younger OBC cohorts, compared to the older ones, and that OBCs face lower labour market discrimination compared to SC-STs, when the average wages of both groups are compared to those of Others'.

Chapter 1

Language use in education and primary schooling attainment: Evidence from a natural experiment in Ethiopia

1.1 Introduction

Cognitive theories suggest teaching in a non-native language at the primary schooling level may have a negative effect on educational outcomes. In Sub-Saharan Africa, with the exception of Tanzania and Ethiopia, no country provides the entire span of primary schooling in a local language, let alone the mother tongue. Africa is also characterised by some of the highest repetition and drop out rates in the world.[1] Whether the mismatch between the language of instruction and language used at home can partly help explain the poor educational outcomes observed in the African continent is a vital question with potentially

important policy implications, which so far has not been wholly addressed by economists.

The existing literature (Angrist and Lavy 1999, Angrist et al. 2008), analysing effects of policy changes in the language used to teach children, has considered changes occurring only at the secondary schooling level. The language of instruction might have more important implications at the primary rather than the secondary schooling level, as mother tongue instruction in the early years of childhood might be necessary to avoid cognitive disadvantages in children (Cummins 1978a, 1979, 1981, Verhoeven 1994). Moreover, these studies are not primarily concerned with the implications of the change in language policy on educational attainment. Angrist and Lavy (1999) concentrate on the effect of the language policy change on labour market outcomes. They find that replacing French with Arabic as the language of instruction, from the 6th grade onwards, led to a substantial reduction in the returns to schooling in Morocco. In our setting as the medium of instruction for secondary schooling and higher education continues to be English, we have reason to believe that such an effect on labour market outcomes might not be relevant for Ethiopia.[2] Angrist et al. (2008) are interested in the effect of the policy change in Puerto Rico, where Spanish replaced English as the language of instruction in secondary schooling, on English language skills and not on educational outcomes. They find that replacing English with Spanish in secondary schooling did not adversely affect English language skills of the individuals exposed to the policy change. A second strand of literature, related to ours, exploits the presence of bilingual programs to estimate the effects of provision of mother tongue education on student achievement (Matsudaira 2005, Chin et al. 2011, Slavin et al. 2011). However, these studies differ from our work in two crucial aspects, the first being that they have been primarily conducted in the context of the United States, where the exposure to the majority language for the language minority students is much higher as compared to in Africa. The difference in exposure to the language at the community

level might be an important factor, as to why results from bilingual studies based in the United States might not be applicable to the African context. The second reason being that the outcome of most bilingual programs have been measured after a period of just 3 to 4 years, whereas the benefits of bilingual schooling may appear often after as late as the 5th grade.[3] This paper aims to fill this gap by providing the first causal estimates of the impact of provision of mother tongue instruction on long run educational attainment, in the context of Sub-Saharan Africa.

We use a change in the medium of instruction policy in primary schooling in Ethiopia to estimate the effects of the change on the years of schooling.[4] Since 1962, the language policy was characterised by the exclusive use of Amharic in primary schooling. Amharic is the language of the second largest ethnic group, the Amharas. In 1994, mother tongue instruction in primary schooling was also introduced for the Oromo people; the largest ethnic group in the country. English was and remains the language of instruction in secondary schooling and higher education.

In order to estimate the effect of the language policy change, we use the fact that an individual's exposure to the new language policy differed by the language group, date of birth and region of residence in the country. We use a difference in differences (D-I-D) estimator that controls for systematic variation in education, both across language groups and cohorts. As Duflo (2001) notes, D-I-D estimators are especially well suited to evaluate policy changes when the reform comes from a variation in a well defined input; in our case the input being a change in the language used to teach children in primary schooling. The design of the empirical exercise allows us to show that the control and the treated groups have similar trends in education attainment before the language policy change. We also show that among the earlier cohorts, who did not benefit from the language policy change, as they had already finished primary schooling, the increase in educational attainment

from one cohort to the other is not correlated to the language policy change in the country. Showing a discontinuity in the educational trend in the first year in which the policy is implemented helps address other competing explanations, such as, omitted changes in schooling or regional conditions, as we would expect such changes to affect also other students who were still in schools but unaffected by the language policy change. Comparing the same language group in the treated and untreated regions in the country, also helps us address the concern that factors, such as, recognition of language rights, increase in group status or differential response of language groups, might be confounding our results. Other possible explanations are also explored to try argue that the effect being found is indeed a result of the provision of mother tongue instruction. The primary focus is on the effect of the language policy change on years of primary schooling and years of schooling. The main result shows that the provision of mother tongue instruction led to an increase of around 0.75 to 1 year of primary schooling and 0.80 to 1.18 years of schooling in the affected cohort, which is $\frac{1}{3}$ of the mean and standard deviation of the years of schooling in the country.

The results show that the change in language policy did not have any significant effect on the enrolment rates and hence the entire increase in the years of schooling is primarily due to the intensive margin of education.[5] The provision of mother tongue instruction, conditional on enrolment, increased the percentage of the sample completing 6 years or more of schooling by 12%. In fact the analysis shows that the provision of mother tongue instruction had a positive effect at all levels of schooling.

Applying our findings of the effect of provision of mother tongue instruction, on primary schooling completion rates, to a set of five African countries suggest potentially large benefits. Estimates suggest that the percentage of population, aged 15 to 49, which completes primary schooling could increase by around 4 percentage points from 29% to 33% in Benin,

7

from 59% to 63.5% in Cameroon and from around 70% to 78% in Ghana. Combining conventional estimates on the rate of return to an additional year of schooling (Psacharopoulos 1994, 2004) with the cost of producing learning materials in local languages suggest that even when these costs are taken into account the potential gains from the introduction of mother tongue instruction remain large. Our findings have important policy implications at a time when increasing budget constraints coupled with surging enrolment rates in Africa imply the need for remedial tools which do not require large capital and infrastructural outlays, but at the same time can increase the quality of education provided.

The rest of the paper is organized as follows. In the next section we describe the language policy change and education trends in Ethiopia. Section III outlines the identification strategy and presents the data. Section IV presents the main results and section V provides robustness checks. Section VI identifies the channel through which language policy works. Section VII discusses the implications of providing mother tongue instruction in other African countries and Section VIII concludes.

1.2   Language and education policy in Ethiopia

Ethiopia is situated in the horn of Africa and with a population of around 80 million it is the second most populous nation in the continent. The population is highly diverse containing more than 80 different ethnic groups. There are more than 90 different languages spoken in Ethiopia and most belong to the Afro-Asiatic language phylum of which three branches are represented, namely, Semitic, Cushtic and Omotic.

The "Oromo" are the largest ethnic group in the country comprising around 33% of the population. Though the most numerous in the country, Oromos can be considered as a minority in terms of political and economic influence they wield as a group both currently and over the course of Ethiopian history.[6] The "Amharas" comprising around 27% of the

8

population are the second largest ethnic group in the country. The other major ethnic groups in the country are the Somali, Tigray, Sidama and Wolaita comprising 6.2%, 6%, 4% and 2.3% of the population, respectively. "Oromigna", the language spoken by the Oromo people belongs to the Cushitic branch and is written using the Latin alphabet. In contrast "Amharic", spoken by the Amhara people, belongs to the Semitic branch and is written using the script Amharic Fidel. The two languages split at the first branch of the Afro-Asiatic language phylum.[7]

1.2.1 Language policy

Ethiopia was a monarchy for most of its modern history and under the imperial rule of Haile Selassie between 1916 and 1974. The medium of instruction policy during the imperial period involved the sole use of Amharic in primary schooling followed by the use of English as the medium of instruction for secondary schooling and higher education.[8]

The emperor was replaced by a Soviet-backed Marxist-Leninist military junta, the "Derg" (meaning council in Amharic), which came to power in 1974. The medium of instruction policy during the "Derg" regime, between 1974-1991, was identical to the one in place during the imperial time. Amharic continued to be the sole medium of instruction in primary schooling followed by the use of English from secondary schooling onwards.

Growing discontent against the "Derg" regime led to the establishment of the Tigray People's Liberation Front (TPLF) in 1975, which merged with other ethnically based opposition parties such as the Oromo Liberation Front (OLF) and the South Ethiopian People's Democratic Coalition (SEPDC) to form the Ethiopia's People's Revolutionary Democratic Front (EPRDF).[9]

The movement against the Junta government reached its peak in May 1991 when the

EPRDF forces stormed Addis Ababa and the "Derg" regime was toppled. The vision of political and cultural autonomy for all ethnic groups, the banner under which the coalition forces had come together, meant that the transitional constitution of 1991 and the constitution of 1994 resulted in the creation of a federal republic. The country was divided into nine regions and two city administrative units along ethnic and linguistic lines. The 9 regions are Afar, Tigray, Oromo, Amhara, Somalia, Benishangul-Gumuz, Southern Nations, Nationalities, and People's Region (SNNPR), Gambella and Harari. The two city administrative units are Addis Ababa and Dire-Dawa, respectively.

The vision of the rights to self-determination led to the introduction of mother tongue instruction in primary schooling for the four major ethnic groups Oromo, Tigray, Sidama and Wolaita starting 1994. Mother tongue instruction in primary schooling for the remaining smaller ethnic groups was to be slowly introduced in the course of the next years. The language policy post 1994 is still characterized by the exclusive use of English as the medium of instruction for secondary schooling and higher education. Pupils from the ethnic group other than the Amharas learn Amharic and English as a subject during the course of primary schooling, while those from the Amhara group only learn English as a subject during primary schooling.

[Insert Table 1.1]

Table 1.1 shows the implementation of language policy by the languages introduced as a medium of instruction in the nine regions and two city administrative units in the country. As can be seen in Table 1.1 the Oromo people entering primary schooling after 1994 gained access to mother tongue instruction in the regions of Amhara, Dire-Dawa, Harari and Oromia.

10

1.2.2 Education policy and trends

During the imperial time, formal schooling remained mostly an urban/semi-urban phenomenon with little or no schools in the rural areas. In the period after the imperial rule, between 1975 and 1989, enrolment increased by around 12%. However lack of investment in education meant that schooling remained out of reach of most rural people.

Post 1994 the education sector was given renewed importance by the EPRDF. The budget for education increased steadily since 1996/97 and in real terms by around 50% in the five-year period after 1995/96. The focus on expanding access to education resulted in almost doubling primary schooling enrolment from 4.5 million in 1996/97 to 8.1 million students in 2001/02. In 2001/02 the total education spending stood at US$ 333 million and was 14% of total public expenditure. Two-thirds were spent on primary and secondary schooling, and the balance on technical and university education.

In per capita terms, however, there had been little increase in education spending. The total expenditure per student increased only by around 5% between 1995/96 to 2001/02. Moreover, when total expenditure is broken down by the components of recurrent and capital expenditure, the main picture that emerges in the words of the 2004 public expenditure review of Ethiopia by the World Bank is the "insufficiency of spending at all levels" (pg. 15). The recurrent expenditure per student at the primary level has decreased by around 20% in real terms over the five-year period of 1996/97 to 2001/02. The share of wages and salaries in the recurrent budget was around 97%. Despite this the pupil teacher ratios (PTR) have steadily deteriorated over the period 1995/96 to 2001/02. The PTR have increased from 32:1 and 33:1 in 1995/96 to around 73:1 and 80:1 in 2001/02 in primary and secondary schooling, respectively. Twenty nine percent of the enrolled population drop out by grade 1 and 55% of the enrolled population by grade 3, implying often for these children lifelong literacy is not achieved.

11

1.3 Identification strategy and data

1.3.1 Identification strategy

As with many policy evaluations, the difficulty in estimating the effect of the provision of mother tongue schooling on education attainment is the lack of a counterfactual. To circumvent this problem the paper employs a D-I-D approach and uses the fact that the date of birth, the language group and the region of residence of an individual jointly determine the exposure to the language policy change.

The main exercise involves comparing the mean years of primary schooling/years of schooling of the Amhara and Oromo group for the older and younger cohorts. For our purpose we define the younger cohort as comprised of individuals who enter primary schooling after the language policy change. The older cohort is defined as comprising of individuals who have finished primary schooling by the time the language policy change was implemented. The older and younger Amhara cohorts had access to mother tongue instruction both before and after 1994, are not directly affected by the policy change, and form our control group of interest. The younger Oromo cohort however gained access to mother tongue instruction after 1994, and comprise our treated individuals. The difference in mean years of primary schooling between the two language groups, for the older cohort is compared with the difference of the younger cohort. This difference in differences can be interpreted as the causal effect of the policy change on the Oromo people, under the assumption that, in the absence of the change in the medium of instruction policy, there would have been no change in the pattern of similar trends in educational attainment for the two groups.

As Duflo (2001) and Strauss and Thomas (1995) note, when a fixed effects estimator is being used to assess the effect of a policy change, the investigator should pay close attention

to the validity of the identifying assumption. The estimate may be capturing something due to the pre-existing differential trends for the two groups or due to the divergence in trends in the post policy period for reasons unrelated to the policy change. Also if the increase in education of the two groups was negatively correlated with the initial levels, then a treatment effect might be observed even if the program had no real effect.

We try and closely analyse the validity of our identifying assumption. The test of our identifying assumption exploits the presence of multiple groups formed by successive cohorts not exposed to the policy change (Duflo 2001, Heckman and Hotz 1989, Rosenbaum 1987). The individuals belonging to the Oromo group, who entered primary schooling before 1994, were not affected by the change in the language policy, and hence we should not expect the educational attainment to vary systematically across the Amharic and the Oromo group for the older cohorts. Figure 1.1 plots the trend in mean years of primary schooling for three older cohorts for the two groups, before the policy change in 1994. The fixed effects estimator allows for the levels across the two comparison groups to be different as long as the trend or the shape of the curve remains the same. We see that the two groups over the 30-year period, of 1964 to 1994, follow parallel trajectories and have very similar trends in primary schooling attainment. In the results section, using a D-I-D estimator, we formally show that the two groups have similar trends before the policy is implemented.

[Insert Figure 1.1]

The data allows for other potential designs to estimate the effect of provision of mother tongue education, though for reasons discussed below we believe the strategy of comparing the Amharas to the Oromos might be the best way to minimize potential estimation bias. The Oromo people gained access to mother tongue education in 4 of the 11 regions in the country. An alternative estimation strategy could involve comparing the older and

younger cohorts of the Oromo group in the treated and untreated regions of the country. This would involve comparing the same language group in different regions whereas our main experiment compares different language groups but in the same regions. There exist disparities in access and schooling infrastructure across regions with the south-western and northern parts of the country being the least developed. In order to minimize the concerns that our results are driven by differential access to schooling infrastructure, we prefer our main experiment to the one comparing the Oromo group in the treated and untreated regions of the country. This said in Section V, we do carry out the exercise of comparing the Oromos in the treated and untreated regions and show that the results obtained are very similar to the ones obtained with our preferred identification strategy.

As noted before, the language policy change involved the introduction of mother tongue instruction for the four major ethnic groups, namely, Oromo, Tigray, Sidama and Wolaita, starting 1994. The Tigray people gained access to mother tongue instruction in only the Tigray region of the country. There is a minimum of 4 to a maximum of 10 observations from the control group Amhara in the Tigray region, implying a difference in differences strategy comparing the Amhara and Tigray groups is not implementable.

The data only allows us to distinguish the population of Ethiopia into four distinct language groups. These are namely the Amhara, Oromo, Tigray and the Others'. The Others' category includes all the remaining language groups clubbed together. As mother tongue instruction for the language groups besides the Oromo, Tigray, Sidama and Wolaita was introduced over the next years after 1994, most individuals who have finished schooling by 2011 from the language group besides the Sidama and Wolaita in the Others' category are untreated in our data. As the data does not allow us to uniquely identify these individuals, comparing the Others' category to the Amhara would provide a lower bound of the real effect of the language policy change. We however create a group called the Non-Amhara,

comprising of individuals from all other language groups besides the Amhara, and also consider the group Others', and carry out a difference in differences strategy comparing the Non-Amhara and Others' to the Amhara group, the results of which are presented in the appendix.

1.3.2   Data

The data comes from the Demographic and Health Surveys (DHS), which are nationally representative data on health, education and demographic trends in developing countries. The data for Ethiopia are from the year 2011 and include information on a nationally representative sample from the 9 regions and two city administrative areas of Ethiopia. Figure 1.2 shows a map of Ethiopia, where the regions and city administrative borders are demarcated. The number of observations in our analysis range from a minimum of 3,210 to a maximum of 11,918 observations.

[Insert Figure 1.2]

The schooling system in Ethiopia involves 8 years of primary schooling followed by 4 years of secondary schooling. The children in Ethiopia normally go to primary school between 7 and 14 years of age. A child born before 1980 and speaking Oromo as her mother tongue was 14 years old in 1994, had already finished primary schooling, and was unaffected by the change in the medium of instruction policy. The data being from the year 2011 implies that the youngest individual who could finish secondary schooling by 2011 was 2 years old in 1994. Similarly the oldest individual who could have been affected by the policy change was 7 years old in 1994. We hence consider the individuals aged 2 to 7 years in 1994 as the younger cohort and the Oromos aged 2 to 7 years in 1994 as the treated individuals.

15

The children speaking Oromo as their mother tongue and who were aged 8 to 12 in 1994 were already in primary school before the policy change was implemented. As the policy involved a change in the medium of instruction, the policy in general was implemented only for children who entered primary schooling from 1994 and thereafter. Implementation of mother tongue instruction in some schools for individuals already in primary schooling and grade repetition and delayed school entry could lead to some of the children aged 8 to 12 in 1994 to benefit from the program. Considering them as completely untreated would then provide us with a lower bound of the true effect of the program. Thus the cohort aged 13 to 20 in 1994, individuals who were completely unaffected by the policy change, are defined to be the older cohort for our purpose.[10] We, however, do show that considering the untreated individuals who were aged 8 to 12 years old in 1994, and were in primary schooling when the policy change happened, leaves our results unchanged. We use the information on age of the individual in 1994, along with his language background i.e. the language spoken as the mother tongue and their region of residence to match it with data on medium of instruction policy implemented in various regions of the country (shown in Table 1.1) to ascertain the impact of provision of mother tongue instruction on the Oromo language group.[11] The descriptive statistics for the younger and the older cohort are shown in Table 1.2.

[Insert Table 1.2]

1.4 Results

1.4.1 Comparison of means

Panel A labelled experiment of interest in Table 1.3 presents the main experiment. There are a total of 5,364 observations with the treated comprising 1,327 observations. Compar-

ing the two groups, for the older cohort aged 13-20 in 1994, shows that the average Amhara and Oromo individual had 2.89 and 1.82 years of primary schooling, respectively. The 1.07 years advantage, for the average Amhara, is due to the Amhara group having had access to mother tongue instruction before 1994 and having been the economic and politically dominant group in the country for the larger part of the 20th century. We observe that the mean years of primary schooling of the younger cohort, aged 2 to 7 in 1994, increases for both the groups due to the spurt in enrolment post 1994. The average years of primary schooling for an Amhara and Oromo individual of the younger cohort stands at 4.16 and 3.81 years, respectively. After gaining access to mother tongue instruction the difference between the average Amhara and Oromo reduces from 1.07 to 0.35 years of primary schooling. The D-I-D or reduction in gap of 0.72 years of primary schooling can be considered as the causal effect of the language policy change. In the next subsection we will check for the statistical significance of the D-I-D calculated.

[Insert Table 1.3]

The causal interpretation as noted before depends on the identifying assumption of similar trends in primary schooling for the two groups in the absence of the policy change. The panel B in Table 1.3 formally presents our control experiment. We consider two cohorts, aged 13-20 and 21-28 in 1994, not exposed to the policy change from the two groups. The identifying assumption of similar trends should imply that the D-I-D should be equal to zero. The panel B indicates that the D-I-D is equal to -0.07 and very close to zero. In the next subsection we show that the difference in difference calculated is indeed statistically insignificantly different from zero.

17

1.4.2 Basic difference in differences regressions

In order to evaluate the difference in differences estimator we run the following reduced form regression:

$$S_{ijkn} = \delta_0 + \delta_1 * D_j * C_k + \delta_2 D_j + \delta_3 C_k + \delta_4 B_k + \delta_5 R_n + \epsilon_{ijk} \tag{1.1}$$

S_{ijkn} refers to the years of primary schooling of individual i, from language group j, of cohort k and in region n. D_j is a dummy variable taking the value 1 if the individual belongs to the Oromo language group and zero otherwise. C_k is a dummy variable which takes the value 1 if the individual belongs to the cohort which was aged 2 to 7 in 1994 and zero otherwise. B_k is a vector of year of birth dummies for the individuals aged 2 to 7 and 13 to 20 in 1994 for each year of birth and R_n is a vector of region dummies.

The results of the main experiment are shown in panel A of Table 1.4. Column (1) does not control for year of birth or region dummies. The language group dummy, which captures the difference in level between the two groups, is negative and significant at the 1% level. It captures the fact that the average Amhara has 1.07 more years of primary schooling as compared to the average Oromo. The cohort dummy captures the time trend of increasing years of primary schooling in the country, and is positive and significant at the 1% level. The coefficient shows that on an average the younger cohort has 1.27 years more of primary schooling due to the increased enrolment post 1994. The main coefficient of interest is δ_1, the one associated with the interaction term between the language and the cohort dummy, and it captures the effect of provision of mother tongue instruction. The calculated D-I-D of 0.72 years is statistically significant at the 1 % level. Column (2) additionally controls for year of birth and region dummies. Additionally controlling for these does not change the significance and in fact increases the value of the point estimate of δ_1. This shows that

18

the provision of mother tongue instruction increased years of primary schooling by 0.75 years in the affected cohort. In column (3) the dependent variable considered is the mean years of schooling instead of mean years of primary schooling, and includes as controls the region and year of birth dummies. The coefficient capturing the effect of the provision of mother tongue instruction is positive and significant and shows that provision of mother tongue instruction increased average years of schooling by around 0.80 years. Comparing the increase with the mean and the standard deviation of the years of primary schooling, for the cohorts aged 2 to 7 and 13 to 20 in 1994 in these 4 regions, implies an increase of about $\frac{1}{3}$ of the value of the mean and standard deviation, respectively. Comparing the increase to the mean years of primary schooling for the entire country implies an increase of around 60%.

[Insert Table 1.4]

The panel B of Table 1.4 presents the results of our control experiment or the placebo test. Column (1) does not control for year of birth or region dummies and shows that the D-I-D of -0.07 years is indeed insignificantly difference from zero. Column (2) controls for the year of birth and region dummies. Additionally controlling for these reduces the size of the point estimate to -0.009 and it remains statistically insignificant. The results of panel B in Table 1.4 provide evidence in support of the assumption that in the absence of the policy change there would have been no divergence in the trend for years of schooling for the two groups.

Comparing the size of the coefficients of the two interactions terms, in panel A and B of Table 1.4, shows that the coefficient in the experiment of interest is about 100 times the size of the coefficient in panel B. Moreover the associated standard errors are nearly identical implying that the insignificant coefficients found in the control experiment are not due to the issue of lower precision in the estimation of the placebo test.

19

1.4.3 Estimating the effect of the policy change for every cohort

The analysis carried out in the previous section compares the cohort aged 13 to 20 in 1994, who are completely untreated, to the cohort aged 2 to 7 in 1994 and who are completely treated. In this subsection we extend our identification strategy to a generalized interaction term analysis to take into account the effect of the policy change for each cohort aged 2 to 21 in 1994.

The relationship between the education (S_{ijkn}) of an individual i, from language group j, in year k, of region n and their exposure to the language policy change can be expressed as follows:

$$S_{ijkn} = \delta_0 + \sum_{l=2}^{l=21} (D_j * d_{il})\delta_{1l} + \delta_2 D_j + \delta_3 B_k + \delta_4 R_n + \epsilon_{ijk}, \tag{1.2}$$

where d_{il} is a dummy that indicates whether individual i is of age l in 1994. The other variables have the same interpretation as in equation (1.1). The omitted dummy category is the individual aged 21 years old in 1994. Each coefficient δ_{1l} can be interpreted as the effect of the language policy change on a given cohort of the Oromo language group. Because children aged 13 and older in 1994 did not benefit from the introduction of mother tongue instruction implies that δ_{1l} should be equal to 0 for $l \geq 13$. We additionally know that all individuals aged 7 and younger in 1994 were exposed to the language policy change implying δ_{1l} should be greater than 0 for $l \leq 7$. The pupils, who were already in primary schooling i.e. between the ages of 8 to 12 in 1994, could have partially benefitted due to implementation of mother tongue instruction in some schools for individuals already in primary schooling or as a result of grade repetition and delayed school entry. The only a priori restriction for $8 \leq l \leq 12$ is that δ_{1l} is greater or equal to zero.

In panel A of Figure 1.3 we plot the estimated coefficients, $\hat{\delta_{1l}}$. Each dot on the solid line corresponds to the coefficient of interaction between the dummy for whether individual i is

20

of age l in 1994 with the Oromo language dummy (the 95% confidence interval is plotted in dashed lines).[12] Each dot thus summarizes the effect of the language policy change on the cohort aged l in 1994 and belonging to the Oromo language group in the treated regions. These reduced form estimates for each year of birth allows us to verify whether $\hat{\delta}_{1l}$ follows the pattern implied by the assumption underlying the identification strategy. As can be seen, these coefficients fluctuate around zero and are statistically insignificantly different from zero for all ages between 20 and 8 and start increasing for ages below 8.

[Insert Figure 1.3]

Panel A in Figure 1.3 shows a discontinuity for the coefficient, $\hat{\delta}_{1l}$, of the cohort aged 7 in 1994. It is seen that the first cohort to benefit from the language policy change has a break from the prevailing trend and is the first coefficient, which is statistically different from zero. The above shows that the policy change did not have any effect on education of cohorts not exposed to it and had a positive effect on the education of all younger cohorts.[13] As Bertrand et al.(2004) stress that one of the factors that has been often overlooked in the use of difference in differences estimators is the problem of serial correlation among the errors. In order to deal with the problem of serial correlation we cluster errors at the level of 231 household clusters. The results shown in column (2) of Table 1.11 in the appendix shows that the clustering errors leaves our results essentially unchanged.

The observed pattern also helps us address the concerns that our estimate may be capturing something due to the omitted changes in schooling and regional conditions. If other omitted changes in schooling conditions or other policy variables were driving our estimate, we would expect it to affect also individuals who were already in primary school or also maybe in secondary schooling.

In the next section to address the concern that it is not the differential response of a particular language group to the general emphasis given to education sector post 1994, or

the recognition of language rights, or increase in group status that is driving our results, we compare the same language group in the treated and untreated regions of the country. We also additionally explore other potential explanations that could be put forth to explain our findings to argue that the effect being found is indeed the result of the language policy change.

1.5   Robustness tests

1.5.1   Comparing the Oromo's in the treated and untreated regions

Post 1994 the Oromo people gained access to mother tongue instruction in 4 of the 11 regions in the country. As a first robustness test we compare the Oromos in the treated and untreated regions of the country. This exercise involves comparing the same language group in different regions whereas the main experiment involved comparing different language groups but in the same regions. There are a total of 3,210 observations with the treated comprising 1,327 observations.

Table 1.5 compares the mean years of primary schooling of the Oromos for the older and the younger cohorts in the treated and untreated regions of the country. The older cohort in the untreated region have on an average 1.22 more years of primary schooling as compared to the older cohort in the treated regions. This is due to the untreated regions including the capital Addis Ababa; the most developed part of the country. After gaining access to mother tongue education, the gap between the average individual in the untreated and treated regions reduces to 0.47 years of primary schooling. The D-I-D suggests that the provision of mother tongue instruction increased average years of primary schooling by 0.75 years in the affected cohort, very similar to the increase of 0.72 years found in Table 1.3. Looking at the Oromo untreated individuals, we see that the younger cohort

gains around 1.24 years of primary education. The increase for the non-treated individuals from the Oromo group is very similar to the increase of 1.27 years observed for the untreated younger Amhara cohort in Table 1.3, although we are looking at two different groups in different regions of the country.

[Insert Table 1.5]

Table 1.6 shows the results of formally evaluating the D-I-D estimator. Column (1) does not control for year of birth or region dummies and shows that the difference in difference of 0.75 years found through comparison of means is statistically significant at the 1% level. Column (2) additionally controls for year of birth and region dummies. Additionally controlling for these increases the size of the point estimate to 1.00 and it remains statistically significant. In column (3) the dependent variable considered is the mean years of schooling instead of mean years of primary schooling, and includes as controls the region and year of birth dummies. The estimate shows that the provision of mother tongue instruction increased average years of schooling by around 1.18 years.

[Insert Table 1.6]

We now as in the Section IV.3 estimate the effect of the policy change for each cohort aged 2 to 21 in 1994. The results are presented in panel B of Figure 1.3.[14] Each dot on the solid line in panel B of Figure 1.3 plots the coefficient of the interaction between a dummy for being a given age in 1994 and the regional dummy, which takes the value one for the regions in which the Oromos were treated. Each dot tells us the effect of the language policy change for a particular cohort for the Oromo people in the treated regions of the country. We again observe a very similar pattern. All the coefficients are very close to zero up until the cohort aged 9 in 1994.[15] The graph has a break in trend for the cohort aged 7 in 1994, the first cohort to be exposed to the language policy change in the country.

The observed break in trend again helps address the concern that it is not the effect of other changes in schooling or regional conditions which is confounding our estimate. Moreover comparing the same language group also helps us rule out the concern that it different groups reacting differently to the general expansionary trend in education that was taking place in Ethiopia, or it is the effect of political changes such as recognition of language rights, or increase in group status that is being captured by our estimate.

In the appendix in Table 1.13 and 1.14 are shown the results of comparing the Non-Amhara to the Amharas. In Table 1.15 are shown the results of comparing the Others' category to the Amharas, in the region where the Sidama and Wolaita gained access to mother tongue instruction. The coefficient again shows that the provision of mother tongue instruction increased average years of primary schooling by around 1 year. These show that the results are robust to considering other language groups and regions of the country. Finally in Table 1.16 is shown the effects of the language policy change by the category of gender, where again we compare the Amharas to the Oromos as in the original experiment of Table 1.3 and 1.4. The estimates suggest an increase of around 1.03 and 0.60 years of primary schooling for boys and girls, respectively.

1.5.2 Other potential confounding factors

Change in the composition of teachers

A potential explanation driving the results could be that it is not the change in the language used to instruct children but the change in the composition of teachers. The existing literature has found weak effects of race, gender and ethnic composition of teachers on student outcomes, and furthermore that these are more likely to matter more for subjective evaluation rather than objective performance of students (Ehrenberg et al. 1995, Klein et al. 2001, Dee 2005). In our context, due to lack of data, we only provide some suggestive

evidence as to why change in the composition of teachers is not likely to affect our result. As we noted in section II, the primary school enrolment rose from around 4.5 million to around 8.1 million and at the same time the pupil-teacher ratio (PTR) increased from around 32:1 to around 73:1. This seems to suggest that as enrolment doubled so did PTR implying that no new teachers were hired. It is interesting to note in this regard that teachers in public schools in Ethiopia are public sector employees, making hiring and firing decisions quite rigid. This potentially seems to suggest that there were no major changes in the composition of public school teachers.[16]

Changes in curriculum or lowering of standards

Another potential explanation that could be put forth is that the effect being found is not due to the provision of mother tongue education but is the effect of a change in the curriculum faced by the Oromos. In this regard it should be noted that the curriculum is designed at the federal level and the regions are responsible for adapting the curriculum to the regions socio-cultural and economic specificities. The report of the Joint Review Mission (JRM) of the Education Sector Development Programme III notes "*Although the JRM did not explore this issue in depth, the impression is that the extent of adaptation which is carried out in practice is relatively limited (mostly translating and adapting examples used)*" (pg.35). The above seems to suggest that the curriculum faced by the different language groups differ only in the language used to prepare the primary school textbooks and not in the content per-se.

Another possible channel which could explain the effect found is that the change in the language policy led to a reduction in standards and children from the language groups which gained access to mother tongue education are now more easily promoted. This would show up in the data as individuals gaining more years of education but in fact this effect

would be actually due to a reduction in schooling quality. This however should imply that the pupils who gain from such a policy should perform worse, than the control group Amharas, in the standardised national assessment tests conducted at the end of primary schooling. However looking at the standardised assessment results from the year 2004 show that this is not the case. The Amhara pupils, in the Amhara region, have a composite score in the subjects of English, Mathematics, Biology, Chemistry and Physics equal to 43, whereas at the same time for the students from the Oromia region, where Oromigna is being used, the composite score is 43.2. Additionally as we show in the next section that the language policy change actually increases the percentage of population completing primary schooling among the treated Oromos by around 30%, as compared to the Amharas. This implies that the change in the composition of the Oromos, if anything, results in now having more pupils from the lower end of the ability distribution taking the standardized national assessment tests. This suggests that correcting for any selection effects would in fact increase the composite score achieved by the Oromos. The fact that the performance of both the language groups without correcting for selection is very similar, seems to suggest that the result is not being driven due to the phenomenon of children being pushed through grades. Moreover as the Oromia region report of the JRM notes *"Teachers and parents are not aware that repetition, far from improving performance, only increases drop-out; nor are they aware of the difficulties of incorporating skills training at primary level"* (pg.32). The above quote if anything seems to suggest the opposite, that grade retention is more of a concern than that of the pupils being pushed through grades.

Differential access to schooling infrastructure

As we noted before the enrolment rates in the country increased sharply post 1994. Another potential confounding factor could be the differential access to schooling infrastructure. In

26

this regard it should be noted that in the main experiment we are looking at the treated and untreated individuals in the same regions, so the question of differential access to infrastructure is partly addressed by this approach. Moreover, our results are robust to restricting our sample to any combination of the four treated regions. Additionally, in the next section we show that there are no systematic differences in enrolment rates across the control and the treated group, after the policy change, and the increase in schooling primarily comes about due to the higher completion and lower dropout rates. This said it should be noted that indeed there could exist within regional variation in access to infrastructure but due to lack of data we are unable to account for this.

The identification strategy, the accompanying robustness tests and the exploration of other potential confounding factors seem to suggest that our causal interpretation provided to the language policy change in the country are reasonable and plausible.

1.6 Identifying the channel of increase in educational attainment

The estimates from the previous exercise show that the provision of mother tongue instruction led to an increase of 0.75 to 1 year of primary schooling in the affected population. The question that we try to address in this section is how much of the increase is due to the extensive margin and the intensive margin of education, respectively. For our purpose, the change in the years of schooling, associated with increased enrolment, due to the change in the institution of language policy, is defined as the extensive margin. The intensive margin is the change in years of schooling, associated with the people who would enrol irrespective of the choice of medium of instruction, but choose different levels of schooling, under the alternative scenarios.

The data allows us to identify whether the individual was ever enrolled in schooling or not. Using the same design as the main experiment in Table 1.3 and 1.4, we continue to compare the Amharas and the Oromos in the regions where the Oromos gained access to mother tongue instruction. We create a dummy variable enrolment equal to 1 in case the individual was ever enrolled in schooling and zero otherwise. Table 1.7 compares the two groups for the cohorts, aged 2 to 7 and 13 to 20 in 1994, where the dependent variable is the proportion of individuals from each group ever enrolled in schooling.

[Insert Table 1.7]

Comparing the cohorts aged 13 to 20 in 1994, we see that the proportion of individuals ever enrolled are 43% and 48% for the Amharas and the Oromos, respectively. The level of enrolment, as discussed in section II, increases sharply for the younger cohort, and stands at 70.9% and 71.1% for the Amharas and the Oromos, respectively. Calculating the D-I-D suggests that language policy resulted in increasing enrolment by around 5% points. We formally estimate the D-I-D estimator by using a Probit model to implement equation (1.1), where now the dependent variable is the dummy variable enrolment.[17] The marginal effects of the Probit regression are shown in column (1) of Table 1.8.[18]

The interaction term, capturing the increased probability of enrolment, due to provision of mother tongue instruction, is very close to zero and insignificant. The language policy seems to have had no significant effect on the probability of enrolment. The benefits of mother tongue instruction on cognitive development have not been well understood in most policy circles and by stakeholders such as parents, as can be seen with the widespread preference and continuing practice of using former colonial languages as a medium of instruction in most African countries. Moreover, the knowledge of the former colonial language or the dominant language of the country, are often seen as a prestigious mark of education and modernity in many countries in Africa, and is cultivated by many as a means to acquire

status.[19] Given the above two factors, it is not surprising that the provision of mother tongue instruction did not have any significant effect on the probability of enrolment for the Oromos.

The above exercise suggests that the entire increase in years of primary schooling can be attributed to the intensive margin. The literature on language, bilingualism and child development has highlighted the role of mother tongue instruction on cognitive development of children (Cummins 1978b, 1979, 1981, 1984, Wong Fillmore 1991, Skutnabb-Kangas and Toukomaa 1976). Although measuring cognitive skills has been a challenging issue, the literature has been able to demonstrate that various measures of cognitive skills are positively correlated with schooling, wages and labour market outcomes (Heckman et.al 2006). The availability of mother tongue instruction resulted in assisting the cognitive development of children, which reduced the cost of obtaining education, and hence increased the completion rates. The fact that the language policy works through the intensive and not the extensive margin can be reconciled by applying a theoretical framework of sequential schooling choice made under uncertainty (Altonji 1993, Zamarro 2004). The pupils have probability distributions defined over the cost of effort. Introduction of mother tongue instruction does not alter the probability distribution of the cost of effort, as it is a new technology over which individuals hold no beliefs. Hence the same sets of individuals enrol under the two language policy settings. However once enrolled, they find that the cost of effort is lower while studying in the mother tongue, leading to an updating of beliefs over the cost of effort, which in turn leads to pupils choosing more years of schooling. We should hence expect that in the future, as people become cognizant of the benefits of mother tongue instruction, the policy would have an effect on the extensive margin as well.

[Insert Table 1.8]

29

The cumulative distribution function (CDF), showing the proportion of total students dropping out at each grade, conditional on enrolment, for grades 1 to 8, is shown in Figure 1.4. Panel A shows the CDF for the two cohorts, aged 13 to 20 and 2 to 7 in 1994, from the Oromo group. We see that that the CDF of the younger cohort stochastically dominates the one of the older cohort. Comparing the younger with the older Oromo cohort shows, conditional on enrolment, 58% of the population now finishes 6 years or more of schooling, as compared to only 37% before. The panel B depicts the CDF for the two cohorts, aged 13 to 20 and 2 to 7 in 1994, from the Amhara group.

[Insert Figure 1.4]

We see that up until grade 5 the two curves almost overlap and after grade 5 in fact the CDF for the older cohort dominates the one of the younger cohort. This decrease in completion, conditional on enrolment, for the Amhara group can be attributed to the reduction in per capita recurrent expenditure per student happening in the country. The panel C in Figure 1.4 depicts the D-I-D of the CDFs. The dot on the 5th year of education, for instance, indicates that provision of mother tongue instruction induced 20% of the sample to complete 6 years of schooling or more as compared to 5 years or less. The curve indicates that the policy change had a positive effect at all levels of primary schooling.[20]

In order to formally estimate the effect of language policy on completion rates, we construct a dummy called "completion". The variable takes value 1 if the individual completed 6 years or more of schooling and zero otherwise.[21] We estimate the regression given by equation (1.1), where now the dependent variable is the dummy completion. This regression is estimated both for the entire sample and only for the individuals who were ever enrolled in schooling. The marginal effects for the restricted and the entire sample are shown in column (b) and (c) of Table 1.8, respectively. The coefficient on the interaction term, capturing the effect of provision of mother tongue instruction on probability

30

of completing 6 years or more of schooling as compared to 5 years or less, is positive and significant. The estimate suggest that the change in language policy, conditional on enrolment, resulted in inducing 12% of the sample to complete 6 years of schooling or more, as compared to 5 years or less, whereas the estimate for the entire sample is 9.2%.

The increase in completion rates is similar to the one found by Jackson (2000). He finds that the use of French in the first two years of primary schooling, instead of Kirundi in Burundi, led to an increase in the dropout rate from around 28% to 40%. Similarly Patrinos and Psacharopoulos (1995), based on a household survey in 1990 in Paraguay, indicate "language strongly influences school attainment and performance." They find that language was the single best predictor of repetition, and the cost of being a Guarani-only speaker is about one year of schooling attainment.

The World Bank (2004) study notes that one of the big problems facing Ethiopia is the large number of students dropping out before finishing grade 3. Dropping out at such early stages implies that for these children little or no effective education is taking place. The problem of high dropout remains prevalent throughout the African continent, implying many resources are spent on educating people who never achieve effective literacy. The use of foreign languages as a medium of instruction might be an important factor driving such high dropout rates observed on the African continent.

The finding that the use of mother tongue as a medium of instruction primarily works through the intensive margin also has other important policy implications. In Ethiopia and other African countries, the surge in enrolment rates has meant that per capita expenditure on students, especially recurrent education expenditure, has been declining. In such a context means of improving quality of education, which do not require massive infrastructural or capital investments are crucial policy tools to ensure effective education for all. As we noted in section II, in the years between 1995/96 and 2001/02, the recur-

rent education expenditure per student in Ethiopia declined by around 20% in real terms. Comparing the younger and the older Oromo cohort, we see that the average years of education obtained, conditional on enrolment, increased from 4.27 years to 5.37 years. In the case of the Amharas, the average years of education, conditional on enrolment, actually decreased marginally from 5.97 to 5.86 years. This small reduction as noted before can be attributed to the reduction in per capita recurrent expenditure per student. The fact the the Oromo people increased their educational attainment despite per capita recurrent expenditure going down, highlights the fact that language policy might be an important policy tool to increase quality and years of schooling in countries with scarce resources.

1.7   Applications to the African continent

In this section, we intend to explore how provision of mother tongue instruction could affect the proportion of population completing primary schooling, in the African context, and the net benefits arising from an additional year of schooling. It should be noted at the outset, that the task associated with extrapolating results found in one specific context, to other countries, comes with its many associated pitfalls and problems. We do not seek to claim that our estimates do not suffer from these problems, but look at the exercise as a way to shed some light and provide some benchmark estimates about the potential benefits of mother tongue education. First, using data from the DHS, for other African countries, we calculate how the provision of mother tongue instruction to all ethnic groups, which comprise 10% of the population or more, would change the percentage of population completing primary schooling.

In the previous section, we saw that the provision of mother tongue instruction, induced 12% of the enrolled sample to complete 6 years or more of education, as compared to 5 years or less. This is the key figure from the previous results that we will employ to

generate the required counterfactuals. The procedure utilized to calculate the change in the proportion of population completing primary schooling is explained in the appendix. The same procedure is applied to a set of five African countries and the results are shown in Table 1.9.

[Insert Table 1.9]

The results suggest that in Benin, where there exist three language groups with population shares greater than 10%, provision of mother tongue instruction to these groups could increase the percentage of population completing primary schooling, for the people aged 15 to 49, from around 29% to 32.5%, an increase of nearly 10% points.

In the case of Burkina Faso, there is only one language group with a population share of greater than 10%, the "Mossi", who comprise 56% of the population. The estimate suggests that the provision of mother tongue instruction, could increase the percentage of population aged 15 to 49 completing primary schooling, from 16% to 19% points for the Mossi, and from 15% to 17.6% for the country as a whole. In the case of Ghana, provision of mother tongue instruction, increases the percentage of population aged 15 to 49 completing primary schooling from 70% to 78% and finally in the case of Gabon and Cameroon the increase is from 66% to 73.2% and 59% to 63.5%, respectively. This application at the face of it suggests potentially large benefits where the percentage of population completing primary schooling increases by as much as 8% points. In what follows we calculate the net present value arising from the gain of one additional year of primary schooling due to the provision of mother tongue instruction.

In order to calculate the net present value of an additional year of schooling, the associated costs of introducing mother tongue instruction have to be taken into account. The task of estimating the costs of producing learning materials through standardised methodology is highly problematic, because this cost depends on a variety of factors such as the

33

state of development of languages to be used, population sizes, attitudes towards use of local languages etc. In what follows, we present some estimates of this cost based on existing studies. The main objective of the exercise is to show that even when these costs are taken into account, the potential gains from the introduction of mother tongue instruction remain large.

Patrinos and Vadwa (1995) analyse the production costs of introducing local language material in the context of Guatemala and Senegal. The estimates for Guatemala are based on 500,000 textbooks developed by Direccion General de Educacion Bilingue Intercultural (DIGBI), for the four majority Mayan languages. The authors estimate that the introduction of Mayan curriculum increased the unit cost of primary education by 9 percent, over the cost of Spanish-only curriculum. This however overestimates costs for the future years, as this includes the curriculum development costs, accounting for 37% of the total cost, which would not have to be borne in the later years. In the case of Senegal, the estimates suggest, whereas the cost of producing a French textbook is US$ 0.35, this increases to US$ 0.84 in the case of textbooks in Wolof. An important point to be noted is the estimates for cost per textbook for French is based on producing around 150,000 books, whereas for Wolof the number of books produced were only 4,140. The authors point out that the per unit cost would decrease significantly as the number of books produced increase, as the associated fixed cost per unit would decrease. They estimate that economies of scale in production can be achieved by printing around 10,000 books and in such a scenario there would be no difference in the cost of a French or a Wolof textbook. Using the above estimates we assume that in the first year there is an increase of 10% in per capita spending per pupil and from the year onwards there is no difference in the cost of provision of local or foreign language instruction.

In order to calculate the return to education, we use the latest estimates of the coeffi-

cient on years of schooling in the Mincerian wage equation from the work of Psacharopoulos and Patrinos (2004).[22] The figures on gross domestic product (GDP) per capita and expenditure per student as percentage of GDP per capita are taken from the World Bank indicators for the latest available year. The GDP per capita are measured in constant 2000 US$.

The net present value of one additional year of education, assuming an individual works for 30 years, is given by:

$$NPV_{ij} = \sum_{t=1}^{30} \frac{(m_j(GDPPC_j))}{(1+r)^t} - (0.10)(p_j GDPPC_j) - (p_j GDPPC_j) \qquad (1.3)$$

where NPV_{ij} refers to the net present value from an additional year of education for individual i in country j. m_j refers to the coefficient on years of schooling from the Mincerian wage equation for country j, $GDPPC_j$ refers to the GDP per capita in country j and p_j is the percentage of GDP per capita spent per student in country j. r refers to the discount rate and is assumed to be equal to 10%. Here note that we assume that the entire increase in cost and the cost of an additional year of schooling are borne by the individual herself in the current period.

As an illustrative exercise we calculate the net present value for an individual from Burkina Faso, Ethiopia and Ghana, respectively. The results are shown in Table 1.10.

[Insert Table 1.10]

The calculation suggests that the net present value of such an investment is equal to about 60% of the yearly per capita income in these countries. These moreover assume that the GDP per capita remains constant over the 30-year horizon and does not take into account any endogenous effects of increase in human capital on the growth rate or other externalities arising from an educated workforce. For instance, Appleton (2000)

35

estimates that a 1-year rise in the average primary schooling of neighbouring farmers is associated with a 4.3% rise in output, compared with a 2.8% effect of own farmer primary education in Uganda. The above evidence seems to suggest that African educationalists and policy makers should reassess current language use in education policies, as introduction of mother tongue instruction might involve sizeable gains and benefits for the majority of the population.

1.8   Conclusion

The paper studies the role of the institution of language use in education on educational attainment. We analyse how provision of mother tongue instruction in Ethiopia, to the ethnic group Oromo, affected their years of schooling. Our estimates suggest that the policy change had a sizeable positive impact and increased mean years of primary schooling by around 0.75 to 1 year and mean years of schooling by 0.80 to 1.18 years in the affected cohorts. The analysis shows that the language policy works primarily through the intensive margin of schooling. The estimate suggests that the change in language policy, conditional on enrolment, resulted in inducing 12% of the sample to complete 6 years or more of education, as compared to 5 years or less. The finding that language policy works through the intensive and not the extensive margin can be rationalized by a theoretical framework of sequential schooling choice under uncertainty.

The importance of education to growth and development of nation states imply that the African continent, which is characterised by the extensive use of the former colonial language in primary schooling, could have potentially large benefits from rethinking its language use in education policy. Applying our findings to a set of African countries show that provision of mother tongue instruction could increase the percentage of population completing primary schooling by as much as 8% points. Even accounting for the costs

of provision show there are still sizeable benefits from the introduction of mother tongue education. The implications of language choices in society extend beyond its effect on educational attainment. The institution of language policy has important implications on health, political participation, and division of power in society. Today most post colonial countries like India, Cameroon, Ghana, South Africa, to name a few, are marked by socioeconomic inequality along linguistic lines. The wider socioeconomic impacts of the institution of language use in education in particular and language choices in society in general remain relatively unexplored and an important area for future research.

Bibliography

[1] Altonji, J. (1993), "The demand for and return to education when education outcomes are uncertain," *Journal of Labor Economics*, 11(1): 48-83.

[2] Angrist, J and Chin, A and Godoy, R. (2008), "Is Spanish-only Schooling Responsible for the Puerto Rican Language Gap?" *Journal of Development Economics*, 85(1-2): 105-128.

[3] Angrist, J. and Lavy, V. (1997),"The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco," *Journal of Labor Economics*, 15: S48-S76.

[4] Angrist, J. and Pischke, S. (2009), *Mostly Harmless Econometrics*, Princeton, NJ: Princeton University Press.

[5] Appleton, S. (2000), "Education and health at the household level in sub-Saharan Africa," Working Paper No. 33, Harvard University Center for International Development.

[6] Bertrand, M. and Duflo, E. and Mullainathan, S. (2004), "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249-275.

[7] Chin, A. and Daysal, N. and Imberman, S. (2012), "Impact of Bilingual Education Programs on Limited English Proficient Students and Their Peers: Regression Discontinuity Evidence from Texas," NBER Working Paper No. 18197.

[8] Cummins, J. (1978a), "The cognitive development of children in immersion programs," *The Canadian Modern Language Review*, 34: 855-983.

[9] Cummins, J. (1978b), "Bilingualism and the development of metalinguistic awareness," *Journal of Cross-Cultural Psychology*, 9: l3l-l49.

[10] Cummins, J. (1979), "Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters," *Working Papers on Bilingualism*, l9: l97-205.

[11] Cummins, J. (1981), "The role of primary language development in promoting educational success for language minority students," in California State Department of Education (Ed.), *Schooling and language minority students: a theoretical framework*, 3-49.

[12] Cummins, J. (1991), "Interdependence of first- and second-language proficiency in bilingual children," in Bialystok, E. (ed.), *Language Processing in Bilingual Children*, Cambridge University Press.

[13] Dee, T. (2005), "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review*, 95(2): 158–165.

[14] Duflo, E. (2001), "Schooling and Labor. Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4): 795-813.

[15] Ehrenberg, G. and Goldhaber, D. and Brewer, J. (1995), "Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from NELS:88," *Industrial and Labor Relations Review*, 48(3): 547-561.

[16] Ethiopian Demographic and Health Surveys (EDHS) (2011).

[17] Education Sector Development Programme III, Joint Review Mission, (2006).

[18] Heckman, J. and Hotz, J.(1989), "Choosing Among Alternative Non Experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," *Journal of the American Statistical Association*, 84(408): 862–74.

[19] Heckman, J. and Stixrud, J. and Urzua, S. (2006), "The Effects of Cognitive and Noncognitive Abilities on Labor Market," *Journal of Labor Economics*, 24(3): 411-82.

[20] Heugh, K. and Benson, C. and Bogale, B and Gebre Yohannis, M.A. (2007), "Final Report: Study on Medium of Instruction in Primary Schools in Ethiopia," Commissioned by the Ministry of Education, Ethiopia.

[21] *International Encyclopedia of Education* (Third Edition), Elsevier.

[22] International Labour Organisation (2010), "Global Wage Report 2010/11: Wage policies in times of crisis."

[23] Jackson, T. (2000), "Equal Access to Education: A Peace Imperative for Burundi," London: International Alert.

[24] Klein, S. and Le, V. and Hamilton, L. (2001), "Does matching student and teacher racial/ethnic group improve math scores?" (Report DRU-2529-EDU), Santa Monica, CA: RAND Education.

[25] *Language and National Identity in Africa*, (2008), Oxford University Press.

[26] MacKinnon, J. G., and White, H. (1985), "Some heteroscedasticity-consistent covariance matrix estimators with improved finite sample properties," *Journal of Econometrics*, 29: 53-57.

[27] Ministry of Education, Ethiopia (2010), "Adequacy and Effectiveness of Public Education Spending in Ethiopia."

[28] Patrinos, H. and Psacharopoulos, G. (1995), "Educational performance and child labor in Paraguay," *International Journal of Educational Development*, 15(1): 47-60.

[29] Patrinos, H. and Vawda, Y. (1999), "Producing Educational Materials in Local Languages: Costs from Guatemala and Senegal," *International Journal of Educational Development*, 19: 287-299.

[30] Patrinos, H. and Psacharopoulos, G., (2004), "Returns to investment in education: a further update," *Education Economics*, 12(2): 111-134.

[31] Patrinos, H. and Velez, E. (2009), "Costs and benefits of bilingual education in Guatemala: A partial analysis," *International Journal of Educational Development*, 29(6), 594-598.

[32] Rosenbaum, P. (1987), "The role of a second control group in an observational study," *Statistical Science*, 2: 292-316.

[33] Skutnabb-Kangas, T. and Toukomaa, P. (1976), "Teaching migrant children's mother tongue and learning the language of the host country in the context of the socio-cultural situation of the migrant family," Tampere: UNESCO Report, University of Tampere, Research Reports, 15.

[34] Strauss, J. and Thomas, D. (1995), "Human Resources: Empirical Modeling of Household and Family Decisions," in Jere Behrman and T. N. Srinivasan, eds., *Handbook of development economics*, North Holland, 3A(9): 1885–2023.

[35] Verhoeven, L. (1994), "Transfer in bilingual development: The linguistic interdependency hypothesis revisited," *Language Learning*, 44: 381- 415.

[36] Wong Fillmore, L. (1991), "When learning a second language means losing the first," *Early Childhood Research*, 6: 323-46.

[37] World Bank. (2005), "Education in Ethiopia: Strengthening the Foundation for Sustainable Progress."

[38] World Bank. (2005), "In their Own Language... Education for All, Education Notes."

[39] World Bank. (2004), " Ethiopia Public Expenditure Review: The Emerging Challenge Volume I and II: Public Spending in the Social Sectors 2000-2020."

[40] Zamarro, G. (2010), "Accounting for Heterogeneous Returns in Sequential Schooling Decisions," *Journal of Econometrics*, 156: 260-76.

Notes

[1]Refer to Sabates, R., Akyeampong, K., Westbrook, J., and Hunt, F. (2010) for a recent review and international comparison of dropout and repetition rates.

[2]It is also useful to note that the services sector (51% of GDP) and tourism are important economic drivers for Morocco, whereas in Ethiopia 80% of GDP is from the agricultural sector. Moreover France remains the most important trading partner highlighting the important role of French in Morocco. In the case of Ethiopia, the major trading partners are China, Germany and Belgium, none of these countries national language is English. The above seems to again suggest that a decrease in labour market returns found by Angrist and Lavy (1999) for Morocco might not be an important factor for Ethiopia.

[3]Refer to Thomas and Collier (2002) for the effect of provision of bilingual schooling on long-term academic achievement. They find that minority language students learning only in English start to show decreases in achievement by the beginning of the 5th grade.

[4]We use the terms language use in education, language policy and medium of instruction policy interchangeably in the paper. For our purpose they should be understood as the language used to teach children in schools.

[5]The enrolment rate in our data stands at around 70% implying there is much scope for increasing years of schooling also through the extensive margin.

[6]Refer to year 2000 report by the Minority Rights Group International tiled Ethiopia: A New Start?

[7]A comparable example would be English and Hindi, which both belong to the Indo-European language phylum, but split at the first branch.

[8]Haile Selassie was born from parents of three Ethiopian ethnicities, the Oromo and Amhara and the Gurage. He decided to adopt his Amharian heritage as the banner under which the centralization of Ethiopia was undertaken. The policy of centralization resulted in the choice of Amharic as the official language.

[9]The "Derg" period was characterized by strong state control. This period has sometimes been referred to as the "Red terror campaign" as wide scale human rights abuses were carried out by the establishment in power against any protesting voices.

[10]The results are not sensitive to the choice of ages of the younger and older cohorts. Expanding the definition of the younger and older cohort to range from -2 to 7 and 8 to 36, respectively, leaves our results essentially unchanged. We additionally in the results section estimate the effect of the language policy for each age in 1994 from 2 to 21 years old.

[11]A potential problem is that the data only lets us identify the current region of residence and not the

actual region where education was obtained. In case the current region of residence is one of the four regions of the country where mother tongue instruction was introduced but the actual region of education was different from the current region of residence, then our estimates would provide a lower bound of the true effect as we wrongly consider untreated individuals as treated. The 2011 DHS data provides no information on childhood place of residence or for how long the individual has resided in the current region. Data on how long the individual has resided in the current region was however collected for the 2005 round of the DHS. Looking at the 2005 DHS round, we see that around 85% of the Oromo men have always lived in the current region of residence and around 70% of the women have always lived in the current region of residence. Internal migration moreover does not bias our estimates if it takes place within the four treated regions of the country. Given that nearly 90% of the Oromo population does indeed live within these four regions further reduces the possibility of potential bias due to internal migration.

[12] The results are also provided in Table 1.11 in the appendix.

[13] The coefficient for the cohort aged 2 and 3 in 1994 are positive but just slightly below conventional significance level. Also as we are looking at a nationally representative sample and given only very few 8 to 12 year olds gained access to mother tongue instruction, the coefficient of zero is to be expected.

[14] The results are also presented in Table 1.12 in the appendix.

[15] The coefficient on the cohort aged 8 in 1994 is just significantly different from zero, but as pointed out before this could be due to implementation of mother tongue instruction in some schools for pupils already in primary schooling or due to delayed entry or grade repetition. Also in column (2) of 1.12 the errors have been clustered at the level of household clusters and comparing with column (1), we can see that it does not affect the results.

[16] It could also be similarly argued that if new teachers were actually hired, which does not seem to be the case, they would in fact be less experienced and with lower training, implying if anything the change in the composition of teachers should go against our findings.

[17] As this is a binary dependent variable with 2 saturated model, we would get identical results using a Logit or a Linear Probability model.

[18] The marginal effects for categorical variables (like our interaction term) shows how $P(Y = 1)$ changes as the categorical variable changes from 0 to 1, holding all other variables at their means.

[19] Language and National Identity in Africa, Oxford University Press 2008.

[20] The policy in fact has a positive effect at all levels of schooling and not just all levels of primary schooling. Results available on request.

[21]The choice of 6 years of education as the cut-off point is chosen as this is the number of years of required to finish primary schooling in most countries and completion of primary schooling by all by 2015 is an important objective in most African states. Using any other year as the cut-off does not change the essence of the result and as mentioned before the language policy changes has a positive effect at all levels of schooling.

[22]We are aware of the problems associated with using Mincerian wage regression such as the endogeneity of post-schooling human capital accumulation, the fact that schooling and training are treated symmetrically in calculating the rate of return to schooling and that the general equilibrium affects are not accounted for. This said most instrumental variable estimates are found to be larger than the ordinary least squares estimates suggesting if anything we are underestimating the benefits of the policy change.

Table 1.1: **Medium of instruction (MOI) policy in Ethiopia by regions**

Region name	Languages implemented as MOI in primary schooling
ADDIS ABABA	Amharic
DIRE DAWA	Amharic, Oromigna, Somali.
AFAR	Amharic and Afar.
AMHARA	Amharic, Awingi, Hamittlena and Oromigna.
BENISHANGUL GUMUZ	Amharic
GAMBELLA	Nuek, Anguak and Meshenger.
HARARI	Amharic, Harari and Oromigna.
OROMO	Amharic and Oromigna
SNNPR	Amharic, Dawro, Gamo, Gedeo, Gofa, Hadiya, Kembata, Kafinono, Kotigna, Sidama and Wolaita.
SOMALI	Amharic and Somali.
TIGARY	Tigrinya.

a. The medium of instruction in primary schooling was Amharic in all the regions prior to 1994.
b. Source: Heugh, K. and Benson, C. and Bogale, B and Gebre Yohannis, M.A. (2007). Final Report: Study on Medium of Instruction in Primary Schools in Ethiopia. Commissioned by the Ministry of Education, Ethiopia. September to December 2006.

Table 1.2: Descriptive statistics for individuals aged 2 to 7 or 13 to 20 in 1994

Variable	Country Average	Amhara group	Oromo	Non- Amhara group
Mean Years of Schooling	4.35	6.27	3.76	3.39
Mean Years of Primary Schooling	3.04	4.31	2.98	2.76
% of Households classified as Middle Class and Below	53	30	47.29	57.43
Age of Household Head	40.44	41.16	39.21	39.72
% of Households with Male Head	80	73	84	79.9
% of Households with Bank Accounts	13	29	9	9.4

47

Table 1.3: **Average years of schooling by language group and cohort for the Amharas and Oromos**

	Years of primary schooling of the Oromo language Group	Years of primary schooling of the Amhara language group	Difference
Panel A: Experiment of Interest			
	(1)	**(2)**	**(3)**
2-7 Years old in 1994	3.81	4.16	-0.35
	(1327, 3.13)	(1430, 3.19)	
13-20 Years old in 1994	1.82	2.89	-1.07
	(1310, 2.62)	(1297, 3.35)	
Difference	1.99	1.27	0.72
Panel B: Control Experiment			
13-20 Years old in 1994	1.82	2.89	-1.07
	(1310, 2.62)	(1297, 3.35)	
21 - 28 Years olds in 1994	1.65	2.65	-1.00
	(108, 2.63)	(368, 1.99)	
Difference	0.17	0.24	-0.07

a. Number of observations and standard deviation in parentheses.
b.The observations are from the regions where the Oromo younger cohort is treated, namely Amhara, Oromia, Harari and Dire-Dawa.

Table 1.4: **Impact of provision of mother tongue instruction on Oromo people**

	(1)	(2)	(3)
Panel A: Experiment of interest: Indviduals aged 2 to 7 or 13 to 20 in 1994.			
(**Youngest cohort aged 2 to 7 in 1994**)			
Cohort Dummy*Oromo Language Group Dummy	0.721***	0.745***	0.797***
	(0.169)	(0.154)	(0.206)
Cohort Dummy	1.270***	1.580***	2.342***
	(0.126)	(0.265)	(0.353)
Oromo Language Group Dummy	-1.075***	-2.868***	-4.394***
	(0.118)	(0.129)	(0.179)
Other Controls	No	Yes	Yes
Observations	5,364	5,364	5,364
R-squared	0.080	0.240	0.424

	(1)	(2)
Panel B: Control experiment: Indviduals Aged 13 to 20 or 21 to 28 in 1994.		
(**Youngest cohort aged 13 to 20 in 1994**)		
Cohort Dummy*Oromo Language Group Dummy	-0.095	-0.009
	(0.180)	(0.153)
Cohort Dummy	0.263*	0.495*
	(0.140)	(0.261)
Oromo Language Group Dummy	-0.979***	-3.373***
	(0.136)	(0.146)
Other Controls	No	Yes
Observations	4,448	4,448
R-squared	0.030	0.309

a. The dependent variable in column (1) and (2) is years of primary schooling and in column (3) is years of schooling.

b. The mean of dependent variable in column (1) and (2) is 3.20 and in column (3) is 4.05.

c. Other controls include year of birth and region dummies.

d. HC_2 standard errors are in parentheses.

e. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

f. The Amhara are the control groups the regions are where the Oromo group is treated, namely, Amhara, Oromo,Harari and Dire-Dawa.

Table 1.5: **Average years of primary schooling by language group and cohort for the Oromo in the treated and untreated areas**

	Years of primary schooling of the Oromo people in treated areas	Years of primary schooling of the Oromo people in the untreated areas	Difference
Experiment of interest			
	(1)	**(2)**	**(3)**
2-7 Years old in 1994	3.54	4.28	-0.47
	(1327, 3.13)	(305, 3.62)	
13-20 Years old in 1994	1.82	3.04	-1.22
	(1310, 2.62)	(268, 3.25)	
Difference	1.99	1.24	0.75

a. Number of observations and standard deviation in parentheses

b. The Oromo Language group are treated in the regions of Amhara, Oromia, Harari and Dire-Dawa and untreated in the regions of Afar, Benishangul Gumuz, Gambella, SNNPR, Somali and Tigray.

Table 1.6: **Regression comparing the Oromo in the treated and untreated regions**

	(1)	(2)	(3)
Cohort Dummy*Region Dummy	0.751***	1.00***	1.178***
	(0.287)	(0.269)	(0.371)
Cohort Dummy	1.240***	1.233***	1.877***
	(0.264)	(0.429)	(0.581)
Regional Dummy	-1.22***	-1.37	-2.553
	(0.212)	(1.65)	(2.472)
Other Controls	No	Yes	Yes
Observations	3,210	3,210	3,210
R-squared	0.104	0.154	0.160

a. The dependent variable in column (1) and (2) is years of primary schooling and in column (3) is years of schooling.

b. Other controls include year of birth and region dummies.

c. HC_2 standard errors are in parentheses.

d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

Table 1.7: **Proportion enrolled by language group and cohort for the Amharas and Oromos**

	Proportion enrolled of the Oromo language group	Proportion enrolled of the Amhara language group	Difference
	(**1**)	(**2**)	(**3**)
2-7 Years old in 1994	0.709	0.711	-0.002
	(1327, 0.45)	(1430, 0.45)	
13-20 Years old in 1994	0.426	0.484	-0.058
	(1310, 0.49)	(1297, 0.49)	
Difference	0.283	0.227	0.056

a. Number of observations and standard deviation in parentheses.
b.The observations are from the regions where the Oromo younger cohort is treated, namely Amhara, Oromia, Harari and Dire-Dawa.

Table 1.8: Probit regression - marginal effects: Impact of provision of mother tongue instruction on Oromo people on enrolment and completion

	(1) Dependent variable Dummy - Enrolled	(2) Only Enrolled Sample Dependent variable Dummy -6 years of schooling or more	(3) Full Sample Dependent Variable Dummy -6 years of schooling or more
Experiment of Interest: Indviduals Aged 2 to 7 or 13 to 20 in 1994.			
(Youngest cohort aged 2 to 7 in 1994)			
Cohort Dummy*Oromo Language Group Dummy	0.0178	0.118***	0.092***
	(0.028)	(0.035)	(0.029)
Cohort Dummy	0.143***	0.125*	0.144***
	(0.047)	(0.066)	(0.047)
Oromo Language Group Dummy	-0.353***	-0.398***	-0.440***
	(0.022)	(0.028)	(0.020)
Observations	5,364	3,147	5,364

a. All specifications include year of birth and region dummies.

b. * , ** and *** significant at 10, 5 and 1 % significance level respectively.

c. The Amhara are the control groups the regions are where the Oromo group is treated, namely, Amhara, Oromia,Harari and Dire-Dawa.

52

Table 1.9: **Impact of provision of mother tongue instruction for African countries on primary schooling attainment**

(1)	(2)	(3)	(4)	(5)	(6)
Country	Current % of pop. completing Primary Schooling	% of pop. completing Primary Schooling with mother tongue provision	Ethnic groups with pop. shares of 10% and more	Share of the Ethnic groups completing Primary Schooling currently	Share of the Ethnic groups completing Primary Schooling with mother tongue provision
BENIN	29%	32.5%	Adja (11.96%), Bariba (9.67%) and Fon (42.66%)	Adja (25.83%), Bariba (11.14%) and Fon (26.98%)	Adja (31.49%), Bariba (13.67%) and Fon (33%)
BURKINA FASO	15%	17.6%	Mossi (56%)	Mossi (16%)	Mossi (19%)
CAMEROON	59.8%	63.3%	Bamblike (20%) and Beti (9.01%)	Bamblike (81%) and Beti (80%)	Bamblike(93.5%) and Beti (91.1%)
GABON	66%	73.32%	Fang (25%), Nzabi-duma (10.2%) and Shira-pun (21%)	Fang (82%), Nzabi-duma (58%) and Shira-pun (69%)	Fang (93.7%), Nzabi-duma (70.39%) Shira-pun (80.34%)
GHANA	70%	78%	Akan (47%), Ewe (13.7%) and Mole-dagiban (16%)	Akan (83.5%), Ewe (76.5%) and Mole-dagiban (40%)	Akan (94.5%), Ewe (86.5%) and Mole-dagiban (46.75%)

a. In column (4) in parenthesis are the population share of each group.

b. In column (5) in parenthesis are the % of population from each group completing primary schooling currently.

c. In column (6) in parenthesis are the % of population from each group completing primary schooling after provision of mother tongue instruction.

Table 1.10: **Net present value arising for an individual from provision of mother tongue instruction**

Country	Discount rate	GDP per capita	Coefficient on years of schooling	% of GDP per capita spent per student	NPV
BURKINA FASO	10%	212	9.6	30%	121
ETHIOPIA	10%	178	8.0	18%	99
GHANA	10%	360	7.1	12%	193
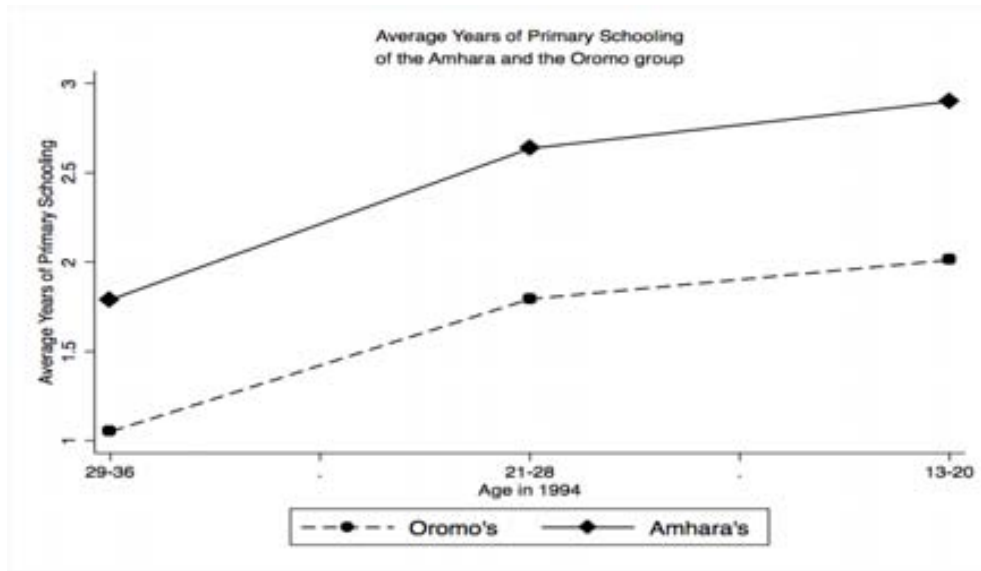
Figure 1.1: **Trends pre-1994 for the Amhara and the Oromo language group**
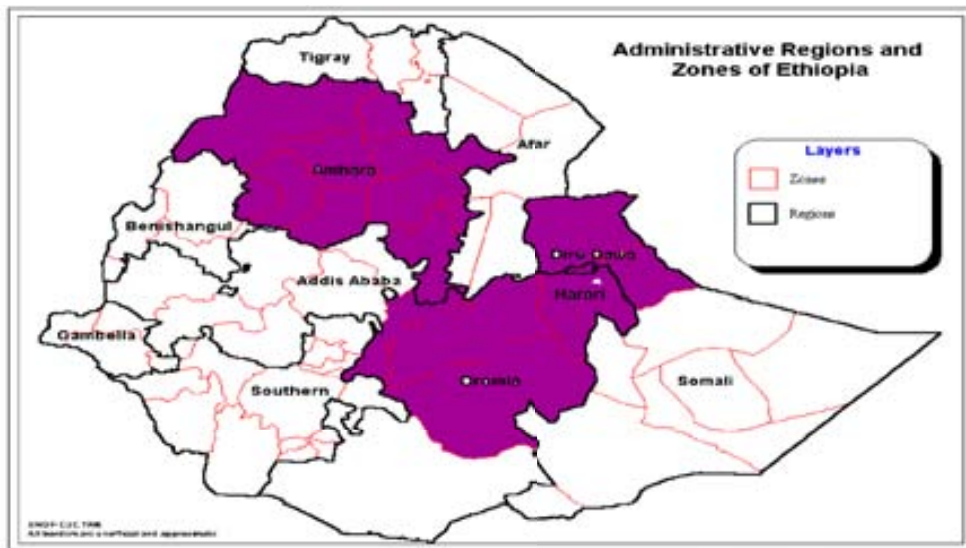
Figure 1.2: **Map showing the regions and city administrative units of Ethiopia**

55

Panel A
Coefficient of interaction age in 1994 with the Oromo language group dummy
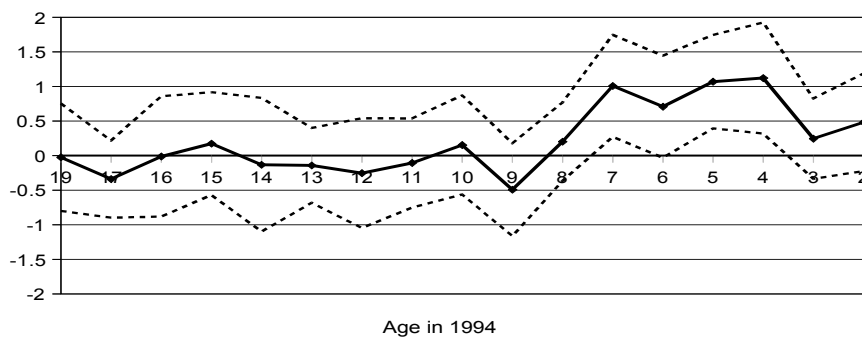
Age in 1994

Panel B
Coefficient of interaction age in 1994 with the regions where the Oromo are treated

Age in 1994
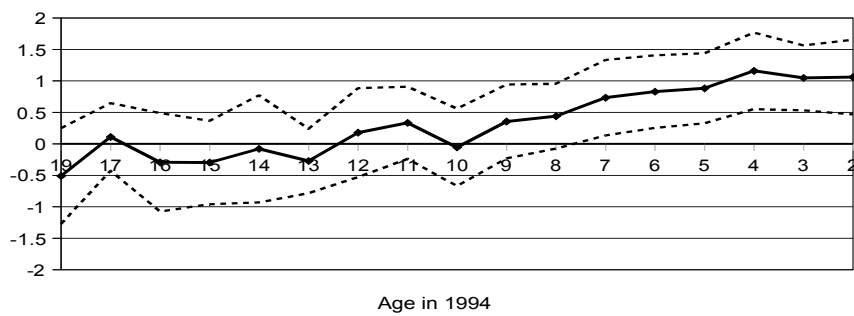
Figure 1.3: **Effect of the policy change for each cohort aged 2 to 21 years in 1994**

Panel A -CDF of Education for Oromo

Panel B - CDF of Education for Amhara

Panel C - Difference in Difference in CDF
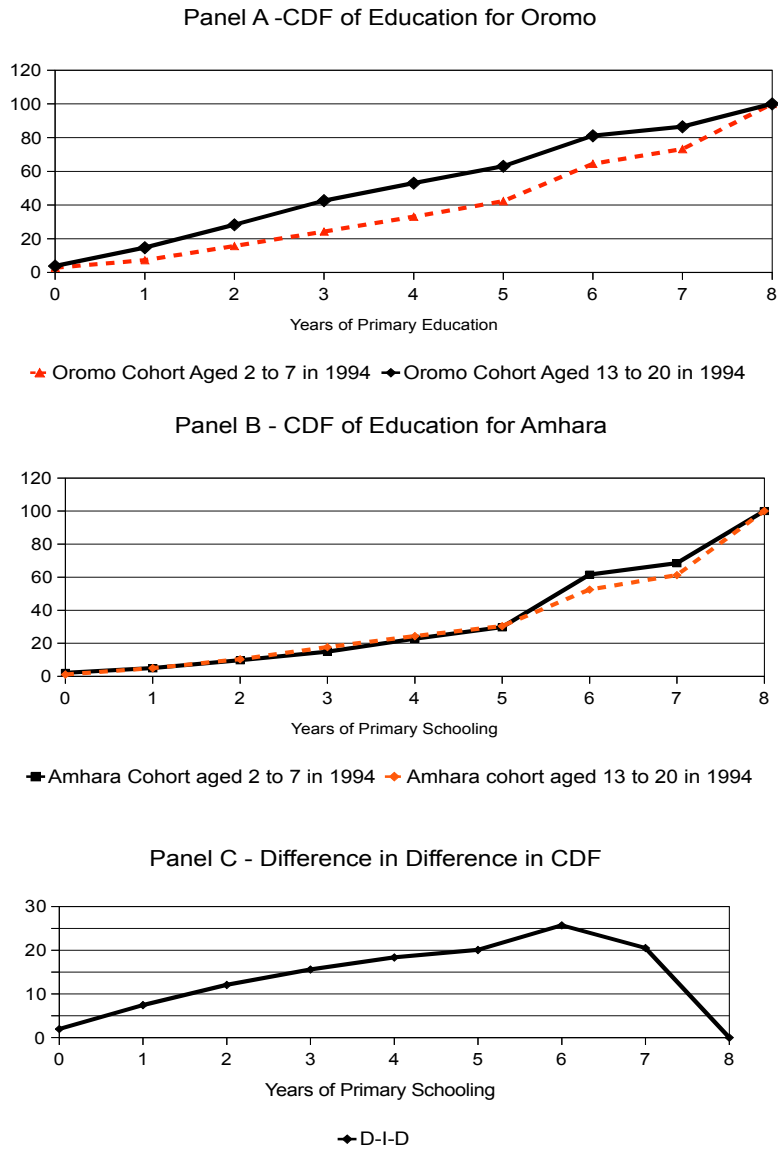
Figure 1.4: **CDFs of education for the Amhara and Oromo and D-I-D of CDFs**

57

1.9 Appendix

1.9.1 Effect on each cohort aged 2 to 21 in 1994 with errors clustered at the household cluster level

Table 1.11: **Impact of provision of mother tongue instruction on Oromo people: Coefficient of interaction between dummies indicating age in 1994 and Oromo language group dummy**

Dependent variable - Number of years of Primary Schooling

Age in 1994	(1)	(2)
2	0.481	0.481
	(0.360)	(0.378)
3	0.244	0.244
	(0.297)	(0.345)
4	1.120***	1.120***
	(0.410)	(0.420)
5	1.068***	1.068***
	(0.344)	(0.365)
6	0.707*	0.707*
	(0.376)	(0.389)
7	1.007***	1.007***
	(0.376)	(0.343)
8	0.201	0.201
	(0.289)	(0.276)
9	-0.493	-0.493
	(0.343)	(0.325)
10	0.153	0.153
	(0.365)	(0.349)
11	-0.107	-0.107
	(0.329)	(0.331)
12	-0.254	-0.254
	(0.405)	(0.420)
13	-0.142	-0.142
	(0.276)	(0.275)
14	-0.133	-0.133
	(0.493)	(0.499)
15	0.172	0.172
	(0.379)	(0.342)
16	-0.014	-0.014
	(0.444)	(0.475)
17	-0.340	-0.340
	(0.284)	(0.253)
19	-0.025	-0.025
	(0.396)	(0.365)
Control Variables	Yes	Yes
Observations	7,953	7,953
R-squared	0.242	0.242

a. All specifications include as controls year of birth and region dummies.

b. in column (1) HC_2 standard errors and in column (2) errors clustered at the level of the 231 household cluster are shown in parentheses.

c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

d. The control group is individuals aged 8 to 21 in 1994.

Table 1.12: **Impact of provision of mother tongue instruction on Oromo people: Coefficient of interaction between dummies indicating age in 1994 and regional dummies where Oromo are treated**

Dependent variable - Number of years of Primary Schooling

Age in 1994	(1)	(2)
2	1.060***	1.060***
	(0.303)	(0.326)
3	1.047***	1.047***
	(0.263)	(0.314)
4	1.159***	1.159***
	(0.310)	(0.354)
5	0.883***	0.883***
	(0.284)	(0.306)
6	0.829***	0.829***
	(0.294)	(0.287)
7	0.733**	0.733**
	(0.306)	(0.329)
8	0.440*	0.440
	(0.263)	(0.289)
9	0.355	0.355
	(0.299)	(0.307)
10	-0.057	-0.057
	(0.314)	(0.324)
11	0.334	0.334
	(0.293)	(0.282)
12	0.178	0.178
	(0.360)	(0.369)
13	-0.273	-0.273
	(0.261)	(0.301)
14	-0.080	-0.080
	(0.434)	(0.478)
15	-0.297	-0.297
	(0.338)	(0.354)
16	-0.294	-0.294
	(0.399)	(0.383)
17	0.109	0.109
	(0.274)	(0.265)
19	-0.512	-0.512
	(0.388)	(0.399)
Control Variables	Yes	Yes
Observations	12,178	12,178
R-squared	0.215	0.215

a. All specifications include as controls year of birth and region dummies.

b. in column (1) HC_2 standard errors and in column (2) errors clustered at the level of the 231 household cluster are shown in parentheses.

c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

d. The control group is individuals aged 8 to 21 in 1994.

1.9.2   Results of robustness tests

Table 1.13: **Average years of schooling by language group and cohort of the Amhara and Non-Amhara people**

	Years of primary schooling of the Non-Amhara Language Group	Years of primary schooling of the Amhara Language Group	Difference
Panel A: Experiment of Interest			
	(1)	**(2)**	**(3)**
2-7 Years old in 1994	3.54	4.31	-0.77
	(4538, 2.66)	(2001, 3.15)	
13-20 Years old in 1994	1.74	3.03	-1.29
	(4794, 2.66)	(1756, 3.32)	
Difference	1.80	1.28	0.52
Panel B: Control Experiment			
13-20 Years old in 1994	1.74	3.03	-1.29
	(4794, 2.66)	(1756, 3.32)	
21 - 28 Years olds in 1994	1.45	2.65	-1.20
	(3284, 2.37)	(1604, 2.84)	
Difference	0.29	0.38	-0.09

a. Number of observations and standard deviation in parentheses
b. The sample includes all regions of the country where the Non-Amharic language group is treated, namely Afar,
Amhara, Benishangul Gumuz, Dire-Dawa, Gambella, Harari, Oromia, SNNPR, Somali and Tigray.

Table 1.14: **Impact of provision of mother tongue instruction on non-Amharic language people**

Dependent variable - Number of years of Primary Schooling	(1)	(2)
Panel A: Experiment of Interest: Indviduals Aged 2 to 7 or 13 to 20 in 1994.		
(**Youngest cohort aged 2 to 7 in 1994**)		
Cohort Dummy*Non-Amharic Language Group Dummy	0.577***	0.553***
	(0.130)	(0.116)
Cohort Dummy	1.210***	1.578***
	(0.114)	(0.191)
Non-Amharic Language Group Dummy	-1.31***	-2.40***
	(0.094)	(0.102)
Other Controls	No	Yes
Observations	11,918	11,918
R-squared	0.091	0.207

Dependent variable - Number of years of Primary Schooling	(1)	(2)
Panel B: Control Experiment: Indviduals Aged 13 to 20 or 21 to 28 in 1994.		
(**Youngest cohort aged 13 to 20 in 1994**)		
Cohort Dummy*Non-Amharic Language Group Dummy	-0.055	0.035
	(0.143)	(0.117)
Cohort Dummy	0.358***	0.672***
	(0.130)	(0.193)
Non-Amharic Language Group Dummy	-1.259***	-2.741***
	(0.107)	(0.113)
Other Controls	No	Yes
Observations	10,281	10,281
R-squared	0.041	0.198

a. Other controls include year of birth and region dummies.

b. HC_2 standard errors are in parentheses.

c. The sample includes all regions of the country where the Non-Amharic language group is treated, namely Afar, Amhara, Benishangul Gumuz, Dire-Dawa, Gambella, Harari, Oromo, SNNPR, Somali and Tigray.

d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

Table 1.15: **Impact of provision of mother tongue instruction on the Others**

	(1)	(2)
Dependent variable - Number of years of Primary Schooling		
Panel A: Experiment of Interest: Indviduals Aged 2 to 7 or 13 to 20 in 1994.		
(**Youngest cohort aged 2 to 7 in 1994**)		
Cohort Dummy*Others' Language Group Dummy	0.983*	1.008*
	(0.570)	(0.558)
Cohort Dummy	0.883	0.637
	(0.552)	(0.683)
Non-Amharic Language Group Dummy	-2.896***	-2.866***
	(0.485)	(0.475)
Other Controls	No	Yes
Observations	1,669	1,669
R-squared	0.141	0.151

a. The other controls include years of birth and region dummies.

b. HC_2 standard errors are in parentheses.

c. The sample includes the region of the country where the Sidama and Wolaita language group are treated, namely SNNPR.

d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

Table 1.16: Impact of provision of mother tongue instruction on Oromo people by gender

Dependent variable - Number of years of Primary Schooling

	(1) Full Sample	(2) Boys Only	(3) Girls Only
Panel A: Experiment of Interest: Indviduals Aged 2 to 7 or 13 to 20 in 1994.			
(Youngest cohort aged 2 to 7 in 1994)			
Cohort Dummy*Oromo Language Group Dummy	0.745***	1.038***	0.609***
	(0.154)	(0.217)	(0.205)
Cohort Dummy	1.580***	0.175	1.954***
	(0.265)	(0.364)	(0.334)
Oromo Language Group Dummy	-2.868***	-2.459***	-3.327***
	(0.129)	(0.180)	(0.178)
Other Controls	Yes	Yes	Yes
Observations	5,364	2,625	2,739
R-squared	0.240	0.247	0.275

a. Other controls include year of birth and region dummies.

b. HC_2 standard errors are in parentheses.

c. The Amhara are the control groups the regions are where the Oromo group is treated, namely, Amhara, Oromia,Harari and Dire-Dawa region.

d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

1.9.3 Methodology applied to calculate effect of mother tongue instruction on a set of African countries

The procedure utilized to calculate the change in the proportion of population completing primary schooling, due to provision of mother tongue instruction, is best explained with the help of an example. The same procedure is then applied to a set of five African countries and the results are shown in Table 1.11.

Benin is used as an example to outline the mechanics underlying the procedure. We consider the population aged 15 to 49 from the DHS data of 2006. The DHS data provides us with information on both the education attainment [1], the language of the respondent and whether the individual was ever enrolled in schooling. Approximately 29% of the population aged 15 to 49 in our sample completes primary schooling or more. Looking at the population shares of the various language groups, there are three groups with population shares of 10% or more. The "Adja " people comprise about 11.96%, the "Bariba" people about 9.67% [2] and the "Fon" about 42.6% of the population. The percentage of population completing primary schooling for the three groups is 25.83%, 11.14% and 26.98%, respectively. We have information on the CDF of education for the three groups both conditional on enrolment and for the whole sample i.e. both individuals who enrolled and did not enrol. The estimate from the previous section suggests that the provision of primary schooling induced 12% of the enrolled sample to complete 6 years of schooling or more as compared to 5 years or less. The sample size of the enrolled population, for instance for the ethnic group Fon, is 7609 individuals. The calculated estimate implies that now 12% or 897 additional individuals complete primary schooling as compared to before. Looking at the entire sample of individuals aged 15 to 49 from the Fon group, which includes 14935 individuals; we see that 10905 individuals were dropping out with 5 years of education or less. The provision of mother tongue, given our estimate, would reduce the number

of individuals dropping out with 5 years of education or less from 10905 individuals to about 10007 individuals (i.e. 10905 minus 897). This would imply that the proportion of population which now completes primary schooling or more increases from around 26.98% to 33%. Doing a similar exercise for the Adja and the Bariba people suggests that the proportion of people completing primary schooling or more would increase from 25.83% and 11.14% to 31.5% and 13.67% respectively. The percentage of population completing primary schooling for the other language group remains unchanged. We hence using the new values of the percentage of population completing primary schooling for the three groups along with their population shares calculate the overall change in the percentage of population completing primary schooling or more. The estimate suggests that the provision of mother tongue instruction to these three groups, which comprise about 64% of the population, would increase the percentage of population completing primary schooling, for the people aged 15 to 49, in the country from around 29% to 32.5%, an increase of nearly 4% points.

Chapter 2

Discrimination Without Taste - How Discrimination Can Spillover and Persist (joint with Christopher Rauh)

2.1 Introduction

The literature of the economics of discrimination was pioneered by the seminal work of Becker (1957, 1971). In the setting envisaged, employers hold a taste for discrimination, such that working with members of a particular group imposes a cost on them, and hence these workers have to compensate the employer by either being more productive or accepting lower wages. The taste based models of discrimination have been mainly criticized on the premise that in the presence of competitive markets employers with taste for discrimination will be driven out of the market. The class of models of statistical discrimination

(Phelps 1972; Arrow 1973; Aigner and Cain 1977; Lundberg and Startz 1983, 1998; Coate and Loury 1993; Mailath et al. 2000) and categorical thinking (Fryer and Jackson 2008) rely on the imperfect observability of worker productivity. In absence of complete information employers base their decision on easily observable characteristics, such as race or gender, to infer the expected productivity of the worker. The third class of models and the theoretical work closest in spirit to ours is that of Akerlof (1976, 1985), where not following the established norm of discrimination against certain groups results in imposition of social sanctions which cause economic losses, and hence make discrimination a rational response.

In this paper we posit a channel of discrimination, where even under perfect observability of individual ability, the absence of discriminatory social norms, and when taste for discrimination has already died out, to discriminate can be the optimal response of principals in the economy. The discrimination in our setting arises due to the complementarity in input provision, which leads to interdependency in payoffs. The coordination failures result from the belief that somebody else might discriminate, imposing losses due to the interdependency in payoffs, which in turn leads people without a taste for discrimination to also discriminate. In the stylized dynamic model there exists a set of principals with a taste for discrimination against a certain group. This taste for discrimination is assumed to die out at some exogenous rate. We outline the decision-making procedure of principals, such that their payoff maximizing actions and beliefs constitute a sub game perfect Nash equilibrium. The model demonstrates how, given certain conditions, actions of all principals, even those without a taste for discrimination, in equilibrium are indistinguishable from those who hold a taste for discrimination. At some finite point, after all principals who had a taste for discrimination no longer exist in society, we show that to discriminate against individuals of the discriminated group can still be an equilibrium response in the game that follows. The model shows how historical existence of discrimination in just one

market can spread across the economy and persist even after no principal in the economy has any taste for discrimination. Here discrimination results from a coordination failure, in contrast to Akerlof (1976, 1985), where it is driven by intentional coordination.

Using self-employment as an example, we illustrate how the discriminatory mechanism might occur due to strategic complementarities and the coordination problem in the production process of self-employment (Basu 2010). In our context, what makes self-employment distinct from wage employment, is the need to establish productive relations with other principals in the economy.[3] The fact that entrepreneurial activities involve establishing links across markets, implies that an entrepreneur's success does not solely depend on his actions and ability, but also on the actions of the other complementary input providers. For instance, a self-employed might require to establish relationships with other complementary input providers such as a distributor and a money lender. The success and return for all three is contingent on the participation of all three in the venture. If one of the complementary input providers, i.e. the moneylender or distributor, believes with a high enough probability that the other is a taste discriminator who will refuse to establish a relation with the entrepreneur, we show that to also discriminate, irrespective of whether one has a taste for discrimination or not, is the optimal action of the other agent, as well. The model shows how the persistent belief regarding the presence of discrimination can sustain a discriminatory equilibrium even after no more taste for discrimination exists in the economy. The coordination failures are driven by the fear that discriminatory actions by others could impose losses, through the inherent complementarity in the production process, leading to principals with no taste for discrimination to also discriminate in equilibrium. For the case of self-employment the main prediction of the model is that individuals who belong to the discriminated group can have lower participation rates in self-employment and/or suffer price discrimination despite equal ability, leading to an over-

all welfare loss. The nature of the coordination failure does not allow for a single principal who does not discriminate to reap the unrealized profits, a possibility traditionally assumed by Becker (1971), therefore, providing a theoretical rational as to why discrimination can persist.

The empirical literature dealing with discrimination and self-employment in the US documents the differences in participation and returns between ethnic groups (Moore 1983; Bailey and Waldinger 1991; Borjas 1986; Fairlie and Meyer 1996, 2000; Fairlie 1996; Blanchflower 2009). In line with our theoretical predictions the above mentioned studies find that in the US 14.4% of white males versus 5.1% of black males are self-employed and blacks are more likely to have loan applications rejected and pay higher interest rates on loans than comparable white males do. In the case of wage employment premarket skills measured by the Armed Forces Qualification Test (AFQT) score have be shown to explain most of the black-white wage gap (Neal and Johnson 1996). However, for the case of self-employment Fairlie (2002) shows that controlling for AFQT test scores does not significantly reduce the black-white gap in the probability of becoming self-employed, suggesting that discrimination might have a role to play in explaining the observed differences in self-employment rates.

Our model can reconcile unexplained features observed in the data for the market for self-employment. For instance, why the socially most disadvantaged groups in India (Schedule Castes and Schedule tribes) are even more relatively underrepresented in urban rather than rural areas in terms of non-farm enterprise ownership, even though discrimination is higher in rural areas (Iyer et al. 2011).[4] Why in Sweden, one of the countries where women's labour force participation rate is very high and only 0.4% of the male population "strongly agree" that men make better business executives than women, has among the lowest level of self-employment for women in the EU.[5] The fact that beliefs about discrim-

ination are higher in urban rather than rural areas in India, and remain high in Sweden concerning women, could be an important explanatory factor.

Our model is also applicable to a range of markets with strategic complementarities. The dominance of particular ethnic groups in certain professions (Greif 1989, 1993; Banerjee and Munshi 2004) might be explained through our mechanism as ethnic enclaves might help secure complementary support from other individuals and overcome coordination failures.[6] Card et al. (2008) assume that when black people move into a neighbourhood, white neighbours with a distaste for blacks will change neighbourhoods. Anticipating a decrease in housing prices, people without a distaste for black neighbours will also sell their property and move. We show that the presence of neighbours with a distaste for black neighbours is not required to trigger the segregating dynamics, the belief is sufficient, hence providing an alternative explanation for the phenomenon of racial tipping points in the United States. The model can also help shed light on phenomena such as the low inter-caste marriage rates observed in India or the choice of life partners from historically discriminated groups, because people might fear discrimination by family or society against their partner.

The persistence of beliefs regarding current presence of discrimination due to the historical existence of discrimination is key to our model. The importance of history, culture, and past events in shaping today's beliefs, behaviour, and outcomes has been demonstrated in the literature (Nunn and Wantchekan 2011; Voth and Voigtlaender 2012; Alesina et al. 2011; Argenziano and Gilboa 2011).

We formally test our theory using data from the General Social Survey from 1972-2010 and by creating proxies for beliefs about and tastes for discrimination, in order to estimate the probability of being self-employed in the US. We find proxies for beliefs about presence of discrimination to be a consistently significant factor in explaining lower levels of self-employment among blacks in the United States. The estimate suggests that reducing

70

the share of people who believe that discrimination exists by 10% points would increase self-employment rates among blacks from 7.3% to 9.3%, an increase of 28%. The results are robust to the inclusion of year and region fixed effects and a variety of controls.

The main contribution of the paper lies in highlighting coordination failure as a channel through which discrimination might exist and persist. Though the outcomes of discrimination operating through various channels might be identical, understanding the mechanism behind it is crucial to be able to devise appropriate policy measures, as we show that traditionally prescribed interventions might be ineffective and, in fact, could result in discriminatory actions against other groups, as well. The weak conditions under which discrimination can spread and persist in our model, suggest that caution should be exercised when thinking about issues related to discrimination. Intervention along the lines of publically relayed signals or use of "role models" might be required to change beliefs to move from an equilibrium, where to discriminate is an optimal response, to one where not to discriminate is the optimal action. The feature of strategic complementarity in actions and the interlinkages across markets shows how the positive (negative) effects of intervention in one market might have significant positive (negative) externalities.

The rest of the paper is organized as follows: section II presents the theoretical model, section III presents empirical support of our theory, section IV evaluates policy tools, and section V concludes.

2.2 The Model

The society consists of individuals i of two types $s \in \{A, B\}$. The types A and B form social groups based on visible characteristics which do not influence performance (e.g. race, religion, ethnicity). Individuals of type A and B belong to the finite, large sets $\mathcal{A}$ and $\mathcal{B}$, respectively.[7] The individuals have an ability a_i, where a is distributed uniformly over $[0, 1]$.

Ability $a_i \in [0, 1]$ reflects productive capacity and is perfectly observable to all. Individuals can either earn wages as workers or choose to become self-employed as entrepreneurs. The productivity, which we assume to be equal to gross income, of individual i of type s in wage employment is given by:

$$W_{i,s} = Y_{i,s} = G_W(a_{i,s}) = a_i, \tag{2.1}$$

while the productivity and gross income of an individual who is self-employed is characterized by:

$$W_{i,s} = Y_{i,s} = G_{SE}(a_{i,s}, C_i) \ where \ \frac{dG_{SE}}{da_{i,s}} \geq 0 \ , \frac{dG_{SE}}{dC_i} \geq 0 \ and \ \frac{dG^2_{SE}}{dC_i da_{i,s}} \geq 0. \tag{2.2}$$

The productivity of the self-employed depends not only on his own ability a_i, but also on C_i. The component C_i captures the productive relations individual i is able to establish with what we call "principals" in the economy, which is necessary for the production process in self-employment. It is assumed that for an entrepreneur to succeed, certain contacts are required to start up a business, while engaging in wage employment simply requires applying one's ability. For example, the entrepreneur might require capital in the form of a loan from a bank (lender) and also may need to have an agreement with a distributor, who will be willing to distribute his goods. The distributor needs to be convinced that the individual will receive a loan in order to produce the goods of the requisite quality at the agreed time, while the lender has to believe that the individual will be able to sell the goods through a distributor. The requirement of productive relations with a distributor and a lender is only for illustrative purposes and could be extended to n-players or include any other contact necessary to setup a successful enterprise (e.g., supplier, landlord to rent office). Thus, both components are necessary and cannot be substituted through ability.

In the coordination game we assume no communication between the potential lender and distributor whose productive relations are complementary to the entrepreneur. The only observables to them are the individuals type and ability, combined with his offer. In order to keep our analysis tractable, we assume that for an individual wanting to become self-employed involves interaction with two principals $p \in \{L, D\}$ (lender and distributor, respectively).

DEFINITION 1. *Throughout the paper we will refer to $s \in \{A, B\}$, those who decide between wage or self-employment, as individuals, and to $p \in \{L, D\}$, the lenders and distributors, as principals.*

We define C_i as

$$C_i = V(c_L, c_D) \ where \ \frac{d^2 V(c_L, c_D)}{dc_L dc_D} \geq 0, \tag{2.3}$$

where c_L and c_D refer to the decisions by the principal $p \in \{L, D\}$ of whether to establish productive relations, which are given by $c_p \in \{0, 1\}$. We only allow for pure strategies, such that they decide whether to lend/sign the contract ($c_p = 1$) or not ($c_p = 0$), and also the amount of their investment they decide upon is normalized to unity. Adding the productivity factor $\lambda > 1$ of self-employment, we can represent the productivity of a self-employed individual by:

$$P_{i,s} = G_{SE}(a_{i,s}, C_i) = \lambda c_L c_D a_i. \tag{2.4}$$

The above functional form exhibits an extreme form of complementarity in the actions of the principals p implying:

$$P_{i,s} = G_{SE}(a_{i,s}, V(c_L, 0)) = G_{SE}(a_{i,s}, V(0, c_D)) = 0. \tag{2.5}$$

The intuition is that establishing a relationship with both principals $p \in \{L, D\}$ is required for the entrepreneur to operate or be successful.

Each period t the individuals i decide whether they want to become workers or attempt to establish productive relations with $p \in \{L, D\}$ in order to become self-employed. The individuals wanting to become entrepreneurs are randomly matched with a pair of principals, i.e. a lender and a distributor, in the market every period. The individual i does not know whether the principals he is matched with have a taste for discrimination or not, while the lender and the distributor can observe i's ability a_i and type s. If the individual simply enters wage employment without entering negotiations he earns a_i. To apply for a loan and the contract of the distributor, they have to offer an amount σ_p to the principal as repayment for the investment. Should this offer be rejected, the individual i enters wage employment earning $(a_i - \delta)$, where δ is a fixed cost arising from the effort exerted. The principal has the opportunity of a risk free investment yielding interest r per unit invested. The bargaining process leading to the offer is defined by a Nash bargaining solution, where the individual i and the principal p share equal bargaining power. This bargaining takes place between only one principal and the individual, without communication between the lender and the distributor.[8] Since the outcome is dependent on what decision the other principal (henceforth denoted by $-p$) takes, p will create a belief concerning the likelihood of the other principal accepting the offer, as well.[9] If p accepts an offer which the other principal rejects, then he is not able to obtain r from the risk free investment in the given period due to his capital being bound, and hence not yielding any interest. Due to non existing possibilities for collusion, random matching and the absence of perfect competition, the interaction between the individual and the principal is characterized by a monopoly versus monopsony, commonly referred to as bilateral monopoly. Both have one shot at earning a surplus compared to their outside option. Assuming equal bargaining power and

linear utility functions in payoffs, the bargaining solution, resulting in offer σ_p, is characterized by the disagreement point $d = (d_i, d_p) = (a_i - \delta, 1 + r)$ and the maximization of $(\sigma_p - 1 - r)(\lambda a_i - \sigma_p - \sigma_{-p} - a_i + \delta) \Rightarrow \sigma_p = \frac{(\lambda-1)a_i - \sigma_{-p} + r + \delta + 1}{2}$. Now assuming i makes the same offer to the lender as well as the distributor, such that $\sigma_p = \sigma_{-p}$, the Nash bargaining solution is defined as:

$$\sigma^N(a_i) = \frac{(\lambda - 1)a_i + r + \delta + 1}{3}. \tag{2.6}$$

This Nash bargaining solution is a function of a_i, which we will denote as $\sigma^N(a_i)$. In order for p to accept this offer we require the amount he expects to be repaid to be at least what he can earn through the risk free investment, such that the lowest ability a' of an individual who could possibly offer him this share is determined by $\sigma^N(a_i) = \frac{(\lambda-1)a_i + r + \delta + 1}{3} \geq 1 + r \Rightarrow$ $a' = \frac{2(1+r)-\delta}{\lambda-1}$. Now looking at the participation constraint of the individual, we require $\lambda a_i - 2\sigma^N(a_i) \geq a_i$

$$\Rightarrow a^* = \frac{2(1 + r + \delta)}{\lambda - 1}. \tag{2.7}$$

Since $a^* > a'$ only individuals with $a_i \geq a^*$ will intend to become entrepreneurs and, as long as there is no taste for or belief about discrimination, will be accepted.[10]

DEFINITION 2. *Let a^* be the individual with the lowest ability who wants to become an entrepreneur in the absence of discrimination.*

Now at time equal to t_0, we assume there is a shock to the taste of a subset of principals in society. We assume that π_0 proportion of principals develop a taste for discrimination equal to $b(> 0)$ against establishing a productive relation with B-type individuals. The taste for discrimination can be understood as a cost/disutility which the principals with taste for discrimination face when they decide to establish a productive relation with a B-type individual in society. The origins of the shock which result in creating a taste for discrimination among a subset of the principals is not the focus of the paper and can

arise due to various reasons. An example could be the incidents of September 11th 2001, which resulted in the creation of a taste for discrimination among various individuals in the United States against Muslims.[11]

The shock to the taste of the principals in society implies that the probability of discrimination occurring has to be taken into account while deciding on the optimal course of action. Observe that individuals, who decide to become entrepreneurs, are matched with a distributor or lender at random. The random matching implies that the probability of meeting a principal with a taste for discrimination, in any period t, is equal to π_t which is their share of total principals in society. Individuals and principals are expected payoff maximizers. Since neither π_t is not common knowledge their decisions are conditioned on beliefs about the share of discriminators amongst the principals which they update through observations of discrimination in the market. We assume that the event which creates a taste for discrimination results in creating a common prior among individuals and principals. The common prior is assumed to have a distribution denoted by η_t, capturing the probability of meeting a principal with a taste for discrimination. The common prior η_t is modelled as having a beta distribution. More specifically it is assumed that the individuals and principals believe that the share of principals with taste b has a beta distribution with parameters α_0 and β_0. Moreover, we denote the density of the distribution η_t by θ. The beta distribution captures the belief regarding the probability of meeting a principal with a taste for discrimination through its expected value, or the mean of the distribution.

ASSUMPTION 1. *The probability parameter capturing the share of lenders with a taste for discrimination equal to b in period t_0 is given by $\theta(\eta_0) \sim beta\ (\alpha_0, \beta_0)$.*

The above distribution implies that the density function associated with facing a dis-

criminator with taste b is given by:

$$\theta(\eta_0) = \frac{(\eta_0)^{\alpha_0-1}(1-\eta_0)^{\beta_0-1}}{beta(\alpha_0, \beta_0)} = \frac{(\alpha_0 + \beta_0 - 1)!}{(\alpha_0 - 1)!(\beta_0 - 1)!}(\eta_0)^{\alpha_0-1}(1-\eta_0)^{\beta_0-1} \qquad (2.8)$$

The beta distribution hence gives us a density on $[0, 1]$, which captures the beliefs held by the individuals and principals regarding η_0. As the individuals and principals will need to decide on optimal actions based on their beliefs and all individuals and principals are assumed to be risk neutral, the individuals and principals use the expected value of the distribution which is given by $E(\eta_0) = \frac{\alpha_0}{\alpha_0 + \beta_0}$. [12]

DEFINITION 3. *An offer of $\sigma^N(a_i) > 1+r$, which is rejected by any principal, is defined as a case of discrimination.*

In the *dynamic* model a principal p exits the market with exogenous probability ω every period. The probability ω is not known to anybody in society. He is replaced by a principal from the same profession, but always without a taste for discrimination, such that at some point no principals with a taste for discrimination will be left. Therefore, if we define the share of principals with a taste for discrimination in period $t = 0$ to be π_0, the probability that an individual i is matched with a principal with taste for discrimination in period T is $\pi_0(1 - \omega)^T$.

The principals and individuals are assumed to use a Bayesian approach to update their beliefs. Now assume that in period 1, n_1 individuals applied and k_1 cases of discrimination are observed in the market. Out of the total of n_1 cases assume that $n_{1b}(\leq n_1)$ cases involve offers such that $1 + r < \sigma^N(a_i) < 1 + r + b$. This implies the total number of people who could be potentially discriminated against is n_{1b}.

ASSUMPTION 2. *It is assumed that all market transactions in terms of the offers made and rejected are common knowledge.*[13]

As out of the potential n_{1b} cases, k_1 cases exhibit discrimination, we can define the posterior density function for the individuals and principals in society. The posterior function for $\theta(\eta_0)$ is given by:

$$\theta(\eta_1|k_1) \sim beta(\alpha_0 + k_1, \beta_0 + n_{1b} - k_1). \tag{2.9}$$

The above outlines the Bayesian belief updating procedure used by individuals and principals regarding the probability of meeting a principal with taste for discrimination. In fact we can denote the posterior distribution for any period $T > t_0$, given the total number of B-type individuals who make offers $\sigma^N(a_i)$, such that $1+r < \sigma^N(a_i) < 1+r+b$ in order to become entrepreneurs and the cases of discrimination observed in the market. The posterior probability density is given by $\theta(\eta_T|\sum_{t=1}^{T} k_t) \sim beta(\alpha_0 + \sum_{t=1}^{T} k_t, \beta_0 + \sum_{t=1}^{T} n_{tb} - \sum_{t=1}^{T} k_t)$. The associated expected value or the point probability estimate used by the individuals and principals to make their optimal decision is given by:

$$E(\eta_T) = \frac{\alpha_0 + \sum_{t=1}^{T} k_t}{\alpha_0 + \beta_0 + \sum_{t=1}^{T} n_{tb}}. \tag{2.10}$$

DEFINITION 4. *Let $\varphi_t = E(\eta_{t-1})$, such that φ_t is the probability that individuals and principals assign to the existence of a principal with taste for discrimination b in period t.*

[Insert Figure 2.1]

Now let us recall, as illustrated in Figure 2.1, the decision of a B-type individual with an offer $1+r < \sigma^N(a_i) < 1+r+b$, so one which could possibly become a victim of discrimination. First the individual decides whether to become a worker or an entrepreneur. If he decides to become a worker then he simply earns a wage equal to his ability a. However, if he attempts to become an entrepreneur he is matched with a lender and a distributor. He

makes an offer σ, which the lender and the distributor decide whether to accept or not. Everybody assigns probability φ to either principal rejecting an offer $1+r < \sigma^N(a_i) < 1+r+b$. At the terminal nodes the payoffs of the individual, the lender L, and the distributor D are illustrated in the mentioned order. We assume the individuals and principals choose among their actions subgame perfectly.

Observe that the shock which creates a taste for discrimination does not affect the A-types in the market. All individuals of the A-type with $a_i \geq a^*$ still offer $\sigma^N(a_i)$ and are accepted. However, the individuals of B-type now take the probability of meeting a discriminator in the market into account while deciding on their optimal course of action. In any period t they calculate their expected payoff from the various actions available to them, namely, *applying for self-employment and offering their Nash bargaining solution, seeking wage employment,* or *applying for self-employment and offering a share to compensate the potential discriminatory principals for their taste for discrimination.* Let us denote by a_b^* the lowest ability type such that $\sigma^N(a_b^*) = 1 + r + b$. All individuals with ability $a_i \geq a_b^*$ offer $\sigma^N(a_i) \geq 1 + r + b$ and will never be discriminated against.

The individuals in the ability range $a^* \leq a_i < a_b^*$ face potential discrimination. Their expected payoff from offering the Nash bargaining solution is given by $(1 - \varphi_T)^2(\lambda a_i - 2\sigma^N(a_i)) + (1 - (1 - \varphi_T)^2)(a_i - \delta)$. They compare the above to the payoff $(\lambda a_i - 2(1 + r + b))$, which they earn by offering $(1 + r + b)$ to each principal and escaping discriminators, to obtaining wage employment and earning a_i.

The principals with no taste for discrimination, when facing an individual in the ability range $a^* \leq a_i < a_b^*$, compare their expected payoff from establishing a productive relation $(1 - \varphi_T)(\sigma^N(a_i)) + \varphi_T(1)$ to their return from the risk free investment $(1 + r)$, and choose the option that gives them the higher expected payoff. As we can observe due to the interdependency in payoffs, the principal while calculating his expected payoff from estab-

lishing a productive relation, also takes into account the fact that the individual might be matched with a second principal, who has a taste for discrimination. Here we can see how discrimination has spilled over to those without a taste for discrimination.

The principals with a taste for discrimination, which is equal to b, reject the Nash bargaining solution when facing an individual in the ability range $a^* \leq a_i < a_b^*$ and accept offers from $a_i \geq a_b^*$.

The decision-making rules of the individuals and principals imply that the probability of becoming self-employed for a B-type individual in any period T will depend upon his ability a_i, the actual share of taste discriminators $\pi_0(1-\omega)^T$, and his beliefs regarding the share of taste discriminators in society φ_T. We can thus express the probability of self-employment for a B-type individual as a function of the above three factors, i.e. $f(a_i, \pi_0(1-\omega)^T, \varphi_T)$. It is easy to see that the probability of self-employment is increasing in ability and declining in the actual share and the belief regarding the proportion of taste discriminators in society, i.e. $f_1 > 0$, $f_2 < 0$, and $f_3 < 0$, where the subscripts refer to the first, second and third argument of the function. In the next section we explicitly test for the predictions of our model using the above function f.

The channel of discrimination that we put forth works on the premise that even once all principals with taste for discrimination die out, to discriminate against members of group B may remain as the optimal action. In what follows we address whether discrimination can exist, and if it can, under what conditions does it exist, for how long does it persist, and in what form does it manifest itself. Let us denote by T^* the first period in which no principals with taste for discrimination remain in the economy. The probability density function for meeting a principal with a taste for discrimination is given by:

$$beta(\alpha_0 + \sum_{t=1}^{T^*-1} k_t, \beta_0 + \sum_{t=1}^{T^*-1} n_{tb} - \sum_{t=1}^{T^*-1} k_t). \tag{2.11}$$

80

The probability point estimate for meeting a discriminator is given by $\varphi_{T^*} = \frac{\alpha_0 + \sum_{t=1}^{T^*-1} k_t}{\alpha_0 + \beta_0 + \sum_{t=1}^{T^*-1} n_{tb}}$.

It is clear that all $B-$type individuals with $a_i \geq a_b^*$, will offer their Nash bargaining solution and be accepted. The form of discrimination and the length for which it will persist after all principals with taste for discrimination have died out will depend on φ_{T^*}. The point probability estimates in period T^* are a function of the initial beliefs (α_0, β_0), the actual share of taste discriminators π_0 and the rate ω at which principals with a taste for discrimination exit the market in every period. If we assume that the initial beliefs are a function of the actual share of taste discriminators i.e. $\alpha_0(\pi_0)$ and $\beta_0(\pi_0)$, the we can write $\varphi_{T^*} = f(\pi_0, \omega)$. Let us moreover denote by a_b as the individual who is indifferent between offering $(1 + r + b)$ to each principal and obtaining wage employment. It is easy to see that a_b is equal to $\lambda a_b - 2(1 + r + b) = a_b$, or $a_b = \frac{2(1+r+b)}{\lambda-1}$. In the proposition that follows we highlight the various forms in which discrimination manifests itself and persists after no principals with a taste for discrimination are left in the multiple equilibria depending on φ_{T^*}.

PROPOSITION 1. *1. Let φ_{T^*} be such that no individual of the B- type in the range $a^* \leq a_i < a_b^*$ prefers the Nash bargaining solution to wage employment at time T^*. In such a scenario all individuals of type B with $a^* \leq a_i < a_b$ will remain as wage employed forever, whereas all the individuals of the A-type with equal ability will become self-employed and enjoy the premium associated with self-employment. All individuals of the B-type with $a_b \leq a_i \leq a_b^*$ will become self-employed, but pay an amount equal to $1 + r + b \geq \sigma^N(a_i)$ forever, and hence will earn strictly less than the A-types with equal ability.*

2. Let φ_{T^*} be such that some individual of B-type in the range $a^* \leq a_i \leq a_b^*$ strictly prefers the Nash bargaining solution to obtaining wage employment. This implies that in the long run all individuals of type B in the range $a^* \leq a_i \leq 1$ will become

81

self-employed paying $\sigma^N(a_i)$, however, the time period for which discrimination will persist will be a function $\varphi_{T^*} = f(\pi_0, \omega)$ and the lowest a_i, who still prefers the Nash solution to wage employment in T^*.

Proof. Proof in appendix. ■

The intuition for the equilibrium outcome where discrimination persists forever crucially depends on whether when the last principal with taste for discrimination dies out, which is the lowest B-type who decides to become an entrepreneur. If all individuals of the B-type, whose Nash bargaining solution is not sufficient to compensate the taste for discrimination (i.e. $1 + r \leq \sigma^N(a_i) < 1 + r + b$), decide to seek wage employment rather than self-employment, due to beliefs about discrimination being prohibitively high will imply that no offers, which are subject to potential discrimination, are ever made. This in turn will imply that the beliefs remain frozen at the current level and hence all individuals with ability levels $a^* \leq a_i < a_b$ will always prefer seeking wage employment and remain workers forever.

On the other hand when the last principal with taste for discrimination dies out, the lowest B-type who decides to become an entrepreneur is one whose Nash bargaining solution is not sufficient to compensate the taste for discrimination (i.e. $1 + r \leq \sigma^N(a_i) < 1 + r + b$) implies discrimination will not persist in the long run. To see this, observe, that as now all principals with taste for discrimination have died out, all offers made by individuals seeking to become self-employed will be accepted. As this includes individuals whose offers could have been subject to potential discrimination, but are not (as no taste discriminators are left), the next period beliefs about discrimination will be lower, after updating of beliefs. As every period all offers will be accepted, in the long run the belief about discrimination will tend to zero.

2.2.1 Welfare Effects

Now let us consider the loss of welfare to society once discrimination persists due to beliefs, despite no taste for discrimination remaining in the economy. Besides the efficiency loss due to productive individuals not becoming entrepreneurs, there is also a redistribution effect. The efficiency loss not only affects the B-type individuals, who become workers instead of entrepreneurs (which would be the social optimum), but also the lenders and distributors, who lose out on opportunities of receiving offers yielding more than the risk free investment r. Additionally, there is a wealth redistribution, as B-types offering $(1+r+b)$ are paying a higher price than the equivalent A-type, from which the lender and distributor are profiting, as they are receiving more than the Nash bargaining solution.[14] Now, assuming that the necessary and sufficient condition of proposition 1.1 hold and discrimination persists, we can quantify the deadweight loss of each period to the B-type as:

$$\sum_{a_i \geq a^*}^{a_b} [(\lambda a_i - 2\sigma^N(a_i)) - a_i], \tag{2.12}$$

while the deadweight loss to the lenders and distributors is:

$$\sum_{a_i \geq a^*}^{a_b} 2[\sigma^N(a_i) - (1+r)]. \tag{2.13}$$

The redistribution from the B-types to the lenders and the distributors is:

$$\sum_{a_i \geq a_b}^{a_b^*} 2[(1+r+b) - \sigma^N(a_i)]. \tag{2.14}$$

PROPOSITION 2. *In an equilibrium as in proposition 1.1, on average lenders and distributors earn lower profits by discriminating.*

Proof. Proof in appendix. ■

In Figure 2.5 in appendix the difference between net earnings of the A and the B-type in function of their ability is illustrated. The dark shaded area is the deadweight loss caused by individuals of the B-type not becoming entrepreneurs due to beliefs about discrimination, whereas the light shaded area illustrates the redistribution caused by the higher price individuals of the B-type are paying in order to escape discrimination when becoming entrepreneurs. In the GSS dataset the pattern of average income in constant dollars of the self-employed by highest educational degree attained exhibits a striking similarity with our theoretical prediction. On average blacks earn less than whites in self-employment for all but those that obtained a graduate degree, which is the highest degree coded in the dataset. The aggregate Figure 2.6 is provided in the appendix.

2.3   Data and Empirics

2.3.1   Data

We use the General Social Survey (GSS) wave from 1972-2010 with more than 50,000 observations along 28 questionnaires to test the predictions of the theoretical framework. The data allows us to construct proxies for the belief and taste for discrimination parameters in our model. Taking the percentage of white people, for each year and region, answering the following questions, we construct two proxies for the taste for discrimination: [15]

1. Whites who answer "yes" to "Do you think there should be laws against marriages of Blacks and Whites?"

2. Whites who are "very" or "somewhat opposed" when asked "What about having a close relative marry a Black person?"

In order to construct a proxy for beliefs regarding discrimination, we take the percentage of the population, for each year and region, answering the following question with "yes":

- "On the average Blacks/African-Americans have worse jobs, income, and housing than White people. Do you think these differences are mainly due to discrimination?"

Unfortunately neither of these questions is asked throughout the entire wave of surveys which restricts our number of observations, but depending on the model specifications, still provides us with sample sizes of up to nearly 20,000 observations. In Figure 2.2 the two measures of taste for discrimination, beliefs about discrimination and the self-employment rates of blacks and whites are plotted from 1972-2010 for those years where we have observations. The discrimination measures are from the GSS dataset, whereas self-employment rates by race are from the Integrated Public Use Microdata Series (IPUMS). Tastes for discrimination seem to decline linearly. The beliefs on the other hand have remained remarkably stable, just as the gap in self-employment rates between blacks and whites.

[Insert Figure 2.2]

Not decomposing by region, beliefs about discrimination among whites peak in 1985 at 45% and reach its lowest point in 2004 at 34%. Our first measure for taste for discrimination among whites declines from 39% in 1972 to 10% in 2002. The second measure declines from 66% in 1990 to 22% in 2010.[16] The figure shows that taste for discrimination has been decreasing nearly linearly over time, whereas beliefs regarding discrimination are sticky in their evolution. When decomposing by year and region, disparities across the US become very clear. Beliefs about discrimination take their highest value in 1993 in New England at 67% and the lowest in West South Central in 2002 at 20%. The first measure for taste peaks in East South Central in 1987 at 71.4%, and takes its lowest value in New

England at 2% in 2002. The second measure for taste also has its highest value in East South Central in 1990 at 81%, while the lowest it takes is in the Pacific region in 2010 at 8%.

2.3.2 Methodology

Following the theoretical model we derive the equation for the estimation of the probability of individuals being self-employed as a function of ability a_i, the proportion π_{tqs} of principals with a taste for discrimination at time t in region q against group s, the proportion φ_{tqs} with beliefs about discrimination at time t, in region q against group s, and a vector of individual characteristics X_i with associated parameter vector γ. As a proxy for ability we use years of schooling. Therefore, the probability of individuals i at time t in region q and of group s being self-employed $se_{itqs} \in \{0, 1\}$ we define as

$$Prob(se_{itqs} = 1 | a_i, \varphi_{tqs}, \pi_{tqs}, X_i) = f(a_i, \varphi_{tqs}, \pi_{tqs}, X_i). \tag{2.15}$$

Using a logit regression we can define the estimated probability as

$$Prob(se_{itqs} | a_i, \varphi_{tqs}, \pi_{tqs}, X_i) = \frac{e^{g(a_i, \varphi_{tqs}, \pi_{tqs}, X_i)}}{e^{g(a_i, \varphi_{tqs}, \pi_{tqs}, X_i)} + 1} \tag{2.16}$$

where $g(a_i, \varphi_{tqs}, \pi_{tqs}, X_i) = \beta_0 + \beta_1 a_i + \beta_2 \varphi_{tqs} + \beta_3 \pi_{tqs} + \gamma X_i + \varepsilon_{itqs}$ and ε_{itqs}, the error term, is a binomially distributed random variable.

The proportion of principals with a taste for discrimination π_{tqs} and the proportion with beliefs about discrimination φ_{tqs} take the value zero for white individuals, i.e. for $s = A$. We restrict our sample to white and black respondents who are not students or retired, while assuming no differences in preferences to become self-employed.[17] In the logit regression estimating self-employment we control for gender, age, age squared, self-categorized family

income when respondent was 16, and whether the father was self-employed.[18]

2.3.3 Results

Table 2.1 shows the results of estimating the model given by equation 2.16. We see that years of schooling, our proxy for ability, and all controls are significant and have the expected sign. We find that either proxy for taste for discrimination explains significantly the event of a black person being self-employed in model (1) and (2) of Table 2.1, only as long as the proxy for beliefs about discrimination does not enter the model. Once beliefs about discrimination enter the model either proxy for taste for discrimination becomes insignificant, as can be seen in model (3) and (4), while the variable representing beliefs about discrimination is significant at the 1% level.

[Insert Table 2.1]

Calculating the expected effect of a decrease in beliefs about discrimination for an average black male in the sample for model (3), we find that by decreasing average beliefs about discrimination by 10% points from 24% to 14%, the probability of the average black male being self-employed increases from 7.3% to 9.3%, which is an increase of 28%. The magnitude of the effect of beliefs about discrimination becomes clear when we calculate the probability of the average black male being self-employed at zero taste for discrimination, but at the remaining 24% beliefs about discrimination. Here the probability of self-employment increases from 7.3% to 7.9%, which is an increase of 8%. The estimation indicates that abolishing taste for discrimination will not be enough to close the gap in self-employment rates. These findings suggest that when beliefs about discrimination are well established this is enough to lower the probability of black people being self-employed significantly, without even requiring a taste for discrimination to be present. In Table 2.2 in the appendix we control for regional and/or year fixed effects and find that beliefs about

discrimination remain significant at the 1% level in all model specifications.

We construct a proxy for statistical discrimination by coding as statistical discrimination the percentage of whites by year and region that assign a value from one to three on a scale from one to seven to the question "Do people in these groups [blacks] tend to be unintelligent or tend to be intelligent?", where one signifies unintelligent to seven signifying intelligent.[19] The question for the first measure for taste for discrimination only coincides for two waves of the survey, such that in the reduced sample we do not have enough variation in our key explanatory variables. Therefore, we exclude the first measure for taste for discrimination from this analysis. When not paired with beliefs we observe that statistical discrimination enters with the expected negative sign and is highly significant, but is insignificant when paired with beliefs about discrimination. Beliefs about discrimination remain significant at a 1% level whether including year and/or regional fixed effects (Table 2.3 in the appendix). As a robustness check we link our measures for tastes and beliefs concerning discrimination and statistical discrimination with the respective years and regions in the Integrated Public Use Microdata Series (IPUMS) of the University of Minnesota allowing for sample sizes of up to 630,000 observations. Here we additionally have information on whether the individual was self-employed and region of residence in the year before. Thus, we can improve our estimation correcting for the transition probability to self-employment, by conditioning on former employment status, a necessity pointed out by Heckman (1981). By using a bivariate probit endogenous regime switching model, as in Cappellari (2002), we remove the partial observability hypothesis of Heckman (1981).[20] As can be seen in Table 2.4 in the appendix we find in both, the selection and the outcome stage, the coefficients of beliefs about discrimination to be negative and significant at the 1% level in all model specifications strengthening the robustness of our findings. In contrast to our findings when estimating with the GSS sample, now both tastes for

discrimination remain significant when paired with beliefs about discrimination. This indicates that tastes for discrimination could still be high enough to prevent black individuals from entering self-employment. The coefficients of beliefs about discrimination are higher, thereby supporting our emphasis on their current importance in the discriminatory prevention of blacks entering self-employment. The covariance ρ between the individual specific error component is positive and significant at the 1% level in all specified models, thereby justifying the model selection based on the assumption that initial self-employment is not exogenous.

2.4 Policy Considerations

The belief driven gridlock put forth by the model, in which discrimination can persist in an inefficient equilibrium, provides opportunities for affirmative action to move the economy to the "good" equilibrium as a focal point in the coordination game. The analysis is restricted to the long run equilibrium where no taste for discrimination remains, but discrimination persists due to beliefs.

Provision of financial subsidies to the B-type with sufficiently high abilities to become entrepreneurs, but who are being discriminated upon, provides a potential solution. With the subsidy they could afford to pay the higher amount, such that beliefs about discrimination would be compensated and their offers would be accepted with certainty. This measure would overcome the problem that beliefs are prohibiting both, the principals from accepting and the individuals from applying. On the downside this provides a solution only as long the subsidy is in place, as this solution does not change beliefs. Moreover, the welfare effect would be negative, as the additional value creation attributed to self-employment sums up to less than the subsidy.[21]

Another method of achieving equality among the A and the B-type would be to discrimi-

nate against the discriminator. By imposing a fine F on principals who reject a B-type that has the same ability and offers the same amount as an A-type that has been accepted in the same period, one could target equal treatment of A and B-types.[22] This equal treatment might come at a high cost, though. If one principal interacts with various individuals in a given period there exists the possibility that principals begin discriminating against the A-type, as well, in order to avoid the fine when rejecting the B-type. Imagine a principal receiving the same offer $\hat{\sigma}$ by two individuals with identical ability $\hat{a}$, but of different types A and B, in the same period. Now if he accepts the A-type and rejects the B-type he will receive $\hat{\sigma} + 1 + r - F$, given that $-p$ accepts the A-type offer, as well. This would only be rational if $\hat{\sigma} - F \geq 1 + r$, because otherwise he would be better off rejecting the A-type, as well. Therefore, discrimination could spill over to the A-type.

By imposing lenders to give an equivalent share of credits at equal conditions as the A-type as observed over past periods to the B-type, lenders would be forced to accept offers by the B-type. This share would have to be benchmarked by total lending in the past conditioned on economic indicators, in order to avoid discrimination against the A-type. This measure by itself would not be sufficient, though, as individuals of the B-type would continue not to apply and distributors would continue to reject out of fear of discrimination. This intervention would have to be communicated publicly, such that it would serve as a signal and would spillover to the beliefs of the B-type and the distributors. To see this in terms of our model, imagine the government announcing publicly and credibly the implementation of this measure. Now there would be no reason for the distributor or the individual of the B-type to assign $\varphi > 0$. The great advantage of this intervention would be that intervening in one market would be enough to correct beliefs in other markets. Once the measure were to be removed, beliefs about discrimination would have vanished and no further discrimination would take place (given no taste for discrimination).

A further possibility to overcome the coordination failure would be the creation of an institution acting as coordination device providing the service of linking pre-screened non-discriminatory lenders and distributors to able blacks wanting to become entrepreneurs. As this could even be a profitable exercise such institutions might automatically arise and be provided by the market itself.

In the above we saw that schemes, such as subsidies or equal treatment regulations, might only address the problem myopically or, even worse, have undesirable consequences (like discrimination of A-types in equilibrium).

2.5 Conclusion

In this paper we show that even once taste for discrimination and statistical discrimination were to cede to exist in society, discrimination can persist not only in form of the cumulative effects of past discrimination, but also due to remaining beliefs making discrimination the best-response. The complementarity of inputs, which leads to interdependency in payoffs, results in coordination failures stemming from the belief that somebody else might discriminate, which in turn leads people without a taste for discrimination to also discriminate.

The outcomes predicted by the model, in terms of participation rates and incomes for the self-employed for the discriminated group being lower, are validated using data from the General Social Survey 1972-2010 of the United States and creating proxies for taste and beliefs regarding discrimination. A simple logit model shows that beliefs about discrimination are significant in explaining the lower probability of becoming self-employed for blacks, even after controlling for individual level characteristics, and region and year fixed effects.

We show that the nature of discriminatory coordination failures does not allow market forces to overcome discrimination, but may require alternative policy tools. The various

mechanisms through which discrimination manifests its dynamic linkages in terms of cross market and intergenerational effects, and the tendency to persist through cumulative and belief based channels, need to be understood and explored in order to develop policies aimed at eradicating discrimination and achieving equal treatment and opportunities.

Bibliography

[1] Akerlof, G.A. (1976), "The Economics of Caste and of the Rat Race and Other Woeful Tales," *The Quarterly Journal of Economics*, 90(4): 599-617.

[2] Akerlof, G.A. (1985), "Discriminatory, Status-based Wages among Tradition-oriented, Stochastically Trading Coconut Producers," *Journal of Political Economy*, 93(2): 265-76.

[3] Alesina, A.F., Giuliano, P. and Nunn, N. (2011), "On the Origins of Gender Roles: Women and the Plough," NBER Working Papers 17098, National Bureau of Economic Research, Inc.

[4] Argenziano, R. and Gilboa, I. (2011), "History as a Coordination Device," *Theory and Decision*, forthcoming.

[5] Arrow, K. (1973), "The Theory of Discrimination," in (O. Ashenfelter and A. Rees, eds.), *Discrimination in Labor Markets*, 3-33, Princeton: Princeton University Press.

[6] Bailey, T. and Waldinger, R. (1991), "Primary, Secondary, and Enclave Labor Markets: A Training System Approach," *American Sociological Review*, 56: 432-45.

[7] Banerjee, A. and Munshi, K. (2004), "How Efficiently is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur," *Review of Economic Studies*, 71(1): 19-42.

[8] Basu, K. (2010) *Beyond the Invisible Hand: Groundwork for a New Economics*, Princeton: Princeton University Press, and New Delhi: Penguin.

[9] Becker, G.S. (1957) *The Economics of Discrimination*, Chicago: University of Chicago Press.

[10] Becker, G.S. (1971) *The Economics of Discrimination*, Chicago: University of Chicago Press.

[11] Blanchflower, D.G. (2009), "Minority Self-Employment in the United States and the Impact of Affirmative Action Programs," *Annals of Finance*, 5(3-4): 361-96.

[12] Borjas, G.J. (1986), "The Self-Employment Experience of Immigrants," Working Paper 1942, National Bureau of Economic Research, Inc.

[13] Cappellari, L. (2002), "Do the 'Working Poor' Stay Poor? An Analysis of Low Pay Transitions in Italy," *Oxford Bulletin of Economics and Statistics*, 64(2): 87-110.

[14] Card, D., Mas, A., and Rothstein, J. (2008), "Tipping and the Dynamics of Segregation," *Quarterly Journal of Economics*, 123(1): 177-218.

[15] Coate, S. and Loury, G.C. (1993), "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *American Economic Review*, 83(5): 1220-40.

[16] Fairlie, R.W. (1996), "The Absence of African-American Owned Business: An Analysis of the Dynamics of Self-Employment," *Journal of Labor Economics*, 17(1): 80-108.

[17] Fairlie, R.W. (2002), "Drug Dealing and Legitimate Self-Employment," *Journal of Labor Economics*, 20(3): 538-67.

[18] Fairlie, R.W. and Meyer, B.D. (1996), "Ethnic and Racial Self-Employment Differences and Possible Explanations," *Journal of Human Resources*, 31(4): 757-93.

[19] Fairlie, R.W. and Meyer, B.D. (2000), "Trends in Self-Employment Among White and Black Men During the Twentieth Century," *Journal of Human Resources*, 35(4): 643-69.

[20] Fryer, R. and Jackson, M. (2008), "A Categorical Model of Cognition and Biased Decision-Making," *The B.E. Journal of Theoretical Economics*, 8(1): Article 6.

[21] Greif, A. (1989), "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders," *The Journal of Economic History*, 49(4): 857-82.

[22] Greif, A. (1993), "Contract Enforceability and Economic Institutions in Early Trade," *American Economic Review*, 83(3): 525-49.

[23] Heckman, J.J. (1981), "The incidental Parameters Problem and the Problem of initial Conditions in estimating discrete Time - discrete Data stochastic Processes and some Monte Carlo Evidence," in (C. Manski and D. McFadden, eds.), *Structural Analysis of Discrete Data with Econometric Applications*, 179-95, Cambridge: MIT Press .

[24] International Social Survey Programme (ISSP): Work Orientations III, 2005, Distributor: GESIS Cologne Germany ZA4850, Data Version 2.0.0 (2009-10-29).

[25] Iyer, L., Khanna, T. and Varshney, A. (2011), "Caste and Entrepreneurship in India," Working Paper 12-028, Harvard University.

[26] Lucas, R.E., Jr. (1978), "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9(2): 508-23.

[27] Lundberg, S. and Startz, R. (1983), "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review*, 73(3): 340-47.

[28] Lundberg, S. and Startz, R. (1998), "On the Persistence of Racial Inequality," *Journal of Labor Economics*, 16(2): 292-323.

[29] Mailath, G.J., Samuelson, L. and Shaked, A. (2000), "Endogenous Inequality in Integrated Labor Markets with Two-Sided Search," *American Economic Review*, 90(1): 46-72.

[30] Neal, D.A. and Johnson, W.R. (1996), "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 104(5): 869-95.

[31] Nunn, N. and Wantchekon, L. (2011), "The Trans-Atlantic Slave Trade and the Evolution of Mistrust in Africa: An Empirical Investigation," *American Economic Review*, 101(7): 3221-52.

[32] Phelps, E.S. (1972), "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62(4): 659-61.

[33] Ruggles, S.J., Alexander, T., Genadek, K., Goeken, R., Schroeder, M.B. and Sobek, M. (2010), "Integrated Public Use Microdata Series: Version 5.0," Minneapolis: University of Minnesota.

[34] Smith, T.W, Marsden, P., Hout, M. and Kim, J. (2011), General social surveys, 1972-2010 Chicago: National Opinion Research Center [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut.

[35] Voth, J. and Voigtlaender, N. (2012), "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," *Quarterly Journal of Economics*, forthcoming.

[36] World values survey (2009), 1981-2008 Official aggregate v.20090901 (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid.

Notes

[1] More specifically the CDF of education for all language groups as well as the overall CDF.

[2] We also include groups which might be just below the 10% level

[3] This is not to claim that wage employment activities do not often involve productive relations with others, but in self-employed activities they are essential.

[4] Observe that coordination failures in urban markets are more likely as they are anonymous, so even if taste for discrimination is higher in the rural than urban settings, it could well be the case that the coordination failures in urban areas outweighs the taste for discrimination effect in rural areas, leading to the outcome observed in the data.

[5] The wave of 2005-2007 of the World Values Survey exhibits that 43% of the Swedish population mention "Discrimination against women and girls" as one of the two most pressing problems facing the country. Moreover Swedish males have the lowest bias against women across all 39 countries in the sample (0.4%), while beliefs about discrimination are the second highest.

[6] This benefit, however, has to be weighed against the restriction on occupational choice that might arise due to ethnic enclaves being effective gate keepers to certain professions.

[7] The assumption of large sets is to ensure that any single individual does not have any market power and collusion cannot take place.

[8] Communication could be incorporated and would generally not change our findings if we assume costly communication or no reputation effects. Without reputation effects moral hazard problems would arise as there would be no gain from admitting when one was not willing to establish the productive relation.

[9] We disregard higher order beliefs, even though they would additionally speed up the contagion-effect.

[10] Similar to Lucas (1978) in our model only the most able want to become entrepreneurs.

[11] The assumption that shocks do not work the other way i.e. people immediately forget the past existence of discrimination due to sudden events today is justified by the literature on trust and beliefs which shows how persistent past beliefs are in shaping todays action. Refer to section I for references.

[12] This is the point at which the density of the distribution takes its highest value.

[13] We relax this assumption in the appendix to allow the individuals and principals to observe only a subset of all the market transactions and show that the results remain essentially unchanged.

[14] This theoretical prediction is consistent with the finding of the US Department of Justice that Countrywide charged more than 200,000 Black and Hispanic borrowers higher fees and interest rates than comparable Whites with similar credit histories between 2004 and 2008 leading to the Bank of America paying

98

a settlement of 335M$.

[15]The reason for segmentation of beliefs and taste for discrimination by regions in the United States can be found in Becker (1971).

[16]The Figures 2.3 and 2.4 in the appendix show the trends in beliefs and taste for the regions "Pacific" and "South Atlantic".

[17]In the International Social Survey on Work Orientation in 2005 we find that 71% of blacks versus 58% of whites in the labour force in the US would choose self-employment if they could choose between different kinds of jobs, suggesting that our estimates might even be underestimating effects of discrimination.

[18]A more detailed specification and justification of the controls and regional segmentation can be found in the appendix.

[19]This measure suffers from the problem that it most likely captures taste based discrimination, as there exists no scientific evidence of racial differences in intelligence.

[20]Model specifications can be found in the appendix.

[21]Observe that if it was not the case, the individuals themselves would pay extra to compensate the belief about discrimination.

[22]This solution assumes that the authorities could estimate ability through information on observables such as education, years of experience, age etc.

Table 2.1: **Logistic regression estimating self-employment in the US**

Dependent variable: Self-employment	(1)	(2)	(3)	(4)	(5)
Belief about discrimination			-2.680***	-4.871***	-3.413***
			(0.952)	(1.764)	(0.479)
Taste for discrimination 1	-1.399***		-0.651		
	(0.293)		(1.192)		
Taste for discrimination 2		-3.880***		0.051	
		(0.816)		(1.423)	
Years of schooling	0.027***	0.021*	0.037***	0.021*	0.026***
	(0.008)	(0.013)	(0.011)	(0.013)	(0.009)
Female	-0.846***	-0.658***	-0.658***	-0.659***	-0.722***
	(0.049)	(0.071)	(0.067)	(0.071)	(0.051)
Age	0.091***	0.098***	0.099***	0.098***	0.095***
	(0.009)	(0.015)	(0.014)	(0.015)	(0.010)
Age squared x 1,000	-0.691***	-0.729***	-0.773***	-0.733***	-0.749***
	(0.096)	(0.153)	(0.139)	(0.153)	(0.106)
Family income at age 16	0.185***	0.129***	0.180***	0.128***	0.154***
	(0.030)	(0.042)	(0.041)	(0.042)	(0.031)
Father was self-employed	0.637***	0.644***	0.575***	0.640***	0.625***
	(0.048)	(0.075)	(0.068)	(0.075)	(0.052)
Pseudo R^2	0.066	0.062	0.060	0.063	0.061
Observations	19584	7704	9417	7704	15882

a. All regressions include a constant.
b. Standard errors are in parentheses.
c. *, ** and *** significant at 10, 5 and 1 % significance level, respectively.
d. Datasource: General Social Survey
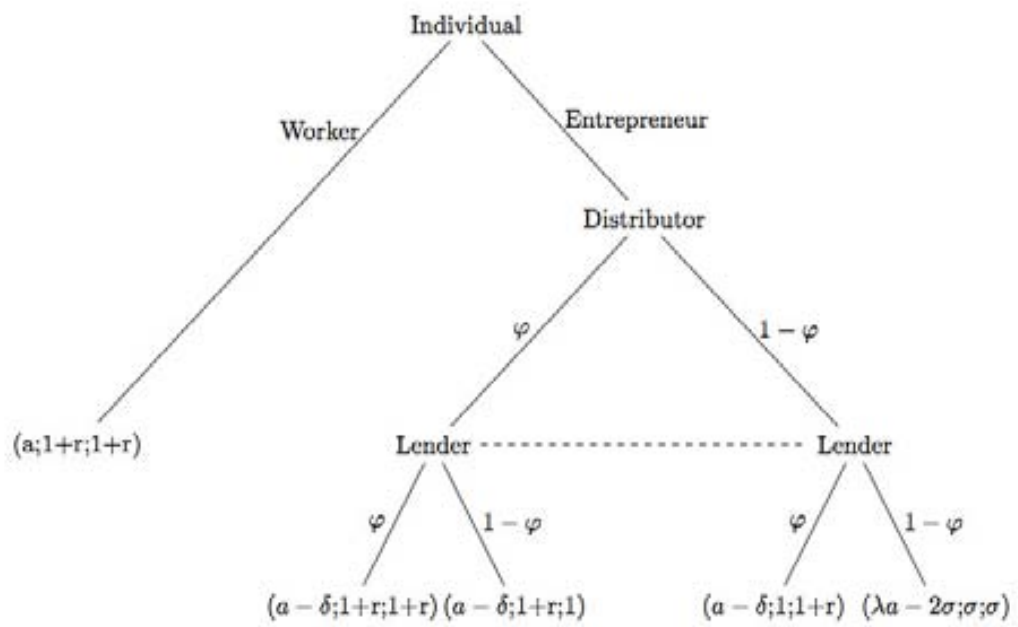
Figure 2.1: **Extensive game form of decision of B-type individual**
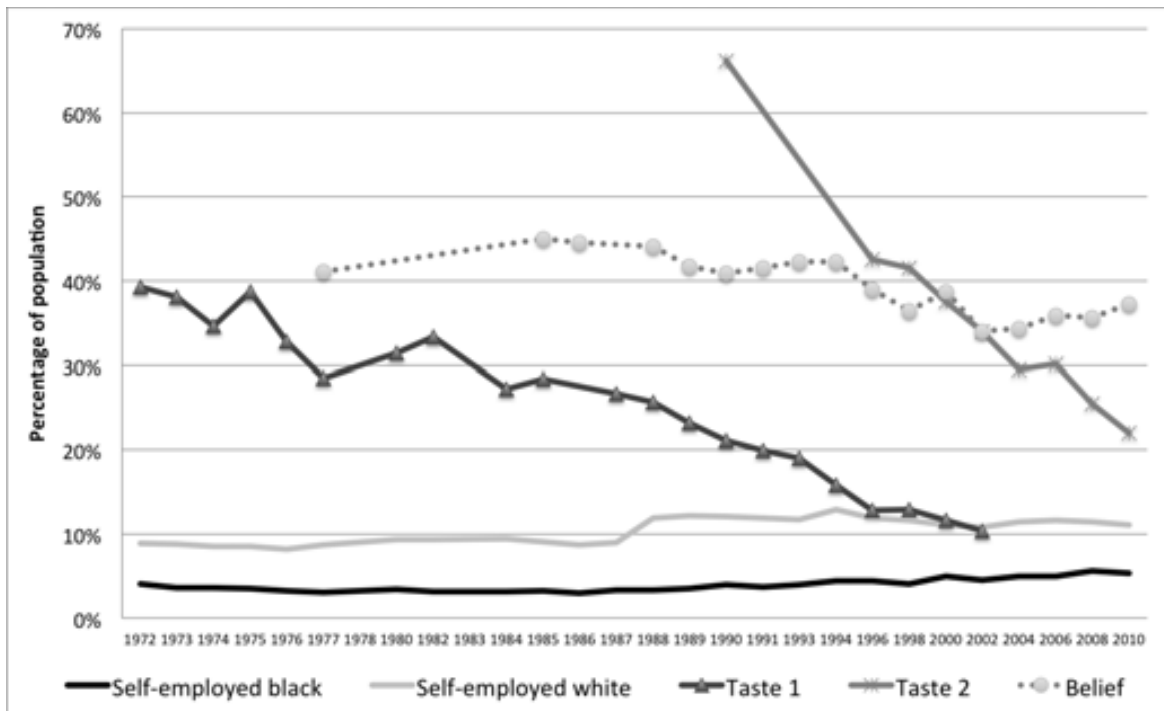
Figure 2.2: **Self-employment rates by race and beliefs and taste regarding discrimination in the US**

2.6 Appendix

2.6.1 Proofs

Proof. Proposition 1

1. First observe that by definition individuals in the ability range $a^* < a_i < a_b^*$ offer $1 + r \leq \sigma^N(a_i) < 1 + r + b$ and hence can potentially face discrimination. The fact that no individual in the range $a^* \leq a_i < a_b^*$ prefers the Nash bargaining solution to wage employment implies $(1 - \varphi_{T^*}^h)^2(\lambda a_i - 2\sigma^N(a_i)) + (1 - (1 - \varphi_{T^*}^h)^2)(a_i - \delta) < a_i$. The left hand side of the expression is strictly increasing in a_i, which implies that if it is not satisfied for a_b^* then it is not satisfied for all $a_i \leq a_b^*$. This will imply that all individual with a_i such that $\lambda a_i - 2(1 + r + b) \geq a_i$ (or all $a_b \leq a_i \leq a_b^*$) will offer a share equal to $(1 + r + b)$ and will be accepted and become entrepreneurs. All individuals with $a^* \leq a_i < a_b$ will be unable to offer a share to compensate the taste of discriminators and hence obtain wage employment. Moreover, note as now either all individuals of the B-type with $a_i \geq a_b^*$ offer either $\sigma^N(a_i) \geq (1 + r + b)$, and all individuals with $a_b \leq a_i \leq a_b^*$ will offer $(1 + r + b)$, implying that from period T^* onwards there will be no offers made within the range of $(1+r)$ to $(1+r+b)$, and hence the beliefs will remain frozen at the current level implying the above equilibrium will persist for ever.

2. Let us denote by a_{lb} as the lowest type individual in the range $a^* \leq a_i < a_b^*$ who prefers offering the Nash bargaining solution to seeking wage employment at T^*. As the Nash bargaining solution is strictly increasing in a_i it will imply all individuals with $a_i \geq a_{lb}$ will offer the Nash bargaining solution in period T^*. Now this means that all individuals in the ability range $a_{lb} \leq a_i \leq a_b^*$ will all offer the Nash bargaining solution and be accepted as no principals with taste for discrimination exist. Assume

this total number of cases for potential discrimination is n_{pot} and the actual cases of discrimination are zero. This implies that the point estimates in the next period $T^* + 1$ for meeting a discriminator is given by $\varphi_{T^*+1} = \frac{\alpha_0^l + \sum_{t=1}^{T^*} k_t}{\alpha_0^l + \beta_0^l + \sum_{t=1}^{T^*} n_{tb} + n_{pot}}$. This implies that $\varphi_{T^*+1} < \varphi_{T^*}$, implying the lowest type who applies in $T^* + 2$ is such that $a_i < a_{lb}$, or generalizing $\varphi_{T^*+t} < \varphi_{T^*}$ for all $t > 0$, or $\frac{d\varphi_t}{dt} < 0$ for all $t > T^*$. This implies at some point $\varphi_T \to 0$, implying all $a_i \geq a^*$ apply and become entrepreneurs and discrimination does not persist in society. The number of periods for which discrimination will persist could be calculated as a function of $\varphi_{T^*} = f(\pi_0, \omega)$ and the lowest a_i who still prefers the Nash solution to wage employment in T^*.

■

Proof. Proposition 2 Now assuming that the set $\mathcal{B}$ is large enough that individuals' abilities can be approximated by a continuous distribution on the interval $[0, 1]$. Proposition 1.1 shows that all individuals of the B-type with $a^* \leq a_i < a_b$ become workers instead of self-employed, while individuals with $a_b \leq a_i < a_b^*$ offer $1 + r + b > \sigma^N(a_i)$. Now we can write the transfer, which lenders and distributors receive through individuals wanting to escape the discrimination space as an integral: $\int_{a_b}^{a_b^*} 2(1 + r + b - \sigma^N(a_i))da = \int_{a_b}^{a_b^*} 2(1 + r + b - \frac{(\lambda-1)a+r+\delta+1}{3})da$. Therefore, the gain to the lenders and distributors would be $\frac{1}{3}((a_b^* - a_b)(4 + 4r + 6b - 2\delta) - (\lambda-1)((a_b^*)^2 - a_b^2))$. Now the loss to lenders and distributors, due to able individuals becoming workers, can be written as: $\int_{a^*}^{a_b} 2[\sigma^N(a_i) - (1 + r)]da = \int_{a^*}^{a_b} 2[\frac{(\lambda-1)a+r+\delta+1}{3} - (1 + r)]da$

$\Rightarrow$ Loss$= \frac{1}{3}((a_b^2 - (a^*)^2)(\lambda - 1) + (a_b - a^*)(2\delta - 4r - 4))$. In order for the loss to be at least as big as the gain we require: $((a_b^* - a_b)(4 + 4r + 6b - 2\delta) - (\lambda - 1)((a_b^*)^2 - a_b^2) \leq ((a_b^2 - (a^*)^2)(\lambda - 1) + (a_b - a^*)(2\delta - 4r - 4))$. Rearranging we get $(4 + 4r - 2\delta)(a_h^* - a^*) + 6b(a_b^* - a_b) \leq ((a_b^*)^2 - (a^*)^2)(\lambda - 1)$. Now substituting $a^* = \frac{2(1+r+\delta)}{\lambda-1}$, $a_b^* = \frac{2(1+r)+3b-\delta}{\lambda-1}$, and $a_b = \frac{2(1+r+b)}{\lambda-1}$ we find that for this to hold the condition is $b \geq \delta$. But since no

104

discrimination exists when $\delta > b$ (because in that case the offer $\sigma^N(a_i) > 1 + r + b$ for all $a_i \geq a^*$) gains can never be greater than losses if we are in an equilibrium as in proposition 1.1. $\blacksquare$

2.6.2   Individuals and Principals only Observe a Subset of Transactions

In this extension of the model we allow individuals and principals to observe just a subset of all market transactions. Relaxing the assumption that individuals and principals observe the entire set of market transactions is to take into account the fact that informational flows might be restricted to specific regions of the country or to members of specific occupations or a combination of both. For instance following Becker (1971), in the empirical exercise we assume that tastes and consequently beliefs are specific to each region of the country. Assume that out of the total k_t cases of discrimination and the total set n_{tb} of Nash bargaining offers with $1 + r \leq \sigma^n(a_i) < 1 + r + b$ the individuals and principals only observe a proportion $\gamma < 1$ of all cases.

Now under the original model, where the individuals and principals view the entire set of market transactions, the expected probability of meeting a principal with a taste for discrimination in period T^* is given by $\varphi_{T^*} = \frac{\alpha_0 + \sum_{t=1}^{T^*-1} k_t}{\alpha_0 + \beta_0 + \sum_{t=1}^{T^*-1} n_{tb}}$. Under the setting where only a γ proportion of cases can be observed the expected probability of meeting a principal with taste of discrimination in period T^* is given by $\varphi_{T^*}(\gamma) = \frac{\alpha_0 + \gamma \sum_{t=1}^{T^*-1} k_t}{\alpha_0 + \beta_0 + \gamma \sum_{t=1}^{T^*-1} n_{tb}}$.

Comparing the expected probability of meeting a principal with a taste for discrimination under the two settings, the point estimate is higher when the individuals and principals view only a subset of the cases if $(1 - \gamma)[\alpha_0 (\sum_{t=1}^{T^*-1}(n_{tb} - k_t)) - \beta_0 \sum_{t=1}^{T^*-1} k_t] \geq 0$.

If the above inequality is satisfied it is easy to see that if discrimination persists in the long run in the original setting, i.e. where the entire set of transactions are observed, it will also do so when only a proportion $\gamma < 1$ of transactions are observed.

If we assume that the above inequality is not satisfied then the expected probability of meeting a principal with taste for discrimination is higher when the entire set of market transactions is observed. Now assume in the original setting discrimination persists in the long run. If $\varphi_{T^*}(\gamma) < \varphi_{T^*}$ is such that no individual in the range $a^* \leq a_i < a_b^*$ still prefers the Nash bargaining solution to wage employment, then even when just a subset of the total transactions are observed discrimination will persist in the long run. On the other hand when the above inequality is not satisfied and γ is low enough such that under the original model no individual in the ability range $a^* \leq a_i < a_b^*$ prefers the Nash bargaining solution, but now some individual under the new setting does, this will imply that whereas in the original setting discrimination persists in the long run, in the setting where only a subset of transactions are observed it does not. The above discussion shows that relaxing the assumption of observability of the entire set of market transactions leaves the set of equilibria essentially unchanged.

2.6.3 Empirical appendix

Description of Controls Used in the Empirical Section

We include years of schooling as a proxy for ability. The argument that the effect of years of schooling might differ for blacks and whites, due to residential segregation and financial constraints forcing blacks into worse schools, does not hold for the given data. When including a dummy for blacks and an interaction term of blacks with years of schooling, the interaction term is insignificant. Gender we include because of the known inequality between men and women in the job market. It would be interesting to see whether it is belief about discrimination against women leading to lower self-employment rates among women, but unfortunately the GSS does not provide the appropriate questions to construct the necessary measures. We add self-categorised family income (far below average, below

average, average, above average, far above average) as a discrete variable taking values 1-5, in order to control for possible inherited wealth or family income, which could help in overcoming credit constraints. The possibility of taking over a family business is controlled for by including a dummy, specifying whether the father was self-employed. The predefined regions are New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain and Pacific.

Specifications of Endogenous Regime Switching Model

The model of self-employment $se_{i,t-1} \in \{0,1\}$ in the year before the survey can be specified as a probit equation:

$$h(s_{i,t-1}) = \beta_0 + \beta_1 a_i + \beta_2 \varphi_{i,t-1} + \beta_3 \pi_{i,t-1} + \gamma X_{i,t-1} + \theta_i + \varepsilon_{i,t-1} = x_i \delta + \theta_i + \varepsilon_{i,t-1}$$

$$\theta i \sim N(0, \sigma_\theta^2), \varepsilon_{i,t-1} \sim N(0, \sigma_\varepsilon^2), d_{i,t-1} = I(s_{i,t-1} = 1)$$

The variable X_i is a vector which determines whether the individual is self-employed or not. θ_i is an individual-specific unobservable component independent from his specified characteristics, and $\varepsilon_{i,t-1}$ is an unobservable random shock independent from his specified characteristics and θ_i. The function $h(.)$ is a monotonic transformation such that the distributive hypotheses on the error terms hold, while I is an indicator function equal to one when d holds, i.e. when the individual is self-employed. By assuming $\sigma_\varepsilon^2 + \sigma_\theta^2 = 1$, we get:

$$Prob(d_{i,t-1} = 1) = Prob(s_{i,t-1} = 1)) = Prob(h(s_{i,t-1}) = h(1)) = \Phi(h(1) - (x_i\delta)) = \Phi(x_i\xi)$$

where Φ is the standard normal cumulative density function, ξ gives the difference between $h(1)$ and the old constant in δ and coefficients associated with individual characteristics in

107

ξ are the same as in δ, but with the opposite sign.

Self-employment status in period t is specified conditionally on outcomes in $t-1$:

$$y(se_{i,t}) = z_i[\kappa_1 se_{i,t-1} + \kappa_2(1 - se_{i,t-1})] + \psi_i + \nu_{i,t}$$

$$\psi_i \sim N(0, \sigma_\psi^2), \nu_{i,t} \sim N(0, \sigma_\nu^2), d_{i,t} = I(s_{i,t} = 1)$$

where z_i is a vector specified characteristics affecting transition probabilities. Error components ψ_i and $\nu_{i,t}$ and the non-linear transformation $y(.)$ have interpretations analogous to those before, but for conditional self-employment. Again for the error components we assume $\sigma_\psi^2 + \sigma_\nu^2 = 1$. Individual-specific errors are jointly distributed as bivariate normal with $cov(\theta_i, \psi_i) = \rho$, whereas random shocks $\varepsilon_{i,t-1}$ and $\nu_{i,t}$ are uncorrelated. Given the assumptions on the errors' distribution, a generic likelihood contribution is:

$$L_i = \Phi_2(k_{i,t} z_i \tau_j, k_{i,t-1} x_i \xi_j; k_{i,t-1} k_{i,t} \rho)$$

$$k_{i,t} = 2d_{i,t} - 1, j = 2 - d_{i,t-1}$$

where Φ_2 is the standard normal cumulative density function and the τ's derive from the κ's in the same manner as ξ derives from δ. The components of τ_1 model the effect of individual characteristics on persistence in self-employment, whereas τ_2 captures the effect of the same characteristics on the probability of becoming self-employed. Therefore, we can summarize

$$Prob(se_{i,t} = 1 | se_{i,t-1}) = se_{i,t-1} \frac{\Phi_2(z_i \tau_1, x_i \xi; \rho)}{\Phi(x_i \xi)} + (1 - se_{i,t-1}) \frac{\Phi_2(z_i \tau_2, -x_i \xi; -\rho)}{\Phi(-x_i \xi)}.$$

As individual specific characteristics we control for the associated age and age squared. We assume individuals' education and gender to be constant across the two periods. Identifying restriction are required in order to estimate the model. In other words we require variable/variables that affect the probability of being self-employed in period $t-1$ but do not affect the transition in and out of self-employment in period t. We use the beliefs about discrimination in period $t-1$ as an instrument so that they are part of the vector X_i but not present in z_i.

2.6.4 Additional figures and tables

Table 2.2: **Logistic regression estimating self-employment in the US with fixed effects**

Dependent variable: Self-employment						
	(1)	(2)	(3)	(4)	(5)	(6)
Belief about discrimination	-2.909***	-4.652***	-2.545***	-4.870***	-2.788***	-4.635***
	(0.964)	(1.796)	(0.947)	(1.764)	(0.959)	(1.797)
Taste for discrimination 1	-0.353		-0.764		-0.446	
	(1.214)		(1.193)		(1.215)	
Taste for discrimination 2		-0.093		0.097		-0.064
		(1.447)		(1.435)		(1.460)
Years of schooling	0.031***	0.021*	0.036***	0.020	0.030***	0.020
	(0.012)	(0.013)	(0.011)	(0.013)	(0.012)	(0.013)
Female	-0.660***	-0.658***	-0.661***	-0.650***	-0.663***	-0.648***
	(0.067)	(0.071)	(0.067)	(0.072)	(0.067)	(0.072)
Age	0.101***	0.098***	0.098***	0.099***	0.100***	0.099***
	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.015)
Age squared x 1,000	-0.795***	-0.727***	-0.772***	-0.734***	-0.794***	-0.728***
	(0.140)	(0.154)	(0.139)	(0.153)	(0.140)	(0.154)
Family income at age 16	0.180***	0.132***	0.178***	0.125***	0.178***	0.128***
	(0.041)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Father was self-employed	0.588***	0.639***	0.576***	0.629***	0.590***	0.629***
	(0.068)	(0.075)	(0.069)	(0.075)	(0.069)	(0.076)
Year fixed effects	Yes	Yes	No	No	Yes	Yes
Region fixed effects	No	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.063	0.064	0.063	0.069	0.065	0.070
Observations	9417	7704	9417	7704	9417	7704

a. All regressions include a constant.
b. Standard errors are in parentheses.
c. *, ** and *** significant at 10, 5 and 1 % significance level, respectively.
d. Datasource: General Social Survey

Table 2.3: **Logit regression estimating self-employment with statistical discrimination**

Dependent variable: Self-employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Belief about discrimination		-4.852***	-4.922***	-4.635**	-4.760***	-4.800***	-5.036***	-4.567**	-4.868***
		(1.769)	(1.421)	(1.801)	(1.444)	(1.766)	(1.395)	(1.801)	(1.418)
Statistical discrimination	-30.238***	1.566	0.801	1.375	0.044	4.392	2.033	4.184	1.226
	(7.213)	(14.188)	(8.434)	(14.295)	(8.573)	(13.465)	(8.086)	(13.586)	(8.229)
Taste for discrimination 2		-0.159		-0.279		-0.513		-0.646	
		(2.390)		(2.415)		(2.377)		(2.407)	
Years of schooling	0.022*	0.021*	0.021*	0.021*	0.021*	0.020	0.020	0.020	0.020
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Female	-0.660***	-0.659***	-0.659***	-0.658***	-0.658***	-0.649***	-0.650***	-0.648***	-0.648***
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)	(0.072)	(0.072)	(0.072)	(0.072)
Age	0.098***	0.098***	0.098***	0.098***	0.098***	0.099***	0.099***	0.099***	0.099***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age squared x 1,000	-0.731***	-0.733***	-0.733***	-0.727***	-0.727***	-0.733***	-0.734***	-0.727***	-0.728***
	(0.153)	(0.153)	(0.153)	(0.154)	(0.154)	(0.153)	(0.153)	(0.154)	(0.154)
Family income at age 16	0.132***	0.128***	0.128***	0.131***	0.132***	0.125***	0.125***	0.128***	0.128***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Father was self-employed	0.647***	0.640***	0.640***	0.639***	0.639***	0.629***	0.629***	0.629***	0.629***
	(0.075)	(0.075)	(0.075)	(0.075)	(0.075)	(0.075)	(0.075)	(0.076)	(0.076)
Year fixed effects	No	No	No	Yes	Yes	No	No	Yes	Yes
Region fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.061	0.063	0.063	0.064	0.064	0.069	0.069	0.070	0.070
Observations	7704	7704	7704	7704	7704	7704	7704	7704	7704

a. All regressions include a constant.
b. Standard errors are in parentheses.
c. *, ** and *** significant at 10, 5 and 1 % significance level, respectively.
d. Datasource: General Social Survey

Table 2.4: **Estimating conditional self-employment in two-stage probit regression**

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A: Outcome of interest						
Beliefs about discrimination	-0.91***	-0.57***	-0.50***	-0.82***	-0.56***	-0.51***
	(0.038)	(0.045)	(0.053)	(0.038)	(0.045)	(0.053)
Taste for discrimination 1	-0.56***		-0.61***			
	(0.067)		(0.069)			
Taste for discrimination 2		-0.47***	-0.35***		-0.42***	-0.34***
		(0.045)	(0.062)		(0.046)	(0.064)
Statistical discrimination		-0.39***			-0.26*	
		(0.14)			(0.15)	
Panel B: Selection stage						
Beliefs about discrimination	-0.89***	-0.64***	-0.62***	-0.81***	-0.62***	-0.60***
	(0.041)	(0.050)	(0.060)	(0.041)	(0.051)	(0.061)
Taste for discrimination 1	-0.54***		-0.57***			
	(0.069)		(0.070)			
Taste for discrimination 2		-0.43***	-0.33***		-0.38***	-0.31***
		(0.046)	(0.063)		(0.047)	(0.065)
Statistical discrimination		-0.25			-0.18	
		(0.15)			(0.15)	
Region fixed effects	NO	NO	NO	YES	YES	YES
Observations	630,895	587,158	587,158	630,895	587,158	587,158

a. Controls: age, age squared, education and a constant.
b. Standard errors in parentheses.
c. *, ** and *** significant at 10, 5 and 1 % significance level, respectively.
d. Datasource: Variables "Beliefs about discrimination", "Taste for discrimination 1", "Taste for discrimination 2", and "Statistical discrimination" are constructed from observations of the General Social Survey. Source of remaining data: Integrated Public Use Microdata Series.
e. In all model specifications the covariance ρ between the individual specific error component is positive and significant at the 1% level, justifying the two-stage sample selection model.
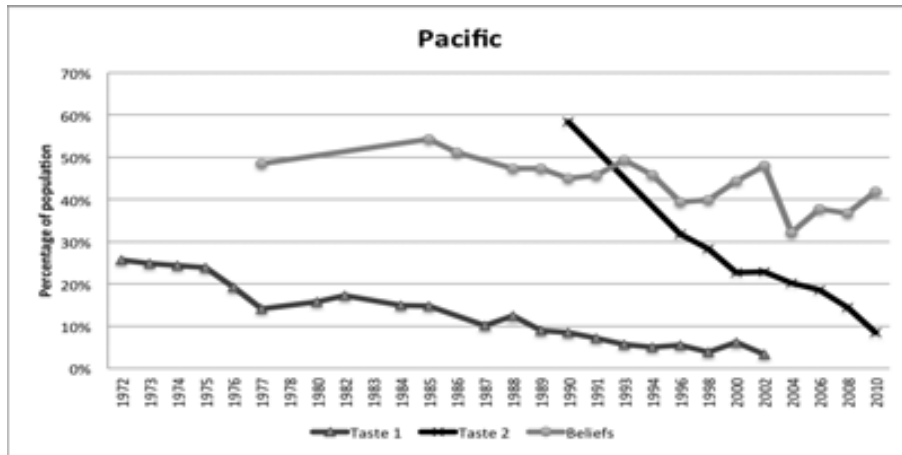
Figure 2.3: **Trends in tastes and beliefs regarding discrimination in the Pacific region**
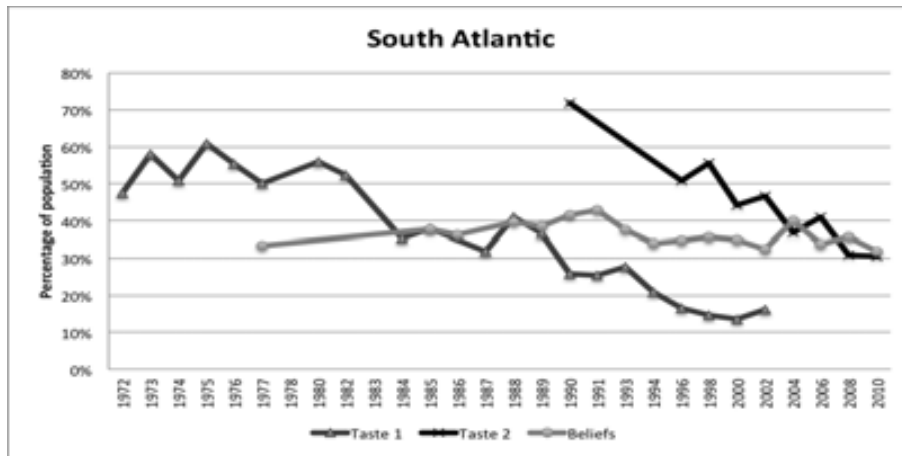
Figure 2.4: **Trends in tastes and beliefs regarding discrimination in the South Atlantic region**
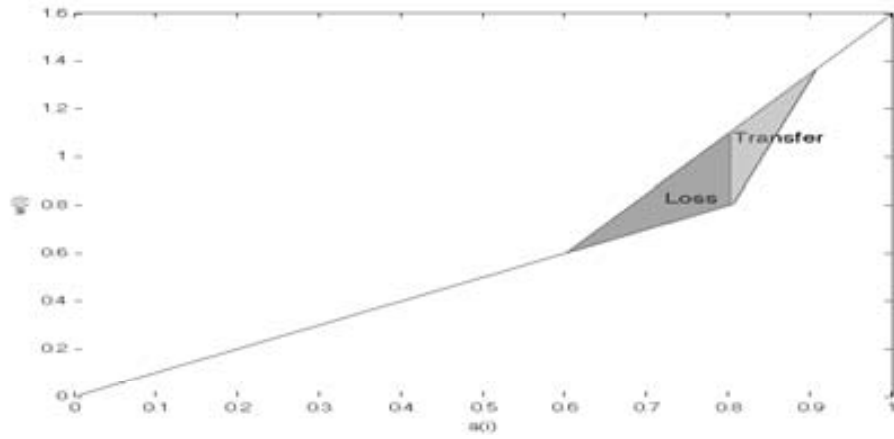
Figure 2.5: **Net wage in function of ability for individuals of type A and discriminated B**

In the specified example individuals of the A-type become self-employed when their ability $a_i \geq 0.6$, whereas only B-type individuals with $a_i \geq 0.8$ become self-employed and those with $0.8 \geq a_i \geq 0.9$ pay higher rates resulting in lower net earnings.
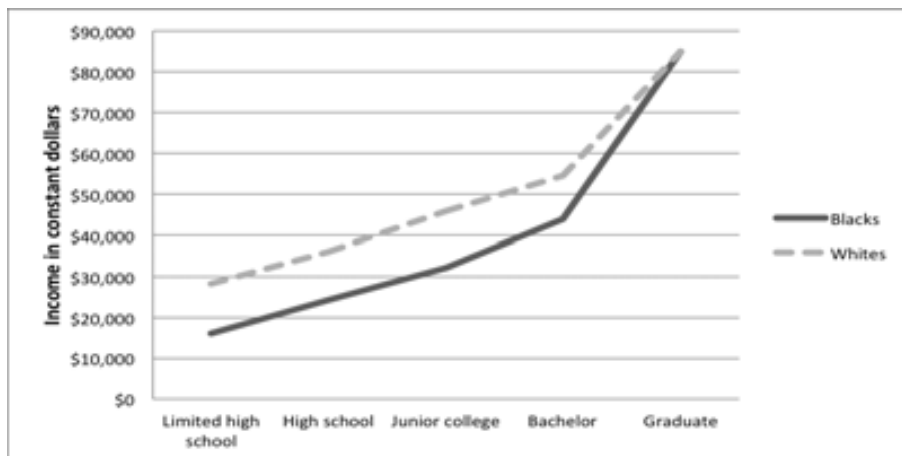
Figure 2.6: **Average income of black and white self-employed by educational degree in the General Social Survey**

Chapter 3

How Backward are the Other Backward Classes? Changing Contours of Caste Disadvantage in India
(joint with Ashwini Deshpande)

3.1 Introduction

The rise of the Other Backward Classes (OBCs) in the political arena since the mid-1980s has been heralded as India's "silent revolution" (Jafferlot, 2003). This political ascendancy has also been viewed as representing a large enough flux in the traditional hierarchies of the caste system, such that we now have "a plethora of assertive caste identities... [that] articulate alternative hierarchies" leading to a scenario where "there is hardly any unanim-

ity on ranking between jatis" Gupta (2004). Indeed, there is no doubt, especially since the 73rd and 74th constitutional amendments in the early 1990s, that the so-called lower castes have become an important force in Indian politics at all levels, local, state and national. Has this change in the political arena been accompanied by a corresponding reshuffling of the traditional economic hierarchies, such as to prevent any meaningful ranking of castes?

The nature and degree of change in the economic ranking between castes, or broad caste groups, is a matter of empirical verification. While there is a large and growing body of work documenting the changes in the standard of living indicators of the Scheduled Castes and Tribes (SCs and STs), as well as the economic discrimination faced by these groups, (see Deshpande 2011, for a review of the recent research), the discussion about the material conditions or the economic dominance of the group of castes and communities classified as the "Other Backward Classes" (OBCs) in India is prompted more by beliefs, or localised case studies, rather than by an empirical analysis of the macro evidence. Part of the reason for this lacuna is the lack of hard data: until the 2001 census, OBCs were not counted as a separate category, while affirmative action (quotas in India) were targeted towards OBCs at the national level since 1991, and at the state level since much earlier. This would be the only instance of an affirmative action anywhere in the world where the targeted beneficiaries of a national programme are not counted as a separate category in the country's census.

Researchers have, therefore, had to rely on data from large sample surveys such as the National Sample Survey (NSS), National Family and Health Survey (NFHS), to mention a few sources, in order to get estimates about the material conditions of the OBCs. The use of this data has generated research which undertakes a broader analysis of various caste groups, OBCs being one of the groups in the analysis, along with the SC- STs and Others', the residual group of the non-SC-ST-OBC population (for instance, Deshpande 2007; Iyer

116

et al. 2013; Madheswaran and Attewell 2007; Zacharias and Vakulabharanam 2011, among others). Others' include the Hindu upper castes and could be considered a loose approximation for the latter, but data constraints do not allow us to isolate the upper castes exclusively. Existing evidence suggests that OBCs lie somewhere in between the SC-STs and the Others', but first, very little is known about their relative distance from the two other categories and second, in order to make a meaningful intervention about the possible links between their political ascendancy and their economic conditions, it is important to trace how their relative economic position has changed vis-a-vis the other two groups over time. Here again, the economic researcher is stymied by the lack of good longitudinal data.

The present paper is an attempt to fill this caveat in the empirical literature by focusing on an important facet of contemporary caste inequalities, viz., the changing economic conditions of OBCs, relative to the other two broad social/caste groups. We use data from two quinquennial rounds of the employment-unemployment surveys (EUS) of the NSS for 1999-2000 and 2009-10 (NSS-55 and NSS-66, respectively), to examine the multiple dimensions of material standard of living indicators, and the changes therein for the OBCs in India, in comparison to SC-STs (for the purpose of this paper, we have pooled the two groups, because despite considerable differences in their social situation, their economic outcomes are very similar), and the Others'. We look at five age cohorts between 25 and 74 years of age in each NSS round, and examine changes in multiple indicators using a difference-in-differences (D-I-D) approach, comparing the three social groups to one another over consecutive cohorts to see how the gaps on the key indicators of interest have evolved over the 60 year period. This allows us to gauge the relative generational shifts between the major caste groups. Our analysis focuses particularly on the OBCs, and compares how the evolution of the different OBC cohorts (in relation to the Others') compares with the evolution of the corresponding SC-ST cohorts to the Others'.

117

Through an analysis based on a comparison of different age cohorts, we are able to build a comprehensive trajectory of change for each of the caste groups since independence, since the oldest cohort in our analysis consists of individuals born between 1926 and 1935, and the youngest cohort consists of those born between 1976 and 1985. Thus, we are able to track outcomes for successive generations of individuals who reached adulthood in the 63 years between Indian independence (in 1947) and 2010.

We start by examining the household level aggregates, such as monthly per capita expenditure (MPCE), proportion of urban population and two landholding measures, and then move to individual indicators, specifically, education, occupation (which focuses on occupation categories as well as the principal activity status and changes in the Duncan dissimilarity index based on activity status) and finally wages and Blinder-Oaxaca estimates of labour market discrimination.

Our main results can be summarized as follows. In a three-fold division of the population between SC-ST, OBCs and Others', we see clear disparities in virtually all indicators of material well-being, with Others' at the top, SC-STs at the bottom and OBCs in between. This confirms the results from several other studies. The average gaps between the Others' and the other two social groups however remain large. MPCE, an indicator of standard of living in developing countries, shows that the average MPCE of the OBCs and SC-ST is 51 and 65 percent of the Others', respectively. Similarly the gap between Others' and OBCs for the composite indicator of years of education remains as large as 2.21, whereas the gap between SC-ST and OBCs is 1.47 years of education. The average wages of the OBCs and SC-ST are seen to be only 42 and 55 percent of the average wage of Others' and the share of labour force employed in white collar prestigious jobs is just one-fourth and one-half the proportion of the Others' employed in white collar jobs.

Breaking down the indicator of years of education, we find evidence of convergence

118

between OBCs and Others' in literacy and primary education, but continued divergence when higher educational categories are considered. In the realm of occupation, the younger cohorts among OBCs seem to be closing the gap vis-a-vis the Others' in terms of access to prestigious white-collar jobs. Based on principal activity status, our calculations of the Duncan Index reveal that OBCs are closer to the Others' (less dissimilar to them) as compared to the SC-STs (who are more dissimilar compared to the Others'). For the category of regular wage/salaried (RWS) jobs we find divergence between the Others' and OBCs and SC-ST except for the very youngest cohort. Looking at average wage gaps for males in the labour force and estimates of labour market discrimination, we find that while average wages of Others' are higher than those for OBCs for all age cohorts, the unexplained (or the discriminatory) component is lower for younger OBC cohorts, compared to the older ones, and that OBCs face lower labour market discrimination compared to SC-STs, when the average wages of both groups are compared to those of Others'.

3.2 The broad picture: household-level indicators

Table 3.1 presents estimates of some indicators of standard of living for three major caste groups: SC-STs considered together, OBCs and Others', for NSS-55 and NSS-66 respectively. The indicators of interest are MPCE, proportion of the group that is urban (percent urban) and two land holding measures: land owned and land possessed.

D-I-D for household-level variables is calculated as:

$$D-I-D_{jk} = [(Indicator_{ijs} - Indicator_{iks}) - (Indicator_{ij(s-1)} - Indicator_{ik(s-1)})] \quad (3.1)$$

where j and k are the two caste groups being compared, for the ith indicator (say MPCE) between survey rounds s and $s-1$.

119

MPCE is shown in nominal terms: Others' have the highest MPCE, followed by OBCs, and then the SC-STs. While the MPCE for each of the groups has expectedly increased in nominal terms, the D-I-D allows us to see the relative gains of groups. Between 1999-00 and 2009-10, we see that the MPCE gap between OBCs and SC-STs has increased by Rs. 173 in favour of OBCs. However, for the OBCs, MPCE has fallen behind that of the Others' by Rs. 428 over the decade. Others' MPCE has increased by Rs. 600 relative to SC-STs over the decade (note that all D-I-D are significant at the 1 percent level). Thus, SC-STs not only continue to have the lowest MPCE, but the other two groups have gained relative to them in terms of MPCE. OBCs have gained relative to SC-STs, but the magnitude of their falling behind Others' is over 2.5 times their gain over SC-STs. Thus, on MPCE, there is no evidence of convergence between Others', either with OBCs or with the SC-STs.

[Insert Table 3.1]

Urbanisation (percent of the group's population which is urban) is an indicator of structural change or of potential integration into the modern, formal sector economy. We see a rise in urban proportions for both OBCs and Others' (at 28 and 43 percent respectively in 2009-10, but virtually no change for the SC-ST population at around 17 percent). Again, looking at relative changes across groups using D-I-D, we find the same pattern as that for MPCE, but the relative gain of Others' over OBCs is only about 2 percentage points. The percentage of population classified as urban for OBCs and Others' increased between 3.3 and 3.55 percentage points relative to SC-STs (note that all D-I-D are significant at the 1 percent level).

The two land holding variables (land possessed and land owned) show sharp disparities in across caste groups in both rounds, with average values for SC-STs slightly over half of the values for Others'. However, in terms of the relative change in these two variables, we see that OBCs marginally fell behind SC-STs by 0.01 hectares for land possessed, but

gained over Others' by close to 0.05 hectares. [1] SC-STs appear to have gained over Others' in both land owned and land possessed by 0.017 and 0.059 hectares respectively. These changes are negligible in magnitude to have any real consequences for standard of living, and are clearly not matched by trends in MPCE.

Overall, at the household level, we see a clear hierarchy in MPCE, such that Others' are at the top, followed by OBCs and then SC-STs. Over the decade, the gap between OBC and SC-STs has increased in favour of the former, and Others' MPCE has increased relative to both SC-STs and OBCs, but the magnitude of gain has been larger vis-a-vis the SC-STs than OBCs.

3.3 Individual-level characteristics: Education

3.3.1 The construction of cohorts

We construct five age cohorts using the age variable in each of the NSS rounds as follows:

[Insert Table 3.2]

From their age, we can determine their birth year (relative to year 2000 and 2010, i.e. the end years of the survey respectively) and thus, over the two rounds we are able to get information for six cohorts, with the oldest being born between 1926-1935 and the youngest cohort of individuals born between 1976-1985. As can be seen from the table above, matching years of birth implies that Cohort 2 in NSS-55 is Cohort 1 of NSS-66, Cohort 3 in NSS-55 is Cohort 2 in NSS-66 and so on.

3.3.2 Years of education

The first indicator of interest that we consider is education. Figure 3.1 plots the evolution of *years of education* for the six cohorts over the two rounds.[2]

We see that all three groups have increased their average years of education over the first five cohorts for both the rounds of the NSS. The oldest cohort aged 65-74 in 2000 (NSS-55) has 0.70 years of education for the SC-STs, 1.14 years for OBCs and 3 years of education for Others'. We see that these increase steadily and stand at 4.52, 6.09 and 8.30 respectively for Cohort 6, aged 25-34 in the year 2010 (Cohort 5 of NSS-66). The average years of education for the OBCs over the 50-year period increases by 4.95 years, whereas it increases by 3.92 years for the SC-ST and 5.3 years for the Others' over the same period.

We calculate the D-I-D over consecutive cohorts defined as follows:

$$D-I-D_{jk} = [(Indicator_{ijn} - Indicator_{ikn}) - (Indicator_{ij(n-1)} - Indicator_{ik(n-1)})] \quad (3.2)$$

where j and k are the two caste groups being compared, for the ith indicator, first for the nth cohort and then for the $n-1^{th}$ cohort (results on the D-I-D and its significance for key indicators of education are presented in the Table 3.13 in the appendix).

The evolution of D-I-D for years of education for the 6 cohorts is shown in Figure 3.2. This shows us that the OBCs lose around 0.36 years of education compared to Others' when we compare Cohort 2 with Cohort 1, i.e. gap between the OBCs and the Others' increases from 1.85 years of education to 2.21 years of education. For the first cohorts who entered schooling after independence, we see again divergence in the country. A comparison of Cohorts 3 and 2, for the OBCs and the Others', shows that the gap between the 2 groups increased by 0.50 years of education. A comparison of Cohorts 4 and 3 for the OBCs and the Others' shows that the gap increased again by 0.08 years, where the D-I-D is insignificantly different from zero (note that all other D-I-D are significant at the 1 percent level). After this we see that the OBCs gain about 0.16 years and 0.42 years of education

when we compare Cohorts 5 and 4 and Cohorts 6 and 5, respectively. The gap between the Cohort 6 of the Others' and OBCs is around 2.21 years of education, increasing from the gap of 1.85 years observed for Cohort 1. The fact that the difference between the two groups, for years of education, has increased when looking at the individuals born between 1926-35 in relation to 1976-85 seems to suggest that overall, in the big picture, convergence seems to be absent.

[Insert Figure 3.2]

An alternative way of comparing the evolution of the gaps could be to compare the oldest cohort who went to school after independence with the youngest cohort. This would imply comparing the cohort born in 1946-55 (Cohort 3) to the ones born in 1976-85 (Cohort 6). This comparison presents a more optimistic picture as the gap between the OBCs and Others' for the cohort born in 1946-55 was 2.71 years of education, which reduces to 2.21 years for the cohort born in 1976-85.

Similarly, when we compare SC-STs with the Others', the picture is not very optimistic. Again we see that the gap in average years of education for the cohort born in 1926-35 (Cohort 1) is around 2.29 years. This gap, in fact, increases to 3.68 years of education, when looking at the last cohort born in 1976-85. This seems to suggest a picture of divergence rather than convergence in the country. Alternatively, comparing the oldest cohort who went to school after independence with the youngest cohort (Cohort 3 and Cohort 6) for SC-STs with the Others' suggests a gain for the SC-STs of 0.08 years of education, which is insignificantly different from zero.

3.3.3 Other indicators of educational attainment

In order to better understand the picture of evolution of the three social groups on educational attainment, we now look at four separate categories of education, namely, the proportion of each cohort literate or more, has finished primary schooling or more, has finished secondary schooling or more and finally is a graduate or has higher education.[3]

For the category literate or more, the proportion of the cohort born in 1926-35 which was literate was 15 percent, 25 percent and 46 percent for SC-STs, OBCs and Others', respectively. This increased to 63, 73 and 86 percent respectively for the cohort born in 1976-85. Looking at the evolution of the OBCs in relation to Others' shows a picture of steady convergence in the country. The gap between the two groups was such that that 21 percent more of the Others' were literate as compared to the OBCs for Cohort 1, and this decreases to 13 percent for Cohort 6. Comparing SC-STs to the Others' also shows a pattern of convergence where the gap reduces from 31 percent more of Others' being literate for Cohort 1 to 23 percent for Cohort 6.[4]

The picture for the category "primary education and more" is very similar to the picture for literacy and more. For the cohort born in 1926-35, the proportion that has primary education or more, stands at 7, 13 and 31 percent for the SC-STs, OBCs and Others', respectively. This increases to 51, 64 and 78 percent respectively for the Cohort 6 born in 1976-85 and aged 25-34. The gap between Cohorts 2 and 6 for the OBCs and the Others' reduces from 20 percentage points to 14 percentage points. Similarly, comparing SC-STs with Others', the gap reduces from 32 percent to 26 percent. The convergence is especially strong for the last 3 cohorts of the OBCs, who gain 8 percentage points relative to the Others'.[5]

The next category of education we examine is all those with "secondary education or more". For the cohort born in 1926-35, 2 percent of SC-STs, 3 percent of OBCs and 13

percent of Others' have secondary education or more. This increases to 19, 30 and 48 percent respectively for Cohort 6 born in 1976-85. The evolution of the OBCs and SC-STs in relation to the Others' suggests that contrary to the earlier categories, the picture for this category of education has been one of divergence rather than convergence. Again, comparing the gap between the two groups for Cohorts 1 and 6 suggests a picture of divergence. 10 percent more of Cohort 1 had secondary education or more for the Others' as compared to the OBCs. This gap, in fact, increases to 18 percent for Cohort 6 born in 1976-85. Similarly, for SC-STs the gap increases from 11 percent more of Others' having secondary education or more for Cohort 1 to about 29 percent for Cohort 6.[6]

For the last category of education, those with a graduate degree or more, for the cohort born in 1926-35, 0.5 percent of SC-STs, 0.4 percent of OBCs and 4 percent of Others' had a graduate degree or more. This increases to 4.7, 9 and 20 percent respectively, for the cohort born in 1976-85 (see Figure 3.3).

[Insert Figure 3.3]

Comparing the gap between the OBCs and Others' for Cohort 2 (which is Cohort 1 in NSS-66) shows that 6 percent more of Others' had a graduate degree and this gap, in fact, increases to 10.5 percent for Cohort 6 born in 1976-85, suggesting divergence in this category of education. The SC- ST with Others' comparison again shows a picture of divergence. The gap between SC-ST and Others' for the cohort born in 1935-46 (Cohort 2) was 7 percent, which increases to 15 percent for the cohort 6 born in 1976-85 (See Figure 3.4 and note that all D-I-D are significant, refer to table 3.13 in appendix).[7]

[Insert Figure 3.4]

3.3.4 The overall picture in education indicators

The overall picture suggests that there seems to be convergence between the Others' and the two socially disadvantaged groups, SC-STs and OBCs when lower categories of educational attainment, namely, literacy and primary schooling are considered. However, this picture is overturned when higher categories of education, viz., secondary schooling or higher, and graduate degree or higher are considered. The composite index of years of education suggests a picture of no change in the gap when the OBCs and Others' are compared for the cohort born in 1936-45 and for the cohort born in 1976-85, and a divergence by 0.36 years when the cohort born in 1926-35 is compared to the one born in 1976-85. This result for the OBCs and Others' is overturned when we compare the oldest cohort (born in 1946-55) that went to school after independence with the youngest cohort that would have finished schooling by 2010 (born in 1976-85). Such a comparison suggests that the OBCs have gained on an average 0.50 years of education, as compared to the Others', over the 40 year period, even though the current gap between the two groups remains as large as 2.21 years of education. On the other hand, comparing SC-STs and Others' for the cohorts born in 1936-45 and 1976-85 suggests that SC-STs fell back by 0.50 years of education more over the 50-year period, and the current gap between the two groups remains as large as 3.70 years of education.

The fact that on the higher categories of education, which would be critical to achieve social mobility, traditional hierarchies have not only persisted but widened over the 50 year period is noteworthy. This indicates that policies targeted towards closing the gaps at the higher education levels are not entirely misplaced, as the lower educational levels are witnessing a convergence between broad caste groups, but higher levels are not, and hence targeted policies would be needed to close those gaps.

3.3.5 The education transition matrix

The above analysis has analysed shifts across birth cohorts. We can go further to examine generational shifts. In order to do that, we go on to construct a matrix which depicts the transitional probabilities of the son's education belonging to a particular education category given the fathers level of education.

We construct six categories of education as follows: 0 representing illiterate; 1 representing literacy but less than primary schooling; 2 representing more than primary schooling but less than secondary; 3 representing more than secondary but lower than higher secondary; 4 representing more than higher secondary but lower than graduate; and 5 representing graduate education and higher. We then match the male head of household's category of education to his son's category of education for the NSS-55 and NSS-66.

The transition matrix provides us easy visual representation of the underlying intergenerational mobility in education for the three social groups. This helps us understand whether the pattern of increasing educational attainment which we observed above is driven by sons of household heads with high education obtaining even higher education (i.e. intergenerational persistence), or is it due to the upward movement of sons whose fathers had low education moving up the ladder (intergenerational mobility).

The transition matrix shown in the table below computes the probability p_{ij} – the probability of a father with education category i having a son in educational category j. A high p_{ij} where $i = j$ represents low intergenerational education mobility, while a high p_{ij} where $i < j$, would indicate high intergenerational education mobility. The last column of the table labelled "size" shows the proportion of fathers in that particular educational category.

So, for instance, from Table 3.3 we see that in NSS-55, the proportion of SC-ST fathers that were illiterate was 59.66 percent. Given that the father was an illiterate, the probabil-

ity of a son from a SC-ST family being illiterate was 40.89 percent, being literate was 11.8 percent, having primary but less than secondary was 31.68 percent, having secondary but less than higher secondary education was 8.9 percent, having more than higher secondary but less than graduate was 4.6 percent, and finally holding a graduate degree or higher was 2.1 percent. Similarly the proportion of OBC fathers who were illiterate was 46.44 percent in 1999-2000. The probabilities of the son being in education categories 0 to 5 were 35.75, 11.58, 34, 11.03, 5.52 and 2.1 percent respectively. Finally, 26.5 percent fathers in the Others' category were illiterate, and probabilities of the son being in categories 0 to 5 were 26.68, 12.14, 38.21, 14.14, 5.6 and 3.2 percent respectively.

[Insert Table 3.3]

Comparing the transitional probabilities of NSS-55 in Table 3.3 with those of NSS-66 in Table 3.4, we first observe that for all three social groups there is an increase in the average proportion of fathers in higher educational categories. For instance, the proportion of fathers with more than primary schooling but less than secondary schooling increases from 17.45 to 22.85 percent, 23.98 to 29.87 percent and 27.87 to 29.80 percent for the SC-STs, OBCs and Others' respectively. We also observe that for sons whose fathers had education category 3, 4 or 5, the probability of the son achieving an educational category equal to or higher than their father increases for all three groups, i.e. intergenerational persistence is high for families with higher levels of education. For instance, for the probability of the father belonging to the education category 3 (more than secondary but lower than higher secondary) and his son belonging to the category 3, 4 or 5 increases from 73.8 to 75.9 percent, 72.8 to 85 percent and 82.1 to 87.8 percent for the SC-STs, OBCs and Others' respectively.

[Insert Table 3.4]

128

Having said this, it should be noted that conditional on fathers' education, sons from the social group Others' are more likely to achieve an education category equal to or higher than their father as compared to SC-STs and OBCs. So, for instance, in 2009-10, for fathers with education category 5 (graduate education and higher), the probability that the son also achieves educational category 5 is 37.8, 33.56 and 54.01 percent for the SC-ST, OBCs and Others', respectively. The reading of the matrix suggest that the ability of highly educated parents to ensure an equivalent or higher education level for their children is best reaped by the Others'. The fact that SC-ST sons have a higher probability to be graduates and above, compared to the OBCs, contingent upon their fathers being graduates suggests that reservations for SC-STs in higher education might be playing a role. The fact that the reservation for SC-ST have been in operation much longer, than for OBCs, could be resulting in producing a greater share of graduates among SC-STs in families where the fathers are also highly educated. It is likely that the SC-ST sons are second-generation beneficiaries of reservations. Also, the calculated transitional probabilities suggest that the conversion of parents' endowment of education into human capital of children is highest for people from the socially privileged, i.e. non-backward groups.

3.3.6 Ordered probit regressions for education categories

We ran an ordered probit regression to calculate the marginal effects of being in five educational categories defined as follows: Education category 1: not literate; category 2: literate, below primary; category 3: primary; category 4: middle; category 5: secondary and above. Table 3.5 shows the probabilities of being in each of these categories for OBCs and SC-STs relative to Others'. We see that all cohorts of OBCs and SC-STs are significantly more likely to be illiterate (category 1) than Others'. The marginal effects rise from Cohort 1 to 3 and decline thereafter, such that between Cohort 1 and 5, the

likelihood of OBCs being illiterate as compared to the Others' reduces from 20.6 percent to 7.2 percent. We see a similar trend for SC-STs as well, but first, their likelihood of being illiterate relative to Others' is higher than that for OBCs and second, the decline in this probability over successive cohorts is lower than that for OBCs.

[Insert Table 3.5]

For higher educational categories, the trend in probabilities changes. For category 2, i.e. literate, below primary, we see that the three youngest cohorts of OBCs show positive marginal effects compared to the Others', indicating convergence. For the next higher category, we see that only the two youngest cohorts of OBCs show positive marginal effects. For the last two educational categories (middle and secondary and above), all cohorts of OBCs are less likely to be in these categories than the Others', confirming the D-I-D result that after the middle school level, we see divergence, rather than convergence in educational attainment.

3.4   Occupation

How does the evolution of differences in educational attainment translate into occupational differences between groups? To start this investigation, we first estimate the number of individuals in the labour force.[8] We then aggregate these individuals into three categories: those with agricultural jobs, blue- collar jobs and white-collar jobs.[9]

In 1999-2000, based on NSS-55, for the first cohort born in 1926-35, the proportion of those in agricultural jobs was 78.85 for SC-ST, 74.55 for OBC and 71.85 for Others'. Over successive cohorts, we see that for all groups, proportion of individuals in agricultural jobs declines, to stand at 51.28, 46 and 35.46 respectively for Cohort 4 in NSS-66 (those who are 35-44 years old in 2010).[10] For blue-collar jobs, proportions for Cohort 1 in NSS-55 for

the three groups are 17.78, 21.68 and 18.97 respectively, which have doubled for Cohort 4 in NSS-66 to stand at 40.4; 41.1 and 39.57 respectively. This illustrates the shift away from agriculture towards secondary and tertiary sectors respectively. We also note that gaps between groups in agricultural occupations are sharper than those for blue-collar jobs. The decline in proportions in agricultural jobs is matched by an increase in proportions with blue-collar and white-collar jobs, reflecting the structural shift in the economy, where the proportion of the population dependent on agriculture is declining over the last several decades.

The other notable feature of the occupational division is of sharp inter-caste disparities in access to these broad occupations. In NSS-55, SC-STs record the highest proportion in agricultural jobs consistently for all cohorts, followed by OBCs and Others'; whereas for white-collar jobs, Others' record the highest proportions for all cohorts, followed by OBCs and then SC-STs. For blue-collar jobs, the picture is mixed, in that OBCs record the highest proportions, followed by Others' and then SC-STs. A decade later, our calculations with NSS-66 reveal a similar pattern in caste disparities, with proportions of different caste groups in blue-collar jobs closer to each other, and with OBCs having a slight edge over the other caste groups. [11]

3.4.1 Evolution of White-Collar Jobs

For the most prestigious white-collar jobs, caste disparities remain substantial: from 3.37 (SC-ST); 3.76 (OBC) and 9.18 (Others') percent respectively for Cohort 1 in NSS-55, the shares of the three groups stand at 8.32; 12.93 and 24.97 respectively for Cohort 4 in NSS-66 (see Figure 3.5). However, we need to examine D-I-D between cohorts across groups in order to understand the relative change between successive generations across the three caste groups.

131

For Cohort 1 (NSS-55), share of OBCs in white-collar jobs is 5.4 percentage points less than the Others' and that of SC-STs is 5.81 percentage points less than the Others'. Looking at Cohort 5 (i.e. Cohort 4 in NSS-66), we find that the gap between OBCs and Others' has increased to 12.04 percentage points and that between SC-STs and Others' has increased to 16.65 percentage points. Thus, the share of OBCs and SC-STs in white-collar jobs has lagged behind that of the Others', but by a greater percentage for the latter.

Looking at the evolution of D-I-D in share of the population in white-collar jobs (Figure 3.6), we see that OBCs in absolute terms are clearly ahead of the SC-STs, although still substantially lower than Others' (the evolution and statistical significance of the calculated D-I-D are shown in Table 3.14 in the appendix). D-I-D between Cohort 2 and 1 reveals that shares of OBCs and SC-STs in white collar jobs further falls behind 2 and 5 percentage points less compared to the Others'. For the SC-ST only Cohort 5 seems to close the gap with the Others' though the gap for the cohort aged 35-44 remains as large as 17 percentage points. SC-STs continue to lag behind Others' in terms of their access to white collar jobs. This is reflected in the overall D-I-D between SC-STs and Others', whether measured as the gap between Cohort 5 and 1 (-11 percentage points) or between Cohorts 5 and 2 (-6 percentage points). The OBCs, on the other hand, are behind the Others' by 5 percent points comparing the gap between Cohort 1, but after Cohort 4 through successive cohorts, continue to gain vis-a-vis the Others'. Thus, while the larger picture (comparing the gap with Others' for cohort 1 (from NSS-55) with a similar gap for cohort 5 (cohort 4 from NSS 66th), suggests divergence, as the gap has increased, focusing on a slice of younger cohorts alters the picture. Their overall D-I-D relative to Others', if measured as the gap between

132

Cohort 5 and Cohort 2 in NSS-66, suggests that OBCs have converged with the Others' proportion by 3.2 percentage points. Given that NSS-66 is the latest survey, the D-I-D evidence from this survey is a clearer indication of the contemporary trends, which suggests that OBCs are catching up with the Others' in access to white collar jobs, whereas SC-STs continue to lag behind. Given the presence of quotas in public sector and government jobs, the continued lagging behind of SC-STs possibly indicates continued gaps in the private sector.

3.4.2 Public sector jobs

We can examine this more directly by looking only at access to public sector jobs, one of the sites for affirmative action, which in India takes the form of caste-based quotas (22.5 percent for SC-ST). Additional 27 percent quotas for OBCs were introduced at the national level (i.e. for central government jobs) in 1990; various state governments introduced state-specific OBC quotas at different points in time after 1950. Public sector jobs, even those at the lowest occupational tier, are considered desirable because most offer security of tenure and several monetary benefits, such as inflation indexation, cost-of-living adjusted pay, provident fund, pensions and so forth. The private sector wage dispersion is larger, so there is a possibility of far greater pay at the higher end, but the private sector is an omnibus category covering very heterogeneous establishments, with large variability in the conditions of work and payment structures.

[Insert Table 3.6]

Looking at Table 3.6 based on NSS-66, we see that SC-ST percentages with access to public sector jobs are consistently higher than those for OBCs, which is at variance with the access to white collar jobs, discussed above. We believe that the difference in the relative picture between SC-STs and OBCs reflects the longer operation of SC-ST quotas. Others' have

133

the highest percentage of public sector jobs across cohorts. The D-I-D reveals that OBCs are catching up, both with SC-STs and Others' (the evolution and statistical significance of the calculated D-I-D are shown in Table 3.14 in the appendix). This is most strikingly true for cohort 3 of NSS-66, born between 1956-1965, individuals who would have been between 35 and 25 years old in 1990 and hence eligible to take advantage of the new quotas. This catch-up continues onwards to cohort 4. We see a similar convergence between SC-ST and Others', which is in contrast to the picture of divergence between SC-ST and Others' in access to white-collar jobs.

Within the public sector, white and blue-collar jobs present different scenarios. The result of quotas can be clearly seen here. Take a representative example. 6.51 percent SC-ST, 13 percent OBCs and 26.29 percent of Cohort 3 of NSS-66 (Cohort 4 of the six cohorts) are in white-collar jobs. But of these, 36 percent of (the 6.51) SC-ST, 21.2 percent OBCs and 24.08 percent Others' are in the public sector. This reveals that there are gaps between caste groups even within the public sector but a much higher proportion of SC-STs owes their access to white-collar jobs to the public sector. If there had been no quotas, the SC-ST access to white collar jobs would not have been as large as 6.51, which is already less than one-fourth the proportion of the Others'. The D-I-D for white collar public sector jobs reveals that OBCs are gaining vis-à-vis both SC-STs and Others', whereas SC-STs are losing vis-à-vis the Others'.

Thus, our suspicion that the lagging behind of the SC-STs in white collar jobs is a result of gaps in the private sector is further confirmed by this picture. Of course, our data do not allow us to identify quota beneficiaries explicitly; hence attributing the catch up to quotas is conjectural. The OBCs' access to white-collar jobs (both public and private), as well as public sector jobs (both blue and white-collar) shows convergence with Others'. A part of this convergence would be due to the operation of quotas but not all of it, since

134

there is convergence between OBCs and Others' in both public and private sectors.

3.4.3   Estimating Probabilities of Job Types

We ran multinomial probit regressions separately for each cohort to estimate the probability of being in one of the three job types (agricultural, blue-collar and white- collar) for the three caste groups. Table 3.7 presents the probabilities (marginal effects) with and without controls for region, sector, and years of education for each cohort for both rounds of NSS.

[Insert Table 3.7]

From the estimates for NSS-66, we see that SC-STs in Cohort 1 are 1.9 times less likely (without controls) and 12.8 times less likely (with controls) be in agricultural jobs compared to Others'. However, SC-STs in Cohorts 2-5 are more likely to be in agricultural jobs compared to Others' in corresponding cohorts. Similarly, OBCs are more likely to be in agricultural jobs compared to Others' in all cohorts (in regressions without controls), but controlling for others explanatory factors, are less likely to be in agricultural jobs.

OBCs, as well as SC-STs, are less likely to be in white-collar jobs compared to Others' in all cohorts, with and without controlling for other explanatory factors. However, Table 3.8 shows us that the marginal effects have by and large declined from the oldest to the youngest cohort, suggesting that the disadvantage of younger cohorts of OBCs relative to Others' appears to have decreased.

[Insert Table 3.8]

Comparing the marginal effects from a similar regression for NSS-55, we see that while OBCs were less likely than Others' to be in white-collar jobs also in 1999-2000, the marginal effects for the NSS-66 cohorts of OBCs are lower, again suggesting that the relative OBC disadvantage might have reduced over the decade between the two surveys. These regressions confirm the D-I-D trends in white-collar jobs for OBCs versus Others'.

135

3.4.4 Duncan's Dissimilarity Index

The NSS divides workers into a few broad categories based on their principal activity status.[12] Thus, this classification is distinct from the one used above, where we aggregated several occupations into three broad types. Using the principal activity status, we calculate the *Duncan Dissimilarity Index* between groups. The value of this index for any two groups (in our case, caste groups) gives the proportion of population that would have to change their activity status to make the distribution of the two groups identical.

Looking at the evolution of the index across cohorts, we find that in 1999-2000, SC-STs are the most dissimilar to the Others', with the dissimilarity *rising* from older to younger cohorts. Between OBCs and Others', Cohorts 3 and 4 are more dissimilar to the Others', as compared to the other three cohorts, and overall, all cohorts taken together, the OBCs are more similar to Others' than they are to SC-STs (See Figure 3.7).

[Insert Figure 3.7]

Data from 2009-10 (see Figure 3.8) reveals that the dissimilarity between SC-STs and Others' continues to look the same as a decade earlier. Between OBCs and Others', too, barring Cohorts 3 and 4, where dissimilarity between the two groups seems to have *increased*, the distribution is similar to what it was in 1999-2000. Again, barring Cohort 4, the OBC distribution is closer to Others' than it is to SC-STs.

[Insert Figure 3.8]

Understanding sources of dissimilarity

There are clear differences in the share of caste groups in the various principal status categories. Across all cohorts, SC-ST proportions in casual wage labour are the highest, followed by OBCs and then by Others'. Mirroring this feature, we find that SC- ST pro-

portions among employers are the lowest across all cohorts, followed by OBCs and then by Others'.

While each of these categories merits a separate analysis, in this paper we focus on two of the important sources of dissimilarity, viz., the proportion of all workers that are regular wage/ salaried (RWS) employees and those doing casual labour. Proportion in RWS jobs is a good indicator of involvement in the formal sector; these jobs are coveted also because of the benefits they confer to the worker, which are typically missing from informal sector or casual jobs (some possible benefits could be inflation-linked indexation, pensions, gratuity, illness cover, group insurance, provident fund and so forth). As Banerjee and Duflo (2011) suggest, job security and regular wages seems to be one of the important aspirations of the poor in India. Thus, the small proportions of SC- STs and OBCs in RWS jobs suggests that this is an important facet of occupational disparity across caste groups.

We see that across all groups, the proportions engaged in RWS jobs have been rising, indicating the greater formalization of jobs. As Figure 3.9 shows, for the Others', there is sharp rise in the proportion in RWS jobs from Cohort 1 to Cohort 4, but the rise is not sustained in the next two cohorts. OBCs and SC-STs too show a much sharper rise from Cohort 1 to Cohort 4, than for the latter two cohorts.

[Insert Figure 3.9]

What is interesting is that the D-I-D in the share of salaried employees across cohorts between groups shows slightly different patterns between NSS-55 and NSS-66. In NSS 55 Cohort 4 and 5 of the OBCs and SC-ST gain relative to the Others'. In NSS 66 only Cohort 5 of the OBCs and SC-ST gain relative to the Others'.[13] Given that NSS-66 is the later survey, we can take the results from this survey as indicating the latest trends. The share of RWS employees by cohort and their evolution of the D-I-D for NSS-66 are shown

in Figure 3.10 and 3.11, respectively.

[Insert Figure 3.10]

Thus, between Cohort 2 and Cohort 1, OBCs fall 9.71 percentage points behind the Others'. This gap consistently increases and finally between Cohort 5 and 4, OBCs gain 3.28 percentage points relative to Others'. The SC-ST versus Others' D-I-D shows the same trend, except that the final cohort gains only 0.62 percentage points relative to the Others'. Over the entire sample period we see that for the OBCs the gap increases from -0.97 percentage points for Cohort 1 to 8.9 percentage points for cohort 5 (born 1966-75). Similarly for the SC-ST the gap increases from 1.5 percentage points for cohort 1 to 14 percentage points for the cohort born in 1966-75. So over the 50 year period there seems to have been divergence in terms of share of RWS between the Others' and OBCs and SC-ST.

[Insert Figure 3.11]

Given the divergence except for the very youngest cohorts in the activity status of RWS, looking at NSS-66 we explore whether the trends in casual labour mirror those of RWS i.e. whether Others' have decreased their share of labour force in casual labour relative to the SC-ST and OBCs.

[Insert Figure 3.12]

From NSS-66 (Figure 3.12) we see that SC-STs not only have the highest proportions in casual labour, this proportion has gone up from 37.66 for Cohort 1 to 50.82 for Cohort 5. The corresponding proportions are 19.74 to 29.94 for OBCs and 8.51 and 18.61 for Others'. Comparing D-I-D across cohorts (Figure 3.13), we see that overall, OBCs' movement across cohorts is not very different from that of Others' (D-I-D between Cohort 5 and 1 is 0.1).

Between Cohorts 4 and 3, the increase in OBC proportion in casual labour is higher than that of Others', but between Cohort 5 and 4, the increase in proportion for Others' is higher than that for OBCs, and for the other cohorts, the increase in OBC proportions is marginally higher, so the net result, comparing OBCs and Others', is that casualisation of labour is proceeding at a similar rate. But between SC-STs and Others', the trend is exactly the opposite, in that SC-ST labour is getting into casual jobs in higher proportions across successive cohorts compared to the Others'. Comparing OBCs and SC-STs, again the rate of casualisation for SC-STs is significantly higher than that for OBCs. Thus, the activity status profiles of the three groups continue to look dissimilar for the three groups, with OBCs closer to the Others' than to SC-STs.

[Insert Figure 3.13]

To sum up the picture seems to suggest that the Others' have increased the proportion of their RWS jobs as compared to the OBCs and SC-ST (except for the youngest cohort). The trend in casualisation of labour is very similar for Others' and OBCs over the period whereas the amount of work force employed as casual labour has increased for the SC-ST relative to the Others'. The two strands of evidence suggest that there has been divergence in the principal activity status between the Others' and the OBCs and SC-ST, with the Others' especially increasing their share of the coveted RWS jobs.

3.5   Wages and labour market discrimination

The average wages for the three caste groups show the expected ranking. In 2009- 10, the average wages were Rs. 660, 848 and 1286 respectively for SC-STs, OBCs and Others' respectively. Interestingly, for OBCs and Others', average wages for Cohort 4 were the highest, as is expected given that this cohort is between 54 and 45 years old, in other

139

words, is at the peak of the earning cycle. However, for SC-STs, average wages for Cohort 3 are higher than for Cohort 4, as can be seen in the Figure 3.14.

[Insert Figure 3.14]

The D-I-D analysis of wages shows that while the gap between OBCs and Others' increases between Cohort 3 and 1 by Rs. 1075, OBCs average wages catch up by Rs. 87 between Cohort 4 and Cohort 3, by Rs. 301 between Cohort 5 and 4 and by Rs. 285 between Cohort 6 and 5. However, in the overall gap (measured as the gap between Cohort 6 and 1), OBCs fall behind the Others' by Rs. 500, but it is clear that younger cohorts of OBCs are catching up with the Others' in terms of average wages. Overall, SC-STs remain further behind the Others' as compared to OBCs (the overall gap between Cohort 5 and Cohort 1 increasing by being Rs. 889), but the two youngest cohorts appear to catch up with the Others' (the evolution and statistical significance of the calculated D-I-D are shown in Table 3.14 in the appendix).

[Insert Figure 3.15]

The kernel density plots for two cohorts of SC-STs (aged 55-64 and aged 35-44) shows a rightward shift in the distribution, confirming that the younger SC-ST cohort is doing better in terms of wages (Figure 3.16). Similar plots for OBCs and Others' (Figures 3.16) do not show this clear rightward shift – the OBC distribution for the younger cohort is flatter and smoother; the Others' distribution retains two peaks but becomes smoother for the younger Cohort.

[Insert Figure 3.16]

140

3.5.1 Blinder-Oaxaca Decomposition

We conduct the Blinder-Oaxaca (B-O) decomposition on the average male wage gap between OBCs and Others' in order to separate the explained from the unexplained component, the basic methodology for which is explained in the Appendix. Based on NSS-66, the results of the B-O decomposition exercise between OBCs and Others' (for males in the labour force) can be seen in Table 3.9.

[Insert Table 3.9]

We see that in regressions which include personal characteristics as controls (years of education, age, age squared, married), for all cohorts between 25 to 74 years, we see that the (geometric) means of wages are Rs. 1254 for Others' and Rs. 830 for Others', amounting to a difference of 51 percent. Adjusting OBC endowment levels to Others' would increase OBC wages by 28.4 percent, but a gap of 17.6 percent remains unexplained. Adding controls for region and sector (rural-urban), the wage difference between OBCs and Others' reduces slightly to 49 percent, with endowment difference now accounting for 30.6 percent and the unexplained component now reduced to 14.4 percent. Adding controls for occupation, the unexplained component further reduces to 9.8 percent. However, whether it is appropriate to add occupational controls is a moot point.

Running similar regressions for each of the cohorts separately, we see from Table 3.9, that the unexplained component is 14.5 percent for the cohort aged 55-64 with personal characteristics as controls, which reduces to 12.1 percent with personal characteristics combined with region and sector and to 3.2 percent with occupation controls included as well. For the cohort aged 45-54, we see that the unexplained component is higher (24 percent) with personal characteristics; 21 with additional sector and region controls and 10.9 percent with further addition of occupation controls. In others words, for all three specifications, the unexplained component of the wage gap for this cohort is higher than for

141

the previous cohort. For the cohort aged 35-44, the respective unexplained proportions are 15, 9.2, and 8.7 – i.e. smaller than for the previous cohort. This reversal or improvement compared to the previous cohort is in line with the evidence from the D- I-D analysis of wage changes across cohorts reported above. The unexplained proportions do not change from this cohort to the next youngest (aged 25-34 years).

Comparing these estimates with the Blinder-Oaxaca decomposition conducted between SC-ST and Others' (Table 3.10) reveals that first, the wage gap for all cohorts considered together is nearly 92 percent (average wage for SC-ST being Rs. 653.86).

[Insert Table 3.10]

Thus, the average wage gap between SC-ST and Others' is a little less than twice the wage gap between OBCs and Others'. Correspondingly, the unexplained portion is 29.8 percent with personal characteristics as controls; with region and sector controls, this reduces to 20 percent, and further to 14.25 percent with controls for occupation included. Thus, all estimates indicate that labour market discrimination against SC-ST is significantly greater than against OBCs, when the wages of these groups are compared to the Others'.

3.6   Conclusion

The findings suggest that the gap between the Others' and OBCs and SC-ST remain large for a variety of important indicators. MPCE and wages of the OBCs and SC-ST are 51 and 65 percent and 42 and 55 percent, respectively, of the average of the Others'. Their shares of labour force employed in white collar prestigious jobs is about one-fourth and one-half the proportion of the Others' employed in white collar jobs. On the other hand their share of labour force employed as casual labour is twice and thrice higher than the Others' for the OBCs and SC-ST, respectively. However, despite significant gaps in the

142

above indicators, we find substantial evidence of catch- up between OBCs and Others' for the younger cohorts (especially in literacy, primary education, access to white-collar jobs, wages), but we find continued divergence in all education categories after the middle school level. This picture is different from the one that emerges after a similar analysis between SC-STs and Others', where the divergence and dissimilarity in all indicators vis-a-vis the Others' is much greater. The only exception is in the education transition matrix: we find that sons of graduate fathers are more likely to be graduates for SC-STs than for OBCs. This could possibly be the result of the longer history of educational quotas for SC-STs in institutes of higher education as compared to that for OBCs. Younger cohorts of OBCs are closer to the Others' than to SC-STs in all indicators, whereas the older cohorts were closer to the SC-STs in several key indicators. What precise factors have contributed to the OBC catch-up needs to be investigated, and we hope to be able to address this in our on going research.

Bibliography

[1] Banerji, A. and Duflo, E. (2011), *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs, New York, USA.

[2] Deshpande, A. (2007), "Overlapping Identities under Liberalisation: Gender and Caste in India," *Economic Development and Cultural Change*, 55(4): 735-760.

[3] Deshpande, A. (2011), *The Grammar of Caste: Economic Discrimination in Contemporary India*, Oxford University Press, New Delhi.

[4] Gupta, D. (2004), (editor), "Caste in Question: Identity or Hierarchy?" *Contributions to Indian Sociology: Occasional Studies 12*, Sage Publications, New Delhi.

[5] Hnatkovska, V. and Lahiri, A. and Paul, S. (2012), "Castes and Labor Mobility," *American Economic Journal: Applied Economics*, 4(2): 274-307

[6] Lakshmi, I. and Khanna,T. and Varshney, A. (2013), "Caste and Entrepreneurship in India," *Economic and Political Weekly*, XLVIII(6): 52-60

[7] Jafferlot, C. (2003), *India's Silent Revolution: The Rise of Low Castes in North Indian Politics*, Permanent Black, New Delhi.

[8] Jann, B. (2008), "The Blinder-Oaxaca decomposition for linear regression models," *The Stata Journal* 8(4): 453-479.

[9] Vakulabharanam, V. and Zacharias, A. (2011), "Caste Stratification and Wealth Inequality in India," *World Development*, 39(10): 1820-1833.

Notes

[1] 1 acre=0.4047 hectares. Land possessed is defined as land (owned+leased-in+neither owned nor leased-in)- land leased out.

[2] The NSS does not have information on years of education. We use the method followed in Hnatkovska et al. (2012) for converting information on educational attainment to years of education. Thus, those with formal schooling were assigned 0 years of education; those with schooling below primary were assigned 2 years; those with primary completed 5 years; those with middle school completed 7 years; those with secondary completed 10 years; those with higher secondary 12 years; those with graduate degrees in technology, engineering, medicine and agriculture 16 years and those with graduate degrees in all other subjects were assigned 15 years.

[3] The detailed tables and charts for all the educational categories are available with the authors upon request. In the interest of space, we are only presenting the figures on years of education and for the educational category "graduate and above".

[4] If we consider the Cohort aged 15-24, i.e. those who should have achieved literacy by the time the survey was done, the gaps further reduce, and the Others' have a lead of 7 percent and 13 percent over the OBCs and SC-ST, respectively.

[5] If we consider the Cohort aged 15-24, i.e. those who should have finished primary schooling by the time the survey was done, the gaps further reduce, and the Others' have a lead of 9 percent and 16 percent over the OBCs and SC-STs, respectively.

[6] Here even if we compare the oldest cohort who went to school after independence (cohort 3), with the youngest cohort who would have finished schooling by 2010 (cohort 6) makes the D-I-D for the OBCs compared to the Others' marginally positive (1 percent) but insignificant, whereas for the SC-STs and Others', it remains negative and significant (gap of 5 percent).

[7] Comparing the oldest cohort that went to school after independence (cohort 3) with the youngest cohort that would have finished schooling by 2010 (cohort 6), the D-I-D for the OBCs and SC-STs compared to the Others' remains negative and significant.

[8] In the NSS EUS, these are all individuals with principal activity status codes between 11 and 81.

[9] We use NCO-68 codes for this classification. Following Hnatkovska et al. (2012), all those with NCO codes between 600 and 699 are classified as being in agricultural jobs; those between 400 and 599 or between 700 and 999 as being in blue-collar jobs; and those between 0 and 399 are classified as having white-collar jobs.

[10]When we trace the evolution of occupations, we focus on Cohort 1 of NSS-55 (the oldest cohort) and compare that with Cohort 4 of NSS-66, which is the second youngest cohort in our data set. The youngest group is cohort 5 in NSS- 66, but these are individuals between 25-34 years of age and might be still be in a state of transition in terms of their occupational choices. Those aged 35-44 years would be more likely settled in their choices.

[11]The appendix shows the table showing the distribution of the labour force across the 3 occupations for the 3 social groups.

[12] The principal activity status has the following categories: own-account worker, employer, helper in household enterprise, regular wage/ salaried employment; casual wage labour in public works; casual wage labour in other types of work.

[13]The graphs for the NSS 55th are are available on request from the authors.

Table 3.1: **Household level indicators: All India**

Indicator	SC-ST	OBCs	Others'
MPCE 55th Round	454.66	534.57	747.50
	(1.639)	(1.969)	(2.903)
MPCE 66th Round	956.68	1209.38	1850
	(5.759)	(5.989)	(11.012)
D-I-D MPCE	172.79***	-427.69***	-600.48***
% Urban 55th Round	17.10	23.96	38.8
	(0.000011)	(0.000011)	(0.00001)
% Urban 66th Round	17.65	27.85	42.9
	(0.000017)	(0.000018)	(0.000025)
D-I-D % Urban	3.34***	-0.21***	-3.55***
Household size 55th Round	4.77	4.94	4.87
	(0.013)	(0.013)	(0.012)
Household size 66th Round	4.45	4.47	4.31
	(0.012)	(0.012)	(0.012)
D-I-D Household size	-0.15***	0.10***	0.24***
Land owned 55th Round	0.44	0.64	0.74
	(0.006)	(0.008)	(0.009)
Land owned size 66th Round	0.43	0.62	0.70
	(0.006)	(0.008)	(0.01)
D-I-D Land owned	0.00	0.02	0.02
Land possessed 55th Round	0.46	0.65	0.73
	(0.006)	(0.008)	(0.009)
Land possessed size 66th Round	0.40	0.57	0.61
	(0.006)	(0.007)	(0.009)
D-I-D Land possessed	-0.02*	-0.08	0.06

a. The standard errors are shown in parenthesis.

b. Note the D-I-D corresponding to the column SC-ST refers to the one calculated comparing OBCs to the SC-ST, the D-I-D in column OBCs compares OBCs to Others' and the D-I-D in column Others' compares SC-ST to Others'.

c. A negative D-I-D in column SC-ST and OBCs implies OBCs are relatively losing ground relatively, a negative D-I-D in column Others' implies SC-ST are relatively losing ground.

d. Land owned and land possessed are in 1000's of hectares.

e. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

Table 3.2: **Birth year of cohorts used from NSS-55 and NSS-66 in the sample**

	Age	Birth year round 55th	Birth year round 66th
Cohort 1	65-74	1926-1935	1936-1945
Cohort 2	55-64	1936-1945	1946-1955
Cohort 3	45-54	1946-1955	1956-1965
Cohort 4	35-44	1956-1965	1966-1975
Cohort 5	25-34	1966-1975	1976-1985

Note: Cohort 1 of NSS round 66th has the same birth years as Cohort 2 of NSS 55th, Cohort 2 of round 66th as Cohort 3 of NSS 66th, Cohort 3 of NSS 66th as Cohort 4 of NSS 66th and finally cohort 4 of NSS 66th as Cohort 5 of NSS 55th. We often combine the 1st cohort of the NSS 55th with the 5 cohorts of NSS 66th. This implies our sample covers the birth years 1926-1985 or sample period of 60 birth years.

Table 3.3: **Educational Transition Matrix, All India - NSS-55**

Transition Matrix for the SC-ST							
	Edu 0	**Edu 1**	**Edu 2**	**Edu 3**	**Edu 4**	**Edu 5**	**Size**
Edu 0	0.41	0.12	0.32	0.09	0.05	0.02	59.66
Edu 1	0.13	0.17	0.44	0.15	0.08	0.03	14.22
Edu 2	0.07	0.06	0.49	0.20	0.11	0.06	17.45
Edu 3	0.03	0.01	0.22	0.29	0.32	0.13	5.11
Edu 4	0.03	0.02	0.19	0.26	0.32	0.19	1.83
Edu 5	0.01	0.00	0.17	0.22	0.33	0.26	1.73
Transition Matrix for the OBCs							
	Edu 0	**Edu 1**	**Edu 2**	**Edu 3**	**Edu 4**	**Edu 5**	**Size**
Edu 0	0.36	0.12	0.34	0.11	0.06	0.02	46.44
Edu 1	0.10	0.12	0.49	0.16	0.09	0.04	17.97
Edu 2	0.06	0.04	0.46	0.23	0.14	0.07	23.98
Edu 3	0.02	0.02	0.23	0.33	0.23	0.17	7.10
Edu 4	0.01	0.02	0.23	0.22	0.28	0.25	2.58
Edu 5	0.00	0.02	0.09	0.18	0.35	0.36	1.94
Transition Matrix for the Others							
	Edu 0	**Edu 1**	**Edu 2**	**Edu 3**	**Edu 4**	**Edu 5**	**Size**
Edu 0	0.27	0.12	0.38	0.14	0.06	0.03	26.50
Edu 1	0.07	0.14	0.42	0.20	0.10	0.08	15.24
Edu 2	0.04	0.03	0.41	0.26	0.15	0.11	27.87
Edu 3	0.01	0.01	0.15	0.28	0.28	0.26	14.95
Edu 4	0.02	0.00	0.10	0.21	0.34	0.33	5.90
Edu 5	0.02	0.00	0.04	0.13	0.32	0.49	9.55

Notes: Each cell ij represents the average probability (for a given NSS survey round) of a household male head with education i having a son with education attainment level j. Column titled "size" reports the fraction of fathers in education category 0, 1, 2, 3, 4, or 5 in a given survey round.

Table 3.4: **Educational Transition Matrix, All India - NSS-66**

Transition Matrix for the SC-ST

	Edu 0	Edu 1	Edu 2	Edu 3	Edu 4	Edu 5	Size
Edu 0	0.23	0.09	0.42	0.13	0.10	0.03	50.12
Edu 1	0.04	0.10	0.55	0.16	0.10	0.05	14.08
Edu 2	0.03	0.03	0.45	0.24	0.20	0.06	22.85
Edu 3	0.01	0.00	0.23	0.27	0.37	0.12	6.38
Edu 4	0.00	0.01	0.12	0.27	0.32	0.27	3.33
Edu 5	0.00	0.11	0.06	0.09	0.36	0.38	3.24

Transition Matrix for the OBCs

	Edu 0	Edu 1	Edu 2	Edu 3	Edu 4	Edu 5	Size
Edu 0	0.19	0.12	0.38	0.16	0.11	0.04	35.66
Edu 1	0.04	0.11	0.43	0.21	0.17	0.04	13.53
Edu 2	0.03	0.02	0.38	0.25	0.22	0.10	29.87
Edu 3	0.01	0.01	0.13	0.28	0.36	0.21	10.57
Edu 4	0.02	0.00	0.13	0.15	0.42	0.28	6.16
Edu 5	0.00	0.00	0.07	0.13	0.47	0.34	4.21

Transition Matrix for the Others

	Edu 0	Edu 1	Edu 2	Edu 3	Edu 4	Edu 5	Size
Edu 0	0.15	0.10	0.40	0.19	0.12	0.05	23.20
Edu 1	0.02	0.08	0.45	0.20	0.16	0.08	9.89
Edu 2	0.02	0.02	0.32	0.28	0.24	0.12	29.80
Edu 3	0.01	0.00	0.11	0.26	0.35	0.27	16.26
Edu 4	0.01	0.00	0.08	0.11	0.45	0.35	8.81
Edu 5	0.01	0.00	0.02	0.08	0.36	0.54	12.04

Notes: Each cell ij represents the average probability (for a given NSS survey round) of a household male head with education i having a son with education attainment level j. Column titled "size" reports the fraction of fathers in education category 0, 1, 2, 3, 4, or 5 in a given survey round.

Table 3.5: Marginal effect of SC-ST and OBC dummy in ordered probit regression for education categories

		ALL COHORTS	COHORT 1	COHORT 2	COHORT 3	COHORT 4	COHORT 5	COHORT 6	Cohort 2 to 1	Cohort 3 to 2	Cohort 4 to 3	Cohort 5 to 4	Cohort 6 to 5	Cohort 5 to 1
Edu 1	SC-ST	0.307***	0.324***	0.347***	0.366***	0.314***	0.268***	0.155***	0.023	0.019	-0.052	-0.046	-0.113	-0.056
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	0.191***	0.208***	0.231***	0.229***	0.196***	0.147***	0.072***	0.023	-0.002	-0.033	-0.049	-0.075	-0.061
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Edu 2	SC-ST	-0.001***	-0.061***	-0.038***	-0.020***	0.005***	0.027***	0.030***	0.023	0.018	0.025	0.022	0.003	0.088
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	0.006***	-0.030***	-0.016***	-0.003***	0.011***	0.021***	0.016***	0.014	0.013	0.014	0.01	-0.005	0.051
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Edu 3	SC-ST	-0.030***	-0.076***	-0.065***	-0.055***	-0.025***	0.004***	0.020***	0.011	0.01	0.03	0.029	0.016	0.08
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	-0.011***	-0.044***	-0.037***	-0.027***	-0.008***	0.011***	0.013***	0.007	0.01	0.019	0.019	0.002	0.055
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Edu 4	SC-ST	-0.067***	-0.062***	-0.074***	-0.082***	-0.071***	-0.051***	-0.044***	-0.012	-0.008	0.011	0.02	0.007	0.011
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	-0.038***	-0.041***	-0.048***	-0.049***	-0.040***	-0.023***	-0.018***	-0.007	-0.001	0.009	0.017	0.005	0.018
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Edu 5	SC-ST	-0.209***	-0.124***	-0.171***	-0.209***	-0.222***	-0.248***	-0.162***	-0.047	-0.038	-0.013	-0.026	0.086	-0.124
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	-0.147***	-0.093***	-0.130***	-0.151***	-0.159***	-0.156***	-0.083***	-0.037	-0.021	-0.008	0.003	0.073	-0.063
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Note: Panel (a) reports the mraginal effects of the SC-ST and OBC dummy in an opdered probit regression of education categories 1 to 5 on a constant and an SC-ST and OBC dummy for each cohort. Panel (b) of the table reports the change in the marginal effects over successive cohorts and over the entire sample period. Standard errors are in parenthesis. * p-value0.10, ** p-value0.05, *** p-value0.01.

Table 3.6: **Evolution on public sector jobs by cohorts**

Social Group	COHORT 1 (1)	COHORT 2 (2)	COHORT 3 (3)	COHORT 4 (4)	COHORT 5 (5)
Share of public sector jobs by cohorts					
SC-ST	2.91	8.02	9.56	7.66	4.76
OBC	0.63	5.69	8.77	5.67	3.85
OTHERS	0.29	10.54	15.07	9.37	5.44
Share of public sector jobs in blue collar jobs by cohorts					
SC-ST	9.01	18.05	18.65	12.86	6.98
OBC	1.11	11.89	14.8	8.06	5.76
OTHERS	0.25	18.31	23.43	12.85	6.9
Share of public sector jobs in white collar jobs by cohorts					
SC-ST	2.29	39.88	35.96	26.35	16.48
OBC	1.58	17.03	21.2	15.97	9.2
OTHERS	1.3	22.15	24.08	15.73	9.02

Note: Cohort 1-5 are from NSS-66.

Table 3.7: Unconditional marginal effect of SC–ST and OBC dummy in ordered probit regression for occupational categories

NSS 55th marginal effects unconditional

		COHORT 1 1	COHORT 1 2	COHORT 2 3	COHORT 3 4	COHORT 3 5	COHORT 3 6	COHORT 4 7	COHORT 4 8	COHORT 5 9	COHORT 5 10
Agricultural Jobs	SC-ST	0.061***	-0.020***	0.158***	0.017***	0.217***	0.008***	0.225***	0.029***	0.206***	0.050***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	0.020***	-0.026***	0.095***	-0.023***	0.133***	-0.027***	0.135***	0.009***	0.112***	0.015***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Blue Collar Jobs	SC-ST	-0.016***	0.031***	-0.048***	-0.001	-0.069***	-0.002***	-0.106***	-0.031***	-0.103***	-0.026***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	0.024***	0.049***	-0.002***	0.034***	-0.006***	0.049***	-0.029***	0.018***	-0.029***	0.012***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
White Collar Jobs	SC-ST	-0.045***	-0.012***	-0.110***	-0.016***	-0.147***	-0.005***	-0.119***	0.002***	-0.103***	-0.024***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	-0.044***	-0.023***	-0.092***	-0.011***	-0.127***	-0.022***	-0.105***	-0.027***	-0.082***	-0.027***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

NSS 66th marginal effects unconditional

		COHORT 1 1	COHORT 1 2	COHORT 2 3	COHORT 3 4	COHORT 3 5	COHORT 3 6	COHORT 4 7	COHORT 4 8	COHORT 5 9	COHORT 5 10
Agricultural Jobs	SC-ST	-0.019***	-0.128***	0.147***	-0.028***	0.175***	-0.015***	0.147***	-0.041***	0.176***	0.024***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	0.001***	-0.059***	0.087***	-0.028***	0.106***	-0.006***	0.098***	-0.015***	0.087***	0.016***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Blue Collar Jobs	SC-ST	0.106***	0.159***	-0.017***	0.067***	-0.014***	0.075***	-0.011***	0.071***	-0.021***	0.049***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	0.061***	0.079***	0.020***	0.058***	-0.002***	0.036***	0.002***	0.037***	-0.005***	0.014***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
White Collar Jobs	SC-ST	-0.087***	-0.031***	-0.131***	-0.038***	-0.161***	-0.061***	-0.136***	-0.030***	-0.155***	-0.073***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	OBCs	-0.063***	-0.020***	-0.107***	-0.030***	-0.103***	-0.029***	-0.100***	-0.023***	-0.082***	-0.030***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

a. Columns (1), (3), (5), (7) and (9) reports the marginal effects on a constant and an SC-ST and OBC dummy for each cohort.

b. columns (2), (4), (6), (8) and (10) reports the marginal effects on a SC-ST and OBC dummy controlling for regional dummies, sector and years of education for each cohort.

c. Standard errors are in parenthesis. * p-value0.10, ** p-value0.05, *** p-value0.01.

Table 3.8: **Changes in unconditional marginal effect of SC-ST and OBC dummy in ordered probit regression for occupational categories**

		Cohort 2 to 1 1	Cohort 2 to 1 2	Cohort 3 to 2 3	Cohort 3 to 2 4	Cohort 4 to 3 5	Cohort 4 to 3 6	Cohort 5 to 4 7	Cohort 5 to 4 8	Cohort 5 to 1 9	Cohort 5 to 1 10
NSS 55th changes in marginal effects											
Agricultural Jobs	SC-ST	0.097	0.037	0.059	-0.009	0.008	0.021	-0.019	0.021	0.145	0.07
	OBCs	0.075	0.003	0.038	-0.004	0.002	0.036	-0.023	0.006	0.092	0.041
Blue Collar Jobs	SC-ST	-0.032	-0.032	-0.021	-0.001	-0.037	-0.029	0.003	0.005	-0.087	-0.057
	OBC	-0.026	-0.015	-0.004	0.015	-0.023	-0.031	0	-0.006	-0.053	-0.037
White Collar Jobs	SC-ST	-0.065	-0.004	-0.037	0.011	0.028	0.007	0.016	-0.026	-0.058	-0.012
	OBC	-0.048	0.012	-0.035	-0.011	0.022	-0.005	0.023	0	-0.038	-0.004
NSS 66th changes in marginal effects											
Agricultural Jobs	SC-ST	0.166	0.1	0.028	0.013	-0.028	-0.026	0.029	0.065	-0.176	-0.024
	OBCs	0.086	0.031	0.019	0.022	-0.008	-0.009	-0.011	0.031	-0.087	-0.016
Blue Collar Jobs	SC-ST	-0.123	-0.092	0.003	0.008	0.003	-0.004	-0.01	-0.022	0.021	-0.049
	OBCs	-0.041	-0.021	-0.022	-0.022	0.004	0.001	-0.007	-0.023	0.005	-0.014
White Collar Jobs	SC-ST	-0.044	-0.007	-0.03	-0.023	0.025	0.031	-0.019	-0.043	0.155	0.073
	OBCs	-0.044	-0.01	0.004	0.001	0.003	0.006	0.018	-0.007	0.082	0.03

a. Columns (1), (3), (5), (7) and (9) reports the change in marginal effects on a constant and an SC-ST and OBC dummy for each cohort.
b. columns (2), (4), (6), (8) and (10) reports the change in marginal effects on a SC-ST and OBC dummy controlling for regional dummies, sector and years of education for each cohort.
c. Note all changes are significant at the 1% level..

Table 3.9: **Blinder-Oaxaca Decomposition: Others versus OBCs: 2009-10**

	Mean wage: Others'	Mean wage: OBCs	Gap	Explained	Unexplained	N
All Cohorts: Controls - personal characteristics(PC)	1254.204	830.4833	51.02	28.4	17.61	29919
Controls: PC, Region and Sector	1251.429	837.2537	49.46	30.62	14.42	28033
Controls: PC, region, sector, occupation	1246.638	836.9334	48.95	34.94	10.38	28033
Cohort aged 55-64: Controls - PC	1422.958	810.3948	75.58	53.35	14.49	2820
Controls: PC, Region and Sector	1402.582	799.2925	75.47	56.57	12.07	2638
Controls: PC, region, sector, occupation	1341.673	812.1725	65.19	60.1	3.17	2638
Cohort aged 45-54: Controls - PC	1527.084	868.8451	75.76	40.77	24.85	7115
Controls: PC, Region and Sector	1510.842	872.5434	73.15	42.64	21.39	6664
Controls: PC, region, sector, occupation	1454.542	894	62.7	46.67	10.92	6664
Cohort aged 35-44: Controls - PC	1273.638	837.951	51.99	32.11	15.04	9568
Controls: PC, Region and Sector	1268.662	848.7306	49.47	36.9	9.18	8978
Controls: PC, region, sector, occupation	1309.48	859.2448	52.399	40.18	8.71	8978
Cohort aged 25-34: Controls - PC	1273.638	837.951	51.99	32.11	15.04	9568
Controls: PC, Region and Sector	1268.662	848.7306	49.47	36.9	9.18	8978
Controls: PC, region, sector, occupation	1309.48	859.244	52.39	40.18	8.71	8978

a. Personal characteristics controlled for are years of education and marital status

Table 3.10: **Blinder-Oaxaca Decomposition: Others versus SC-ST: 2009-10**

	Mean wage: Others'	Mean wage: SC-ST	Gap	Explained	Unexplained	N
All Cohorts: Controls - personal characteristics(PC)	1254.204	653.8629	91.81	47.69	29.86	29374
Controls: PC, Region and Sector	1251.429	657.9229	90.2	58.41	20.06	27321
Controls: PC, region, sector, occupation	1246.638	657.5139	89.59	65.94	14.25	27321
Cohort aged 55-64: Controls - PC	1422.958	634.8243	124.14	84.325	21.6	2667
Controls: PC, Region and Sector	1402.582	644.172	117.73	86.55	16.71	2490
Controls: PC, region, sector, occupation	1341.673	642.8613	108.7	91.01	9.25	2490
Cohort aged 45-54: Controls - PC	1527.084	687.7152	122.05	76.26	25.97	7007
Controls: PC, Region and Sector	1510.842	690.5461	118.78	83.1	19.48	6521
Controls: PC, region, sector, occupation	1454.542	716.7164	102.94	94.24	4.48	6521
Cohort aged 35-44: Controls - PC	1273.638	693.2092	83.73	44.75	26.92	9357
Controls: PC, Region and Sector	1268.662	688.2011	84.344	58.34	16.41	8660
Controls: PC, region, sector, occupation	1309.48	663.3888	97.39	62.67	21.33	8660
Cohort aged 25-34: Controls - PC	1273.638	693.2092	83.73	44.75	26.92	9357
Controls: PC, Region and Sector	1268.662	688.2011	84.34	58.34	16.41	8660
Controls: PC, region, sector, occupation	1309.48	663.388	97.39	62.67	21.33	8660

a. Personal characteristics controlled for are years of education and marital status
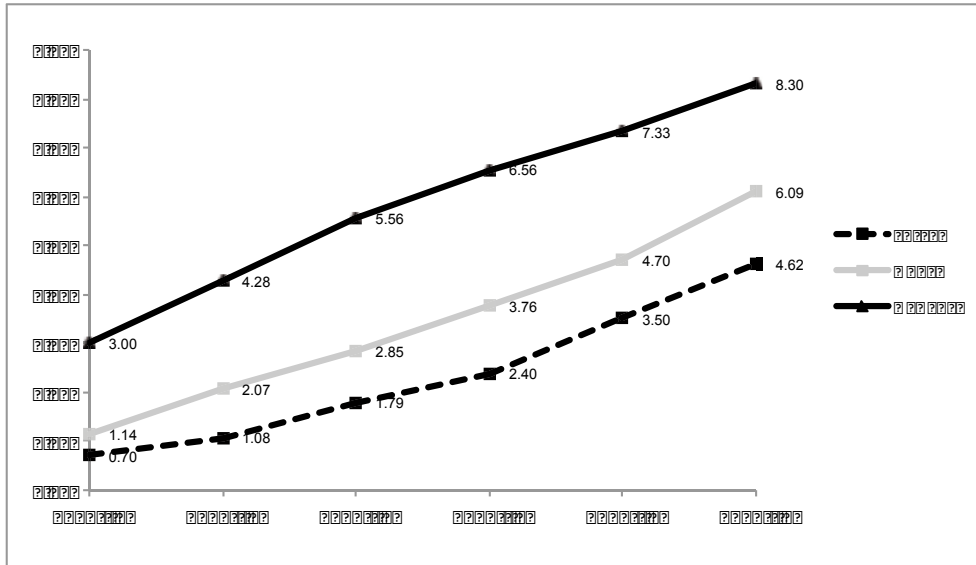
Figure 3.1: **Years of education across cohorts**
Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.
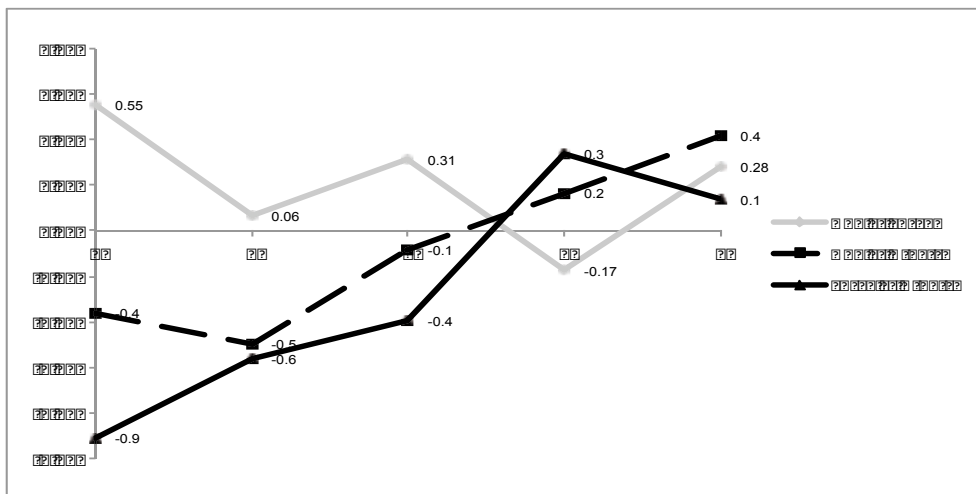

Figure 3.2: **Evolution of D-I-D for years of education across consecutive cohorts**
Note: A negative D-I-D for the line comparing OBCs to Others or the OBCs to the SC-ST implies OBCs relatively losing ground whereas a positive value implies convergence in case of comparison with the Others and divergence when compared to SC-ST. A negative value D-I-D line comparing Others to SC-ST implies the SC-ST are relatively losing ground when compared to the Others and positive value implies convergence.

156
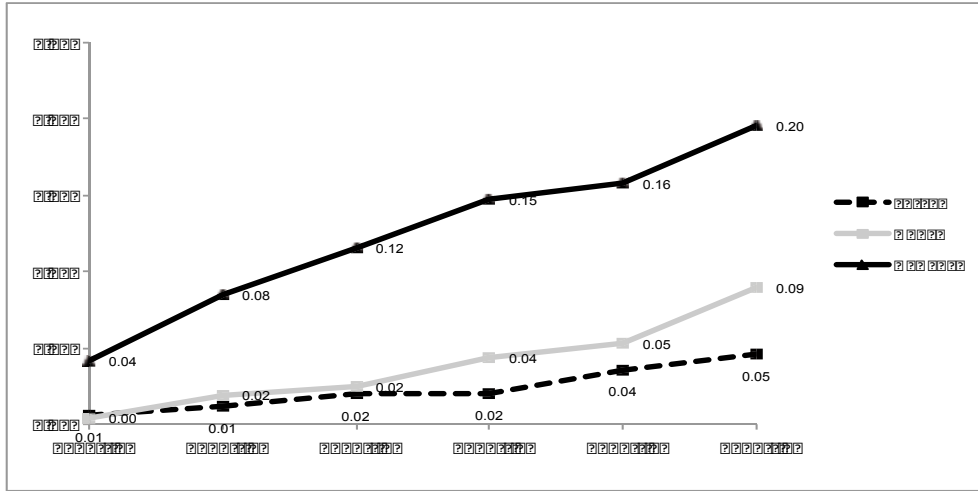
Figure 3.3: **Proportions of different cohorts that have a graduate degree or more of education**
Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.
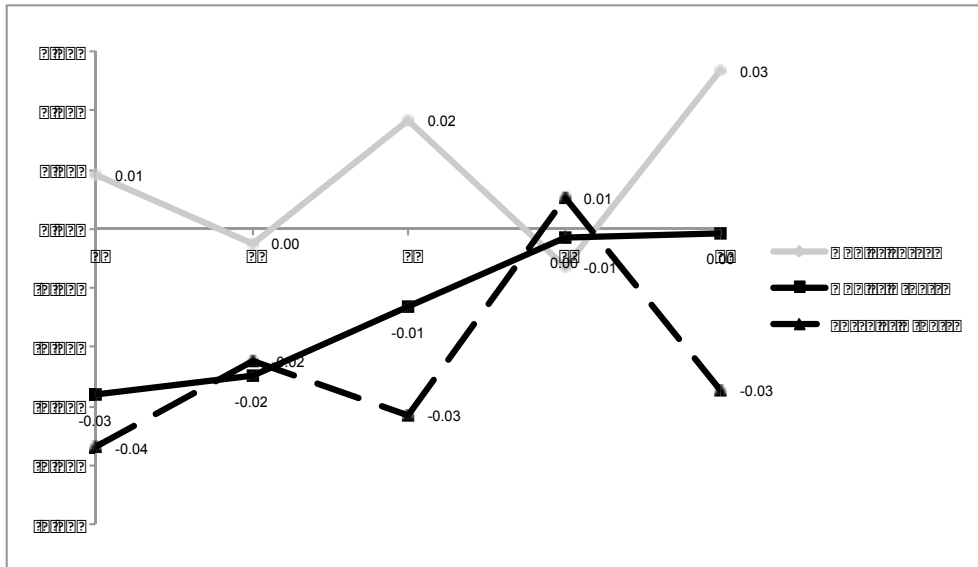

Figure 3.4: **Evolution of D-I-D for graduates and more across consecutive cohorts**
Note: A negative D-I-D for the line comparing OBCs to Others or the OBCs to the SC-ST implies OBCs relatively losing ground whereas a positive value implies convergence in case of comparison with the Others and divergence when compared to SC-ST. A negative value D-I-D line comparing Others to SC-ST implies the SC-ST are relatively losing ground when compared to the Others and positive value implies convergence.
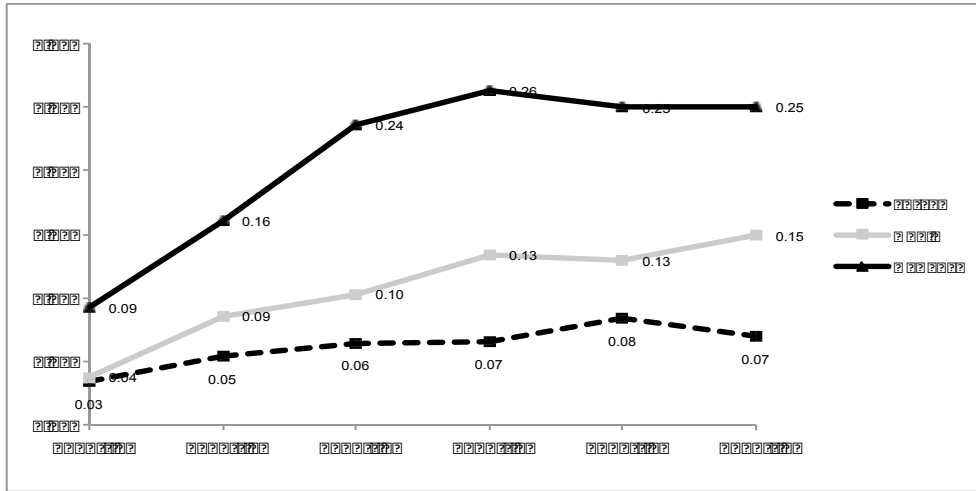
157

Figure 3.5: **Proportion in white-collar jobs across cohorts**
Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.
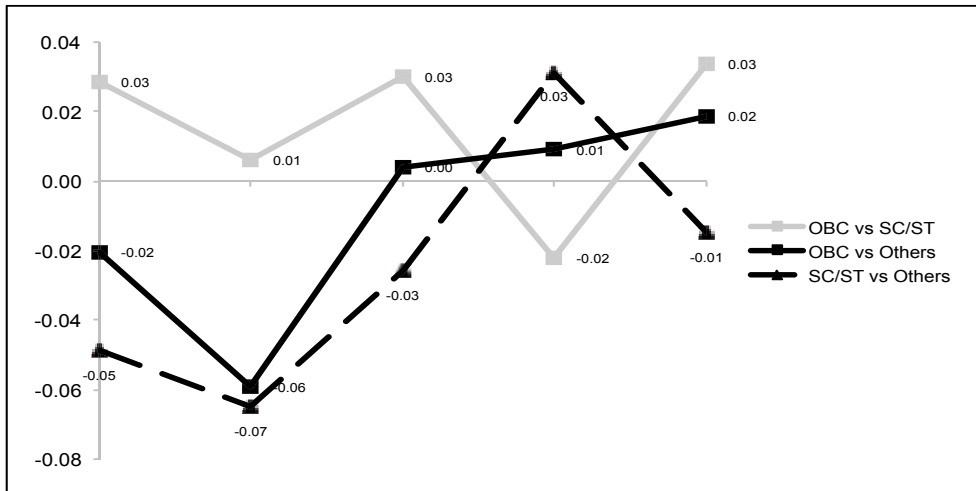

Figure 3.6: **Evolution in D-I-D in white collar jobs**
Note: A negative D-I-D for the line comparing OBCs to Others or the OBCs to the SC-ST implies OBCs relatively losing ground whereas a positive value implies convergence in case of comparison with the Others and divergence when compared to SC-ST. A negative value D-I-D line comparing Others to SC-ST implies the SC-ST are relatively losing ground when compared to the Others and positive value implies convergence.
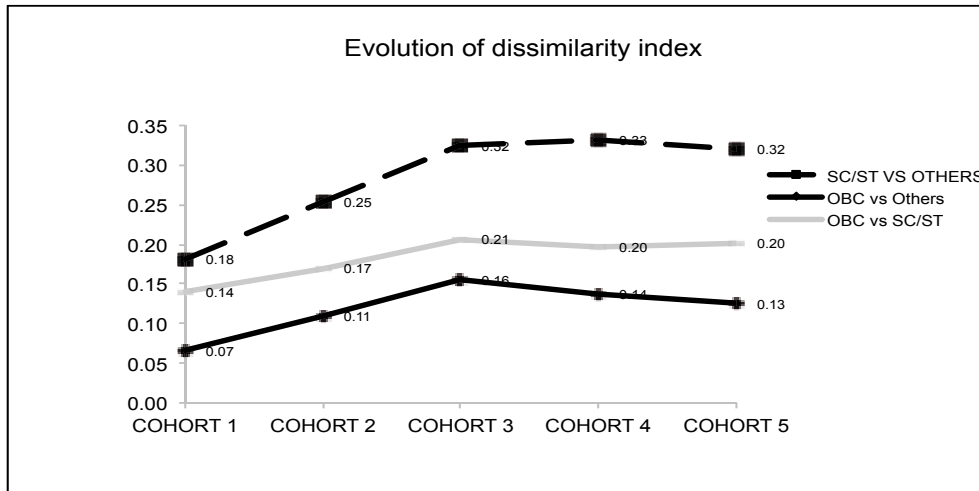
158

Figure 3.7: **Duncan Dissimilarity Index 1999-2000**
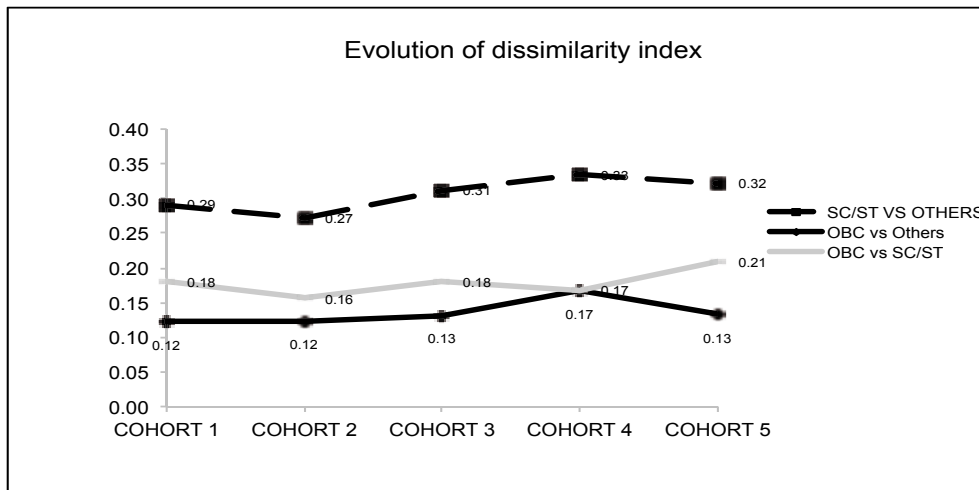Note: These represent Cohort 1 to 5 from the NSS 55th.

Figure 3.8: **Duncan Dissimilarity Index 2009-2010**
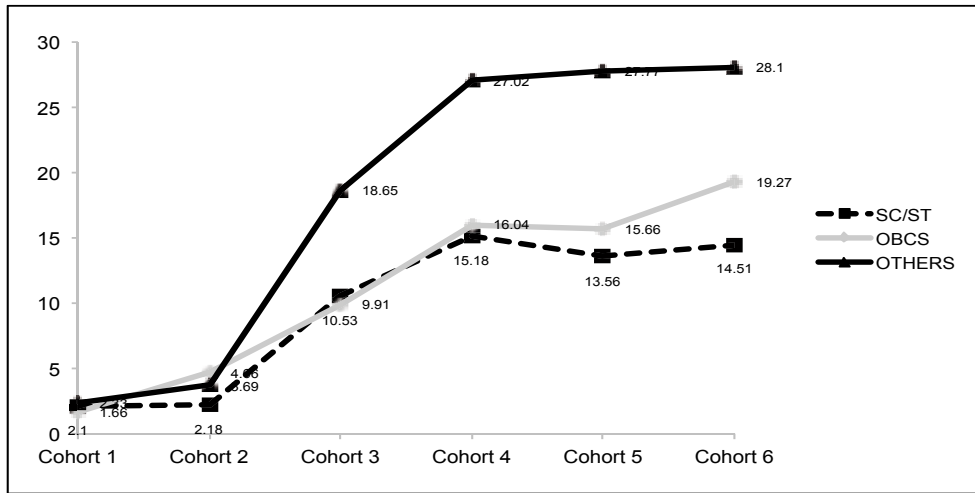Note: These represent Cohort 1 to 5 from the NSS 66th.

159

Figure 3.9: **Regular wage/salaried employees by cohort**

Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.
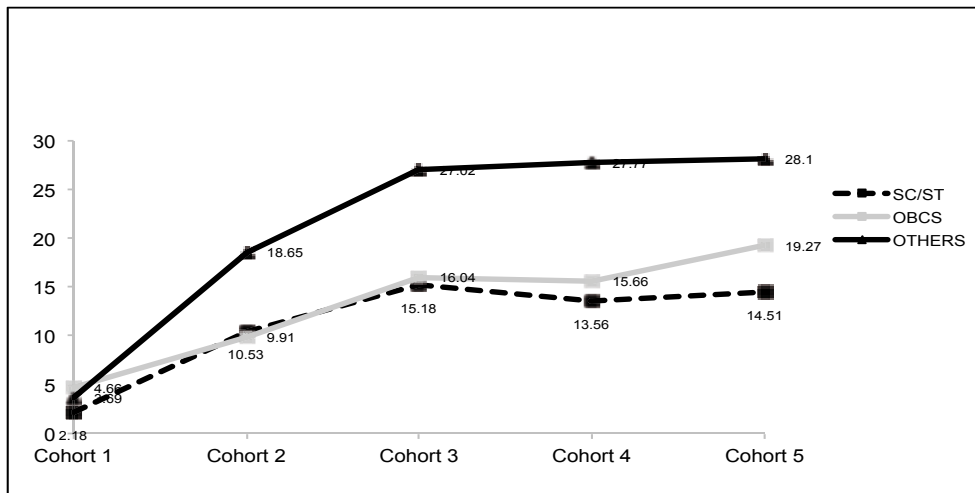

Figure 3.10: **Regular wage/salaried employees by cohort - NSS-66**

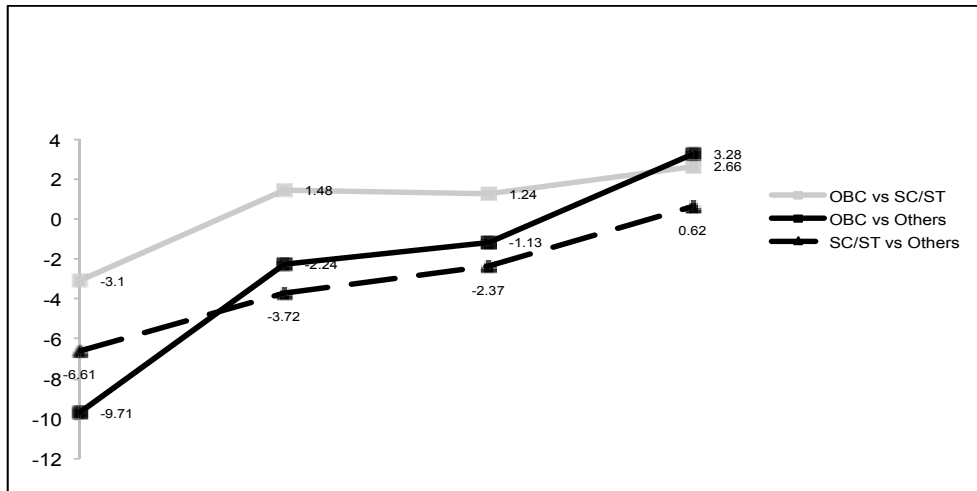Note: These represent Cohort 1 to 5 from the NSS 66th.

160

Figure 3.11: **Evolution of D-I-D in regular salaried employees**
Note: A negative D-I-D for the line comparing OBCs to Others or the OBCs to the SC-ST implies OBCs relatively losing ground whereas a positive value implies convergence in case of comparison with the Others and divergence when compared to SC-ST. A negative value D-I-D line comparing Others to SC-ST implies the SC-ST are relatively losing ground when compared to the Others and positive value implies convergence.
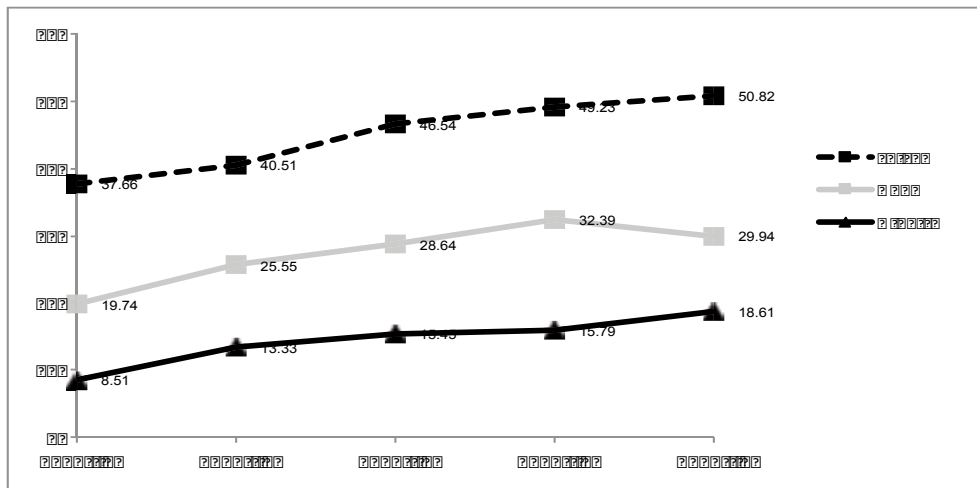
Figure 3.12: **Share of casual labour in workforce by cohort 2009-10**
Note: These represent Cohort 1 to 5 from the NSS 66th.

161

Figure 3.13: **Evolution of D-I-D in casual labour force share 2009-10**

Note: A negative D-I-D for the line comparing OBCs to Others or the OBCs to the SC-ST implies OBCs relatively losing ground whereas a positive value implies convergence in case of comparison with the Others and divergence when compared to SC-ST. A negative value D-I-D line comparing Others to SC-ST implies the SC-ST are relatively losing ground when compared to the Others and positive value implies convergence.
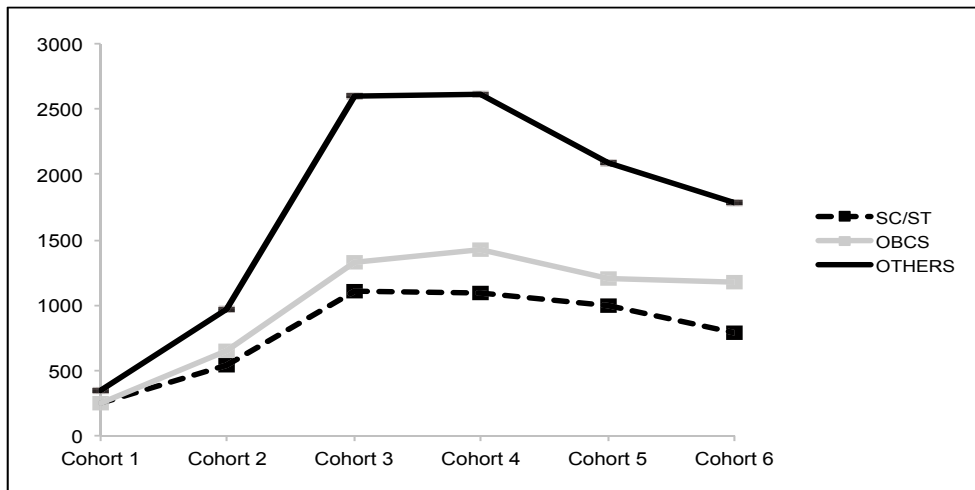

Figure 3.14: **Wages by cohort**

Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.
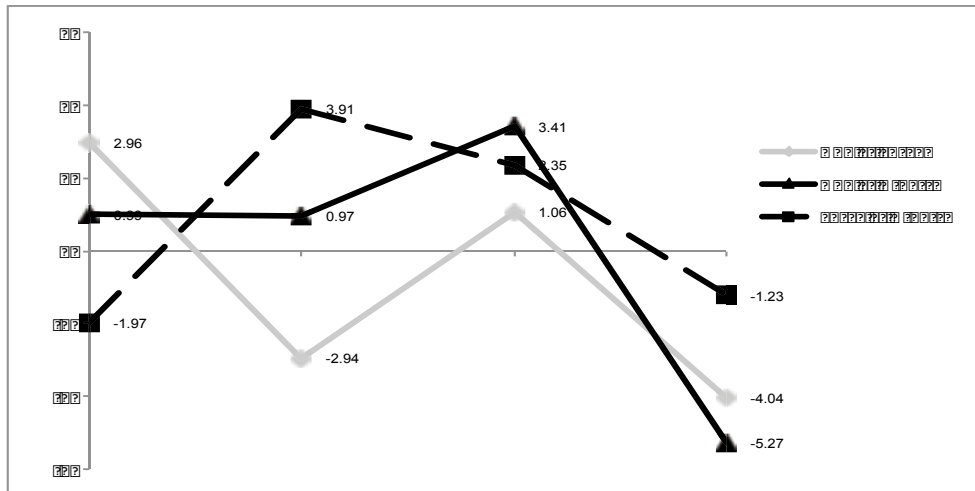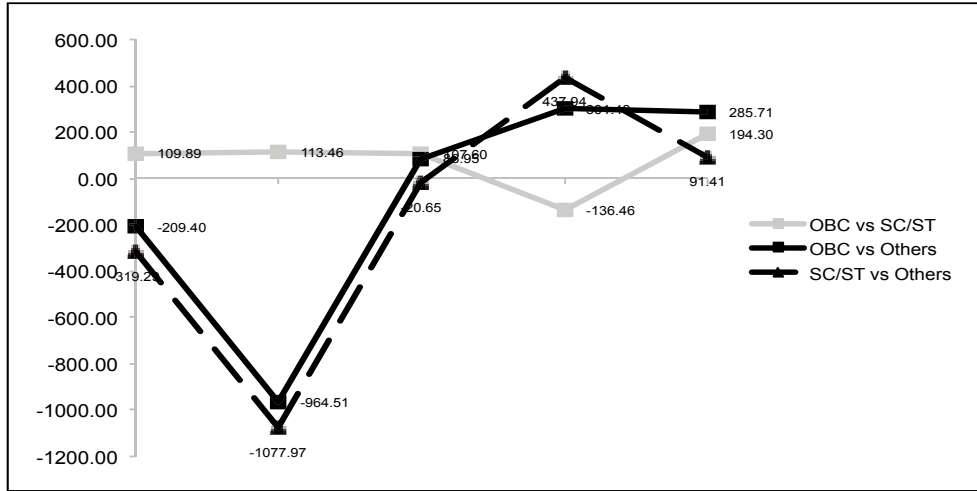
Figure 3.15: **Evolution of D-I-D in wages**

Note: A negative D-I-D for the line comparing OBCs to Others or the OBCs to the SC-ST implies OBCs relatively losing ground whereas a positive value implies convergence in case of comparison with the Others and divergence when compared to SC-ST. A negative value D-I-D line comparing Others to SC-ST implies the SC-ST are relatively losing ground when compared to the Others and positive value implies c
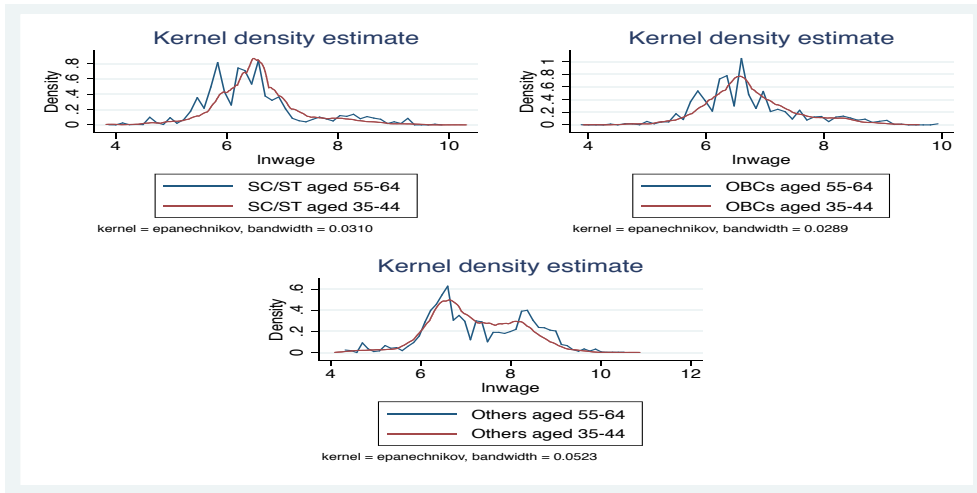
Figure 3.16: **The wage distributions of SC-ST, OBCs and Others' for 2009-10**

3.7   Appendix

3.7.1   The Blinder-Oaxaca Decomposition Methodology

The detailed methodology can be found in Jann (2008). In this appendix we explain the method intuitively for those not inclined to go into the technical details. In two independently written pioneering papers, Blinder (1973) and Oaxaca (1973) outlined the econometric methodology to decompose the average wage gap between two groups into two components: the "explained" component, or the part of the wage gap which can be explained by human capital or endowments (the wage-earning characteristics), and the "unexplained" component. The latter is interpreted as a measure of labour market discrimination as it is the part of the wage gap that remains unaccounted for after all the wage-earning characteristics are accounted for. The basic belief behind this approach is that wages differ both because of productivity or skill differences between groups as well as because the market treats the same characteristics differently. What can be observed are only the actual wage differences; the B-O method artificially separates the endowment/productivity differences from the treatment or the rate of return effect. The basic Blinder- Oaxaca method suggests substituting the estimated rates of returns from one group into the estimated wage equation of the other group to construct counterfactual wage distributions (if there are two groups being compared, as in our paper, there are two counterfactual wage distributions which get constructed). However, this leads to question of which counterfactual wage distribution would prevail in the absence of discrimination and one possible alternative to estimating two separate counterfactuals is to estimate a pooled model over both groups to get the reference coefficients (which are supposed to represent the non-discriminatory wage structure). We use the pooled method in the present paper.

164

3.7.2 Additional figures and tables

Table 3.11: **Evolution on educational indicators across cohorts**

Social Group	COHORT 1 (1)	COHORT 2 (2)	COHORT 3 (3)	COHORT 4 (4)	COHORT 5 (5)	COHORT 6 (6)
Years of education						
SC-ST	0.703	1.076	1.789	2.396	3.504	4.618
OBCS	1.145	2.068	2.846	3.764	4.697	6.093
OTHERS	2.997	4.281	5.557	6.558	7.327	8.304
Proportion of cohort literate or more						
SC-ST	0.148	0.201	0.294	0.382	0.504	0.626
OBCS	0.25	0.337	0.43	0.53	0.625	0.731
OTHERS	0.462	0.553	0.634	0.729	0.781	0.86
Proportion of cohort with primary schooling or more						
SC-ST	0.079	0.116	0.205	0.272	0.39	0.52
OBCS	0.13	0.234	0.319	0.416	0.508	0.636
OTHERS	0.312	0.437	0.549	0.635	0.706	0.784
Proportion of cohort with secondary schooling or more						
SC-ST	0.02	0.034	0.066	0.091	0.141	0.187
OBCS	0.03	0.082	0.114	0.156	0.214	0.299
OTHERS	0.133	0.205	0.306	0.362	0.414	0.48
Proportion of cohort with higher secondary schooling or more						
SC-ST	0.008	0.021	0.032	0.04	0.071	0.099
OBCS	0.011	0.032	0.049	0.076	0.104	0.16
OTHERS	0.063	0.122	0.175	0.225	0.259	0.312
Proportion of cohort with graduate degree or more						
SC-ST	0.006	0.012	0.02	0.02	0.036	0.046
OBCS	0.004	0.019	0.025	0.043	0.053	0.089
OTHERS	0.042	0.085	0.115	0.147	0.158	0.195

Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.

Table 3.12: Evolution on indicators of interest across cohorts

Social Group	COHORT 1 (1)	COHORT 2 (2)	COHORT 3 (3)	COHORT 4 (4)	COHORT 5 (5)	COHORT 6 (6)
Wages by cohorts						
SC-ST	244.171	539.399	1103.313	1094.255	1001.993	790.764
OBCS	244.45	649.571	1326.945	1425.486	1196.763	1179.836
OTHERS	346.864	961.386	2603.267	2614.855	2084.648	1782.01
Proportion employed in white collar jobs by cohorts						
SC-ST	0.034	0.054	0.064	0.065	0.083	0.069
OBC	0.038	0.086	0.103	0.133	0.129	0.149
OTHERS	0.092	0.16	0.236	0.263	0.25	0.25
Proportion employed in agricultural collar jobs by cohorts						
SC-ST	0.7885	0.657	0.6547	0.5658	0.5128	0.5358
OBCs	0.7455	0.6701	0.5852	0.491	0.4598	0.4425
OTHERS	0.7185	0.6569	0.4835	0.3755	0.3546	0.35
Proportion employed in blue collar jobs by cohorts						
SC-ST	0.1778	0.2895	0.281	0.369	0.404	0.3951
OBCs	0.2168	0.244	0.3121	0.3756	0.411	0.4086
OTHERS	0.1897	0.1827	0.2802	0.3616	0.3957	0.3996
Proportion employed in RWS jobs by cohorts						
SC-ST	0.021	0.0218	0.1053	0.1518	0.1356	0.1451
OBCS	0.0166	0.0466	0.0991	0.1604	0.1566	0.1927
OTHERS	0.0233	0.0369	0.1865	0.2702	0.2777	0.281
Proportion employed in casual labour jobs by cohorts						
SC-ST	0.3493	0.3766	0.4051	0.4654	0.4923	0.5082
OBCS	0.1674	0.1974	0.2555	0.2864	0.3239	0.2994
OTHERS	0.0922	0.0851	0.1333	0.1545	0.1579	0.1861

Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.

Table 3.13: Evolution of D-I-D on selected educational indicators

	D-I-D (COHORT(2-1)) (1)	D-I-D (COHORT(3-2)) (2)	D-I-D (COHORT(4-3)) (3)	D-I-D (COHORT(5-4)) (4)	D-I-D (COHORT(6-5)) (5)
Years of education					
OBC vs Others'	-0.36***	-0.5***	-0.08	0.16***	0.42***
	(0.08)	(0.10)	(0.08)	(0.07)	(0.06)
SC-ST vs Others'	-0.91***	-0.56***	-0.39***	0.34***	0.14***
	(0.13)	(0.11)	(0.08)	(0.07)	(0.06)
Literacy or more					
OBC vs Others'	-0.004	0.012	0.005	0.042***	0.028***
	(0.008)	(0.01)	(0.008)	(0.006)	(0.005)
SC-ST vs Others'	-0.038	0.011	-0.006	0.069***	0.044***
	(0.015)	(0.011)	(0.008)	(0.007)	(0.006)
Graduate or more					
OBC vs Others'	-0.028***	-0.025***	-0.013***	-0.001	-0.001
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
SC-ST vs Others'	-0.037***	-0.022***	-0.032***	0.005***	-0.027***
	(0.006)	(0.006)	(0.005)	(0.004)	(0.004)

Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, so covering the birth years 1926-85.

Table 3.14: Evolution of D-I-D on selected indicators

	D-I-D (COHORT(2-1)) (1)	D-I-D (COHORT(3-2)) (2)	D-I-D (COHORT(4-3)) (3)	D-I-D (COHORT(5-4)) (4)	D-I-D (COHORT(6-5)) (5)
Wages					
OBC vs Others'	-209.4**	-964.51***	86.95	301.48***	285.71***
	(101)	(221)	(102)	(64)	(60)
SC-ST vs Others'	-319.29***	-1077.97***	-20.65	437.94***	91.41
	(123)	(229)	(98)	(63)	(91)
Proportion employed in white collar jobs					
OBC vs Others'	-0.02***	-0.059***	0.004	0.009	0.019***
	(0.01)	(0.01)	(0.09)	(0.007)	(0.007)
SC-ST vs Others'	-0.049***	-0.065***	-0.026**	0.031***	-0.015**
	(0.015)	(0.015)	(0.10)	(0.007)	(0.007)
Proportion employed in public sector jobs					
OBC vs Others'	-0.0519****	-0.0145****	0.026****	0.0211****	
	(0.009)	(0.007)	(0.005)	(0.004)	
SC-ST vs Others'	-0.0514****	-0.0299****	0.038****	0.0103****	
	(0.011)	(0.008)	(0.006)	(0.004)	

Note: Cohort 1 is Cohort 1 of NSS-55 and Cohort 2-6 are Cohort 1-5 of NSS-66, for the indicator of wages and white collar jobs. The proportion employed in public sector jobs is for the first 5 cohorts of NSS-66.