The London School of Economics and Political Science

Development Writ Small

Thomas William O'Keeffe

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, July 2018

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of approximately 35000 words.

Acknowledgements

I am extremely grateful to the Palanpur team for welcoming me with open arms, particularly Himanshu, Ruth Kattamuri and Nicholas Stern. My thanks go to Sarthak Gaurav for all his work on the data and his limitless patience with my questions about it. Gajanand Ahirwar provided excellent research assistance and insight on the ground.

I am indebted to Robin Burgess, my supervisor, for his advice and support throughout. Gharad Bryan, Greg Fischer, Camille Landais, Gerard Padro i Miquel and seminar participants at LSE have provided extremely useful comments and advice at various points of the process.

I acknowledge financial support from an ESRC Studentship.

Clare Balboni, Florian Blum, Milad Khatib-Shahidi, Stephan Maurer, Jonathan Pinder, Francesco Sannino, Matthew Skellern, Eddy Tam, Torsten Walter, Guo Xu and many others, have been great colleagues and friends during my time at LSE. They have all contributed immeasurably to both my research and experience of the PhD process.

Many thanks must go to my family, both old and new. Lesley, Paul and Emily - you have always been there for me. Barbara, Andrzej and Leonard - spędziliśmy wspólnie już wiele wspaniałych chwil - nie mogę się doczekać kolejnych.

Lastly, and most importantly I must thank my wife Natalie who is the most inspiring and supportive person I know, without her none of this would be possible, and to Dzidzia who provided the best motivation, and the best distraction.

Abstract

This thesis is concerned with using micro-level data to examine important features of the process of development which occur on a much larger scale. Using a uniquely long and detailed dataset for a single village in India, allied with data from other sources, we explore what development at the level of a village can tell us about development at the level of a state or country. In the first chapter we introduce the village setting of this thesis - Palanpur, describe the data, and document the broad features of development experienced by the village over the course of 60 years. We focus on changes in employment, education, migration within the village - and relate these to the development of India or other areas where appropriate. The overriding picture is one of a village which has been touched by the outside world. The Green revolution initiated sustained growth in agricultural productivity. Large numbers have moved out of subsistence agriculture into non-agricultural pursuits, many of these outside the village. There have been substantial increases in education, migration, and income levels - similar in magnitude to other areas of India.

The second chapter investigates how structural transformation, the reallocation of economic activity from agriculture to manufacturing and services, is experienced for economic entities smaller than countries. Despite a vast macroeconomic literature concerning structural transformation for countries along their development path there is little evidence on the nature of structural transformation at a more microeconomic level. Firstly, we document the stylised facts of structural transformation from the empirical macroeconomic literature. Secondly, we show that these stylised facts are consistent with India's development experience over more than 100 years. We then proceed to document how these empirical facts map onto progressively smaller geographic areas

within India. Finally we demonstrate that these features of structural transformation hold true even at the level of a single village in India. The pattern of sectoral reallocation in terms of both income and employment shares is strikingly similar and consistent with the extant stylised facts at all levels. This result has important implications for the way we should think about the complementarity of agricultural and non-agricultural development.

The third chapter explores the role of employment networks within the process of development in rural India. The relevant networks we examine are caste and extended family networks, called dynasties. We first establish that there exist job networks in non-agricultural employment for individuals working outside the village. These networks have large effects, and these effects are larger for extended family networks. We then demonstrate that these job networks exhibit competition from fellow network members. As a placebo test we confirm smaller or non-existent network effects for another type of employment believed to be less prone to job referral networks. The second part of this chapter then tests if these dynasty network effects observed for outside employment are consistent with a model of labour market network dynamics. The data are consistent with the model and display both a negative competition effect and a positive information effect. Dynasty network cohorts who arrive in the labour market prior to workers have a positive effect on their employment prospects but those who arrive at the same time have a negative impact. The chapter finishes with some evidence on the potential long run implications of these networks.

Contents

List of Tables

List of Figures

Chapter 1

Palanpur: Development Writ Small

This thesis is concerned with using micro-level data to examine important features of the process of development which occur on a much larger scale. Using a uniquely long and detailed dataset for a single village in India, allied with data from other sources, we explore what development at the level of a village can tell us about development at the level of a state or country. In this chapter we introduce the village setting of this thesis - Palanpur, describe the data, and document the broad features of development experienced by the village over the course of 60 years. We focus on changes in employment, education, migration within the village - and relate these to the development of India or other areas where appropriate. The overriding picture is one of a village which has been touched by the outside world. The Green revolution initiated sustained growth in agricultural productivity. Large numbers have moved out of subsistence agriculture into non-agricultural pursuits, many of these outside the village. There have been substantial increases in education, migration, and income levels - similar in magnitude to other areas of India.

1.1. Introduction

In recent years there has been a focus on panel data with large numbers of individuals over short time periods which are used to answer many questions in development. In this thesis we ask what we can learn about the process of development from another type of dataset: one which examines the same individuals over a long time period. There are interesting and important features of development which by their nature take time, and therefore cannot be examined without this time dimension. The Palanpur data upon which this thesis is based contains individual level data that span 1957 to 2015.

The reason for the general focus on panels which span short time periods in development is obvious. One cannot simply collect 60 years of data at a given time, it takes 60 years to collect it. This is typically longer than an academic career, thus the incentives are not strong for academics to establish panels which are focused on long run questions. In developed countries some of this space is occupied by national statistical offices who have the capacity to invest in long run panels. The longest running of these is the United States Panel Study of Income Dynamics which began in 1968[1], more than 10 years after the commencement of the Palanpur data. Other countries have household panels which begin usually in the 1980s and 1990s - for example Germany, United Kingdom, Canada and Japan. However, in developing countries there are generally greater concerns for national governments than collecting panel data, even if the country has the state capacity to do so. Notable household panel data which exist in developing countries include the Indonesian Family Life Survey which began in 1993, the KwaZulu-Natal Income Dynamics Survey (KIDS) of KwaZulu-Natal province in South Africa which began in 1993 and was extended in 2008 to the National Income Dynamics Survey (NIDS) which covers

[1] https://psidonline.isr.umich.edu

the whole country.

Those panel studies which are closest to the Palanpur data are those in India and Bangladesh. The Additional Rural Incomes Survey and Rural Economic and Demographic Survey (ARIS-REDS) panel from the NCAER includes four rounds between 1971 and 2006[2]. The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) Village Level Studies began in 1975 with a sample of 240 households from six villages in two Indian states. It has been extended at different times to include households from a further fourteen villages. However, these surveys are samples of households within villages rather than a census, so they are prone to losing households through death or splitting of households (Foster and Rosenzweig (2002)). Similar in spirit, if not in content, is the Health and Demographic Surveillance System in Matlab, Bangladesh, which began registering all births, deaths and migrations in 149 villages in 1966. While there have been occasional censuses undertaken, there is only detailed economic data for a representative sample in 1996.

We are fortunate to have access to a panel of detailed information from a single village in India beginning in 1957 and ending in 2015. This is a longer timespan than the World Bank's World Development Indicators. It is also longer than 61% of the country-level GDP series from the Penn World Tables data (version 9), Feenstra et al. (2015).

These data are incredibly detailed on many dimensions. They consist of 7 surveys which are undertaken as censuses of the village. Data collected by enumerators include demographics, land ownership, cultivation, land rental, outstanding loans, occupation, labour contracts in which village residents are engaged, wage rates, income, nutrition and consumption. Although the exact dimensions upon which data were collected varies

[2]See Foster and Rosenzweig (2002); Foster and Rosenzweig (2004) or the NCAER website for details of this panel: http://www.ncaer.org/data.php

by survey, the demographic, employment, land ownership, and education data are available for the full period. There is also a huge amount of qualitative data and observations, including diaries kept by researchers for many months.

During the period covered by the Palanpur data, India has experienced a huge amount of change. In 1957 India was a young independent country, seven years after completion of their constitution, with a GDP per capita of 944 in 2011 international dollars[3]. While by 2015 India was one of the world's fastest growing economies, and the 3rd largest in PPP terms[4]. GDP per capita in 2014 was 5,452 in 2011 international dollars. The development of India in this period has not just been economic. Life expectancy at birth has risen from 39.65 between 1955-1960 to 67.57 between 2010-2015, in the same period under-five mortality has fallen from 255 to 52 per 1,000 live births[5]. Coincident to this improvement in health outcomes, India has experienced a demographic transition - total fertility as measured by children per woman has declined more than 50 percent[6]. Palanpur is uniquely positioned to explore the causes and consequences of this tremendous change at the micro level.

These data are able to help us to answer a number of questions which are important to development economists. Have the increases in economic development transferred to the rural poor? Have these improvements in well-being changed the income distribution? What have been the causes and consequences of these changes in relative income? What determines who migrates away? Has this changed over time? The data also allow us to truly examine intergenerational effects, not just look at contemporaneous effects of parental states as is common in the literature due to data limitations. It also may be

[3]Output-side real GDP per capita from PWT 9.0
[4]Feenstra et al. (2015)
[5]United Nations (2017)
[6]1955-1960: 5.90. 2010-2015: 2.44 from United Nations (2017)

that comparisons across space, in lieu of available longitudinal data, result in different answers due to confounding locational fundamentals or other confounders about which we are unaware. The use of a single location over a long time period avoids this potential pitfall.

This chapter describes in detail the setting and data for this thesis, and the experience of this village over nearly 60 years of study. The following section introduces Palanpur and describes the broad changes it has experienced during the study period. Clearly we are not the first to examine this panel. Palanpur data have been the source of many important contributions to the understanding of Indian development, and development more generally. These are discussed in Section 1.3. The subsequent sections examine Palanpur in more detail, with emphasis on its employment (Section 1.4), education (Section 1.5) and migration (Section 1.6) patterns during the 60 year study period. Section 1.7 concludes and introduces the following chapters of this thesis which focus on specific questions which can only be answered with data such as these.

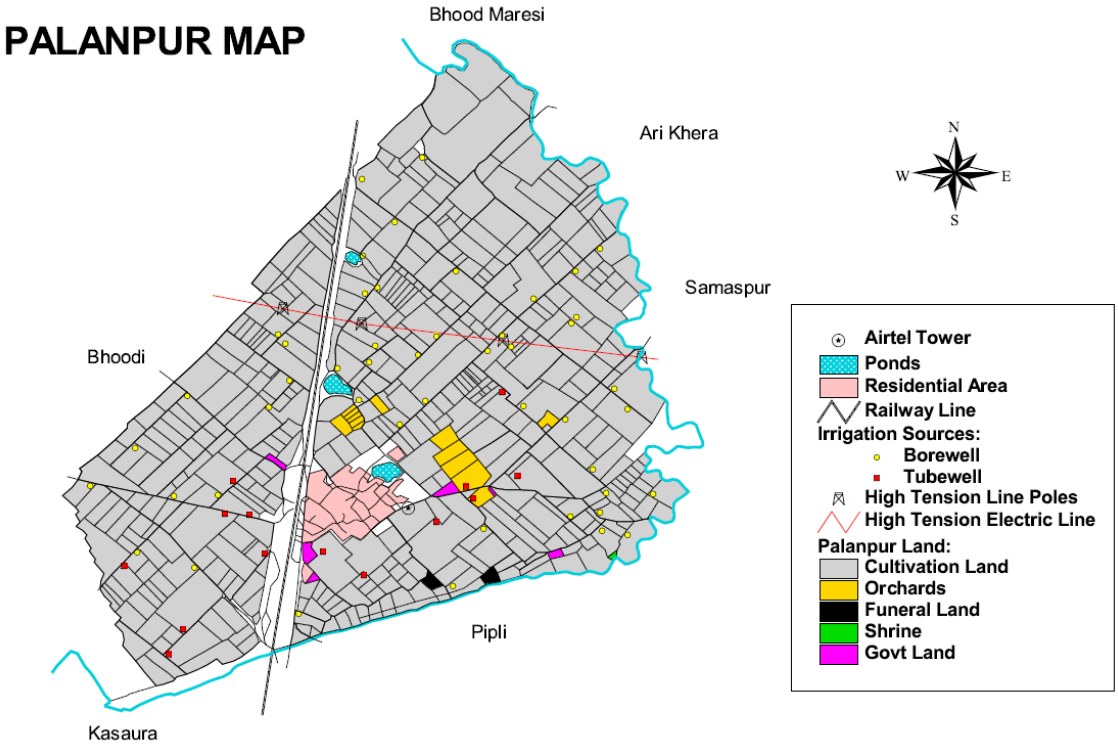
1.2. Palanpur 1957-2015

Palanpur is probably the most famous of India's 640,930 villages[7]. It has been studied by social scientists since 1957, when the Agricultural Economics Research Centre (AERC) of the University of Delhi began their first survey. The village has since been visited by surveyors in every decade, in 1963/64, 1974/75, 1983/84, 1993, 2008/10 and 2015[8]. The data is now of such length that it includes a number of great-grandchildren of original 1958 survey members.

[7] Ministry of Rural Development, 2012. http://mospi.nic.in/Mospi_New/upload/SYB2014/ch42.html (Accessed 4/8/16)

[8] For simplicity, from now on I will refer to the surveys as 1958, 1964, 1975, 1984, 1993, 2009 and 2015 respectively.

Figure 1.1.: Palanpur in 2009



Note: GIS Map of Palanpur in 2009

Source: Palanpur Survey Team. Constructed by Ruchira Bhattacharya, Himanshu and Ashish Tyagi, reproduced with their permission.

Palanpur was not chosen because it was unique, it was just one of many villages chosen to be part of the AERCs continuous village surveys. The reason we remember it and study it still is that Christopher Bliss and Nicholas Stern were looking for baseline data to examine the impacts of the "green revolution", and went back to re-survey Palanpur in 1974. Nicholas Stern has been involved in every survey of Palanpur since. While the intention of the original surveys was the study of agriculture, they collected data on many aspects of the village. Bliss and Stern chose Palanpur because it was not dominated by "an unusual craft or crop", and because it was "large enough to encompass a certain amount of variety but not too large to permit a detailed study including all its households." They also required wheat to be a major crop in the village, and for the village to be accessible from Delhi but not too close to ensure independence from its economy. They hoped that tenancy would be common since "that subject was one of their main interests." They also required independent living arrangements for the team of four researchers.

The data contain a huge amount of qualitative and quantitative data. Enumerators spent an enormous amount of time in the village conducting the surveys. The surveys were conducted as censuses of the village. There is detailed information on demographics, land ownership, cultivation, land rental, outstanding loans, occupation, labour contracts in which village residents are engaged, wage rates, income, nutrition and consumption. Plus a number of the enumerators kept detailed diaries which are full of qualitative data.

Palanpur is located in Moradabad district in Uttar Pradesh, around 205km from Delhi. There is a train station in the village, evident in Figure 1.1, this line connects directly to Delhi in around 6 hours. Although much closer are Chandausi 13km away, a town of approximately 114,254; and Moradabad 31km away, the district capital with a population

of 889,810.[9] Not every train stops at the station, but many do - enough to allow villagers the possibility to commute daily to Chandausi or Moradabad. Chandausi contains the market at which most of Palanpur's produce is sold. When it was first surveyed Palanpur was a village of 534 people and has grown to one with a population of 1,299 by 2015. The longevity of the data means that we observe in some instances four generations of the same family. The details of this population growth by survey, as well as the corresponding growth in the number of households is shown in Table 1.1.[10] Also included in Table 1.1 are details on the evolution over the course of the survey period of variables relating to land holdings and usage, agriculture in the village, income and educational attainment.

Noticeable in Table 1.1 is the growing importance of the non-farm sector, both as a source of employment and income. In this thesis we define the non-farm sector following Lanjouw and Lanjouw (2001). Thus the non-farm 'sector' includes all economic activities except agriculture, livestock, fishing and hunting. Agricultural production remains important in Palanpur throughout the 60 year survey period (in 2015 more than 70% of village households are still engaged in farming their own land), although the average area of land cultivated by village households has fallen considerably, as well as the total area of land cultivated by village households. The average number of acres of land owned by households is less than 30% of its original 1958 amount. All these landholdings

[9] Government of India (2011) - http://www.censusindia.gov.in/2011-prov-results/paper2/data_files/India2/Table_2_PR_Cities_1Lakh_and_Above.pdf

[10] In most of the Palanpur work to date the village has 528 people in the 1958 survey spread among 100 households. This is because the original survey questionnaires are missing for 3 households. However, all 3 are surveyed again in 1964 and report the households as unchanged. So while there is no data on their income, debts, agricultural output etc. for 1958, we know how they relate to others in the village. Since the focus of Chapter 3 is primarily the importance of networks we include these individuals since they would still form part of other villagers networks in 1958. It is also possible that other populations or estimates may differ from those previously calculated using the Palanpur data given that this thesis is based on new cleaning and matching of the data. Any errors that remain are solely the responsibility of the author.

lie within 3km of the village.

Table 1.1 also demonstrates the growing importance of a market for land. This despite leasing land being illegal (at least during some of the period) and illegally leased-out land being liable to confiscation. Individuals are still happy to be landlords and to talk about it freely, Bliss and Stern (1982):p.24. Previous examinations of Palanpur have shown that most tenancy in Palanpur takes the form of sharecropping, and also that households are happy to lease "down caste ladder", Bliss and Stern (1982):p.25. That is from high caste households to lower or scheduled caste households, but not the other way round. Leasing of land also takes place within caste. For a more detailed discussion of Caste in Palanpur see Chapter 3, Section 3.3.

The unique elements of the Palanpur data, aside from it's longevity, are that the data are collected as a census of the village, and it is possible to track how each household or individual are related to each other, if at all. The data is taken as a series of snapshots of the village over 60 years and there is no systematic attempt to create a timeline of events which would enable the creation of a yearly panel from this data, since that was not the purpose of the original studies.

However, due to the census nature of the data we know not only the individuals present in the village who are enumerated in each survey. We also know about those members of Palanpur who migrate out of the village, even if they are not present in any survey. This is because households are asked about any members who are now not present when they were in previous surveys, and also about any births in the household in intervening years. Thus we know basic demographic information about everyone who was at one time a member of the village, although we do not have income or employment data for those migrants who have left permanently. As a result the sample of unique individuals who are present in at least one of the Palanpur surveys is 2,752, but

the number of unique individuals in what we refer to as the migration sample is 2,831.

Despite being much poorer than India on average[11], the development experience of Palanpur over the last 60 years has been remarkably similar to that of India as a whole. Village average real incomes per capita have increased at an annualised growth rate above 1.5% per year from 1958 to 2009. Despite having a population in 2009 almost 1 million times smaller than India, in Chapter 2 we show that both India and Palanpur are consistent with the stylised facts of structural transformation described by the literature across countries[12]. That is, Palanpur has experienced large reallocations of resources away from agriculture into other sectors. Table 1.1 shows this at the household level in the large increase in households who are engaged in non-farm activities, together with the increase in the percentage of income which comes from non-farm sources. Chapter 2 shows that this is also true for employment shares, in 1958 56.6% of males are employed in agriculture as their primary occupation, by 2015 this is 17.3%. Together with, and necessary for, this sectoral reallocation has been increasing agricultural productivity. Table 1.1 illustrates this increase in productivity in terms of real output values, but it is also true in yield terms. Cereal yields in kilograms per hectare increased from 559 in 1958 to 2,958 in 2009.

[11] The average income in the village in 2009 is below the Indian average GDP per capita of the early 1960s.
[12] See Herrendorf et al. (2014); Buera and Kaboski (2012b); Buera and Kaboski (2012a) for these stylised facts across countries.

Table 1.1.: Village Descriptives

	1958	1964	1975	1984	1993	2009	2015
Demography							
Population	534	583	793	968	1121	1246	1299
Households	103	106	117	143	193	233	213
Mean HH Size	5.18	5.50	6.78	6.77	5.81	5.35	6.10
Land							
Total Land Cultivated	474.59	454.28	392.15	433.80	365.43	354.71	.
Land owned (acres) by HH	4.48	4.17	3.60	2.91	1.95	1.35	1.28
% Landless HH	12.62	12.26	9.40	20.28	21.76	16.74	23.47
% Land Leased	9.95	11.94	23.59	34.27	26.44	37.95	.
Agricultural							
% HHs in Non-Farm	16.50	26.42	39.32	60.84	45.08	60.94	57.75
% HHs in Own Cultivation	82.52	87.74	79.49	66.43	66.32	71.67	70.42
% HHs in Agricultural Labour	25.24	14.15	23.08	17.48	15.03	9.87	2.35
Land Productivity	118.59	128.03	280.69	273.68	.	400.21	.
Labour Productivity	365.70	412.32	871.96	510.82	.	555.53	.
Income							
Real HH Income Per Capita	190.36	211.46	278.87	235.42	.	413.52	.
Income: % Non-Farm	13.30	17.51	17.19	31.85	.	46.60	.
Education							
Mean Max Education	1.39	1.86	2.10	2.45	2.66	2.83	3.19
Mean HH Education	1.11	1.28	1.29	1.37	1.54	1.73	1.96
% HHs Educated Class 9+	1.94	14.15	19.66	41.96	48.19	55.79	81.69

Note: Variable definitions are in Appendix Table A.1.

Source: Author's Calculations

1.3. Contributions of the Palanpur Data

We are not the first to use the Palanpur data, and surely not the last. There are many publications which use at least some elements of the data, beginning with Ansari (1964) who reports on the first survey. The most comprehensive accounts of Palanpur are contained in the two books published to date, Bliss and Stern (1982) and Lanjouw and Stern (1998), and there is a third book which is forthcoming.

Bliss and Stern (1982) details the village over the first three surveys, and is focused on the functioning of rural markets and the behaviour of farmers. The great contribution of this book, aside from beginning the creation of such a long run dataset, was to increase the understanding of sharecropping contracts as a response to imperfect or missing input markets. While Lanjouw and Stern (1998) focus primarily on what they describe as the three main drivers of change in Palanpur - population growth, technological change in agriculture, and increased employment opportunities outside the village - using surveys up to and including 1993. However, they also document changes in poverty, inequality and nutrition, plus the workings of the labour, tenancy and credit markets within the village. The book contains more than we could possibly summarise here. The forthcoming third book will extend this description of the village and its changes into the 21st century.

In addition to these books, Palanpur has spawned a number of journal articles concerning different aspects of rural life. Bliss and Stern (1978a) and Bliss and Stern (1978b) made important contributions to early nutrition-based efficiency wage theory and evidence.

Lanjouw (1999) uses Palanpur data from the 1975 and 1984 surveys to estimate a general equilibrium model of land leasing. This is then used to show that village markets are

not complete or efficient, and that knowledge about the relative farming skill of farmers is common information in the village.

Other papers have examined the experience of agricultural labourers within Palanpur. Baland et al. (1999) use Palanpur data to explain the coexistence of daily-wage and piece-rate contracts in agriculture. Their key finding is that the most able and least able labourers work on piece rates, as they can choose their own level of effort. Mukherjee and Ray (1991) use Palanpur 1984 survey data to test the implications of their model of wage bargaining of agricultural labourers in the presence of seasonality.

Sharma and Drèze (1996) present a case study based on the census of tenancy contracts collected as part of the 1984 survey. They make a contribution to the understanding of sharecropping as the predominant tenancy arrangements in rural India.

Himanshu et al. (2013) and Himanshu, Joshi et al. (2016) document the increasing diversification of household incomes associated with non-farm employment. They show that this increase in non-farm employment has been accompanied by increasing inequality and mobility, as well as falling poverty rates.

Relatedly there have also been contributions to the study of inequality from the Palanpur data. Lanjouw and Rao (2011) highlight the flaws in standard between group inequality measurement using caste groups in Palanpur. They show that caste is an important determinant of inequality within the village, and this inequality does not seem to be driven by education or land endowment differences across castes.

There are also a series of more recent working papers which focus on different elements of village life up to the 2009 survey: Agriculture (Himanshu and Tyagi (2011a)), Tenancy (Himanshu and Tyagi (2011b)), Education (Kattumuri and Tiwari (2011)), Non-farm Diversification (Himanshu, Lanjouw, Mukhopadhyay et al. (2011)), Poverty, Inequality and Mobility (Himanshu, Bakshi et al. (2011a)), Women in Palanpur (Sinha and

Coppoletta (2011)), and Nutrition (Sinha (2011)).

The Palanpur data also contribute three chapters to Himanshu, Jha et al. (2016) which endeavours to document the history and results of village studies in India, beginning with the "Slater villages" in Tamil Nadu in 1916.

1.4. Employment in Palanpur

Historically the villagers of Palanpur, like any other village in rural India, relied on subsistence agriculture for their livelihood. However, as we have seen in the previous section, the importance of non-agricultural sources of income has been growing.

The change in occupational structure in Palanpur has mainly come through an increase in the opportunities outside the village. The percentage of males of 18 or over working outside the village increased from 6.2% in 1958 to 35.9% in 2015. While the percentage of males of 18 or over working in non-agricultural pursuits within the village has increased from 11.1% to 21.9%. Many of the jobs outside the village have emerged in larger villages or the towns of Moradabad and Chandausi, 31kms and 14kms away respectively. These jobs have developed as the towns have increased in size. Moradabad, the district capital, was just over 200,000 people in 1975[13], approximately 444,000 people in the 1991 census, and had grown to almost 900,000 people in the 2011 census. Chandausi, while smaller than Moradabad, has also grown significantly during the last 60 years[14]. It is also the primary market for agricultural products from Palanpur. Both Moradabad and Chandausi are on the train line which runs through Palanpur. As of 2017 there are 5 trains a day to Moradabad and 4 trains a day to Chandausi from Palanpur. The journey

[13] Bliss and Stern (1982)

[14] Chandausi population: 83000, 1991: 103,757, 2001: 114,383, 2011 from Census data. Estimated 1961 population 45,000 from supplement to the Gazette of India, September 16, 1961.

times are between 30 mins (to Chandausi) and 1 hour and 10 mins (to Moradabad). Thus, they allow for members of Palanpur to commute to either of these towns. Himanshu, Joshi et al. (2016) discuss the effect of this employment outside Palanpur on others in the village "A common perception among the villagers was that employment outside of Palanpur was definitely more remunerative than in Palanpur, as the living standards of these families had improved. This also served as a source of inspiration for many others who wanted to go out of the village in search of employment." This impression is supported by the available wage data. For example, in 2009 an individual working in agricultural labour within the village would earn Rs. 100 per day for their labour, while work in the Moradabad railway yard would earn them, on average, Rs. 200 per day. This is not a skilled form of employment, but is a popular form of casual labour undertaken by village residents.

There is an important distinction to be drawn at this point about the types of jobs which are available to Palanpur residents. The key distinction is that of employment being "regular". To a first approximation a regular job is one which involves monthly rather than daily wage payments. These jobs are sought after since they offer numerous benefits. "Employees with such jobs usually enjoy a modicum of employment security, and, in most cases, substantially higher earnings than casual labourers." Lanjouw and Stern (1998):p.128. Examples of these regular jobs in which members of Palanpur are employed are factory work (steel, textiles, cement, biscuits), teaching, bank cashier, insurance agent, railway employee. These jobs form a significant proportion of all jobs; the percentage of adult males who are present in the village in regular employment across all surveys is 14.0%. The rate of employment in these jobs has not risen a great deal during the survey period, from 7.4% to 10.7%. However, this disguises a hump-shape in regular employment rates which peaks in 1984 at 21.7%. The explanation for this feature

is not clear, although one obvious candidate is changing migration rates for those who are more likely to become regular workers.

The type and location of employment of Palanpur residents are not independent. In fact, 71.3% of all regular jobs done by an adult Palanpur resident in any survey are outside the village.

1.5. Education in Palanpur

Education in Palanpur, much like education across the whole of India, has risen substantially throughout the survey period. India's primary completion rate was 98% in 2014, up from 40% in 1971[15]. For Palanpur it is not possible to generate comparable numbers since we do not have enrolment, only self-reported schooling. We can calculate the percentage of Palanpur residents aged between 12 and 18 who are educated to class 6 or above (primary ends after class 5). This increases from 25.8% in 1958 to 45.8% in 2015. There has been much discussion about whether this increase in enrolment rates across India has led to real increases in educational attainment, see Bosworth et al. (2006) for example. The same concerns are held anecdotally within Palanpur. In all the following we abstract from educational quality and focus only on education completed. We will however focus primarily at the high end of educational achievement and create an indicator variable, since the distinction between 5 and 7 years of schooling, for example, is not likely to be important in this context.

Measures of education are self-reported school years in the Palanpur data, and to allow for comparability across all surveys these data are collapsed into bins of education level - illiterate, class 1 to 5, class 6 to 8, class 9 to 12, class 12 or above. For those individuals

[15]UNESCO Institute for Statistics - SE.PRM.CMPT.ZS variable.

Table 1.2.: Male Education in Palanpur by Year

	1958	1964	1975	1984	1993	2009	2015
% Illiterate	80.00	73.10	70.34	62.63	50.97	29.70	21.65
% Class 1 to 5	12.78	12.69	11.44	6.40	8.08	23.10	24.33
% Class 6 to 8	3.89	7.11	8.05	12.79	15.88	19.04	18.97
% Class 9 to 12	1.11	6.09	6.36	16.16	20.89	23.10	25.00
% > Class 12	0.00	1.02	3.39	2.02	1.95	5.08	8.26
% Missing Education	2.22	0.00	0.42	0.00	2.23	0.00	1.79
Individuals	180	197	236	297	359	394	448

Note: Percentage of males over 14 years of age by education level and survey. Only those present in Palanpur are included.

Source: Author's Calculations

who appear as 14 or older in at least one survey, 16.2% of them are educated to class 9 or above. This overall number hides important differences between the educational attainment by sex. Males who appear in at least one survey when older than 14 are over 5 times more likely to be educated to class 9 than equivalent females, 27.5% and 5.4% respectively. A more detailed look at the education of Palanpur is contained in Table 1.2 for males, and Table 1.3 for females. These tables both show strong increasing trends in literacy[16].

There is a complication when considering the educational attainment of the females in Palanpur. Due to the process of patrilocal exogamy, and early marriages in this part of India, many girls born in Palanpur disappear from the education sample before they have their final education enumerated. For example, someone who appears as an 8 year old in the 1975 survey and gets married at 16, moving to their new husband's village, will

[16]This is true even if we assume all individuals with missing educational data are illiterate.

not appear in the 1984 survey. So we would not capture her final education level. Thus, many of the women we sample are those educated in other villages or towns who move in when they marry a Palanpur male. While the level of education of these women and the nature of assortative matching are interesting questions in themselves, they are not the ones we focus on in this thesis. As such, we will mostly be restricting our samples to males in what follows.

A complication remains when we are considering male educational attainment, that is of migration. If migration is selective and migration patterns change over time then we are not capturing the true trends in education for those born in Palanpur, but for those living in the village. We discuss the nature of migration in Palanpur in the following section.

Table 1.3.: Female Education in Palanpur by Year

	1958	1964	1975	1984	1993	2009	2015
% Illiterate	96.18	98.84	93.36	93.23	87.96	75.70	59.95
% Class 1 to 5	1.27	0.58	3.79	3.01	2.16	11.25	16.35
% Class 6 to 8	0.00	0.00	1.42	1.50	5.25	7.93	8.53
% Class 9 to 12	0.00	0.58	0.00	2.26	2.78	3.84	4.27
% > Class 12	0.00	0.00	0.00	0.00	0.62	1.02	1.66
% Missing Education	2.55	0.00	1.42	0.00	1.23	0.26	9.24
Individuals	157	173	211	266	324	391	422

Note: Percentage of females over 14 years of age by education level and survey. Only those present in Palanpur are included.

Source: Author's Calculations

1.6. Migration in Palanpur

As in all rural development settings, migration is a key feature of Palanpur society. 32.7% of adult males have migrated out of Palanpur in at least one survey. Foster and Rosenzweig (2008) find 21.2% of males aged between 10-24 in 1982 migrate out of their village between 1982 and 2002 in Bangladesh[17]. Using the NCAER REDS survey for India they find 23.7% of males aged between 10-24 in 1982 migrate out of their village by 1999. Table 1.4 shows the results of the same calculations when performed on the Palanpur data. We find that for periods roughly similar to the 17-20 year periods looked at by Foster and Rosenzweig (2008) the migration percentage for males in Palanpur ranges from 14.7% to 32.2%, with a clearly increasing trend.

Table 1.4.: Percentage of Out-Migrants

	1958-1975	1964-1984	1975-1993	1993-2009
Male	14.67	17.39	22.76	32.22
Female	51.72	51.61	55.91	72.07

Note: Includes individuals present in Palanpur who are aged 10-24 in the earlier of the two surveys examined, who have left Palanpur by the later survey.

Source: Author's Calculations

Table 1.4 also shows the same calculations done for females aged between 10-24 in Palanpur between the same surveys. These range between 51.6% and 72.1%, percentages which are also in line with those from Foster and Rosenzweig (2008). They find that 56.3% of equivalent females in Bangladesh migrate out of their villages, while in India it was 86.9%. These numbers are unsurprisingly higher than for males due to the practice of patrilocal exogamy - females move to the village of their husband when they get married.

[17] Using the Bangladesh Nutrition Surveys of 1982 and 2002.

However, the percentage of females leaving the village appears to have increased in more recent years. This may be because females are now not just migrating for marriage but also for employment or educational opportunities, or it may be that there is an increase in the number of households who are migrating out of the village *en masse* and this includes females who previously would have remained in the village. Table 1.5 shows the number of households which migrate out of Palanpur as a collective between surveys. This appears to indicate that it is this increased propensity for households to migrate out between 1993 and 2009 which has led to the sharp increase in female migration seen between 1975-1993 and 1993-2009 in Table 1.4. No doubt this increased propensity for collective household migration also explains some of the increase in the male migration rates between these two surveys. However, it is impossible for us to disentangle these increases into individual and collective migration since we do not have information on the migration destinations of all individuals.

Table 1.5.: Households Migrating Out

	1958-1964	1964-1975	1975-1984	1984-1993	1993-2009	2009-2015
Households	8	14	6	13	39	21
Individuals	21	52	18	78	204	122

Note: Households: Unique households which migrate out of the village as a whole. Individuals: Number of unique individuals within these migrating households.
Source: Author's Calculations

Foster and Rosenzweig (2008) find in their samples that out-migration is selective with respect to schooling. In both countries they examine the schooling attainment (in years) is higher for migrants than non-migrants. In the Palanpur data however, at least initially, it appears that non-migrants have higher educational attainment. Table 1.6 shows mean education for males between 15-24 in an initial survey by migration status

of these individuals during a later survey. As shown above there is an increasing trend in educational attainment irrespective of migration status. However, this increasing trend seems to be much stronger for migrants than non-migrants. The result of this is that despite non-migrants having higher educational attainment in earlier surveys, by the most recent surveys it is migrants who have marginally higher education levels. This is interesting as it suggests the nature of migration out of the village may be changing during the study period. However, these results should be interpreted with caution since Wilcoxon and Fisher exact tests fail to reject a difference in the means and medians of these two groups for each pair of surveys considered.

Table 1.6.: Education by Migration Status

	1958-1975	1964-1984	1975-1993	1993-2009
Migrants	1.13	1.33	2.50	2.67
Non-migrants	1.93	2.20	2.47	2.56

Note: Includes males present in Palanpur who are aged 15-24 in the earlier of the two surveys examined. Migrants are those who have left by the later survey, non-migrants are those who are still present in the village. Education is measured in years completed, collapsed into 5 categories (Illiterate, Class 1-5, Class 6-8, Class 9-12, Above Class 12). These categories are given numerical values in order of attainment (e.g. Illiterate is 1). The table reports mean education by migration status.

Source: Author's Calculations

1.7. Summary

Palanpur is not very different to thousands of other Indian villages. The major difference has been the presence of investigators intermittently over 60 years which has provided us with a unique opportunity to study the development of India. From a young post-

colonial democracy to the rapidly expanding service-led economy of the present day.

This chapter has demonstrated how Palanpur has evolved through this process, and how it is not a peculiar village in any obvious way. We have shown that the changes in employment, education and migration patterns are broadly consistent with data from other sources for India. With this in mind we now progress to using these data to tackle two questions of central importance to our understanding of economic development which can only be answered with data of this length and detail.

Chapter 2 examines the process of structural transformation within Palanpur, how this compares to sectoral reallocation as documented in the literature at a national level, and what this can tell us about structural transformation in development more generally. In chapter 3 we explore the importance of networks within the village labour market in light of this movement away from agriculture, and the long run effects of these networks for their members.

References from Chapter 1

Ansari, N., 1964. Palanpur: a study of its economic resources and economic activities. *Continuous Village Survey*, (41).

Baland, J.-M., Drèze, J. and Leruth, L., 1999. Daily wages and piece rates in agrarian economies. *Journal of Development Economics*, 59(2), pp.445–461.

Bliss, C.J. and Stern, N.H., 1982. *Palanpur: the economy of an indian village*. Oxford University Press.

Bliss, C. and Stern, N., 1978a. Productivity, wages and nutrition: part i: the theory. *Journal of Development Economics*, 5(4), pp.331–362.

Bliss, C. and Stern, N., 1978b. Productivity, wages and nutrition: part ii: some observations. *Journal of Development Economics*, 5(4), pp.363–398.

Bosworth, B., Collins, S.M. and Virmani, A., 2006. Sources of Growth in the Indian Economy. *India Policy Forum*, 3(1), pp.1–69.

Buera, F.J. and Kaboski, J.P., 2012a. Scale and the origins of structural change. *Journal of Economic Theory*, 147(2). Issue in honor of David Cass, pp.684–712.

Buera, F.J. and Kaboski, J.P., May 2012b. The rise of the service economy. *American Economic Review*, 102(6), pp.2540–69.

Feenstra, R.C., Inklaar, R. and Timmer, M.P., October 2015. The next generation of the penn world table. *American Economic Review*, 105(10), pp.3150–82.

Foster, A.D. and Rosenzweig, M.R., 2002. Household division and rural economic growth. English. *The Review of Economic Studies*, 69(4), pp. 839-869.

Foster, A.D. and Rosenzweig, M.R., 2004. Agricultural productivity growth, rural economic diversity, and economic reforms: india 1970-2000. English. *Economic Development and Cultural Change*, 52(3), pp. 509-542.

Foster, A.D. and Rosenzweig, M.R., 2008. Chapter 47 economic development and the decline of agricultural employment. In: Schultz, T.P. and Strauss, J.A. eds. Vol. 4, Handbook of Development Economics. Elsevier, pp.3051–3083.

Government of India, 2011. *Census of india 2011.*

Herrendorf, B., Rogerson, R. and Valentinyi, Á., 2014. Chapter 6 - growth and structural transformation. In: Aghion, P. and Durlauf, S.N. eds. *Handbook of economic growth.* Vol. 2, Handbook of Economic Growth. Elsevier, pp.855–941.

Himanshu, Bakshi, I. and Dufour, C., 2011a. *Poverty, inequality and mobility in palanpur: observations from palanpur*. LSE Asia Research Centre, (Working Paper 45).

Himanshu, Jha, P. and Rodgers, G., eds., 2016. *The changing village in india: insights from longitudinal research*. eng. First edition. Oxford University Press. ISBN: 9780199461868.

Himanshu, Joshi, B. and Lanjouw, P., June 2016. Non-farm diversification, inequality and mobility in palanpur. *Economic and Political Weekly*, 51(26-27), pp.43–51.

Himanshu, Lanjouw, P., Mukhopadhyay, A. and Murgai, R., 2011. *Non-farm diversification and rural poverty decline: a perspective from indian sample survey and village study data*. LSE Asia Research Centre, (Working Paper 44).

Himanshu, Lanjouw, P., Murgai, R. and Stern, N., September 2013. Nonfarm diversification, poverty, economic mobility, and income inequality: a case study in village india. *Agricultural Economics*, 44(4-5), pp.461–473.

Himanshu and Tyagi, A., 2011a. *Change and continuity: agriculture in palanpur*. LSE Asia Research Centre, (Working Paper 48).

Himanshu and Tyagi, A., 2011b. *Tenancy in palanpur*. LSE Asia Research Centre, (Working Paper 47).

Kattumuri, R. and Tiwari, D.K., 2011. *Developments in education in palanpur, a village in uttar pradesh*. LSE Asia Research Centre, (Working Paper 49).

Lanjouw, J.O. and Lanjouw, P., 2001. The rural non-farm sector: issues and evidence from developing countries. *Agricultural Economics*, 26(1), pp.1–23.

Lanjouw, J.O., 1999. Information and the operation of markets: tests based on a general equilibrium model of land leasing in india. *Journal of Development Economics*, 60(2), pp.497–527.

Lanjouw, P. and Rao, V., 2011. Revisiting between-group inequality measurement: an application to the dynamics of caste inequality in two indian villages. *World Development*, 39(2). Ethnicity and Ethnic Strife, pp.174–187.

Lanjouw, P. and Stern, N., June 1998. *Economic Development in Palanpur over Five Decades*, OUP Catalogue 9780198288329. Oxford University Press.

Mukherjee, A. and Ray, D., 1991. Wages and involuntary unemployment in the slack season of a village economy. *Journal of Development Economics*, 37(1), pp.227–264.

Sharma, N. and Drèze, J., 1996. Sharecropping in a north indian village. *The Journal of Development Studies*, 33(1), pp.1–39.

Sinha, D., 2011. *Nutrition status in palanpur*. LSE Asia Research Centre, (Working Paper 51).

Sinha, D. and Coppoletta, R., 2011. *The invisible half - women's status in palanpur*. LSE Asia Research Centre, (Working Paper 50).

United Nations, 2017. *World population prospects: the 2017 revision*. Online version.

Chapter A

Appendix to Chapter 1

A.1. Variable Definitions

continued overleaf

Table A.1.: Variable Definitions

Variable	Definition
Population	Individuals who are living in the village during the survey
Households	Households with any members living in the village during the survey
Total Land Cultivated	Land cultivated, in acres, by households who are present in the village during the respective survey
Land owned (acres) by HH	Land owned, in acres, by households who are present in the village
% Landless HHs	Percentage of households which own no land
% Land Leased	Land leased in by households present in the village as percentage of total land cultivated by village residents
% HHs in Non-Farm	Percentage of households which have at least one member present in the village who works in a primary or secondary occupation which is in the non-farm sector
% HHs in Own Cultivation	Percentage of households which have at least one member present in the village who cultivates land owned or leased by the household
% HHs in Agricultural Labour	Percentage of households which have at least one member present in the village who works for a wage in agriculture
Land Productivity	Real value of agricultural output (in 1960 Rupees) per acre of land cultivated
Labour Productivity	Real value of agricultural output (in 1960 Rupees) per agricultural worker (cultivator or labourer)
Real HH Income Per Capita	Real household income per household member present in the village
Income: % Non-Farm	Percentage of household income from non-farm sources
Mean Max Education	Maximum level of education achieved by any individual present in HH averaged across all households present in year
Mean HH Education	Mean level of education within HH, averaged across all households present in year
% HHs Educated Class 9+	The percentage of households present in the village which have at least one member who is educated at or above class 9
Education	Education is coded as 1-5. 1: Illiterate, 2: Able to read/write OR education between class 1 and 5, 3: Education between class 6 and 8, 4: Education between class 9 and 12, 5: Education beyond class 12

Note: Variable definitions for variables included in Table 1.1.

Chapter 2

The Fractal Nature of Structural Transformation

Structural transformation, the reallocation of economic activity from agriculture to manufacturing and services, is one of the most pervasive features of economic development. A vast literature exists on this phenomenon for countries, and there is a more limited microeconomic literature on the linkages between agricultural development and non-agricultural development. However, to our knowledge there is no evidence on how the nature and experience of structural transformation differs for entities smaller than nations. In this chapter we examine the experience of a village, a district, a state and a country in light of the stylised facts of structural transformation in the cross-country literature. We show that the pattern of sectoral reallocation is strikingly similar and consistent with the extant stylised facts at all levels. This, as yet undocumented, feature of structural transformation has important implications for sectoral complementarity and development policy.

2.1. Introduction

Structural transformation is one of the most widely studied phenomena in the history of Economics as an academic discipline. From the seminal contributions of Lewis (1954), the observations of Kuznets (1973), through to modern contributions such as Duarte and Restuccia (2010) - there innumerable papers concerning structural transformation in the macroeconomic literature. Equally many observers have noted the increasing importance of the non-agricultural sector in rural areas of developing countries, see Lanjouw and Lanjouw (2001) for a cross-country review. However, to our knowledge, no attempt has been made to reconcile these macro and micro findings explicitly, and the microeconomic foundations of the transition from an agricultural economy to an industrial economy are not especially well understood. Thus, in this chapter we ask how similar the experience of structural transformation is at different geographical and population levels. Does a country experience structural transformation in the same way as a village, a district or a state?

It seems to be commonly accepted in the economics literature that structural transformation and urbanisation are simply two sides of the same coin, and that movement out of agriculture must occur in unison with rural-urban migration. The occurrence of these two features together is obviously reasonable, and such coincidence is supported by a number of papers (e.g. Fajgelbaum and Redding (2014); Michaels et al. (2012); Young (2013)). However, rural-urban migration is not a necessary feature of development. The locus of economic activity need not change just because individuals get richer, even if the nature of the economic activity changes. We show an important example of a setting where development has not been accompanied by mass emigration. Individuals change sector of employment from agriculture into manufacturing or services but they

remain resident in a rural area.

In this thesis we define structural transformation as the reallocation of economic activity across agriculture, manufacturing and services. This structural transformation is characterised by Kuznets (1973) as "Major aspects of structural change include the shift away from agriculture to non-agricultural pursuits and, recently, away from industry to services; a change in the scale of productive units, and a related shift from personal enterprise to impersonal organization of economic firms, with a corresponding change in the occupational status of labor." We examine these features of structural transformation, in light of more recent cross-country work by Herrendorf et al. (2014); Buera and Kaboski (2012b); Buera and Kaboski (2012a), for the village of Palanpur, the district of Moradabad, the state of Uttar Pradesh and the country of India. We follow the macroeconomic literature on structural transformation by using manufacturing to refer to the sector of the economy which does not include agriculture or services.[1]

The focus of our study is India. Structural transformation in India has been documented by Bosworth et al. (2006) and Verma (2012), among others, who use a growth accounting framework to examine the contribution of broad sectors to overall growth in India GDP from 1960-2004. Their findings are consistent with the description of structural transformation by Kuznets above. They emphasise the service led growth of India from 1980 onwards, but Bosworth et al. (2006) argue that the role of services in the Indian growth experience is overstated. Bosworth and Collins (2008) compare the growth experience of India to that of China.

However, even if we know that at the country level there is a sectoral shift away from agriculture it does not necessarily follow that we will see comparable reallocations at all

[1]Strictly speaking it is probably more proper to refer to this sector as industry (as do the World Development Indicators, World Bank (Various years)) and consider manufacturing to be simply the largest part of this sector. However, the convention is long established.

levels of the economy. If agricultural and non-agricultural development are substitutes then it may be that smaller geographic areas specialise in the sector in which they have comparative advantage. We would then expect to see differential patterns in structural transformation for different snapshots of the economy. Foster and Rosenzweig (2004); Foster and Rosenzweig (2008) use a sample of households in villages in rural India who are surveyed three times in a 28 year period to explore the relationship between growth in agricultural technology and the non-farm sector. They find large increases in income from the non-farm sector in their villages, and that these increases are largest where agricultural productivity growth is smallest. This is consistent with the finding from Bustos et al. (2016) who show that the adoption of a land-augmenting technology in Brazil led to a reduction in the industrial sector, while in contrast the adoption of labour-saving agricultural technology led to growth in the industrial sector. Thus, if it is agricultural productivity growth which is the driving force behind this sectoral reallocation then the spatial distribution of sectors, and therefore the congruence of sub-samples of this changing economy, will depend on the factor-bias of the technical change[2].

Our examination of structural transformation at the village level uses data from Palanpur. The Palanpur data are uniquely suited to this purpose as they cover an extremely long time period, from 1957 to 2015. These 60 years span a period of rapid change and development within India. The other feature of the Palanpur data which allows us to explore structural transformation at the village level is that it was undertaken as a full census of the village. Usual panel data, particularly in rural areas of developing countries, face the dual challenges of (selective) attrition and household division. Most survey designs do not retain in their sample any or some sub-divisions of households which do

[2]See Gollin et al. (2002) andGollin et al. (2007) for models in which agricultural productivity growth is central to structural transformation.

not remain intact between survey rounds. These household divisions are not uncommon: in the ARIS-REDS panel data 30% of the original sample farm households had divided by the time of resampling 12 years later[3]. In the Palanpur data 64% of the original survey households have split at least once. Foster and Rosenzweig (2002) show using the ARIS-REDS panel data for a representative sample of rural India that this household division is non-random. It depends, for example, on land holdings, within household inequality in schooling, and marriages. Thus if one examines the nature of structural transformation using a typical survey which samples households within villages then this non-random nature of division, and therefore attrition, will lead to biased estimates of the relative sizes of sectors. The Palanpur data avoid this problem because all household splits are sampled in subsequent surveys as long as they remain in the village. Due to the nature of the survey we observe who migrates out of the village but do not observe their outcomes. However, this is the same as when we examine sectoral distributions at any level of aggregation - those who emigrate from India are no longer captured in the data for sectoral allocations at the country level. It is also true of most household surveys in developing countries which do not specifically focus on migration.

We combine our data on Palanpur with sector level data for Moradabad district, Uttar Pradesh and India. Armed with these data we compare the nature of structural transformation for each of these geographic areas with the stylised facts documented for cross-country data by Herrendorf et al. (2014).

We show that where the data are available these stylised facts hold true at all levels. The share of total employment which is agricultural has declined for India, Uttar Pradesh, Moradabad district and Palanpur. Agricultural productivity has increased. Income

[3]The Additional Rural Incomes Survey and Rural Economic and Demographic Survey conducted by the National Council of Applied Economic Research. See Foster and Rosenzweig (2002); Foster and Rosenzweig (2004) or the NCAER website for details of this panel: http://www.ncaer.org/data.php

per capita has risen. Economic activity has relocated from agriculture into both manufacturing and service sectors within Palanpur. For low levels of development this increase in the growth of manufacturing is faster than for services, and the employment share of services is bounded away from zero - both of these features are found in the cross-country data. We also find that the ratio of the manufacturing and service sectors evolves in a way which is strikingly similar to the cross-country averages as described in Buera and Kaboski (2012a). The manufacturing sector initially grows faster than the service sector, but the continuing expansion of the service sector eventually reduces the ratio - resulting in a hump-shape. The novel observation that this occurs even within a village economy seems to run counter to the prevailing perception in the literature which focuses on services as the principal component of non-farm employment in rural areas.

The results of this chapter are consistent with the agricultural technological change experienced by India, and its component parts, via the Green Revolution being labour-saving. Our findings also suggest that agricultural and non-agricultural development are complements. This complementarity is in contrast to the results of Foster and Rosenzweig (2004); Foster and Rosenzweig (2008) who find that agricultural and non-agricultural growth are substitutes. This complementarity means that development policies aimed at increasing agricultural productivity aid non-agricultural development which is important because non-agriculture is the more productive sector. This suggests that the large amounts of money spent on agricultural technological innovations and higher productivity crops have improved the incomes of many of the world's rural poor. However, our focus is a single village and we must learn how widespread the features of local development documented here are before we can draw strong policy conclusions.

This chapter contributes to the literature by forming a bridge between the microe-

conomic studies of the process of development, such as those reviewed in Foster and Rosenzweig (2008), and the vast macroeconomic and theoretical literatures which build upon the models of those such as Matsuyama (1992) to examine the nature of development at the national level. To our knowledge this is the first explicit examination of the nature of structural transformation at the level of a single village. The unique nature of the Palanpur data enables us to conduct this exercise.

We believe this chapter contributes to a nascent literature on the spatial dimension of the exit of rural workers from agriculture. Our description of India through the development experience suggests that there exist rural areas of developing countries where it is not a relocation of from rural to urban environments, but rather a reallocation of sectoral activity within rural areas. This illustrates the important distinction between urbanisation and structural transformation. While the two are clearly related concepts (Michaels et al. (2012)), and urbanisation may be the consequence of structural transformation if there are agglomeration externalities, they are distinct economic phenomena. In a context such as this it is therefore not necessary for there to be rural-urban migration together with reduced agricultural activity. It appears that the existence of adequate transport infrastructure and sufficiently local manufacturing or services allows development of the non-agricultural sector without large migrant flows. This is consistent with the large rural-urban wage gap and low mobility which are a feature of India, Chauvin et al. (2017); Munshi and Rosenzweig (2016).

Related to this is another strand of the literature which considers transport costs, infrastructure and development. Gollin and Rogerson (2014) calibrate a model to answer policy questions about the impact of transport costs on structural transformation.[4] Their

[4]They model location and occupation type as perfect complements, and thus don't distinguish urbanisation and structural transformation, however the intuition of their model will still hold if we allow rural workers to live and work in different regions.

model suggests that the extent of structural transformation is driven by improvements in agricultural productivity and transportation productivity, and these productivity improvements are complementary.[5] While there have been improvements in the transportation infrastructure around Palanpur during the study period, mostly through improved road quality in surrounding areas, it has always had access to the railway network. This relatively good transportation infrastructure, and corresponding relatively low transport costs, may be an important explanatory factor in Palanpur's structural transformation experience in the face of increasing agricultural productivity. If there were more comparable data to the Palanpur data it would be interesting to see whether this fractal nature of structural transformation holds in all settings, or if there is heterogeneity by village transportation infrastructure (or travel time to nearest urban location). If this feature of fractal structural transformation is common only to those areas with sufficiently good transport infrastructure then this suggests that existing estimates of the impact of transport infrastructure (e.g. Donaldson (2018); Michaels (2008)) may be underestimates since they do not include the long run general equilibrium increases in local incomes which come through sectoral reallocation.

This result has implications for development policies. It is suggestive of a policy avenue to reduce rural poverty through reducing the urban-rural wage gap and overcome the low levels of migration in India compared with other countries of similar levels of development (Munshi and Rosenzweig (2016)) - increased transport infrastructure allows rural areas to benefit from more productive non-agricultural pursuits in urban centres, while allowing individuals to remain part of their traditional rural social insurance networks which constrain usual forms of migration.

[5] Other papers which examine the importance of transport costs in development include Atkin and Donaldson (2015); Herrendorf et al. (2012); Storeygard (2016).

The following section describes the stylised facts of structural transformation from the cross-country literature, and the data we use. We then examine how well these stylised facts fit the data for India (in section 2.3), Uttar Pradesh (in section 2.4), Moradabad District (in section 2.5), and Palanpur (in section 2.6). Section 2.7 discusses the similarities and differences in the development experience of each of the geographic areas. Finally section 2.8 concludes.

2.2. Stylised Facts and Data

It has long been recognised, starting with the seminal work of Lewis (1954), that structural transformation - the reallocation of economic activity across agriculture, manufacturing and services - is one of the most pervasive elements of economic development. Kuznets (1973) listed it as the third of six main features of modern economic growth. This transition away from agriculture is necessary because individual demand for agricultural goods is relatively price and income-inelastic as incomes increase, and because of the presence of land as a fixed factor in agricultural production which limits its capacity for labour. Herrendorf et al. (2014) present the stylised facts of structural transformation using cross-country data from a variety of sources and find that it is characterised by: the share of agriculture in total employment, value added, and final consumption expenditure falling with development; the share of services in total employment, value added, and final consumption expenditure rising with development; the share of manufacturing in total employment, value added, and final consumption expenditure is n-shaped (or hump-shaped) with development.

In addition to these broad sectoral shifts Herrendorf et al. (2014) show that for low levels of development the value added share in agriculture is considerably lower than

the employment share. In the service sector both the employment share and the nominal value added share are bounded away from zero even at very low levels of development; the lowest value added share of services is around 20% and the lowest employment share is around 10%. This is consistent with work by Broadberry et al. (2013) who show that even in 14thcentury Britain the share of male workers in the service sector was just below 20%[6].

Equally it is clear that, at the level of the world, agricultural productivity growth is essential for sustained economic development. Increases in agricultural productivity which are labour-saving lead to industrial growth (Bustos et al. (2016)). Additionally, the work of Buera and Kaboski (2012b) shows an acceleration in the rate of relative growth of the service sector in the later stages of the process of structural transformation. Buera and Kaboski (2012b) document this stylised fact for a panel of 31 countries, who use 9,200 in 1990 international dollars as their threshold, based on evidence from the United States. They argue that this acceleration in the share of services at higher incomes appears to be a common feature of structural transformation, and that it is driven by growth in skill-intensive services. This feature is consistent with the findings of Herrendorf et al. (2014), who show that this acceleration in the service sector share tends to coincide with the peak manufacturing share. In their data this threshold occurs at approximately a log GDP per capita of 9 in 1990 international dollars, which is 8,103 in 1990 international dollars.

Buera and Kaboski (2012a) describe a new stylised fact in their panel of countries, that for low levels of GDP per capita the manufacturing sector expands more quickly than does the service sector. They show the average ratio of manufacturing to services share in value added increases from below 0.2 to a peak of approximately 0.75 before declining

[6]For females this share was even higher.

again as the manufacturing share declines past the threshold described above.

2.2.1. Data

When trying to examine structural transformation across countries data comparability is one of the primary issues faced. This is no less complicated when trying to document features of this transformation across different geographical areas. Although since we are looking within a country we do not need to focus on accurate purchasing power parity adjusted estimates.

In the structural transformation literature employment shares are typically calculated using hours worked by sector. We cannot employ these measures due to data availability at lower levels of disaggregation. Hence, we focus on workers by sector and abstract from hours worked. This ignores possible changing patterns of hours worked by level of development, and may mean that our employment shares do not exactly match the percentage of labour inputs employed in each sector. This could be an issue if the hours worked in agriculture change differentially by area in consideration. However, we have no reason to believe this is happening within a single country.

Given concerns about the under-representation of female employment in the Palanpur data, and that this under-representation is probably most prevalent in the agricultural sector, we use the percentage of male workers working in agriculture from the total male population when examining employment shares.

To allow for comparability of our agricultural employment share numbers across levels of aggregation, we use data from the Census of India, a decadal census last conducted in 2011, from which we have obtained population, cultivator and agricultural labourer counts for India, Uttar Pradesh and Moradabad district. We sum those who

work as cultivators or agricultural labourers to generate the total number of workers in the agricultural sector, and population to construct a percentage share. As we use total population and not number of workers as is typical within the literature, the numbers we will find will be lower than one might expect. However, we take advantage of estimates of the labour force participation rate to show these numbers are comparable with those in the literature when we inflate them.

We use this same narrow definition of agricultural employment for the Palanpur data. We count only those who are employed in cultivating their own land, or who work as agricultural labourers. There exist at all levels a number of services, many of which relate to agriculture. One example of this type of service within Palanpur is tractor rental. So while the labour of the tractor owner is engaged in the supporting the agricultural sector and increasing its productivity, it is not counted in the figures for agricultural employment. However, in constructing our employment data in this way we allow for direct comparability across all levels of observation.

Within the cross-country literature both production and consumption measures of sectoral shares are used, value added shares and final consumption expenditure shares respectively[7]. The decision for us over which measure to use is made easy, since there are no data on final consumption expenditure shares for any of our levels of interest. Therefore, we use production measures of sectoral shares in what follows. Data on Indian sectoral shares of GDP come from the Indian National Accounts produced by the Indian Central Statistical Organisation (CSO) via Reserve Bank of India (2016-17), these are measured at current prices at factor cost.[8] For Uttar Pradesh we use estimates of state

[7] Herrendorf et al. (2014) show where these two different measures can give different quantitative results, particularly for the manufacturing sector.

[8] See Bosworth et al. (2006) for a detailed description on the construction of India's GDP and sectoral share estimates.

domestic product and the percentage of state domestic product by sector from Ozler et al. (1996) for 1961-1993 and data from the CSO from 1994 onwards. There are no GDP or income estimates available at the district level for Moradabad. At the village level the Palanpur survey measures income in 5 of the 7 surveys. Therefore we use percentage of income from cultivation as a measure of the relative size of the agricultural sector.

Agricultural yield data for India for 1961-2014 come from the World Bank (Various years) World Development Indicators[9]. Detailed information on crop yields in Palanpur are collected in the 1958, 1963, 1975, 1984 and 2009 surveys. For comparability with other data sources we use only the yields and cultivated area from the farming of wheat, Bajra (pearl millet) and Paddy - the cereal crops which are grown in the village - to calculate cereal yields in kilograms per hectare. Land productivity measures for cereal for Uttar Pradesh and Moradabad come from ICRISAT (2015), again in kilograms per hectare. These are available from 1966-2011.

One of the stylised facts described in Herrendorf et al. (2014) is the n-shaped trend in the share of manufacturing, both in terms of employment and income shares. Due to data constraints it is hard to examine these manufacturing shares at levels below that of a country. However, we have coded occupations for the Palanpur data as manufacturing or services. Thus we will use these employment shares to consider the existence of this hump-shape in manufacturing. We do not expect a pronounced hump since Palanpur is not yet at the level of development needed to reach the peak of manufacturing employment share, typically over a log GDP per capita of 9 (in 1990 international dollars)[10] in the cross-country data. A comparison of the price level between countries and villages is inherently difficult and unreliable. Bearing this in mind, log GDP per capita of 9 is

[9] using series AG.YLD.CREL.KG
[10] Herrendorf et al. (2014)

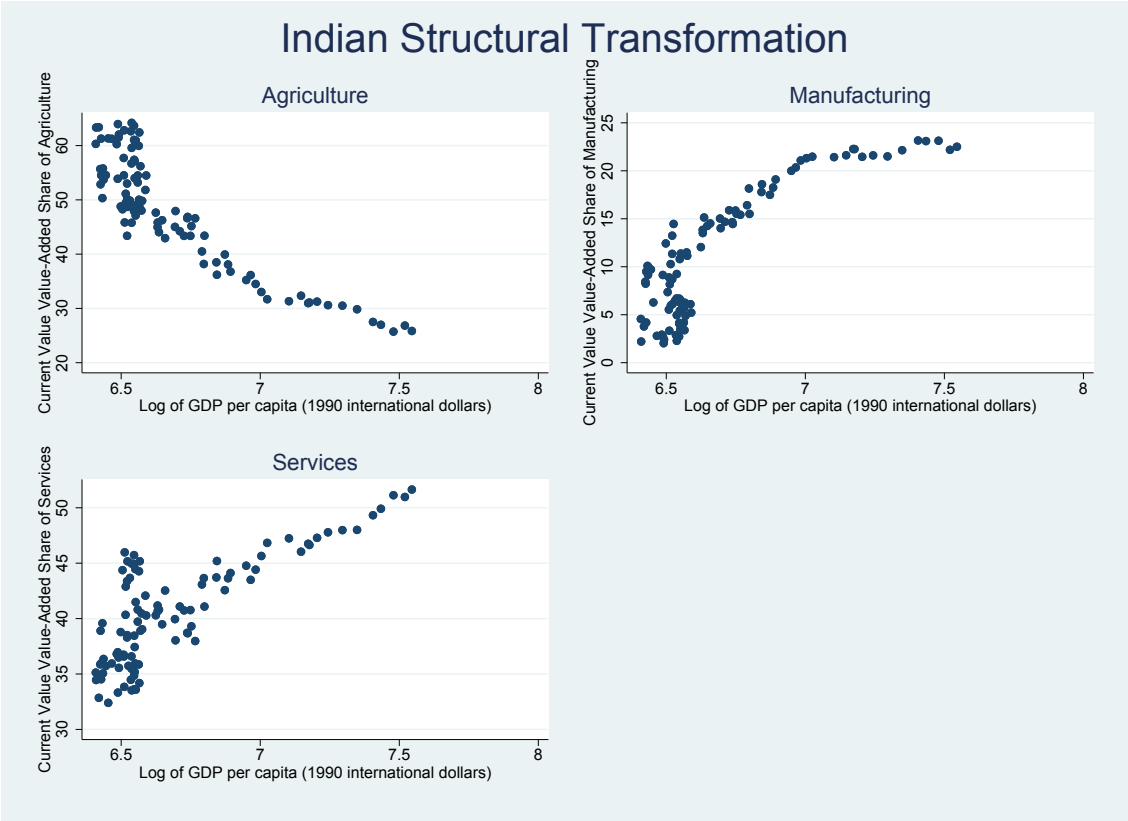
8,103 in 1990 international dollars, while the village average income in 2009 (the highest observed) is approximately 688 in 1990 international dollars[11], well below the observed threshold in the data. Thus, we should not see any decline in the manufacturing sector share even if the pattern is consistent with the cross-country data. In fact, in the Maddison 2018: Bolt et al. (2018) dataset India has yet to reach this level of GDP per capita by 2016.

2.3. India

We now turn our attention to the development experience of India. Due to the work of Sivasubramoniam (2000) there is long time-series of comparable data on sectoral shares available for India. These data are available for 1900-2000. Combining this with estimates of per capita GDP from Maddison (2010) we can see from Figure 2.1 that India's experience is in many ways typical of a country during the process of development. The value added share of agriculture in current prices is initially above 50% but as Indian per capita income grows this share falls to below half of its 1900 level. This activity moves out of agriculture into both manufacturing and service sectors. However, consistent with the cross-country stylised facts the manufacturing sector expands more quickly than the services sector for low levels of development. For the first 40 years of the sample the annualised growth rate of the manufacturing share is around 3%, while that of services is 0.7%. As India has developed this manufacturing sector growth rate has slowed, while the service sector has kept growing. We also observe, consistent with the cross-country evidence, that the service share is strongly bounded away from zero even at the lowest

[11] Using the World Bank International Comparison Program Database and the CPI for India from the WDI. 1 1990 International dollar is worth 5.417 1990 Indian Rupees at PPP. Consumer Price Index for India from WDI - 1960 = 2.58792717, 1990 = 22.8740634

Figure 2.1.: Sivasubramoniam (2000) GDP shares against GDP per capita



Note: Development Experience of India - 1900-2000
Source: GDP shares: Sivasubramoniam (2000). GDP per capita: Maddison (2010)

levels of development. In fact Broadberry et al. (2015) estimate a non-zero service sector share in 1600.

Next we examine the experience of India over the last 60 years, a timeline comparable with the Palanpur data. We find, using the World Bank World Development Indicators (WDI) and CSO data, that since 1960 India has been characterised by increasing population and agricultural productivity, falling agricultural land per capita and a smaller share of workers employed in agriculture. All while experiencing four fold increase in real GDP per capita. Figure 2.3 panel 3 shows this increase in GDP per capita (in constant

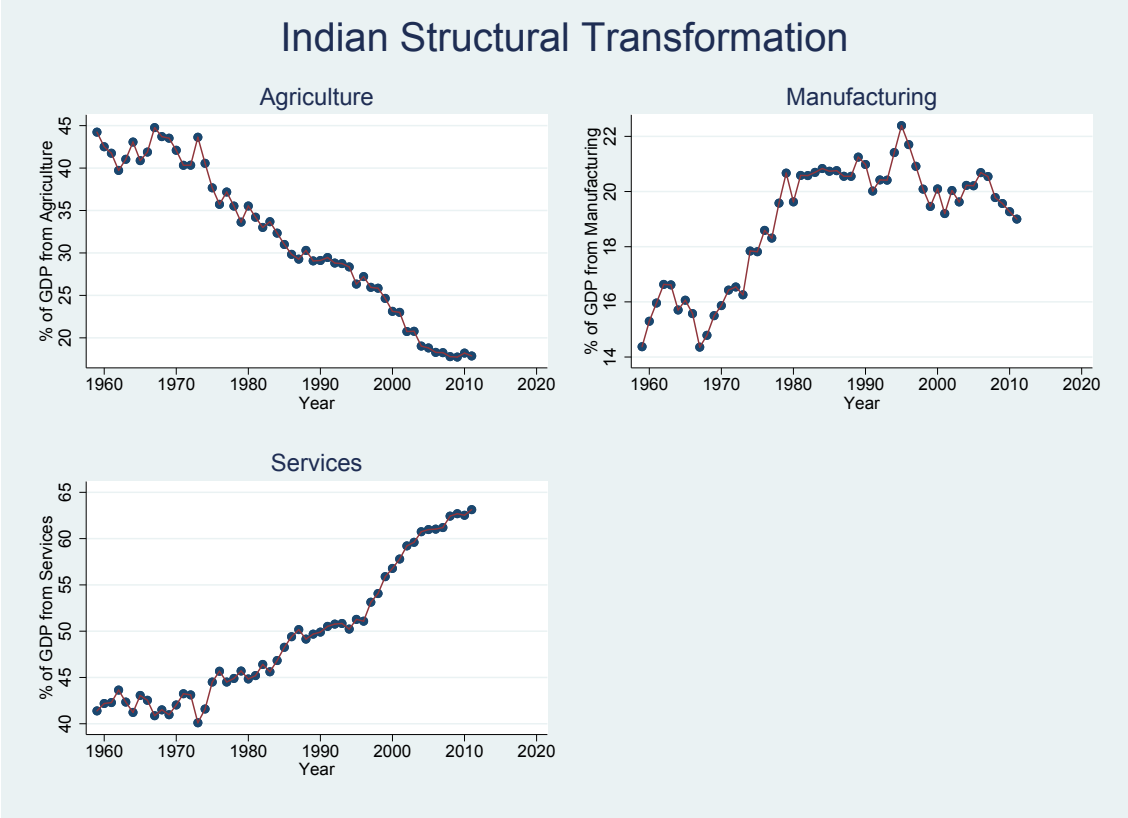
2009 Rupees). Over this period India has been one of the fastest growing economies in the world. Much has been written about the liberalisation undertaken by India beginning in 1991[12] leading to this increased growth, although graphically it seems that the increased growth rate predates these reforms. This is consistent with the findings of Aghion et al. (2008), who examine the dismantling of the license raj beginning in 1985, and Rodrik and Subramanian (2004), who argue that growth was triggered in the 1980s by a shift to a pro-business approach by the national government.

India's agricultural productivity has risen dramatically throughout the period, see Figure 2.4 Panel 4. This threefold increase in cereal yields (kilograms per hectare) clearly starts in the mid-1960s with the work of Norman Borlaug and the Green revolution, and has continued unabated with the introduction of further high-yielding varieties. Concurrently the agricultural value added per hectare has increased from 15,212 rupees to 77,263 rupees (constant 2009 rupees) between 1961-2010, an annualised growth rate of 3.37%.

Thus during this period we see India has increased incomes and agricultural productivity rapidly, so we would expect it to also be undergoing a shift of resources out of agriculture and into other sectors. Using data from Reserve Bank of India (2016-17) for 1959-2011, Figure 2.2 Panel 1 illustrates the reduction in income share (in current prices) from agricultural sources: India has experienced a 26 percentage point fall in agricultural sector share. Along with this reduction, Figure 2.2 shows the contemporaneous increase, by 22 percentage points, in the share of services in GDP. In this same period the share of manufacturing in GDP has risen almost 5 percentage points. However, these overall numbers disguise an apparent sharper increase which occurs roughly in the 1970s followed by a longer period where the share remains more constant, aside

[12]Reducing tariffs, reducing barriers to entry, increased privatisation and liberalising foreign investment.

Figure 2.2.: Sectoral Shares of Indian GDP



Note: Components of Gross Domestic Product (at Factor Cost) in Current Prices
Source: Reserve Bank of India (2016-17) Handbook of Economic Statistics

from a transitory peak in 1995. The impression one gets from the graph is that the share of manufacturing is falling at the end of the period, although it seems too short and small a decline to draw any strong conclusions. This pattern is consistent with the stylised fact of Herrendorf et al. (2014) who describe an n-shaped pattern for manufacturing during the process of structural transformation. Interestingly though they find in their sample of countries using multiple data sources that this period of peak manufacturing share tends to start when log GDP per capita in 1990 international dollars reaches approximately 9. In the case of India we find that this peak manufacturing share begins around 1979, with the maximum share occurring in 1995. India's per capita GDP in 1979 was 895.34 international dollars and 1,553.35 in 1995, from Maddison (2010). These are 6.8 and 7.3 log points respectively, far below the threshold found in cross-sectional data by Herrendorf et al. (2014).[13] However, consistent with their findings of a threshold we see that there is an acceleration in the growth in the service share around the time of the peak share in 1995. This stylised fact is also documented in Buera and Kaboski (2012b), who argue that this acceleration in the share of services at higher incomes appears to be a common feature of structural transformation, and that it is driven by growth in skill-intensive services. Whether this acceleration in India is also due to that is beyond the scope of this chapter[14]. Again however, this apparent threshold is well above what we observe here. The level of manufacturing share is also much lower than is typical for its peak (mean 40% in Buera and Kaboski (2012b) panel). Whether this is an indication that India is experiencing this transition to a primarily service based economy earlier than

[13]The graphs in Figure 2.2 are replicated, as in Herrendorf et al. (2014), with log GDP per capita in 1990 international dollars from Maddison (2010) on the x-axis, in Appendix Figure B.1. These show the same qualitative and quantitative pattern as those graphs with year on the x-axis, as one would expect when examining just one country. Although this exercise reduces the sample to end in 2008 due to the Maddison (2010) data.

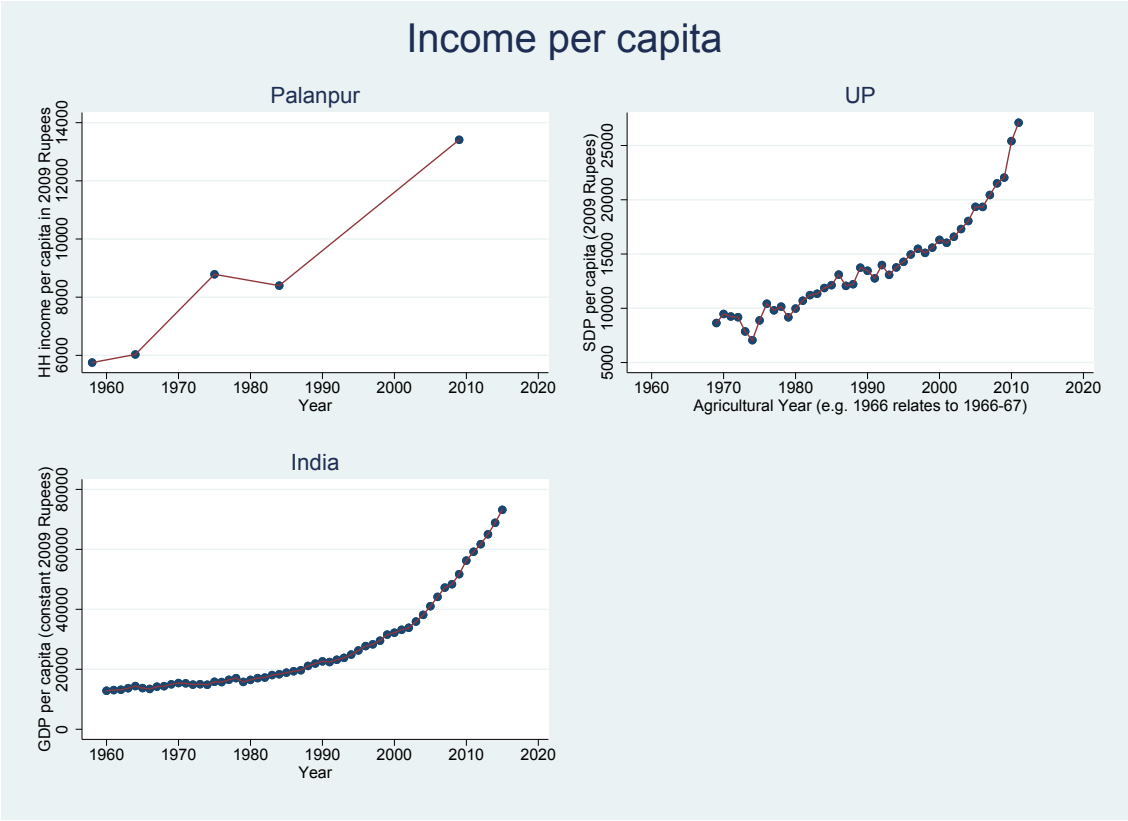
[14]For a discussion of the role of education in the Indian growth experience see Bosworth et al. (2006) for example.

other countries have previously or this is a transitory acceleration in the growth of the service share on the equilibrium path is worthy of further research.

Indian employment in agriculture data is only available since 1994 from the WDI, which shows that in this time it has fallen from 60.5% of total employment to 47.1% of total employment in 2012[15]. Instead we use data from the decadal Census of India starting in 1961 which contain counts of people cultivating their own land or working as agricultural labourers. We sum these counts and report them as a percentage total population. Figure 2.7 Panel 4 contains this data for India, interpolated log-linearly for graphical purposes. For comparability with other levels of disaggregation we look at this percentage for males only. From 1961 to 2011 the share of the male population employed in agriculture falls from 39.6% to 27.4%. For males and females combined these figures are 31.3% and 22.3% respectively. These numbers are naturally lower than those typically seen (including those from the WDI) because we are examining percentages with total population as the denominator, rather than employed population. Unfortunately it appears these data are simply not available in the length of time-series we need here. Using the average[16] employment to population ratio for those aged 15 or over from the ILO ILOSTAT database for India[17], 56.9%, to adjust our population shares for these different populations, we find that this estimate of agricultural share of total employment falls from 55.1% in 1961 to 39.4% in 2011. These numbers are much more in line with other estimates of this agricultural sector share for India. They are still an underestimate because the ILO employment ratio uses only those over 15 years old and our population includes all ages. Obviously this method of estimation ignores changes in the employment to population ratio so we do not have much confidence in the trend for these data.

[15]Similarly ILO estimates are available only from 1991, and show a fall from 62.7% to 44.3% in 2017.
[16]for all available years in our sample, 1994-2011
[17]WDI variable name: SL.EMP.TOTL.SP.ZS

Figure 2.3.: Income per Capita in Palanpur and Beyond



Note: Palanpur: Household income per capita in constant 2009 Rupees. Uttar Pradesh: SDP per capita in constant 2009 Rupees. India: GDP per capita in constant 2009 Rupees. GDP data not available for Moradabad District.

Source: Palanpur: Author's Calculations. India: World Bank (Various years). Uttar Pradesh: Ozler et al. (1996) and Indian Central Statistical Organisation, deflated using Uttar Pradesh CPIAL. UP population data interpolated from census years.

It is just reassuring that estimates using this technique are reasonably close to existing estimates, so we have confidence in that our census data is showing us the true trend in agricultural employment share.

The Indian development experience mirrors well the cross country evidence from Buera and Kaboski (2012a) of an initially more rapidly expanding manufacturing sector. The data from Sivasubramoniam (2000) demonstrate that the ratio of manufacturing to service share in current prices is approximately 0.06 in the early years of the 20th century. This ratio rises to a peak of 0.47 in the mid-1980s, and falls to 0.43 by the end of the sample in 2000. While data from Reserve Bank of India (2016-17) Handbook of Economic Statistics, from 1959-2011, show the clear hump-shape documented by Buera and Kaboski (2012a). Rising from 0.35 in 1959 to a peak of 0.46 in 1981, and subsequently dropping to 0.3 in 2011. It is possible to infer this hump-shape from examining Figure 2.2.

2.4. Uttar Pradesh

Now we aim to show that these stylised facts of structural transformation also hold for the State of Uttar Pradesh, where data are available. The Net State Domestic Product (SDP) Data used for Uttar Pradesh comes from Ozler et al. (1996) for 1960-1993 and the post-1993 data is collected from the India Central Statistical Organisation. We use net SDP because it allows us to extend the data back to 1960 for current prices. However, when we use the Consumer Prices Index for Agricultural Labour for Uttar Pradesh to convert the series to comparable 2009 constant rupees we can only start the series from 1969.

Firstly we examine a necessary condition for the process of structural transformation,

growing income per capita. We use net State Domestic Product (SDP) measured in constant 2009 rupees in Figure 2.3 Panel 2, which demonstrates a more than doubling of real per capita incomes in Uttar Pradesh between 1969 and 2011. It is more difficult to interpret because of the reduced time-series but it appears that there is an increase in per capita SDP growth in the late 1970s .

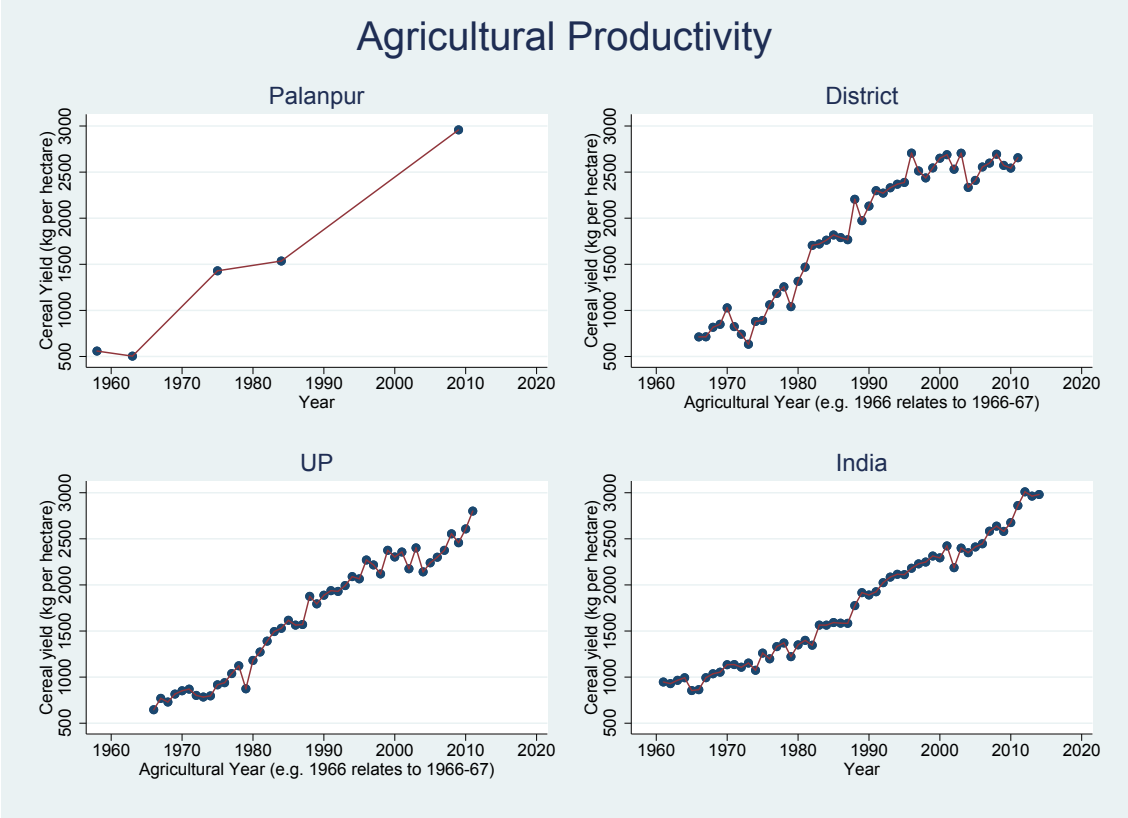
Figure 2.4 Panel 3 uses data from Ozler et al. (1996) and the CSO on state level agricultural yields to show that Uttar Pradesh has experienced large increases in agricultural productivity in line with those across the whole of India. Cereal yields have increased 333% between 1966 and 2011. It appears that relative to India as a whole the increase in agricultural productivity occurs later, in the mid-1970s rather then the late 1960s.

Also from Ozler et al. (1996) and the CSO are estimates of SDP for Uttar Pradesh by sector. Using these we construct a series of estimated percentage of current price SDP from Agriculture for 1960 to 2014. The agricultural sector share falls from 59% to 27% over this time period, as displayed in Figure 2.5 Panel 2. This decline in the GDP share of agriculture appears to begin in the first half of the 1970s and remain roughly constant until the decline seemingly abates around 2010.

Figure 2.7 Panel 3 demonstrates the relative size of the agricultural sector in Uttar Pradesh when measured by employment share. Employment type data in Uttar Pradesh comes from the ICRISAT (2015) data for Census of India years. Using these data, and interpolating log-linearly between decadal census, we find that agricultural employment of males as a percentage of total male population in Uttar Pradesh has fallen from 53% in 1961 to 28% in 2011.[18] Again, as with India, these numbers are an underestimate when considering the employed population as the denominator. If we use the same India-

[18]These same data for overall agricultural employment as a percentage of total population are 35.3% in 1961 to 19.43% in 2011

Figure 2.4.: Agricultural Productivity in Palanpur and Beyond



Note: Cereal Yields by year in Kilograms per hectare.

Source: Palanpur: Author's Calculations, Moradabad District: ICRISAT (2015), Uttar Pradesh: Ozler et al. (1996) and India Central Statistical Organisation, India: World Bank (Various years)

wide average employment to population ratio of over 15s in lieu of any reliable time-series estimates at state level, the male agricultural employment share decreases from 66% in 1961 to 35% in 2011.[19] Although these are still underestimates given we use the employment to population ratio for those over 15, these estimates for the employment share in Uttar Pradesh are higher than the agricultural share as measured by its share of GDP - consistent with the stylised facts of structural transformation.

2.5. Moradabad District

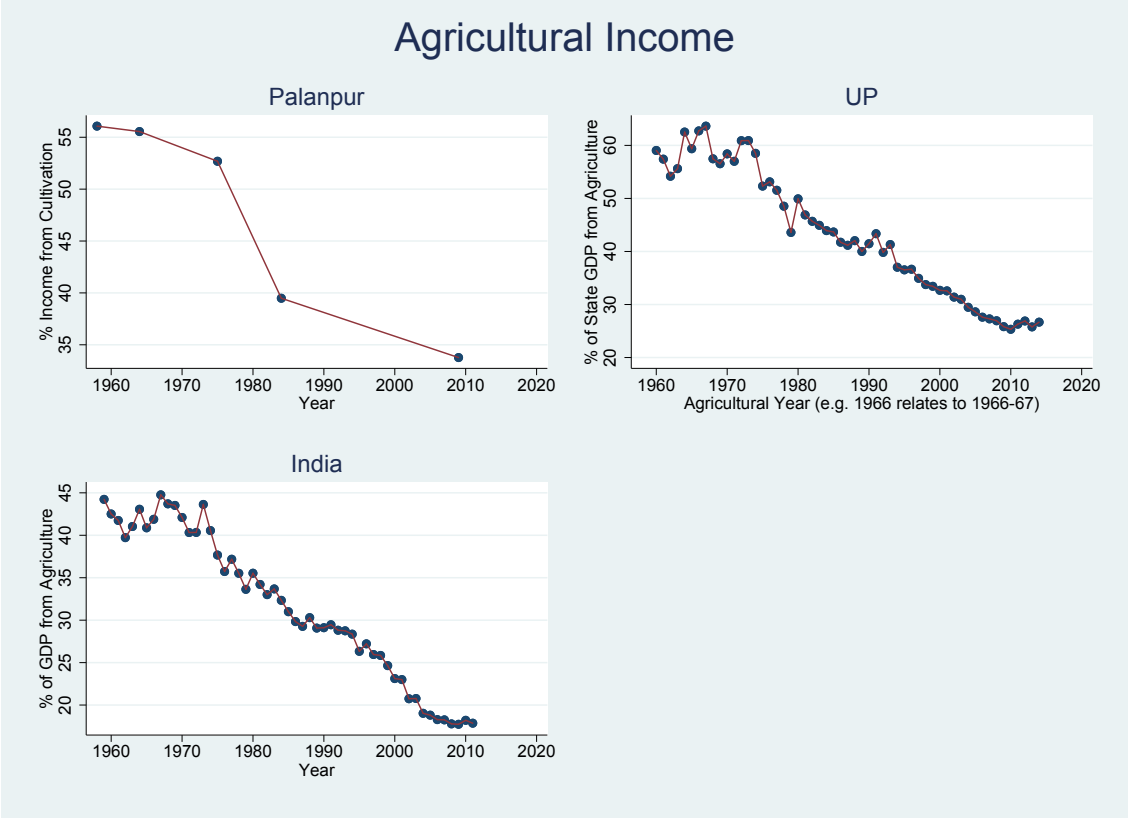
In this subsection we try to show as many of these stylised facts for Moradabad district as possible. Certain district level data is available from ICRISAT (2015). Unfortunately, GDP and GDP shares by sector are not available at the district level. However, we can show the reallocation of resources away from the agricultural sector in the district using employment shares. Employment type data in Moradabad District is available from ICRISAT (2015) for Census of India years. Using this we find that male agricultural employment as a percentage of total male population in Moradabad district has fallen to 25% in 2011 from 75% in 1961, see Figure 2.7 Panel 2 for graphical evidence. If we include male and female agricultural employment as a percentage of total population these numbers are 16% and 42% respectively. Just as for India and Uttar Pradesh above, because we use total population we underestimate the usual measure of the agricultural employment share. There are no estimates of the employed to population ratio for Moradabad district, so again we use the India level estimate as before. Thus, the estimated agricultural employment shares at the district level become 94% in 1961 and 32% in 2011 for males, and 74% in 1961 and 28% in 2011 overall. We make no claim for the accuracy of these es-

[19] Or 62% to 34% for the full population.

timates, they are merely indicative of the transition of the Moradabad district economy shifting from a state where agriculture is responsible for the vast majority of employment to one where it is significantly less important. However, due to data restrictions we cannot determine into which sectors these individuals move.

During this period the cereal yield in kilograms per hectare in Moradabad district has increased from 712.5 in 1966 (the earliest year for which the data are available) to 2,656.3 in 2010. This increase in agricultural productivity is shown in Figure 2.4 Panel 2. Once more this evidence is consistent with the argument of Bustos et al. (2016) who show that increasing agricultural productivity is labour-saving and causes reallocation of labour away from agriculture.

Figure 2.5.: Agricultural Income in Palanpur and Beyond



Note: Palanpur: Percentage of income from cultivation by year (current prices),
India: Agriculture as a percentage of GDP (current prices),
Uttar Pradesh: Agriculture as a percentage of state GDP (current prices).
District level income not available.

Source: Palanpur: Author's Calculations, India: Reserve Bank of India (2016-17),
Uttar Pradesh: Ozler et al. (1996) and India Central Statistical Organisation

2.6. Palanpur

Now we narrow our focus even further to the level of a village, Palanpur, and explore how many of these stylised facts of structural transformation for countries map onto the experience of a village. The development of Palanpur is clear from Figure 2.3 Panel 1, which shows the increasing trend in real household income per capita in constant 2009 rupees. Despite this 133% increase in real per capita incomes through the study period[20], these villagers are relatively poor. Although not directly comparable, their income level in 2009 is below the India level per capita GDP of the early 1960s.

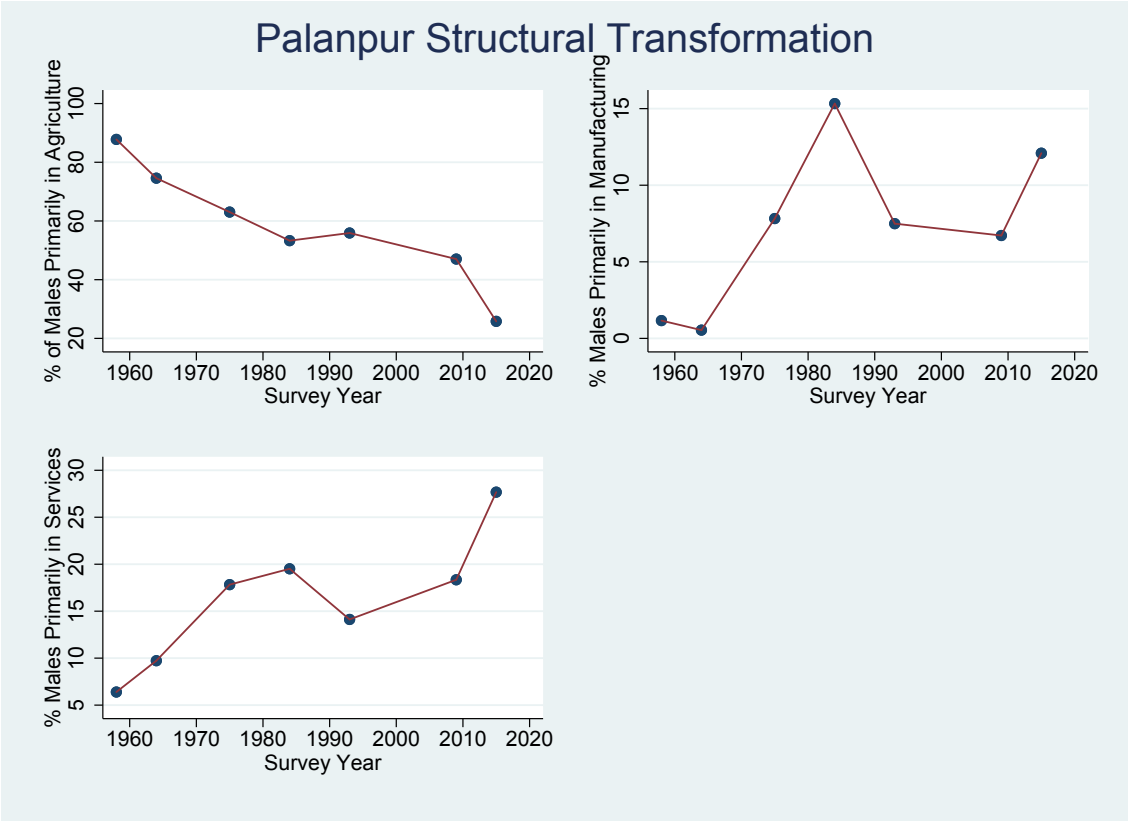
Contemporaneous to this increase in real incomes Figure 2.4 panel 1 shows the increasing agricultural productivity in Palanpur over the survey period. Cereal yield in kilograms per hectare increased from 559 in 1958 (low of 504 in 1963) to 2,958 in 2009. In monetary terms agricultural productivity increased from 120 rupees per acre in 1958 to 445 in 2009, in constant 2009 rupees. Due to the nature of the data it is impossible to determine exactly when this increasing trend in agricultural productivity began. All we can say for certain is that it began some time between 1964 and 1975 since the yields roughly tripled between these surveys. Figure 2.5 panel 1 shows that while agricultural productivity has increased the percentage of income from cultivation in Palanpur has decreased, from 56.1% in 1958 to 33.7% in 2009.

Employment data for Palanpur residents was collected in all surveys. Thus we use this measure of economic activity to describe the experience of structural transformation within Palanpur. Figure 2.6 shows the percentage of males who are over 15 and present in the village employed by sector within each survey.[21] These panels use the

[20] Rs5,753 in 1958, Rs13,424 in 2009

[21] These percentages do not add up to 100 since these individuals may be engaged in study or unemployed. However, unemployment is a nebulous term in this setting and is not consistently coded so we use the

Figure 2.6.: Employment in Palanpur by Sector



Note: Palanpur Structural Transformation using Employment by Sector. Included are all males 15 or over present in the village.

Source: Palanpur: Author's Calculations.

individual's primary occupation, that is the one in which they spend most time. The figure illustrates many of the stylised facts of structural transformation. The percentage of individuals employed in agriculture is strongly decreasing from 88% to 26% over the full study period. If instead we consider not just primary occupation but all employment this decline is less stark but still substantial - from 89% in 1958 to 50% in 2015. Comparing these initial employment rates with the percentage of income from cultivation in 1958 from Figure 2.5 it is evident that the labour allocation in agriculture is higher than its income share, by approximately 32 percentage points. While this definition of income from agriculture is narrow, even if we expand it to its fullest extent and include everything not considered part of the non-farm sector, as in Table 1.1, then the income share from agriculture (almost 87%) is still lower than the employment share. It also must be noted that this measure of agricultural income share contains many occupations which we consider to be services in our employment shares, for example tractor rental. As such, we consider these figures as consistent with the cross-country stylised fact that the agricultural income share is considerably lower than the agricultural employment share. That is, the village labour resources are engaged in the least productive sector.

Figure 2.6 also illustrates two other features of structural transformation, a rising share of employment in both services and manufacturing. The share of services in employment rises more consistently than that of manufacturing throughout the study period. The share of services in employment is also bounded away from zero even at the beginning of the period when Palanpur is poorest. In fact, if one considers not just primary occupation the share of working age males who are employed in services rises to 11% - rather than 6.4% for primary occupation. Once again this is consistent with the cross-country

whole working age population as the denominator.

stylised facts who find a lower bound of around 10% for the service sector employment share.

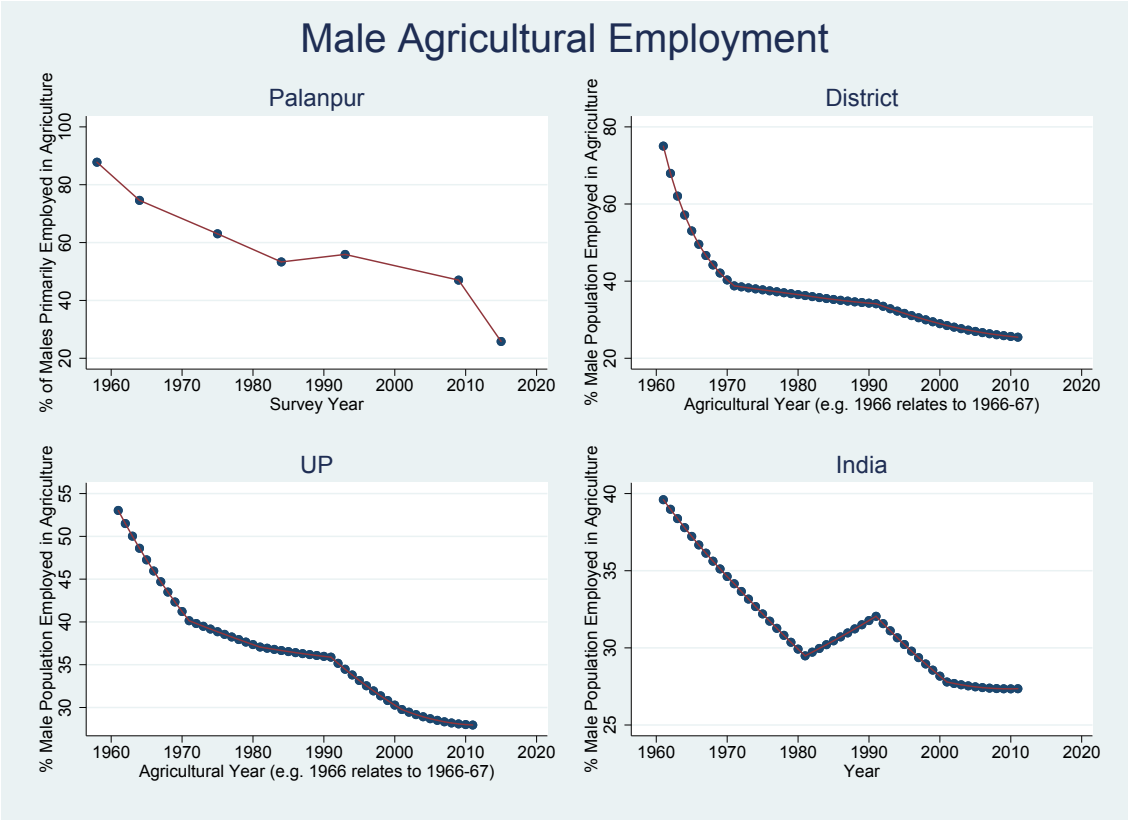
The share of employment in manufacturing expands more quickly than does the share of employment in services initially. The ratio of employment in these sectors increases from 0.06 in 1964 (0.18 in 1958) to a peak of 0.79 in 1984. These ratios and their pattern are eerily similar to the averages from a panel of 31 countries found in Buera and Kaboski (2012a). In their work, they describe the ratio of manufacturing to services using value-added shares in current prices rather than the employment shares we use here. They see initial average ratios of approximately 0.1 rising to a peak of 0.75 and subsequently falling to 0.4. Our respective figures for the ratios in Palanpur are 0.06, 0.79 and 0.44. However, the income per capita at which they occur is much lower. These ratios are obviously subject to a degree of noise given a relatively small sample size in Palanpur, but we believe this similarity is worthy of further investigation in other areas.

The other stylised fact for the manufacturing share in the cross-country data is the existence of a hump-shape through development. There is minimal evidence of a strong consistent decline in the manufacturing share of employment in the Palanpur data. However, this is also true of India over the same period. The pattern in Palanpur suggests more a strong increase followed by a more constant share - although there is a reasonable degree of noise. As noted previously, the typical threshold for this decrease in manufacturing share to be observed in the cross-country data is roughly 8000 in 1990 international dollars, and Palanpur income is well below this level so we do not expect to see a falling manufacturing share.

Returning now to the share of employment in the agricultural sector, we construct a measure of employment in agriculture which is directly comparable with those figures we have calculated from the India Census data for other areas. The percentage

of Palanpur members present who are employed in agriculture in some form (either as a cultivator or as an agricultural worker) has fallen from 30.8% to 19.0% from 1958 to 2015. Those who are employed in agriculture as their primary occupation has fallen from 30.5% to 10.6%. These numbers are much lower than may be expected in this context. We think it is highly likely that employment rates for women are underestimated in the Palanpur data. Especially in agriculture, where anecdotally many women help rear livestock and collect milk amongst other domestic work. As a result of this apparent under-reporting we focus on the employment patterns of males. For males only we find that the equivalent rates are from 57.3% in 1958 to 34.5% in 2015 for males who are employed in agriculture, and 56.6% in 1958 to 17.3% in 2015 for males who are employed in agriculture as their primary occupation, shown in 2.7. These figures are much more in line with the district and state figures.

Figure 2.7.: Employment in Agriculture in Palanpur and Beyond



Note: Palanpur: Percentage of males employed primarily in agriculture.
India, Uttar Pradesh and Moradabad District: Percentage of male population employed in agriculture. Data for India, UP and Moradabad are interpolated log-linearly from census data in 1961, 1971, 1981, 1991, 2001 and 2011.

Source: Palanpur: Author's Calculations.
India, Uttar Pradesh and Moradabad District: ICRISAT (2015)

2.7. Comparison between Palanpur, Moradabad District, UP and India

Despite India having a population of 1.2 billion in 2011[22], Uttar Pradesh a state of 200 million in 2011[23], Moradabad a district of 4.7 million in 2011[24] and Palanpur being a village of 1,299 people in 2009, and being vastly different entities, their development experience has been remarkably similar over the last 60 years. The three for which the data are available have experienced annualised growth in income per capita between 1.5% and 3% over the full period of data availability - between 40 and 51 years.[25] Along this development path we see almost identical growth in agricultural productivity for each entity, beginning the period with cereal yields around 500 kg per hectare and ending at approximately 3,000 kg per hectare. The rates and timing of growth also appear similar across each entity. The one exception is perhaps a slowing of the growth rate of district level productivity from the mid-1990s onwards. We consider these two elements - income per capita and agricultural productivity growth - to be necessary for structural transformation.

Data are available on the agricultural employment share for each level of observation, Figure 2.7. The trend in each of the series is decreasing. Focusing on just the share of agriculture in male employment we see, unsurprisingly, that India has the lowest initial level - 40% - since it has the highest level of income per capita throughout. The largest percentage point decrease, 50 percentage points, occurred at the district level. However, Palanpur experienced the largest percentage decrease of 69% - the others declined

[22] 1,210,854,977 in Government of India (2011)
[23] 199,812,341 in Government of India (2011)
[24] 4,772,006 in Government of India (2011)
[25] Palanpur: 1958-2009 - 1.74%. UP: 1969-2009 - 2.37%. India: 1960-2009 - 2.88%.

between 31 and 67 percent. Interestingly it appears that the decline in agricultural employment share was particularly rapid in Uttar Pradesh during the 1960s, and that this rapid decrease happened contemporaneously in Moradabad district and Palanpur. This drop to between 35 and 40 percent is followed in each case by a much more gradual downward trend in agricultural share in employment. The reason for this sharp decrease may be the Green Revolution which began in India in the mid-1960s with the introduction of new high-yielding varieties of wheat. This sharp trend break in 1971 is less apparent in the India data, and appears to occur in the 1980s. Perhaps this can be explained by other Indian states being slower to adopted these new, more productive, varieties of wheat and thus experiencing labour-saving technology related declines in agricultural employment more recently. There is certainly a large degree of variation in the increase in yields, and variation in when this increase began, by state. Foster and Rosenzweig (1996) show for 12 states ratios of yields per acre between 1961 and 1981 which vary from 1.1 to 2.8. Graphically the variation in timing of the increases is evident. Uttar Pradesh seems to be a relatively early adopter, among these 12 states, and is around the median productivity growth over the period.

Those geographic entities for which the data on the share of income from agriculture exist - Palanpur, Uttar Pradesh and India - all experienced between a 22 and 32 percentage point decline in the income share of the agricultural sector[26], Figure 2.5. In percentage terms this fall is largest for India since it starts from a lower base. What is noticeable is the similarity in trend for each of the series, initially a slow downward trend which increases quite sharply in the early to mid-1970s, followed by a flattening of the trend in the mid-2000s. This decrease occurs after the decrease in agricultural employment share through the 1960s and seemingly coincident with the increase in agricultural productiv-

[26]Palanpur: 22, Uttar Pradesh: 32, India 26 percentage points

ity. One limitation of the data used in this chapter is that the sectoral employment data is decadal, so it is impossible to be more specific about when this sharp decline in the agricultural employment share began.

Another feature of the data is that, despite our measure of the share of employment in agriculture being an underestimate of the more common measures, initially for Palanpur the share of employment in agriculture is higher than the share of income from agriculture. This is one of the stylised facts shown by Herrendorf et al. (2014) for low levels of development. The employment share levels for Uttar Pradesh and India are only 5 percentage points below the level of income shares at the beginning of the period. Additionally, when we use our estimate of the labour force participation rate to adjust these estimates to the usual measures of employment share we find estimates which are 7 and 11 percentage points higher than the income shares.

The ratio of manufacturing to services shares also behave similarly at all levels for which we have the data. The cross-country literature displays a clear hump-shape pattern for this ratio - the three values of this ratio which describe the beginning, peak and end of this hump-shape are 0.1, 0.75 and 0.4 on average. When examining only India this hump-shape is also clear, described by 0.06, 0.47 and 0.3 as beginning, peak and end respectively. We do not have shares of income by sector in Palanpur, so we use employment shares by sector. However, even when doing this we find similar ratios of manufacturing to services. The initial, peak and final ratios are 0.18, 0.79 and 0.44 respectively. For the manufacturing share of employment to rise initially faster than the service share of employment within a rural village runs counter to the impression one gets by reading the economic literature and popular press about rural India. There are many discussions about rural non-farm employment. However, the perception seems to be that this non-farm employment is typically in the service sector, and most commonly

it takes the form of services which support the agricultural sector. Possibly this perception is driven by the fact that this service provision takes place within a village itself, while work in the manufacturing sector typically takes place outside a village. Although obviously as an income source it can play an important role in the village economy.

Thus far all the stylised facts of structural transformation have held up at every level of disaggregation we have examined in this region of India. We believe this is an interesting and as yet undocumented feature of structural transformation. This apparent fractal nature of structural transformation means that we may be able to learn about macro features of sectoral reallocation by studying micro level data. We investigate one of these micro-level features of structural transformation, who are those transitioning out of agriculture and into other sectors and why, in Chapter 3.

The similarities of the structural transformation experience across levels of India is evidence in support of the complementarity of the agricultural and non-agricultural sectors. All areas have experienced large increases in agricultural productivity but this does not seem to have led to crowding out of non-agricultural activity, as found in Foster and Rosenzweig (2004). The data seem more consistent with the labour-saving agricultural productivity increases freeing up labour resources to change sectors as described in Bustos et al. (2016).

2.8. Conclusion

A single village cannot possibly claim to be representative of a country of 1.2 billion people, a state of 200 million people or even a district of 5 million people[27]. However, the experience of Palanpur is remarkably similar to that of rural India, Uttar Pradesh

[27] Government of India (2011)

and of Moradabad district as a whole. We find, where data are available, that all levels demonstrate declines in both agricultural income share and employment share, and that for low levels of development the employment share is greater than the income share. The share of manufacturing in output is initially strongly increasing, but this growth slows so the share of manufacturing is roughly constant after this initial increase. The decline in this share which defines the hump-shape in the classic cross-country stylised facts is not strongly evident yet. This is likely because India, and the areas within it, have not yet achieved the manufacturing share or level of per capita GDP which are typically associated with the decline in manufacturing share[28]. However, there is some indication that this decline may be happening earlier for India than is characteristic in the cross-country evidence. This is supported by an acceleration in the growth rate of the service sector which is coincident with a small decline in the manufacturing sector - a feature which is evident in the cross country data of Buera and Kaboski (2012b). This aspect of the Indian experience is worthy of further investigation[29].

We also describe a striking commonality between a village economy and the cross-sectional country-level evidence from Buera and Kaboski (2012a) - the ratio of the manufacturing to service share displays a hump-shape of comparable size and magnitude. To our knowledge this feature of a village economy has not previously been uncovered. That the manufacturing sector within a village economy would grow faster initially than the service sector is antithetical to the common perception. However, we can document this for only one village and further work to establish if this is a widespread feature of local economies is required.

An analysis of structural transformation writ small such as this is only possible due

[28] The mean of peak manufacturing share in Buera and Kaboski (2012a) is 40%.

[29] Although it may be complicated by the fact that India have recently changed their method of calculating GDP and are yet to release comparable retrospective data.

to the nature of the Palanpur data. Data for small geographic areas within developing countries over the extremely long time periods needed to explore the process of development are rare. There are a number of panel datasets which cover decades in developing countries, for example the ARIS-REDS and ICRISAT in India, the KIDS in South Africa. However, none of these span the length of the Palanpur data and these are typically samples of households within areas. They cannot give an accurate representation of the full local economy for all points in time due to attrition and household division. The Palanpur data does not face these problems as it is a census of the village.

Palanpur is a village which is well-connected to other areas and larger urban conurbations. Much of the manufacturing and service sector employment by Palanpur residents actually takes place outside the village. We have no reason to believe that Palanpur is unique in this regard. In fact, Foster and Rosenzweig (2008) use a model in which a factory sector is external to the village economy when examining rural non-farm employment for their sample of villages. However, it is not clear how many of these features of structural transformation would hold in a more closed village economy. Although it seems reasonable to imagine that there would still be a sector of non-tradable services to bound the service sector away from zero at the very least.

Related to this we find that an increase in agricultural productivity combined with good pre-existing transport links have led to a reallocation of labour out of agriculture, consistent with the models of Gollin and Rogerson (2010); Gollin and Rogerson (2014). However, we document a feature which is not captured by their models, that of commuting - and thus lower rates of migration away from rural areas. Further work is needed to examine the necessary conditions for this rural structural transformation to occur - although the data requirements for examining this question would be extensive.

Our contribution is an explicit analysis of structural transformation for economies

lower than the level of countries which to our knowledge has not been performed previously in the literature. This attempt indicates avenues for future research. One possible avenue we wish to highlight would be the estimation of a model of structural transformation using village level data. How closely matched the estimated structural parameters from this exercise are to the evidence at the level of countries are would be extremely interesting.

References from Chapter 2

Aghion, P., Burgess, R., Redding, S.J. and Zilibotti, F., September 2008. The unequal effects of liberalization: evidence from dismantling the license raj in india. *American Economic Review*, 98(4), pp.1397–1412.

Atkin, D. and Donaldson, D., July 2015. *Who's getting globalized? the size and implications of intra-national trade costs*. National Bureau of Economic Research (Working Paper Series, Working Paper 21439).

Bosworth, B. and Collins, S.M., March 2008. Accounting for growth: comparing china and india. *Journal of Economic Perspectives*, 22(1), pp.45–66.

Bosworth, B., Collins, S.M. and Virmani, A., 2006. Sources of Growth in the Indian Economy. *India Policy Forum*, 3(1), pp.1–69.

Broadberry, S., Campbell, B.M. and Leeuwen, B. van, 2013. When did Britain industrialise? The sectoral distribution of the labour force and labour productivity in Britain, 1381-1851. *Explorations in Economic History*, 50(1), pp.16–27.

Broadberry, S., Custodis, J. and Gupta, B., 2015. India and the great divergence: an anglo-indian comparison of gdp per capita, 1600-1871. *Explorations in Economic History*, 55, pp.58–75.

Buera, F.J. and Kaboski, J.P., 2012a. Scale and the origins of structural change. *Journal of Economic Theory*, 147(2). Issue in honor of David Cass, pp.684–712.

Buera, F.J. and Kaboski, J.P., May 2012b. The rise of the service economy. *American Economic Review*, 102(6), pp.2540–69.

Bustos, P., Caprettini, B. and Ponticelli, J., June 2016. Agricultural productivity and structural transformation: evidence from brazil. *American Economic Review*, 106(6), pp.1320–65.

Chauvin, J.P., Glaeser, E., Ma, Y. and Tobio, K., March 2017. What is different about urbanization in rich and poor countries? cities in brazil, china, india and the united states. eng. *Journal of Urban Economics*, 98, pp.17–49.

Donaldson, D., April 2018. Railroads of the raj: estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5), pp.899–934.

Duarte, M. and Restuccia, D., 2010. The role of the structural transformation in aggregate productivity*. *The Quarterly Journal of Economics*, 125(1), pp.129–173.

Fajgelbaum, P. and Redding, S.J., June 2014. *External integration, structural transformation and economic development: evidence from argentina 1870-1914*. National Bureau of Economic Research (Working Paper Series, Working Paper 20217).

Foster, A.D. and Rosenzweig, M.R., 1996. Technical change and human-capital returns and investments: evidence from the green revolution. English. *The American Economic Review*, 86(4), pp. 931-953.

Foster, A.D. and Rosenzweig, M.R., 2002. Household division and rural economic growth. English. *The Review of Economic Studies*, 69(4), pp. 839-869.

Foster, A.D. and Rosenzweig, M.R., 2004. Agricultural productivity growth, rural economic diversity, and economic reforms: india 1970-2000. English. *Economic Development and Cultural Change*, 52(3), pp. 509-542.

Foster, A.D. and Rosenzweig, M.R., 2008. Chapter 47 economic development and the decline of agricultural employment. In: Schultz, T.P. and Strauss, J.A. eds. Vol. 4, Handbook of Development Economics. Elsevier, pp.3051–3083.

Gollin, D., Parente, S.L. and Rogerson, R., 2007. The food problem and the evolution of international income levels. *Journal of Monetary Economics*, 54(4), pp.1230–1255.

Gollin, D., Parente, S. and Rogerson, R., May 2002. The role of agriculture in development. *American Economic Review*, 92(2), pp.160–164.

Gollin, D. and Rogerson, R., April 2010. Agriculture, roads, and economic development in uganda. eng. *NBER Working Paper Series*, p.15863.

Gollin, D. and Rogerson, R., March 2014. Productivity, transport costs and subsistence agriculture. eng. *Journal of Development Economics*, 107, pp.38–48.

Government of India, 2011. *Census of india 2011.*

Herrendorf, B., Rogerson, R. and Valentinyi, Á., 2014. Chapter 6 - growth and structural transformation. In: Aghion, P. and Durlauf, S.N. eds. *Handbook of economic growth*. Vol. 2, Handbook of Economic Growth. Elsevier, pp.855–941.

Herrendorf, B., Schmitz Jr., J.A. and Teixeira, A., 2012. The role of transportation in u.s. economic development: 1840-1860. *International Economic Review*, 53(3), pp.693–716.

ICRISAT, 2015. *Meso level data for india: 1966-2011, collected and compiled under the project on village dynamics in south asia*.

Kuznets, S., 1973. Modern economic growth: findings and reflections. *The American Economic Review*, 63(3), pp.247–258.

Lanjouw, J.O. and Lanjouw, P., 2001. The rural non-farm sector: issues and evidence from developing countries. *Agricultural Economics*, 26(1), pp.1–23.

Lewis, W.A., 1954. Economic development with unlimited supplies of labour. *The Manchester School*, 22(2), pp.139–191.

Maddison 2018: Bolt, J., Inklaar, R., Jong, H. de and Zanden, J.L. van, 2018. *Rebasing 'maddison': new income comparisons and the shape of long-run economic development*. Maddison Project Working paper. version 2018.

Maddison, A., 2010. *Statistics on world population, gdp and per capita gdp, 1-2008 ad*.

Matsuyama, K., 1992. Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 58(2), pp.317–334.

Michaels, G., 2008. The effect of trade on the demand for skill: evidence from the interstate highway system. *Review of Economics & Statistics*, 90(4), pp.683–701.

Michaels, G., Rauch, F. and Redding, S.J., 2012. Urbanization and structural transformation *. *The Quarterly Journal of Economics*, 127(2), pp.535–586.

Munshi, K. and Rosenzweig, M., January 2016. Networks and misallocation: insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1), pp.46–98.

Ozler, B., Datt, G. and Ravallion, M., 1996. *A database on poverty and growth in india*. http://go.worldbank.org/WTTXWLQKI0: Policy Research Department, World Bank, (mimeo).

Reserve Bank of India, 2016-17. *Handbook of statistics on indian economy*. https://m.rbi.org.in//Scripts/publications.aspx. accessed 17 Feb 2018.

Rodrik, D. and Subramanian, A., March 2004. *From "hindu growth"to productivity surge: the mystery of the indian growth transition*. National Bureau of Economic Research (Working Paper Series, Working Paper 10376).

Sivasubramoniam, S., 2000. *National income of india in the twentieth century*. eng. Oxford: Oxford University Press. ISBN: 0195650506.

Storeygard, A., 2016. Farther on down the road: transport costs, trade and urban growth in sub-saharan africa. *The Review of Economic Studies*, 83(3), pp.1263–1295.

Verma, R., 2012. Can total factor productivity explain value added growth in services? *Journal of Development Economics*, 99(1), pp.163–177.

World Bank, Various years. *World development indicators*. data.worldbank.org/data-catalog/world-development-indicators. Washington, D.C. : World Bank, accessed 9 Feb 2017.

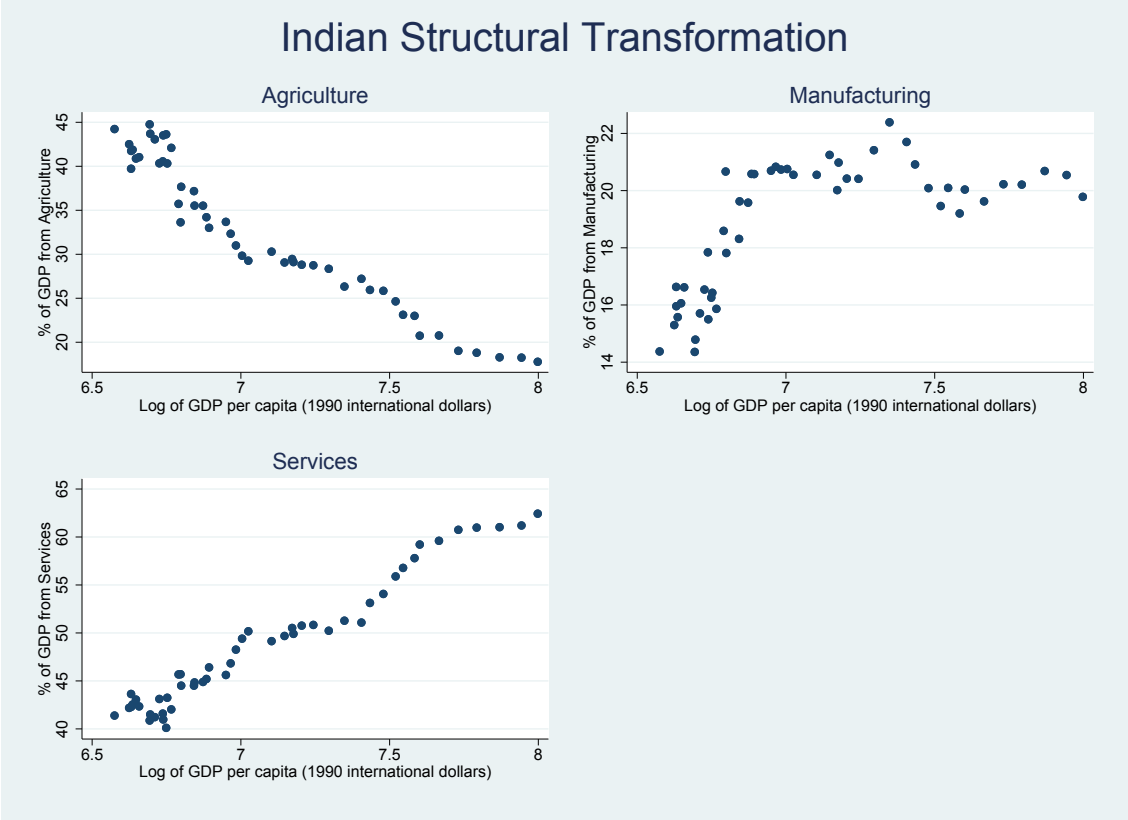
Young, A., 2013. Inequality, the urban-rural gap, and migration*. *The Quarterly Journal of Economics*, 128(4), pp.1727–1785.

Chapter B

Appendix to Chapter 2

B.1. Appendix Figures

Figure B.1.: GDP shares against GDP per capita



Note: Replication of Herrendorf et al. (2014) graphs using India data

Source: GDP shares: Reserve Bank of India (2016-17). GDP per capita: Maddison (2010)

Chapter 3

The Long Run Effects of Networks: Evidence from Palanpur

The presence of social networks within labour markets is well-established. However, the vast majority of empirical estimates of labour market network effects are contemporaneous or short run estimates. In contrast to this we examine the long run effects and dynamics of labour market networks. We show, for two different networks, short and long run effects on employment within a single village in India, Palanpur. Familial ties have stronger network effects in the labour market than do caste ties. The size of an agent's cohort is initially negatively associated with employment, but larger cohorts are beneficial for network members who arrive in the job market later. These networks have effects which span decades. We show that these networks may lead to greater inequality across networks through increased out-migration.

3.1. Introduction

Humans are, by their very nature, social animals. We form partnerships and groups in friendship, love, to achieve common goals, and have developed more than 7,000 languages[1] with which to communicate within these groups. Some of these groupings are large, e.g. nationalities, some of them are small, e.g. marriages. It is not surprising then that economists have found these ubiquitous social networks to be important in a wide variety of settings too numerous to list here. Within developing countries, where formal institutions are less common or less effective, these traditional institutions of social networks may be even more important.

In this chapter we focus on two of these social networks, caste and family networks, in a single village in India, Palanpur, for which we have rich data throughout a period of substantial change in the nature of employment and of rising income levels over a span of nearly 60 years. We explore the effects that these networks have on the distribution of employment within the village, both in the short and long run. Two types of employment are considered, one for which *a priori* we expect to see smaller network effects - known as *regular* jobs. Consistent with this prior, employment opportunities outside the village are more prone to job networks than these other regular jobs, and also that the networks for families are stronger than those for castes. We then extend the estimation of these short-run effects to examine the dynamics of these networks, exploiting information on the entire village population at all points in time within our study period to examine the effects of these networks in outside employment for cohorts within the village. We show that the dynamics of these networks for outside workers are consistent with a dynamic model from Beaman (2012) - there exists a competition effect within a cohort, and an

[1] Simons and Fennig (2018)

information effect of job referrals. We show again that the network effects are stronger for dynasty networks than caste networks.

In the final part of the chapter we present evidence that these dynasty-level networks have long run implications. We show that these networks lead to increased outmigration of individuals from the village. Given the evidence in the literature on the positive effects of migration on the migrant (for example Bryan et al. (2014) in Bangladesh find induced seasonal migrants experienced 30% increase in consumption), and on the inefficiently low migration rate in this setting (Munshi and Rosenzweig (2016)) these networks are likely to lead to increased inequality across networks in the long run.

The caste network is a feature unique to India. Genetic evidence suggests that it is a societal feature which began up to 4000 years ago, but it still has a great deal of practical and salient influence in present day India. Caste networks have been shown to be important for social insurance (Mazzocco and Saini (2012); Munshi and Rosenzweig (2016)), business activity (Munshi (2011)), schooling choices (Munshi and Rosenzweig (2006)), trade (Anderson (2011)), and resource misallocation (Banerjee and Munshi (2004)); amongst other things. Extended family networks are less widely studied in the economics literature, perhaps due to problems identifying them in typical household surveys. One paper which does show the importance of familial networks in a developing country is Angelucci et al. (2010), who show that a conditional cash transfer scheme only raises secondary school enrolment rates for those households which are part of an extended family network.[2] While Bandiera and Rasul (2006) show that technology adoption decisions are more correlated within self-reported family and friends groups than religion-based networks in Mozambique.

[2]Loury (2006) also discusses the effect of extended family members on educational outcomes for younger members.

The existence of networks within labour markets is well established by direct evidence from survey responses. Topa (2011) reviews the literature on labour market referrals, mainly from developed countries, and describes reported referral rates from surveys of job seekers - he concludes roughly half of all jobs are found through personal or informal contacts. In developing countries the same networks are reported to be present. For example Munshi and Rosenzweig (2006) find, for a sample of parents in Mumbai, that the percentage of males who received help from a relative or caste member in finding their first job was 68 percent of working-class men, and 44 percent of men in white-collar jobs.

As well as direct evidence from surveys the literature contains many studies which try to estimate these networks effects explicitly. Bayer et al. (2008) uncover neighbourhood referral effects for individuals living in the same city block: a one standard deviation increase in potential referrals raises hours worked per week by between 0.3 and 1.8 hours for men. Consistent with the survey evidence from Munshi and Rosenzweig (2006) above, they find the estimated referral effects are stronger for less educated and younger workers. Since residential locations are not randomly assigned they estimate these effects locally relative to nearby city blocks in a reference group. The critical assumption is that people choose residential locations by choosing a location within a group of blocks, and cannot choose a precise city block.

To avoid having to make these kinds of assumptions other authors have explored network effects in labour markets using natural experiments. Laschever (2013) examines how employment outcomes are affected by peers within a network of draftees into the US army during World War I. He finds, for employment in the 1930 census, that each additional employed network member increases a veteran's own likelihood of employment by 0.8 percentage point.

There are a number of papers which estimate employment network effects using migrant groups, usually to the US. Munshi (2003) studies Mexican migrants to the US, using rainfall in a migrants origin community as an instrument for the size of the destination network, he finds that migrants are more likely to be employed in a higher paying non-agricultural job when their network is larger. This effect is driven by established migrants[3] within the network, and new migrants[4] have an insignificantly positive effect. However, this weakly positive response of employment is at odds with the theory of Calvó-Armengol and Jackson (2004) and Beaman (2012), which argue that within network competition could reduce employment.

Beaman (2012) examines this empirically with data from the International Rescue Committee on refugees to the US arriving between 2001 and 2005. She finds that the number of network members who arrive in the current and previous year decrease the probability of employment after 90 days in the US for a new arrival, but that a larger number of network members arriving 2 to 4 years previously increases the employment probability of new arrivals. This is compelling evidence of the presence of competition and information effects in job networks for migrants. However, it is not clear whether migrants are representative of non-migrants in the degree to which they rely on networks for obtaining jobs. Our work speaks to this as we show network effects for a non-migrant population in a setting which is very different to the one faced by migrants to urban locations in the US. We also have the advantage of observing our network members over their lifetimes. The limitation of our setting is that we do not have a natural candidate for an instrument or natural experiment. We are limited to exploiting a variety of fixed effects to control for unobservables. Although one useful feature of our setting

[3]Community members located at the migrants destination for 4 or more years
[4]Community members at the destination for 1-3 years.

is that, unlike choices over which city to move to or which neighbourhood to live in, our networks of observation are exogenously defined by birth and immutable, given the feature of endogamous marriage within castes.

Another strand of the literature related to this chapter is that which concerns intergenerational mobility and (nuclear) family networks, beginning with Becker and Tomes (1979). Most similar to our work is that of Magruder (2010) who finds that fathers are a useful network member for sons in South Africa. He estimates that growth within a father's industry of 10 percent is associated with a 3-4 percent increase in the son's employment probability when they are network members[5]. However, while this effect is intergenerational from father to son we still consider them as short run network effects since he is exploiting heterogeneity within a 3 year panel.

A small number of papers focus on the potential long run impacts of networks. Munshi and Rosenzweig (2006) show that the caste network may lead to a dynamic inefficiency within schooling choices in Mumbai. They argue that caste occupation norms and caste referral networks led to lower-caste male network members missing out on the increasing returns to English language schooling which accompanied service sector growth in India during the 1990s. These schooling choices locked these males into traditional caste occupations and could lead to intergenerational occupational persistence. In stark contrast to this Munshi (2011) illustrates a setting where these same caste networks can assist in avoiding occupational traps, the Indian diamond industry. He shows, consistent with a dynamic model, that intergenerational occupational mobility is highest in those networks with the worse outside options.

The final part of this chapter is in the spirit of these papers on long run effects. We

[5]In his paper he exploits differences between co-located father-son pairs and geographically separated pairs.

show that the familial job networks in our setting induce members to out-migrate from the village. Given the large rural-urban wage differentials within India[6][7], and the potential income benefits to households of even temporary migration (Morten (forthcoming) finds that households increase their income by 30% in the years in which they send a temporary migrant to the a city), these inducements to migrate could have economically significant impacts on the income of familial networks and the distribution of income across these networks. Our evidence is also consistent with work by Akram et al. (2017) who find that a randomised experiment in Bangladesh which induces migration also induces others connected to migrants to migrate as well.

The following section briefly introduces some of the relevant details about Palanpur which are important for this chapter (Section 3.2), for a fuller description of the village see Chapter 1. Following that we describe the caste networks within Palanpur (Section 3.3) and their changing fortunes (Section 3.4), as well as the dynasty networks (Section 3.5) and the distribution of outside employment across these networks (Section 3.6). In Section 3.7 we describe the size and construction of our competition networks. The estimates of the short run network effects are in Section 3.8. Estimation of the network dynamics and evidence of migration effects are in Section 3.9 and Section 3.10 respectively. Finally, Section 3.11 concludes.

[6]Munshi and Rosenzweig (2016) show the urban wage is over 47 percent higher than the rural wage. Chauvin et al. (2017) find urban earnings are 122 percent higher than rural earnings in India

[7]Herrendorf and Schoellman (2018) find significant wage gaps between agricultural and non-agricultural sectors in a panel of 13 countries, including India. They also show using panel data that those who switch between agriculture and non-agriculture, even after accounting for human capital differences, in the United States, Brazil and Indonesia, earn higher wages.

3.2. Palanpur

In this section we describe some of the features of Palanpur which are pertinent to our study of networks. For a more detailed discussion of Palanpur see Section 1.2.

The evolution of Palanpur within the study period has been a declining importance of agriculture, in terms of both employment and income. In place of this agricultural work residents have moved into both services and manufacturing. Much of this employment takes place outside the village but this "outside employment" as we term it is not necessarily accompanied by migration. There is a train station in the village, which connects directly to Delhi in around 6 hours. Although much closer are Chandausi 13km away, a town of approximately 114,254; and Moradabad 31km away, the district capital with a population of 889,810.[8] Not every train stops at the station, but many do - enough to allow villagers the possibility to commute daily to Chandausi or Moradabad.

These outside jobs are not homogeneous, both in terms of location and type. Some are in manufacturing, some are services. However, these outside jobs are notably better paying than agricultural labour. Even for the most difficult manual labour types of these outside jobs - working in the Moradabad railway yard unloading trains - earn on average Rs. 200 per day in 2009, compared to the prevailing agricultural wage of Rs. 100 per day. This appears to be a fact which is not unique to Palanpur or more recent years. Himanshu et al. (2013) report evidence from India National Sample Survey (NSS) data that the ratio of mean casual non-farm to agricultural wages for 1983-2010 is between 1.3 and 1.5.

However, there is an important distinction to be drawn within these non-agricultural jobs between what are commonly known as "regular" jobs and other types of jobs.[9] Reg-

[8] Government of India (2011) - http://www.censusindia.gov.in/2011-prov-results/paper2/data_files/India2/Table_2_PR_Cities_1Lakh_and_Above.pdf

[9] For further discussion of employment for Palanpur residents see Section 1.4.

ular jobs are to a first approximation jobs which involve monthly rather than daily wage payments. These jobs are sought after since they offer numerous benefits. "Employees with such jobs usually enjoy a modicum of employment security, and, in most cases, substantially higher earnings than casual labourers." Lanjouw and Stern (1998):p.128. From the NSS data for the whole of India Himanshu et al. (2013) report that the ratio of mean regular to casual non-farm wages for 1983-2010 is between 2.0 and 2.4. Many, though not all of the regular jobs held by Palanpur residents in the study period are outside the village, approximately 70%. These jobs are more formal than the other jobs held by Palanpur residents. Thus, we believe they are less likely to be subject to the effects of job networks since they are more likely to have a formal hiring process. Although a formal hiring process clearly does not rule out job referrals, as illustrated by the evidence from the US.

The unique elements of the Palanpur data, aside from their longevity, are that the data are collected as a census of the village, and it is possible to track how each household or individual are related to each other, if at all. The data are taken as a series of snapshots of the village over 60 years and there is no systematic attempt to create a timeline of events which would enable the creation of a yearly panel from this data, since that was not the purpose of the original studies.

However, due to the census nature of the data we know not only the individuals present in the village who are enumerated in each survey. We also know about those members of Palanpur who migrate out of the village, even if they are not present in any survey. This is because households are asked about any members who are now not present when they were in previous surveys, and also about any births in the household in intervening years. Thus we know basic demographic information about everyone who was at one time a member of the village, although we do not have income or em-

ployment data for those migrants who have left permanently. As a result the sample of unique individuals who are present in at least one of the Palanpur surveys is 2,752, but the number of unique individuals in what we refer to as the migration sample is 2,831. This feature of the surveys allows us to construct reliable networks for those in the village at any point in their life.

Since we will be examining the networks of village members at different points in their life it is important to address one concern about the Palanpur data, the reliability of the age data for teenage females. "It is, we were told, quite common for people to lie about the ages of girls in the range 11-15 or even to conceal their existence." Bliss and Stern (1982):p.15. However, when examining the density of ages of females in the village we find no evidence for this. There is no hole in the distribution of age of females present in the village, which we would expect to see if females in the early teen range are consistently being hidden from enumerators or if their ages are being manipulated. Also, the distribution of female ages is very similar to the age distribution of males present in the village[10]. An examination of the cumulative distribution functions of age by sex further reassures that any differences in the distributions come from females outlasting males[11]. However, despite these reassurances we use only males in much of our analysis due to the persistent nature of social norms in this environment where women do not do work other than domestic work or caring for livestock, or other cultivation work.

There are two levels of social network which are salient for residents of Palanpur in their everyday lives. These are the members of their caste, and the members of their family. In the following sections we discuss the nature and experience of these networks within Palanpur.

[10] A Kolmogorov-Smirnov equality of distributions test gives a test statistic of 0.030, which means we fail to reject the null hypothesis of equality at the 10% significance level (p-value 0.107).
[11] The CDFs are shown in Appendix Figure C.1

3.3. Caste in Palanpur

Genetic evidence suggests that the process of endogamous marriage, a fundamental feature of the caste system in India, dates from between 1,900 and 4,200 years ago (Moorjani et al. (2013)). This feature of endogamy continues to persist, surveys suggest that almost 95 percent of Indian marriages continue to occur within *jati* (Munshi (2014)).[12] As well as endogamous marriages, there are other traditional features of castes - hierarchy, restrictions on inter-dining and inter-mingling.[13] In the most extreme cases a minimum physical distance is expected to be maintained between members of different *jatis*. Thus *jatis* are not just extended networks of individuals linked by marriage or blood, but also salient community networks in day-to-day interactions. In Indian villages it is typical for houses belonging to members of the same caste to be located in clusters. Srinivas (1976) describes how individuals in his study village would always ask a visitor about his *jati* since they "regarded that bit of information as essential in order to learn about his occupation, diet and life-style. How they behaved towards him also depended to some extent on his *jati*" *(page 164)*.

As a consequence of this history, and marriage practices, caste networks tend to be

[12]For a discussion about the relationship between caste, *varna* and *jati* see Srinivas (1962) and Béteille (1996). In this thesis the terms caste and *jati* are used interchangeably.

[13]For a detailed account of how these features affect caste relations in a rural Indian village see Srinivas (1976) chapter 6.

larger than those typical in other countries based on clan (Luke and Munshi (2006) in Kenya), relatives (Wang (2013) in China), or neighbourhood (Munshi (2003) in the United States), since they span many villages and consist of thousands of individuals. In this thesis however we focus only on the network of caste members who are resident in Palanpur at some point since 1958.

There are members of 16 castes present in Palanpur in at least one survey; of these, 11 castes are represented in every survey. In the majority of Palanpur publications to date 8 main (more populous) castes are considered: Thakur, Murao, Dhimar, Gadaria, Passi, Jatab, Dhobi, Teli.[14] The Dhobis and Telis are Muslims, however "The Muslims of Palanpur also divide themselves into hereditary occupation groups that have many of the basic contemporary features of Hindu castes, e.g. endogamy and hierarchy" (Lanjouw and Stern (1998):p.26). They identify first as Dhobis or Telis rather than as Muslims, in much the same way as the Thakurs identify first as Thakurs, not as Hindus. These features of Muslim society in Palanpur mean that we treat Dhobis and Telis as castes in the usual Hindu sense. Additional to these 8 main castes are 8 smaller castes which range between 1 and 23 individuals in each survey: Nai, Kayasth, Balmiki[15], Lohar, Khati, Brahmin, Badhai, Carpenter[16] - as well as 4 individuals whose caste is unknown. Table 3.1 shows the size of each caste in each survey, as well as the total number of individuals of each caste in the village over the span of the surveys. It shows that the Muraos, a traditional cultivating caste, have been the most populous caste throughout most of the survey period. However, the Thakurs, ancestrally a fighting caste, are of a similar size

[14] More properly these castes are in fact *jatis*, or subcastes, but for convenience we refer to them as castes throughout. These are the groups that villagers self-identify as when asked their caste, and are the caste distinctions which are important for interactions among individuals and households within the village.

[15] Known in previous Palanpur publications as either Bhangi or Harijan

[16] There is no record of the Hindi word for this caste in the original questionnaires.

to the Muraos in all periods, and are actually largest in 1984. Thakurs belong to the Kshatriya varna[17], only they and the Brahmin household present in the 1993 survey are from the "twice-born" castes. Other members of Palanpur are from the Sudra varna or the "untouchable" castes. Thus the Thakurs are the "highest" caste in the caste hierarchy within the village, excluding the Brahmin father and son in 1993. The lowest castes in the hierarchy are those belonging to the scheduled castes - the Jatabs, Balmikis and Passis. Although in Palanpur the Passis are endowed with better social status than either the Jatabs or Balmikis[18].

That there are no Brahmins in the village for the majority of the survey period is not uncommon for villages of this size in Uttar Pradesh. As Bliss and Stern (1982) note: "When the need is felt for a Brahmin to perform a religious ceremony he comes in from a nearby village." Anderson (2011) discusses the origins of village caste composition in Uttar Pradesh and Bihar, and the phenomenon by which villages in this region come to have no upper castes present. In their data they find that 42 percent of villages contain no upper caste individuals.

Given the importance of caste networks in Indian life it is not surprising that these networks have been found to be the level at which networks are drawn to serve a variety of functions. There are many papers which illustrate the importance of caste networks in individual outcomes. Castes have been shown to be important for social insurance (Mazzocco and Saini (2012); Munshi and Rosenzweig (2016)), business activity (Munshi (2011)), schooling choices (Munshi and Rosenzweig (2006)), trade (Anderson (2011)), and resource misallocation (Banerjee and Munshi (2004)); amongst other things.

[17] The 4 ancient varna divisions of Hindu society in order of ritual status are Brahmin, Kshatriya, Vaishya and Sudra.

[18] "having come to this village from far off they do not seem to have imported with them their caste status and enjoy a definitely better position than they had in their ancestral homeland"Ansari (1964)

Table 3.1.: Caste Sizes by Year

	1958	1964	1975	1984	1993	2009	2015	Unique Present	Unique Total
Thakur	104	125	172	218	279	287	302	599	627
Murao	119	132	183	218	290	302	345	670	675
Dhimar	56	52	59	75	86	112	106	236	244
Gadaria	42	45	68	83	88	91	89	188	188
Dhobi	6	2	22	30	31	46	59	94	94
Teli	47	57	71	93	109	143	125	264	277
Passi	61	75	77	85	60	19	27	176	189
Jatab	71	71	97	119	133	198	201	384	393
Lohar	1	0	0	0	0	0	0	1	1
Nai	10	7	9	4	6	8	11	24	24
Badhai	0	0	8	9	10	12	23	29	32
Kayasth	8	7	12	14	8	5	5	20	20
Khati	5	7	0	0	0	0	0	12	12
Carpenter	0	0	3	0	0	0	0	3	3
Balmiki	5	3	12	20	19	19	6	46	46
Brahmin	0	0	0	0	2	0	0	2	2
Unknown	0	0	0	0	0	4	0	4	4
Total	535	583	793	968	1121	1246	1299	2752	2831

Note: Individuals present by caste for each survey year. Unique present is the number of unique individuals who are present in the village in at least one year. Unique Total is the number of unique individuals who are named as part of a household in at least one year, even if absent.

Source: Author's Calculations

3.4. Patterns of Change for Castes

Unsurprisingly, given the importance of caste in these economic decisions, castes have not evolved equally in Palanpur. This can be seen in Table 3.2, which shows mean real per capita incomes for selected castes (those present in all survey rounds for which income data were collected).[19] Income data were not collected for the 1993 and 2015 survey rounds. Incomes are collected at the household level and then divided by the number of members present in the household, enabling them to be reported as per capita annual incomes in 1960-61 rupees.[20] These income numbers are designed to be comparable across survey rounds and have been constructed using the same methodology.[21] However, there are caveats associated with the use of reported income data from surveys - especially when much of the income is from home production.[22] There is also unavoidably a degree of noise which comes with surveying incomes only 5 times in 50 years, thus we avoid focusing on small or short trends, or inferring too much from exact magnitudes. Obviously these incomes can only be observed for those individuals and households which remain in the village, and must therefore be viewed with this in mind. We consider migration decisions and patterns in Section 3.10.

Table 3.2 contains the real per capita income by year of each caste which is present in all years of the survey. We see from Table 3.2 that some castes have become poorer on average relative to the village as a whole over the 50 years examined, e.g. Dhobis and Passis. While others have become relatively more prosperous on average, e.g. Kayasth.

[19] Incomes for all castes are included in Appendix Table C.1.

[20] The income from agricultural labour, non-agricultural labour, regular employment, self-employment, jajmani, remittances, renting, farming, non-farm and livestock sources are all included. Nominal incomes are deflated using the Consumer Price Index for Agricultural Labourers (GoI) for Uttar Pradesh.

[21] For details on this see Gaurav (2013) and Surendra (2013), which are part of the Palanpur data documentation.

[22] A discussion of these issues as they related to the Palanpur data is in Himanshu, Bakshi et al. (2011b)

Table 3.2.: Mean Real Income Per Capita by Caste-Year

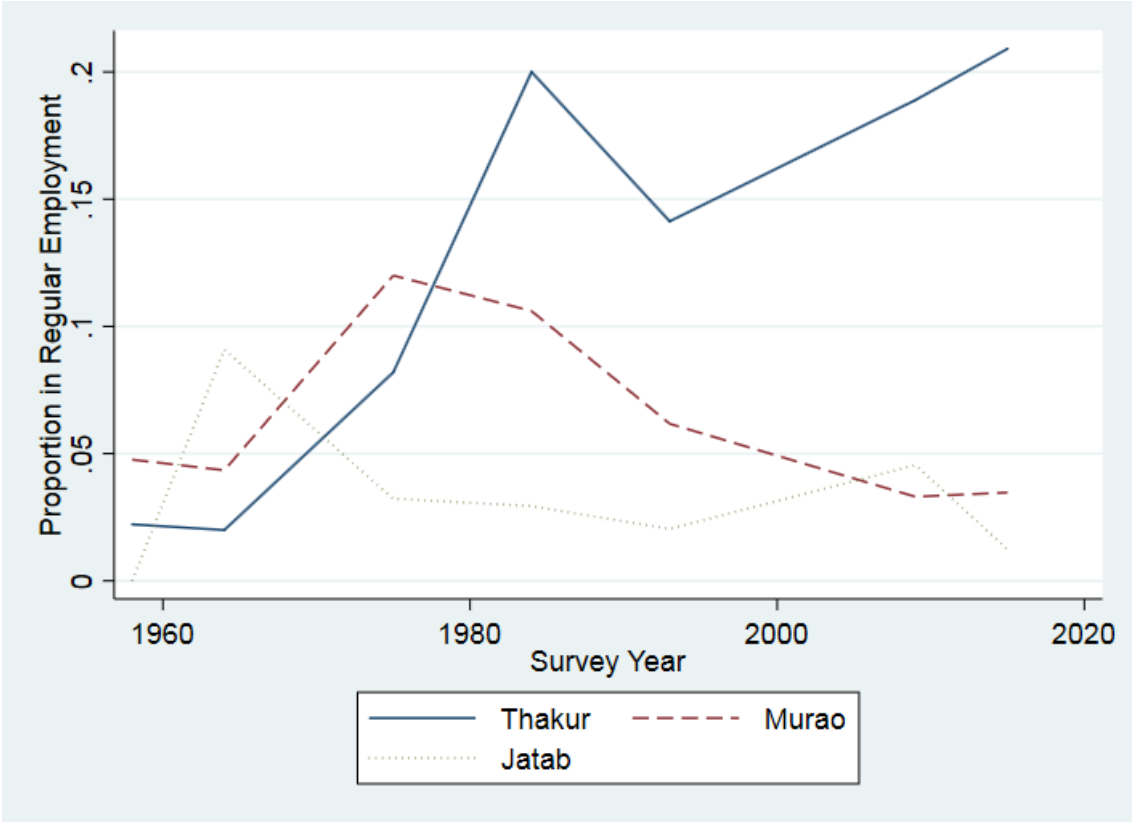
	1958	1964	1975	1984	2009	Total
Thakur	203.4	245.9	326.5	266.3	474.3	333.6
Murao	280.2	252.7	332.8	330.8	456.1	353.7
Dhimar	121.3	97.71	221.6	210.7	343.0	223.7
Gadaria	195.5	228.4	278.2	220.7	502.8	308.4
Dhobi	249.1	802.0	157.3	196.0	162.4	187.8
Teli	112.2	114.3	237.8	228.6	538.8	308.9
Passi	191.8	260.5	263.8	228.3	315.5	242.9
Jatab	152.9	129.2	208.5	117.6	298.3	203.8
Kayasth	167.1	233.2	251.3	290.5	544.0	277.6
Village Average	190.9	204.5	278.2	241.9	421.2	291.4

Note: Includes only those castes which were present in all surveys when income data were collected. Values are mean per capita annual incomes of households, they are in 1960-61 Rupees. Income is not available for the surveys in 1993 and 2015. Totals and Village Averages are weighted by individuals present. Village Averages include all castes.

Source: Author's Calculations

Interestingly for us, examining the two largest castes by population, the income of Thakurs has caught up to (and even overtaken in the most recent data) that of the Muraos. Consistent with this Thakurs have higher consumption expenditure than Muraos in the 2009 survey.

Figure 3.1.: Regular Employment by Adult Males of Selected Castes



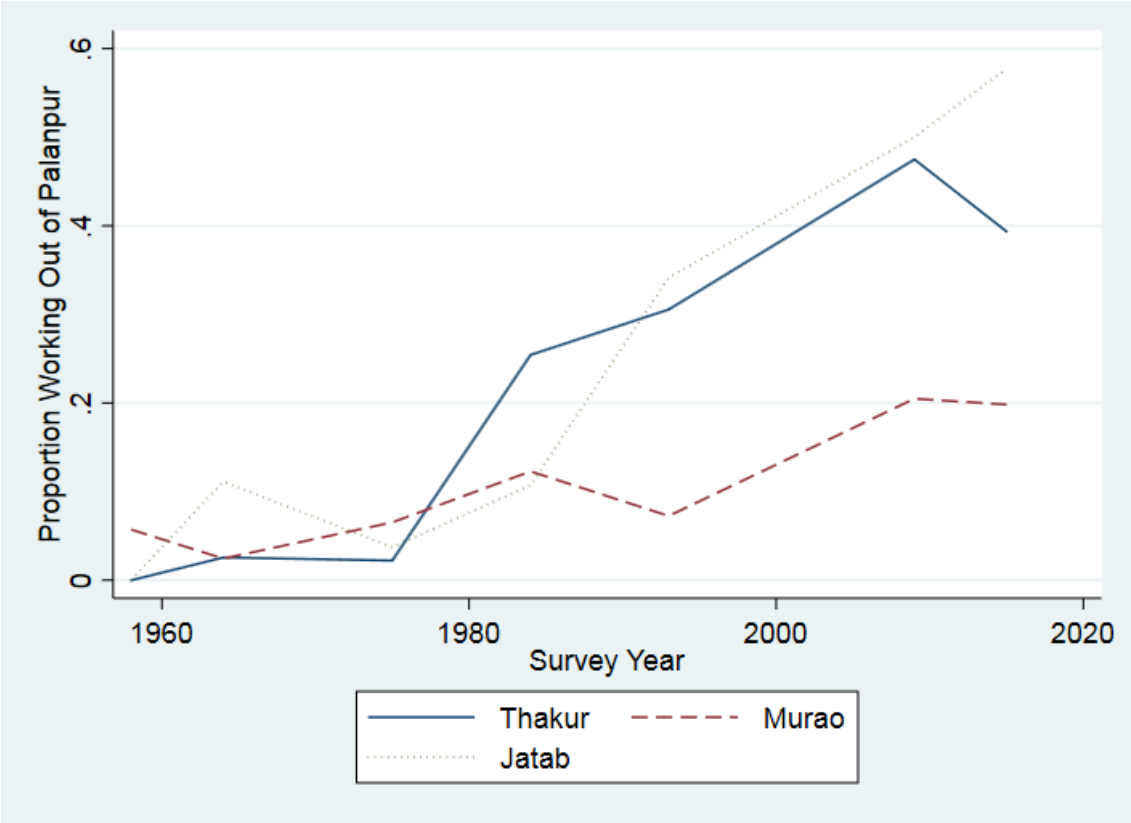
Note: All males 18 and over in migration sample included.
Source: Author's Calculations

One possible explanation for the differing fortunes of castes is the relative difference in employment patterns. Focusing on the 3 most populous castes, Figure 3.1 shows the

participation in regular employment by caste and year for adult males. It is apparent that Thakurs have been much more likely, than Muraos or Jatabs, to be employed in regular jobs in most of the surveys. Figure 3.2 displays the participation of these same castes in all employment which is outside the village[23]. Again Thakurs are more likely than Muraos to be working outside the village for the majority of the survey period, but now Jatabs (a scheduled caste) are roughly as likely as Thakurs to be employed outside Palanpur. The sample in these figures includes all male adults in the migration sample, that is it includes even those who have left Palanpur as long as they are still alive. We do this to avoid any trends being driven by selective migration of castes.

[23]This does not include those who migrate away from Palanpur permanently. Always when referring to outside employment we mean employment outside the village for those who live in the village.

Figure 3.2.: Outside Employment by Adult Males of Selected Castes



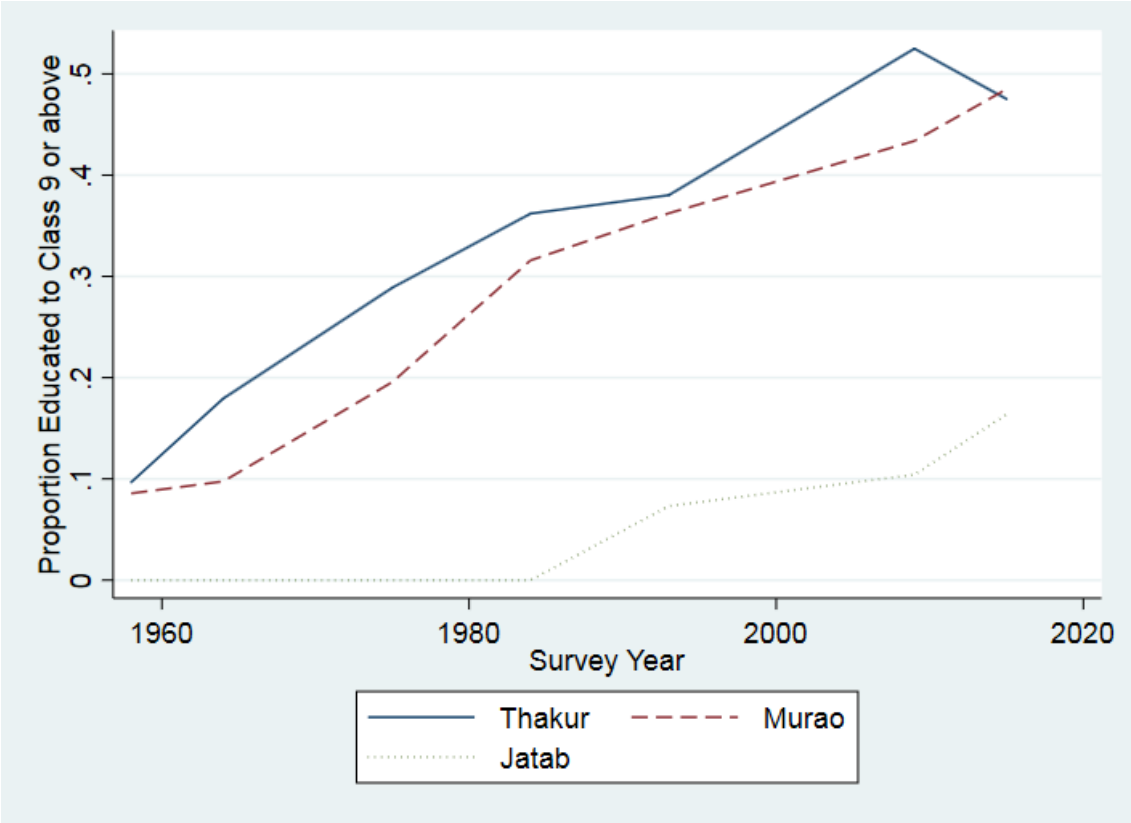
Note: All males 18 and over in migration sample included.
Source: Author's Calculations

These figures combined suggest that there is a degree of specialisation of employment by caste. Thakurs tend to work outside the village and are over represented in regular employment, Jatabs are seemingly more likely to work outside the village in more casual jobs, and Muraos are over-represented in work within the village. The traditional caste occupation of the Murao is agriculture, it is possible this caste heritage makes them more reluctant to venture into other types of employment. Another alternative explanation

is that all castes are attracted by these higher paying jobs (which tend to be outside the village) but individuals from some castes are able to make this transition because of connections they have. Then once these network connections are established they perpetuate a specialisation of employment type within these networks. Consistent with this evidence is Mukhopadhyay (2011), who finds using only 1983, 1993 and 2008 Palanpur data that even conditional on land ownership Thakurs are more likely to work outside than Muraos in 1993 and 2008 (he does not test if this is significant for all years combined).[24]

[24]Lanjouw and Stern (1998):p. 266 contains probits for 1975, 1984 and 1993 (again they do not combine years). They find that Murao, Thakur and Passi dummies are insignificant, but that Jatabs are significantly less likely to work outside during 1975 and 1984 - conditional on land ownership, education and male household size.

Figure 3.3.: Education by Adult Males of Selected Castes



Note: All males 18 and over present in the village included
Source: Author's Calculations

One potential confounder to this explanation is education. Perhaps, in the example of Muraos above, it is a lack of investment in education which means these regular jobs or outside jobs are unattainable for this caste. Figure 3.3 suggests that this is not likely to be the case here. It shows the proportion of the adult males in each of the 3 most populous castes who have achieved what we refer to as the highest levels of education. That is they are educated to class 9 or above. In absolute terms this is not a high level of education - it

represents completion of at least 4 years of secondary school - however, only around one-third of Palanpur males are educated to this standard and above. Figure 3.3 illustrates the general increasing trend in education in Palanpur, and also that Thakurs are not more likely than Muraos to be highly educated. In fact, Jatabs remain less likely than the average villager to be highly educated, but this has not stopped them taking up job opportunities outside the village.

This suggests to us that networks in employment may be important within Palanpur. However, while caste networks have been shown to be important in an Indian context, they are not the only potential network. In the next section we discuss another candidate, that of an individual's dynasty.

3.5. Dynasties in Palanpur

We define a dynasty as all of those households which descend from the same root household which was enumerated in the 1958 survey.[25] If households migrate into the village and their relationship to existing households is known then these households are added to the same dynasty. If in-migrating households have no ties to existing households, or they are unknown, then the households begin their own unique dynasty. Thus, the number of dynasties includes the 102 emanating from the original households in the 1958 survey, as well as 29 new dynasties formed by migrant households.

This definition of dynasty obviously ignores any relationships between the original 103 households, except for the two which merge in 1964. However, there is no simple or clearly exogenous way of identifying the links between the remaining households. Plus ignoring these links will bias our results towards zero if these links really matter in

[25] There is one exception to this, 2 households in the 1958 survey remerge in 1964, so are treated as one dynasty.

individuals networks. Therefore we have 131 dynasties in our data. Given the practice of endogamous marriage within Palanpur and India more widely, these dynasties are also a subset of an individual's caste network.

Table 3.3 shows the number of households which are present by caste in each of the survey years. 102 of the 103 original households present in 1958 are dynasties, but the additional dynasties by caste are shown by the difference between this number in the 1st column of Table 3.3 and the last column, named Unique Dynasties. Thus we find that the most populous castes are also those which have the most dynasties, there are 25 Thakur and 24 Murao dynasties respectively.

These dynasties range in size from 1 person households to dynasties of 103 unique individuals. The biggest dynasty in any single year has 63 individuals present across 10 households. The mean dynasty is 21 unique individuals present in at least one survey, this mean has a standard deviation of 18.1. There is also considerable variation in dynasties within surveys. The mean size of non-zero dynasty-years is 9.87 unique individuals, with a standard deviation of 8.4.

Given the variation in dynasty size, there must also be variation in the degree of relatedness of individuals within a dynasty. Obviously a dynasty will, in most cases, include an individual's parents, children, siblings, but it will often include nieces, nephews, grandparents and even cousins or further relations. Familial ties are important in many ways during an individual's life, we can observe this simply by looking at our own or our family's experience. Additionally there are countless papers in almost every area of science, medicine or social science which highlight a potential impact of familial ties. We do not list them here, but see Munshi (2014) for an introduction to the economics of these informal networks in the context of developing countries.

In the context of Palanpur dynasties are extremely important. Often many generations

Table 3.3.: Households by Caste-Year

	1958	1964	1975	1984	1993	2009	2015	Unique Dynasties
Thakur	17	19	25	30	48	57	55	25
Murao	21	25	28	27	44	55	54	24
Dhimar	10	9	8	13	15	22	19	11
Gadaria	9	9	10	12	14	15	15	9
Dhobi	2	1	3	4	4	8	8	5
Teli	8	9	12	16	20	21	17	11
Passi	13	17	11	15	16	6	5	14
Jatab	16	13	14	19	24	39	33	18
Lohar	1	0	0	0	0	0	0	1
Nai	2	1	1	1	1	1	1	2
Badhai	0	0	1	1	1	2	2	1
Kayasth	1	1	1	2	2	1	1	1
Khati	1	1	0	0	0	0	0	2
Carpenter	0	0	1	0	0	0	0	1
Balmiki	2	1	2	3	3	5	3	4
Brahmin	0	0	0	0	1	0	0	1
Unknown	0	0	0	0	0	1	0	1
Total	103	106	117	143	193	233	213	131

Note: Households present by caste for each survey year. Unique dynasties is the number of unique root households who are present in the village in at least one year.

Source: Author's Calculations

of one family will live in one household and eat the same food. This sort of close interaction and reliance on each other can foster very close ties between individuals. Thus given that we believe that there is some evidence of networks in Palanpur occupations within caste, supported by previous evidence supplied by the literature (e.g. Munshi and Rosenzweig (2006); Munshi and Rosenzweig (2016)), it would seem imprudent not to consider dynasties to be an important level at which networks may form. This is consistent with anecdotal evidence from Palanpur. For example, "Caste-based networks seem to influence access to non-farm jobs. In many cases, access to job opportunities is facilitated by network ties with family members, friends and neighbours. It is through such networks that information about job opportunities is disseminated and that employment in non-farm activities is secured." from Himanshu, Joshi et al. (2016).

3.6. Distribution of Outside Employment

Here we present some more evidence of networks in employment within Palanpur. We look at the distribution of outside employment for males across castes and dynasties.

Table 3.4 Panel A clearly shows a high degree of clustering of outside employment outcomes within both caste and dynasty-year pairs. We observe that 66% of dynasty-year pairs (33% of caste-year pairs) have no outside workers in them. If jobs were allocated randomly then each individual would have a probability of outside employment of just over 14%. Given the average dynasty-year size of 5.24 males, we would expect to observe 47% of dynasty-year pairs to have zero outside workers. The probability of observing 66% dynasty-year pairs of zero is essentially zero[26]. The story is similar for castes, given an average caste-year size of 38.97 males we would expect to see no castes without outside

[26] $\text{Prob} = 7.65e^{-19}$

workers, whereas we observe a third of them don't.

If jobs were allocated randomly with no regard to networks, then one would expect the proportion of individuals in these jobs to be constant across all networks, and the standard deviation would be zero. Although this is only true if the jobs are assigned randomly across years. Given an increasing trend in outside employment evident in the data this is clearly not the case. Thus we de-mean all network-year pairs relative to the yearly mean observed for all networks. We test the equality of the proportions between networks using this adjusted series, weighted by network members. We reject the hypothesis of equal proportions of outside workers within dynasties at any level of significance[27]. We conduct the same test for castes and reject the null of equality at a 0.22% significance level.

Clearly outside workers are not randomly assigned to between castes or dynasties. However, this doesn't guarantee that this observed pattern is the result of networks. It may be, as mentioned earlier that the ability (for example education) distribution is skewed towards some networks. We showed some evidence that this is not true for the example of education in section 3.4 for some castes, but we will return to a discussion of this issue later.

Also in Table 3.4, in Panel B, is the distribution of migrants across castes and dynasties. Once again the pattern is clearly one of non-random assignment. Both caste and dynasty-year pairs with no migrants or all migrants are overrepresented in the data. Using the same adjustment for the time trend in migration as we did for outside employment, we reject the equality of proportions of migrants in network-year pairs for both caste and dynasty. The significance levels at which we can reject are even smaller than those for outside employment.

[27] $\text{Prob} = 2.56e^{-13}$

Table 3.4.: Distribution of Outcomes Within Networks

Panel A	Outside Employment								
	Net-Year Pairs	Mean	S.D.	None %	All %	Min	Max	% 1 s.d >mean	% 1 s.d <mean
Caste	87	0.12	0.13	33.3	0	0	0.50	21.8	0
Dynasty	653	0.12	0.21	65.8	1.99	0	1	17.0	0
Panel B	Migration								
	Net-Year Pairs	Mean	S.D.	None %	All %	Min	Max	% 1 s.d >mean	% 1 s.d <mean
Caste	108	0.23	0.31	31.5	9.26	0	1	14.8	0
Dynasty	834	0.21	0.35	63.3	12.6	0	1	17.0	0

Note: All Males present in village included in Panel A. All males in migration sample included in Panel B. Net-Year Pairs: Number of network-year pairs. None %: is the percentage of network-year pairs for which none of the males are engaged in the activity in question All %: is equivalently defined for network-pairs where all males are engaged in the activity in question.

Source: Author's Calculations

3.7. Network Sizes

In the previous sections we have shown the potential importance of caste and dynasty networks in the labour market in Palanpur. Briefly in this section we describe in more detail the sizes of these networks and the construction of individual level network variables.

Firstly it is important to describe the samples used to for the analysis that follows. There are two primary samples used, one which we refer to as the "present sample", and one referred to as the "migration sample". The present sample contains for each survey year, as the name suggests, all those individuals who are present in Palanpur at the time of the survey. They may work outside the village or migrate away for short periods for work, but they primarily reside in the village. The migration sample includes all those in the present sample, plus those who are not present in the village but at some point in time were members of a household in Palanpur[28]. All individuals alive in a year are included in the sample for that year. If it is unknown whether an individual who has left the village is alive or dead we treat them as alive until they reach an age of over 65, as this is the rural life expectancy in India for 2006-2010[29]. The present sample consists of

[28] These individuals need not be present in the village during any survey to be recorded in our data. During each survey households are asked about any members who have left the village between surveys, and some basic demographic details about these individuals.

[29] http://censusindia.gov.in/vital_statistics/SRS_Based/Introduction.pdf

2,752 distinct individuals, but the migration sample comprises 2,831 distinct individuals.

Basic network (both caste and dynasty) variables are created simply by counting the numbers of individuals in each sample during each survey, overall and by sex. We also generate network variables which are the counts of an individual's network members engaged in particular types of job in a given year. For example, an individual may have 3 members of his caste working outside the village in 1975. For these variables we exclude the individual whose network is being considered. Therefore the variable represents the potential contacts in that employment type for the individual.

One quirk of the Palanpur data is that it is a series of snapshots of a village which are often a long time apart. This means that in each subsequent survey we observe people at very different times in their lives. In contrast to many panel surveys which are annual. Due to this, and the census nature of the data, we can reconstruct the demographics of the village to any point since 1958. We use this to construct truly individual network variables which can represent the size of an individual's network at any time in their life. Most commonly we use the size of an individual's network on their first appearance as an adult ($\geqslant 18$) in what follows.

3.7.1. Caste Networks

In Table 3.5 we present summary statistics for some of our caste network variables. We focus only on males because they are the primary participants in the labour market within a household. This table shows us that these caste networks are fairly large relative to the size of the village. The mean caste network size is nearly 9% of the total number of distinct individuals. This is driven by a number of very large castes, and there is substantial variation in the size of caste networks. The size of caste networks

Table 3.5.: Caste Network Sizes for Males

	Mean	Median	S.D.	Min	Max	Obs
Caste Network	243.2	200	157.4	1	526	4391
Caste Network of Males	104.7	84	65.2	1	227	4391
Caste Network at 18	191.0	147	143.0	0	525	3458
Caste Outside Employment Network	15.1	9	15.8	0	50	4391
Caste Regular Employment Network	7.54	5	9.00	0	35	4391
Caste Non-Farm Self-Emp Network	5.80	3	6.17	0	21	4391
Caste Highly Educated Network	30.4	19	29.7	0	90	4391

Note: All males in migration sample included calculations. Caste Network and Network of Males variables include all individuals in network. Those variables which are occupation-specific do not include the member whose network is being considered, and thus are the potential contacts for the member in question.

Source: Author's Calculations

at 18 is smaller because of population growth at the village level, and there are fewer observations since we do not create this variable for individuals who do not reach 18 in the survey period. Since our focus is the labour market and males are the primary participants in the labour market in our context it is important to note the average male has a network of approximately 100 males within his caste. Not all of these males live within the village so this is a large and potentially widely dispersed network of contacts. There are an average of 15 outside workers, 8 in regular employment and 6 self-employed in non-farm occupations. These averages disguise substantial variation in the typical occupations within these networks, there are networks with much greater or lesser contact with these jobs types.

3.7.2. Dynasty Networks

The dynasty networks summarised in Table 3.6 are generated in the same way as the caste network variables. These networks also show substantial variation in size, and in employment of different types. These dynasty networks are much smaller than the caste networks, approximately 10% of the size. Thus now, the network of labour market contacts within a dynasty is on average 8.85[30]. This is a considerably smaller network than one's caste network but the kinship ties between them are much closer on average.

It is apparent when comparing the caste and dynasty networks that employment types are more dispersed in dynasties than castes. The coefficient of variation of these employment networks for dynasties vary between 1.44 and 1.96, while for caste employment networks these range between 1.05 and 1.19. This larger specialisation in dynasties than castes is also present for education, coefficient of variation of 1.42 versus 0.98.

[30] subtracting the individual whose network we are considering

Table 3.6.: Dynasty Network Sizes for Males

	Mean	Median	S.D.	Min	Max	Obs
Dynasty Network	22.0	17	17.2	1	85	4391
Dynasty Network of Males	9.85	8	7.43	1	36	4391
Dynasty Network at 18	16.4	12	15.1	0	82	3458
Dynasty Outside Employment Network	1.42	0	2.05	0	9	4391
Dynasty Regular Employment Network	0.73	0	1.43	0	9	4391
Dynasty Non-Farm Self-Emp Network	0.60	0	1.12	0	6	4391
Dynasty Highly Educated Network	2.88	1	4.08	0	21	4391

Note: All males in migration sample included calculations. Dynasty Network and Network of Males variables include all individuals in network. Those variables which are occupation-specific do not include the member whose network is being considered, and thus are the potential contacts for the member in question.

Source: Author's Calculations

3.7.3. Competitive Networks

We are not only interested in the size of individuals networks. Another factor we think will be relevant to the effect of networks on job attainment is the degree of competition for a job which is faced by an individual within his network. We believe this is captured well by a count of the number of network members an individual has within a small number of years of them. That is, an extra network member who is a year older than you is much more of a direct competitor for any potential jobs which become available than one who is 10 years older or younger than you. We describe the nature of this competition within both caste and dynasty networks in Table 3.7, panels A and B. These display the mean and median age difference between the individual whose network is being measured and the closest member of the same sex in their network. As expected caste, as the larger of the two network sizes, is more densely populated and therefore has lower average differences between closest network members.

Also in both panels of Table 3.7 are counts of the number of male members within 4 years (inclusive) of an individual within their network. It is these network members, males within 4 years, who we consider to be the competitive network of an individual[31]. As can be seen, the number of competitors in both networks is increasing over time due to population growth, and this is mirrored by the falling mean and median age differences within networks. The median male in our sample has one competitor within his dynasty, but 13 within his caste. However, *a priori* these competitors may be a hindrance or a benefit to an individual. They may take a job referral which may otherwise have been passed to the individual in question, or they may be the source of a job referral to

[31] Appendix Table C.2 contains an extended set of descriptives for caste networks, including descriptives for different definitions of a competition network. Appendix Table C.3 contains these same descriptives for dynasty level competition variables.

Table 3.7.: Competitive Network Descriptives - Males

Panel A	Caste							
	1958	1964	1975	1984	1993	2009	2015	All
Mean age difference	1.47	1.30	0.93	0.74	0.65	0.48	0.45	0.73
Median age difference	0	0	0	0	0	0	0	0
Mean % with competitor within 4 years	89.6	91.0	92.5	96.0	96.6	97.5	97.7	95.4
Mean: count competitors within 4 years	7.42	8.09	11.9	14.0	18.4	20.3	22.3	16.6
Median: count competitors within 4 years	6	7	10	12	14	17	21	13
Panel B	Dynasty							
	1958	1964	1975	1984	1993	2009	2015	All
Mean age difference	12.8	12.8	9.37	7.21	5.66	4.36	3.90	6.73
Median age difference	7	7	5	4	3	2	2	3
Mean % with competitor within 4 years	33.4	36.0	46.9	57.4	64.4	72.2	74.7	60.4
Mean: count competitors within 4 years	0.46	0.55	0.87	1.07	1.53	1.81	2.07	1.40
Median: count competitors within 4 years	0	0	0	1	1	1	2	1

Note: All Males in migration sample included. The Mean and Median age difference refer to the age difference between an individual and the closest male member in their network in the migration sample. Panel A contains competitive networks for males within Caste, Panel B contains those competitive networks within Dynasty.

Source: Author's Calculations

the individual. Which of these effects dominates is an empirical question.

The next section explores the effects of these networks within the labour market in Palanpur.

3.8. Event Study

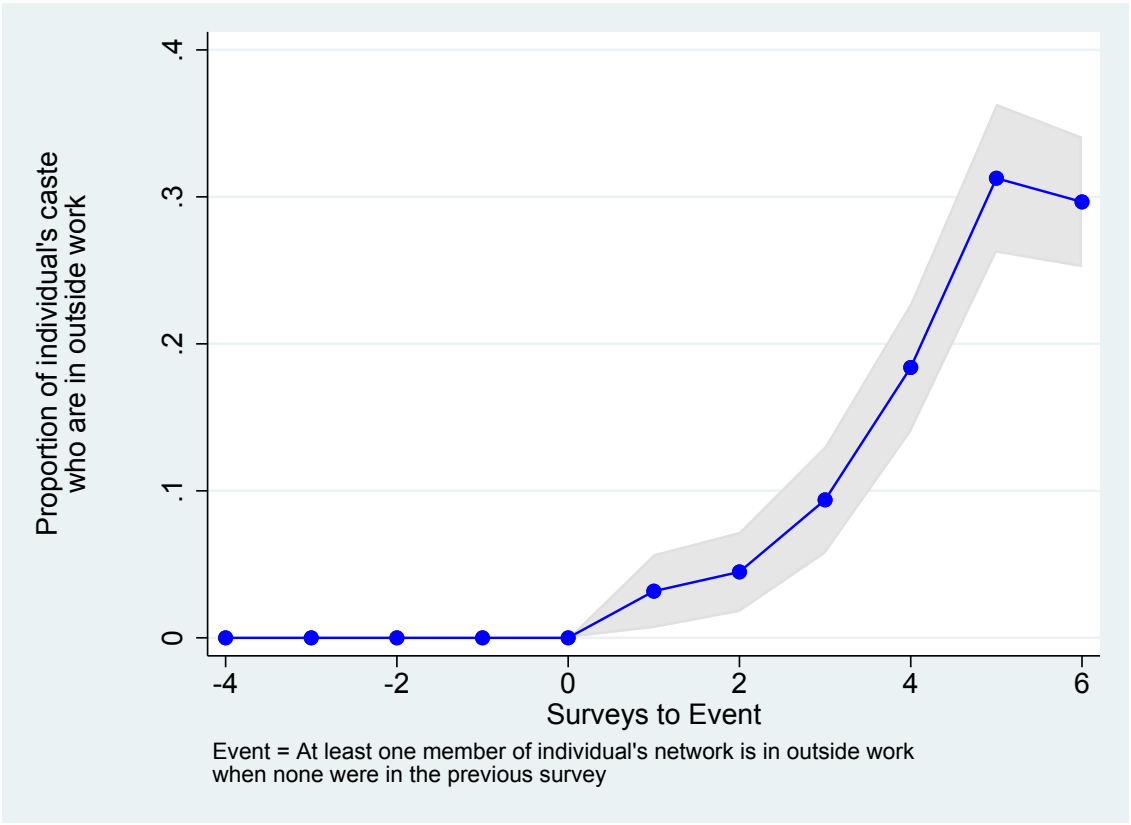
3.8.1. Outside Employment

In this section we present graphical evidence for these network effects in the form of an event study. The event we use here is that at least one member of an individual's network is in outside work in the current survey when none were in the previous survey[32]. Our event study examines the effect of this event on the proportion of an individual's network who are in outside employment, after removing the members of the network who received these initial outside jobs. For the graphical evidence we restrict the sample to include only the first occurrence of an outside worker within the network of interest. That is, the event captured by the graph is the first appearance of an outside worker within an individual's network[33].

Figure 3.4 displays this event study for caste networks. We see that while the appearance of an outside worker within one's caste network increases the likelihood that other caste members are employed outside the village in subsequent surveys, there is not a sharp jump in the probability initially after the event. However, in Figure 3.5, which shows the same event study for dynasty networks, there is a large jump in the probability of outside employment in the first survey after a dynasty member gains outside employment. This large increased proportion of the network which is employed outside

[32] $y_{nt} \geqslant 1 \cap y_{n(t-1)} = 0$. Where y_{nt} is outside employment within network n in survey t.
[33] $y_{nt} \geqslant 1 \cap y_{n(t-j)} = 0 \; \forall j > 0$

Figure 3.4.: Event Study Graph - Outside Employment - Caste



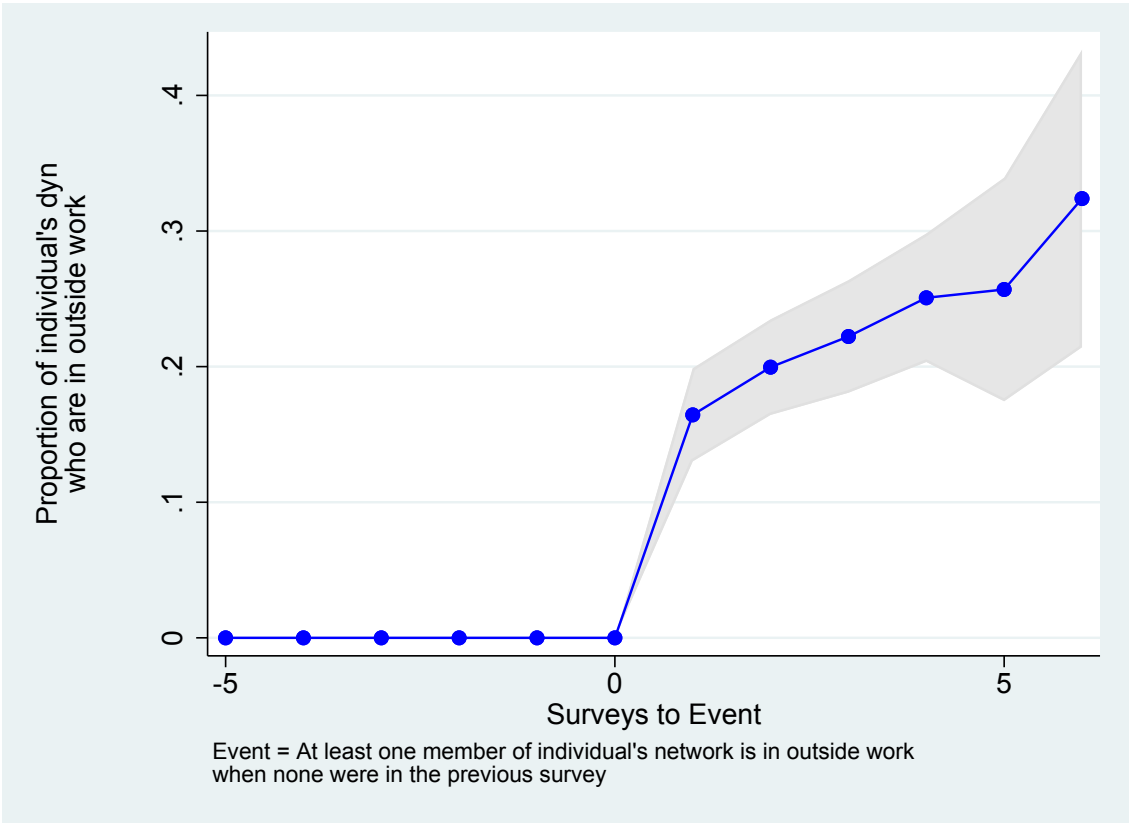
Note: Includes all males in migration sample. Figure includes only the first appearance of an outside worker within an individual's caste network. The event occurs at 0 on the x-axis.
Source: Author's Calculations

the village is sustained over all remaining surveys. We think this is interesting and compelling evidence for the existence of network effects in outside employment. Due to the periodic nature of the Palanpur data these are not simply the same individuals remaining in outside employment. The surveys are 6 to 16 years apart, the nature of a dynasty network can be markedly different in this time span - children are born, households sub-divide, members migrate or die.

After the graphical evidence we now turn to some simple regression specifications to explore these network effects in outside employment further. Table 3.8 contains a version of this event study for caste networks. Our data permit us to include caste, dynasty and individual fixed effects - however since these are subsets of one another we can only include one at a time - and also year fixed effects, which can also be thought of as survey fixed effects, are included in all specifications. Once again we include only males in our sample due to the lack of a female labour market outside the village. We include only males who are 14 or older in any survey since those who never appear in our data as older than this will almost certainly not obtain employment outside the village at such a young age. For this section we only include the sample of individuals who are living in the village, since we consider those resident in the village to be more indicative of the network of competitors which an individual faces in trying to obtain one of these jobs. We will discuss migration further when we examine the long run effects of these networks later in section 3.9.

In the regression version of this event study we include all occurrences of outside workers within an individual's network. The event variable is defined as equal to one if there are *any* outside workers (not including the individual's own employment) within an individual's network, and zero otherwise. The regressions are of the form:

Figure 3.5.: Event Study Graph - Outside Employment - Dynasty



Note: Includes all males in migration sample. Figure includes only the first appearance of an outside worker within an individual's dynasty network. The event occurs at 0 on the x-axis.
Source: Author's Calculations

$$y_{idct} = \alpha + \beta E_{idct} + \gamma(E_{idct} \times C_{idct}) + \delta_t + \phi_{(i)(d)(c)} + \epsilon_{idct} \tag{3.1}$$

Where E_{idct} is the event variable for an individual, i, of caste c, and dynasty d, in survey t. y_{idct} is a dummy variable for the employment status of individual i. $y_{idct} = 1$ if individual is employed outside the village in survey t, and zero otherwise. Thus β captures the job network effect - if it is positive (as we would expect) then the presence of outside workers within an individual's network is associated with an increased likelihood of the individual being employed outside the village.

The regressions include survey fixed effects, δ_t, and one of individual, dynasty or caste fixed effects, $\phi_{(i)(d)(c)}$. They also include an interaction between the event variable and C_{idct}, a variable designed to capture the degree of *within-network* competition that an individual faces. Here we define a competitor as another male member of the network who is aged within 4 years of the individual under consideration. Then the degree of competition faced by an individual, C_{idct}, is the count of such members within their network. If the job networks within Palanpur behave as is typical of the theoretical literature (Calvó-Armengol and Jackson (2004)) then job information is passed through employed individuals to their contacts. Thus if an employed individual has more contacts then each of those contacts is less likely to receive the job referral. As a consequence we expect the job network effect to be mitigated by this form of competition. In our empirical specification, Equation 3.1, this mitigation is captured by γ which we expect to be negative - and the total job network effect of an outside worker in an individual's network is given by $(\beta + \gamma C_{idct})$.

Table 3.8.: Outside Employment Caste Networks

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside
Event	0.0468**	0.0239	0.0107	0.00465	0.00922	0.0226
	(0.0223)	(0.0292)	(0.0309)	(0.0307)	(0.0283)	(0.0295)
Competition Interaction	−0.00218**	−0.00486**	0.0000792	0.000524	−0.00557***	−0.00564***
	(0.000930)	(0.00221)	(0.00104)	(0.00103)	(0.00184)	(0.00205)
Caste Network			−0.0000767	−0.000218	−0.00105***	−0.00155**
			(0.000381)	(0.000401)	(0.000384)	(0.000763)
Dynasty Network			0.00117	0.000789	−0.00278	0.00157
			(0.00116)	(0.00175)	(0.00188)	(0.00236)
Caste members working outside					0.00303*	
					(0.00158)	
Dynasty member working outside					0.0344***	
					(0.00790)	
Caste members highly educated						0.00290
						(0.00253)
Dynasty member highly educated						−0.00970
						(0.0102)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	No	No	Yes	No	No	No
Dynasty FE	No	No	No	Yes	No	No
Individual FE	No	Yes	No	No	Yes	Yes
Observations	2949	2826	2949	2949	2826	2826
Castes	13	12	13	13	12	12
Dynasties	116	110	116	116	110	110
Individuals	923	800	923	923	800	800
Mean Dep Var	0.163	0.167	0.163	0.163	0.167	0.167
R^2	0.145	0.488	0.169	0.229	0.501	0.492
P-Value of Comp Interaction	0.021	0.030	0.939	0.613	0.003	0.007

Note: Event: There is at least one individual in outside employment in individuals caste. Competition Interaction: Event variable interacted with count of caste members in an individual's competitive network (within 4 years of age) Dependent variable is dummy variable for working in outside employment. Only males who are over 14 years of age in at least one survey are included. Standard errors are clustered at the dynasty level.

Table 3.8 shows that the presence of an outside worker in an individual's caste is associated with a $0.47 - 4.7$ percentage point increase in the likelihood of that individual working outside in that survey, although only one of these coefficients is significantly different from zero. Given a mean of outside work of 16.7% in this sample, that is a 13.5% increase in the probability of working outside in our preferred specification, column (6). We interpret this as the job network effect of castes. This job network effect in outside work is mitigated, as predicted, by the degree of competition faced by an individual within their network. In the regressions in Table 3.8 the interaction between caste competition and the event variable shows us that the presence of an extra competitor in your caste network reduces this job network effect by 0.56 percentage points in our preferred specification.

Columns (1) and (2) in Table 3.8 present the simplest form of these regressions, with only age controls, column (1) has only year fixed effects, while column (2) contains individual fixed effects also. Progressively we add controls and different levels of fixed effects. In columns (3)-(6) we control for both caste and dynasty size to help control for differential migration or demographic trends of these networks. Column (5) includes controls for the number of outside workers within caste and dynasty. These are obviously bad controls if, as we contend, job networks exist in outside employment since they will be co-determined with the dependent variable. Their inclusion does not change the point estimates of our event and competition interaction variables. It does however reduce the significance of the event variable because they are correlated with it by construction. However, the event effect remains positive even with these controls for the degree of interaction with outside jobs within an individual's caste and dynasty. Additionally the presence of more network members working outside is positively associated with individual outside employment for both networks, as one would expect in the pres-

ence of network effects. Our preferred specification is that given in column (6) of Table 3.8, since this includes individual fixed effects which remove any endogeneity from fixed characteristics of the individuals, such as ability. Column (6) also controls for the total size of the caste and dynasty networks of an individual in case there is an interaction between the referral behaviour of networks and their size. Additional controls for the number of highly educated (completed class 9 or above) members of the individuals caste and dynasty are also included - these are intended as proxies for the overall level of ability within networks, and to control for any endogeneity coming from differential referral behaviour by more or less able networks. The inclusion of these variables leaves the coefficients of interest unchanged in magnitude and significance.

Table 3.9.: Outside Employment Dynasty Networks

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside
Event	0.168***	0.113***	0.138***	0.0693**	0.0605	0.120***
	(0.0273)	(0.0314)	(0.0273)	(0.0319)	(0.0366)	(0.0318)
Competition Interaction	−0.0219**	−0.0255*	−0.0158*	−0.0171**	−0.0265**	−0.0250*
	(0.00928)	(0.0143)	(0.00808)	(0.00755)	(0.0133)	(0.0143)
Caste Network			−0.0000540	−0.000154	−0.00103***	−0.00143**
			(0.000301)	(0.000350)	(0.000361)	(0.000660)
Dynasty Network			−0.0000227	0.000738	−0.00222	0.000435
			(0.000952)	(0.00170)	(0.00172)	(0.00209)
Caste members working outside					0.00280*	
					(0.00152)	
Dynasty member working outside					0.0291***	
					(0.00923)	
Caste members highly educated						0.00256
						(0.00223)
Dynasty member highly educated						−0.00940
						(0.00886)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	No	No	Yes	No	No	No
Dynasty FE	No	No	No	Yes	No	No
Individual FE	No	Yes	No	No	Yes	Yes
Observations	2949	2826	2949	2949	2826	2826
Castes	13	12	13	13	12	12
Dynasties	116	110	116	116	110	110
Individuals	923	800	923	923	800	800
Mean Dep Var	0.163	0.167	0.163	0.163	0.167	0.167
R^2	0.169	0.493	0.186	0.233	0.501	0.497
P-Value of Comp Interaction	0.020	0.077	0.053	0.025	0.049	0.082

Note: Event: There is at least one individual in outside employment in individuals dynasty. Competition Interaction: Event variable interacted with count of dynasty members in an individual's competitive network (within 4 years of age) Dependent variable is dummy variable for working in outside employment. Only males who are over 14 years of age in at least one survey are included. Standard errors are clustered at the dynasty level.

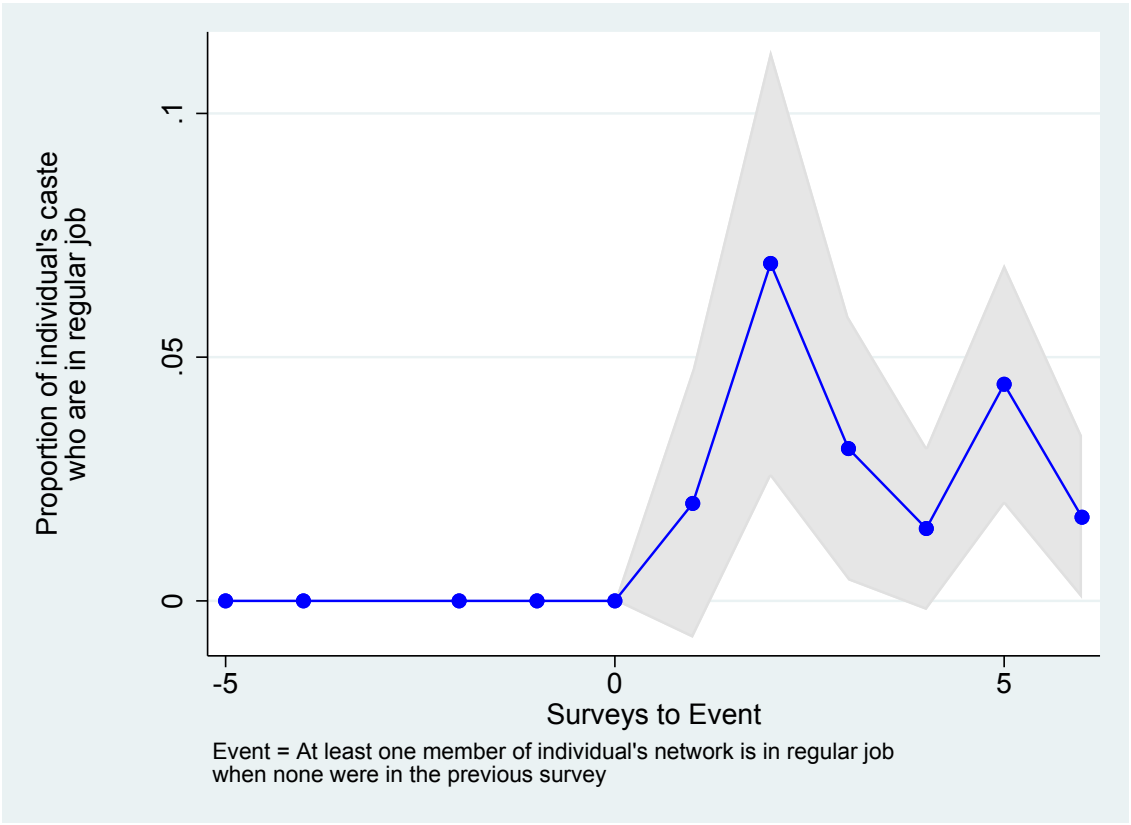
Table 3.9 contains the same regressions for outside employment network effects, but this time using the dynasty as our network of analysis. It shows a much larger job network effect. The presence of an outside worker in dynasty is associated with a $6.1-16.8$ percentage point increase in the likelihood of an individual working outside the village. In our preferred specification, column (6), this is a 71.86% increase in the probability of working outside the village. Just as with caste networks this effect is mitigated by the degree of competition faced by an individual. An extra competitor in an individual's dynasty network decreases this probability of outside employment by 2.5 percentage points in column (6). These effect sizes are larger than those for caste networks. This is unsurprising given that the connectedness of contacts in a dynasty network will be closer than for a caste network.

3.8.2. Regular Employment

Turning our attention now to an event study using a different type of employment, regular employment. Introduced in Section 3.2, this definition of a job is common in rural India but to a first approximation it means that wages are paid monthly rather than daily. These jobs are relatively highly remunerated, and are thus highly sought after. While there are regular jobs in the village (e.g. railway workers) most, 71.3%, of all regular jobs done by an adult Palanpur resident in any survey are outside the village. We believe that these jobs are less likely to exhibit the features of employment networks due to certain requirements for the position, or a selection process which happens at a higher level of authority. However, due to the wide variety of jobs which are still included in this job category it is likely that some of them may permit referrals.

Figure 3.6 contains our event study done in the same way as for outside employment

Figure 3.6.: Event Study Graph - Caste



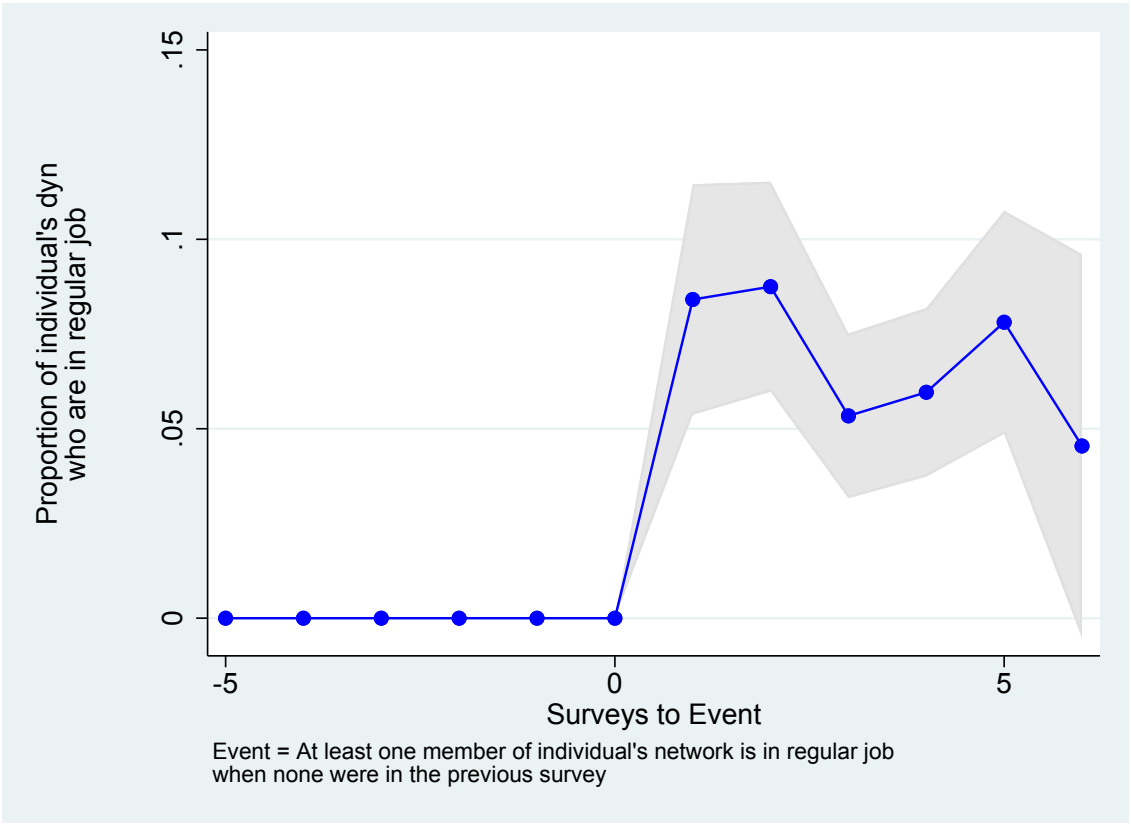
Note: Includes males in migration sample. Figure includes only the first appearance of a regular worker within an individual's caste network. The event occurs at 0 on the x-axis.
Source: Author's Calculations

previously, this time done for regular employment within caste networks. Now the event illustrated is the first appearance of a regular employee within an individual's network. While there is an increase in the proportion of regular employees within a caste, after the initial event of one member obtaining an outside job, this increase is not a sharp increase in the initial survey after the event. It also appears from the figure that this increase is not sustained.

In comparison, Figure 3.7 illustrates the event study for regular employment within dynasties graphically. The increase in the proportion of regular employees within the dynasty happens sharply with the first appearance of at least one regular employee. It also appears that, despite a dip, this increase is sustained within dynasties in subsequent surveys. These figures taken together suggest that dynasty networks are more important for regular employment labour markets than are castes.

Examining these short run job networks using the same regression specifications (Equation 3.1) as we did previously with outside employment opportunities. Except this time with a dummy for employment in a regular job as the dependent variable. Table 3.10 shows the regressions for caste networks. The presence of a regular employee in caste is negatively associated with an individual working in regular employment. This indicates a lack of a job referral network for regular jobs in Palanpur, so we observe a negative effect because these jobs are not plentiful so if a network member obtains one of these jobs then that is one which you can now no longer obtain. Consistent with this an extra competitor within ones network is associated with a lower probability of regular employment.

Figure 3.7.: Event Study Graph - Dynasty



Note: Includes all males in migration sample. Figure includes only the first appearance of a regular worker within an individual's dynasty network. The event occurs at 0 on the x-axis.
Source: Author's Calculations

Table 3.10.: Regular Employment Caste Networks

	(1) Regular Emp	(2) Regular Emp	(3) Regular Emp	(4) Regular Emp	(5) Regular Emp	(6) Regular Emp
Event	−0.0136	−0.0833*	−0.125***	−0.128***	−0.0899**	−0.0728*
	(0.0336)	(0.0428)	(0.0407)	(0.0433)	(0.0419)	(0.0420)
Competition Interaction	−0.000162	−0.00535***	−0.000553	−0.000455	−0.00534***	−0.00570***
	(0.000658)	(0.00177)	(0.000849)	(0.000943)	(0.00146)	(0.00173)
Caste Network			0.000717***	0.000569**	−0.000715**	−0.000860**
			(0.000251)	(0.000268)	(0.000293)	(0.000377)
Dynasty Network			0.00119	0.00154	0.00238	0.00213
			(0.00117)	(0.00149)	(0.00183)	(0.00218)
Caste members in regular job					0.00716***	
					(0.00251)	
Dynasty member in regular job					0.0201	
					(0.0139)	
Caste members highly educated						0.00250**
						(0.00117)
Dynasty member highly educated						0.00402
						(0.00733)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	No	No	Yes	No	No	No
Dynasty FE	No	No	No	Yes	No	No
Individual FE	No	Yes	No	No	Yes	Yes
Observations	2949	2826	2949	2949	2826	2826
Castes	13	12	13	13	12	12
Dynasties	116	110	116	116	110	110
Individuals	923	800	923	923	800	800
Mean Dep Var	0.091	0.093	0.091	0.091	0.093	0.093
R^2	0.058	0.480	0.112	0.181	0.492	0.483
P-Value of Comp Interaction	0.806	0.003	0.516	0.630	0.000	0.001

Note: Event: There is at least one individual in regular employment in individuals caste. Competition Interaction: Event variable interacted with count of caste members in an individual's competitive network (within 4 years of age) Dependent variable is dummy variable for working in regular employment. Only males who are over 14 years of age in at least one survey are included. Standard errors are clustered at the dynasty level.

The dynasty level job network effect regressions are in Table 3.11. In these regressions the event of a dynasty member obtaining a regular job is associated with between a −1.5 to 11.6 percentage point increase in the probability of individual regular employment. This is a large range of effect sizes and two are insignificantly different from zero at the 10% level of significance. In our preferred specification (column 6) the effect size is 5.5 percentage points and is marginally significant. This apparent increase in the regular employment probability does not seem to be related to the number of competitors within an individual's dynasty. The effect of an extra competitor is insignificantly different from zero in all specifications, although consistently slightly negative as we would expect.

The coefficients on the competition interaction variable are at least an order of magnitude smaller than those for outside employment, while the outside employment jobs are only twice as numerous. So outside workers appear to be more networked than regular employees. This is plausible, since a number of the regular jobs require formal training, e.g. teachers, while others are selected at a higher level than the village, e.g. Public Distribution Scheme worker. In comparison many of the outside jobs which are not regular are more casual in nature.

Table 3.11.: Regular Employment Dynasty Networks

	(1) Regular Emp	(2) Regular Emp	(3) Regular Emp	(4) Regular Emp	(5) Regular Emp	(6) Regular Emp
Event	0.116***	0.0595**	0.0686**	−0.0154	0.00390	0.0548*
	(0.0269)	(0.0288)	(0.0292)	(0.0304)	(0.0324)	(0.0289)
Competition Interaction	−0.00990	−0.0153	−0.00885	−0.00926	−0.0121	−0.0177
	(0.00776)	(0.0111)	(0.00808)	(0.00817)	(0.0107)	(0.0114)
Caste Network			0.000671***	0.000586**	−0.000586*	−0.000811**
			(0.000210)	(0.000282)	(0.000299)	(0.000389)
Dynasty Network			0.000651	0.00212	0.00247	0.00185
			(0.00105)	(0.00178)	(0.00175)	(0.00194)
Caste members in regular job					0.00708***	
					(0.00251)	
Dynasty member in regular job					0.0221	
					(0.0162)	
Caste members highly educated						0.00266**
						(0.00111)
Dynasty member highly educated						0.00404
						(0.00661)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	No	No	Yes	No	No	No
Dynasty FE	No	No	No	Yes	No	No
Individual FE	No	Yes	No	No	Yes	Yes
Observations	2949	2826	2949	2949	2826	2826
Castes	13	12	13	13	12	12
Dynasties	116	110	116	116	110	110
Individuals	923	800	923	923	800	800
Mean Dep Var	0.091	0.093	0.091	0.091	0.093	0.093
R^2	0.085	0.476	0.113	0.177	0.486	0.479
P-Value of Comp Interaction	0.205	0.172	0.275	0.259	0.264	0.124

Note: Event: There is at least one individual in regular employment in individuals dynasty. Competition Interaction: Event variable interacted with count of dynasty members in an individual's competitive network (within 4 years of age) Dependent variable is dummy variable for working in regular employment. Only males who are over 14 years of age in at least one survey are included. Standard errors are clustered at the dynasty level.

The results in this section suggest that the relevant social division for labour market networks in this context is families and not castes. The magnitudes of these short run job network effects are larger for dynasties in both subsets of jobs we examine, and the effect of competition from others is more pronounced at this level. While caste networks also appear to have these short run job network effects, particularly for outside employment, and are important determinants of outcomes in the labour market, if we ignore the role of family ties and dynasties then we will be missing an important social network in this context.

In order to try and identify the causal impact of the networks on job attainment this section has relied on a series of fixed effects. The survey design permits the inclusion of individual fixed effects, these remove any unobserved time-invariant characteristics of individuals which may be correlated with the probability of attaining an outside job, such as ability. The inclusion of year fixed effects removes differences in average probability of working outside the village for different surveys. However, there is still a threat to identification if there are time-varying unobserved characteristics which determine whether an individual works outside the village, and if these unobserved characteristics are correlated with the event of a member of ones network being employed outside the village when there were none previously.

This section has been largely focused on associations between individuals in Palanpur in the same survey, and the correlations in their outcomes in the job market. While these labour market network effects, and the level at which they operate, are interesting, the primary advantage of the Palanpur data is the length of time over which we have data. Thus we now turn our attention to the presence of long run effects of these networks of caste and dynasty in labour market outcomes.

Long-run Job Network Effects

3.9. Dynamics of Job Networks in Palanpur

From the previous section we know that a member of one's dynasty becoming employed outside the village between surveys is associated with a 70% increase in individual probability of outside employment. This is the short run effect of job networks. However, we are also interested in the long-run effects of these job networks. To aid in this we use a framework from Beaman (2012) to illustrate intuition and provide testable implications.

3.9.1. Theoretical Framework

The model from Beaman (2012) has an OLG framework where each agent lives and works for S periods. An agents cohort, c, is defined as the period they enter the labour market, in the context of Palanpur this cohort consists of the individuals who reach adulthood (finish school) at or around the same time. Each cohort is of size N_c, all individuals within a cohort are identical. We assume that all individuals within a network are connected, the network being the dynasty or caste in this context. If an agent i in cohort c is employed outside village at end of period t, $s^t_{ic} = 1$. $s^t_{ic} = 0$ if agent i is not employed outside the village.

There is an exogenous breakup rate of b, which is the probability of losing a job at the beginning of the period. An agent hears about a job with probability a, independent of

other agents[34]. If the agent is not employed outside the village, they accept the job. If they are already employed, they pass on to an unemployed network member. The constant probability with which jobs arrive, a, means that the total number of jobs available scales up as the size of the network increases. However, this assumption is reasonable if the agent's network is small with respect to the world. In the context of Palanpur this seems entirely justified since we are considering an individual's caste or dynasty network in relation to the job market in larger villages or towns in the surrounding area. The employment rate, s^t_{ic}, is the probability of outside employment for an individual, i, in cohort c, in period t.

$$s^t_c = a + r^t \quad \text{if } c = t \tag{3.2}$$

Where r^t is the probability of receiving job information through a network member. Thus for individuals already in the market for at least one period, the probability of being employed outside is:

$$s^t_c = (1-b)s^{t-1}_c + (1-(1-b)s^{t-1}_c)(a + r^t) \quad \text{if } c \leqslant t \leqslant c + (S-1) \tag{3.3}$$

Given a network size in a given period, N_k, the probability of receiving job information through a network member is given by

$$r^t = \sum_{k=t-S+1}^{t-1} \frac{aN_k(1-b)s^{t-1}_k}{\sum_{k=t-S+1}^{t} N_k - (1-b)\sum_{k=t-S+1}^{t-1} N_k s^{t-1}_k} \tag{3.4}$$

[34]Propositions 1 and 2 which we use are robust to allowing the job information arrival rate to be state dependent, i.e. depend on employment status. See Beaman (2012)

Where the numerator is the total number of jobs that are available in network to be passed and the denominator is the number of potential recipients. The timing of the model is as follows: exogenous breakup, learn about jobs, accept jobs if unemployed, pass jobs onto unemployed network member if employed, network members accept jobs, work, end of period. The model gives rise to 2 propositions which are useful in the context of Palanpur.

Proposition 3.1. *For all values* $0 < a < 1$ *and* $0 < b < 1$*, an increase in cohort size* N_j *decreases* s_c^j *for all* c*.*

Proposition 3.2. *The impact of an increase in* N_j *on* s_k^k *is monotonically increasing from period* $k = j$ *to period* $k = j + S - 1$

That is, an increase in the size of the current cohort, j, should decrease the current employment rate for all cohorts in proposition 1. This is the *competition effect* within the model, since all new members of the labour market are competing for the job information held by existing members of the labour market. Proposition 2 states that an increase in the size of an entering cohort negatively impacts cohorts near it (negative from Proposition 1), but that this effect becomes positive for cohorts sufficiently far apart. Eventually the increase in cohort j will increase the employment rate for cohorts arriving later. This is because the fraction of individuals employed in cohort j through the referral channel increases with each period, when their employment rate is sufficiently high they become an asset to their network. This is called the *information effect.*

The model in Beaman (2012) describes further features and propositions of this model. However, we only require these two propositions to inform our intuition and the empirical strategy in the following section.

3.9.2. Empirical Strategy

These two propositions together imply that an increase in the size of a cohort negatively impacts the employment prospects of cohorts close to the larger cohort, but the now larger cohort eventually increases the probability of employment for cohorts who arrive sufficiently later, i.e. the larger cohort becomes an asset to the network when it's employment rate has increased sufficiently. This comes from the two competing effects, the competition effect and the information effect. The *competition effect* is negative and predicts that having a larger cohort decreases the probability of an individual obtaining outside employment. The *information effect* is the increase in the employment rate which comes from an increase in the likelihood of a referral due to the increase in network size. These two effects give rise to a testable implication.

$$y_{idct} = \alpha + \gamma_1 N_{idct} + \gamma_2 N_{idc(t-1)} + \gamma_3 N_{idc(t-2)} + X_{idct}\beta + \delta_{ct} + \phi_d + \epsilon_{idct} \quad (3.5)$$

Where y_{idct} is the employment status of individual i in dynasty d and caste c in time t. N_{idct} is the size of individual i's cohort. $N_{idc(t-k)}$ is the size of the cohort which arrived in the labour market k periods before individual i.

Having a larger cohort decreases the probability of an individual obtaining outside employment. This competition effect decreases with the distance between two cohorts. Instead the information effect of referrals begins to dominate. Thus the γ coefficients from estimating Equation 3.5 will be more positive for network cohorts which are further apart, that is for higher order γ_g's. The implication of the model is that $\gamma_3 > \gamma_2 > \gamma_1$ in the estimation of equation 3.5.

The model gives no prediction about when the positive information effect will begin to dominate, so multiple γ coefficients are consistent with the model. However, there should be a unique crossing from negative to positive.

Obviously these testable implications are derived from a model which assumes that the probability of learning about a job is constant and independent across agents. In the context of Palanpur this is possibly not a realistic assumption - it seems more likely that those individuals who work outside the village are more likely to learn about new job opportunities outside the village than those who live and work inside. However, if this was incorporated in the model it would simply mean a stronger information effect. The benefits of having an outside worker in your network would be greater, leading to greater concentration of outside jobs. This would not change the testable implication. The other issue is the assumption that all network members are identical. In our setting these network members are clearly related to each other in more subtle ways than simply being in the same network. One would reasonably believe that an agent would be more likely to refer a job an unemployed individual to whom they were more closely related than one to whom they were more distantly related. This information is available to us in the Palanpur context, and could potentially be used. However, the sample size is limited and does not permit an examination at a more disaggregated level for reasons of power.

When we take these testable implications to the Palanpur data we are faced with an additional problem. The Palanpur data, and therefore individual's occupations, have been collected only seven times over 60 years. This complication means that we don't necessarily observe the networks of individuals at the time of entry to the job market. However, we do observe the entire village and those who migrate out. This enables us to generate networks for an individual at any age. As a result we don't use an individual's

actual entry into the job market as the definition of their cohort, rather we use the individuals within 4 years of their age in our sample when they first appear in a survey as an adult. We calculate the networks at their first appearance as an adult since we believe this is most representative of the cohort with which he joined the labour market[35] We calculate the older cohort variables by counting the number of males within an villager's network (caste or dynasty) who are within a range of ages older than the villager on the villagers first appearance in the data as an adult. All males in these cohorts are counted, whether they have left Palanpur or not in any survey. All older cohorts are made such that they do not overlap.

We use a linear probability model to estimate equation 3.5 to allow us to include fixed effects. We know that probit specifications are unreliable for large numbers of fixed effects, see Greene (2004).

3.9.3. Outside Employment Results

Tables 3.12 and 3.13 show the results of a set of Beaman (2012) specifications with outside employment as the dependent variable for caste and dynasty networks respectively. The independent variables used are individual-level competition network variables which represent the number of male network members within their dynasty who are within 4 years of certain ages, all defined relative to the individuals age. These relative ages are 0, 9, 18, 27 and 36 years difference in Tables 3.12 and 3.13. They are defined in this way so that there is no intersection between these variables, that is a network member cannot be in 2 of these groups.

[35] In theory we could construct this network for an individual at any age, e.g. 18. In practice it makes almost no difference since we include all out-migrants, the only differences would be created by the death of network members between an individual being 18 and their first appearance as an adult. Deaths in late childhood and early adulthood are, thankfully, very rare.

Table 3.12.: Outside Work Beaman Caste Regressions

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside	(7) Work Outside
Network members in Cohort	−0.00261	−0.00200	−0.00224	0.00147	−0.00193	0.0000575	0.00189
	(0.00304)	(0.00292)	(0.00313)	(0.00313)	(0.00301)	(0.00304)	(0.00326)
Cohort $5 \leqslant age \leqslant 13$ above	0.00767**	0.00619*	0.00818**	0.0102***	0.00632*	0.0103***	0.0112***
	(0.00363)	(0.00346)	(0.00365)	(0.00339)	(0.00357)	(0.00326)	(0.00363)
Cohort $14 \leqslant age \leqslant 22$ above	0.00517	0.00425	0.00326	0.00130	0.00434	0.00540	0.00154
	(0.00476)	(0.00466)	(0.00483)	(0.00509)	(0.00481)	(0.00480)	(0.00511)
Cohort $23 \leqslant age \leqslant 31$ above	0.000774	0.0000261	0.000402	−0.00490	−0.0000389	−0.00375	−0.00478
	(0.00536)	(0.00543)	(0.00531)	(0.00512)	(0.00560)	(0.00524)	(0.00517)
Cohort $32 \leqslant age \leqslant 40$ above							−0.00763
							(0.00865)
Year FE	Yes	Yes	No	No	Yes	Yes	No
Caste FE	Yes	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	No	Yes	No
Caste-Year FE	No	No	Yes	No	No	No	No
Dynasty-Year FE	No	No	No	Yes	No	No	Yes
Dynasty-Year Trend	No	No	No	No	Yes	Yes	No
Observations	1904	1891	1896	1695	1891	1891	1695
Castes	15	14	13	13	14	14	13
Dynasties	128	115	125	105	115	115	105
Individuals	791	778	788	748	778	778	748
Mean Dep Var	0.251	0.252	0.252	0.265	0.252	0.252	0.265
R^2	0.151	0.251	0.204	0.434	0.440	0.507	0.434
P-Value of Older Cohorts	0.119	0.260	0.109	0.030	0.273	0.020	0.050

Note: Dependent variable is dummy variable for working in outside village. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the caste level. Cohorts are calculated at an individual's first appearance as an adult in the data.

The results for caste in Table 3.12 show that the effect of an individual's network is negative, albeit statistically insignificant, in 4 of the 7 specifications, and insignificantly positive in the others. These results are suggestive of the negative competition effect for caste networks. The earlier caste cohorts are positively associated with outside employment on the whole. This association is strongest with the cohort who are between 5 and 13 years older - which is significantly different from zero at the 10% level at least in all specifications. However, earlier cohorts than this appear to have smaller positive associations with employment outside, and are statistically indistinguishable from zero. A one standard deviation increase in the members of all earlier cohorts (from ages 5-31 in regression 4) increases the probability of outside employment by 5.28 percentage points, which relative to the sample mean is a 19.92% increase.

Table 3.13.: Outside Work Beaman Dynasty Regressions

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside	(7) Work Outside
Network members in Cohort	0.00846	0.0110	0.00684	0.0151	0.0111	0.00869	0.0156
	(0.0149)	(0.0167)	(0.0156)	(0.0166)	(0.0172)	(0.0179)	(0.0168)
Cohort $5 \leqslant age \leqslant 13$ above	0.0113	0.0221*	0.0113	0.0196	0.0224*	0.0220*	0.0199
	(0.0126)	(0.0124)	(0.0128)	(0.0121)	(0.0128)	(0.0129)	(0.0120)
Cohort $14 \leqslant age \leqslant 22$ above	0.0497**	0.0570**	0.0507**	0.0574*	0.0573**	0.0653**	0.0563*
	(0.0207)	(0.0227)	(0.0202)	(0.0297)	(0.0234)	(0.0278)	(0.0296)
Cohort $23 \leqslant age \leqslant 31$ above	−0.0196	−0.0134	−0.0205	0.000708	−0.0130	0.0000814	−0.00130
	(0.0201)	(0.0193)	(0.0207)	(0.0189)	(0.0198)	(0.0185)	(0.0220)
Cohort $32 \leqslant age \leqslant 40$ above							−0.00828
							(0.0331)
Year FE	Yes	Yes	No	No	Yes	Yes	No
Caste FE	Yes	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	No	Yes	No
Caste-Year FE	No	No	Yes	No	No	No	No
Dynasty-Year FE	No	No	No	Yes	No	No	Yes
Dynasty-Year Trend	No	No	No	No	Yes	Yes	No
Observations	1904	1891	1896	1695	1891	1891	1695
Castes	15	14	13	13	14	14	13
Dynasties	128	115	125	105	115	115	105
Individuals	791	778	788	748	778	778	748
Mean Dep Var	0.251	0.252	0.252	0.265	0.252	0.252	0.265
R^2	0.157	0.259	0.211	0.437	0.446	0.511	0.437
P-Value of Older Cohorts	0.064	0.017	0.049	0.155	0.021	0.065	0.256

Note: Dependent variable is dummy variable for working in outside village. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.

Table 3.13 contains the dynasty level results of the same specifications. The pattern of coefficients is much more consistent with the implications of the model than those for caste networks. The sizes of the coefficients for the 14-22 years older cohort are larger than those for the 5-13 years older cohort. These earlier cohorts are jointly significant and positively associated with increased outside employment probability in 5 of the 7 specifications, 3 at the 5% significance level, 2 at 10%. A one standard deviation increase in the size of all older cohorts within a dynasty (ages 5-31 in specification 4) leads to a 7.27 percentage point increase in the outside employment probability. This is 27.44% increase relative to the sample mean.

When we construct these competition variables for different competition sizes the pattern that emerges is qualitatively similar. Network members who are between 12 and 20 years older than the individual in question seem to have the greatest impact on the probability of employment outside the village[36]. Appendix Tables C.4, C.5 and C.6 demonstrate the robustness of these results to different sizes of cohort.

However, in the Table 3.13 specifications we do not see the negative competition effect present for the individual's own cohort - instead the coefficients are all insignificantly positive. One possible explanation for this pattern is that, as shown in section 3.8.1, the short run job network effects appear to be much stronger for dynasties than castes. So this, combined with the feature of the Palanpur data whereby individual employment may be observed long after the job has been obtained, could mean that the point at which larger cohorts become beneficial to the network is very early after their entry into the labour market. Another piece of evidence consistent with this explanation is illustrated when an individual's cohort is split into 2 parts, older and younger than the individual. Table 3.14 performs this test by including an additional variable which is the

[36]Remember that this is not only the effect of fathers, all male dynasty members are included.

Table 3.14.: Outside Employment Beaman Dynasty Regressions - Split Cohort

	(1) Work Outside	(2) Work Outside	(3) Work Outside
Network members in Cohort	−0.000236	−0.000169	0.00259
	(0.0229)	(0.0236)	(0.0238)
Older Members in Cohort	0.0219	0.0219	0.0247
	(0.0282)	(0.0291)	(0.0283)
Cohort $5 \leqslant age \leqslant 13$ above	0.0218*	0.0221*	0.0198*
	(0.0122)	(0.0126)	(0.0118)
Cohort $14 \leqslant age \leqslant 22$ above	0.0577**	0.0580**	0.0587**
	(0.0222)	(0.0230)	(0.0289)
Cohort $23 \leqslant age \leqslant 31$ above	−0.0124	−0.0120	0.00244
	(0.0194)	(0.0200)	(0.0189)
Year FE	Yes	Yes	No
Dynasty FE	Yes	No	No
Dynasty-Year FE	No	No	Yes
Dynasty-Year Trend	No	Yes	No
Observations	1891	1891	1695
Castes	14	14	13
Dynasties	115	115	105
Individuals	778	778	748
Mean Dep Var	0.252	0.252	0.265
R^2	0.259	0.446	0.438
P-Value of Older Cohorts	0.013	0.016	0.116

Note: Dependent variable is dummy variable for working in outside village. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.

count of all of the male dynasty members who are between 0 and 4 years older than the individual. This means the original network cohort variable now shows the effect of those network members who are younger than the individual, and the effect of older cohort members is the sum of the two coefficients. We see that the effect of younger members is much smaller than the effect of older cohort members, and is negative for two of the three specifications. This result is consistent with the existence of job networks in outside employment since younger members likely move into the job market after the individual in question, therefore they are not in an individual's network of contacts when they join, thus the only effect from younger members should be the competition effect. However, members of a cohort who are just older may also be able to refer an individual to a job via the information effect. If we take these coefficients seriously then what we see is that in the Palanpur context the information effect is much stronger than the competition effect. This helps to explain why in most specifications all the coefficients are positive and there is no crossing of zero where the information effect overwhelms the competition effect. Given the size of the village relative to the job markets where most of this outside employment takes place it is perhaps not surprising that the effect of one extra competitor in ones network is not that important on the margin.

One feature of all the results seems to be that cohorts which arrive substantially earlier have a smaller impact on outside employment rates, and in some cases negative although insignificant[37]. Empirically this seems to happen for cohorts who are more than 25 years older. We hypothesise that it is plausible that network members who arrive in the job market considerably earlier than an individual will have left the job market by the time they are needed to help a new network member get a job, or that it is possible that their

[37]Interestingly this feature also seems to be present to some extent in the results of Beaman (2012). Although in her context this occurs for cohorts who arrive 3 to 4 years earlier, rather than some 25 years in our context.

own connections are no longer useful due to the changing nature of outside employment in this context of structural transformation. Perhaps a more complicated model would predict a hump-shaped pattern of this type. However, the simple model we use abstracts from such possibilities.

As a robustness test we ran the regressions in Table 3.12 and 3.13 with a dummy for regular employment as the dependent variable. Given the evidence on the lack of network effects for regular employment we have shown for short run networks we do not expect a pattern consistent with the model to emerge from these regressions. Reassuringly when we do so there is only one coefficient which is significantly different from zero at the 10% level of significance, out of 29 cohort coefficients. Plus the coefficients are largely inconsistent with the pattern predicted by the testable implication of the model[38].

One concern with this approach is that there is no information on in or out migration from Palanpur before the 1958 survey. So accurate cohorts cannot be obtained for those who would have entered the labour market before this period. However, out-migration rates for young males in this time period was low. As we show in Section 1.6, the percentage of 10-24 year old males migrating out of Palanpur between 1958 and 1975 is 14.7%. This migration rate is increasing, and is thus likely to have been lower before 1958. Plus this is migration over a long time period, an equivalent is to 14.7% of males migrating out between 1941 and 1958 - which, even if we assume no population growth between these two years is only approximately 12 individuals. This estimate is an upper bound given the likely presence of population growth and an increasing migration trend. Even so, spread over each of the castes or dynasties this small number of individuals is unlikely to make much difference empirically. Additionally, when we replicate the results from this section using data only from the 1975 survey onwards the qualitative results remain

[38]These tables are contained in the appendix, as Table C.7 and Table C.8

the same.

The results from these two robustness tests are reassuring. However, the limitation of our setting remains that we do not have a candidate for a natural experiment, to provide exogenous variation in cohort sizes. Instead we rely on fixed effects to help us remove any potential endogeneity. Our data permit caste, dynasty, year, caste-year and dynasty-year fixed effects. They do not permit individual fixed effects because we are using cohorts at a given age, which do not vary within an individual. Reassuringly, the exact combination of fixed effects we use does not seem to change the qualitative nature of our estimates. In our most restrictive specification we use dynasty-year fixed effects, thus we are comparing the outside employment outcomes of males from within the same dynasty who appear in the same survey at different ages and thus have different cohort sizes. Therefore if there is variation in individual characteristics which make one more likely to undertake outside employment, e.g. ability, and this is correlated positively (negatively) with older (current) cohort sizes conditional on current (older) cohort sizes within dynasty in the same survey, our results would be affected by this endogeneity. It is difficult to think of a plausible mechanism through which you could get this set of correlations, but we cannot rule it out.

3.10. Migration

Migration is a common feature of Palanpur. Section 1.6 demonstrated the patterns of migration out of Palanpur over the survey period and showed they are consistent with migration rates across rural India. However, the pattern of migration is not the same across castes or dynasties. The proportion of males in castes who migrate away from the village at some point in their life ranges from 19.5% to 100%, for dynasties this range

is $0 - 100\%$[39].

We hypothesise that the networks in outside employment that we have shown operate over many decades in Palanpur may have long run effects on individual and dynasty outcomes through their effect on migration patterns. This hypothesis is supported by some descriptive statistics from the village - a male in Palanpur is 6.5 percentage points more likely to be an out-migrant if their dynasty network contained an outside worker two surveys previously, this is a 32% increase[40].

Such an empirical feature may occur through a number of mechanisms. One potential mechanism is that observing closely a network member who is not very different from you may alter an individual's *perceived* returns to migration through learning about wage rates and types of work or accommodation away from the village. It is ultimately these *perceived* returns upon which individual migration decisions are based. This information may also change individuals *perceived* returns to education[41]. There are obviously a number of other candidate explanations, for example effects on aspirations or expectations from seeing achievements by those similar to you as demonstrated by Jensen (2012) and Jensen and Oster (2009). Unfortunately we will not be able to distinguish these mechanisms from each other in this chapter[42].

An example network mechanism which could lead to higher levels of migration in dynasties with larger numbers of outside workers is that individuals who work outside gain contacts from outside the village through work or travel, these new contacts expand

[39] Mean=50.2, Std. Dev. 36.2 for dynasties. Mean=58.7, Std. Dev. 31.6 for castes.

[40] This feature is robust to larger distances between surveys.

[41] See Jensen (2010) for a paper which demonstrates the importance of perceived returns to schooling.

[42] Akram et al. (2017) show that migrants who know migrants induced to migrate by their RCT in Bangladesh are also more likely to migrate temporarily in the same season, and this effect does not exist when households do not know migrants who received their treatment. This is an interesting result, however since it is concerned with contemporaneous co-migration it is more likely to be about mitigating the risk of temporary seasonal migration in their setting.

the network through which dynasties can obtain referrals - and the jobs being referred may be more diverse. Consistent with this if a male was employed outside the village two surveys previously they are 13.85 percentage points more likely to have migrated away from the village (a 58.57% increase).

The census of households in Palanpur includes asking about births and deaths of household members in the period between surveys. Households are also asked about individuals living outside the village who are related to the household head. From this we can impute which individuals migrate out of Palanpur. We do not observe the occupation of all migrants. However, we can use migration as an outcome. We cannot show that migrants are better off, however, there is a wealth of evidence (Munshi and Rosenzweig (2016); Morten (forthcoming)) that migration is inefficiently low in this setting and migration is welfare improving on average (Bryan et al. (2014); Garlick et al. (2016)).

To explore these long run effects we again create cohort variables, but this time we use counts of those individuals within a dynasty who ever hold outside employment within 8 year cohorts. We run regressions of the following form:

$$m_{idct} = \alpha + \gamma_1 N_{idct} + \gamma_2 N_{idc(t-1)} + \gamma_3 N_{idc(t-2)} + X_{idct}\beta + \delta_{ct} + \phi_d + \epsilon_{idct} \quad (3.6)$$

Where m_{idct} is a dummy variable for the out-migration status of individual i in dynasty d and caste c in time t. N_{idct} is the count of individuals within i's cohort who ever hold outside employment. $N_{idc(t-k)}$ is the count of individuals who ever hold outside employment in the cohort which arrived in the labour market k periods before individual i. We focus on individuals who have ever held outside employment since our focus is on the long run effects of these networks, we are less interested in a contemporaneous

effect. Ideally we would include counts of individuals within a cohort who have worked outside between an individual's birth and date of migration (or date of survey). However, our data does not permit us to determine the exact date of migration and employment outside for all individuals. The use of ever holding outside employment as a proxy is likely to lead to increased noise, sadly this is unavoidable.[43]

If one of the candidate mechanisms discussed above is present then we would expect to see a positive correlation between the presence of older network members in outside employment and migration of younger members - that is positive values for γ_2, γ_3. We expect γ_1 to also be positive in our setting, because Palanpur is small relative to other labour markets and towns to which individuals would migrate, there is unlikely to be a competition effect in migration[44]. The size of γ_1 is ambiguous. If we could use our preferred measure we would expect $\gamma_2 > \gamma_1 > 0$, since any cohort members employed outside would likely have shorter periods of outside employment over which to alter perceived returns. Given we are using our time-invariant proxy this may not be true.

Table 3.15 reports the results of regressions based on equation 3.6. We see that in the most restrictive specifications, with dynasty-year fixed effects, individuals with more outside workers older than them in their dynasty are significantly more likely to migrate out of the village. This association is strongest for cohorts between 5 and 13 years older. In contrast, the number of own cohort members who obtain outside employment is not significantly correlated with individual level migration decisions in any specification, and appears both negatively and positively. Columns 8 and 9 show that our results are unchanged by controlling for individuals who achieve high levels of education.

[43] We include only adult males since we are interested in those who migrate for work, many of those who migrate before the age of 18 are undoubtedly part of the movement of whole households.

[44] In fact Akram et al. (2017) show a complementarity in migration decisions in their study. They argue that this co-migration helps to mitigate the risk inherent in temporary seasonal migration.

As shown in the previous section, these dynasties create outside workers through network effects and here there is a positive association between outside workers within a dynasty and migration of younger generations. We believe this is suggestive evidence that these dynasty level job networks which help individuals obtain jobs also have long run effects because they induce dynasty members to migrate out of the village. Given the existing evidence on the returns to migration in the Indian context, this suggests that these job networks can lead to long run increases in inequality across dynasties.

Despite the absence of a source of clearly exogenous variation and a small sample size these results demonstrate an interesting long run consequence of job networks. It is unlikely that this is the only consequence of these networks, and we think this is a previously unexplored research area worthy of further work.

Table 3.15.: Migration Effects of Outside Employment

	(1) Migrated	(2) Migrated	(3) Migrated	(4) Migrated	(5) Migrated	(6) Migrated	(7) Migrated	(8) Migrated	(9) Migrated
outside workers in Cohort	−0.00961 (0.0179)	−0.000304 (0.0175)	−0.00944 (0.0182)	0.0237 (0.0174)	−0.0000297 (0.0180)	0.0133 (0.0172)	0.0242 (0.0176)	0.0243 (0.0174)	0.0138 (0.0171)
Cohort workers $5 \leqslant age \leqslant 13$ above	−0.00493 (0.0202)	0.00615 (0.0205)	0.00308 (0.0199)	0.0496** (0.0203)	0.00658 (0.0211)	0.0313 (0.0204)	0.0497** (0.0204)	0.0465** (0.0199)	0.0280 (0.0199)
Cohort workers $14 \leqslant age \leqslant 22$ above	0.00323 (0.0202)	0.00513 (0.0171)	0.00931 (0.0197)	0.0259 (0.0177)	0.00538 (0.0175)	0.0174 (0.0181)	0.0261 (0.0180)	0.0229 (0.0181)	0.0140 (0.0184)
Cohort workers $23 \leqslant age \leqslant 31$ above	−0.0100 (0.0264)	−0.0222 (0.0240)	−0.0108 (0.0258)	−0.00802 (0.0214)	−0.0222 (0.0246)	−0.0216 (0.0218)	−0.00668 (0.0217)	−0.00553 (0.0224)	−0.0192 (0.0225)
Cohort workers $32 \leqslant age \leqslant 40$ above							0.0191 (0.0362)		
Highly Educated								−0.0731** (0.0294)	−0.0807*** (0.0303)
Year FE	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Caste FE	Yes	No	No	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	No	Yes	No	No	Yes
Caste-Year FE	No	No	Yes	No	No	No	No	No	No
Dynasty-Year FE	No	No	No	Yes	No	No	Yes	Yes	No
Dynasty-Year Trend	No	No	No	No	Yes	Yes	No	No	Yes
Observations	2362	2361	2350	2095	2361	2361	2095	2095	2361
Castes	17	17	14	13	17	17	13	13	17
Dynasties	130	129	127	107	129	129	107	107	129
Individuals	826	825	823	788	825	825	788	788	825
Mean Dep Var	0.203	0.203	0.203	0.200	0.203	0.203	0.200	0.200	0.203
R^2	0.166	0.313	0.214	0.491	0.454	0.543	0.491	0.495	0.547
P-Value of Older Cohorts	0.977	0.784	0.910	0.056	0.792	0.235	0.095	0.091	0.328

Note: Dependent variable is dummy variable for migration. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data. The Cohorts here are a count of all individuals in dynasty within an age range relative to one individual who have ever held outside employment

3.11. Conclusion

In this chapter we have shown evidence of the presence of labour market networks in a village in rural India. These networks are important determinants of job attainment in the short run. Having a dynasty network contact employed outside the village leads to a 70% increase in ones own likelihood of working outside the village. However, the structure of ones familial network is an important determinant of individual outcomes. Those with an additional competitor within their dynasty cohort are approximately 15% less likely to obtain an outside job in the short run. Using a subset of jobs for which we believe *a priori* job networks are less likely to determine employment we find smaller short run network effects and insignificant effects of competition consistent with this prior. These effects are large relative to a number of other studies of these job networks[45]. This is not particularly surprising given that most labour market network estimates are from developed countries, where formal sector employment rates are higher and other job finding institutions are better established. Unlike other studies of network effects we exploit the uniqueness of the Palanpur data to illustrate the dynamic effects of these networks on labour market outcomes over 60 years, during a time of rapid changes in the nature of labour within this context due to structural transformation. To the best of our knowledge there have not been any other studies of labour market networks over such a long time horizon. Magruder (2010) examines intergenerational networks in South Africa, but does so using a 3 year panel so his findings are more analogous to our short run findings with specified relationships rather than taking the network as a whole.

[45]The most directly comparable are those of Laschever (2013) who finds the effect of additional employed network member of 0.8 percentage points, while ours are an order of magnitude greater. However, it must be noted he examines employment generally and 93% of his sample are employed. In contrast, less than 20% of our sample is employed outside the village.

We use a model from Beaman (2012) to explore these dynamics and inform our empirical strategy. In her original paper she uses this framework to examine labour market dynamics in a 5 year panel. We have extended this analysis to much longer time series and shown that it remains a useful framework with which to study these dynamics. We find that dynasty and caste networks exhibit both the competition effect from network members as well as the information effect of job referrals. However, the referral contacts who are important for individual employment prospects seem to differ for the two networks.

We also use this framework to illustrate that these dynasty networks have long run consequences for their members. Younger male members of those dynasties with more outside workers are more likely to migrate away from Palanpur. We know from the migration literature that migrants experience large increases in income and consumption on average. Thus these dynastic labour market networks lead to divergent outcomes for future dynasty members, and are as such a source of persistent inequality.

Our work also relates closely to the work of Munshi and Rosenzweig (2006); Munshi and Rosenzweig (2016); Munshi (2011); Anderson (2011) who demonstrate the importance of caste networks in India. We find evidence consistent with these papers, that castes have important labour market network effects. However, we also find that focusing just on castes as a source of networks in an Indian context will overlook networks with even stronger social ties and larger labour market network effects. There appear to be important differences in the dynamics of these labour market job networks for castes and dynasties. Caste networks appear to lose their value as a link to obtain employment more quickly than do dynasty networks. Also, the job information sharing effects appear to be much stronger for dynasties than castes.

While we have exploited various fixed effect estimation strategies to try and reassure

that we are capturing the true effect of job networks, we cannot hope to control for all possible sources of unobserved individual time-varying characteristics and this remains a limitation of our study. We do not have a source of explicitly exogenous variation to exploit in this setting. However, we believe the benefits to estimating the effects of employment networks over the long run in a previously understudied area of rural labour markets within a developing country outweigh these potential endogeneity concerns.

Palanpur is a well connected village, it is not clear whether these effects would be the same in a village which did not have access to transport allowing for commuting. One would imagine that where commuting is more costly (in either a monetary or temporal sense) it would be less common. This may mean that network effects take longer to become established. Although it is conceivable that these similar network referral effects may operate in the more long distance temporary migration which is common across India (Morten (forthcoming)), this is certainly an interesting avenue for further research. However, that is not to say that Palanpur's experience is unique. Travelling by train in India one sees countless villages near the railway line. Additionally, more and more villages are being connected to the expanding road network (Datta (2012)).

Another interesting question raised by our work is the mechanism by which this exposure to outside employment may lead to increased migration. Candidate explanations include the role of aspirations or expectations of the kind described by Jensen (2012) and Jensen and Oster (2009), revealed information about the returns to education as in Jensen (2010), or simply an extension of the possible network from which they can now receive job information - whether this be work colleagues or extended social network contacts in the place of work.

There are avenues for extension of the work within this chapter. One could try to exploit the range or density of an individual's network. For example, examining the

role of the proportion of weak versus strong ties within networks (in the spirit of Montgomery (1992)). Alternatively, using the literal range of networks - those members who migrate away to other areas, Moradabad versus small nearby villages may have different effects[46]. However, the obvious problem with these extensions is the available sample size, there are only so many castes or dynasties within Palanpur.

References from Chapter 3

Akram, A.A., Chowdhury, S. and Mobarak, A.M., October 2017. *Effects of emigration on rural labor markets*. National Bureau of Economic Research (Working Paper Series, Working Paper 23929).

Anderson, S., January 2011. Caste as an impediment to trade. *American Economic Journal: Applied Economics*, 3(1), pp.239–63.

Angelucci, M., Giorgi, G.D., Rangel, M.A. and Rasul, I., 2010. Family networks and school enrolment: evidence from a randomized social experiment. *Journal of Public Economics*, 94(3-4), pp.197–221.

Ansari, N., 1964. Palanpur: a study of its economic resources and economic activities. *Continuous Village Survey*, (41).

Bandiera, O. and Rasul, I., 2006. Social networks and technology adoption in northern mozambique*. *The Economic Journal*, 116(514), pp.869–902.

Banerjee, A. and Munshi, K., 2004. How efficiently is capital allocated? evidence from the knitted garment industry in tirupur. *The Review of Economic Studies*, 71(1), pp.19–42.

[46]There is some existing evidence (e.g. Campbell et al. (1986)) that those with more geographically dispersed networks have higher socio-economic statuses

Bayer, P., Ross, S.L. and Topa, G., 2008. Place of work and place of residence: informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116(6), pp.1150–1196.

Beaman, L.A., 2012. Social networks and the dynamics of labour market outcomes: evidence from refugees resettled in the u.s. *The Review of Economic Studies*, 79(1), pp.128–161.

Becker, G.S. and Tomes, N., 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6), pp.1153–1189.

Béteille, A., 1996. Varna and jati. *Sociological Bulletin*, 45(1), pp.15–27.

Bliss, C.J. and Stern, N.H., 1982. *Palanpur: the economy of an indian village*. Oxford University Press.

Bryan, G., Chowdhury, S. and Mobarak, A.M., 2014. Underinvestment in a profitable technology: the case of seasonal migration in bangladesh. *Econometrica*, 82(5), pp.1671–1748.

Calvó-Armengol, A. and Jackson, M.O., June 2004. The effects of social networks on employment and inequality. eng. *The American Economic Review*, 94(3), pp.426–454.

Campbell, K.E., Marsden, P.V. and Hurlbert, J.S., 1986. Social resources and socioeconomic status. *Social Networks*, 8(1), pp.97–117.

Chauvin, J.P., Glaeser, E., Ma, Y. and Tobio, K., March 2017. What is different about urbanization in rich and poor countries? cities in brazil, china, india and the united states. eng. *Journal of Urban Economics*, 98, pp.17–49.

Datta, S., 2012. The impact of improved highways on indian firms. *Journal of Development Economics*, 99(1), pp.46–57.

Garlick, J., Leibbrandt, M. and Levinsohn, J., June 2016. *Individual migration and household incomes*. National Bureau of Economic Research (Working Paper Series, Working Paper 22326).

Gaurav, S., November 2013. *Comparability of cultivation income in palanpur across the study rounds for panel data construction*. LSE India Observatory.

Government of India, 2011. *Census of india 2011*.

Greene, W., 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal*, 7(1), pp.98–119.

Herrendorf, B. and Schoellman, T., April 2018. Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics*, 10(2), pp.1–23.

Himanshu, Bakshi, I. and Dufour, C., 2011b. *Poverty, inequality and mobility in palanpur: observations from palanpur*. LSE Asia Research Centre, (Working Paper 45).

Himanshu, Joshi, B. and Lanjouw, P., June 2016. Non-farm diversification, inequality and mobility in palanpur. *Economic and Political Weekly*, 51(26-27), pp.43–51.

Himanshu, Lanjouw, P., Murgai, R. and Stern, N., September 2013. Nonfarm diversification, poverty, economic mobility, and income inequality: a case study in village india. *Agricultural Economics*, 44(4-5), pp.461–473.

Jensen, R., 2010. The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics*, 125(2), pp.515–548.

Jensen, R., 2012. Do labor market opportunities affect young women's work and family decisions? experimental evidence from india. *The Quarterly Journal of Economics*, 127(2), pp.753–792.

Jensen, R. and Oster, E., 2009. The power of tv: cable television and women's status in india. *The Quarterly Journal of Economics*, 124(3), pp.1057–1094.

Lanjouw, P. and Stern, N., June 1998. *Economic Development in Palanpur over Five Decades*, OUP Catalogue 9780198288329. Oxford University Press.

Laschever, R., June 2013. *The doughboys network: social interactions and labor market outcomes of world war I veterans*. Unpublished manuscript.

Loury, L.D., 2006. All in the extended family: effects of grandparents, aunts, and uncles on educational attainment. *The American Economic Review*, 96(2), pp.275–278.

Luke, N. and Munshi, K., 2006. New roles for marriage in urban africa: kinship networks and the labor market in kenya. *The Review of Economics and Statistics*, 88(2), pp.264–282.

Magruder, J.R., 2010. Intergenerational networks, unemployment, and persistent inequality in south africa. *American Economic Journal: Applied Economics*, 2(1), pp.62–85.

Mazzocco, M. and Saini, S., 2012. Testing efficient risk sharing with heterogeneous risk preferences. *The American Economic Review*, 102(1), pp.428–468.

Montgomery, J.D., 1992. Job search and network composition: implications of the strength-of-weak-ties hypothesis. *American Sociological Review*, 57(5), pp.586–596.

Moorjani, P., Thangaraj, K., Patterson, N., Lipson, M., Loh, P.-R., Govindaraj, P., Berger, B., Reich, D. and Singh, L., 2013. Genetic evidence for recent population mixture in india. *The American Journal of Human Genetics*, 93(3), pp.422–438.

Morten, M., forthcoming. Temporary migration and endogenous risk sharing in village india. *Journal of Political Economy*.

Mukhopadhyay, A., 2011. *Stepping out of palanpur: employment outside palanpur*. LSE Asia Research Centre, (Working Paper 46).

Munshi, K., 2003. Networks in the modern economy: mexican migrants in the u. s. labor market. *The Quarterly Journal of Economics*, 118(2), pp.549–599.

Munshi, K., 2011. Strength in numbers: networks as a solution to occupational traps. *The Review of Economic Studies*, 78(3), pp.1069–1101.

Munshi, K., November 2014. Community networks and the process of development. *Journal of Economic Perspectives*, 28(4), pp.49–76.

Munshi, K. and Rosenzweig, M., 2006. Traditional institutions meet the modern world: caste, gender, and schooling choice in a globalizing economy. *American Economic Review*, 96(4), pp.1225–1252.

Munshi, K. and Rosenzweig, M., January 2016. Networks and misallocation: insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1), pp.46–98.

Simons, G.F. and Fennig, C.D., 2018. *Ethnologue: languages of the world, twenty-first edition*. Online version: `http://www.ethnologue.com`. Dallas, Texas.

Srinivas, M.N., 1962. Varna and caste. *Caste in Modern India and Other Essays*, pp.63–9.

Srinivas, M.N., 1976. *The remembered village*, 26. University of California Press. ISBN: 0520029976.

Surendra, V., November 2013. *Non-farm and total income: comparability across rounds*. LSE India Observatory.

Topa, G., 2011. Chapter 22 - labor markets and referrals. In: Benhabib, J., Bisin, A. and Jackson, M.O. eds. Vol. 1, Handbook of Social Economics. North-Holland, pp.1193–1221.

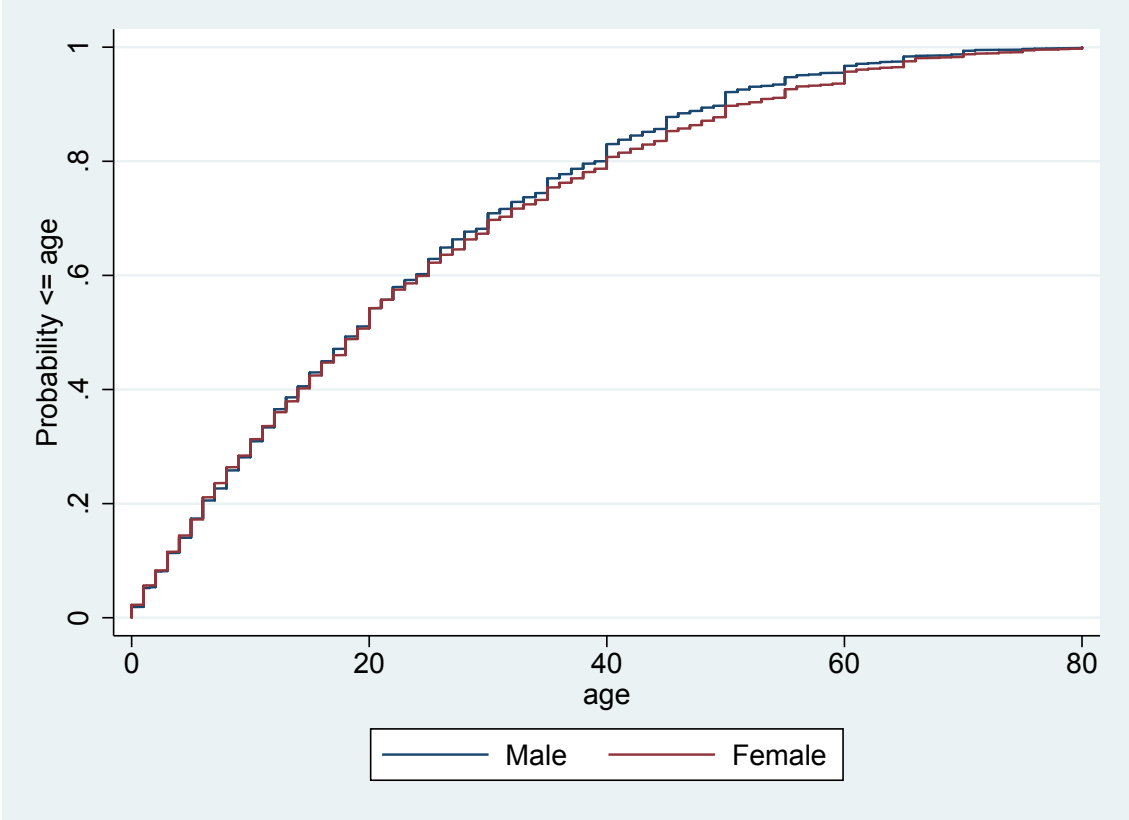
Wang, S.-y., 2013. Marriage networks, nepotism, and labor market outcomes in china. *American Economic Journal: Applied Economics*, 5(3), pp.91–112.

Chapter C

Appendix to Chapter 3

C.1. Appendix Figures

Figure C.1.: CDF of Age by Sex



Note: All individuals present in village included

Source: Author's Calculations

C.2. Caste Descriptives

Table C.1.: Mean Real Income Per Capita by Caste-Year

	1958	1964	1975	1984	2009	Total
Thakur	203.4	245.9	326.5	266.3	474.3	333.6
Murao	280.2	252.7	332.8	330.8	456.1	353.7
Dhimar	121.3	97.71	221.6	210.7	343.0	223.7
Gadaria	195.5	228.4	278.2	220.7	502.8	308.4
Dhobi	249.1	802.0	157.3	196.0	162.4	187.8
Teli	112.2	114.3	237.8	228.6	538.8	308.9
Passi	191.8	260.5	263.8	228.3	315.5	242.9
Jatab	152.9	129.2	208.5	117.6	298.3	203.8
Lohar	254.2	.	.	.	.	254.2
Nai	63.43	61.66	.	101.4	197.0	105.1
Kayasth	167.1	233.2	251.3	290.5	544.0	277.6
Khati	28.97	29.15	.	.	.	29.08
Carpenter	.	.	171.3	.	.	171.3
Balmiki	.	227.6	204.4	163.6	0	118.7
Brahmin	.	.	.	.	.	.
Unknown	.	.	.	.	872.0	872.0
Village Average	190.9	204.5	278.2	241.9	421.2	291.4

Note: Values are mean per capita annual incomes of households, they are in 1960-61 Rupees. Income is not available for the surveys in 1993 and 2015. Members of the Brahmin caste were only present in 1993, when no income data were gathered.

Source: Author's Calculations

C.3. Network Descriptives

continued overleaf

Table C.2.: Competitive Network Descriptives - Caste

Panel A	Caste							
	1958	1964	1975	1984	1993	2009	2015	All
Mean age difference	1.47	1.30	0.93	0.74	0.65	0.48	0.45	0.73
Median age difference	0	0	0	0	0	0	0	0
Mean % with competitor within 1 years	81.8	78.6	85.7	88.8	90.6	92.3	92.6	88.9
Mean: count competitors within 1 years	2.77	2.86	4.14	4.88	6.34	7.03	7.75	5.76
Median % with competitor within 1 years	100	100	100	100	100	100	100	100
Median: count competitors within 1 years	3	2	3	4	4	6	7	4
Mean % with competitor within 2 years	87.2	87.3	90.2	93.7	94.8	96.2	96.6	93.6
Mean: count competitors within 2 years	4.40	4.75	6.83	8.01	10.6	11.5	12.7	9.48
Median % with competitor within 2 years	100	100	100	100	100	100	100	100
Median: count competitors within 2 years	4	4	5	6	8	11	11	7
Mean % with competitor within 3 years	89.3	89.7	91.4	94.5	96.2	97.5	97.7	94.9
Mean: count competitors within 3 years	5.89	6.47	9.43	11.1	14.6	16.1	17.7	13.1
Median % with competitor within 3 years	100	100	100	100	100	100	100	100
Median: count competitors within 3 years	5	6	7	9	11	14	15	10
Mean % with competitor within 4 years	89.6	91.0	92.5	96.0	96.6	97.5	97.7	95.4
Mean: count competitors within 4 years	7.42	8.09	11.9	14.0	18.4	20.3	22.3	16.6
Median % with competitor within 4 years	100	100	100	100	100	100	100	100
Median: count competitors within 4 years	6	7	10	12	14	17	21	13

Note: All Males in migration sample included. The Mean and Median age difference refer to the age difference between an individual and the closest male member in their network in the migration sample.

Source: Author's Calculations

Table C.3.: Competitive Network Descriptives - Dynasty

Panel B	Dynasty							
	1958	1964	1975	1984	1993	2009	2015	All
Mean age difference	12.8	12.8	9.37	7.21	5.66	4.36	3.90	6.73
Median age difference	7	7	5	4	3	2	2	3
Mean % with competitor within 1 years	8.06	7.67	12.9	19.5	27.7	33.9	37.9	25.1
Mean: count competitors within 1 years	0.090	0.080	0.13	0.26	0.35	0.45	0.55	0.33
Median % with competitor within 1 years	0	0	0	0	0	0	0	0
Median: count competitors within 1 years	0	0	0	0	0	0	0	0
Mean % with competitor within 2 years	22.4	21.4	28.4	36.1	44.6	53.4	57.8	42.5
Mean: count competitors within 2 years	0.27	0.26	0.41	0.51	0.78	0.96	1.14	0.73
Median % with competitor within 2 years	0	0	0	0	0	100	100	0
Median: count competitors within 2 years	0	0	0	0	0	1	1	0
Mean % with competitor within 3 years	29.9	30.4	41.8	46.9	56.3	64.5	67.9	53.3
Mean: count competitors within 3 years	0.38	0.40	0.66	0.78	1.17	1.39	1.61	1.07
Median % with competitor within 3 years	0	0	0	0	100	100	100	100
Median: count competitors within 3 years	0	0	0	0	1	1	1	1
Mean % with competitor within 4 years	33.4	36.0	46.9	57.4	64.4	72.2	74.7	60.4
Mean: count competitors within 4 years	0.46	0.55	0.87	1.07	1.53	1.81	2.07	1.40
Median % with competitor within 4 years	0	0	0	100	100	100	100	100
Median: count competitors within 4 years	0	0	0	1	1	1	2	1

Note: All Males in migration sample included. The Mean and Median age difference refer to the age difference between an individual and the closest male member in their network in the migration sample.

Source: Author's Calculations

C.4. Robustness of Long Run Network Effects

continued overleaf

Table C.4.: Robustness of Long Run Effects to Alternative Network Sizes

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside	(7) Work Outside	(8) Work Outside	(9) Work Outside
Members in 1 year Cohort	0.0472	0.0458	0.0397	0.0457	0.0298	0.0444	0.0479	0.0492	0.0484
	(0.0326)	(0.0314)	(0.0324)	(0.0323)	(0.0326)	(0.0328)	(0.0321)	(0.0319)	(0.0330)
Cohort $2 \leqslant age \leqslant 4$ above	0.00575	0.00804	0.00659	0.00812	0.00426	0.0106	0.0154	0.0159	0.0161
	(0.0248)	(0.0250)	(0.0250)	(0.0258)	(0.0276)	(0.0247)	(0.0246)	(0.0246)	(0.0234)
Cohort $5 \leqslant age \leqslant 7$ above	0.0272	0.0325	0.0157	0.0324	0.0232	0.0187	0.0237	0.0262	0.0303
	(0.0248)	(0.0258)	(0.0274)	(0.0267)	(0.0272)	(0.0271)	(0.0264)	(0.0265)	(0.0259)
Cohort $8 \leqslant age \leqslant 10$ above	0.00505	−0.000843	−0.0164	−0.000674	−0.0116	−0.0172	−0.0217	−0.0207	−0.0161
	(0.0246)	(0.0209)	(0.0195)	(0.0216)	(0.0214)	(0.0194)	(0.0196)	(0.0199)	(0.0218)
Cohort $11 \leqslant age \leqslant 13$ above						0.0455	0.0466*	0.0460	0.0486*
						(0.0289)	(0.0279)	(0.0278)	(0.0275)
Cohort $14 \leqslant age \leqslant 16$ above							0.0920***	0.0927***	0.0965***
							(0.0306)	(0.0317)	(0.0319)
Cohort $17 \leqslant age \leqslant 19$ above								0.0333	0.0359
								(0.0461)	(0.0469)
Cohort $20 \leqslant age \leqslant 22$ above									0.0506
									(0.0392)
Year FE	Yes	Yes	No	Yes	Yes	No	No	No	No
Caste FE	Yes	No	No	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	Yes	No	No	No	No
Dynasty-Year FE	No	No	Yes	No	No	Yes	Yes	Yes	Yes
Dynasty-Year Trend	No	No	No	Yes	Yes	No	No	No	No
Observations	1902	1889	1695	1889	1889	1695	1695	1695	1695
Castes	14	13	13	13	13	13	13	13	13
Dynasties	127	114	105	114	114	105	105	105	105
Individuals	790	777	748	777	777	748	748	748	748
Mean Dep Var	0.251	0.252	0.265	0.252	0.252	0.265	0.265	0.265	0.265
R^2	0.151	0.252	0.432	0.441	0.505	0.433	0.438	0.439	0.441
P-Value of Older Cohorts	0.670	0.607	0.840	0.633	0.843	0.403	0.049	0.070	0.088

Note: Dependent variable is dummy variable for working in outside village. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.

Table C.5.: Robustness of Long Run Effects to Alternative Network Sizes

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside	(7) Work Outside	(8) Work Outside	(9) Work Outside
Network members in Cohort	0.0372**	0.0390**	0.0446**	0.0392*	0.0351	0.0423**	0.0418**	0.0441**	0.0437**
	(0.0180)	(0.0194)	(0.0200)	(0.0200)	(0.0223)	(0.0201)	(0.0202)	(0.0204)	(0.0204)
Cohort $3 \leqslant age \leqslant 7$ above	0.00375	0.0152	0.0109	0.0153	0.0142	0.0168	0.0174	0.0190	0.0176
	(0.0186)	(0.0206)	(0.0221)	(0.0213)	(0.0222)	(0.0205)	(0.0205)	(0.0205)	(0.0203)
Cohort $8 \leqslant age \leqslant 12$ above	0.000940	0.00118	−0.00149	0.00144	−0.00195	0.00332	0.00463	0.00363	0.00400
	(0.0226)	(0.0205)	(0.0184)	(0.0211)	(0.0195)	(0.0199)	(0.0197)	(0.0195)	(0.0198)
Cohort $13 \leqslant age \leqslant 17$ above	0.0517**	0.0668**	0.0606**	0.0671**	0.0699**	0.0625**	0.0633**	0.0616**	0.0634**
	(0.0237)	(0.0256)	(0.0287)	(0.0265)	(0.0291)	(0.0295)	(0.0299)	(0.0298)	(0.0303)
Cohort $18 \leqslant age \leqslant 22$ above						0.0513	0.0525	0.0511	0.0522*
						(0.0317)	(0.0317)	(0.0313)	(0.0312)
Cohort $23 \leqslant age \leqslant 27$ above							0.00965	0.00700	0.00927
							(0.0247)	(0.0244)	(0.0246)
Cohort $28 \leqslant age \leqslant 32$ above								−0.0154	−0.0112
								(0.0210)	(0.0237)
Cohort $33 \leqslant age \leqslant 37$ above									0.0212
									(0.0360)
Year FE	Yes	Yes	No	Yes	Yes	No	No	No	No
Caste FE	Yes	No	No	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	Yes	No	No	No	No
Dynasty-Year FE	No	No	Yes	No	No	Yes	Yes	Yes	Yes
Dynasty-Year Trend	No	No	No	Yes	Yes	No	No	No	No
Observations	1904	1891	1695	1891	1891	1695	1695	1695	1695
Castes	15	14	13	14	14	13	13	13	13
Dynasties	128	115	105	115	115	105	105	105	105
Individuals	791	778	748	778	778	748	748	748	748
Mean Dep Var	0.251	0.252	0.265	0.252	0.252	0.265	0.265	0.265	0.265
R^2	0.156	0.257	0.437	0.445	0.509	0.440	0.440	0.440	0.440
P-Value of Older Cohorts	0.165	0.045	0.199	0.054	0.121	0.221	0.305	0.352	0.401

Note: Dependent variable is dummy variable for working in outside village. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.

Table C.6.: Robustness of Long Run Effects to Alternative Network Sizes

	(1) Work Outside	(2) Work Outside	(3) Work Outside	(4) Work Outside	(5) Work Outside	(6) Work Outside	(7) Work Outside	(8) Work Outside
Network members in Cohort	0.0177	0.0211	0.0143	0.0289*	0.0213	0.0226	0.0308*	0.0302*
	(0.0159)	(0.0163)	(0.0155)	(0.0159)	(0.0168)	(0.0166)	(0.0158)	(0.0157)
Cohort $4 \leqslant age \leqslant 10$ above	0.00155	0.0136	0.00432	0.0129	0.0139	0.0148	0.0120	0.0108
	(0.0135)	(0.0128)	(0.0131)	(0.0131)	(0.0132)	(0.0141)	(0.0133)	(0.0134)
Cohort $11 \leqslant age \leqslant 17$ above	0.0371*	0.0502**	0.0358**	0.0530**	0.0505**	0.0572**	0.0514**	0.0530**
	(0.0193)	(0.0203)	(0.0180)	(0.0235)	(0.0210)	(0.0237)	(0.0233)	(0.0239)
Cohort $18 \leqslant age \leqslant 24$ above	0.0399*	0.0524**	0.0397*	0.0540**	0.0527**	0.0615**	0.0517**	0.0535**
	(0.0215)	(0.0215)	(0.0208)	(0.0257)	(0.0222)	(0.0241)	(0.0255)	(0.0252)
Cohort $25 \leqslant age \leqslant 31$ above							−0.0119	−0.00834
							(0.0138)	(0.0155)
Cohort $32 \leqslant age \leqslant 38$ above								0.0163
								(0.0305)
Year FE	Yes	Yes	No	No	Yes	Yes	No	No
Caste FE	Yes	No	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	No	Yes	No	No
Caste-Year FE	No	No	Yes	No	No	No	No	No
Dynasty-Year FE	No	No	No	Yes	No	No	Yes	Yes
Dynasty-Year Trend	No	No	No	No	Yes	Yes	No	No
Observations	1904	1891	1896	1695	1891	1891	1695	1695
Castes	15	14	13	13	14	14	13	13
Dynasties	128	115	125	105	115	115	105	105
Individuals	791	778	788	748	778	778	748	748
Mean Dep Var	0.251	0.252	0.252	0.265	0.252	0.252	0.265	0.265
R^2	0.157	0.260	0.210	0.439	0.447	0.512	0.439	0.440
P-Value of Older Cohorts	0.087	0.012	0.062	0.068	0.015	0.028	0.094	0.133

Note: Dependent variable is dummy variable for working in outside village. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.

C.5. Robustness of Long Run Effects - Regular Employment Placebo

continued overleaf

Table C.7.: Robustness of Long Run Effects to Regular Work as Placebo - Caste

	(1) Regular Job	(2) Regular Job	(3) Regular Job	(4) Regular Job	(5) Regular Job	(6) Regular Job	(7) Regular Job
Network members in Cohort	−0.00127	−0.00000367	−0.00369*	−0.000633	0.0000185	−0.000730	−0.0000290
	(0.00222)	(0.00223)	(0.00216)	(0.00211)	(0.00229)	(0.00215)	(0.00229)
Cohort $5 \leqslant age \leqslant 13$ above	0.00229	0.00337	−0.00187	0.000732	0.00340	0.00154	0.00211
	(0.00240)	(0.00260)	(0.00265)	(0.00287)	(0.00268)	(0.00300)	(0.00287)
Cohort $14 \leqslant age \leqslant 22$ above	0.00758**	0.00587	0.00482	0.00211	0.00585	0.00333	0.00265
	(0.00370)	(0.00357)	(0.00346)	(0.00368)	(0.00367)	(0.00372)	(0.00381)
Cohort $23 \leqslant age \leqslant 31$ above	0.00610	0.00240	0.00525	−0.0000861	0.00235	−0.000844	−0.0000329
	(0.00377)	(0.00382)	(0.00365)	(0.00402)	(0.00394)	(0.00395)	(0.00397)
Cohort $32 \leqslant age \leqslant 40$ above							−0.0102
							(0.00714)
Year FE	Yes	Yes	No	No	Yes	Yes	No
Caste FE	Yes	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	No	Yes	No
Caste-Year FE	No	No	Yes	No	No	No	No
Dynasty-Year FE	No	No	No	Yes	No	No	Yes
Dynasty-Year Trend	No	No	No	No	Yes	Yes	No
Observations	2362	2361	2350	2095	2361	2361	2095
Castes	17	17	14	13	17	17	13
Dynasties	130	129	127	107	129	129	107
Individuals	826	825	823	788	825	825	788
Mean Dep Var	0.118	0.118	0.118	0.119	0.118	0.118	0.119
R^2	0.070	0.156	0.117	0.326	0.255	0.330	0.328
P-Value of Older Cohorts	0.044	0.105	0.272	0.924	0.122	0.704	0.661

Note: Dependent variable is dummy variable for working in regular employment. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.

Table C.8.: Robustness of Long Run Effects to Regular Work as Placebo - Dynasty

	(1) Regular Job	(2) Regular Job	(3) Regular Job	(4) Regular Job	(5) Regular Job	(6) Regular Job	(7) Regular Job
Network members in Cohort	0.0111	0.000986	0.00911	−0.00544	0.000840	−0.00817	−0.00426
	(0.00993)	(0.0111)	(0.0101)	(0.0120)	(0.0114)	(0.0119)	(0.0122)
Cohort $5 \leqslant age \leqslant 13$ above	0.0117	0.0116	0.00822	−0.00147	0.0115	0.00325	−0.000723
	(0.00908)	(0.00857)	(0.00863)	(0.00847)	(0.00879)	(0.00811)	(0.00855)
Cohort $14 \leqslant age \leqslant 22$ above	0.0255	0.0333*	0.0233	0.0217	0.0334	0.0288	0.0195
	(0.0206)	(0.0201)	(0.0200)	(0.0234)	(0.0206)	(0.0221)	(0.0236)
Cohort $23 \leqslant age \leqslant 31$ above	−0.00142	−0.00166	−0.00541	−0.0126	−0.00174	−0.0110	−0.0171
	(0.0115)	(0.0143)	(0.0116)	(0.0153)	(0.0148)	(0.0155)	(0.0148)
Cohort $32 \leqslant age \leqslant 40$ above							−0.0196
							(0.0205)
Year FE	Yes	Yes	No	No	Yes	Yes	No
Caste FE	Yes	No	No	No	No	No	No
Dynasty FE	No	Yes	No	No	No	Yes	No
Caste-Year FE	No	No	Yes	No	No	No	No
Dynasty-Year FE	No	No	No	Yes	No	No	Yes
Dynasty-Year Trend	No	No	No	No	Yes	Yes	No
Observations	2362	2361	2350	2095	2361	2361	2095
Castes	17	17	14	13	17	17	13
Dynasties	130	129	127	107	129	129	107
Individuals	826	825	823	788	825	825	788
Mean Dep Var	0.118	0.118	0.118	0.119	0.118	0.118	0.119
R^2	0.071	0.156	0.119	0.329	0.255	0.333	0.329
P-Value of Older Cohorts	0.246	0.035	0.380	0.715	0.043	0.524	0.607

Note: Dependent variable is dummy variable for working in regular employment. Includes all males in migration sample who are over 18 years of age. Standard errors are clustered at the dynasty level. All regressions include controls for age and age squared. Cohorts are at the dynasty level. Cohorts are calculated at an individual's first appearance as an adult in the data.