The London School of Economics and Political Science

Spatial aspects of the economic development process

Massimiliano Calì

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Philosophy, London, September 2009

Declaration

I certify that the thesis I have presented for examination for the MPhil/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 warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

In recent years the spatial dimension of economic development has attracted increasing interest in the development field. However there is still little analysis and evidence of the ways many spatial dimensions interact with other economic dimensions in the development process. This thesis aims to help filling this gap by bringing a geographical perspective into development economics frameworks. It is empirical in nature and uses data on different sub-national units from India and Uganda. The work is structured around four main papers (divided into six chapters).

The first paper analyses two important aspects of the Indian urbanisation process. First it finds a U-shaped relation between rural-urban disparities in living standards and income per capita across Indian states in the Post-Independence period. Second, it shows that the urbanisation process in India has been characterised by convergence in the 20th century: smaller towns grow faster than large ones.

The second paper examines the role of the agricultural sector in influencing the shape of the urban system. The analysis suggests that the elasticity of rural-urban labour supply increases both urban primacy and the urbanisation rate in Indian states during the Post-Independence period.

The third paper tests for the impact of urban growth on rural poverty using a sample of Indian districts in the period 1981-1999. It finds that urbanisation reduces poverty surrounding rural areas. This effect is largely attributable to positive spillovers from urbanisation rather than to the movement of the rural poor to urban areas.

The final paper examines the determinants of rising returns to schooling in Ugandan districts during the 1990s. The findings suggest that both educational supply and demand factors influenced the wedge between skilled and unskilled labour. Moreover while trade opening reduced this wedge, pro-market reforms increasing inter-district trade raised returns to education in districts relatively abundant in skilled labour.

Acknowledgements

Although the thesis is obviously the product of an inherently individual effort, in a way it is also the result of a collective process. In fact these pages would not have seen the light without the comments, the suggestions, the dedication and the passion of several people during these long and exciting five years.

I am extremely indebted to my supervisor Henry Overman for his guidance and support throughout the entire PhD. His thorough comments and suggestions have always been spot on and have improved enormously the work in all its aspects. Equally important have been his constant support, encouragement and the speed with which he has examined all the material that I have ever sent him. These factors have made the development of the thesis a much smoother and more enjoyable (i.e. less painful) process. Henry has not been the only person in the department of Geography to have contributed to this work. The interactions with the academic staff have improved various parts of the thesis. In this respect I am particularly grateful to Steve Gibbons, Ian Gordon, Frédéric Robert-Nicoud, Andrés Rodríguez-Pose and Olmo Silva. Similar thanks also go to my PhD colleagues: my co-author Carlo Menon who I hope will take responsibility for the eventual errors in chapter 4 (;-)), Rosa Sanchis-Guarner, Filipe Lage de Sousa, Roberto Picchizzolu, Jayaraj Sundaresan, Achileas Tsamis.

Pursuing a PhD while working would not have been feasible without the support of an exceptional team of colleagues at the Overseas Development Institute (ODI). Not only have they always guaranteed the necessary flexibility, but they have also helped me develop lines of research relevant to my PhD within the ODI work, often providing useful and pertinent comments. In this respect I am grateful in particular to Dirk Willem te Velde, who has been instrumental to achieve the necessary synergies between my work and the PhD. Thanks also go to Sheila Page, whose thorough (and direct) comments have enhanced the rigour of my research, and to Edward Anderson for the extensive discussions on the Ugandan chapters. I am also indebted to Kunal Sen, whose passion for development economics (and for the Indian economy) is partly responsible for my decision of pursuing a PhD after my Masters.

A special thank you goes to my parents, and to my brother Fabrizio: they have shared quite a bit of the burden of the thesis' process guaranteeing constant enthusiasm and unabated support from day 1. Their support went as far as asking me about specific questions about the power of some of the instruments used in the thesis! I also would like to thank Scimo not only for his spaghetti carbonara, but possibly more for the passionate discussions (maybe save those on religion) and constant encouragement that contributed to the thesis more than he probably thinks. And I am also grateful to Andrés for las hermosas charlas part of which are embedded in chapters 2 and 3.

Last but by no means least, the biggest thank you goes to my partner Frenci. She has been a constant source of inspiration and encouragement: a loving, caring, hilarious, passionate, witty, tough, sweet companion, friend, and adviser. From the intuitions related to urbanisation in India to the patience used in coping with the absence of my mind lost in PhD related thoughts, her contributions to this work are too many to list here or even remember all. This thesis would not have been possible without her presence in my life. To her go my gratitude and love as well as this thesis, which is dedicated to her.

All the remaining errors, omissions and the opinions expressed in the following are obviously solely of the author.

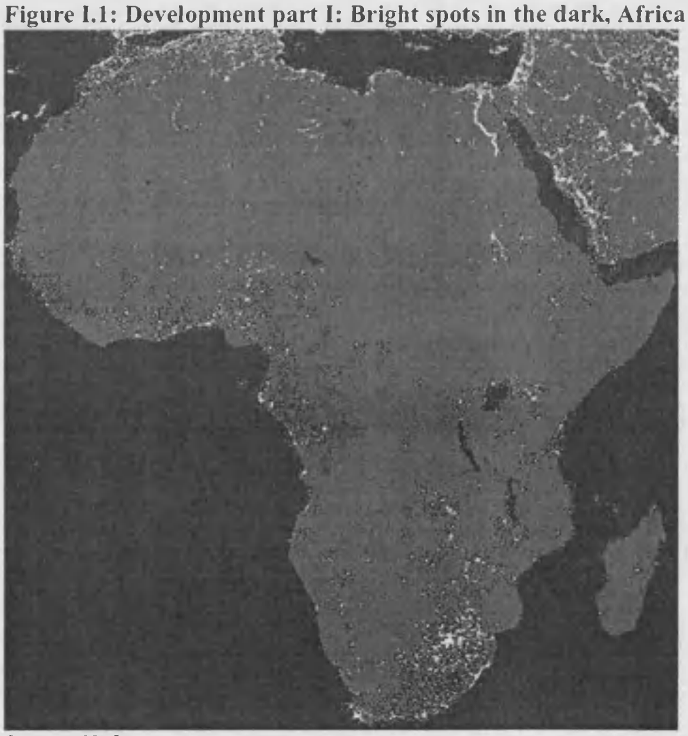
TABLE OF CONTENTS

Introduction

Development from space

Even a cursory look at pictures of the earth at night from a satellite reveals a number of interesting patterns about the spatial distribution of economic activity in the development process.[1] Some of the patterns that emerge from such an inspection represent the focus of investigation of this work. Figure I.1 shows a satellite picture of the African continent at night in 2008. Most of the continent is in the dark. If we exclude North Africa (whose areas on the coast and around the Nile river are fairly lit), there are a few scattered bright spots mainly in correspondence with urban areas and only a handful of larger areas consistently lit: the Eastern part of South Africa, the Western Cape, the oil fields in the coast of Gabon, the greater Lagos and the mines in northern Zambia. Sub-Saharan Africa can arguably be considered as the first stage of economic development in which the level of economic activity is generally low except in a few (relatively isolated) locations.

Figure I.1: Development part I: Bright spots in the dark, Africa at night (2008)



Source: NASA

[1] This relies on the assumption that electricity utilisation is a good approximation of the intensity of economic activity (Henderson et al., 2009 and Sutton et al., 2007).

A jump ahead of a few decades (and a few US dollar thousands per capita) could turn the African picture into something close to Figure I.2. This depicts South America, which is a continent with similar factor endowments (in terms of e.g. natural resources wealth, population density) but with a substantially higher level of income per capita than sub-Saharan Africa.[2] With the majority of countries in the middle-income status according to the World Bank classification, South America could represent the next level of economic development of sub-Saharan Africa (see Wood (2002) for a careful development of this argument). The light is more spread than in Africa and it is very intense in a number of areas (e.g. mega-cities, coastal area of Chile, the state of Sao Paulo, the province of Buenos Aires). However many territories still remain in the dark (including also the unpopulated Amazonia which will hopefully continue to stay dark even in the future).

Figure I.2: Development part II: Spreading the light, South America at night (2008)

Source: NASA

[2] Population density is a very important factor in determining the intensity of light, thus taking areas with similar density is crucial for night-light comparability purposes. South America has a land mass of 17.8 million Km^2 with an estimated population of 392 million; sub-Saharan Africa has a land mass of 24.3 million Km^2 and an estimated population of 800 million.

While the vast majority of land in sub-Saharan Africa is in the dark, in South America dark areas coexist with remarkably lit ones. This seems to be consistent with an uneven pattern of development whereby economic activity and income become increasingly concentrated as countries progress from low to middle income status (see evidence in World Bank, 2009). Also, geographic differences in living standards within countries diverge before converging. These are the stylized facts, based on the experiences of successful developers over the last two centuries (World Bank, 2009). In the first stages of development as countries GDP per capita grows, people and production become concentrated in some parts of countries, so-called leading areas. This concentration appears to slow or stop at per capita incomes between US$10,000 and US$15,000 (World Bank, 2009). Both urban economics and New Economic Geography (NEG) develop models that are consistent at different spatial scales with such uneven pattern of development (e.g. Henderson, 1988, Venables, 2004). In the initial stages of development agglomeration forces (such as urbanisation economies or demand linkages) induce the clustering of economic activity in a few (or one) locations. As economies develop, such areas become more congested and activity tends to spread to other areas as well.

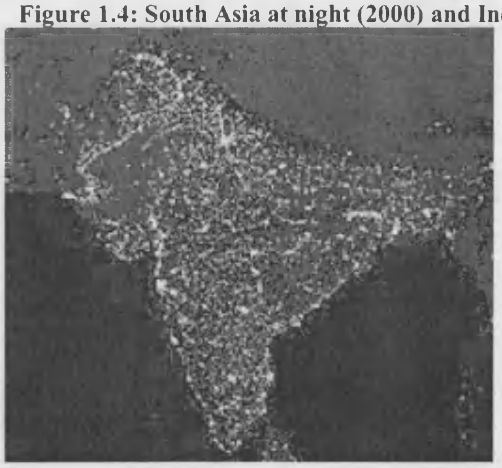
Figure I.3: Development part III: "Everything is illuminated", North America at night (2008)

Source: NASA

The (spatial) convergence part of this process is shown in Figure I.3 depicting North America, where most land is in the light ("Everything is illuminated" to borrow a title of a contemporary novel – Safran Foer, 2002). This (spatially) unbalanced pattern of development was highlighted already by Williamson (1965), although the literature has not devoted much attention to it until recently. Chapter 1 of the thesis tries to help redress this balance by looking at the relation between development and spatial inequality from a relatively new angle (see below).

Now take a look at the left hand-side of Figure I.4, which is the picture of South Asia at night. Light is stronger and more widespread than in the other developing regions. Nonetheless the region is home to the largest number of poor people in the world and its income per capita is lower than any South American countries. This picture reveals the other dimension (other than income per capita) underlying the distribution of economic activity across space: the distribution of population. The very high population density of South Asia (it has a population 50% larger than Africa with a land mass almost ten times smaller) rather than its income drives the relatively intensity of the light. This high density is reflected in the high density of towns as well, as shown in the right-hand side panel of Figure I.4 where the dots represent the towns of India in 2001. There is a remarkably strong relation between the dots and the light. This relation provides compelling prima facie evidence of the key role of urban areas as centres of economic activity even in a fairly unurbanised country as India. This stylised fact is at the heart of this work, a substantial part of which is dedicated to the investigation of the causes and consequences of urbanisation (chapters 2-4).

Figure 1.4: South Asia at night (2000) and Indian distribution of towns (2001)



Source: NASA and Government of India (2001)

Although the light in Figure I.4 is more intense in certain areas (e.g. Indian mega-cities of Mumbai, Delhi, Calcutta and Chennai, the Bengali coast) it is substantially more widespread than in the other regions. Still geographical differences do exist but appear to be more important at lower spatial scales where the darkness of rural areas contrasts with the shining large urban ones. This contrast but also the interconnectedness between urban and rural areas will be the focuses of the thesis' analysis on India (chapters 1-4).

These pictures underscore three key spatial dimensions of the development process: the unevenness in the spatial distribution of economic activity, the importance of urbanisation in determining such distribution, and the relevance of the spatial scale to properly evaluate the inequality of this distribution. The thesis aims to address all of these dimensions from different perspectives, using a variety of quantitative empirical methods and data (on individuals, urban areas, districts and states). Before turning to its description it is useful to briefly examine how the economics literature has incorporated these spatial features in development in order to appreciate the gaps that this thesis helps to fill.

The Economic geography of development

Despite the apparent importance of these spatial features in development the economics literature on developing countries has been fairly silent on them until very recently. One reason for this silence is that space has not featured in any major way in mainstream economics in general until the rise of New Economic Geography (NEG) spurred by the seminal work of Krugman (1991). While recognising the importance of NEG in modelling the role of space in economic development, the thesis is closer to other strands of literature, which are more appropriate to capture the types of spatial interactions treated here. For example the modelling of the urbanisation process follows the urban economics literature in the spirit of the works of Henderson (1985, 1988). The assumptions of this literature appear to be more relevant than the NEG ones at smaller spatial scales where local externalities are expected to play a major role (Combes et al. 2005). On the other hand the NEG assumptions make this class of models better suited to fairly large spatial units (regions, countries or even groups of countries) for which long distance market interactions should play an important role while short distance effects become of secondary relevance. These assumptions have oriented the recent flurry of NEG based empirical analyses of developing countries (e.g. Hanson, 1997, Lin, 2005, Hering and Poncet, 2006) towards the regional dimension within countries, which represents a larger spatial scale than that most of the thesis is concerned with. In terms of the pictures above, this empirical literature addresses questions on which regions within a country or which macro-regions the light is likely to concentrate in.[3] The present work departs from these questions and rather focuses on a set of complementary issues, which need to be addressed at different (usually more refined) spatial scales. These include the disparities between rural and urban areas, the nature of the urbanisation process and its relation with the rural sector and the inequality within sub-national units. Given the focus on the interconnectedness between rural and urban areas, the thesis draws also on the literature on rural-urban migration inspired by the traditional development economics literature (e.g. Lewis, 1954, Harris and Todaro, 1970, Lall et al., 2006 for a review).

As such, this work lies at the crossroad between urban and development economics. It is somewhat complementary to the increasing efforts of the literature to fill some of the analytical and empirical gaps on the ways spatial dimensions interact with other economic dimensions in development (e.g. the UNU/WIDER research project on "Spatial disparities in human development" co-ordinated by Ravi Kanbur and Tony Venable - Kanbur and Venables, 2005 and the World Development Report - World Bank, 2009).[4] This work aims to help fill these gaps and is guided by three main general questions: how have welfare related variables been varying across sub-national units within countries? What drives this spatial variation? What is the impact of government policies on this variation? The following chapters seek to answer these questions by bringing a geographical perspective into traditional development economics frameworks. Although empirical in nature, the thesis builds analytical frameworks in which to ground the results of the empirical analyses. Some of the frameworks developed also introduce innovative elements, which allow the reconciliation of the empirical findings with the theory.

[3] Examples of questions addressed by this literature include the effects of distance to product and factor markets, of reductions in trade costs on the distribution of income and production (mainly) across regions within countries.

[4] For example an important difference with the UNU/WIDER project includes the thesis' focus (Chapter 5-6) on income inequality within spatial units which complements the UNU/WIDER work on inequality across units.

Looking inside rather than across countries

All of the chapters in the thesis are based on intra-country rather than (the more traditional) cross-country analyses. There are two main reasons for this focus. First, although some spatial features in the process of development may be common across countries (as highlighted by the pictures above), "the most striking pattern that emerges from the data on the spatial inequality of developing countries is its varied nature." (Kim, 2007). This suggests that national level factors tend to play a disproportionately large role in shaping the patterns of spatial inequality in developing countries. These factors, which may include among others the type of political system, the factor endowment and the history of a country, tend to be deeply embedded in the national structures and hence are hard to modify in the short run. On the other hand a within country analysis such as that employed throughout the thesis, allows controlling for national level variables and isolating the determinants of the variation of spatial patterns across sub-national units. Such determinants are likely to be more easily modifiable in the short-run and their analysis may be more amenable to policy determination. Second, and related to the previous point, within country analyses are able to evaluate the effects of policies affecting spatial inequalities by exploiting the differential effects across sub-national units.

However this type of intra-country approach poses also special challenges especially when it is applied to developing countries. Macro data at the sub-national level is not often readily available for developing countries. India is an exception in this respect as it enjoys long-spanning detailed data collection processes on a variety of socio-economic variables. I exploit such richness of data in the subsequent analyses, using data derived from the national Statistical Census (Government of India, various years), from household and agricultural surveys. This allows the thesis to focus on a number of different spatial scales, including states, districts and cities, which are adequate to address different types of questions (chapters 1-4). But most developing countries have much less data available for analysis than India, especially among least developed countries (LDCs). These countries often rely on externally funded data collection processes that try to fill the data gaps. One such country is Uganda, where the World Bank has carried out two household surveys (Living Standards Measurement Survey) over the 1990s. I use them by aggregating individuals and community level data to construct datasets at fairly refined spatial levels (regional and district levels).

This allows me to analyse labour market inequalities in a LDC context from new angles and through innovative methodologies (chapters 5-6).

Perhaps a more important drawback of the intra-country analysis is the natural limitation that such analysis imposes on the researcher in terms of the number of observations. That is the case for instance in the empirical analyses based on Indian states (chapters 1 and 3), where consistent data is available for the 16 major states. In the context of an econometrics framework such as that employed throughout the thesis, this relatively small N has at least two implications. First, a large part of the variation in the data comes from the time dimension, unlike in most cross-country analyses, which mainly rely on the large cross-sectional variation for identification. One challenge related to this is that the instrumental variable estimation eventually employed to correct for the potential endogeneity of some regressors needs to rely on time (and cross-section) varying instruments. This condition rules out the use of time invariant geographical instruments, which have proven increasingly popular in the modern literature. A further complication is that the use of datasets with small N does not fulfil the conditions needed to employ the General Method of Moments (GMM) estimation, which is another popular way to correct for the endogeneity of the regressors (Arellano and Bond, 1991; Blundell and Bond, 1998). This method is only efficient asymptotically and is therefore suitable for samples with large N and small T. Second, in the absence of fairly high frequency data (e.g. annual), the overall number of observations is relatively small. For instance as population data in India is collected every ten years, the urbanisation analysis across Indian states relies only on five distinct periods (chapter 3). This reduces the theoretical maximum number of observations to 80 (and in practice they are often slightly less). This creates a further challenge in terms of the estimates' precision, increasing the difficulty of obtaining significant coefficients relative to other Indian states' analyses based on a large number of time series observations (e.g. Datt and Ravallion, 1996, Besley and Burgess, 2000, 2002 and 2004, Rud, 2009). Despite these challenges I make an effort to seriously tackle the estimation issues facing the analyses throughout the thesis, without (hopefully) compromising on the rigour typical of good quality research.

As mentioned, the testing grounds of the thesis' intra-country analyses are two developing countries representative of different phases (and possibly types) of development: India and Uganda. The former is a low income country with a relatively low level of urbanisation.[5] Given the sheer size of its population it is the largest provider of new urban dwellers, although it is urbanising fairly slowly. Between 1981 and 2001 it has added 126 million urban dwellers and is expected to add a further 280 million urban dwellers by 2030 (see Chapter 4). Understanding the causes and implications of such massive urbanisation process is crucial for the future welfare of the country (which is also home to the largest number of poor in the world). The first four chapters of the thesis focus on India and address these issues through new types of empirical analyses.

Overview of the thesis

Chapter 1 provides a description of two important (but under-researched) aspects of the urbanisation process using data on Indian states and cities. First it looks at the evolution of rural-urban inequality across Indian states in the Post-Independence period and at its relation with economic development. This analysis departs from existing ones not only in that it concentrates on sub-national units, but also in that it seeks to establish a relation between rural-urban disparities and economic development (measured by income per capita). This approach tackles one important shortcoming of most existing literature on the relation between inequality and economic growth which does not separate out the different components of inequality (i.e. rural-urban, intra-urban and intra-rural). The results of the analysis support the idea of a U-shaped relation between rural-urban disparities in socio-economic indicators and income per capita. The second part of the chapter examines the urban-counterpart of the rural-urban inequality analysis of the first part, by analysing the (population) growth of Indian urban areas in the 20th century. It explores the determinants of the growth (in size) of Indian cities, focusing in particular on the question of convergence in growth rates. In the context of an urbanising country the analysis of convergence is important as it answers the question of whether larger cities grow slower than smaller ones. As policy-makers are concerned that large cities especially in developing countries are growing too large, the absence of convergence, or even the presence of divergence may support the idea of the state's intervention to tilt the balance of the urbanisation process in favour of smaller cities. However the evidence suggests a strong non linear pattern of convergence

[5] In fact India has just become in 2008 a lower-middle income country according to the World Bank; however it has been a low income country throughout the entire period of the analysis here.

process: city size is inversely associated with subsequent growth up to a point after which the relationship becomes positive.

This may fuel the policy-makers' concern with respect to the growth of mega-cities, which may expand beyond their efficient scale, while the forces towards agglomeration make it difficult to rebalance the growth across urban areas (Venables, 2006). As such there may be a case for policy intervention to decentralize activity, "but we remain woefully ignorant about what works and what doesn't" (*ibid.*, p. 20). Chapters 2 and 3 try to contribute to fill this gap. They focus on the role of the agricultural sector (i.e. the main push factor of the urbanisation process) in influencing the rate of urbanisation and the distribution of urban population across cities, and particularly in the largest city. Such a focus is novel to the literature in search of the determinants of urban concentration (and urban primacy), e.g. Ades and Glaeser, 1995; Davis and Henderson, 2003. This literature has tended to concentrate on pull factors instead (i.e. the factors pulling people into urban areas). This focus is consistent with urban systems models, which do not pay much attention to rural-urban interaction and to the structural shift in economic activities from agriculture in rural areas to manufacturing and services in cities (Kim, 2007).[6] For policy makers in developing countries, these models of urban inequality are likely to prove inadequate guides for understanding urban inequality. Chapter 2 develops an analytical framework trying to incorporate rural-urban migration into an urban economics setting in order to find a relationship between the elasticity of rural-urban labour supply and the urbanisation pattern. The urban economics model determines the urban labour demand curve while conditions in the rural sector determine the urban labour supply.

Chapter 3 provides an empirical analysis based on this framework using both the share of urban population and measures of urban concentration as dependent variables. The analysis is performed using an intra-country panel data from India for the Post-Independence period rather than the traditional cross-country analysis. Such a strategy allows one to control for the effects of national level variables and to better isolate the impact of agricultural variables on urbanisation. Using Indian states as the unit of analysis guarantees that the assumption of inter-state immobility of labour, a necessary condition for the analysis, is reasonably realistic. A number of rural variables, including

[6] The exception is Puga (1998), who takes into account the rural sector in a two-city model using an NEG framework. Its predictions refer to the values of transport costs for which different urban equilibria arise. Chapter 2 discusses the empirical challenges involved in implementing such predictions.

agricultural land productivity, agricultural wage and the demographic composition of the rural population are used to proxy the elasticity of labour supply which is not observed directly. These variables are also instrumented via rainfall levels and land reform legislation to address their potential endogeneity to urban variables. The results support the idea that the elasticity of rural-urban labour supply has a positive (and causal) effect on urban primacy and on the urbanisation rate.

Chapter 4 provides a symmetrical perspective to the previous chapters, by examining the impact of urbanisation on poverty in surrounding rural areas with a sample of Indian districts in the period 1981-1999. Again this is an under-researched question but a relevant one in a period of increasing urbanisation in most developing countries. This analysis becomes more important when considering that most of the world's poor reside in rural areas, where the incidence of poverty is higher than in urban areas. In particular, with over 316 million of $1/day rural poor in 2002, India is home to 36% of the world's rural poor (the largest number in the world). The chapter represents one of the first efforts to map the channels through which an expanding urban area may affect poverty in surrounding rural areas. In particular it distinguishes between first and second-round effects of urbanisation and tries to disentangle them empirically. The former involve only a statistical association between urbanisation and changes in rural poverty due to the change in residency of some rural poor (who may or may not be lifted out of poverty in their move to the urban areas). On the other hand, second-round effects capture the impact of the urban population growth on the rural rate of poverty. Such a relationship is causal in nature and should tell us how good or bad urbanisation is for rural poverty. The results suggest that urbanisation has a substantial and systematic poverty reducing impact in surrounding rural areas, which is mainly attributable to second round rather than first-round effects.

The first chapters concentrate on the spatial distribution of population and income across areas, with a special focus on urban areas. In this sense they are related to spatial disparities *between* areas. However most (around two third) of the economic inequality in developing countries is determined *within* areas rather than *between* them (Kanbur and Venables, 2005, Elbers et al., 2005 and Demombynes et al., 2003). The last part of the thesis (chapters 5 and 6) focuses on this type of within area inequality by using one of its most popular measures, i.e. labour market inequality (as measured by returns to schooling). It examines the case of Uganda in the nineties, which provides an empirical setting complementary to the Indian one. Uganda is a LDC with lower levels of income per capita than India. It has much smaller population and area. In terms of the night light pictures above, this implies that unlike India, Uganda is almost entirely in the dark. But similarly to India Uganda underwent a substantial economic liberalisation process (and experienced sustained economic growth) during the nineties. This process has been associated with rising inequality at the national level (as highlighted by the findings in chapter 5), a result which apparently seems inconsistent with the predictions of traditional trade theory. The district level analysis undertaken in chapter 6 helps reconcile the rising returns to schooling with this trade theory.

Chapter 5 develops the empirical framework to measure labour market inequality via returns to schooling, both at national and sub-national levels. This measure of inequality is interesting also in relation to the importance given by the government and the donor community alike to investment in education to raise the rate of growth and poverty reduction in many sub-Saharan African countries (including Uganda). Through the framework the chapter examines the evolution of disparities in Uganda in the nineties at the national and regional level showing in both cases a steep rise in returns to schooling (with substantial variation across regions). Chapter 6 uses this empirical framework to analyse the inequality within districts in the same period. It tests for the effects of both educational supply and demand related factors on the wedge between skilled and unskilled labour, finding that both factors are important to explain such wedge. The estimation controls for the likely endogeneity of the educational variables, instrumenting them through a set of variables based on communities' distance to primary and secondary schools. Importantly, the chapter also examines the likely impact of some policy shocks, i.e. the trade liberalisation and the facilitation in cross-district trade, on such inequalities. The results suggest that while trade opening reduced this wedge (in line with standard H-O theory), pro-market reforms increasing trade across districts raised it in districts with a relatively larger skilled labour force.

In a world where policy-makers are increasingly interested in spatially focused policies such as migration restrictions, political decentralisation and development of export promotion zones, increasing our understanding of spatial issues is a key step for adequate policy formulation. The hope is that this work may represent a small contribution on the route to doing this.

Chapter 1. Urbanisation, inequality and development: Evidence from Indian states and towns

1.1. Introduction and scope of the work

The process of economic development is generally accompanied by the structural transformation of the economy, whereby the rural population mainly employed in agriculture turns into an urban one shifting towards industry and eventually services. Although such a process is a well established fact, we still lack a clear understanding of its relations with the spatial dimensions of inequality. In particular, two important and complementary aspects of this inequality in the transition process have received relatively little attention so far. The first relates to the inequality between rural and urban areas. What changes in welfare do rural and urban areas experience in the process of economic development? Is there a systematic relation between rural-urban inequality and income per capita growth? The urban counterpart of this question focuses on the way different urban areas are affected in the rural-urban transition. Are small and large urban areas differently affected along this transition? Is there a convergence or a divergence process in cities' growth rates as a country urbanises?

Answers to these questions are important to gain a better understanding of the nature of both the rural-urban and the within urban transitions. This chapter addresses both types of questions empirically by employing different types of data. I use the difference in direct measures of wellbeing between rural and urban areas in order to examine the relationship between rural-urban disparities and per capita income. As these direct measures of welfare are not available for cities, I use population growth as a proxy of wellbeing for urban areas, as argued by Glaeser et al. (1995). I investigate the different questions of the empirical analysis using Indian states over the Post-Independence period and Indian towns over the 21st century as the units of analysis. In this way I can exploit the richness of contexts within the Indian sub-continent, controlling for many of country level unobservables that undermine the robustness of inferences from cross-countries studies. India is currently experiencing the type of structural transformation process mentioned above. Although its rate of urbanisation is not particularly rapid, the sheer size of its population makes the country the world's largest source of new urban dwellers in the next decade (according to UN, 2008). Thus the findings from this analysis may be particularly relevant in assessing the possible implications of such a massive urbanisation process. This analysis provides also a characterisation of this urbanisation process in India, which is useful to put the results of the subsequent chapters in context.

1.1.1. Rural-urban inequality and economic development

The relationship between income distribution and economic development was first identified by the seminal work of Kuznets (1955). His work hypothesised an inverted U shaped relationship between income and inequality: the initial stage of a country's economic development would be associated with rising inequalities up to a point (during the middle-income stage of development), after which inequalities would decrease with income per capita.

I find it useful to split income inequality along a spatial dimension into a *within-sector* and a *between-sectors* component. The latter relates to the income disparity between the rural and the urban sector while the former refers to intra-urban and intra-rural inequalities. This chapter is concerned with rural-urban inequality, which usually explains the majority of a country's inequality in the early stages of the development process. For instance Kanbur and Zhang (1999) find that over 70% of overall inequality in China was explained by the rural-urban component over the period 1983-1995.[7] Frankema (2006) argues that this component of inequality mimics the relationship between overall inequality and income growth. In the initial phase of development the urban sector expands due to rapid urban labour productivity growth. This widens the rural-urban income gap as the increase in rural productivity is more sluggish. After peaking, rural-urban dualism declines and eventually dissolves in the long run as rural labour productivity catches up following rural-urban migration and technology and demand spill-overs from the urban sector. Evidence from Latin American countries in the 20th century supports this type of relation (Frankema, 2006).

Such a relation has important similarities with the one between regional inequality and economic development within a country. The work in this area has been

[7] This share was over 80% for inland areas.

inspired by the analysis of Williamson (1965), who found that regional disparities would bear a typical inverted U shaped relation with income per capita. NEG models are suitable to provide the theoretical intuition for this pattern (e.g. Fujita et al. 1999; Venables, 2004). This relies on the tension between centripetal forces towards concentration of economic activity and centrifugal forces towards dispersion. The former can be driven for instance by labour mobility (i.e. a large market creates jobs and workers' expenditure makes the market larger, as in Krugman 1991) or input-output linkages (i.e. firms create the market for other firms, as in Venables 1996). Such centripetal forces would prevail in the early stages of development, causing a process of cumulative causation which reinforces the initial advantage of the more advanced region/location. This concentration process would continue until the level of economic activity reaches a threshold, after which the congestion costs from agglomeration (centrifugal forces) would offset the centripetal forces, dispersing economic activity again.[8] These models may provide some micro-foundations for explaining the divergence-convergence hypothesis of the rural-urban model.

While the relationship between regional inequalities and economic development has been the subject of increasing interest, relatively little empirical evidence is available on the rural-urban inequality-income relation. Although the analysis of rural-urban dualism has enjoyed increasing attention in the recent development literature (as wished by Bourguignon and Morrison in 1998), this has tended to focus on a handful of countries, and particularly China (Park, 2007, Knight et al., 2006). Importantly, the main focus of these studies has been the evolution of rural-urban inequality over time (e.g. Ravallion and Chen, 2007, Ferreira et al., 2008, Sahn and Stifel, 2003) or its relation with the urbanisation process (Park, 2007, Lu and Chen, 2006). On the other hand the relationship between rural-urban inequality and economic development has received little attention. Arguably, this is also due to the failure to separate the sub-components of the inequality-income relationship. Most literature concentrates on the relation between inequality and economic growth without distinguishing the different components of inequality (i.e. rural-urban, intra-urban and intra-rural). This may lead to overlook countervailing forces hidden in the catch-all income inequality variable, which may help explain why the inverted U-shaped Kuznets' hypothesis is still empirically

[8] Junius (1996) uses this NEG framework to derive the inverted U-curve between economic concentration and development at the country level.

controversial, although it seems to fit a number of countries' development processes, including the recent growth experience of China.[9]

The first part of this chapter analyses the relationship between rural-urban inequality and economic development in Post-Independence India. It is worth noting from the outset that there is no presumption to identify any causal relation in this context, for the main objective is to examine whether a systematic association between economic development and an important component of inequality does exist empirically. Understanding the nature of this relation would be relevant in order to assess whether the economic development process has an urban, a rural or a neutral bias.

1.1.2. Size and growth of cities

The second part of the chapter examines the urban-counterpart of the rural-urban inequality analysis of the first section, by analysing the (population) growth of Indian urban areas in the 20th century. It explores the determinants of the growth (in size) of Indian cities, focusing in particular on the question of convergence in economic growth. This question links back to the literature in the tradition of Baumol (1986) and Barro (1991). The findings from this literature firmly support the idea of convergence between sub-national units (e.g. US States) (Quah, 1996), while the evidence on cross-country convergence is much weaker. In a neoclassical world, these results may be explained by the limited mobility of factors of production, and capital in particular, across countries, and by the free factors' mobility within countries.[10] Analysing convergence across cities in one country, where capital and labour are quite mobile at least within states in one country, provides an interesting testing ground for the convergence hypothesis. Income growth is a natural measure of productivity growth across countries as labour is immobile. In contexts where labour is mobile as across Indian cities (within a state or a district) the situation is different (Glaeser et al.,1995). To the extent that internal

[9] A vivid illustration of this point is provided by Frankema (2006), who finds that the persistent personal income inequality of Latin America in the 20th century is also associated to declining rural-urban inequality.

[10] Lucas (1990) provides two sets of explanations for the apparent paradox that capital doesn't flow to countries where it is relatively scarce. One has to do with differences in fundamentals between countries that influence the production function, e.g. technology, institutions, human capital; the other is related to imperfections in international capital markets (e.g. asymmetric information, risk of expropriation). These differences are usually much smaller within a country.

migration responds strongly to growth opportunities, population growth captures the extent to which cities are becoming increasingly attractive labour markets. On the other hand Glaeser et al. (1995) show that income growth in a city is associated not only with productivity growth, but also with a decline in the quality of life. Therefore according to this view income growth is a less straightforward measure of urban success than population growth. The latter is thus a more natural variable to use when testing for convergence across cities within a country.

In the context of an urbanising country the analysis of convergence is important as it answers the question of whether larger cities have grown at a different rate than smaller ones. To the extent that policy-makers fear that large cities especially in developing countries are growing too large, the absence of convergence, or even the presence of divergence may support the idea of the state's intervention to tilt the balance of the urbanisation process in favour of smaller cities.[11] This has been, for example, the rationale for the Integrated Development of Small and Medium Towns programme (IDSMT), implemented by the Indian government since 1979.[12]

Despite the potential importance of the question of convergence in cities' growth, not much empirical evidence has examined it so far. In general the focus of the empirical literature on cities' growth has rather been on the determinants of city size, such as human capital, unemployment, natural conditions and the industrial base (e.g. Glaeser et al., 1995, Glaeser and Shapiro, 2001, Shapiro, 2006, Rappaport, 2007, de Mata et al., 2007). A recurrent finding of this literature is that human capital enhances the growth prospect of a city while unemployment has the opposite effect. In the United States cities with warmer weather have grown more than average in recent decades (Glaeser and Shapiro, 2001, Rappaport, 2007), while local government spending – except on highways – was associated with lower growth. Only a few studies devote some attention to the question of convergence. Using a cross-section of US cities Glaeser et al (1995) find little evidence of convergence in population growth rates between 1960 and 1990 (and somewhat more robust evidence for the period 1950-1970). Similarly Eaton and Eckstein (1997) find that initial population is unrelated to

[11] Scott and Storper (2003: 581) argue for instance that urbanisation patterns in developing countries have generated "macrocephalic urban systems consisting of a few abnormally large cities in each country".
[12] This programme supported the growth of small and medium sized towns through public investments. especially in small infrastructure projects.

subsequent city growth in France and Japan. On the other hand da Mata et al. (2007) find some evidence of convergence for a sample of Brazilian cities between 1980 and 2000. The present focus on India is particularly interesting in light of the existing literature, as it characterises the growth of cities during the urbanisation process of the country. This provides a rather different context to that of a fully urbanised country (as in the studies mentioned above), where the majority of city growth is driven by inter-city migration. As much of the empirical literature on cities' growth is based on fully urbanised countries, generalising its findings to (developing) countries in the midst of the urbanisation process may be problematic. One partial exception to it is the work by Henderson and Wang (2007), who examine urbanisation patterns across countries between 1960 and 2000. They find that increasing urbanisation is accommodated by the growth of small and medium sized cities. Furthermore, democratisation seems to help smaller cities grow relative to larger ones as it levels the playing field across cities. However, as for all cross-country studies, the extent to which these results are driven by time varying unobservable differences across countries is open to question. By focussing on an individual country this chapter aims to examine some cities' growth determinants isolating national level factors.

Alongside convergence, given the data available it is possible to examine some other determinants of city growth as well, such as geography, climate and proximity to large agglomerations. These factors have been shown to be potentially important predictors of city growth in fully urbanised contexts. Glaeser et al. (2001), and Glaeser and Shapiro (2003) find that warm and dry weather was positively associated with city population growth in the United States at the end of the 20th century. As noted by Rappaport (2007) nice weather in the US has become a more important consumption amenity probably due to broad-based rising per capita income. There are reasons to expect that weather conditions may not be as important determinants in the case of India as they are in more developed economies. First, average incomes in India are still very low by international standards and therefore weather may still not be considered as an important consumption amenity by the majority of the population. Second, the variation of weather conditions across India is not as large as in the US, as the vast majority of Indian urban areas are concentrated in tropical and semi-tropical geo-climatic areas. Another determinant of city growth that can be tested in the analysis is proximity to large agglomerations used as a proxy for market potential. Da Mata et al (2007) find that market potential (measured as the distance discounted sum of incomes of all metropolitan areas for any city) is an important determinant of city growth in Brazil. In particular, being close to large urban markets raises the city's productivity.

Finally, I can also test the extent to which growth rates are persistent over time, a finding that is consistently supported in the case of US cities (e.g. Glaeser et al., 1995, Glaeser and Shapiro, 2001). Section 1.3 will explore these features using a panel of Indian cities in the 20^{th} century.

1.2. Rural-urban disparities

The basic idea of this section is to test whether a relationship exists between economic development and rural-urban inequality, and what shape it eventually takes. To the extent that economic progress is associated with urbanisation this analysis is also related to that of Chapter 4, which looks at the impact of urbanisation on rural poverty.

I use three families of indices to measure the disparities in welfare between rural and urban areas across Indian states over time:

1) Poverty based measures

2) Consumption based measures

3) Health based measures

The use of such indicators rather than income based ones is due to the lack of data on rural and urban incomes at the state level in India. It is also motivated by another consideration. Although poverty- and consumption-based indicators of inequality are obviously inter-related with more traditional measure of inequality (calculated on absolute incomes), their use may add a further interesting angle to the debate. That is the case to the extent that poverty reduction can arguably be considered a more important policy objective than inequality per se. For instance, as argued by Eastwood and Lipton (2000, p. 21) transfers among those above the poverty line that "reduce inequality without touching poverty should be of second-order concern". Finally, the use of a health related indicator allows examining the disparities in one of the key dimensions of wellbeing.

I construct two different indicators of rural-urban disparity based on poverty measures: the difference in the *headcount index* between rural and urban areas

(H_1=H_{rur}-H_{urb}); and the difference in the *poverty gap* (PG_1=PG_{rur}-PG_{urb}) – Appendix 1.1 describes the construction of these indices.[13] The rural headcount index measures the share of rural people below the rural poverty line in total rural population, and so does the urban headcount index. Thus their difference is conveying information on the differential incidence of poverty between rural and urban areas. However, the index does not contain any information on the severity of poverty, as it assigns the same weight to all poor. The poverty gap instead measures the mean distance from the poverty line as a proportion of the poverty line, thus taking into account how far from the poverty line the poor are on average. As argued by Eastwood and Lipton (2001), indicators using the poverty gap index are better at capturing relative rural poverty than those using the headcount index.[14] This is because the latter does not capture any changes in the poverty of persons below the poverty line unless they cross such line. I find it useful to use both indicators as they convey different insights. The results are slightly different between the two, especially as far as socio-demographic determinants are concerned, as shown below.

I also use the ratio of the rural to the urban mean per capita monthly expenditure as the consumption measure (ME_2=ME_{urb}/ME_{rur}). Finally, I proxy the difference in health conditions with the rural-urban difference in death rates per 1000 people (D_1=D_{rur}-D_{urb}). Death rates convey important information about development in their own right. They are likely to be determined by a large array of socio-demographic and economic factors. I try to control for as many of these factors as possible. All of these indicators except ME_2 are based on variables that measure negative attributes, such as poverty and death rates. I modify ME_2 (i.e. rural expenditure is subtracted from the urban one) so as to make all indicators increasing in the rural-urban gap in living standards.

[13] The rural-urban division is made according to the Census definition of urban areas. The Census in 1991 classifies towns as all the statutory places with a municipality, corporation, cantonment board or notified town area committee, or, alternatively, places satisfying simultaneously the following three criteria: i) a minimum population of 5,000; ii) at least 75 per cent of male working population engaged in non-agricultural pursuits; and iii) a density of population of at least 400 per sq. Km. This definition has changed slightly over the time considered but in a consistent way across states, thus year effects control for this potential issue.

[14] Eastwood and Lipton (2001) actually use ratios instead of differences in indices. When I tried to use ratios the results are similar to those with differences although slightly less robust.

The basic approach is to estimate the following panel data model:

$$h_{st} = \alpha_s + \gamma_t + \beta_1 y_{st} + \beta_2 y_{st}^2 + \beta_3(\Delta y_{st} / y_{st-1}) + \Gamma X_{st} + \varepsilon_{st} \qquad (1.1)$$

where h_{st} is some measure of rural-urban disparities as described above in state s at time t, y_{st} is state real income per capita, X_{st} is a vector of socio-demographic controls, α_s is state fixed effects and γ_t is year effects. Since the dependent variable has also negative values, all variables in (1) are in levels (rather than in log). I estimate equation (1.1) for a panel of the sixteen major Indian sates (in the 1958-2002 period). The cross-sectional dimension of the dataset is relatively small implying that an important part of the identification strategy comes from the time variation. Such a strategy has become increasingly popular in the empirical literature in development economics, a relevant part of which uses the same Indian states as in this study (e.g. Datt and Ravallion, 1996, Besley and Burgess, 2000, 2002 and 2004, Rud, 2009). I cannot use all Indian states due to data limitations but those in the sample account for over 95% of total Indian population in 2001.

In such a context fixed effects estimation appears to be more appropriate than random effects, as the states considered are not randomly selected (they are the largest ones) and states' unobservables are not likely to be systematically related to the right hand side variables.[15] Moreover the Hausman test rejects the null of non-systematic difference between fixed and random effects estimators. Endogeneity can arguably affect specification (1.1), especially as far as income per capita is concerned. For instance both income per capita and rural-urban gap could be driven by the same unobserved shocks (e.g. monsoon failure or a policy reform). However as mentioned earlier the analysis does only seek to examine whether a systematic association between rural-urban disparities and income per capita exist empirically, without establishing any causality. Nevertheless I try to control for a large array of factors that may bias the β_1 and β_2 coefficients in (1), and I also use the lagged values of income per capita that may partly reduce the endogeneity bias.

The X vector in (1) includes all those socio-demographic characteristics of the rural and urban populations likely to affect both rural and urban welfare measures. In

[15] I obtain similar results to those detailed in the main text estimating the model through GLS modelling the error term as an AR(1) process allowing for state-specific autocorrelation. Results are available upon request.

particular, they include the age structures of the population and the sex ratio, which may influence the relative productivity (which in turn affects poverty). The potential effect of the sex ratio on productivity is rooted in the differential productivity between males and females in economies dominated by agriculture. This difference has been one of the major social factors to explain the "missing women" (Sen, 1992) problem in South Asia, and in India in particular (Agnihotri et al., 2002, Das Gupta, 2005). The supposedly lower female productivity in agriculture (relative to males) contributes to skewing the intra-household allocation of resources in favour of males. This includes food and healthcare provision determining higher infant mortality rates for females than males. As such, the sex ratio would be expected to have an effect in rural but not in urban areas. By the same token, ceteris paribus working age population are usually more productive than older people and infants. The share of scheduled caste population, which has historically had above average poverty incidence, is a further control. I also include the growth rate of GDP in (1.1) to capture eventual dynamic effects of economic growth on rural-urban disparities. Finally, I also test whether the land reform enacted in the Post-Independence period has had any impact on rural-urban disparities. As shown by Besley and Burgess (2000) such legislation has had a substantial poverty reducing effect in rural areas, thus it is expected to reduce the rural-urban poverty gap.

1.2.1. Data

The data for the income and consumption based measures come from the World Bank dataset prepared by Ozler, Datt and Ravallion (1996), and further updated by the same authors (see Appendix 1 for a description of the methodology to construct those indices). The same dataset provides also state-wise income data, which have been updated until 2002.[16] Data for death rates come from various years of the Indian Census, and so do the demographic data. Data on land reform legislation come from Besley and Burgess (2000), who coded all the relevant acts passed by the state parliaments to implement the land reform since 1957. The reform was implemented under the 1949 Indian Constitution, according to which states are granted the powers to enact (and implement) land reforms. Each state parliament implemented the reform through autonomous acts. There are significant differences in the intensity with which

[16] The data has been updated by the Economic Organisation and Public Policy Programme at the London School of Economics.

states have enacted the various types of land reform legislation over time. Such differences have been captured by Besley and Burgess (2000) who construct a yearly cumulative land reform variable, by adding up individual land reform acts between 1957 and 1992 in the major Indian states. They classify each act into four types of legislation: tenancy reform, abolition of intermediaries, land ceiling legislation, land consolidation legislation. Using this cumulative variable, the authors show that land reform had a significant impact in reducing poverty (and increasing agricultural wage) across states over time.

Table 1.1 presents the summary statistics for the rural-urban disparity and income variables. Interestingly, not all welfare indicators are worse in rural than urban areas at any point in time, as shown by the negative minimum values for the headcount, poverty gap and death rate differences. However the average difference in headcount poverty rates between rural and urban areas is 8 percentage points (first row, column 2), indicating a substantial (albeit variable) gap between rural and urban areas across Indian states.

Table 1.1: Summary statistics for the main variables

	Obs.	Mean	Std. Dev.	Min	Max
Rural-urban poverty headcount difference	622	8.06	10.79	-21.14	50.06
Rural-urban poverty gap difference	562	2.42	4.03	-11.03	14.73
Rural-urban mean consumption ratio	562	1.38	0.21	0.90	2.41
Rural-urban death rate difference (per 1000 people)	448	4.13	2.04	-3.90	12.30
Per capita GDP in constant 1980 prices (Rs '000)	679	1.93	0.96	0.62	6.22
Annual GDP growth	663	0.03	0.09	-0.34	0.47
Urban share (% of total population)	816	21.37	8.19	4.06	43.86

Source: Indian Census and Datt et al. (1996).

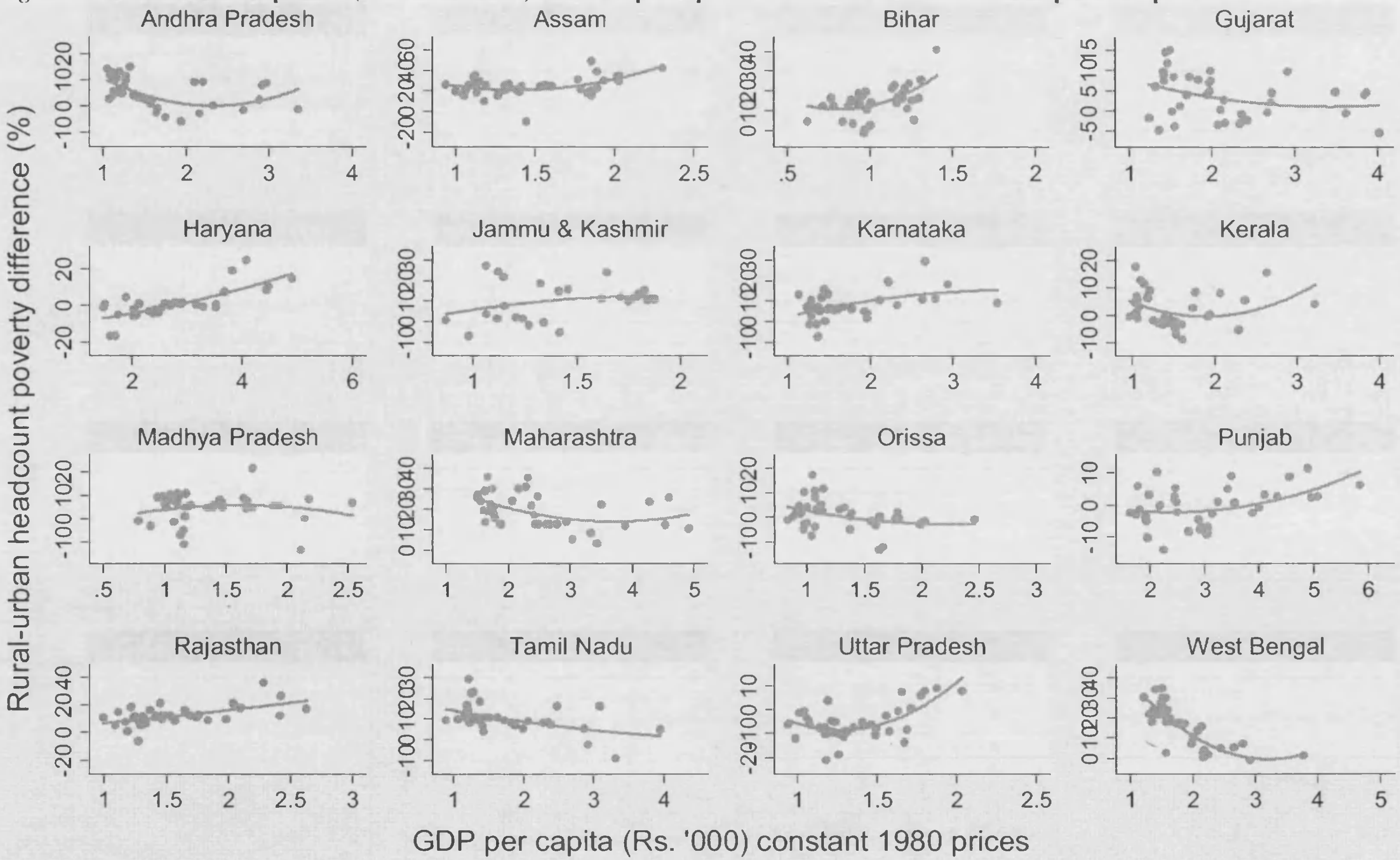
1.2.2. Graphical evidence

Before proceeding with the regression analysis, it is useful to graphically explore the relationships between the main variables. Figure 1 shows the relationship between rural-urban disparities (using the headcount index) and income per capita for each Indian state over the period 1958-2002. I fit the line for each state by using OLS with a quadratic function.

A quite clear U-shaped relationship emerges for most states, although a few states (i.e. Jammu & Kashmir, Madhya Pradesh and Karnataka) have the opposite inverted U-shaped pattern and Orissa, Rajasthan and Tamil Nadu have linear patterns. These patterns have been corroborated by state-level regression analysis (not shown here), which confirms that rural-urban poverty difference is slightly increasing in per capita GDP and decreasing in its squared term for Jammu & Kashmir, Madhya Pradesh and Karnataka. Similarly the squared term appears to be insignificant for Orissa, Rajasthan and Tamil Nadu. This U-shaped pattern emerges quite vividly when using the other consumption and health based measures of inequality as well (Figures A1.1 and A1.2 in the Appendix).[17] These stylised facts may suggest a pattern of economic development accompanied by a reduction in rural-urban inequalities over time (with an eventual slight increase for certain states). This pattern is the product of a monotonic increase in GDP per capita over time and a mixed evolution – with often a decreasing pattern – of rural-urban poverty difference, as shown in Figure 1.2. This plots the evolution of GDP per capita and rural-urban disparities over time. Let us turn to a more formal scrutiny of the relationship to test to what extent this U-shaped relationship holds when controlling for other factors.

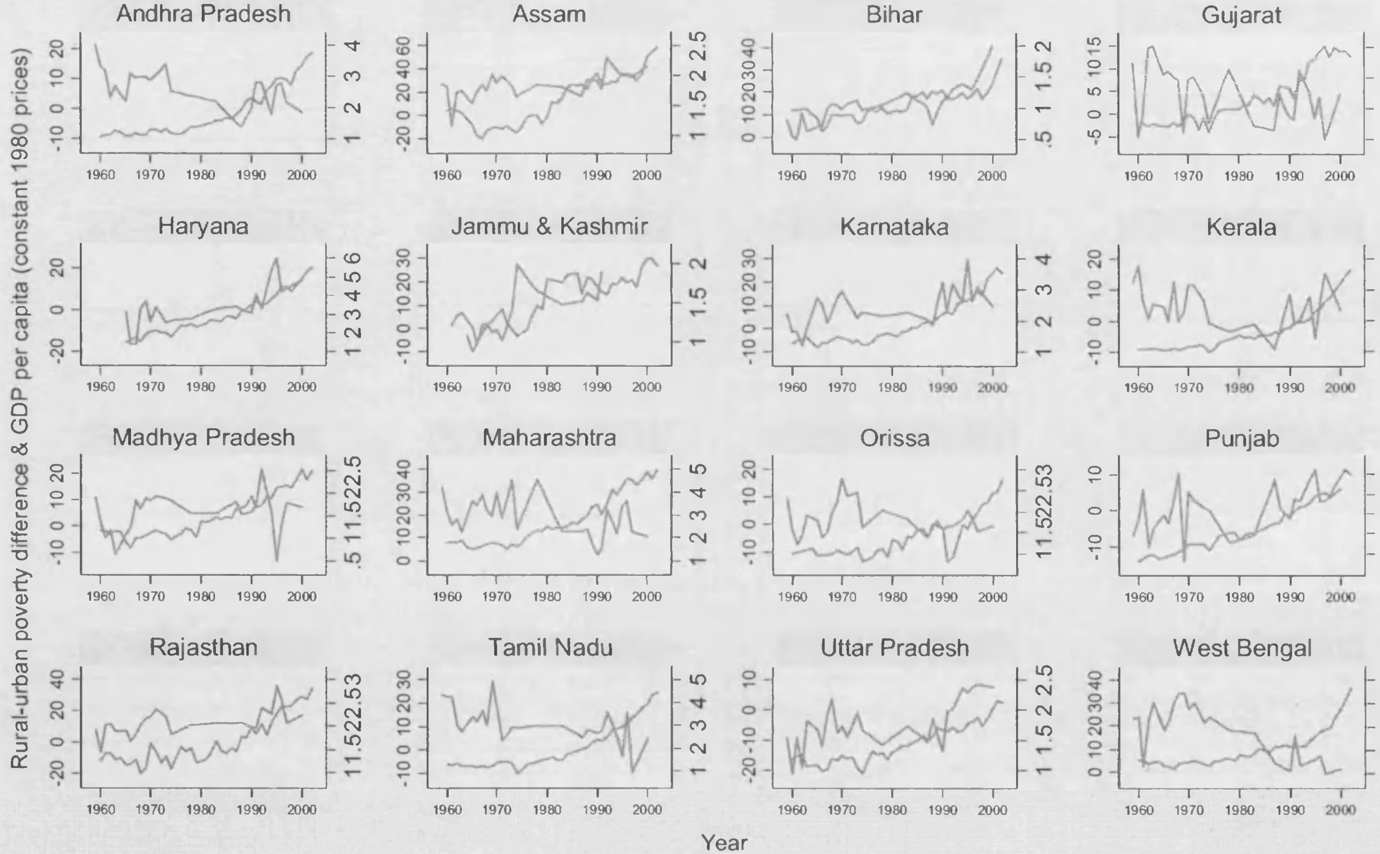
[17] Note that due to cross-states variation in data availability over time, some states have less data points than others.

Figure 1.1: Rural-urban disparities (headcount index) and GDP per capita in Indian states in the Post-Independence period



Note: Poverty difference is measured as the difference between the poverty headcount index in rural areas and that in urban areas (both measured in percentage terms)

Figure 1.2: Evolution of rural-urban disparities and real income across Indian states over time, 1958-2002



Note: Rural-urban poverty difference is measured as the difference between the poverty headcount index in rural areas and that in urban areas (both measured in percentage terms). Poverty difference is on the left y axis and GDP per capita is on the right y axis.

1.2.3. Regression analysis

Table 1.2 presents the results of the regressions based on equation (1.1), which provide support for the U-shaped relation emerging from the graphs. The difference in the headcount poverty index between rural and urban areas decreases as income rises up to a point after which it starts increasing again. In particular in the baseline regression without controls (column 1) a 10% increase in real per capita GDP (from the mean of 1,930 Rs.) is associated with a reduction of 3.1 percentage points in rural-urban difference in the headcount index.[18] The trough in this difference is reached for a value of GDP per capita of 4,450 Rs. (at 1980 prices) after which the rural-urban headcount difference starts rising. Note that only 3 out of 16 states (Haryana, Maharashtra and Punjab) had income per capita higher than this level in 2000. Column 1 further suggests that rural-urban inequality increases in the speed of per capita income growth.

These results are robust to the inclusion of a number of socio-demographic controls (column 2). The share of the population in working age (15-59 years) in urban areas is positively associated with rural-urban headcount poverty difference while the coefficient of the same share in rural areas has a negative sign. The reverse is true for the share of population over 60, consistently with the relatively lower productivity of the older age group. The female/male ratio in rural areas is associated with an increasing rural-urban poverty headcount difference, while the ratio in urban areas has an insignificant (but positive) effect. This is consistent with a lower female productivity in agriculture but not in the urban sector where the incidence of poverty among females is not significantly different than among males. As expected, the presence of scheduled caste influences the headcount difference through its positive association with poverty rates, although the only significant effect is in urban areas. These results are broadly in line with the expected effects of socio-demographic factors on poverty.

[18] This coefficient is obtained considering that a 10% increase in GDP per capita at the mean is 0.193 and for its squared term it is 0.463. Therefore $(\beta_1 * \Delta GDP/GDP)+(\beta_2 * \Delta GDP/GDP) = -21.71 * 0.193 + 2.44 * 0.463 = -3.06$

Table 1.2: Rural-urban disparities and income per capita across Indian states, 1958-2002

	(1) headcount difference	(2) headcount difference	(3) headcount difference	(4) headcount difference	(5) headcount difference	(6) headcount difference	(7) headcount difference
GDP pc	-21.71***	-17.67***	-20.30***		-20.83***	-20.83***	-19.95***
	(3.26)	(3.68)	(3.63)		(3.95)	(3.95)	(4.07)
GDP pc sq.	2.44***	2.41***	2.30***		2.85***	2.85***	2.84***
	(0.43)	(0.46)	(0.46)		(0.55)	(0.55)	(0.55)
GDP growth	12.26***	10.95***	13.31***	4.91*	11.44***	11.44***	10.96***
	(3.29)	(3.02)	(3.17)	(2.63)	(2.99)	(2.99)	(3.01)
Population		-12.84*	-13.61**	-20.17***	-9.53	-9.53	-8.87
('00000)		(6.71)	(6.63)	(7.16)	(7.90)	(7.90)	(7.91)
Rural 15-59		-1.49***	-1.54***	-2.35***	-0.85**	-0.85**	-0.93**
		(0.39)	(0.41)	(0.43)	(0.39)	(0.39)	(0.41)
Urban 15-59		0.47***	0.50***	0.39**	0.14	0.14	0.12
		(0.17)	(0.16)	(0.17)	(0.19)	(0.19)	(0.19)
Rural 60+		2.67	3.29**	0.77	2.27	2.27	2.13
		(1.81)	(1.66)	(1.63)	(2.01)	(2.01)	(2.02)
Urban 60+		-5.31***	-5.23***	-3.53*	0.05	0.05	0.48
		(1.88)	(1.91)	(2.01)	(2.34)	(2.34)	(2.41)
Fem/male		159.56***	154.15***	160.62***	73.49*	73.49*	76.65*
rural		(40.95)	(43.47)	(46.58)	(43.97)	(43.97)	(44.35)
Fem/male		44.58	-5.37	-22.88	27.18	27.18	22.37
urban		(29.09)	(28.70)	(29.52)	(31.46)	(31.46)	(32.03)
Shr scheduled		0.89			-0.08	-0.08	-0.13
caste rural		(0.74)			(0.79)	(0.79)	(0.81)
Shr scheduled		-2.74***			-1.75**	-1.75**	-1.69**
caste urban		(0.66)			(0.73)	(0.73)	(0.75)
Share urban		0.24	-0.35	-0.60	-0.87*	-0.87*	-0.82
(% tot. pop.)		(0.41)	(0.38)	(0.42)	(0.49)	(0.49)	(0.50)
Cumul. land					-1.46***	-1.46***	-1.44***
reform (t-4)					(0.30)	(0.30)	(0.30)
Year trend						1.03***	1.00***
						(0.19)	(0.21)
Trend *x* Post-						0.002*	0.004
1991						(0.001)	(0.003)
Urban share *x*							-0.11
Post-1991							(0.18)
Observations	577	544	544	544	514	514	514
States	16	15	15	15	15	15	15
R-sq. (within)	0.646	0.743	0.726	0.705	0.773	0.773	0.773

*All regressions include state and year effects; Robust standard errors (Huber-White method); * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; all variables are in levels; Jammu & Kashmir is excluded from columns 2-8 due to lack of data.*

Finally, the degree of urbanisation (as measured by the share of urban population in total) is associated with a widening rural-urban headcount poverty difference, although this is not significant at conventional levels. Such a positive association is surprising for at least two reasons. First, the evidence at the global level shows that the process of urbanisation has recently been accompanied by a steep reduction in poverty rates only in rural areas (Ravalion et al., 2007). Second, the evidence presented later on in Chapter 4 suggests that the urbanisation process in India has had a significant poverty reducing effect on rural areas in the period 1981-1999. Given these stylised facts, an increasing urban share of total population would be expected to be associated with a declining rural-urban poverty headcount difference. Interestingly this is the case once I exclude the scheduled caste variables from the set of controls. Column 3 replicates the specification in column 2 without the scheduled caste variables: in this case the coefficient of urban share becomes negative but not significant. This suggests that the urbanisation process may be associated with an increasing concentration of scheduled caste in urban areas (i.e. scheduled caste are more likely to migrate from rural to urban areas than other rural dwellers). This increasing concentration drives the widening rural-urban poverty difference effect of the urban share variable. Results in column 4 show that the relation between urban share and rural-urban headcount poverty difference is partly mediated via GDP per capita. To see this note that the exclusion of GDP per capita from the right hand side variables (as in column 4) almost doubles the magnitude of the urban share coefficient (cf. column 4 with column 3). This result is consistent with the positive correlation between urbanisation and GDP per capita.

The main results are also robust to the inclusion of the cumulative land reform variable (lagged four years as in Besley and Burgess, 2000), which has a negative and significant effect on rural-urban disparities as expected (column 5). Interestingly the inclusion of this variable has the effect of increasing the size and the significance of the GDP per capita coefficients and turns the urban share coefficient to negative and significant even when including the scheduled caste controls (cf. column 5 with column 2). This influence of the land reform on the urban share coefficient can be interpreted as follows. Other things being equal, states which implemented the land reform more thoroughly tended to have faster declines in rural poverty (Besley and Burgess, 2000) and a lower pace of urbanisation (as the evidence in Chapter 3 confirms). Therefore failing to control for the effects of land reform may generate a bias towards a positive association between urban share and rural poverty and thus between urban share and rural-urban poverty difference as well (as in column 2).

The year 1991 marked a watershed in India's economic policy with a wave of pro-market reforms that included a sharp trade liberalisation (Topalova, 2005) and the dismantling of a large part of the state controlled licensing system in manufacturing (the so-called license Raj, see Aghion et al., 2008) among others. It is worth examining whether such important policy shocks have had a different impact on rural and urban areas. To do this I add a time trend and an interaction between this time trend and a post-1991 dummy to the specification in column 5 (see column 6). The results suggest that rural-urban headcount poverty difference increased faster after 1991 relative to their long-term trend, i.e. positive coefficient of the interaction between the trend and the post-1991 dummy.[19] The size of this positive coefficient increases (but becomes less significant) when adding the interaction between the urban share of population and the post-1991 dummy, which tests for the differential impact of urbanisation on rural-urban poverty difference after 1991 (column 7). The positive coefficient of this interaction term indicates that urbanisation has had a larger negative effect on the headcount difference after 1991, although such a differential impact is not significant at conventional levels. These results are in line with those of Topalova (2005), who finds a more adverse effect of trade liberalisation in rural than in urban areas.

The main results are robust also to the use of the other measures of rural-urban disparities described above, as shown in Table 1.3. Income per capita maintains the same U-shaped relation with rural-urban disparities using either the difference in poverty gap or the ratio of real mean consumptions as the dependent variable (columns 1-4). The urbanisation coefficient is not significant when using the poverty gap difference while it is negative (but not significant at standard levels) when using the consumption ratio measure. As in Table 1.2, the exclusion of the scheduled caste controls makes the urban share coefficient more negative (cf. column 3 with column 4). The coefficients of the other controls are generally in line with those of Table 1.2. Interestingly, the female-male ratio in urban areas is consistently positive and significant with both dependent variables. This suggests two stylised facts. First the incidence of severe poverty in urban areas is lower among females than males. Second, females seem to have a higher level of expenditure than males in urban areas.

I next test the rural-urban disparity-per capita income relation using the death rate differential as the dependent variable (column 5). Again, the U-shaped relation

[19] The results are not as clear when I run the same regressions using earlier years (i.e. 1988, 1989 and 1990) as break points (not shown here), confirming that the effective break seems to have occurred in 1991.

between disparities and income emerges strongly using the full set of controls (column 5).[20] The coefficient of the share of urban population is positive but not significant, which may suggest a composition of rural-urban migrants skewed towards relatively more healthy individuals (e.g. as those are the ones potentially more able to withstand the change in living conditions from rural to urban areas). The share of elderly - over 59 – increases the death rates in both rural and urban areas, while the share of working age population reduces it in urban areas but interestingly increases it in rural areas. The latter result may be driven by the relatively high mortality rates among people above 35-40 years of age in rural areas. Also, ceteris paribus a higher share of female in the rural population reduces the disparities in death rates, and the opposite is true for urban areas. This can be related to females having higher life expectancy than males and possibly to the role of women in improving children's healthcare provision by tilting household spending towards social expenditures, and health services in particular. Interestingly, the share of scheduled caste in urban population increases the death rate differential, implying that the scheduled caste population has a relatively lower death rate than the rest of the population.

To summarise, the results support the idea of a U-shaped relation between rural-urban disparities in socio-economic indicators and income per capita. These disparities decrease as income per capita grows for low income levels of economic development until they reach a trough and then they start to increase again. Only a few state-year observations in our dataset appear to lie on the right of this trough. This pattern seems to be consistent with the one characterising countries at the very early stages of development (e.g. LDCs), as highlighted by World Bank (2009, fig. 2a) on a cross-section of countries. This evidence further suggests that there is a tendency for rural-urban disparities to increase in the following transition from low to middle income and then to decrease again after reaching the middle-income stage. As most Indian states in the second half of the 20th century had income per capita typical of the bottom part of today's low income countries, this analysis may be representative of the very early stages of the transition. If that is the case, living standards in rural and urban areas may now diverge as states develop and will eventually return to converge when states reach

[20] The results are different for random effects estimation, but the Hausman test of random vs. fixed effects estimator strongly rejects the null of no systematic difference between coefficients estimated using the two methods. Therefore RE estimation may yield biased coefficients.

much higher levels of incomes. A further result is that urbanisation seems to play some role in shaping this inequality, part of which is driven by its relationship with GDP per capita. It is associated with declining rural-urban disparities in poverty and consumption, while it is positively (but not significantly) associated with death rate differentials.

Table 1.3: Rural-urban disparities and income per capita across Indian states, robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PG diff	PG diff	Mean con. ratio	Mean con. ratio	Death rate diff.	headcount difference	headcount difference
	FE	FE	FE	FE	FE	FE	FE IV
GDP pc	-8.15***	-8.24***	-0.84***	-0.84***	-1.27**		-27.63***
	(1.74)	(1.74)	(0.12)	(0.12)	(0.54)		(9.41)
GDP pc sq.	1.17***	0.99***	0.10***	0.08***	0.20***		3.50***
	(0.24)	(0.24)	(0.02)	(0.02)	(0.06)		(1.03)
GDP growth	5.74***	6.34***	0.26***	0.31***	1.14		8.64
	(1.52)	(1.54)	(0.09)	(0.09)	(0.71)		(13.17)
GDP $pc_{(t-2)}$						-15.89***	
						(4.41)	
GDP pc $sq._{(t-2)}$						2.63***	
						(0.68)	
GDP $growth_{(t-2)}$						2.58	
						(3.01)	
Population	-2.28	-3.52	-0.72***	-0.82***	1.97	-9.98	-5.99
('00000)	(4.14)	(4.30)	(0.22)	(0.24)	(1.30)	(8.47)	(8.33)
Share urban	0.07	-0.02	-0.01	-0.02	0.06	-0.93*	-1.10**
	(0.19)	(0.20)	(0.01)	(0.01)	(0.06)	(0.50)	(0.48)
Rural 15-59	-0.34**	-0.39**	0.01	0.01	0.48***	-0.94**	-0.43
	(0.16)	(0.16)	(0.01)	(0.01)	(0.09)	(0.45)	(0.55)
Urban 15-59	0.09	0.09	0.01**	0.01**	0.09***	0.01	0.07
	(0.09)	(0.08)	(0.00)	(0.00)	(0.03)	(0.19)	(0.19)
Rural 60+	1.18	1.16	0.06	0.04	1.48***	0.34	2.00
	(0.91)	(0.79)	(0.04)	(0.04)	(0.25)	(1.98)	(2.19)
Urban 60+	-0.38	-0.21	-0.11*	-0.08	-1.44***	1.24	0.96
	(1.17)	(1.04)	(0.06)	(0.06)	(0.30)	(2.33)	(2.25)
Fem/male rural	13.29	9.60	2.08	1.72	4.03	101.13**	87.29**
	(19.38)	(20.19)	(1.39)	(1.45)	(9.08)	(45.11)	(41.62)
Fem/male urban	25.29*	16.71	2.83***	1.98***	10.65	-18.32	-13.98
	(13.30)	(12.08)	(0.78)	(0.76)	(6.62)	(29.08)	(27.95)
Shr sched. caste	-0.57*		-0.06***		-0.13	-1.25*	-0.99
rural	(0.29)		(0.02)		(0.10)	(0.70)	(0.70)
Shr sched. caste	-0.04		0.01		0.44***	-1.07	-0.67
urban	(0.33)		(0.02)		(0.12)	(0.81)	(0.79)
Cumulative land	-0.41***	-0.47***	-0.02**	-0.03***		-1.38***	-1.65***
reform (t-4)	(0.14)	(0.13)	(0.01)	(0.01)		(0.29)	(0.33)
Obs.	484	484	484	484	412	486	486
States	15	15	15	15	15	15	15
R-squared	0.707	0.701	0.832	0.825	0.852	0.786	0.785

*All regressions include state and year effects; Robust standard errors (Huber-White method); * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; Jammu & Kashmir is excluded due to lack of data; in column 7 per capita GDP, per capita GDP squared and per capita GDP growth are instrumented with their values lagged two years.*

As mentioned above these results are to be interpreted as statistical relationships rather than in any causal way. This is mainly due to the potential endogeneity of the main regressors. The typical method to address this problem is instrumental variable (IV) estimation. However it is arguably very challenging to find appropriate instruments that explain GDP per capita but do not influence rural and urban areas in any differential way (which represents the necessary exclusion restriction for any such instruments). In the absence of instruments satisfying both conditions, researchers often employ the General Method of Moments (GMM) estimation, developed by Arellano and Bond (1991) and Blundell and Bond (1998). However this method is only efficient asymptotically and is suitable for samples with large N and small T. In this case I have a small N (16 states) with a relatively large T (between 36 and 45 years). Therefore the GMM estimation may not be appropriate for our purposes.

Instead I employ two different specifications using lagged regressors, which may be indicative of the robustness of the GDP per capita coefficients although they do not properly tackle these endogeneity concerns. The results from these specifications (reported in the last two columns of Table 1.3) provide support to the robustness of the coefficients in Table 1.2. In column 6 I run the same specification as in Table 1.2, column 5 but lagging the GDP per capita, GDP per capita squared and GDP per capita growth variables two years. Lagged variables tend to be less endogeneous than the contemporaneous ones, although their persistence over time suggests this is not a satisfactory way to deal with their endogeneity. The U-shaped relationship between rural-urban headcount poverty difference and GDP per capita remains although the coefficient of the linear negative term is lower. On the other hand the coefficient of GDP growth becomes insignificant. In column 7 I employ IV estimation by instrumenting the three GDP per capita regressors by their values lagged two years. Again, this is a not fully satisfactory way of tackling endoegeneity as it is based on weakly exogenous instruments. The results confirm the U-shaped relationship between rural-urban disparity and GDP per capita (while the growth coefficient is insignificant) and the absolute size of both coefficients is larger than in the corresponding FE estimation (cf. Table 1.2 column 5). It is also reassuring that the coefficients of GDP per capita and GDP per capita squared from Table 1.2, column 5, lie in between those in column Table 1.3, columns 6 and 7.

1.3. Convergence and the determinants of towns' growth

While the first part of the chapter is concerned with the way the economic progress is shared between rural and urban areas, this second part relates to how the growth in wellbeing differs across the urban sector, and in particular between towns of different size. This represents the urban-counterpart of the rural-urban inequality analysis of the previous section and is based on an analysis of the (population) growth of Indian towns and urban agglomerations in the 20th century (compiled from the Indian Census). It is important to understand the distinction that the Census makes between towns and urban agglomerations (UAs), as this will feature in the analysis. Towns' definition is described above (see footnote 13). According to the Indian Census (Government of India, 2001) an UA is "a continuous urban spread constituting a town and its adjoining urban outgrowths or two or more physically contiguous towns together and any adjoining urban outgrowths of such towns".[21] In practice the most important agglomerations usually comprise a core large town, surrounded by a number of smaller towns. Sometimes the difference in population between the main town and the UA may be substantial. For example Calcutta UA had a population of 13.2 million in 2001 while the town of Calcutta had 4.6 million. The former comprises over 100 towns, many of which have been incorporated into the UA over time. The incorporation of a new town may bias the analysis as it provides a source of growth which is lumpy and has little relation with the socio-economic characteristics of the UA. Therefore throughout the analysis I try to focus on towns (which are not subject to this problem), while analysing separately UAs.

All towns and urban agglomerations with a population over 10,000 in 1991 are included, while the coverage for urban areas below 10,000 is patchy.[22] This translates into an average of almost 2,500 observations per period for a total of 11 periods (i.e. 1901-2001 with a ten-yearly frequency). Table 1.4 provides summary statistics for the two main variables used in the analysis: population and ten-year population growth rate.

[21] Examples of outgrowth may include railway colonies, university campuses, port areas. These may develop outside the town's statutory limits but within the revenue limits of a village or villages contiguous to the town or city. Each such individual area may not satisfy the minimum population limit to qualify as an independent urban unit but the Census may decide to club it with the town as a continuous urban spread.

[22] Note that the towns' coverage is slightly smaller than that of Chapter 5, which uses data only for the 1971-2001 period rather than for the entire 20th century as here.

The latter is computed using the formula: $g = (u_t / u_{t-10})^{1/10} - 1$ where u_t is population at time t. Both the number of towns and urban agglomerations and their average population increase over the century following India's urbanisation process. Interestingly, the process intensifies over time, as it is indicated by the increase in the mean growth rate of urban areas over the 20[th] century, at least until 1981, after which there is a slight drop in the growth rate.

Table 1.4: Summary statistics for cities' population and population growth, 1901-2001

	Population			Growth rate		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
1901	1,297	17,505	46,007			
1911	1,332	17,291	50,444	1,242	-0.29%	2.79%
1921	1,419	17,634	53,768	1,307	0.33%	2.38%
1931	1,550	19,328	56,662	1,400	1.34%	1.89%
1941	1,716	23,647	83,238	1,537	1.82%	2.26%
1951	2,015	28,183	113,532	1,685	2.08%	2.83%
1961	2,165	34,856	143,645	1,807	2.36%	2.72%
1971	2,534	41,645	183,891	2,152	2.75%	2.39%
1981	3,061	49,506	227,400	2,490	3.22%	2.26%
1991[a]	4,200	50,397	246,537	3,061	2.71%	2.58%
2001[b]	3,845	67,658	322,523	3,845	2.16%	2.18%

Note that the number of observations for the growth rate is lower than that for the towns' population in the corresponding year due to the disappearance of a number of towns from one decade to the next.
a. The mean value for 1991 is not strictly comparable to that of the other years due to the wider cities' coverage; b. the distribution of town for the year 2000 is slightly skewed towards larger towns due to data availability.

The analysis focuses on testing whether the urban areas in India have evolved according to a process of convergence during the 21[st] century. Glaeser et al. (1995) assume a linear influence of city's size on subsequent rate of growth. This is the approach experimented initially, by running the following cross-sectional specification for each year t:

$$g_{it} = \alpha + \beta_1 \ln(u)_{it-10} + \Gamma Z_{it} + \varepsilon_i \qquad (1.2)$$

where u_{it} is population of urban area (town or urban agglomeration) i at time t, g_{it} is the annual rate of growth (as defined above) of town (or UA) and Z is a vector of geographical characteristics, including distance to the state capital, a dummy for the presence of a river in the town, distance to the nearest large town (i.e. above 100,000), level of rainfall, maximum and minimum temperature. The first control captures the possible benefits that a town enjoys by being close to the seat of the state power. The presence of a river may have an ambiguous effect on growth in that it may facilitate transportation but may also constrain the physical expansion of the urban area. The proximity to a large town controls for the degree of access to a sizeable market (both for goods and inputs), but also for the possible diversion of migrants (away from the city under investigation) that the presence of a nearby large town may generate. The climatic controls may capture eventual migrants' preferences for certain weather conditions, as it is the case in the United States with the recent migration towards warmer urban areas (Rappaport, 2007). The test for convergence is captured by the coefficient β_1, with a positive sign indicating a tendency towards divergence (i.e. larger towns grow faster than smaller ones).

Table 1.5 presents the results obtained by running specification (1.2) over a number of different periods. The population coefficient is not significantly different from zero for the period 1991-2001 (columns 1-3). Adding the squared term of town size (column 4) makes the linear population variable significant, suggesting that city size exerts a significant negative influence on subsequent growth, although the intensity of the effect diminishes with size. In particular a town's population appears to have a negative effect on growth for towns up to 62,000 in 1991 after which the effect becomes positive. Only 538 towns (14% of the total) in the sample considered were larger than this threshold in 1991. The analysis further highlights that the distance from the state capital negatively affects the town's growth prospects, and so does the presence of a navigable river.[23] The negative coefficient on distance to the state's capital may have two non-mutually exclusive interpretations: it can indicate the positive impact for a city's prospects of being close to the seat of the political power; it could also represent the effect of market potential on cities' growth, as state capitals are usually large markets as well. The latter effect is more clearly driving the positive coefficient on the distance to the closest large town (i.e. with over 100,000 population). The negative effect on a town's growth prospects of being situated by a river could be related to the physical constraints to growth imposed by the presence of the river or to the danger of flooding which may induce people to settle in towns without rivers. However, these explanations would need further research to be verified. These effects of geographical variables are robust to the inclusion of both state (25 states, column 2) and district effects (443 districts, column 3), which control for local conditions likely to affect urban

[23] This is a dummy variable which takes the value of 1 if the town has a navigable river.

growth, such as policy variables, market characteristics, location. On the other hand the influence of weather related factors on urban growth is explained by state effects (cf. column 1 and 2) and these controls may be excludable according to an F-test. As they are not available for several towns, I exclude these variables in the rest of the analysis.

Table 1.5: The effects of population size on towns' subsequent growth, 1961-2001

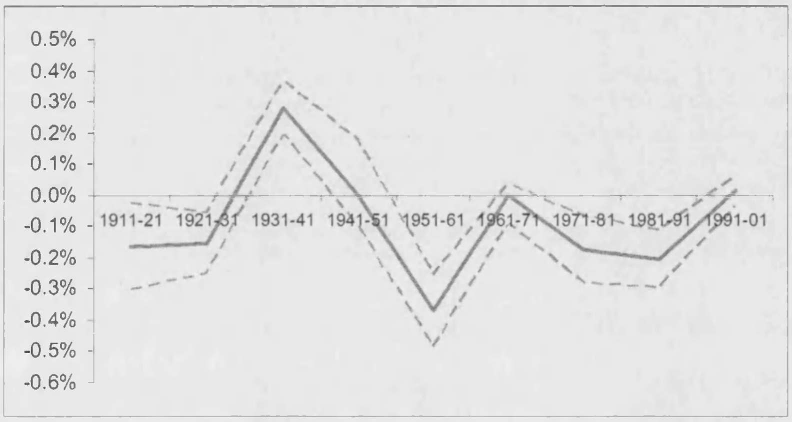
	(1)	(2)	(3)	(4)	(5)	(6)
	1991-2001		**1991-2001**		**1981-1991**	
Pop_{t-10} (log)	-0.0001	-0.00003	0.0001	-0.022**	-0.002**	-0.026
	(0.0005)	(0.0005)	(0.0005)	(0.009)	(0.001)	(0.021)
Pop_{t-10} (log) sq.				0.001***		0.001
				(0.000)		(0.001)
Distance state capital (log)	-0.004***	-0.004***	-0.005***	-0.004**	-0.009***	-0.009***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Distance large town (log)	-0.002***	-0.002***	-0.003***	-0.003***	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
River	-0.005***	-0.004***	-0.004***	-0.004***	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rainfall (log)	-0.001**	0.0053				
	(0.001)	(0.001)				
Max temp (log)	-0.002	0.0003				
	(0.003)	(0.003)				
Min temp (log)	-0.001	-0.0002				
	(0.001)	(0.001)				
State effects	NO	YES	YES	YES	YES	YES
District effects	NO	NO	YES	YES	YES	YES
Observations	3320	3320	3797	3797	3019	3019
R-squared	0.042	0.064	0.209	0.215	0.294	0.299

	(7)	(8)	(9)	(10)	(11)	(12)
	1971-1981		1961-1971		1971-2001	
Pop_{t-10} (log)	-0.002*	-0.050*	-0.0003	-0.045***		
	(0.001)	(0.028)	(0.001)	(0.012)		
Pop_{t-10} (log) sq.		0.002*		0.002***		
		(0.001)		(0.001)		
Distance state capital (log)	-0.006***	-0.004*	-0.004**	-0.003*	-0.005***	-0.004***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Distance large town (log)	-0.001	-0.001	-0.001	-0.002	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
River	-0.002	-0.002	-0.002	-0.002	-0.003***	-0.004***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Pop_{t-30} (log)					-0.001**	-0.054***
					(0.001)	(0.011)
Pop_{t-30} (log) sq.						0.003***
						(0.001)
District effects	YES	YES	YES	YES	YES	YES
Observations	2452	2452	2116	2116	2230	2230
R-squared	0.243	0.273	0.283	0.294	0.344	0.408

*Robust standard errors (Huber-White method) in parenthesis; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; dependent variable: annual growth rate in population in the period indicated.*

The convergence hypothesis is confirmed for the periods 1981-91 (columns 5-6), 1971-81 (columns 7-8), and 1961-71 (columns 9-10), and in the first case the convergence process appears to be more linear than non-linear. The non-linear U-shaped convergence holds also for the earlier periods (1911-61) as well as shown in Table A1 in the Appendix. The convergence hypothesis of towns' growth seems to hold for the entire Post-Independence period, as the plot of the (linear) effect of population size on subsequent growth between 1911 and 2001 shows (Figure 1.5).[24] However, in the last period the convergence process seems to be slowing down and whether it will be reverted remains an open question. The U-shaped relation between growth and initial size holds also when considering long-term urban growth. In the period 1971-2001, the growth reducing effect reverts for towns larger than 39,000, which represent 20% of the towns in the 1971 sample (columns 11 and 12).

Figure 1.5: The effect of city size on subsequent growth rate, 1911-2001



The solid line is derived from the coefficients of linear term in city's population in OLS estimation with geographical and district controls; the dashed lines represent confidence intervals

1.3.1 Robustness checks

I next perform some tests to check whether the main results are robust to different samples and regressors. The results of these checks are presented in Table 1.6. First, I run specification (1.2) for the first period 1901-11, which has a much smaller sample of towns as the urban system was much less developed (column 1). The results are pretty similar to the other periods in terms of U-shaped type of convergence with the trough of the effect for towns of around 18,000, or 21% of the sample in 1901. However

[24] I use the linear coefficients of the population variable on subsequent growth as they can be more easily represented than two coefficients (i.e. linear and non-linear).

the geographical controls have opposite signs than in the other periods. Distance from the state capital has no significant effect, which is likely to be due to the fact that the configuration of Indian states was very different under the British rule and states had different (and generally lower) powers than in the Post-Independence period. As a matter of fact the negative effects of distance to the state capitals become more important after 1931 and in particular since the decade 1941-51 when the modern states' configuration started to emerge after the Independence (see Table A1 in the Appendix). This may hint at the importance of the political explanation (i.e. being close to the seat of the power) to account for the significant effect of this variable in recent periods. Also, being located by a river has a positive effect on the growth prospect in 1901, which suggests that transportation channel was important to spur the development of the urban areas during that period, unlike the present day. This positive effect of being next to a river does not instead hold for the subsequent decades (1911-61) as reported in Table A1 in the Appendix.

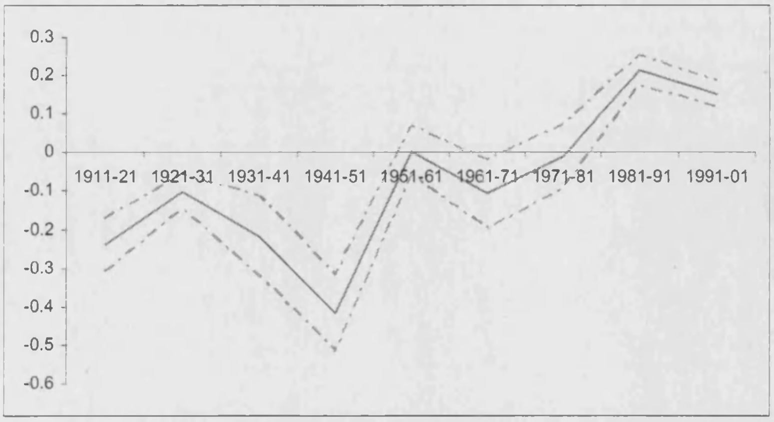
Table 1.6: The effects of population size on towns' growth, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period	1901-11	1991-01	1991-01	1991-01	1981-91	1971-01	1991-01	1981-91
Sample	All	>100 yr old	>20 yr old	Same as in (3)	>20 yr old	>60 yr old	>50 yr old	>50 yr old
Pop_{t-10} (log)	-0.137***	0.002***	-0.034*	-0.031*	-0.029*			
	(0.049)	(0.000)	(0.018)	(0.018)	(0.016)			
Pop_{t-10} (log) sq.	0.007***		0.002*	0.001*	0.001*			
	(0.002)		(0.001)	(0.001)	(0.001)			
Growth(10) $_{t-10}$			0.151***		0.211***			
			(0.035)		(0.040)			
Pop_{t-30} (log)						-0.025***		
						(0.009)		
Pop_{t-30} (log) sq.						0.001***		
						(0.000)		
Growth(30) $_{t-30}$						0.113**		
						(0.051)		
Pop_{t-50} (log)							-0.021***	-0.021***
							(0.006)	(0.007)
Pop_{t-50} (log) sq.							0.001***	0.001***
							(0.000)	(0.000)
Distance state capital	0.003	-0.005***	-0.004**	-0.005***	-0.003*	-0.004***	-0.005***	-0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
River	0.007**	-0.005***	-0.004***	-0.005***	-0.003**	-0.002	-0.005***	0.000
	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance large town		0.001	-0.001	-0.001	0.001	0.001*	0.002*	0.001**
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs.	1222	1163	2673	2673	2452	1470	1838	1676
R-squared	0.452	0.542	0.281	0.258	0.348	0.490	0.374	0.405

*Robust standard errors (Huber-White method) in parenthesis; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; dependent variable: annual growth rate in population in the period indicated.*

When I restrict the 1991-2001 analysis to the sample of towns which have existed since 1901, the coefficient of population becomes linear and positive, thus reverting the previous results (column 2). This suggests that it is the relatively new towns that drive the non-linear convergence process found in Table 5: they are on average smaller than the old towns (the median town in the latter had a population of 42,000 in 1991 against 13,000 for those towns which did not exist in 1901), but tend to grow faster than them (2.2% vs. 2.0% on average in 1991-2001). Column 3 tests for the persistence in growth rates over time. Towns which grew relatively fast in 1981-91 did also well in the following decade. This result is consistent with previous studies for the US (e.g. Glaeser et al., 1995, Glaeser and Shapiro, 2001), but both the extent of persistence and its explanatory power are lower than in those studies. Here if a town grew 1 percent more quickly in the 1980s, it grew on average 0.15 percent more quickly in the 1990s as well, while the same proportion was 0.58 for the US (Glaeser and Shapiro, 2001). The other coefficients are little affected by the introduction of past growth rates (cf. column 4, which uses the same sample as in column 3). The persistence in growth rates also applies to the 1981-91 period, when the coefficient is even higher (0.21, see column 5), and it holds also when using the growth rate over thirty years (i.e. growth in 1971-2001 on growth in 1941-1971, column 6). However, such persistence appears to be a relatively recent phenomenon in India. As shown in Figure 1.6, past growth rates have become to be positively correlated with future ones only since 1971-81. And before then, the relationship was negative and significant: towns that grew faster in a period would grow relatively more slowly in the next period.

Figure 1.6: Persistence in cities' growth rate, 1911-2001



The solid line is derived from the coefficients of lagged growth rate in OLS separate estimation for each decade with population, geographic and district controls; the dashed lines represent confidence intervals.

The last two columns of Table 1.6 check whether the U-shaped effect of population on growth is robust to using larger lags of population. This check addresses the potential simultaneity bias that may arise as the dependent variable is calculated using also the main regressor. The results using population lagged 50 years are similar to those using a 10 year lag. Interestingly, the population coefficients are practically identical for 1981-91 (column 8) and 1991-2001 (column 7).

1.3.2 Urban agglomerations

Following the discussion in section1.2.1, I also examine the convergence hypothesis separately for UAs. I apply the analysis to the larger UAs (i.e. above 50,000 people), as these incorporate a collection of urban outgrowths, that make many of them differ substantially from the core town. Table 1.7 presents the results, which are rather different from those for towns.[25] In particular, the non-linear effect of population for 1991-2001 is reversed and has now more the shape of an inverted-U: the effect is positive but diminishing with size, although it is not significant at the standard level (column 2). On the other hand the U-shaped relationship emerges again for the period 1981-91 and when using growth over a longer period of time as the dependent variable (columns 3 and 4). In columns 5 and 6 I restrict the analysis only to outgrowths of UAs: the results are similar to those for towns, although they are less robust for 1991-2001. This suggests that individual outgrowths of UAs follow a similar convergence pattern to independent towns. The real difference in terms of the effects of population is between entire UAs and towns. Among the former there is a tendency for larger UAs to grow faster than smaller ones in the last decade, although at some point this size effect eventually vanishes. Middle-large UAs grew relatively fast in the nineties, while the very large UAs did not fare equally well, which may suggest that they have already reached a point where congestion costs offset the benefits of agglomeration. Finally, the effect of past growth on subsequent growth is positive and similar to that for towns in the period 1991-2001 (column 7), while it is insignificant for the preceding period (column 8).

[25] These regressions do not include any control variables, as geographical variables are not available for UAs and district effects would have exhausted the degrees of freedom.

Table 1.7: The effects of population size on urban agglomerations' growth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Large UAs	Large UAs	Large UAs	Large UAs	UA outgrowth	UA outgrowth	Large UAs	Large UAs
Period	1991-01	1991-01	1981-91	1971-01	1991-01	1981-91	1991-01	1981-91
Pop_{t-10} (log)	0.0001	0.026	-0.066		-0.013	-0.045***	0.027	-0.070
	(0.001)	(0.021)	(0.045)		(0.015)	(0.017)	(0.021)	(0.044)
Pop_{t-10} (log) sq.		-0.001	0.002		0.001	0.002**	-0.001	0.002
		(0.001)	(0.002)		(0.001)	(0.001)	(0.001)	(0.002)
Pop_{t-30} (log)				-0.024				
				(0.017)				
Pop_{t-30} (log) sq.				0.001				
				(0.001)				
$Growth(10)_{t-10}$							0.161	0.083
							(0.114)	(0.121)
State Effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	91	91	89	98	730	686	89	89
R-squared	0.36	0.37	0.36	0.27	0.08	0.14	0.39	0.37

*Robust standard errors (Huber-White method)in parenthesis; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; dependent variable: annual growth rate in population in the period indicated*

1.3.3 Fixed effect analysis

Given the type of data available I can also test for convergence with a fixed effect specification. This would allow controlling for all time invariant town-specific factors which may affect the growth pattern of a town's population in each time period. For that I use the following variant of equation (2):

$$g_{it} = \alpha_i + \gamma_t + \beta_1 \ln(u)_{it-10} + \beta_2[\ln(u)_{it-10}]^2 + \varepsilon_{it} \qquad (1.3)$$

as in (2) g and u are the growth rate and the town's population respectively, α are towns' effects and γ are year dummies. Note that the presence of the lagged population term in this FE framework is likely to generate endogeneity. In the absence of adequate instruments, I cannot deal with it through GMM estimation, as this uses lagged values of the regressors to instrument them. In (1.3) this strategy would generate perfect collinearity as $g_{it}=\ln(u)_{it}- \ln(u)_{it-10}$. As I am unable to deal with the likely endogeneity in this case, an extra note of caution is needed when interpreting the subsequent fixed effect results.

Table 1.8 presents the results of using specification (1.3). The first column pools all the periods together and the U-shaped convergence is highly significant. In fact the linear term is much more relevant than before and the trough after which towns' population size starts to exert a positive impact on subsequent growth is 139 million, which is much larger than any Indian town. These results are robust to the exclusion of influential observations (column 2).[26] They also hold (although with slightly reduced coefficients) when lagging population by thirty years (column 3). In this case using only the linear term for population gives a highly significant (and negative) coefficient (column 4). However, the share of within group variation explained by the model is considerably lower than with the non-linear specification, suggesting that the latter is more appropriate. The results are valid for both the pre-Independence (1901-1941) and Post-Independence (1951-2001) periods (columns 5-6), although in the former period the negative effect of size on growth for small towns is larger than in the Post-Independence period but then it vanishes more rapidly. The U-shaped relation between growth and initial size applies also to the 1981-2001 period (column 7).

The results are robust to considering UAs and their outgrowths as well. Columns 8 and 9 test the same model as in column 2 for both large UAs and all UAs, obtaining similar results to those for towns. The U-shaped relationship between growth and initial size also holds when considering UAs outgrowths only (column 10) as well as for large UAs in the 1981-2001 period and in the Post-Independence period (columns 11-12).[27]

These fixed effect results suggest that as a town (or UA) grows in size, its rate of growth slows down relative to the rate of growth experienced when its size was smaller. This *within group* result is statistically more important than the cross-sectional one, as it emerges from two facts. First, the population variables in Table 1.8 explain a much larger part of the within group than of the between groups variation. This can be seen by noting that the within group R-squared is over 100 times larger than the between group R-squared across all specifications in Table 8 (not shown here). Second the intensity of the size effect calculated for the median town is larger in the FE analysis than in any of the previous cross-sectional regressions. The within group component is more important for the Pre- than for the Post-Independence period. This is consistent with the findings in figure 6: during the Pre-Independence period a higher growth rate is associated with

[26] I exclude those observations, for which the town either shrunk by more than 5% in any ten-year period or grew by more than 20%.

[27] These results are also robust to the inclusion of state-year effects, which control for time varying state-specific urban systems. In addition, the findings are equally valid using a balanced panel, i.e. conditional on the existence and the statistical reporting of towns in every year between 1951 and 2001. Results of these further robustness tests are available from the author upon request.

lower growth rate in the subsequent decade, while the persistence of growth rates emerges only in the latter part of the 20th century.

Table 1.8: The effects of population on cities' growth, fixed effects, 20th century

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All towns	All towns	All towns	All towns	All towns	All towns
Period	1901-2001	1901-2001	1901-2001	1901-2001	Pre-Ind	Post-Ind
Pop_{t-10} (log)	-0.075***	-0.045***			-0.161***	-0.046***
	(0.007)	(0.004)			(0.021)	(0.006)
Pop_{t-10} (log) sq.	0.002***	0.001***			0.005***	0.000
	(0.000)	(0.000)			(0.001)	(0.000)
Pop_{t-30} (log)			-0.033***	-0.011***		
			(0.003)	(0.001)		
Pop_{t-30} (log) sq.			0.001***			
			(0.000)			
Fixed eff.	YES	YES	YES	YES	YES	YES
Year eff.	YES	YES	YES	YES	YES	YES
Influential obs.	YES	NO	NO	NO	NO	NO
Observations	20526	20352	13064	13064	5387	14965
Towns/UAs(No.)	4207	4196	2558	2558	1560	4196
R-sq. within	0.310	0.286	0.594	0.112	0.490	0.284
	(7)	(8)	(9)	(10)	(11)	(12)
Sample	All towns	UAs > 50,000	All UAs	UA outgrowths	UAs > 50,000	UAs > 50,000
Period	1981-2001	1901-2001	1901-2001	1901-2001	1981-2001	Post-Ind
Pop_{t-10} (log)	-0.105***	-0.035***	-0.039***	-0.058***	-0.189***	-0.106***
	(0.018)	(0.009)	(0.007)	(0.010)	(0.045)	(0.023)
Pop_{t-10} (log) sq.	0.001*	0.000	0.001***	0.001***	0.004**	0.002*
	(0.001)	(0.000)	(0.000)	(0.000)	(0.002)	(0.001)
Fixed eff.	YES	YES	YES	YES	YES	YES
Year Eff.	YES	YES	YES	YES	YES	YES
Inflluen. obs.	NO	NO	NO	NO	NO	NO
Observations	6878	1743	2863	3747	317	954
Towns/UAs(No.)	4189	237	374	872	228	237
R-sq. within	0.457	0.311	0.325	0.248	0.631	0.518

*Robust t-statistics in parenthesis; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; dependent variable: annual growth rate of population in each decade in the period indicated.*

1.4 Conclusions

This chapter has described some aspects of both the rural-urban and the within urban transitions in the Indian economic development process. The results of the former analysis performed on Indian states support the idea of a U-shaped relation between rural-urban disparities in socio-economic indicators and the level of income per capita. Such disparities decrease as income per capita grows but at the diminishing rate until they reach a trough and then there is some indication that they may start to rise again.

This is the mirror-image to the traditional U-shaped inequality-income curve hypothesised by Kuznets (1955). One way to reconcile these rural-urban results with the national ones à la Kuznets on income inequality is to put the Indian states in the international context. The rural-urban disparities-income relationship in Indian states seems consistent with the one characterising countries at the very early stages of development (World Bank, 2009). As most Indian states in the second half of the 20th century had income per capita typical of the bottom part of today's low income countries, this analysis may be representative of the very early stages of the low to lower-middle income transition. The international evidence further suggests that there is a tendency for rural-urban disparities to increase with income in the following transition from low to middle income and then to decrease again after reaching the middle-income stage (World Bank, 2009). If that is the case, living standards in rural and urban areas may now diverge as Indian states develop and will eventually return to converge when states reach higher levels of incomes. Two further findings of the analysis may help policy-makers prepare for (and eventually mitigate) these changes. First, land reform has a negative association with rural-urban disparities helping rural areas to catch up relative to urban areas. Second, rural-urban disparities decrease with urbanisation. Hence to the extent that the relation between urbanisation and rural-urban disparities follows the same path as in the past, urbanisation may dampen the eventual future increase in disparities.

In addition using a large dataset of Indian towns for the 20th century, I found a tendency towards convergence in growth rates among Indian towns across all decades of the century. Smaller towns grow faster than larger ones, although this pattern holds until a certain size threshold after which growth rate becomes an increasing function of size. This finding contrasts with the concern of policy-makers about urban concentration with large towns growing too quickly relative to small and medium ones.[28] On the other hand very large cities seem to be less subject to this negative size effect on growth than large ones, which may be of some concern with respect to the Indian mega-cities. Unlike the evidence for US towns, which shows clear persistence in growth rates (Glaeser et al., 1995), Indian towns exhibit a mean reversing process in

[28] This fear stimulated for instance the Integrated Development of Small and Medium Towns programme launched by the Indian Government in the sixth five year plan to foster the growth of small and medium towns through public investment projects.

growth rates (i.e. the growth rate in one period is negatively correlated to the period before) in each decade of the century except in the last two. This suggests that in the first half of the century the determinants of city growth in India had little to do with time invariant factors, which would have determined persistence in growth rates. This pattern changed in the last two decades of the 20^{th} century, when rapidly growing towns have eventually maintained their sustained rate of growth over time and vice-versa for low growing towns. This may suggest that similarly to urban areas in the US some persistent characteristics of towns (such as sectoral and population composition) may have become important determinants of their rate of growth. To the extent that such characteristics become to dominate towns' growth rate, initial size may cease to be a significant determinant of subsequent growth (as in the US). But further data would be needed to test for this hypothesis with Indian towns.

Appendix 1.1
Methodological note to the construction of poverty measures

The poverty headcount ratio and the poverty gap index are two standard Foster Greer Thorbecke (FGT) measures of poverty. FGT poverty measure for a given population is defined as:

$$H_{\alpha}^{i} = \int_{0}^{z_i} \left(\frac{z_i - y}{z_i} \right)^{\alpha} f(y)dy$$

where z_i is the poverty line in the area i (with i = [rural, urban]), and $f(y)$ is the distribution function of monthly per capita expenditure (in this case), with the population ordered in ascending order of y (i.e. starting from the poorest).

Headcount Index

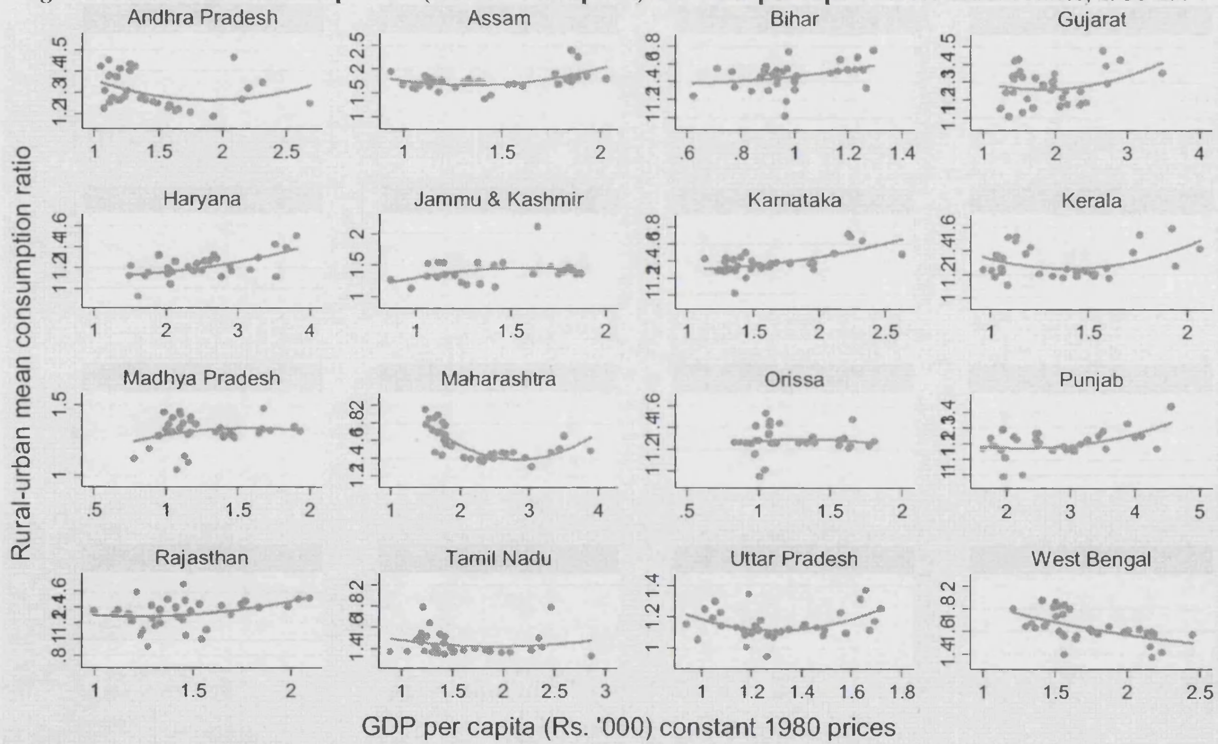
The headcount ratio is computed by setting α=0, thus it represents the proportion of the population below the poverty line in a certain geographical unit (poverty rate). The poverty lines used by the dataset are those recommended by the Planning Commission (1993) and are as follows. The rural poverty line is given by a per capita monthly expenditure of Rs. 49 at October 1973-June 1974 all-India rural prices. The urban poverty line is given by a per capita monthly expenditure of Rs. 57 at October 1973-June 1974 all-India urban prices (see Datt (1995) for further details on the rural and urban cost of living indices and the estimation of poverty measures).

Poverty Gap Index

This is computed by setting α=1 and is defined as the mean distance below the poverty line as a proportion of the poverty line where the mean is taken over the whole population, counting the non-poor as having zero poverty gap. That is the mean shortfall from the poverty line (counting the non poor as having zero shortfall), expressed as a percentage of the poverty line.

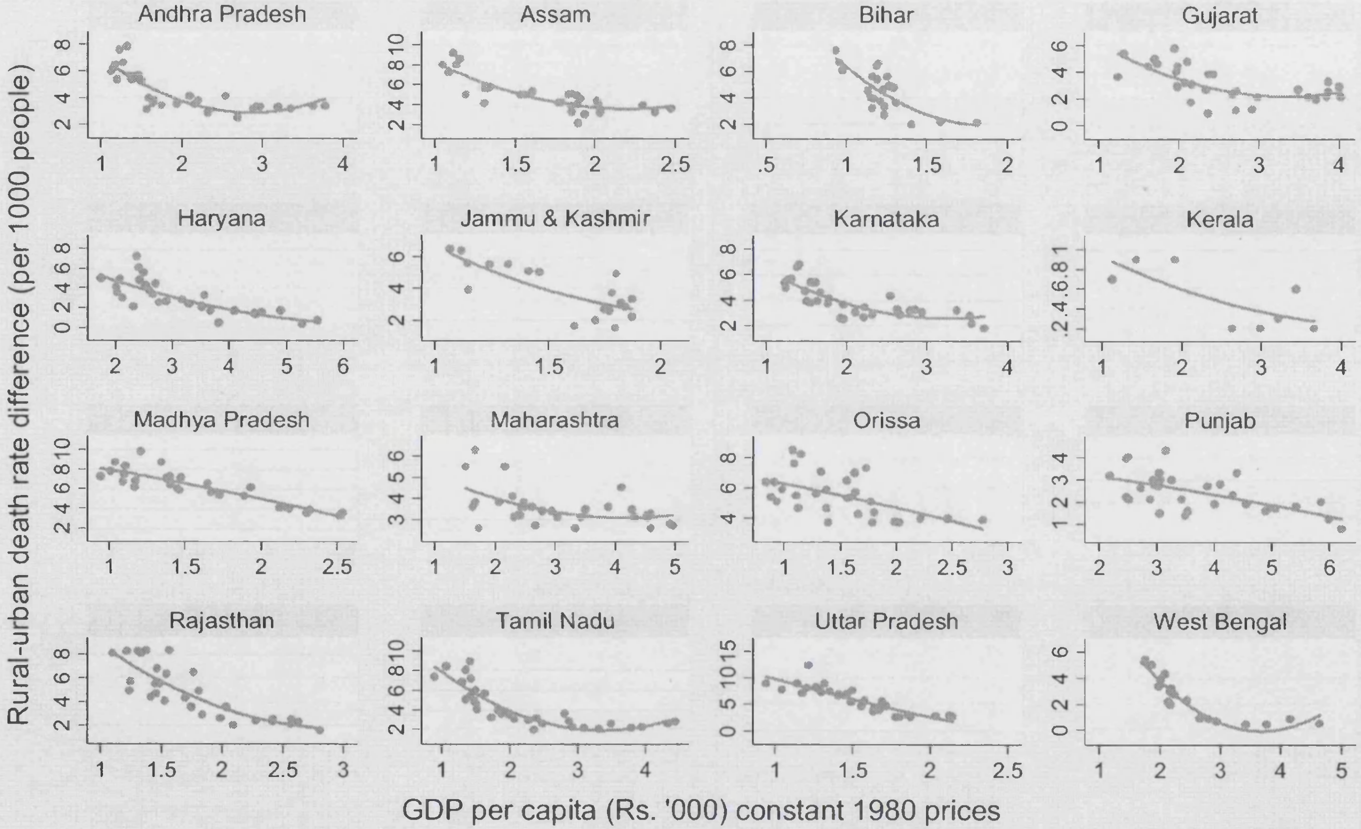
Appendix 1.2
Other Figures and Tables

Figure A1.1: Rural-urban disparities (Mean consumption) and GDP per capita across Indian states, 1958-1994



Note: The X axis measures the ratio between the average expenditure in Rs. in urban areas (measured at all India urban prices) and the average expenditure in rural areas (measured at all India rural prices). A ratio higher than 1 indicates that the urban dwellers consume on average more than the rural ones. An increasing ratio therefore implies a widening gap in living standard between urban and rural areas and vice-versa.

Figure A1.2: Rural-urban death rates disparities and GDP per capita across Indian states, 1971-2001



Note: The X axis measures the difference in the number of deaths per 1000 people in rural and urban areas

Table A1: The effect of city size on subsequent growth rate, 1911-1961

	(1) 1911-21	(2) 1921-31	(3) 1931-41	(4) 1941-51	(5) 1951-61
Pop_{t-10} (log)	-0.115***	-0.039**	-0.041**	-0.043*	-0.078***
	(0.026)	(0.018)	(0.018)	(0.024)	(0.030)
Pop_{t-10} (log) sq.	0.006***	0.002**	0.002**	0.002*	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance state	-0.002	-0.000	-0.003*	-0.008***	-0.003
capital (log)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
River	-0.001	-0.001	-0.004*	0.001	-0.000
	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)
District Eff.	YES	YES	YES	YES	YES
Observations	1287	1380	1517	1663	1780
R-squared	0.382	0.321	0.305	0.345	0.353

*Robust standard errors in parenthesis; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level; dependent variable: annual growth rate in population in the period indicated.*

Chapter 2. On the connection between the agricultural sector and urbanisation patterns in a developing country

2.1. Introduction

For the first time in its history the world is estimated to have become more urban than rural in 2007. The world's urban population was estimated at 3 billion in 2003 and is predicted to rise to 5 billion by 2030 (UN, 2005). Most of this growth will take place in developing regions, many of which are (and will be) experiencing substantial increases in urbanisation rates. While a significant amount of literature has focused on the costs and the benefits of the urbanisation process in developing countries, the investigation of the factors shaping it has received less attention. In particular, key questions on *what* socio-economic factors determine urban hierarchies in the process of development, and *how* they do so, are left largely unaddressed. As noted by Overman and Venables (2005), this gap is in the development economics as well as in the urban economics literature. The former has mainly approached urbanisation issues in the context of rural-urban migration, although this has occurred without any significant analysis of intra-city dynamics. The urban economics approach has concentrated on city-level dynamics of output production (e.g. Henderson, 1985 and 1988), mainly focussing on pull factors (city's labour demand) and usually abstracting from the modelling of rural-urban migration.

This chapter and the next one aim to help fill this gap, by focussing on the role of push factors in the urbanisation process in an economy experiencing rural-urban migration (i.e. a developing country). In particular the work tries to combine insights from the urban economics and the development economics literature, focusing on an explicit role for the agricultural sector in the determination of the urban hierarchy and urbanisation rate. More importantly, chapter 3 uses a novel empirical strategy based on intra-country (rather than cross-country) analysis to test for the effects of agricultural variables on urban primacy and on urbanisation in the light of the theoretical framework developed in this chapter. The test is applied to the case of India, by constructing a dataset of Indian state-wise urban agglomerations over the 1951-2001 period. The remainder of this chapter is organised as follows: the next section puts the analysis in the context of the existing literature; section three develops a simple analytical framework linking equilibrium city size and urban population to the agricultural sector; section four concludes summarising the predictions of the model which will help guide the empirical analysis in chapter 3.

2.2. Related literature

What do we currently know about the connections between the rural sector and urbanisation in the development process? A long standing tradition in development economics (see inter alia Lewis, 1954; Ranis and Fei, 1961; Harris and Todaro, 1970) has highlighted the importance of the structure of the agricultural sector (e.g. shape of the production function, unemployment, factors' endowment) as the main push factor in the urbanisation process. This type of analysis has concentrated on the determinants of rural-urban migration. Lewis (1954) describes the urbanisation process as one fuelled by rural people who migrate to the urban area without affecting the rural wage (due to the over-supply of labour in the rural areas). In Lewis' view this is a healthy process as it reduces the excess labour supply in the rural area, and provides labour to the expanding industrial/urban sector. This model assumes full employment in the urban sector throughout the process of development, which contrasts with the empirical observation of rural-urban migration in the presence of open urban unemployment. To resolve this issue Harris and Todaro (1970) model rural-urban migration as a function of the gap between the expected net urban wage (i.e. the product of the urban wage in the formal sector and the probability of being employed in that sector) and the rural wage. Therefore the former can be higher than the latter even in the presence of unemployment.

These works have been key to understand the rural-urban migration process in developing countries; however they generally assume a perfectly elastic labour supply curve. This is in contrast with more recent evidence on developing countries, reviewed in Lall et al. (2006), which shows that the labour supply curve is upward-sloping in the agricultural wage (e.g. Brueckner, 1990 and Ravallion and Wodon, 1999). Moreover, as better micro-level data emerged, the search for determinants of rural-urban migration has expanded, including the role of other factors, such as social networks, skills and education.

Despite the substantial effort in understanding the migration mechanism underlying the labour supply curve, the literature has been little concerned with linking this curve to the process of city formation. An exception is Brueckner (1990), who models the relationship between rural-urban migration and city growth using an urban economic framework. He finds theoretically and empirically (using a small cross-section of developing countries) that the size of the largest city of a country and the urbanisation rate are both inverse function of the rural-urban income ratio. While this finding points towards the importance of the urban-rural wage gap in determining the pace of urbanisation, it does not inform on the role of this gap in determining the type of urbanisation pattern.

A somewhat complementary approach is the one taken by the urban economics literature, which is instead focused on the modelling of the formation and development of cities (see Henderson, 1985 and 1988). These typically emerge as a consequence of increasing returns to agglomeration of economic activity. These models derive conditions for equilibrium city size determined by the intersection of the urban labour demand and supply curves. The former is the result of the tension between productivity enhancing agglomeration effects and productivity reducing congestion costs (both increasing in the city's population). The labour supply curve is usually treated as exogenous and it is assumed to be perfectly elastic in wage. This is justified by two assumptions underlying most models: net wages are equalised across urban areas and cities' population changes only via inter-city migration (along with natural population growth). This latter assumption is consistent with the study of fully urbanised systems typical of high income countries, which do not experience rural-urban migration any longer. Although a number of urban papers have started to focus on developing countries (e.g. Venables, 2004, Overman and Venables, 2005, Au and Henderson, 2006a and 2006b, da Mata et al., 2007, Duranton, 2007), the implications of rural-urban migration for the distribution of the urban population across cities are still largely unclear. This chapter tries to shed some light on these implications by combining insights from the development economics and the urban economics literature.

I focus in particular on urban primacy, i.e. the share of urban population concentrated in the largest city. There are a number of reasons that justify this focus. First in the developing world urban primates often represent a major urban concern for policy-makers. These primates often have the highest intra-city disparities in the country, such as in the case of Nairobi, Johannesburg and Lagos (World Bank, 2009). In India itself Delhi's rapid population growth (from 1.4 million in 1950 to 15.6 million today) has been accompanied by an increase in the number of slum clusters from 200 to 1,160 (World Bank, 2009). A lot of urban primates in the developing world have areas within the city with far higher incidence of severe poverty than in rural areas (Satterthwaite, 2004). In addition urban primates tend to foster the concentration of income and production in a highly localised area often contributing to generate a high degree of spatial inequality within the country. Greater Cairo produces 50 percent of Egypt's GDP, using just 0.5 percent of its land area and politicians generally view such economic imbalance disapprovingly (World Bank, 2009). A variety of policies have been tried to slow down the rate of growth of urban primates (e.g. migration restrictions, change of the capital, investment in alternative cities).

A second reason to focus on primacy is that the largest city is usually a relevant deviation from the remarkable empirical regularity of the city size distribution defined by the Zipf's law. This states that the city size distribution is approximated by a Pareto distribution with exponent equal to minus one. Gabaix (1999) shows that Zipf's law can be generated when the urban growth satisfies Gibrat's law for the mean and variance of the growth rate in steady-state (i.e. the mean and variance of cities growth rate are independent of city size).[29] The largest city is often an outlier in the distribution and especially in developing countries it tends to be larger than Zipf's law would predict (see evidence in Soo, 2005). Finally, the use of urban primacy has the advantage of requiring little data relative to other measures based on the entire city size distribution. This is desirable in contexts where data on cities' population over time may not be as complete as necessary.

The literature in search of the determinants of urban primacy can be divided into two major strands. One has tried to find a deterministic relationship between the level of economic development and that of urban concentration in a country, hypothesising an inverted U-shaped relationship (e.g. Williamson, 1965, Wheaton and Shishido, 1981,

[29] Such a property is confirmed by Eaton and Eckstein (1997) for France and Japan and by Ioannides and Overman (2003) for the US. Duranton (2006) provides also an economic rationale for Zipf's law through a model where city growth is driven by innovations generated by R&D activities.

Alperovich, 1992) but without reaching any undisputed conclusion. A more analytic (and fruitful) approach is the one which focuses on country-specific pull factors, e.g. political and institutional factors as the main determinants of urban primacy. Ades and Glaeser (1995) develop a model where a country's population is attracted to the capital city to maximise its benefits through rent-seeking activities. They show theoretically and empirically that countries with dictatorial and politically instable regimes and where the largest city is the capital tend to develop higher levels of urban primacy, while transport infrastructure investments and trade openness reduce urban primacy. This last finding is in line with the theoretical predictions of Krugman and Livas Elizondo (1996), who explain urban primacy in developing countries as the result of an inward-looking process of industrialisation. The empirical analysis by Moomaw and Shatter (1996) and Davis and Henderson (2003) provide further evidence on the influence of political factors on the degree of urban primacy.

As mentioned above these analyses of urban primacy do not include an explicit role for the push factor in the urbanisation process as they concentrate only on the pull factors. A relevant exception is Puga (1998), who uses a new economic geography model to show that a higher wage elasticity of labour supply from the rural sector widens the range of transport costs values for which a primate city equilibrium emerges. This chapter adopts a similar focus to Puga's (concentrating on the role of push factors to explain urban concentration) while it takes a different approach for the reasons explained below (although the conclusions here may be reconciled with those of Puga).

2.3. A simple theoretical framework

The objective of this section is to develop a highly stylised model which links the conditions in the agricultural sector to the urbanisation pattern and to the determination of urban primacy in particular. Using a simple urban economics framework, I will highlight some channels through which such a role may be warranted. The basic intuition is that the conditions in the agricultural sector shape the urban labour supply curve, especially during the early stages of the urbanisation process.[30] As noted by Duranton (2008), the available evidence suggests that the slope of the supply curve

[30] Note that I use the terms "city" and "urban area" interchangeably in the paper, unless otherwise specified (as in the empirical section).

seems to be determined mainly by the conditions in the rural sector (i.e. the decision whether or not to migrate from rural to urban areas depends more on the conditions in the rural than in the urban sector). This influence of the rural conditions would then play a key role in the determination of the equilibrium level between the costs and benefits of urban agglomeration.

I describe a stylised urban economics model, focussing on the labour supply curve, rather than on the labour demand (as in most urban economics literature) in determining the equilibrium city size. The model is developed in a partial equilibrium framework considering only the equilibrium in the urban sector while treating the rural sector as exogenous. In particular, the net rural wage, which captures the conditions of the agricultural sector, is exogenous to the model, i.e. it is not affected by the changes in the urbanisation rate. This assumption reflects Lewis' (1954) idea of "unlimited labour supply" in the rural sector. In such a context rural-urban migration mops up the excess labour supply in the rural areas without affecting rural wages.[31] This is a situation close to that of the rural sector in India in the Post-Independence period, which is the focus of the empirical analysis in the next chapter.[32] This model is used to derive ways in which the elasticity of labour supply may impact on the degree of urbanisation and importantly of urban primacy. The framework should provide a rationale for the empirical work in chapter 3, although it makes unambiguous predictions only on the relation between the elasticity of labour supply and the degree of urbanisation (but not for urban primacy).

Assume a closed region with two potential locations for cities, 1 and 2 and two sectors: a modern urban sector and a traditional agricultural sector. The size of the city depends on the number of firms-workers in it. Both sectors use labour to produce their goods. Labour is imperfectly mobile between the rural and the urban sector, but is perfectly mobile between cities. This type of labour mobility arises for example if there is a fixed cost to moving from the rural to the urban sector (e.g. learning a different technology of production; adjusting to a new lifestyle, etc.), which an urban dweller

[31] In Lewis' framework the high fertility of the rural sector determined rural labour to be employed beyond the point where the marginal product was zero. Therefore rural-urban migration did not produce any changes in the agricultural wage.
[32] I also discuss below the possible implications for the framework of relaxing the assumption of the exogeneity of the rural wage.

does not face when moving between cities. I analyse below the effects of relaxing this assumption. Labour is entirely concentrated in the agricultural sector in the beginning.

2.3.1 Urban labour demand

The urban production function is location specific. As in Venables (2004), let us assume that labour productivity has two components: a location specific component which is linked to the city's characteristics such as its geography or administrative status (e.g. access to the sea, seat of the political power). This component acts like a parameter shifter for the city's labour productivity function. The second component captures the idea that the productivity in the city is driven by its number of workers. Thus changes in productivity are a function of city size only. The aggregate production in urban area *i* is: $Q_i = f_i(n)$ for i=1, 2, where n is the number of workers in the city (equal to the city's population), with $\partial Q_i / \partial n > 0$ and importantly $\partial^2 Q_i / \partial n^2 > 0$. The latter condition describes the key feature of the urban production function, i.e. increasing returns to labour at the city level. A worker in city *i* is paid its marginal product, i.e.:

$$w_i^U = \partial Q_i / \partial n = w_i(n) \qquad (2.1)$$

where the location specific component determines the differences between cities in the earning function for any level of city size *n*. The increasing returns at the city level implies that the nominal wage is an increasing function of city size, i.e. $\partial w_i / \partial n_i > 0$. Since a great deal of attention has been devoted to determine and model the micro-foundations of such an urban production function, I don't delve into any details of its functioning here. One can think for instance of how learning mechanisms operate more effectively in denser environments (i.e. more populated cities), where more frequent direct interactions between economic agents favour the creation, diffusion and accumulation of knowledge.[33]

Expression (2.1) determines the local nominal wage in *i*; in order to obtain the city demand for labour I need to add the costs of living to (2.1) so as to compute the net wage function. These costs can be thought of as related to the increasing congestion as

[33] See Duranton and Puga (2004) for a complete description of the forces shaping the urban production function.

the city expands, such as the costs of commuting to the central business district, or the housing costs (in the presence of a relatively sticky supply of land). These are usually assumed to be an increasing and strictly convex function of *n*. For the sake of simplicity, assume that the cost function is the same across locations, so that a worker will face the cost function $c(n)$, with $c'(n) > 0, c''(n) > 0$.[34] The net wage for a worker in city *i* would then be determined as:

$$\omega_i^U(n) = w_i(n) - c(n) \qquad (2.2)$$

Expression (2.2) presents the stylised version of the trade-off between agglomeration and congestion forces at work when a city expands (i.e. when *n* grows). As in standard urban economics let us assume that the shapes of $w_i(n)$ and $c(n)$ are such that ω_i increases for small levels of *n*, it peaks and then decreases.

Following Combes et al. (2005), it is effective to study the net wage curves through diagrams that for each city focus on the inverse demand for labour rather than on output per worker as a function of the city's workforce. The inverse demand relates the net wage of workers to the size of the city's population (i.e. its workforce). The inverted U-shaped solid curves in the upper part of Figure 2.1 show the net wage curves for city 1 and city 2 (we assume that $\omega_1^U(n) > \omega_2^U(n) \forall n$).[35] The horizontal sum of the two net wage curves represents the total labour demand of the urban sector (D_L). In Figure 2.1, only the decreasing part of the D_L curve is represented as this is the portion of the curve which coincides with stable cities' equilibria.

2.3.2 Urban labour supply

Given the demand for labour a labour supply curve is needed in order to find the equilibrium urbanisation level and cities' size. Cities will face an horizontal labour supply curve as labour is perfectly mobile within the urban sector. This implies that the equilibrium level of urbanisation will determine the size of the two cities given the net wage curves. Also, the net wage will be equalised across locations in equilibrium.

[34] Introducing a location specific cost function would not change the results in any significant way,

[35] This is because ω_1 *(n)* is augmented by a location specific externality, which increases labour productivity for each level of city size. One could think of such externality as determined for example by the advantage of being close to the political power, in case City 1 were the capital, and/or by its geography.

The total urban labour supply curve is determined solely by rural-urban migration, as the rural sector provides the labour force needed to the urban sector to industrialise.[36] Such an assumption reflects the idea that during the process of urbanisation, and in its early stages in particular, most of the population growth in urban areas occurs through rural-urban migration.[37]

As migration is a function of the urban-rural net wage gap, the urban labour supply curve can be modelled as a function of net rural and urban wages:

$$N^S = N(\omega^U, \omega^A) \qquad (2.3)$$

with $\partial N^S / \partial \omega^U > 0$; and $\partial N^S / \partial \omega^A < 0$, where ω^U is the net wage in urban areas and ω^A is the net wage in rural areas. As mentioned above, ω^A is exogenous in the model as the large pool of idle labour in the rural sector guarantees the rural wage to be unaffected by rural-urban migration. Expression (2.3) states that the supply of labour from the agricultural to the urban sector is a positive function of the urban net wage and a negative function of the rural net wage. The hypothesis of an upward sloping labour supply curve is supported by a large empirical literature (see Duranton, 2007 and Lall et al., 2006 for a review). The upward sloping labour supply curve is determined by the imperfect mobility of labour between the rural and the urban sectors. For any urban-rural wage gap, there will be only a certain number of rural dwellers willing/able to move to the urban areas. The flatter the labour supply curve the higher this number will be. One can think of various mechanisms determining this imperfect inter-sectoral labour mobility, such as idiosyncratic preferences for rural areas, high fixed costs of rural-urban migration, etc. It is important to understand how ω^A affects this curve, as this is the main identifying assumption of the empirical analysis in chapter 3. It seems reasonable to assume that N^S is a function of the ratio between the urban and the rural net wages rather than of their difference. In other words changes in ω^A would affect the slope of the labour supply curve, i.e. its elasticity with respect to the net urban wage. To see this take one special case of (2.3), i.e.

$$N^S = \frac{\omega^U(n)}{\omega^A} \qquad (2.3')$$

[36] The assumption underlying this labour supply curve is that there is no other source of labour for the urban sector, i.e. fertility rate is equal to death rate in the urban areas.

[37] This assumption seems warranted for low income countries, most of which have recently been (or are currently) undergoing such process.

the slope in (2.3') is determined by $(\partial\omega^U / \partial n)\times\frac{1}{\omega^A}$; therefore an increase in ω^A would make N^S steeper, i.e. less elastic.

Why is this assumption warranted? The wage ratio, unlike the wage difference, captures the concept of relative gap, which in the presence of imperfect rural-urban labour mobility seems a more appropriate determinant of the elasticity of labour supply than the absolute gap. Consider the same monthly net wage difference between the urban and rural sector, say 1000 Rupees, with two different (average) net rural wages, say 100 and 10,000 Rupees.[38] In the former case the expected increase in wage for a rural-urban migrant is 1000% while in the latter case it is only 10%. Given the decreasing marginal utility of income, the latter case should provide a lower incentive to migration than the former. This feature is captured by the wage ratio as the determinant of the labour supply. This discussion implies that in the remainder of the chapter (and in the next one as well) I will treat changes to the net rural wage as affecting the slope of the labour supply curve. It is worth noting that the main results of the model would hold even when modifying this assumption, by making the changes in rural wage shift the labour supply curve but not its slope.

2.3.3 Equilibrium

Plotting this labour supply curve in the same $[\omega,n]$ space (for a given value of ω^{A}') allows one to find the equilibrium level of total urban population. In Figure 2.1 this is determined by point A (where the urban labour demand and urban labour supply curves intersect). At this point the equilibrium urban net wage is ω', which determines the equilibrium cities' size given each city's net wage curve. Hence, the equilibrium net wage-size would be $[\omega',N_1]$ for city 1 and $[\omega',N_2]$ for city 2, with $N_1 > N_2$. Note that net wages are equalised across urban areas due to the perfect labour mobility assumption. It is useful to transpose the same situation to the two cities diagram in the bottom part of Figure 2.1 in order to highlight changes in the relative size of the cities. This diagram has two origins (O_1 and O_2), each of which representing the origin for one city. The width of the diagram represents the total urban population (O_1-O_2) and it changes

[38] Technically the urban wage facing a potential migrant is an expected rather than an actual wage. In the presence of urban unemployment this would be the product of the urban wage and the probability of obtaining a job in the urban sector.

according to the level determined in the upper part of the diagram. The equilibrium is determined at point E, where the cities' net wage curves and the cities' labour supply curve S_L intersect. The population of city 1 is O_1-P=N_1 and that of city 2 is O_2-P=N_2.

2.3.4 Comparative statics

Consider now a situation with a higher level of agricultural wage, ω^{A}''. The labour supply curve is now steeper (i.e. less elastic), and it is represented by a dashed line in the upper part of Figure 2.1. It intersects the labour demand curve at point A', which corresponds to a lower level of urban population (O_1-O_2') and to a higher level of net urban wage (ω'') than before. Given the same net wage curves for cities, the equilibrium in the two cities' diagram (lower part of Figure 2.1) would be at point E'. The cities' population would be O_1-P'=N_1' and O_2'-P'=N_2' for city 1 and city 2 respectively (with $N_1' < N_1$ and $N_2' < N_2$). Therefore a higher level of rural wage generating a steeper labour supply would be associated with a lower level of urban population. What is the degree of urban primacy associated with this higher level of ω^A? Ruling out the possibility of having unstable equilibria (i.e. considering only the decreasing portions of the sum of the two net wage curves), the answer will depend on the relative slope of the two curves. It is useful to look at the upper part of the diagram to see why this is the case. The condition for a lower elasticity of labour supply (a steeper S_L) to be associated with a less concentrated urban system is that: $\frac{N_1'}{N_1'+N_2'} < \frac{N_1}{N_1+N_2}$, or equivalently:

$$\frac{N_1}{N_2} > \frac{N_1'}{N_2'} \qquad (2.4)$$

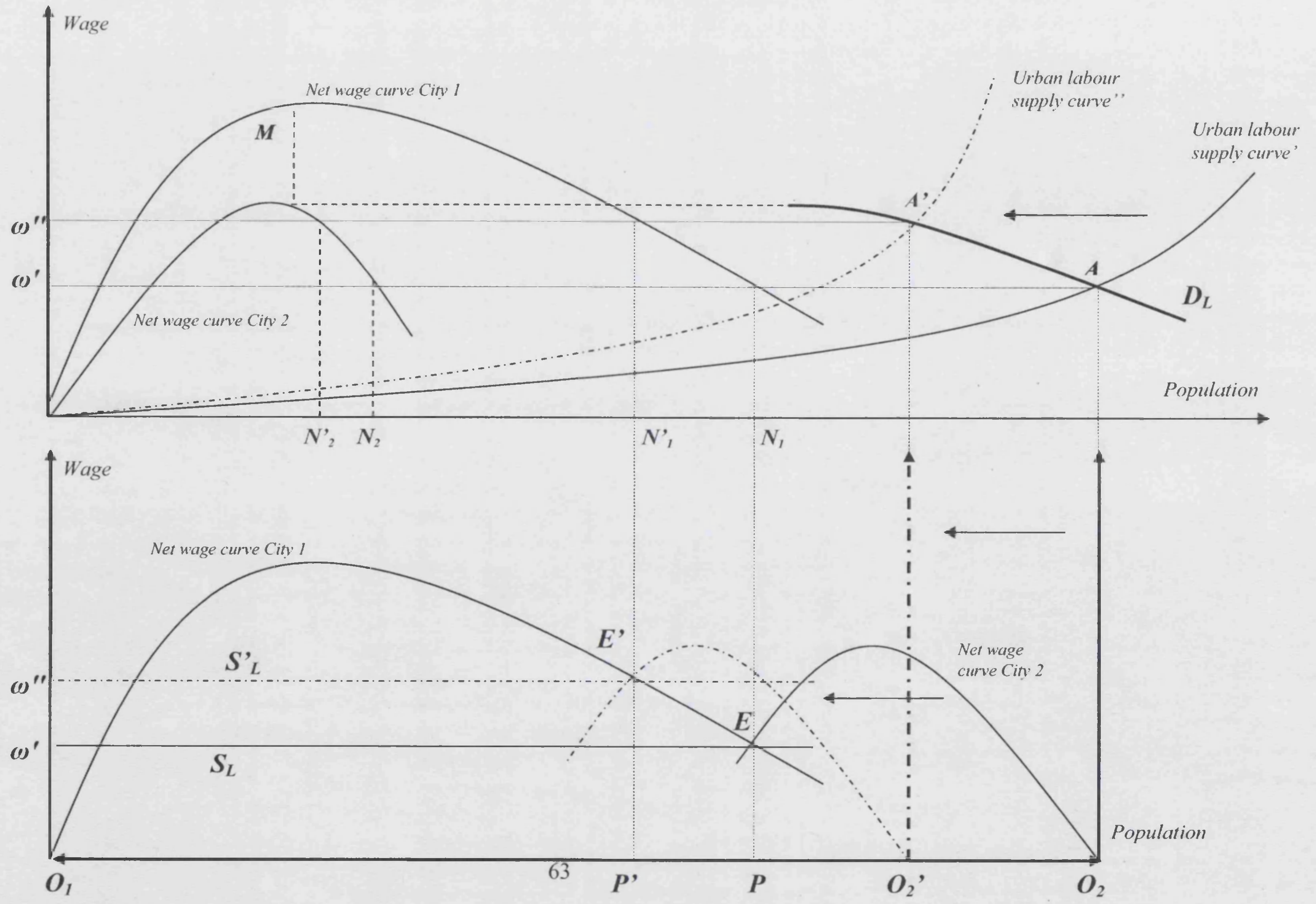
Taking the inverse function of $\omega_i^U(n)$ from expression (2.2) as $n_i = g_i(\omega)$ (and taking away the superscript U to save clutter), expression (2.4) may be re-written as: $\frac{g_1(\omega')}{g_1(\omega'')} - \frac{g_2(\omega')}{g_2(\omega'')} > 0$, or equivalently $g_1(\omega')g_2(\omega'') > g_1(\omega'')g_2(\omega')$ (as $g_i(\omega'') > 0 \forall i$). If we consider this condition at the margin using the Taylor formula (given that $g_i(\omega + \Delta\omega) = g_i(\omega) + \frac{\partial g_i(\omega)}{\partial \omega} d\omega$ for i=1,2), (2.4) could be re-written as:

$$\frac{\partial g_1(\omega')}{\partial \omega} < \frac{N_1}{N_2}\frac{\partial g_2(\omega')}{\partial \omega} \qquad (2.5)$$

Let us analyse this condition in more detail. As $\frac{\partial g_i(\omega')}{\partial \omega} < 0$ (we are moving along the decreasing part of the net wage curves), the absolute value of the first term must be greater than that of the second term for (2.5) to be satisfied. This means that the inverse of the net wage curve for city 1 (expressing n_1 as a function of ω) needs to be steeper than the inverse of the net wage curve for city 2. Or equivalently, the net wage curve for city 1 needs to be flatter than that for city 2. This is intuitive by examining Figure 2.1: a higher value of ω^U should determine *a more than proportionate* reduction in city 1 equilibrium population relative to city 2 population for condition (2.4) to be verified. From expression (2.5) it is apparent that this more than proportionate decrease in population (i.e. the ratio between the slope of $g_1(\omega)$ and $g_2(\omega)$ evaluated at ω') should be at least as large as $N_1/N_2>1$.

This is the case represented in Figure 2.1, where a steeper labour supply curve is indeed associated with a reduction in urban concentration. Therefore the predictions of this framework with respect to the effects of a shift in the elasticity of labour supply on urban primacy ultimately rest upon the intensity with which congestion costs (summarised by $c(n_i)$) operate realtively to agglomeration forces (summarised by $w(n_i)$) in each of the two cities.

Figure 2.1: Rural sector and urbanisation with perfect mobility of labour



On the other hand the predictions with respect to the degree of urbanisation are unambiguous: the higher the elasticity of labour supply, the lower the urbanisation rate. The intuition underlying this result lies in the role of the agricultural sector as a provider of labour during the industrialisation process. If the supply of labour from the rural areas is very elastic, the demand for labour by the urban sector can be satisfied by a large pool of rural workers willing to migrate to the urban areas even in the presence of a small wage gap. For any given urban net wage curve, this allows the urban sector to expand relative to a situation with an inelastic labour supply curve.

Finally, let us briefly examine how the framework would change if the assumption of an exogenous net rural wage were relaxed. A standard way used to model this is to make the marginal product of labour in the rural sector vary with the number of rural dwellers. In particular given a fixed supply of land, the marginal productivity of labour is diminishing with the size of the rural labour force. In terms of the framework sketched above, denoting total rural population with L, the net rural wage would be expressed as: $\omega^A = \omega^A(L-n)$ with $\partial\omega^A / \partial n > 0$.[39] In this case expression (2.3') could be re-written as: $N^S = \dfrac{\omega^U(n)}{\omega^A(L-n)}$, or expressing it in terms of its inverse function:

$$\omega^U = s(n) \times \omega^A(L-n) \qquad (2.6)$$

where $s(n)$ is the expression for ω^U as a function of the supply of labour to the urban sector. In the light of this framework let us re-consider the situation depicted in Figure 2.1 with equilibrium at point A'. At this point the urban population is lower than at point A. As ω^A is a positive function of n, ceteris paribus its value would be lower due to the feedback effect via a higher rural population. This means that the slope of the labour supply curve would be flatter than that of its corresponding curve in the case ω^A were exogenous (i.e. *urban labour supply curve''* in Figure 2.1). The extent of this adjustment depends on the functional form of $\omega^A(L\text{-}n)$.[40]

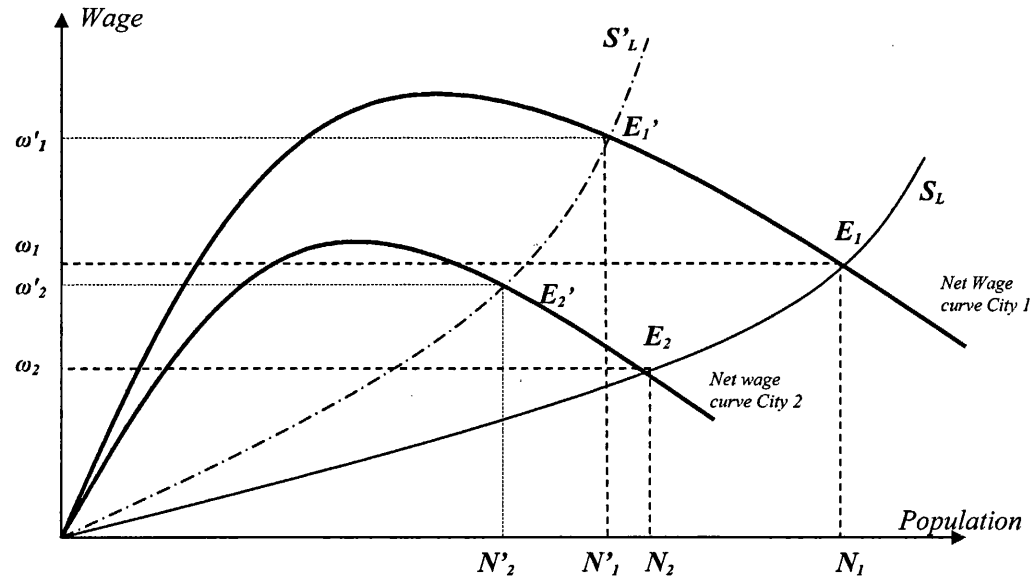
[39] Note that the assumption of an exogenously determined rural net wage may be reconciled with this function allowing ω^A to be fixed for $n<n^*$.

[40] A large enough adjustment could make the labour supply curve even flatter than the *urban labour supply curve'* in Figure 2.1, reversing the functioning of the model depicted in the Figure.

2.3.5 Equilibrium with imperfect labour mobility between cities

How does the story change if we assume some imperfections to labour mobility between cities as well? As in the case of rural-urban migration this would be the case if there are substantial costs of migration as a function of distance between cities (e.g. due to transport costs, language and cultural barriers). As a consequence workers would need a large wage gap in order to relocate from one city to the other in the same way as when they move from the rural to the urban sector. In terms of the model, this assumption implies that the labour supply curve to each city $S_L=s(n)$ (defined in the $[n,\omega]$ space for any given value of w_A) would be upward sloping rather than horizontal as in Figure 2.1 (and as in most urban economics literature). This upward sloping S_L is represented in Figure 2.2 along with the two net wage curves. I assume the standard functional form such that $\partial s/\partial n > 0$ and $\partial^2 s/\partial n^2 > 0$. This time the equilibrium city size and net wage are obtained where S_L intersects each of the wage curves: point E_1 for city 1 and point E_2 for city 2.

Figure 2.2: Rural sector and urbanisation with imperfect labour mobility



As in the previous framework, the location specific advantage makes city 1 larger than city 2. Because labour is not perfectly mobile, city 2 still manages to attract part of the rural population (e.g. rural workers who live near city 2 and who are not able to bear the costs of moving to city 1), despite the higher net wage in city 1 (in equilibrium $\omega_1 > \omega_2$). The equilibrium city sizes are obtained by solving the following expression for n: $s(n_i) = w_i(n_i) - c_i(n_i)$ for $i=1,2$. With a steeper labour supply curve S_L' (due to a higher ω^A) the equilibrium city size of city 1 would be $N_1' < N_1$ (point E_1' in Figure 2.2). Similarly for city 2 the equilibrium populationwould be $N_2' < N_2$ (point E_2'). Again in equilibrium $\omega_1' > \omega_2'$ (with $\omega_1' > \omega_1$ *and* $\omega_2' > \omega_2$). As in Figure 2.1, total urban population is lower for a steeper labour supply curve. But in this case the resulting change in urban primacy depends directly on the slope of the S_L (i.e. by the changes in the rural sector), other than on the shape of the net wage curves.

2.4 Conclusions

The basic hypotheses arising from the two frameworks presented above are summarised in Table 2.1.

Table 2.1: Predictions from the models

	Urban primacy	Urbanisation
Model with perfect labour mobility	Depends on net wage curves	Positive function of elasticity of labour supply
Model with imperfect lab. mobility	Depends on net wage curve and on labour supply curve	Positive function of elasticity of labour supply

Therefore the predictions are unambiguous in relation to the rate of urbanisation, while they depend on the shape of the net wage curves (and on the labour supply curve in the case of the second framework) in relation to urban primacy. The empirical analysis in the next chapter will test the former hypothesis and shed light on the sign of the latter. In particular, it will investigate the effects of changes in the agricultural sector (proxy for changes in the labour supply curve) on urbanisation outcome (and urban primacy in particular), controlling for changes in the net wage curves.

The mechanisms presented in the frameworks above are not far in spirit from those arising from the NEG model developed by Puga (1998), where the elasticity of rural to urban labour supply increases the range of transport costs for which an equilibrium with an urban primate arises. This implies that other things being equal the higher the elasticity of labour supply the more a higher urban primacy equilibrium is likely to emerge. There are however some differences between the two approaches, which may have implications for the empirical analysis. Differently from the present paper, Puga (1998) develops a NEG general equilibrium model, with constant urban population, which distributes either only in one city or evenly across two cities. This very nature of the model (where cities have the same production functions) generates these 'extreme' outcomes, which are obviously absent in the real world. Importantly unlike Puga's model this approach does not require the use of data on transport costs, which are often unavailable, in the empirical implementation. Finally the elasticity of labour supply in Puga's model is obtained as a ratio between other variables (i.e. share of manufactures in expenditure and labour share in the agricultural production function), which have a direct (independent) effect on the range of transport costs for which the primate equilibria arise. This may make the identification of the impact of the elasticity of labour supply more problematic at the empirical level.

Chapter 3. Agricultural productivity and urban primacy: Empirical evidence from Indian states

This chapter develops the empirical analysis based on the framework presented in chapter 2. As mentioned above the rate of urbanisation is the only dependent variable on which chapter 2 has an unambiguous prediction. However I focus mainly on the effects of agricultural variables on urban concentration (and urban primacy in particular) rather than on the urbanisation rate. This is for two main reasons. First, there exists more knowledge on the determinants of urbanisation than on the determinants of urban primacy. In particular the degree of urbanisation is intrinsically and positively related to the process of economic development. Second, as argued in chapter 2 urban primacy is a major urban policy issue in developing countries unlike the degree of urbanisation per se, which is perceived to be structurally linked with the process of economic development. Therefore policy-makers are relatively less prone to influence it than in the case of the growth of an urban primate.

The main dependent variable in the analysis is urban primacy, i.e. the share of the largest city in total urban population; but other measures of urban concentration are used as well that are based on the upper tail of the city-size distribution. In addition, I also apply the same empirical analysis to other urban concentration measures derived from a larger set of urban areas than only the upper tail. This allows one to test whether the findings about the impact of the agricultural sector on urban concentration extend beyond the upper tail of the city size distribution. The rest of the chapter is organised as follows: the next section describes the empirical methodology and the data used to test the hypotheses; section two presents the results; and section three concludes.

3.1. Methodological framework

In order to bring the theoretical framework of Chapter 2 to the data, I employ a methodology which departs from the traditional empirical analyses on the determinants of urban concentration, such as Davis and Henderson (2003), Moomaw and Shatter (1996) and Ades and Glaeser (1995). I use sub-national units (i.e. states within a federal country) rather than countries as the unit of analysis, studying the variation of urbanisation and urban concentration measures across units and over time.

3.1.1 The Indian context

I apply the empirical test to the case of India, as it has a number of features that make it particularly amenable to this type of empirical analysis. First, it is a federal country composed of several states with a fairly high degree of political autonomy, which allows for some state-wise variability in policy variables.

Second, the size of the major states is similar in terms of both population and geographical extension to that of medium-large countries. The average population of the 16 major states considered for the analysis in 2001 was 61,921,484 (Government of India, 2001).[41] If it were a country, it would rank number 20 (between Thailand and United Kingdom) out of 236 (CIA, 2003). Even the least populous state, Jammu & Kashmir with 10,069,917, would rank above the median country (number 70). The average size of the 16 states is 189,573 Km^2 (Government of India, 1991), which would rank number 88 among the largest countries in the world between Senegal and Syria (CIA, 2003). The smallest state is Kerala that with 38,863 Km^2 would rank number 137 (slightly below the median).

The vast size and population of Indian states along with their differences in terms of languages, culture and social norms appear to have limited the mobility of labour across states. Cashin and Sahay (1995) find that the response of migration to income differentials across states was similar to the weak responsiveness of population movements to income differentials across the countries of Europe. Similarly, Topalova (2005) finds extremely limited labour mobility across Indian regions between 1983 and 2000. The World Bank (2009) confirms that the share of inter-state migrants in total migrants in India in the nineties has been small even relatively to other Asian countries such as Vietnam and more recently China.[42] This relative inter-state immobility of labour is a necessary condition for the empirical test to be meaningful. If that were not the case changes in the labour supply curve in one state may be reflected on

[41] The states considered for the analysis are: Andhra Pradesh, Assam,, Bihar, Gujarat, Haryana, Jammu & Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, West Bengal. Together they represent over 97% of Indian population.

[42] Note that circular and temporary migration, which are an important part of internal migration in India, are not included in the statistics underlying these findings as these types of migrants do not change residence in the process of migration. Therefore the Census data consider them as non-migrant. This should not be a problem for this chapter's analysis as the dependent variables are constructed on the basis of Census data.

urbanisation patterns in other states as well.

Finally, Indian urbanisation experienced an important growth over the period considered with its rate increasing from 17% in 1950 to 27.7% in 2000 (UN, 2006). Over 20 million people moved from rural to urban areas only in the 1990s accounting for 30% of national urban growth (Lall et al., 2006). As a comparison, migration from rural areas accounted for about 25% of urban growth in the 1980s and 1990s in Africa.

3.1.2 The intra-country analysis

The intra-country methodology has various advantages over cross-country regressions in search of the determinants of urban concentration. First, it allows better control for country-specific unobservable factors, such as macro-institutions, culture, politics and history. Although - as argued above - these factors differ among states, such differences are likely to be less significant than those between countries, as the states are set in a single institutional context, with a common history and similar languages. These unobservable factors are likely to be relevant determinants of urbanisation patterns. Most empirical studies found national level poltical and policy variables to be major determinants of urban primacy (e.g. Ades and Glaeser, 1995 and Davis and Henderson, 2003). This hints at the potential role that other related unobservable factors, such as the quality of political institutions, or country-specific political shocks, may play in determining urbanisation patterns. For example suppose that a government increased the level of centralisation of the country's political set up without changing any formal institution. This may lead to higher returns to rent-seeking activities close to the seat of the power and thus to a higher concentration of population, which would not be picked up in cross-country analyses. Davis and Henderson (2003) try to deal with the problem that the error term is likely to be correlated with both political variables and income through IV estimation. However concerns remain as to what extent it is possible to properly deal with country-specific time variant unobservables captured in the error terms.

Second, using sub-national units avoids the problem of the artificial differences in the definition of urban area across countries, which may systematically bias cross-country urban concentration analysis. This may be the case if the urban definition criteria changed in a way that is correlated to other explanatory variables of urban concentration. As such criteria may be related to a country's total population, this may represent a source of systematic bias in cross-country regressions. Satterthwaite (2006) finds that between 20 and 40 percent of the population in many nations lives in settlements that could be considered to be either rural or urban according to the definition employed.[43]

Finally, by controlling for national level variables (e.g. types of political system, degree of regional autonomy, trade policy, level of political stability), which in cross-country regressions take up most explanatory power of urban primacy, this methodology allows one to focus on agricultural variables as well. It is important to isolate the effect of these variables, as they may have a direct policy relevance.

The main drawback of the intra-country methodology relatively to the cross-country one is the relatively small number of cross-sectional observations (i.e. 16 states). Unlike the analysis in chapter 1, the analysis in this chapter does not have many observations in the time series dimension, although it spans a long period of time (1961-2001). That is due to the fact that population data are available only every ten years, thus the baseline regressions are run on 5 time periods only. This creates a further challenge for the precision of the estimates, increasing the difficulty of obtaining significant coefficients relative to those Indian states' analyses based on a large number of time series observations (e.g. Datt and Ravallion, 1996, Besley and Burgess, 2000, 2002 and 2004, Rud, 2009). Given these challenges, the significance of the coefficients in the following analyses should signal an even more robust relationship than that found by analyses with a large time series dimension.

3.1.3 Specifications

The basic empirical strategy is to test for the effects of agriculture and socio-demographic variables on measures of urban concentration and on the urbanisation rate across a panel of states over five decades in the Post-Independence period.[44] I employ a reduced form which incorporates the effects of changes in the agricultural sector on the

[43] For instance India is a predominantly rural nation according to its Census. But most of India's rural population lives in villages between 500 and 5,000, which would be classified as "urban" by some national urban definitions, thereby turning India into a predominantly urban nation (Satterthwaite, 2005).
[44] The analysis does not address issues of intra-distribution dynamics (as for example in Overman and Ioannides, 2001), as the urbanisation measures that I use are aggregated ones and abstract from the pattern of individual cities.

labour supply curve and the impact of the latter on urban concentration. I use two basic specifications to test for the effects of rural variables on urban concentration. First I run a pooled OLS regression of the type:

$$y_{st} = \beta_0 + \beta_1 rur_{st-i} + \beta_2 X_{st} + \beta_3 t + \varepsilon_{st} \qquad (3.1)$$

Second, I also employ a fixed effect (FE) model as follows:

$$y_{st} = \beta_0 + \beta_1 rur_{st-i} + \beta_2 X_{st} + \beta_3 t + \alpha_s + \varepsilon_{st} \qquad (3.2)$$

where y_{st} is some measure of urban concentration in state s at the period t, rur_{st-i} is the rural variable of interest lagged *i* years, *X* is a set of control variables, α_s is state fixed effects and *t* is year effects. The choice of the lag should reflect the time that changes in a variable takes to exert their influence on the dependent variable. I experiment with different lags using the average value of the agricultural variable between *t-i* and *t-1*. The use of the average value of a variable is also preferred as it limits the extent to which the results are subject to yearly fluctuations. As shown below the results obtained are robust to the choice of the lag. These empirical specifications are similar to those used by the empirical literature described in chapter 2, with the main difference being the addition of the rural variables on the right hand side. I also test for the effects of the rural variables on urbanisation by using the same specifications with the share of urban population in total population as the dependent variable and a similar set of controls as in Davis and Henderson (2003).

3.1.4 Dependent variables

The main dependent variables are for each state and time period the urban primacy ratio and the share of urban population in total. The former is the share of national urban population residing in the most populous urban area (*PRIMACY*). Considering an urban system of *N* cities, this is defined as:

$$UP = \frac{P_1}{\sum_{i=1}^{N} P_i} \qquad \text{with } 1/N < UP < 1$$

where P_i is the population of the ith largest urban area.

For robustness purposes I also test some of the results using the following other three indices of urban concentration (again higher values of the indices represent more concentrated urban systems):

1) the ratio of the largest urban area to the sum of the next three cities,

RATIO4: $R_4 = \frac{P}{\sum_{i=2}^{4} P_i}$ with $1/3<R_4<\infty$.

2) the Herfindhal index of urban concentration (*HERF*), which is calculated as[45]:

$$H = \sum_{i=1}^{N} (P_i / P_{tot})^2 \qquad \text{with } 1/N<H<1$$

3) the Herfindhal index of urban concentration conditioning for the total population of the urban areas considered (*HERF2*), defined as:

$$H_2 = \sum_{i=1}^{n} \left(\frac{P_i}{\sum_{i=1}^{N} P_i}\right)^2 \qquad \text{with } n/N^2<H_2<1$$

where n is the number of urban areas considered.

Although all indices are measures of urban concentration and tend to be highly correlated (due to the important influence exerted by the largest urban area), it is worth discussing some differences between them. *PRIMACY* and *RATIO4* take into account only the very upper tail of the urban areas' distribution and measure urban primacy: the first one measures the extent to which the urban population of a state is concentrated in the largest urban area; the second measures the extent to which the urban population is uniformly distributed in the upper tail of the distribution (i.e. whether other large cities have developed besides the largest one). The Herfindhal index usually takes into account the entire distribution of the variable on which is based. I have consistent series of data for the 20 largest urban areas for each state in every period (i.e. n=20), which cover at least 70% of the entire urban population. As it is evident from the formulae, $H<H_2$, *as* H_2 is adjusted for the fact that the largest ten urban areas account for less than the entire population of a state. However it is likely that even the Herfindhal indices

[45] Some authors (Wheaton and Shishido 1981 for example) consider the inverse of this index as a measure of "urban decentralisation".

may be disproportionately influenced by the largest urban area. This does not allow to properly identify the impact of the agricultural variables over the entire city-size distribution. We discuss below ways to address this issue.

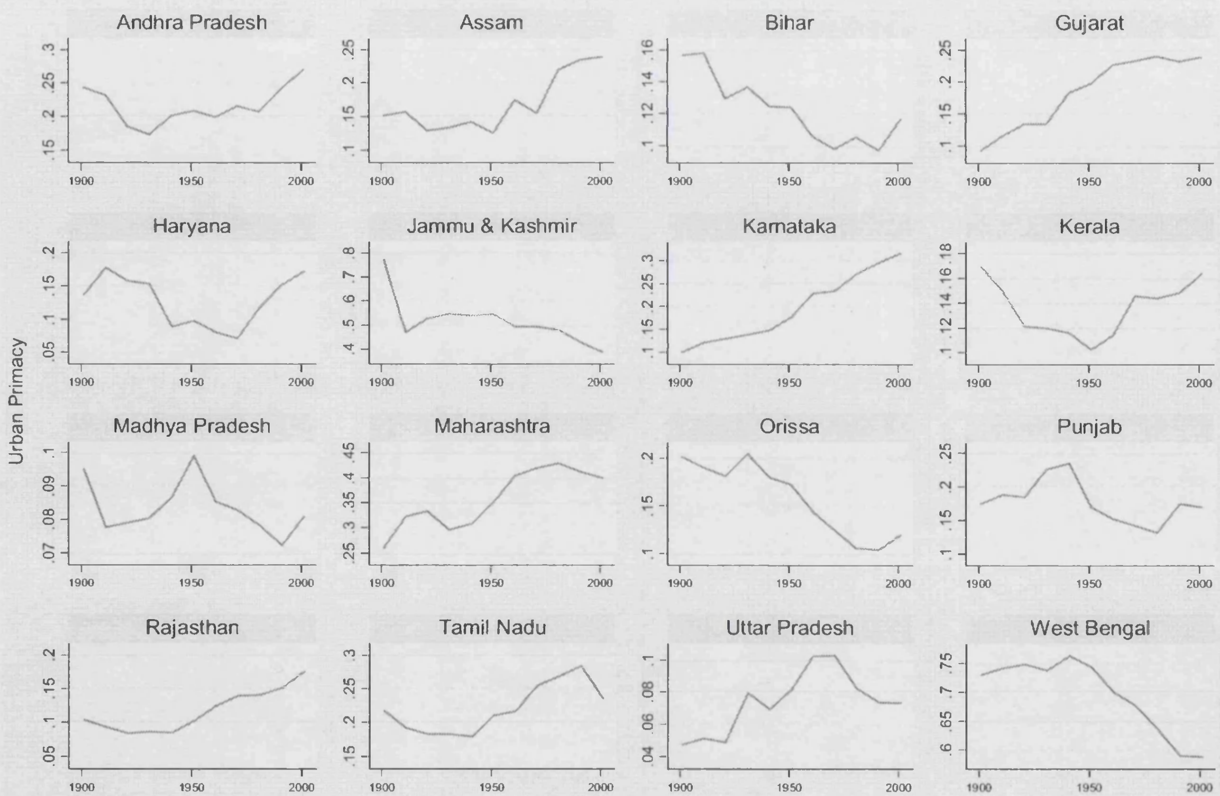
Figure 3.1 shows the evolution of urban primacy for the main 16 India states over the period 1901-2001.[46] There is a large variability of trends and levels across states. For example states like Assam, Gujarat, Karnataka, Maharashtra and Rajasthan have increased their urban primacy during most of last century, while states like Orissa and Bihar have had opposite trends. This variability should make the investigation of the determinants of urban primacy amenable to longitudinal analysis.

3.1.5 Explanatory variables

The explanatory variable of interest in the model is the wage elasticity of the rural to urban labour supply curve (i.e. the slope of the S_L curve in chapter 2). As there is no direct measure of such elasticity I consider indirect measures on the basis of the channels through which the conditions of the rural sector may shape the slope of the S_L curve.[47] In this way I can derive an empirical specification through which the impact of agricultural and socio-demographic variables can be meaningfully tested. Because of the indirect nature of this empirical test, I try to use different types of variables to make the interpretation of the results as robust as possible. There are a number of variables that could be used to proxy the elasticity of labour supply to the urban sector, although I will argue that land productivity is the most appropriate one. For the sake of simplicity, in the following discussion I present the hypothesis of the possible effect of the agricultural variables on urban primacy assuming that a flatter S_L curve determines higher urban primacy.

[46] Although I use only urban concentration measures for the period 1951-2001, I find it useful to plot the evolution of urban primacy also for the pre-Independence period as it provides a more complete overview of long term trends.
[47] Chapter 2 already discussed one such channel, i.e. net agricultural wage.

Figure 3: The evolution of urban primacy across main Indian states in the 20[th] century



Source: Author's elaboration on Indian Census (various years)

The main variable I use is land productivity, measured as yield (in '000 real Rupees) per hectare of sown land (*YIELD*).[48] This is a measure intrinsically related to agricultural income but it has more observations available than wage for Indian states. Moreover as argued below it is potentially less subject to the endogeneity problem than agricultural wage. I construct averages for agricultural yield over different number of years. I use the ten years average (i.e. $Avg(10)Yield_t = (\sum_{i=0}^{9} Yield_{t+i})/10$) as the main regressor as this allows to incorporate all the information of the explanatory variable up to the period before the year which the dependent variable refers to (as the urban variables are available every ten years). As shown below the results are unaffected by the number of years over which the variable is averaged.

To the extent that *YIELD* is positively related to agricultural income, its reduction is likely to be associated to an increase in the number of rural workers willing to move to the urban area for any given level of urban wage. Hence, the lower the agricultural productivity, the flatter the labour supply schedule should be. On the other hand, in the presence of fixed costs of rural-urban migration, a lower rural income may allow fewer potential migrants to meet these fixed costs. Therefore according to this channel a decrease in *YIELD* would be associated with a steeper labour supply curve. However available evidence from developing countries suggests that the latter channel appears to be second-order relative to the other one, i.e. the decrease in the urban-rural wage gap. For example Barrios et al. (2006) find that the decrease in rainfall in Sub-Saharan Africa (SSA) between 1960 and 1990 determined a considerable increase in rural-urban migration. As the decline in rainfall diminished agricultural productivity in SSA, the results support the idea of a negative relationship between *YIELD* and rural-urban migration. Da Mata et al. (2007) find that a 1% increase in surrounding rural income opportunities decreases a city population by 6.9%, concluding that city populations are very sensitive to rural earning opportunities. Hence a rise in *YIELD* is expected to be associated with a steeper labour supply curve. There is however a potential problem of endogeneity with this variable. If a low level of *YIELD* is associated with a high rate of rural-urban migration, the latter may reduce the rural population density. In the absence of unused resources in the rural sector, this may

[48] Agricultural yield in real Rupees is computed by dividing the net state agricultural product by the agricultural price index and then divide this product by the net sown area.

affect the agricultural productivity of land. Given the modest level of rural-urban migration relative to the rural population in India, such a migration is unlikely to have had a significant impact on outcomes in rural areas, though it may have some impact on urban areas (Topalova, 2005).[49] Notwithstanding this, I also include rural population as a control. In addition I test for the effects of *YIELD* on urban primacy through IV estimation as well, as explained below.

As a robustness check, I use another agricultural variable in place of *YIELD*: agricultural wage in constant 1980 Rupees (*WAGE*). Again, there are different ways in which this variable may influence the elasticity of labour supply. Similarly to *YIELD,* a lower level of *WAGE* could determine a flatter labour supply curve, as there would be more individuals willing to move for any given level of urban wage. Moreover, to the extent that the average wage in the rural area is an inverse function of its supply of labour, a higher density of rural population per arable unit of land is likely to be associated with a lower rural wage. This would imply a higher responsiveness of migration to any given urban-rural wage gap. The opposite argument related to the fixed costs of migration discussed above should apply to *WAGE* as well. Similarly to the case of *YIELD* this channel should be second order determining a negative relationship between *WAGE* and the elasticity of labour supply. This variable is likely to suffer from a more serious endogeneity problem than *YIELD*, as rural wages should respond to variations in the urban sector in terms of both prices and demand for labour and goods. This is confirmed by the evidence on India provided in chapter 4. I add the share of urban population as a control to deal with this problem and again test the effects of *WAGE* through IV estimation.

The last set of variables used to proxy for the elasticity of labour supply is the one related to the demographic characteristics of the rural population. It is quite uncontroversial that migrants tend to be concentrated among the young adults, and rural-urban migration in India confirms this pattern.[50] Hence, other things being equal a higher proportion of young adults are likely to be related to a flatter S_L curve. This is

[49] The author calculates that the median urban sector of a district has one fifth of the population of the median rural sector of a district. Given that 7.6% of the median urban district population were rural migrants during the 1990s, this means only 1.6% of the median rural district population migrating to the city.

[50] See for instance Joshi and Lobo (2003) for an analysis of all India's rural-urban migration, and Oberai et al. (1989) for case studies of the states of Bihar, Kerala and Uttar Pradesh.

measured by the share of the rural population in the 15-34 age range (*SHARE 15-34*). Analyses of migration in India (e.g. Joshi and Lobo, 2003) also highlight that men are the main economic migrants (i.e. moving to cities in search of a job). I use the female-male ratio in the young population (15-34) – *FEM-MALE 15-34* – to test the impact of gender balance on urbanisation pattern. Note that such variables may be endogenous to urban concentration patterns as rural areas with an elastic labour supply may experience an intense out-migration flow of young adults, which may deplete their stock in the following period. Again such a concern may be not so problematic in India, given the relatively low shares of out-migration from the rural areas.

I also include a number of controls, which comprise most of the standard explanatory variables of urban concentration. First, although the intra-country analysis control for (fixed and time varying) national level political factors, state-level politics may still play a role in determining the degree of primacy. A capital city dummy (i.e. value of 1 in any year when the largest urban area is the capital city) should capture the effect of the political advantage on the degree of primacy. I can include this variable even in the fixed effect model as it has some variation over time due to a number of events in Indian modern history. First, some Indian states have changed their capital throughout the period considered (such as Assam, whose capital changed from Shillong to Guwahati in 1973). In addition, the largest urban area in some states has changed over time, as in Kerala where the capital city Trivandrum has been the largest city only until 1954 (then it was overtaken by Cochi). Finally, some states have been created during the period considered, thus generating a new capital, as in the case of Gujarat (created in 1960). The controls include also real income per capita to capture any structural relationship between this variable and primacy, rural population and the share of urban population (plus its squared term due to non-linearities) to partly control for possible reverse causation (i.e. from the urban sector to rural variables).

Both *YIELD* and *WAGE*, as average measures, do not convey any information on the distribution of agricultural incomes across the population of the state. For example non wage earners (e.g. unemployed, landowners) are not included in the calculation. Moreover the concentration of land ownership may affect the way in which changes in *YIELD* may translate into income changes to the rural workforce. In order to control for these factors I include in some of the regressions a measure of land ownership concentration, i.e. the Gini coefficient of land concentration (*GINILAND*).[51]

3.1.6 Instrumental variables

I use two sets of instruments to tackle the endogeneity of the agricultural variables discussed above: rainfall data and land reform legislation. Rainfall is a key input in the agricultural production function. There is substantial evidence that drops in rainfall tend to damage agricultural productivity, especially in developing countries (see the Intergovernmental Panel on Climate Change, 2001 for a review). In line with a growing literature (e.g. Miguel et al., 2004, Brückner and Ciccone, 2008), I exploit rainfall variation as an exogenous shock to the agricultural sector. In particular, good precipitations should boost agricultural productivity and income, and via that reducing the propensity of the rural population to migrate to urban areas.[52]

Land reform is likely to represent an exogenous shock to the agricultural sector, which has an impact on agricultural variables (wage and productivity), but as I argue below it is not related to a state's urban concentration. The reform has been described in chapter 1 and the intensity of the reform has been captured by Besley and Burgess (2000) who coded each state's acts into four types of legislations, and added them to obtain cumulative land reform variables. Using these cumulative variables, the authors show that land reform had a significant impact in reducing poverty (and increasing agricultural wage) across states over time.

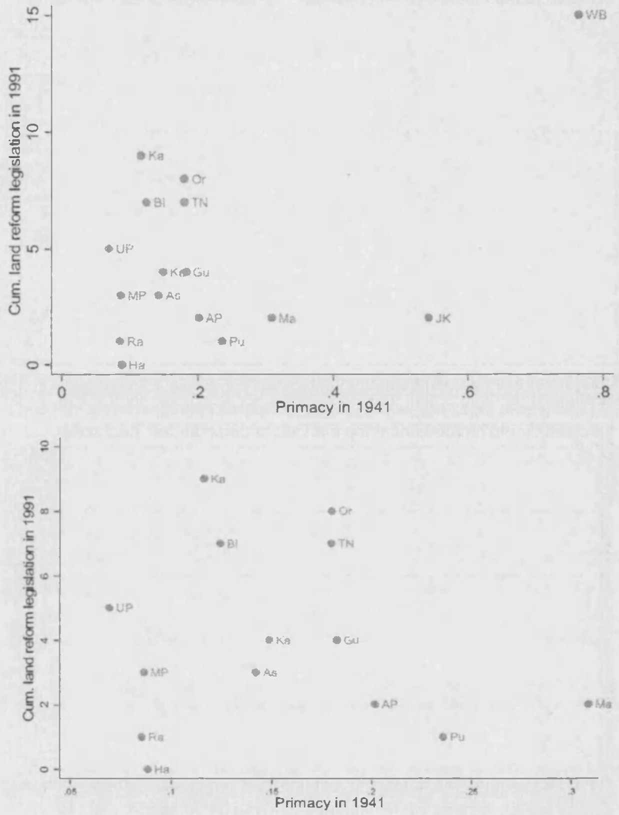
As the authors note, the land reform was mainly designed to reduce rural poverty. Theories of urban bias suggest that as the elites in developing countries tend to be concentrated in the urban areas, this leads to a policy discrimination against rural areas (Lipton, 1976). Therefore the urbanisation rate may influence land reform legislation. Including the share of urban population in the controls should address this concern. But that does not address the potential reverse causality problem, i.e. the degree of urban primacy as a potential determinant of land reform implementation. A plausible story undermining the exogeneity of the land reform as an instrument could be

[51] As this variable slightly reduces the number of observations, I only include it in a few specifications.
[52] A similar channel has been tested for a large panel of developing countries by Barrios et al. (2007).

the following. As a high degree of primacy is usually associated with both political and economic power concentrated in the primate city, a high primacy could be associated with a poor implementation of land reform legislation to ensure the supply of a large pool of cheap labour from the country-side to satisfy the urban primate's labour demand. However this is not the case when looking at Figure 3.2 which plots the degree of urban primacy in 1941 (well before the implementation of the land reform) against the cumulative land reform legislation 50 years after (upper quadrant). West Bengal has the highest level of urban primacy in 1941 and the highest intensity in land reform implementation at the end of the period, opposite to the predictions according to the story above. This positive association is most likely to be spurious. The high primacy (due to the historical role of Calcutta as the capital of the British Indian empire, see below) was not the driver of the aggressive land reform legislation, which was a feature of West Bengal's hard left governments in the Post-Independence period. Dropping the two influential observations (West Bengal and Jammu & Kashmir) makes the lack of any relationship between the two variables clearer (lower quadrant).

I employ the IV estimation also for the urbanisation analysis to deal with the engogeneity of the agricultural variables, which is even clearer in this case. Although land productivity (as measured by *YIELD*) is less subject to this problem than agricultural wage, it is still an important concern to deal with. For instance urbanisation is associated with an increase in urban-rural remittances, as argued in chapter 4. To the extent that these are used to invest in agricultural equipment, the increased capital-land ratio would raise land productivity. On the other hand, Rozelle, Taylor, and deBrauw (1999) found that there are incentives for household members to curtail their own labour efforts and use supplemental income from remittances to maintain their standards of living. This may be a consequence of migrant workers' typical inability to monitor the use of their transfers. Also, urbanisation is usually associated with an increase in the demand for agricultural products especially of higher value (see Parthasarathy Rao et al., 2004 on India). This may provide the incentive to using agricultural land more intensively (see Chapter 4). Likewise it is immediate to see the potential endogeneity of agricultural wage to urbanisation as well. Other than the channels described above for *YIELD*, another important one undermining the exogeneity in the case of *WAGE* is the following. To the extent that urbanisation reduces the rural labour supply, this may increase agricultural labour productivity (but not necessarily land productivity), given the fixed land supply and diminishing marginal returns to land. This may pose upward pressure on rural wages.

Figure 3.2: Primacy and land reform legislation



Source:Author's elaboration on Indian Census and Besley and Burgess (2000)

3.1.7 Data

I construct the urban concentration measures described above for the 16 main states from the Indian Censuses (1951-2001). These have a ten year frequency. The boundaries of the Indian states are kept as in 1991 as the states have changed boundaries over the period considered. I also complement the Census data with UN (2006), which compute the population figures for over 3,000 urban agglomerations and for total urban population in all countries in 5 year intervals (i.e. 1955, 1960, etc.). The computations are inter-censual years interpolations and extrapolations based on the difference between the growth rate of each city and the growth rate of the population of the rest of the country and on the difference between the urban and the rural growth rates (see

Appendix 3.1 for a more detailed explanation). These data have already been used to examine the determinants of urbanisation across countries (Davis and Henderson, 2003). Although the main results are based on the restricted sample from the Census, it is reassuring that the findings are consistent with those obtained using the larger sample, which includes data from UN (2006) as well.

The agricultural variables (including the land reform variables) and most of the other control variables (socio-economic, manufacturing, rainfall) come from Ozler, Datt and Ravallion (1996) and have been updated by the Economic Organisation and Public Policy Programme (EOPP) at the London School of Economics.[53] I have updated the agricultural yield, the rainfall data and the land reform data up to the early 2000s.[54] The demographic data (age and gender) are taken from the Censuses with some state level aggregations. I also use a capital city dummy (i.e. the value of 1 if the largest urban area is the capital of the state) constructed from a variety of sources including the Indian Census and Playne et al. (2006). Table 3.1 shows the descriptive statistics for the main variables in the analysis for the Census years (1961-2001).

Table 3.1 Descriptive statistics for the main variables (Census years, 1961-2001)

	Obs	Mean	Std. Dev.	Min	Max
Primacy	80	0.228	0.152	0.070	0.701
Ratio4	80	1.708	2.348	0.367	13.557
Herf	80	0.251	0.153	0.104	0.771
Herf2	80	0.092	0.104	0.015	0.492
urb. share	80	0.225	0.085	0.063	0.439
Avg(10)YIELD	64	0.268	0.153	0.062	0.834
Avg(3)YIELD	78	0.275	0.174	0.056	1.013
YIELD	79	0.267	0.180	0.054	1.016
Avg(10)WAGE (log)	56	1.607	0.644	0.394	3.102
WAGE (log)	55	1.914	0.731	0.871	3.331

Source: Indian Statistical Census and Datt et al. (1996) updated by EOPP

[53] Available at: http://sticerd.lse.ac.uk/eopp/research/indian.asp

[54] Agricultural yield is computed by dividing the real net state agricultural product (for which data are available from the Reserve Bank of India) by the net sown area (for which data are available from a variety of sources and are obtained through Indiastats). State-wise rainfall data come from the Compendium of Environment Statistics, 2002, published by the Central Statistical Organisation. The cumulative land reform data through the World Bank (2006) report on India "Land policies for growth and poverty reduction".

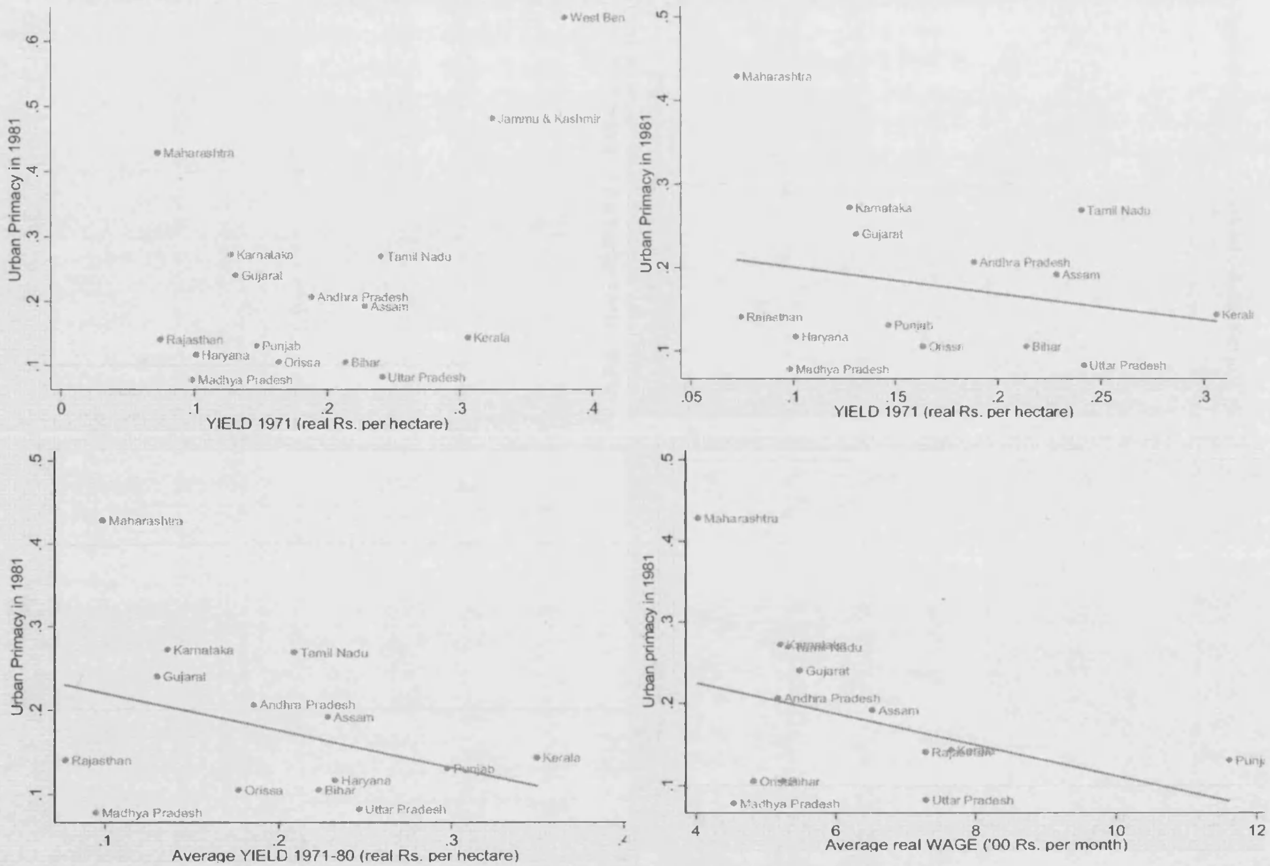
3.2. Results

3.2.1 Cross-state relationship

I first check the relationship between urban concentration and the main explanatory variables across states only. The upper-left quadrant of Figure 3.3 shows a positive relationship between the level of urban primacy in 1981 and *YIELD* in 1971. However, this relationship appears to be driven by two outlier states (Jammu and Kashmir and West Bengal). There may be good reasons why these state don't obey the relationship observed for the other states. Jammu & Kashmir has a very mountainous territory with the lowest level of fertile land in all states - only 4.7% of total land is agricultural land (Government of India, 2001). This is likely to make the distribution of land an unimportant variable in determining urbanisation patterns in Jammu & Kashmir. The state has had the most equal distribution of land in the sample during the post-independence period with a relatively high level of urban concentration. West Bengal's urban system is dominated by Calcutta, which used to be the capital of British India during the entire period of British colonisation. Therefore its growth may have been driven mainly by its political attractiveness (for rent-seeking activities) rather than by any other factors. As a matter of fact West Bengal has the highest urban primacy of all states throughout the period considered, despite relatively high land productivity.

Taking the two outliers out a negative relation (significant at the conventional 5% level) between *PRIMACY* and *YIELD* emerges (upper right panel of Figure 3.3). This means that a higher agricultural productivity of land is associated with a lower level of urban primacy, suggesting that urban primacy may be a positive function of the elasticity of labour supply. The relationship between the two variables holds also when using the average value of *YIELD* between 1971 and 1980 (lower left quadrant) and for other years as well (not shown here). The cross-section relationship between *PRIMACY* and *WAGE* ten years before is also negative (lower-right quadrant). This suggests that a higher elasticity of labour supply may be associated with more concentrated urban systems.

Figure 3.3: Urban primacy vs. agricultural income variables



Note: Primacy is computed as the share of state's urban population concentrated in the largest urban area
Source: author's elaboration on Indian Census and Ozler et al. (1996)

But more systematic evidence than these simple correlations is needed to establish any meaningful systematic connection between urban concentration and rural variables. Let us turn to it in the next section.

3.2.2 Urban concentration

I carry out the longitudinal analysis with data spanning over the period 1951-2001, testing first the contemporaneous relationship of income-related agricultural variables with urban primacy. Table 3.2 shows the results of regressing urban primacy separately on *YIELD* and on the log of *WAGE.* I choose these two different functional forms for the two variables, as the analysis suggests that these are the most appropriate in terms of the significance of the results.[55] Both variables have negative sign consistently throughout the various specifications, although they are not always significant according to the standard levels. *YIELD* has a negative but not significant relationship with *PRIMACY* in the basic OLS regression with year effects, the set of controls and dummies for West Bengal and Jammu and Kashmir (column 1). This relationship remains negative and becomes significant (at the 10% level) when adding state effects (column 2). This suggests that *YIELD* may be correlated with some state invariant characteristics, which dampen the negative relation between *YIELD* and *PRIMACY.* The coefficient of *YIELD* increases (and so does its significance) when adding the Gini coefficient of land concentration (column 3), suggesting that the distribution of land matters in driving the impact of *YIELD* on *PRIMACY.* The coefficient of *YIELD* is not affected when I use the larger sample including the UN (2006) data (column 4). On the other hand the opposite holds for WAGE: the coefficient is negative and significant only in the specification without fixed effects (column 5), while it remains negative but becomes insignificant adding state effects (columns 6-7). Similarly to YIELD, this suggests that *WAGE* may be correlated with some state invariant characteristics, which inflate the negative relation between *WAGE* and *PRIMACY.*

It is worth commenting briefly on the coefficients of the set of controls as well. Not surprisingly, being the capital city of the state increases the degree of urban primacy after ten years; the level of primacy appease to rise with that of income per

[55] The results are not qualitatively affected by changing the functional forms (e.g. log for YIELD) although their significance decreases (not shown here).

capita, which is consistent with the fact that most Indian states have not yet reached a middle-income status (when primacy should start decreasing with income growth). The relation between GDP per capita and urban primacy becomes less significant in the regression with *WAGE*, suggesting that *WAGE* is somewhat more correlated with GDP than *YIELD*. The coefficient of rural population is always insignificant, while urbanisation seems to have a mild inverted U-shaped relation with primacy in the specification with YIELD and fixed effects (column 2).[56] Adding *GINILAND* makes this relationship not significantly different from zero (columns 3 and 4). Urbanisation has a positive but not significant relationship with primacy in the FE regression using *WAGE* as the main regressor. Finally GINILAND is negatively associated with primacy, i.e. the more concentrated the land ownership the lower the degree of primacy.

Table 3.2: Contemporaneous regressions of urban primacy on agricultural variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Census	Census	Census	Census + UN	Census	Census	Census + UN
	Primacy	primacy	primacy	Primacy	primacy	primacy	primacy
YIELD	-0.036 (0.055)	-0.102* (0.062)	-0.199** (0.076)	-0.218*** (0.069)			
WAGE (log)					-0.077** (0.030)	-0.003 (0.022)	-0.011 (0.012)
$Capital_{t-10}$	0.086*** (0.018)	0.026 (0.020)	0.022** (0.008)	0.022** (0.009)	0.054*** (0.019)	0.027** (0.011)	0.024** (0.011)
GDP pc (log)	0.085* (0.043)	0.158*** (0.043)	0.083** (0.034)	0.042* (0.023)	0.121* (0.063)	0.073** (0.033)	0.040 (0.024)
Rural pop (log)	0.005 (0.016)	0.137 (0.107)	-0.042 (0.107)	-0.135 (0.103)	-0.010 (0.014)	0.005 (0.194)	-0.067 (0.155)
Urban share	0.381** (0.156)	1.195 (0.855)	-0.320 (0.684)	-0.668 (0.559)	-0.977*** (0.352)	0.423 (0.519)	0.178 (0.475)
Urban share squared		-1.382 (0.964)	1.047 (1.300)	1.037 (0.979)	3.168*** (0.860)		
GINILAND			-0.491*** (0.090)	-0.283*** (0.056)		-0.398 (0.250)	-0.219 (0.182)
Jammu & Kashmir	0.256*** (0.041)						
West Bengal	0.397*** (0.025)				0.444*** (0.022)		
State Effects	NO	YES	YES	YES	NO	YES	YES
Observations	76	76	57	147	55	55	138
R-sq. (within)	0.872	0.384	0.592	0.453	0.914	0.329	0.296
Nr. of States	16	16	15	15	15	14	14

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all regressions include year effects. Agricultural wage data for Jammu and Kashmir and Haryana are not available.*

[56] This result is comparable to that in Davis and Henderson (2003) in their IV estimation.

I next regress the measures of urban concentration on the main explanatory variables lagged a number of years. Table 3.3 presents the results of the regressions using *Avg(10)YIELD* lagged ten years as the main explanatory variable. The FE estimation better accounts for the historical urbanisation trends of the individual states than the simple OLS, hence I only employ the FE estimation here.[57] The coefficient of *Avg(10)YIELD* is negative and significant in the specification with the standard set of controls (column 1). Its size implies that a 10% increase in *Avg(10)YIELD* is associated with a 2% reduction in primacy. The coefficient increases slightly (i.e. 2.4% reduction in primacy) when adding *GINILAND* as a further control (column 2).[58] The coefficients of *Avg(10)YIELD* are slightly larger than the corresponding ones in Table 3.2., which imply a reduction between by 1.2% and 1.9% in primacy for a 10% increase in *YIELD* (cf. columns 2 and 3). This confirms that there may be some time lag for the display of the effects of *YIELD* on primacy. This negative effect is robust to the use of the larger sample (column 3), to employing the average value of *YIELD* over the preceding three-year period (column 4) and to using the simple value of *YIELD* lagged one year (column 5). The specification using the average value of *WAGE* over ten years confirms the pattern of influence of agricultural variables on *PRIMACY* (column 6). These negative coefficients are also robust to the use of other measures of urban concentration as dependent variables (*RATIO4* and *HERF2* in columns 7 and 8 respectively) in a specification with a full set of controls. As these measures take into account a wider spectrum of the city-size distribution, this finding may suggest that the effects of *YIELD* are operating on the upper tail of city size distribution more generally rather than only on urban primacy. This issue is explored further below.

The results in Table A1 in Appendix 3.2 confirm that the negative impact of the agricultural variables on urban concentration is robust to using different lags, to calculating average values of *YIELD* and *WAGE* over five and three years respectively and to using the other three measures of urban concentration.[59] The Table reports also the coefficients of the simple values of *YIELD* lagged one year, whose coefficients have lower magnitude and significance than the average values. This may suggest that year-

[57] I also test the effects of *YIELD* and *WAGE* on *PRIMACY* (as in Table 2) using random effects instead of FE estimation, largely obtaining similar results (results available upon request).

[58] This elasticity is obtained by considering a 10% increase in *Avg(10)YIELD* at the mean (i.e. increase in 0.0268) and multiplying it by the coefficient in column 1, i.e. -0.175. The result (-0.0047) is 2% of the mean of *Primacy* (0.229).

[59] The Table includes year and state effects and the full set of controls except *GINILAND*.

to-year fluctuations tend to bias the coefficients downwards. In addition, I also check the robustness of the results to the influence of any specific state. I run the same regressions in columns (2) and (6) excluding a different state at a time (i.e. each regression for *YIELD* and *WAGE* uses 15 and 13 states respectively). The results in all cases – not shown here for reasons of space – are qualitatively unaffected.

Table 3.3: The effects of agricultural variables on urban concentration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Census	Census	Census + UN	Census	Census	Census	Census	Census
	Primacy	Primacy	Primacy	Primacy	Primacy	Primacy	Ratio4	Herf2
Avg(10) $yield_{t-10}$	-0.175*	-0.217**	-0.194*				-5.800*	-0.181**
	(0.089)	(0.096)	(0.103)				(3.128)	(0.069)
Avg(3) $yield_{t-3}$				-0.179***				
				(0.058)				
$Yield_{t-1}$					-0.104*			
					(0.059)			
Avg(10) log $wage_{t-10}$						-0.066*		
						(0.037)		
$Capital_{t-10}$	0.062***	0.059***	0.035**	0.028	0.030	0.016	0.984***	0.027**
	(0.013)	(0.012)	(0.012)	(0.019)	(0.021)	(0.021)	(0.332)	(0.009)
GDP pc (log)	0.160***	0.124***	0.067**	0.147***	0.143***	0.049	1.676*	0.086***
	(0.045)	(0.040)	(0.031)	(0.041)	(0.045)	(0.032)	(0.826)	(0.026)
Rural pop (log)	0.147	0.002	-0.094	0.112	0.136	0.002	3.376	0.063
	(0.107)	(0.083)	(0.096)	(0.105)	(0.110)	(0.097)	(2.440)	(0.055)
Urban share	1.315	0.423	-0.192	0.813	1.146	0.540	33.602	0.665
	(0.760)	(0.563)	(0.598)	(0.843)	(0.900)	(1.007)	(21.079)	(0.480)
Urban share squared	-1.602*	-1.046	-0.556	-1.083	-1.341	-0.831	-28.223	-0.814
	(0.896)	(0.763)	(0.734)	(1.004)	(0.987)	(1.024)	(19.298)	(0.549)
$GINILAND_{t-10}$		0.096	0.011					
		(0.114)	(0.101)					
Obs.	64	60	165	77	78	68	64	64
R-sq. within	0.499	0.519	0.405	0.437	0.362	0.260	0.469	0.495
Nr. of states	16	15	15	16	16	14	16	16

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all regressions include year and state effects; data for Haryana and Punjab not available before 1965, as they were a single state; Avg(10) and Avg(5) indicate the number of years over which the average of a variable is calculated.*

The findings point to an important role of the agricultural sector in influencing the patterns of urban primacy (and possibly of urban concentration) across Indian states. In particular, to the extent that the assumptions on the relationship between the agricultural variables and the elasticity of labour supply are correct, the latter exerts a positive influence on the level of urban primacy.

I check further the robustness of these findings to using the other proxies of the elasticity of labour supply, which are based on demographic variables (Table 3.4). The results are consistent with the finding that a steeper rural-urban labour supply curve decreases the degree of urban primacy. In particular the main variable *SHARE15-34* lagged one year (which should increase the elasticity of labour supply) exerts a positive and significant influence on *PRIMACY* (column 1). The *FEM-MALE 15-34* variable appears to have a very mild and non significant negative effect on *PRIMACY* (column 2). This suggests that females are less likely to migrate to urban areas than males for work but this effect is possibly dampened by the female dominated migration flow for marriage reasons (which involves the movement to the husband's place of residence).[60]

Table 3.4: The effects of demographic determinants on urban primacy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Census primacy	Census primacy	Census primacy	Census + UN primacy	Census primacy	Census primacy	Census Primacy
Share 15-34$_{t-1}$	0.658**		0.775***	0.686**		0.462	0.354
	(0.235)		(0.239)	(0.270)		(0.374)	(0.378)
Fem/Male 15-34$_{t-1}$		-0.198	-0.270	-0.242		-0.130	-0.037
		(0.285)	(0.287)	(0.283)		(0.195)	(0.270)
Share 15-34					0.740***		
					(0.199)		
Fem/Male 15-34					-0.284		
					(0.252)		
Avg(10) yield $_{t-10}$						-0.312***	-0.252**
						(0.071)	(0.099)
Avg(10) log wage$_{t-10}$							-0.026
							(0.041)
Other controls	YES	YES	YES	YES	YES	YES	YES
Observations	74	74	74	198	76	62	56
Nr. of states	0.206	0.183	0.236	0.276	0.275	0.649	0.562
R-squared	16	16	16	16	16	16	14

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all regressions include year and state effects; control variables include capital city dummy lagged 10 years; real GDP per capita (log), rural population (log), share of urban population and its squared term.*

Including both variables in the same specification strengthen their significance suggesting that they ought to be included simultaneously (column 3). The results are also robust to using the larger sample with UN data (column 4). As demographic data are also obtained from the Census, they refer to the same year of observation as the urban data. This means that they need to be interpolated in order to obtain the value of the variable lagged one year. In order to use the actual data, I regress *PRIMACY* on both variables in the same period. Their coefficients and level of significance are little

[60] In 1991 for instance the totel stock of rural-urban migrants was 18 million for male vs. 22 million for female (Census of India 2001).

affected (column 5).

Finally, I check how the results vary when including all the determinants of the elasticity of rural to urban labour supply. Although these variables are likely to be correlated with each other, they also capture partially different channels of influence on the elasticity of labour supply. In column 6 I add *Avg(10) YIELD* to the specification in column 3. The coefficients of the demographic variables are halved although they maintain their sign (but *SHARE15-34* becomes insignificant). On the other hand the *Avg(10)YIELD* coefficient is negative and significant and is statistically larger than the corresponding one in Table 3.3 (cf. column 1). The inclusion of *Avg(10) WAGE* (column 7) reduces slightly the other coefficients, but *Avg(10)YIELD* remains significant. The coefficient of *WAGE* is negative but insignificant, which suggests that the effects of the agricultural sector on primacy are mainly captured by *YIELD.*

IV Estimation

Although the previous results hold over a variety of specifications, using different dependent variables, explanatory variables and samples, some concerns still remain over their robustness. In particular, all the agricultural variables employed are potentially subject to endogeneity bias, as the type of urbanisation pattern may be likely to influence the rural sector. The error term in (3.1) and (3.2) may then be correlated with the rural variable, thereby determining inconsistent estimates. I instrument *YIELD* and *WAGE* with rainfall levels and land reform legislation to deal with this potential endogeneity.

Table 3.5 reports the results of the first stage FE regression obtained by regressing various measures of *YIELD* and *WAGE* on the different instruments for the periods in which the second stage is run (i.e. every ten years). I average rainfall over the same period as the dependent variables and run a contemporaneous specification using state and year effects, land reform legislation plus the other set of controls included in the second stage. The analysis yields the surprising result that rainfall exerts a negative, albeit moderate, effect on *Avg(10) YIELD* (column 1). A 10% increase in the average rainfall levels in the decade is associated with a reduction in yield per hectare of sown area by 2%. This effect is somewhat puzzling as rainfall is one of the main inputs in any agricultural production function and thus deserves further analysis. One possible explanation for it is that in a flood-prone country like India too much of a good thing may become a curse. Indeed between 1988 and 1998 (except in 1992-93) India enjoyed consistently above normal levels of rainfall during the monsoons (which are responsible for 80% of rainfall in India). At least three quarter of the districts experienced above average levels of rain in each of those years (Indian Ministry of Finance, 1997 and 2001). Excessive rain can be harmful to agricultural production. For example in 1996 it caused floods in Rajasthan and Haryana, West Bengal and Bihar and extensive damage in Andhra Pradesh (Indian Ministry of Finance, 1997). This is also consistent with evidence from Besley and Burgess (2002) showing that precipitations well above the average of the period are associated with floods in Indian states. In a context where almost all of the land received enough rainfall, it may be possible that those areas that received comparatively more rainfall may have been penalised by floods and similar disruptions. This hypothesis receives some indirect support also from the results in column 2, which excludes the latest decade (i.e. the nineties) from the analysis.[61] The rainfall coefficient turns now positive (albeit it is not significant) suggesting that the negative effect of rainfall on agricultural productivity over the entire period is driven by its effect in the nineties. The change in impact of rainfall on *YIELD* in the nineties may not only be due to the above normal level of rainfall in India during this period but possibly also to the diminished dependence of agriculture on rainfall over the last decade. This is due to better irrigation systems and to the lower importance of rain-fed agriculture in India (e.g. rice accounted for a lower proportion of total production value in the nineties than in earlier decades).[62]

Unsurprisingly the cumulative land tenancy reform variable exerts a positive and significant effect over both *Avg(10) YIELD* (column 1-4) and the simple *YIELD* (column 5-6). Note that as $Avg(10)YIELD_t$ is computed over the period t-$t+9$ I find it convenient to use the cumulative land reform variable in period t so that its influence is measured over the entire decade *Avg(10)YIELD* refers to. The results hold using different lags and/or using the average of the tenancy reform variable (not shown here). The results are also robust to using the extended samples including the years for which

[61] Note that the *Avg(10)*YIELD variable is calculated as the average of YIELD over the following 10 years, therefore *Avg(10)YIELD* for the year 1991 is computed over the decade of the nineties.
[62] The difference may also be partly due to the different sources used to construct the rainfall data series. The update for the latest decade (1991-2001) has been undertaken using another source than that used in the original dataset (see data section).

also UN population data are available (column 3) as well as yearly data (column 4). The effect of rainfall on the simple value of *YIELD* is still negative but becomes less significant over the entire period of analysis (1960-2000) - column 5. Importantly, excluding the latest decade turns the relationship between *YIELD* and rainfall into an inverted U-shaped one (column 6). An increase in rainfall is beneficial for agricultural productivity in Indian states up to a point, after which it becomes damaging. This is consistent with the hypothesis presented above and suggests that using ten year-averages tends to make the relationship more linear than it is with annual values.

Table 3.5: The effects of rainfall and land reform on agricultural variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Census	Census	Census + UN	Yearly data	Census	Census	Census
	Avg(10) Yield	Avg(10) Yield	Avg(10) Yield	Avg(10) Yield	Yield	Yield	Avg(10) log Wage
Avg(10) log rainfall	-0.059**	0.041	-0.049**	-0.042*			-0.037
	(0.025)	(0.041)	(0.021)	(0.021)			(0.055)
Cumul. land tenancy reform leg.	0.027**	0.010*	0.023**	0.017*			-0.017
	(0.010)	(0.005)	(0.009)	(0.008)			(0.017)
Rainfall (log)					-0.050	0.307*	
					(0.042)	(0.151)	
Rainfall (log) sq.						-0.026*	
						(0.013)	
Cumul. land ten. reform leg.$_{t-3}$					0.018**	0.008*	
					(0.007)	(0.005)	
Cumul. abolition of intermediaries leg.							0.080**
							(0.038)
Period	1961-91	1961-81	1961-91	1961-91	1960-2000	1960-90	1961-91
Other controls	YES	YES	YES	YES	YES	YES	YES
Observations	61	45	168	532	75	59	54
Number of states	16	15	16	16	16	15	14
R-squared (within)	0.863	0.860	0.856	0.843	0.785	0.788	0.974

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; data for Haryana and Punjab not available before 1965, as they were a single state; Avg(10) indicates the number of years over which the average of a variable is calculated.*

The determinants of *Avg(10)WAGE* are slightly different (column 7). First, the negative effect of rainfall is reduced and becomes insignificant. Second the cumulative land tenancy reform legislation exerts a negative but not significant effect, while the cumulative abolition of intermediaries legislation performs better than the tenancy reform variable to influence *WAGE* with a positive and significant coefficient.

Table 3.6 presents the results of the IV estimation. The first stage F-statistics obtained from the standard IV FE estimation suggests that the instruments are weak in most specifications. To alleviate concerns that using weak instruments may generate unreliable results, I run most of the regressions using the Limited Information Maximum Likelihood (LIML) method. Stock and Yogo (2005) show that this method performs better than standard IV when instruments are weak. I also use the small sample correction whenever the size of the sample is smaller than 100 observations.[63] In addition, I make sure that all specifications are not overidentified according to the standard Hansen J test. The joint null hypothesis of this test is that the instruments are valid (i.e. uncorrelated with the error term) and that they are correctly excluded from the estimated equation. The following Tables report the values of the test statistics distributed as a χ^2 (all values reported do not reject the null hypothesis at least at the 10% level).[64]

Following the results in Table 3.5, the specifications with measures of YIELD as dependent variable use (log of) rainfall and cumulative land tenancy reform legislation as instruments. This allows to maximise the predictive power of the instruments. The coefficient of *Avg(10)YIELD* in the specification with all controls but *GINLAND* is negative and significant at the 1% level (column 1), and its magnitude is roughly double than in the OLS estimation (cf. Table 3.3, column 2). This suggests that the endogeneity tends to bias the estimates downwards. The coefficient is unaffected by the addition of *GINILAND* as a further control (column 2). This result holds also when using the land tenancy reform legislation as the only instrument (column 3). This shows that the results are not driven by the surprising (albeit possible) effect of rainfall on *YIELD*. The result is robust to the specification using the larger sample including UN (2006) data as well, and its first stage explanatory power is higher than in ther restricted sample (column 4).

[63] This correction adjusts the covariance matrix for the number of fixed effects g. The adjustment is (N-g-K), where K is the number of regressors, as opposed to the large-sample covariance matrix which has the adjustment (N-g).
[64] Under the null, the test statistic is distributed as a χ^2 in the number of (L-K) overidentifying restrictions (where L is the total number of instruments and K is the number of regressors in the second stage). A rejection casts doubt on the validity of the instruments. In the case of our analysis (i.e. 2SLS estimator), the test statistic is equivalent to Sargan's statistic, calculated as N*R-squared from a regression of the IV residuals on the full set of instruments (see Hayashi, 2000).

Table 3.6: The effects of agricultural variables on urban primacy, IV estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	primacy	primacy	primacy	primacy	primacy	primacy	primacy	Primacy
Sample	Census	Census	Census	Census + UN	Census Interpol.	Census	Census	Census
Method	LIML IV	LIML IV	LIML IV	LIML IV	FE IV	LIML IV	LIML IV	FE IV
Avg(10) $yield_{t-10}$	-0.441***	-0.449***	-0.429***	-0.452***	-0.594***			
	(0.128)	(0.137)	(0.126)	(0.085)	(0.075)			
Avg(3) $yield_{t-3}$						-0.453**		
						(0.220)		
$Yield_{t-1}$							-0.516**	
							(0.246)	
Avg(10) agr. $wage_{t-10}$								-0.337*
								(0.175)
$Capital_{t-10}$	0.048***	0.050**	0.056***	0.023**	0.019***	0.023*	0.026	0.019
	(0.016)	(0.020)	(0.020)	(0.009)	(0.005)	(0.013)	(0.016)	(0.029)
GDP pc (log)	0.116***	0.117***	0.162***	0.065***	0.036***	0.146***	0.173***	-0.121
	(0.035)	(0.036)	(0.042)	(0.015)	(0.010)	(0.054)	(0.064)	(0.092)
Rural pop (log)	-0.056	-0.067	0.136	-0.146***	-0.176***	-0.044	0.002	-0.237
	(0.097)	(0.097)	(0.129)	(0.050)	(0.033)	(0.090)	(0.138)	(0.184)
Urban share	-0.623	-0.672	0.058	-0.927***	-0.986***	-0.776*	-0.827	-0.882
	(0.436)	(0.439)	(0.529)	(0.234)	(0.158)	(0.429)	(0.629)	(0.658)
GINILAND $_{t-10}$		0.103						
		(0.125)						
Small sample corr.	YES	YES	YES	NO	NO	YES	YES	YES
Period	1961-01	1961-01	1961-01	1961-01	1961-01	1951-01	1951-01	1961-01
Observations	60	60	64	168	532	73	74	56
Nr. of states	15	15	16	16	16	15	15	14
First Stage								
Excluded Instruments	Rain, CLR1	Rain, CLR1	CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1, CLR2
F-statistics	6.23	6.38	7.74	15.07	27.62	4.24	2.79	3.49
> Stock-Yogo cr. val.	15%	15%	20%	1%	5%	25%	30%	25%
Hansen Overid. test	0.68	0.87	-	1.59	4.02	1.06	0.79	1.39

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; > Stock-Yogo critical value indicates the level of significance of the minimum critical values according to the Stock-Yogo weak identification test lying below the F-statistics of the regression. Data for Haryana and Punjab not available before 1965, as they were a single state; all regressions include state and year effects. Instruments: Rain is log of monthly rainfall (mm) in the same period as the endogenous variable; CLR1 is cumulative tenancy reform legislation (contemporaneous to the endogenous variable if the latter is averaged over ten years, lagged one year with Avg(3)Yield and lagged 3 years with Yield); CLR2 is cumulative abolition of intermediaries legislation; Avg(10), Avg(3) indicate the number of years over which the average of a variable is calculated.*

In column 5 I also run the same regression using yearly data with interpolated urban data. The use of yearly data in the first stage (which is instead based on actual rather than interpolated data) magnifies the instruments power, which is now well above the standard critical values suggested by Stock and Yogo (2005). This suggests that part of the reason for the relatively weak predictive power of the instruments in some of the specifications has to do with the limited sample size. The higher instrument power (and the larger number of observations) has the effect of increasing the significance of the coefficient as well. This result should further alleviate the concern that weak instruments are driving the significance of the coefficients in the small baseline sample. A higher instrument power would probably increase further (rather than reduce) the significance of the second stage coefficients. The results are again robust to using different length over which average values of *YIELD* are computed (three years in the case of column 6) as well as to using the simple value of *YIELD* (column 7) as dependent variables. The negative coefficient holds also in the specification with (log of) *WAGE* averaged over 10 years as the main regressor (column 8). In this case the increase in the size of the coefficients relative to the corresponding FE estimation (cf. Table 3.3, column 6) is larger than that for *YIELD*, in line with the supposed stronger endogeneity of *WAGE* than *YIELD*. IV results appear to be highly robust also to the use of the other three measures of urban concentration as dependent variables (results reported in Table A2 in Appendix 2). Hence the findings of the IV estimation largely confirm the robustness of the negative effects of the agricultural variables on primacy.

3.2.3 Urbanisation

The frameworks in chapter 2 predict also that a steeper labour supply curve would reduce the level of urbanisation. Let us turn to the test of this prediction. Table 3.7 shows the results of the regressions of the share of urban population in total population on *YIELD* and *WAGE*. The results are consistent with the prediction of a positive impact of the elasticity of labour supply on the level of urbanisation (i.e. negative and significant coefficient for *YIELD* and *WAGE*). Column 1 presents the basic specification with year and fixed effects and a parsimonious set of control variables (so as to maximise the number of observations), including (the log of) total population, income per capita and its square term. The coefficient of *Avg(10) YIELD* is negative and significant, while income has a U-shaped relationship with urbanisation and total population exerts a negative effect on urban share. The negative coefficient of *YIELD* is robust to including a number of further controls, similar to those used by Davis and Henderson (2003) in their urbanisation regressions (column 2). Such controls include agriculture's share of GDP, the ratio of manufacturing to services in GDP and land area interacted with population. In addition I also include GINILAND lagged ten years as an extra control. This should capture the role of the distribution of land in providing the incentives for rural-urban migration. Somewhat differently from Davis and Henderson, I use the percent urban in the state as the dependent variable rather than total urban population. This variable is a more direct measure of urbanisation than urban population with total population as a control. As I don't have data on urban wages, following Davis and Henderson (2003) I use sector composition measures to reflect the outcomes of urban versus rural wages and prices at any stage of development. In particular the manufacturing to service ratio may capture a related measure to urban wages since basic retail and personal services are usually found in rural villages and towns, while manufacturing production is mainly concentrated in urban areas. Therefore an increase in the ratio of manufacturing to services may be associated with higher urbanisation.[65] In this way I hope to isolate the effects of agricultural variables via changes in the elasticity of labour supply (rather than via the direct effects on urban-rural wage gap). The inclusion of these controls reduces the number of observations but the coefficient of *YIELD* is qualitatively unaffected. The coefficient remains negative and significant also when using *Avg(3)YIELD* (column 3), the simple value of *YIELD* lagged one year (column 4) and the large sample including UN data (column 5). The results also indicate that the higher the concentration of land distribution the lower the urbanisation rate after ten years (columns 2, 5 and 7). This seems to suggest that a higher concentration of land provides less incentives for rural-urban migration.[66]

The results of the regressions using *Avg(10) WAGE* (lagged 10 years) as the main explanatory variable are slightly different than for *YIELD* (columns 6-7). In particular, the effect of *WAGE* on urbanisation is non linear (U-shaped). Urbanisation first decreases as *WAGE* increases and then it increases together with *WAGE*. This effect is significant when including a parsimonious set of controls (column 6), while it becomes insignificant with the larger set of controls possibly due in part to the small sample size available (column 7). The weaker results for *WAGE* may also indicate a more important endogeneity problem than for *YIELD*, which may bias the coefficient downwards, reducing its significance.

[65] I also try to use the share of manufacturing in GDP as a control instead of the ratio of manufacturing to services, obtaining the same results.
[66] However, when including the contemporaneous value of GINILAND in the regression, its coefficient becomes positive (but not significant). In the analysis I keep the same lag as the agricultural regressor main objective of the inclusion of *GINLAND* is to control for the effects of *YIELD* on urbanisation via land distribution.

Table 3.7: The effects of agricultural variables on urbanisation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Census	Census	Census	Census	Census + UN	Census	Census
	Urb. share	Urb. share	Urb. share	Urb. share	Urb. share	Urb. share	Urb. share
Avg(10) $yield_{t-10}$	-0.145***	-0.188*			-0.144**		
	(0.037)	(0.098)			(0.064)		
Avg(3) $yield_{t-3}$			-0.216***				
			(0.070)				
$Yield_{t-1}$				-0.189**			
				(0.082)			
Avg(10) Agr. $Wage_{t-10}$						-0.086**	-0.040
						(0.031)	(0.030)
Avg(10) Agr. $Wage_{t-10}$ sq.						0.017***	0.010
						(0.005)	(0.007)
GDP pc (log)	-0.902***	-0.441	-0.388	-0.016	-0.390***	-0.577***	-0.337
	(0.166)	(0.327)	(0.230)	(0.316)	(0.123)	(0.133)	(0.290)
GDP pc (log) squared	0.056***	0.028	0.027*	0.004	0.025***	0.036***	0.025
	(0.010)	(0.021)	(0.014)	(0.020)	(0.008)	(0.009)	(0.018)
Total pop (log)	-0.219***	-0.147	-0.034	-0.026	-0.155*	-0.248***	-0.037
	(0.054)	(0.138)	(0.092)	(0.089)	(0.087)	(0.076)	(0.103)
Area x ln tot. pop (x mln)		0.0002*	0.0001	0.0002*	0.0002*		0.0002**
		(0.0001)	(0.0001)	(0.0001)	(0.0001)		(0.0001)
Manuf./ Services		0.010	0.011	0.014	0.022		0.032
		(0.031)	(0.025)	(0.034)	(0.016)		(0.047)
Agricultural share in GDP		0.076	0.061	0.097	0.060		0.023
		(0.086)	(0.068)	(0.091)	(0.040)		(0.078)
$GINILAND_{t-10}$		-0.166**			-0.085*		-0.220**
		(0.072)			(0.044)		(0.082)
$GINILAND_{t-3}$			-0.034				
			(0.074)				
$GINILAND_{t-1}$				0.002			
				(0.073)			
Observations	64	45	56	57	135	68	52
Nr. of states	16	15	15	15	15	14	14
R-sq. (within)	0.912	0.913	0.941	0.938	0.919	0.909	0.918

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all regressions include year and state effects; data for Haryana and Punjab not available before 1965, as they were a single state; X in Avg(X) indicates the number of years over which the average of a variable is calculated.*

IV Estimation

As discussed above the agricultural variables are likely endogenous due to both reverse causality and omitted variable. I employ IV estimation using the same instruments as in Table 3.6 to deal with this issue.

Table 3.8: The effects of agricultural variables on urbanisation, IV estimation

	(1) Urb share	(2) Urb share	(3) Urb share	(4) Urb share	(5) Urb share	(6) Urb share	(7) Urb share	(8) Urb share
Sample	Census	Census	Census	Census + UN	Census	Census	Census	Census
Method	LIML IV	LIML IV	LIML IV	LIML IV	LIML IV	LIML IV	LIML IV	LIML IV
Avg(10) $yield_{t-10}$	-0.207***	-0.387*	-0.266***	-0.428***				
	(0.069)	(0.230)	(0.079)	(0.159)				
Avg(3) $yield_{t-3}$					-0.253**			
					(0.106)			
$Yield_{t-1}$						-0.385*		
						(0.222)		
Avg(10) Agr. $Wage_{t-10}$							-0.177**	-0.178
							(0.076)	(0.374)
Basic controls	YES	YES	YES	YES	YES	YES	YES	YES
Extra controls	NO	YES	NO	YES	YES	YES	NO	YES
Small sample corr.	YES	YES	YES	NO	YES	YES	YES	YES
Observations	60	45	64	135	56	57	56	42
Nr. of states	15	15	16	15	15	15	14	14
First stage regression								
Excluded Instruments	Rain, CLR1	Rain, CLR1	CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1, CLR2	Rain, CLR1, CLR2
F-statistics.	11.04	2.49	10.24	4.30	5.20	2.04	3.93	1.25
> Stock-Yogo critical value	5%	30%	15%	20%	15%	35%	20%	10%
Hansen Overid test	3.15	0.79	-	1.00	0.37	1.09	1.02	2.92

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; > Stock-Yogo critical value indicates the level of significance of the minimum critical values according to the Stock-Yogo weak identification test lying below the F-statistics of the regression. All regressions include state and year effects; data for Haryana and Punjab not available before 1965, as they were a single state. Basic controls include: log of total population, log of GDP per capita and its squared term. Extra controls include: agriculture's share of GDP, the ratio of manufacturing to services in GDP, land area interacted with population and GINILAND (lagged the same number of years as the main agricultural regressor). Instruments: Rain is log of monthly rainfall (mm) in the same period as the endogenous variable; CLR1 is cumulative tenancy reform legislation (contemporaneous to the endogenous variable if the latter is averaged over ten years, lagged one year with Avg(3)Yield and lagged 3 years with Yield); CLR2 is cumulative abolition of intermediaries legislation; Avg(10), Avg(3) indicate the number of years over which the average of a variable is calculated.*

Again the IV estimation broadly confirms the FE results, confirming also that the endogeneity tends to bias the coefficients downwards (Table 3.8). This is the case for various measures of *YIELD* (columns 1-6) as well as for *Avg(10)WAGE* (columns 7-8). The negative coefficient of *Avg(10)YIELD* holds using the basic controls (i.e. total population, GDP per capita and its squared term, column 1) as well as the extra controls (column 2) as in Table 3.7. It is also robust to instrumenting *Avg(10)YIELD* only with the cumulative land tenancy reform (column 3) and to using the extended sample with UN data as well (column 4). The results hold also when using *Avg(3)YIELD* (column 5) and the simple value of *YIELD* (column 6) as the proxies for the elasticity of labour supply. Results in column 7 suggest that the coefficient of *Avg(10)WAGE* becomes linear (and remains negative) in the IV estimation with the basic set of controls. Adding the extra controls leaves the size of the coefficient unaffected but reduces the precision of the estimation (column 8). This result is driven by the reduction of the sample for which the extra controls are available (42 observations only) rather than by the effect of the controls per se.[67] Therefore the omission of the extra controls should not generate any important bias in the other coefficients.

3.2.4 Beyond the upper tail of the distribution

The preceding findings provide quite substantial evidence that lowering the elasticity of labour supply reduces the level of urban primacy and of urbanisation in Indian states. However, the empirics so far have not analysed whether changes in the labour supply curve have an impact on the entire city size distribution of the state, or they mainly influence the largest city. Two of the four measures of urban concentration employed (*UP* and *Ratio4*) have only the largest city's population as the numerator. The Herfindhal indices take into account the next nineteen largest cities as well, but the largest city (especially if it is a mega-city) is likely to exert a dominant influence on their value given the way the index is constructed. Moreover, the analytical framework of chapter 2 does not have clear predictions on this distinction, as it is based on a two city system. While the results using Herfindhal indices hint at the possibility that the rural variables affect urban concentration as well as urban primacy, more precise analysis is needed in order to shed some light on this issue. Works by Rosen and Resnik (1980) and more recently by Soo (2005) provide some evidence that developing countries tend to have more concentrated urbanisation patterns than the Zipf's law would predict, i.e. fewer and larger cities. Yet, we lack any systematic evidence to assess whether developing countries are actually following a different urbanisation

[67] I test this by running the same regression as in column (7) with and without these controls keeping the same sample as in column (8). Results available upon request.

pattern than that followed by developed ones and what role push and pull factors may have in this respect.

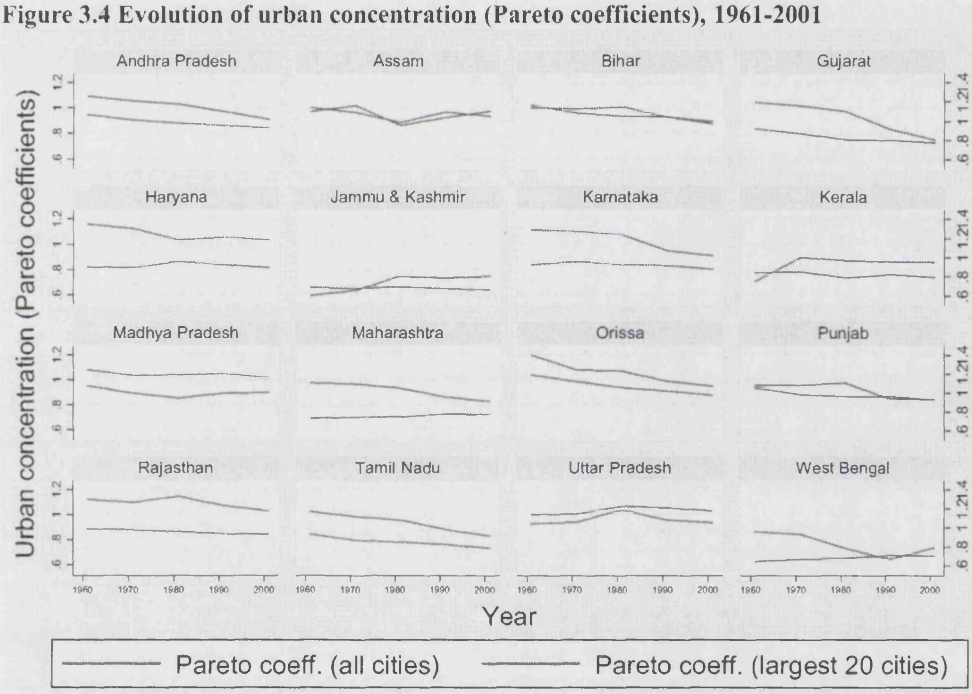
The literature has used the Pareto coefficient as the most convenient measure of urban concentration across the entire city size distribution (e.g. Parr, 1985; Alperovich, 1992, Gabaix, 1999). This is the coefficient α in the relation underlying the size distribution of cities: $y_i = Ax_i^{-\alpha}$, which can be re-written as:

$$\ln y_i = \ln A - \alpha \ln(x_i) \qquad (3.3)$$

where y_i is the number of cities with population equal or above that of city i, x_i is the population of city i and A is a constant. The Pareto coefficient is a standard measure of the city-size distribution and it is inversely related to urban concentration. This coefficient assigns the same 'importance' to all cities, therefore the largest one is not over-represented as in the other urban concentration measures used above. This feature should allow the Pareto coefficient to capture a concept closer to urban concentration than to urban primacy.

I calculate the Pareto coefficients in each Census year for every state and use them as dependent variables in regressions of the type of (3.2). If changes in the elasticity of labour supply not only affect urban primacy but also urban concentration, then the rural variables (proxies for the elasticity of labour supply) should be significant and with the opposite sign as in the regressions above (as higher Pareto coefficients indicate lower concentration). Since I have consistent data across states and over time for the twenty largest cities, I calculate the Pareto coefficients using only these cities (α_{20}). I also derive the coefficients using all cities I have data for in each year (α). There is a wide cross-state variation in the number of cities for which data are available as the urban population differs significantly across states. For example in 1991 I include data on 682 cities for Uttar Pradesh, 198 for Andhra Pradesh and only 57 for Jammu and Kashmir. Moreover, as the urban population increases substantially between successive Census years, the number of cities for which data are available increases over time too. The only exception is the year 2001 for which data on fewer urban areas are available. Therefore I exclude this year from the regressions with α as dependent variable. Figure 3.4 plots the evolution of both concentration indices between 1961 and 2001. Except for Haryana and Uttar Pradesh α_{20} lies always below α, indicating that the upper tail of the city-size distribution is more concentrated (i.e. skewed towards larger cities) than the entire distribution. The evolution of the two measures is obviously correlated although there are several significant exceptions to such correlation, e.g. Kerala in the sixties, Haryana and Rajasthan in the seventies, Gujarat after 1981, West Bengal in the nineties. These differences are at the heart of the different results I obtain in the regressions below.

Figure 3.4 Evolution of urban concentration (Pareto coefficients), 1961-2001



Source: Author's elaboration on Indian Census

The results from these regressions are reported in Tables 3.9 and 3.10 and suggest that a higher elasticity of labour supply increases urban concentration only in the upper tail of the city-size distribution. Table 3.9 presents the results using α_{20} as the dependent variable. *Avg(10)YIELD* reduces urban concentration (i.e. increases the value of the Pareto coefficient) in the simple FE regression for the 1961-2001 period (column 1) as well as for the 1961-1991 period (column 2), although the coefficients are not significant at standard levels. This contrasts with the significant (and negative) effect of agricultural variables on primacy and on the other urban concentration measures in the Tables above. Given the difference between the Pareto coefficients and the other urban concentration measures discussed above, the lack of significance suggests that a substantial part of the effects of the elasticity of labour supply on urban concentration is driven by its effects on the largest city. This is also confirmed by using the share of urban population of the second largest city as the dependent variable. The coefficients of the agricultural variables in this case are negative but never significant (not shown here - results available upon request).

Table 3.9: The effect of agricultural variables on urban concentration (using α_{20})

	(1) FE zipf20	(2) FE zipf20	(3) LIML IV zipf20	(4) LIML IV zipf20	(5) LIML IV zipf20	(6) LIML IV zipf20	(7) LIML IV zipf20
Avg(10) $yield_{t-10}$	0.126	0.252	0.283*	1.062			
	(0.108)	(0.238)	(0.150)	(0.722)			
Avg(3) $yield_{t-3}$					0.326**		
					(0.157)		
$Yield_{t-1}$						0.383*	
						(0.206)	
Avg(10) Agr. $Wage_{t-10}$							0.198
							(0.119)
Controls	YES	YES	YES	YES	YES	YES	YES
Period	1961-01	1961-91	1961-01	1961-91	1958-01	1960-01	1961-01
Small sample correction	NO	NO	YES	YES	YES	YES	YES
Observations	64	48	60	45	73	74	56
Nr. of states	16	16	15	15	15	15	14
R-squared	0.450	0.392	0.437	0.263	0.350	0.237	0.429
First Stage							
Excluded Instruments			Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1, CLR2
F-test			9.11	1.53	4.24	2.79	3.36
> Stock-Yogo cr. val.			10%	35%	20%	30%	25%
Hansen Overid test			0.65	0.66	0.97	0.39	2.26

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; > Stock-Yogo critical value indicates the level of significance of the minimum critical values according to the Stock-Yogo weak identification test lying below the F-statistics of the regression. Data for Haryana and Punjab not available before 1965, as they were a single state; all regressions include state and year effects. Control variables include capital city dummy lagged 10 years; real GDP per capita (log), rural population (log), share of urban population and its squared term. Instruments: Rain is log of monthly rainfall (mm) in the same period as the endogenous variable; CLR1 is cumulative tenancy reform legislation (contemporaneous to the endogenous variable if the latter is averaged over ten years, lagged one year with Avg(3)Yield and lagged 3 years with Yield); CLR2 is cumulative abolition of intermediaries legislation; Avg(10), Avg(3) indicate the number of years over which the average of a variable is calculated.*

The IV results (columns 3-7) suggest again that the endogeneity may bias the coefficient of *YIELD* downwards (column 3). Its size is now twice as large as that in column 1 and is significant at the 10% level. However the effect of *Avg(10)YIELD* on α_{20} appears to be comparatively smaller than the corresponding effect on *primacy*. While a 10% increase in *Avg(10)YIELD* reduces *primacy* by 4.9% (according to Table 3.6, column 1), it increases α_{20} by only 0.8%. This difference confirms again that the effect of *YIELD* operates particularly on the relative size of the largest city. I also exclude the last decade from the analysis for reasons of comparability with the results using α as the dependent variable. This increases the coefficient almost four-fold and the standard error by even more (column 4), making the coefficient significant at the 15% level only. The direction of the results holds also when using *Avg(3)YIELD* (column 5), the simple value of *YIELD* lagged one year (column 6) and *Avg(10)WAGE* as the main regressor.

These results do not hold when α is used as dependent variable (Table 3.10). The coefficient on *YIELD* becomes consistently insignificant across the various specifications both in the FE (columns 1-2) and in the IV estimations (columns 3-5). And the same applies to *WAGE* (column 6). This suggests that the agricultural variables have a different effect on α_{20} and α. What may explain such a difference? The answer must lie in the different ways of constructing the two coefficients: α_{20} captures the characteristics of the upper tail of the city-size distribution, while α includes a much wider range of cities. For example the influence of the largest city is likely to be greater for α_{20} than for α (as the largest city represents 10% of all observations over which the Pareto coefficient is derived). The other reason for the difference between the results may be related to the varying number of cities over which α is calculated. This is especially true across states as discussed above: in states with data on a large number of cities, each city has a lower influence than in states where data is available on a few cities. This cross-states difference of course disappears for the calculation of α_{20}.

As the reasons for the difference between the results for α and α_{20} are not clear, these findings need to be interpreted with caution. To the extent that such a difference reflects the fact that the two Pareto coefficients capture different portions of the city-size distribution, the results may provide further confirmation that only the upper tail of the distribution is influenced by changes in the elasticity of labour supply. If the difference between α and α_{20} is mainly related to the larger variation in the number of cities over which the values of α are calculated, then one would need a better understanding of the properties of the two Pareto coefficients in relation to the concept of urban concentration (e.g. which of the two coefficients captures the concept of urban concentration most effectively).

Table 3.10: The effect of agricultural variables on urban concentration (using α)

	(1) FE α	(2) FE α	(3) LIML IV α	(4) LIML IV α	(5) LIML IV α	(6) LIML IV α	(7) LIML IV α
L10.avg10yield_real	0.002		-1.971				
	(0.348)		(1.925)				
L3.avg3yield_real		0.121		-2.299			
		(0.182)		(1.624)			
l1yield_con					-3.163	-0.343	
					(4.462)	(0.636)	
L10.lavg10wage_80							-0.228
							(0.270)
Controls	YES	YES	YES	YES	YES	YES	YES
Period	1961-91	1958-91	1961-91	1958-91	1960-91	1960-01	1961-91
Small sample corr.	NO						
Observations	48	61	45	58	59	74	42
Number of state	16	16	15	15	15	15	14
R-squared	0.698	0.513					
First Stage							
Excluded Instruments			Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1, CLR2
F-test			1.53	1.74	0.331	2.79	1.12
> Stock-Yogo cr. val.			35%	35%	50%	30%	40%
Hansen Overid test			0.47	0.12	0.06	1.40	0.16

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; > Stock-Yogo critical value indicates the level of significance of the minimum critical values according to the Stock-Yogo weak identification test lying below the F-statistics of the regression. Data for Haryana and Punjab not available before 1965, as they were a single state; all regressions include state and year effects. Control variables include capital city dummy lagged 10 years; real GDP per capita (log), rural population (log), share of urban population and its squared term. Instruments: Rain is log of monthly rainfall (mm) in the same period as the endogenous variable; CLR1 is cumulative tenancy reform legislation (contemporaneous to the endogenous variable if the latter is averaged over ten years, lagged one year with Avg(3)Yield and lagged 3 years with Yield); CLR2 is cumulative abolition of intermediaries legislation; Avg(10), Avg(3) indicate the number of years over which the average of a variable is calculated.*

3.2.5 Discussion

The previous results provide compelling evidence that rural variables which influence the labour supply curve to the urban sector have been important determinants of urbanisation patterns across Indian states in the post-Independence period. Other things being equal, a steeper rural-urban labour supply tends to reduce urbanisation levels. Importantly and perhaps more surprisingly, it also reduces the degree of urban concentration in the upper tail of the city-size distribution. The largest city appears to benefit the most from an increase in the propensity of the rural workforce to migrate to urban areas. In terms of the two-city models presented in chapter 2 these results suggest that the descending part of the largest city's (city 1) net wage curve is flatter than that of the other city. In other words, congestion costs kick in relatively slowly in city 1. Therefore an elastic labour supply would ensure a constant flow of labour especially towards the largest city.

In the context of the model with imperfect labour mobility between cities (Figure 2.2) this flatter net wage curve arises due to the location specific externality φ, which raises the productivity of labour in city 1 above that of city 2. For example this can be thought of as the political advantage of being close to the seat of power, so as to exert rent seeking more effectively in a setting where the return to this activity is particularly high. Similarly, it can be consistent with a model where there is a limited provision of certain public goods across cities (e.g. due to scarce resources), such as good tertiary education or health services. In particular, some of these services may be provided only in the largest city. To the extent that entrepreneurs care about the provision of such services (unlike unskilled labour) they would disproportionately locate in that city. This implies that as an economy grows and more entrepreneurs set up new firms and existing firms grow (i.e. the industrialisation process), urban population growth would be increasingly accommodated in the largest city. If the probability of starting a firm and/or if the growth prospects of an existing firm depend on the availability of labour from the rural sector, then the more elastic the supply of labour is, the more this will fuel both urbanisation and the relative growth of the largest city. Therefore it is the interaction between the conditions in the rural sector and those in the different urban areas that drive the results here.

Such a sketch of model may help explain the empirical findings for Indian states in this chapter. In the period examined these states were in the early stages of the urbanisation process and most of them can be classified as low income countries. These conditions are similar to those of a number of sub-Saharan African countries today, where some evidence is emerging of an urbanisation pattern skewed towards large cities (Behrens and Polo Bala, 2006). Whether this type of story can apply only to the early stages of urbanisation and economic development remains an open question and further research is needed to shed light on this question.

What does appear clearer is that the effect of rural factors seems to play a role only on the upper tail of the city-size distribution without affecting the overall distribution. This finding is consistent with the stability of the rank size rule over time and across countries, which seem to be independent of socio-economic factors.

3.3. Conclusions

This chapter has analysed empirically the role of the rural sector in the determination of urbanisation patterns, and of urban primacy in particular, on the basis of the framework developed in chapter 2. The findings point towards the positive influence of the elasticity of labour supply (proxied by different sets of variables) on urban primacy and on urbanisation levels. The results hold for a variety of specifications using an array of dependent variables, of rural variables and controls; and for different sets of urban data. Results from IV estimation suggest that these results are causal in nature. The methodology using an intra-country rather than the usual cross-country analysis should add further robustness to the findings.

The results along with the framework of chapter 2 suggest that the generally higher level of concentration of urban systems in developing countries is determined by the combination of both pull and push factors rather than only by pull factors (via political variables) as in traditional models of urban primacy. Whether the role of the rural sector acts mainly on the largest urban area, on the upper tail of the distribution (as the findings of this chapter seem to point towards), or on the entire system of cities is still open to question.

Finally, the results of the analysis may provide a new angle to the policy debate on urban decongestion. To the extent that the rural sector is an important determinant of urban primacy (as these chapters suggest), reducing the intensity of the push factors in rural areas may help reducing (or decreasing the rate of growth in) the level of urban concentration. Analytically, this would be similar to the effects of migration restrictions, which also act on the labour supply curve making it steeper). However, restrictions contemporaneously reduce the welfare of potential migrants by preventing labour to move where it is more productive.

This may lead to a possible revaluation of the role of agricultural policies as urban de-concentration policies alongside other policies based on the action on pull factors, such as development of poles in remote areas, and political de-congestion policies (e.g. movement of the capital city, increasing the political decentralisation process).[68]

[68] Countries such as Egypt, Brazil, Korea, Mexico, and China have pursued medium size city programs designed to forestall increased growth of larger cities (World Bank, 2009 and Davis and Henderson, 2003).

Appendix 3.1
Procedures to estimate urban populations by the UN Population Division

The United Nations Population Divisions compute the both total urban population and individual urban agglomerations' population. The explanation of the computation methods is based on UN (2004).

Urban population estimates
The proportion of the population living in urban areas is estimated by country or area for the period 1950-2030 in five-year intervals. Once values of the proportion urban at the national level are established for the 1950-2030 period, they are applied to the estimates and projections of the total national population of each country or area derived from *World Population Prospects: The 2002 Revision* (United Nations, 2003) so as to obtain the corresponding urban population for 1950 to 2030. Calculation of the proportion urban during the estimation period involves interpolation between recorded figures and extrapolation back to 1 July 1950 when the earliest of recorded figures refer to a later date. Such interpolation or extrapolation to 1950 is based on the *urban-rural ratio (URR)*, defined as the ratio of the urban to the rural population, that is:

$$URR(t) = U(t)/R(t) \qquad \text{(A.6)}$$

where U(t) and R(t) denote the urban and the rural populations at time t, respectively. The urban-rural ratio at time t is directly related to the percentage urban ($PU(t)$) since

$$PU(t) = URR(t)/[1+URR(t)] \qquad \text{(A.7)}$$

Letting $rur(t,n)$ denote the growth rate of the urban-rural ratio between time t and $t+n$, it follows
that

$$rur(t,n) = \ln(URR(t+n)/URR(t))/n \qquad \text{(A.8)}$$

where, substituting URR for its equivalent according to (A.6), one obtains

$$rur(t,n) = [\ln(U(t+n)/R(t+n)) - \ln(U(t)/R(t))]/n = [\ln(U(t+n)) - \ln(R(t+n)) - \ln(U(t)) + \ln(R(t))]/n =$$
$$[\ln(U(t+n)/U(t)) - \ln(R(t+n)/R(t))]/n = u(t,n) - r(t,n) \qquad \text{(A.9)}$$

where $u(t,n)$ denotes the growth rate of the urban population between t and $t+n$, and $r(t,n)$ is the growth rate of the rural population between the same time points. That is, the growth rate of the urban-rural ratio is equivalent to the difference between the growth rates of the urban and the rural populations. Therefore, $rur(t,n)$ is known as the urban-rural growth difference and it is the basis for the interpolation and extrapolation of the proportion urban. Thus, if T is any time point within the intercensal period $(t, t+n)$,

$$URR(T) = URR(t)\exp[rur(t,n)(T-t)] \qquad \text{(A.10).}$$

The use of (A.10) for interpolation and extrapolation purposes implies that rur is assumed to remain constant during each intercensal period and during the period 1950 to the reference date of the second observation available. Once an estimate of $URR(T)$ is available, it can be converted to $PU(T)$ by using equation (A.7).

Urban agglomerations estimation

Estimates and projections of the population of cities with an estimated of 100,000 or more within the 1950-2000 period are considered, provided data on their population size is available from a census or population register. For the *2003 Revision*, a total of 3,284 cities or urban agglomerations was considered. Because countries take population censuses at different times, the actual dates of observation vary from city to city, although they are usually identical for cities within a particular country. Consequently, just as with the estimates of the proportion urban, the first step in preparing estimates and projections of city populations consists in estimating the population size of all cities for the same dates in the past.

To estimate the population of cities on 1 July of the years 1950, 1955, 1960 and so on, the procedure is similar to that described for the proportion urban. However, in this case the interpolation or extrapolation is based on the difference between the growth rate of a city minus the growth rate of the population of the rest of the country. Specifically, if one considers the ratio of the city population at time t, $C(t)$, to the population of the rest of the country, $RES(t)$, that is

$$CRR(t) = C(t)/RES(t) \qquad \text{(A.11)}$$

where $RES(t) = P(t) - C(t)$ and $P(t)$ is the total population of the country at time t, then the growth rate of CRR between t and $t+n$, denoted by $rcr(t,n)$, is

$$rcr(t,n) = [\ln(CRR(t+n)) - \ln(CRR(t))]/n \qquad \text{(A.12)}$$

which is equivalent to

$$rcr(t,n) = c(t,n) - res(t,n) \qquad \text{(A.13)}$$

where $c(t,n)$ is the growth rate of the city's population between t and t+n, and $res(t,n)$ is the growth rate of the rest of the country's population between t and $t+n$. Then, the value of CRR for any time T within the period $(t, t+n)$ is given by:

$$CRR(T) = CRR(t)\exp[rcr(t,n)(T-t)] \qquad \text{(A.14)}$$

The same equation can be applied to obtain extrapolated values of CRR when T is outside the intercensal period $(t, t+n)$ and that period is the closest to T. Then, because the proportion of the total population living in the city at time T, $PC(T)$, is equivalent to:

$$PC(T) = CRR(T)/[1+ CRR(T)] \qquad \text{(A.15)}$$

that proportion can be calculated for time T and multiplied by an independent estimate of the country's population to obtain the population of the city at time T. Such independent estimate is obtained from the country-level estimates published in *World Population Prospects: The 2000 Revision* (United Nations, 2001).

Appendix 3.2

Table A1: Using different lags and dependent variables

	(1) *Ratio4*	(2) *Herf*	(3) *Herf2*	(4) *Primacy*	(5) *Ratio4*	(6) *Herf*	(7) *Herf2*
Avg(3) $yield_{t-3}$	-5.081*	-0.166**	-0.161***				
	(2.493)	(0.066)	(0.047)				
Avg(5) $yield_{t-5}$				-0.158**	-4.927*	-0.162**	-0.155***
				(0.067)	(2.479)	(0.069)	(0.050)
Obs.	77	77	77	64	64	64	64
R-sq. (within)	0.459	0.445	0.497	0.516	0.474	0.532	0.506
Nr. of states	16	16	16	16	16	16	16

	(8) *Ratio4*	(9) *Herf*	(10) *Herf2*	(11) *Ratio4*	(12) *Herf*	(13) *Herf2*	(14) *Primacy*
$Yield_{t-1}$	-3.372	-0.093	-0.107**				
	(2.154)	(0.065)	(0.047)				
Avg(10) log $wage_{t-10}$				-1.884	-0.070*	-0.052*	
				(1.216)	(0.036)	(0.029)	
Avg(3) log $wage_{t-3}$							-0.038
							(0.026)
Obs.	78	78	78	68	68	68	54
R-sq. (within)	0.392	0.351	0.406	0.343	0.328	0.285	0.273
Nr. of states	16	16	16	14	14	14	14

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all regressions include state effects, year effects and the full set of controls as in the Tables in the main text (except GINILAND); data for Haryana and Punjab not available before 1965; Avg(5) indicates that the value of the variable is averaged over 5 years.*

Table A2: Robustness for IV regression using other measures of urban concentration

	(1)	(2)	(3)	(4)	(5)	(6)
	Ratio4	*Herf*	*Herf$_2$*	*Ratio4*	*Herf*	*Herf$_2$*
Method	LIML IV	LIML IV	LIML IV	LIML IV	LIML IV	LIML IV
Avg(10) Agr.	-17.431**	-0.463***	-0.424***			
Yield_{t-10}	(7.639)	(0.160)	(0.116)			
Yield_{t-1}				-22.953*	-0.675	-0.510**
				(13.756)	(0.495)	(0.221)
Controls	YES	YES	YES	YES	YES	YES
Small sample correction	YES	YES	YES	YES	YES	YES
Observations	60	60	60	74	74	74
Nr. of states	15	15	15	15	15	15
First Stage						
Excluded Instruments	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1	Rain, CLR1
F-test	6.23	6.23	6.23	2.79	2.79	2.79
> Stock-Yogo cr. val.	15%	15%	15%	25%	25%	25%
Hansen Overid. test	1.46	1.29	0.54	0.88	0.95	0.65

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all regressions include state and year effects; data for Haryana and Punjab not available before 1965, as they were a single state; control variables include capital city dummy lagged 10 years; real GDP per capita (log), rural population (log), share of urban population and its squared term. Instruments: Rain is log of monthly rainfall (mm) in the same period as the endogenous variable; CLR1 is cumulative tenancy reform legislation (contemporaneous to the endogenous variable if the latter is averaged over ten years, and lagged 3 years with Yield); Avg(10) indicate the number of years over which the average of a variable is calculated.*

Chapter 4. Does urbanisation affect rural poverty? Evidence from Indian districts[69]

4.1. Introduction

The typical transformation of an economy from agricultural and mainly rural to industrial and predominantly urban in the process of development has long been a well established fact (Lewis, 1954; Kuznets, 1955). However, the direct implications of this transformation on the economic welfare of the population during this process remain less apparent. In particular, what happens to surrounding rural areas when a city grows? Does the area's population receive economic benefit from it and if so, to what extent? In a period of increasing urbanisation in most developing countries, answers to these questions can have important implications for development policies.

There is still little knowledge about the actual economic impact of urbanisation on rural areas. This chapter represents one of the first efforts to fill this gap, as it tries to measure the impact of urbanisation on rural poverty in the Indian context. The chapter uses district-level panel data between 1981 and 1999 to show that urbanization has been an important determinant of poverty reduction in rural areas. In our preferred estimations, we find that an increase of 100,000 urban residents in the representative district (21% increase from the mean) implies a decrease of between 3 and 6 percentage points in the incidence of poverty in the district's rural areas.

This analysis becomes more important when considering that most of the world's poor reside in rural areas, where the incidence of poverty is higher than in urban areas across all developing regions. In 1993 rural areas accounted for 62% of the world population and for 81% of the world's poor at the $1/day poverty line; in 2002 after a period of intensive urbanisation the same figures stood at 58% and 76% respectively (Ravallion et al., 2007).[70] The process of urbanisation (which mostly concerns the developing world) has been accompanied by an unequal distribution of the global reduction in poverty rates. Between 1993 and 2002 while the number of $1/day poor in

[69] This Chapter is co-authored with Carlo Menon.
[70] In fact the actual poverty line used by Ravallion et al. (2007) is $1.08/day; to save clutter we refer to it as the $1/day poverty line.

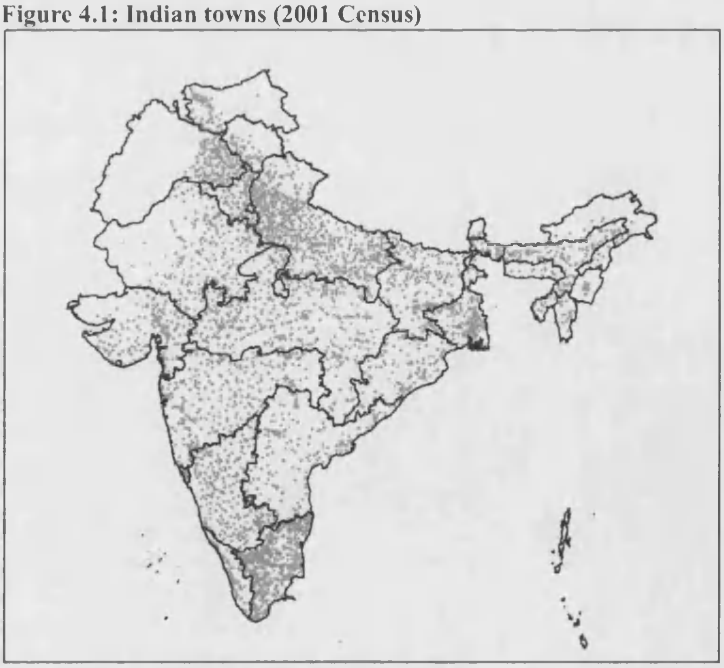
rural areas declined by 100 million, that of urban poor increased by 50 million. Ravallion et al. (2007) explain this "urbanisation of poverty" through two related arguments.[71] First, a large number of rural poor migrated to urban areas, thus ceasing to be rural poor and either they have been lifted out of poverty in the process (through a more productive use of their work) or they have become urban poor. This is a direct (or 'first-round' in Ravallion et al. (2007) terminology) effect of urbanisation on rural poverty. Second, the process of urbanisation also impacts the welfare of those who remain in rural areas through second-round effects. The overall impact of urbanisation on rural poverty is substantial but, in the absence of data on the poverty profile of rural-urban migrants, it is not possible to distinguish between the two effects. We mainly focus on these second-round effects, trying to control for the direct effects of urbanisation on rural poverty.

Distinguishing between first and second-round effects is important. The former involves only a statistical association between urbanization and changes in rural poverty due to the change in residency of some rural poor (who may or may not be lifted out of poverty in their move to the urban areas). This entails no causal link. On the other hand, second-round effects capture the impact of the urban population growth on the rural rate of poverty. Such a relationship is causal in nature and tells us how good or bad urbanisation is for rural poverty. In a developing country context, understanding this relationship is particularly important because most of the population in these countries will continue to be rural for at least another decade and for another three decades in the LDCs.[72] This figure, along with the recognition that poverty has a higher incidence in rural than urban areas, suggests that it is on this rural non-migrant population that the implications of urbanisation will be most important for global poverty reduction in the near future.[73] The focus on developing countries is essential given that almost the entire future population growth in urban areas (94% in 2005-2030) is predicted to take place in developing countries (UN, 2008).

[71] The term "urbanization of poverty" was first introduced by Ravallion (2002).
[72] Based on calculations on UN (2008) data, developing countries are expected to become more urban than rural in 2018 and LDCs in 2045.
[73] This does not deny the importance of urban poor in global poverty. In fact these represent a substantial and increasing share of poor globally (although still lower than rural poor). However, estimating the effects of urbanisation on urban poverty would require another model altogether and it is left to the future research agenda.

We consider Indian urbanisation at the district-level for the period 1981-1999. During this period the country urbanised at a relatively slow rate: the urban population was 23.3% of the total in 1981 and 27.8% in 2001 (Government of India, 2001). However, given the sheer size of the Indian population, this moderate increase turned into a massive rise in the absolute number of urban dwellers (126 million). This represents an increase of almost 80% in the urban population over this period. These figures mask a large variability in urbanisation patterns at the sub-national level; states have urbanised at very different rates. Among the major states, Tamil Nadu increased its share of urban population from 33% to 44% between 1981 and 2001, while Bihar maintained the same urbanisation rate over this period (13%). The differences are also evident in absolute terms: Uttar Pradesh increased its urban population by 28 million people (+140%); at the other extreme West Bengal increased its urban population by only 8 million (+56%). Not only are the urbanisation dynamics different, but so is the geographical spread of urban areas. Figure 4.1 shows that the density of towns is concentrated in Northern India, roughly in the area along the Ganges river and in the South-East (Tamil Nadu in particular). Other areas, such as Andhra Pradesh, Madhya Pradesh and the North-West have significantly lower densities.

Figure 4.1: Indian towns (2001 Census)



Note: the State of Delhi is excluded from the map
Source: Authors' elaboration on data from Indian Census 2001, and data on city spatial coordinates from Indian Gazetteer and GPSvisulizer.com.

This variability (both in levels and in changes) is even more remarkable at the district level, as the left hand-side map in figure 2 shows. For instance, a district like Idukki in Kerala increased its urban population by 13,000 (+29%) between 1981 and 2001, while the urban population in Rangareddi (Andhra Pradesh) increased by 1.6 million (+416%) and in Pune (Maharashtra) by 2.4 million (+130%) over the same period. In the subsequent analysis we try to exploit this variability to identify the impact of urbanisation on rural poverty.

In this period India also provides an interesting case in terms of the policy environment and economic performance because the country experienced structural changes in economic policy, rate of growth, and poverty levels. After a long period of economic planning and import substitution industrialisation, the government started reforming the economy toward a more liberal regime in 1991. This change was brought about by the external payment crisis due to the government's deficit spending. Possibly helped by the liberalisation of the economy, economic growth took off since the mid-1980s, and more evidently since 1993, having increased more rapidly than in the 1960s and 1970s (Datt and Ravallion, 2002). Despite disagreements on the extent to which economic growth increased the welfare of India's poor, poverty in India declined steadily in the 1990s, particularly in rural areas (Kijima and Lanjouw, 2003). The geography of the decrease in the share of poor, however, is extremely variegated, as the right hand side map in figure 4.2 shows. While in many districts more than 30% of rural population was lifted out of poverty between 1983 and 1999, for around a quarter of them the share of poverty has remained roughly constant or has even worsened over the same period.

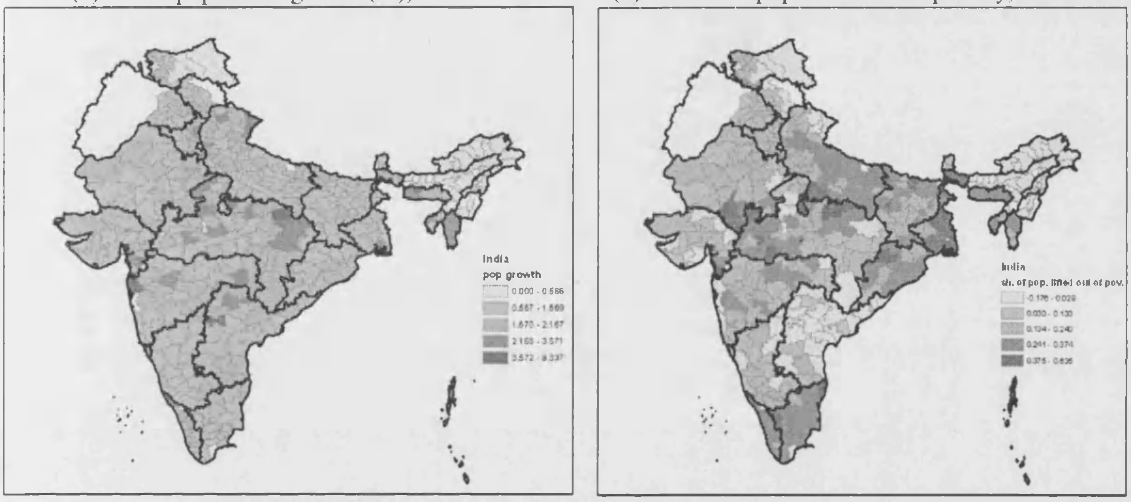
This chapter's geographical focus is particularly important as India is the country with the largest number of both rural and urban poor. Its number of $1/day rural poor in 2002 was over 316 million, representing 36% of the world's rural poor. Moreover, its urbanisation process is still in its infancy with only 28% of the population being urban in 2000. The country is expected to add a further 280 million urban dwellers by 2030.[74] Thus estimating the impact of urbanisation on rural poverty in India

[74] This is based on authors' calculations on UN (2008).

may help identify the potential effects of this expected massive growth of urban population on the world's largest stock of rural poor.

Figure 4.2 – Urban population growth (%) and poverty reduction, by district 1981-99

(a) Urban population growth (%), 1981-97 (b) % of rural pop. lifted out of poverty, 1983-99



Note: the map (b) reports the difference between the district poverty share in 1983 and 1999. E.g., a value of 0.30 means that in 1983 the share of poor rural population was 0.3 bigger than in 1999. The State of Delhi is excluded from the map
Source: Authors' elaboration on Indian Census and NSS (various rounds).

4.2. Urbanization and rural poverty: Channels

Why would the increase in urban population have an impact on poverty in surrounding rural areas? There are various ways in which urbanisation and rural poverty are linked. We can distinguish between a simple composition effect due to migration of poor from rural to urban areas (first round effect), and a spillover effect due to positive externalities of urbanization on surrounding urban areas (second round effect). In the following, we analyse the main mechanisms through which the latter effect may take place. Then we discuss the way in which we can try to isolate second-round from first-round effects.

4.2.1 Second round effects

There are at least six main indirect channels through which urban population growth may affect rural poverty in surrounding areas: backward linkages, rural non-farm employment, remittances, agricultural productivity, rural land prices and consumer prices.

Backward linkages: An expanding urban area (both in terms of population and income) will generate an increase in the demand for rural goods. For perishable products and in general for those products without spatially integrated markets (e.g. due to high transportation costs), such a demand will typically be met by surrounding rural areas; while the other agricultural products could be provided by locations farther away. This is linked to an idea that goes back to von Thünen's (1966) theory of concentric circles of agricultural specialisation around cities that is determined by the size of transport costs. Rural locations close to urban areas specialise in high transportation cost goods, while locations farther away specialize in lower transport cost commodities. The farther one moves away from cities the more likely it is for rural communities to be self-subsistent in both agricultural and non-agricultural commodities. This is similar to the pattern found by Fafchamps and Shilpi (2003) for Nepal.

This channel is likely to operate via an *income* as well as a *substitution* effect. The former is related to the increased demand for agricultural goods due to higher incomes in urban areas relative to rural areas. Such a higher income is usually explained by urbanisation economies: urban areas have denser markets for products and factors, which raise labour productivity and wages over the level of rural areas (see Fujita et al., 1999). The substitution effect relates to the increased share of higher value added products in total agricultural demand typical of more sophisticated urban consumers. Empirical evidence confirms this composition effect. Parthasarathy Rao et al., 2004 found that Indian districts with an urban population over 1.5 million have a significantly higher share of high value commodities than the other districts. Thanh et al. (2008) show that per capita consumption of high value fruit in Vietnam has increased faster in urban than in rural areas over the nineties.

Rural non-farm employment: Expanding urban areas may also favour the diversification of economic activity away from farming, which typically has a positive effect on incomes (see e.g. Berdegue et al., 2001; Lanjouw and Shariff, 2002). This effect is particularly important in rural areas surrounding the cities. Three concomitant effects may explain such increased diversification. First, proximity to cities may allow part of the peripheral urban workforce to commute to the city to work. This in turn generates suburban non-farm jobs in services, such as consumer services and retail trade, which are needed by the growing commuter population. Second, as cities provide dense markets to trade goods and services more efficiently, rural households close to cities may afford to specialise in certain economic activities (based on their comparative advantage), relying on the market for their other consumption and input needs (Fafchamps and Shilpi, 2005). This more extensive specialisation should boost productivity and income (Becker and Murphy, 1992). Third, proximity to urban areas stimulates non-farm activities instrumental to agricultural trade (which is increased by urbanization), such as transport and marketing. Recent evidence from Asia provides strong support for the effect of cities in stimulating high return non-farm employment in nearby rural areas (see Fafchamps and Shilpi, 2003 on Nepal, Deichmann et al., 2008 on Bangladesh and Thanh et al., 2008 on Vietnam). On the other hand, and consistent with this line of argument, isolated rural communities do not tend to specialise and rely on subsistence activities dominated by farming. The growth of urban areas would raise the share of rural areas that are close enough to cities to develop a substantial non-farm employment base.

Remittances: Remittances sent back to rural households of origin by rural-urban migrants constitutes another potentially important second-round effect of urbanization on rural poverty. The vast majority of rural-urban migrants (between 80% and 90%) send remittances home although with varying proportions of income and frequency (Ellis, 1998). To the extent that urbanization is (partly) fuelled by rural-urban migration, this growth may be associated with larger remittance flows to the rural place of origin. The positive effects of remittances in reducing resource constraints for rural households as well as providing insurance against adverse shocks (as their income is uncorrelated with risk factors in agriculture) have been shown by the literature (Stark, 1980, Stark and Lucas, 1988). On the other hand the migrant's family often provides economic supports (monetary or in kind) to the migrant during his initial stay in the urban area. This support aimed at covering the fixed costs of migration can be interpreted as an investment whose main return is the counter urban-to-rural remittances flow which is received afterwards (Stark, 1980). This urban-to-rural remittance flow may somewhat reduce the net resources transferred to rural areas by urban workers.

Agricultural (labour) productivity: As discussed in chapter 3, urbanisation and rural poverty can also be linked by the changes in rural labour supply that accompany the urbanisation process. To the extent that rural-urban migration reduces the rural labour supply, this may increase (reduce the decrease of) agricultural labour productivity, given the fixed land supply and diminishing marginal returns to land.[75] This may pose some upward pressure on rural wages. There is indeed evidence some areas of India of out-migration from rural areas being associated to higher wages in sending areas (Jha, 2008).

Rural land prices: The growth of cities may increase agricultural land prices (owned by farmers) in nearby rural areas due to the higher demand for agricultural land for residential purposes. This may generate increased income for landowners through sale or lease, or through enhanced access to credit markets, where land acts as collateral. Some evidence from the US indicates that expected (urban) development rents are a relatively large component of agricultural land values in US counties which are near or contain urban areas (Plantinga et al., 2002). The impact on rural poverty through this channel depends on the way this increased income is distributed across the rural population. Typically, if land is very concentrated, this channel is likely to benefit a few landowners, potentially restricting access to waged agricultural employment for the landless population. To illustrate, let us assume the extreme case of all rural land concentrated in the hands of one landowner, who employs labour to cultivate it. If the growth of the nearby city pushes the price of the land above the expected value of the discounted stream of profits from cultivating the land, the landowner will sell it. This would leave all the agricultural labourers in the district unemployed. The net effect on poverty will depend on the extent to which the new use of the land will be able to absorb labour (e.g. via construction-related employment). However, given the constraints to the reallocation of agricultural labour across sectors and the high labour intensity of agriculture, we would expect the net effect on rural poverty to be adverse (i.e. increase in rural poverty) when land is highly concentrated (and vice-versa).

[75] In fact Eswaran et al. (2008) show that land to labour ratios decreased in most states in India over 1983-1999 as rural population growth rate more than offset rural-urban migration. In this case our argument would become: to the extent that rural-urban migration reduces the growth of the rural labour supply, this may reduce the decrease of agricultural labour productivity.

Consumer prices: because the growth of a city is associated with lower consumer prices, this may benefit surrounding rural consumers who have access to urban markets. This effect may be due to increased competition among a larger number of producers in the growing urban area as well as to thicker market effects in both factors' and goods' markets (e.g. Fujita et al., 1999).

A further potential channel may relate to early arguments made by Jacobs (1969) and Dore (1987) that agriculture in rural areas surrounding cities also benefits from spillover effects in technology and marketing. However, to the best of our knowledge, no specific evidence has been provided in support of this view yet.

Table 4.1 summarises the expected net effects of these second-round channels on rural poverty as well as their expected reach on rural areas according to the discussion above. The total net effect of urbanization on rural poverty is predicted to be negative (i.e. poverty reducing) with the bulk of the effects being felt at a relatively small distance to the urban area (in surrounding rural areas). The next sections will detail the methodology used to test these hypotheses by measuring this total net effect in the case of Indian districts.

Table 4.1: Ex-ante second-round effects of urbanization on rural poverty

	Predicted net effect	Expected reach of the effect
Backward linkages	**Negative**	Nearby rural
Share of non-farm employment	**Negative**	Peri-urban
Remittances	**Negative**	Rural
Changes in agricultural productivity	**Negative**	Rural
Rural land prices	**Pos/Neg** (depending on land concentration)	Nearby rural
Consumer prices	**Negative**	Nearby rural

Note: Reach of the expected effect is defined in descending order of distance from the urban area as: Rural; Nearby rural and Peri-urban.
Source: Authors' elaboration on the basis of main text.

4.2.2. Disentangling first and second round effects

As discussed above, we are particularly interested in estimating the second-round effects of urbanization on rural poverty. To do this we first need to disentangle the two effects and then to identify an appropriate way to control for the first round effects in the empirical analysis. This section deals with the former task. Let us assume N distinct geographical units (districts), each with population P_{it} at time t, split between urban (P_{it}^{U}) and rural areas (P_{it}^{R}), with $i \in [1, N]$. We can characterise the incidence of poverty (H_{it}^{R}) in rural areas in district i at time t as a function of the urban population of the district and a series of other characteristics of the district (such as its total population, specific policies, etc.), represented by the vector X:

$$H_{it}^{R} = f(P_{it}^{U}, X_{it}) + \varepsilon_{it} \qquad (4.1)$$

Let us assume that natural growth rate is zero and the only changes in the rural-urban split of the population are determined by one (or both) of these two phenomena: intra-district rural-urban migration or rural areas becoming urban (either because they are encompassed by an expanding urban area or because their population has grown sufficiently to upgrade from the status of village to that of town).[76] Define α_t as the number of rural poor divided by rural population (i.e. headcount poverty in rural areas), σ_t as the number of *poor* rural-urban migrants divided by the number of rural *poor* and λ_t as the number of rural *poor* in villages which become urban areas divided by the number of rural *poor*. Define also γ_t as the number of rural-urban migrants divided by rural population at time t and φ_t as the number of rural dwellers residing in villages

[76] This does not consider the possibility of inter-district migration, nor of urban-rural migration. The latter is relatively unimportant in influencing the rural-urban split of the population in a country like India. The stock of urban-rural migrants represented less than 1.4% of total population in the majority of Indian districts in 1991, with mean equal to 1.7% (based on the Indian districts database at the University of Maryland – see below). Inter-district migration represents instead a substantial share of total migration, in particular rural-urban. In 1991 it accounted for less than 34% of total migration for the majority of Indian districts (with mean equal to 37%); the share of inter-district migration in total rural-urban migration was even larger in 1997 (median 46%, average 49%). However, the empirical analysis below rejects the relevance of this type of migration in determining rural poverty. We could reconcile this finding with the model presented here by assuming that the distribution of inter-district migrants in both the sending and the recipient districts follows the rural-urban distribution of the those district's populations.

which become urban at time *t* divided by rural population at time *t* (with $\gamma_t \geq \alpha_{t-1}\sigma_t$ and $\varphi \geq \alpha\lambda$). We can then re-write (4.1) as:

$$H_{it}^R = \frac{\overbrace{\alpha_{t-1}P_{it-1}^R}^{\substack{\textit{Rural}\\ \textit{poor at t-1}}} \times [1 - \overbrace{(\sigma_t + \lambda_t - \sigma_t\lambda_t)}^{\substack{\textit{Share rural poor turning}\\ \textit{urban between t-1 and t}}}]}{\underbrace{P_{it-1}^R}_{\textit{Rural pop at t-1}} \times [1 - \underbrace{(\gamma_t + \varphi_t - \gamma_t\varphi_t)}_{\substack{\textit{Change in rural pop}\\ \textit{between t-1 and t}}}]} + g(P_{it}^U, X_{it}) + \varepsilon_{it} \tag{4.2}$$

The first term on the right hand side of (4.2) defines the first-round effects of the growth of urban population on rural poverty. Its numerator represents the number of rural poor at time t as if the change in this number (between t and t-1) were only due to the change of status of those rural poor (at *t-1*) becoming urban dwellers at *t* (through parameters σ_t and λ_t). The denominator represents the total rural population at *t*.

The condition under which this first-round effect decreases rural poverty incidence (ceteris paribus) is $\frac{\alpha_{t-1}P_{it-1}^R(1-\sigma_t - \lambda_t + \sigma_t\lambda_t)}{P_{it-1}^R(1-\gamma_t - \varphi_t + \gamma_t\varphi_t)} < \frac{\alpha_{t-1}P_{it-1}^R}{P_{it-1}^R}$. Ignoring the terms $\sigma_t\lambda_t$ and $\gamma_t\varphi_t$ as they are likely to be very small and the subscripts to save clutter this condition becomes:

$$\sigma + \lambda > \gamma + \varphi \tag{4.3}$$

The key variables here are the poverty distributions of both rural-urban migrants and dwellers of rural-urban transitional areas relative to the poverty distribution of the rural population. Expression (4.3) states that if the distribution of migrants is skewed towards low income individuals – i.e, the incidence of poverty is higher among migrants than non migrants – and if the poverty incidence in rural villages that become urban is higher than that in total rural population of the district then rural-urban migration will directly reduce rural poverty. Recent cross-country evidence by Ravallion et al (2007) seem to be consistent with the validity of condition (4.3). They find a sizeable negative effect of urbanisation on the incidence of rural poverty and concomitantly an increase in the number of urban poor with urbanisation. Although they cannot isolate the direct effects of rural-urban migration, their findings would be hard to reconcile without condition (4.3) holding. Although there is no evidence establishing empirically the relative size of the parameters in (4.3), some studies find that those rural areas on the outskirts of (large) urban areas may benefit economically from this vicinity (e.g. Fafchamps and Shilpi, 2003 for Nepal). This may imply lower levels of rural poverty in those peri-urban areas about to be incorporated into urban areas, i.e. $\lambda < \varphi$. This means that the poverty incidence among rural-urban migrants needs to be substantially higher than that among rural non-migrant population for expression (4.3) to be verified, i.e. $\sigma > \gamma + (\varphi - \lambda)$.

As the main aim of this chapter is to estimate the size and direction of the second-round effects of urbanization on rural poverty, we can re-arrange (4.2) to control for the direct effects of urbanisation as well as for other covariates of rural poverty:

$$H_{it}^{R}(P_{it}^{U} | \sigma_{it}, \gamma_{it}, \lambda_{it}, \varphi_{it}, X_{it}) = h(\sigma_{it}, \gamma_{it}, \lambda_{it}, \varphi_{it}) + g(P_{it}^{U}, X_{it}) + \varepsilon_{it} \quad (4.4)$$

This expression represents the basis of the empirical analysis described in the next section. Effectively we need to estimate the partial derivative of H_{it}^{R} with respect to P_{it}^{U}. The channels described above should underlie the second-round effects that we are trying to capture through this partial derivative.

4.3. Empirical methods

Using a district-level analysis, we try to systematically assess whether and to what extent urbanisation in Indian districts during the 1981-1997 period has affected rural poverty in those districts. In order to evaluate the eventual effects of urbanization on the people in extreme poverty, we also use specifications of rural poverty which try to isolate changes in the intensity of poverty for the very poor.

We argue that districts are an appropriate spatial scale for such an analysis in India as all of the first and second-round channels described above are likely to display most of their effects within the district's boundaries. This is consistent with the theoretical discussion above, arguing that the effects of city growth are concentrated in surrounding rural areas. Various pieces of specific evidence on India confirm that this is likely to be the case.

First, evidence suggests that intra-district migration in India is a large component of total rural-urban migration. According to the Census (Government of India, 1991), 62% of the total stock of permanent internal migrants was intra-district in 1991, although a share of this stock was composed of women migrating for marriage reasons.[77] However, a consistent part of internal migration in India is not captured by the Census because it does not involve change in residence. This may include various forms of temporary migration, such as seasonal and circular as well as commuting. Such migration may account for an important part of income generation and livelihoods in several rural areas (Deshingkar and Start, 2003, and Deshingkar, 2005). Due to its temporary nature, this migration is likely to be short-distance. In a recent survey of a number of rural villages in two Indian states, Deshingkar and Start (2003) reported that in a number of villages several households were commuting daily to nearby urban locations (although this movement was not registered in the migration data) and in one village, one entire caste took up casual labouring in the urban sector. This does not deny the existence of long-distance migration in India, which in fact was increasing during the nineties (Jha, 2008). However, long distance rural-urban migration is mainly directed to a few growing metropolitan areas, such as Mumbai, Delhi, Bangalore and Chennai, which are excluded from the analysis.[78] Notwithstanding the importance of intra-district migration, in the empirical section we also test the robustness of the results against the relative size of the intra-district migrant population.

Second, during the period of analysis (1981-1999) most perishable agricultural goods' markets do not appear to be well integrated at the national or even at the state level in India. This is due to relatively poor transport infrastructure networks and lack of appropriate technology (such as cold storage facilities).[79] Agricultural produce is often sold in nearby towns and even most trade in livestock tends to occur at a short distance. This is due to lack of infrastructure, which brings livestock marketing costs to distant

[77] This is in line with Topalova (2005), who finds limited labour mobility across Indian regions between 1983 and 2000.
[78] We exclude them either because the district which contains them does not have any rural area (e.g. Delhi, Urban Bangalore) or because the effects of their growth are likely to extend well beyond the boundaries of their district.
[79] Infrastructure endowments have to certain extent been upgraded since then.

markets up to 20-30 percent of the sale price (Chandra Mohan Reddy, 2000). As a result, most transactions in live animals take place within the same district (Birthal, 2005). Thus we would expect a consistent share of agricultural trade to occur at a small distance, making districts a suitable spatial scale to capture a substantial part of the first two channels above as well. In line with these ideas, some studies have performed district level analyses to try to capture demand-side effects on agriculture. Parthasarathy Rao et al. (2004) for instance analyse the effects of urbanisation on agricultural diversification into high value commodities, such as fruit, vegetables, dairy products, using districts as the unit of analysis.

There is also emerging evidence of increases in land prices in peri-urban and rural areas surrounding urban agglomerates. Land values in those areas may be well above the discounted future stream of income from agricultural activity, inducing several landowners to sell the land (Jha, 2008).[80]

The core idea of the empirical analysis is to assess the effects of urbanization on rural poverty at the district level over time. For that we estimate equation (4.4) trying to control for the direct effects of urbanisation as well as for other determinants of rural poverty. We use the basic specification:

$$H^R_{dt} = \beta_0 + \gamma_d + \beta_1 P^U_{dt-j} + \beta_2 \left[(\sigma_{dt} + \lambda_{it})/(\gamma_{dt} + \varphi_{it})\right] + \chi X_{dt} + \varepsilon_{dt} \qquad (4.5)$$

where H^R_{dt} is a measure of rural poverty in district d at time t, γ is district fixed effects, P^U_{dt-j} is the urban population of district d at time t-j (where $j \in [0,2]$), $\left[(\sigma_{dt} + \lambda_{it})/(\gamma_{dt} + \varphi_{it})\right]$ is a term capturing the direct effects of urbanization on rural poverty, i.e. the term $h(\sigma_{it}, \gamma_{it}, \lambda_{it}, \varphi_{it})$ in (4.4), and X is a vector of controls, which include other variables likely to have independent impact on rural poverty. The district's

[80] All of this evidence seems to be roughly consistent with Fafchamps and Shilpi (2003), who find that in Nepal the effects of proximity on rural areas peters out beyond a four hour radius (in travel time) around cities. Using the boundaries of Indian districts as in 1987, the average district size in our analysis is around 7,300 Km^2. If we approximate the district with a circle, a city located in the centre of it would be at around 50 Km from the boundary of the district. It is plausible that in several districts this distance could be covered in about three to four hours on rural Indian roads during the period considered.

urban population is computed as $P_{dt}^{U} = \sum_{i=1}^{N_d} u_{it-j}^{d}$, where u_{it-j}^{d} is the population of town i in district d at time t-j (where $j \in [0,2]$) and N_d is the number of cities in district d. Given the above discussions, we would expect $\beta_1<0$ and $\beta_2<0$.

4.4. Data and variables

Data to run specification (4.5) comes from three main sources: district level measures of poverty are available from various rounds of the Indian household survey data (National Sample Surveys), which have been appropriately adjusted by Topalova (2005) for the 1983-84, 1987-88, 1993-94 and 1999-2000 rounds of the NSS.[81] Other district level data, such as population composition come from the Indian districts database at the University of Maryland (which has been extrapolated from the original data in the Indian Census).[82] Data on town populations are available from various rounds of the Indian Census. In addition, for crop production volumes and values we use the district level database for India available with International Crops Research Institute for semi-Arid Tropics (ICRISAT) from 1980 to 1994 and recently updated by Parthasarathy Rao et al (2004) up to 1998.[83]

The district classification has been modified during the period of analysis, as some districts have been split into two units. Topalova (2005) created a consistent classification by aggregating the 2001 districts originated from the splitting into the district division of 1987. We conform to this re-aggregation and modify the original population and demographic data accordingly.

Dependent variables: We use two standard Foster Greer Thorbecke (FGT) measures of poverty as dependent variables: the poverty headcount ratio and the poverty gap index. These have already been defined in Appendix 1.1. Both measures are increasing in poverty, i.e. a higher value means a higher level of poverty.[84]

---

[81] Although each survey was carried out over two years, we refer to them with the first of the two years.

[82] Available at www.bsos.umd.edu/socy/vanneman/districts/codebook/index.html

[83] The original source of this data is the Government of India, Directorate of Economic and Statistics, Ministry of Agriculture and Cooperation.

[84] In the subsequent analysis we also run some specifications using poverty rate as a control and poverty gap as the dependent variable. This tries to capture a concept more closely related to extreme poverty, as

Population variables: as mentioned in chapter 1 the Census 1991 (and 2001) classifies towns as all the statutory places with a municipality, corporation, cantonment board or notified town area committee, or, alternatively, places satisfying simultaneously the following three criteria: i) a minimum population of 5,000; ii) at least 75 per cent of male working population engaged in non-agricultural pursuits; and iii) a density of population of at least 400 per sq. Km. This is consistent with the classification of the 1981 Census, except for condition iii), which required a minimum population density of 1000 per sq. Km. The year effects should anyway control for eventual problems of statistical consistency of urban data over time. The NSS uses the Census definition to classify urban vs. rural areas, thus ensuring the consistency of data across sources.

There were 5179 towns that met these criteria in 2001. We calculated the total urban population at the district level, by summing the figures for towns. Due to its peculiar nature, we excluded from the dataset the State of Delhi and the districts of the other megalopolises, Calcutta, Chennai, Bangalore and Mumbai; we also excluded three other districts due to an extraordinary increase in urban population in the period under study, which is extremely likely to be imputable to errors in the data: Anantapur in Andhra Pradesh, Kanniyakumari in Tamil Nadu, and Thane in Maharashtra.

As population data are available only with a ten-year frequency (1971, 1981, etc.), we estimate the values for the year 1997 by non linear interpolation in order to conduct the analysis for three rounds of the NSS. We first estimate the yearly growth rate in the period 1991-1997, calculating a weighted average of the growth rate of the 1981-1991 and 1991-2001 periods; we then calculate the 1997 population applying the estimated growth rate to the 1991 level.[85] In this way we try to reduce the potential endogeneity of the urban population to rural poverty interpolated only using the 1991-2001 growth rate. The main results are also robust to using interpolated 1997 data based only on the 1991-2001 growth rate (results available upon request).

---

it nets out the share of poor (poverty rate) from the share of the poor weighted by each poor's distance from the poverty line (poverty gap).

[85] The exact specification adopted is the following: $pop(1997) = pop(1991)*[1+yg(1981\text{-}1991)*0.3+ +yg(1991\text{-}2001)*0.7]^6$, where yg(t-T) is the yearly growth rate of the period t-T.

There are 431 districts in Topalova's (2005) original dataset, 409 of which have a positive urban population (at least for one of the three time periods); total population figures are available for only 363 of these, therefore constituting our main sample of analysis; in the year 2001, this sample accounts for a total population of 1,000,053,152 of which 270,153,691 are urban residents, corresponding to 97% and 94% of the Indian total respectively.

Controls: Following the discussion in section 4.2, we would need data on the poverty profile of rural urban-migrants (σ_{dt}) and of dwellers of areas which are rural at t-1 and become urban at time t (λ_{it}) in order to properly estimate β_2 in expression (4.5), i.e. the direct effects of urbanization on rural poverty. Unfortunately this data is not available, thus we proxy for it by including variables measuring the extent to which migrants (and dwellers of rural areas turning into urban areas) are over- or under-represented among the poor (σ_t) relative to the whole rural population (γ_t).[86] We use two types of such variables.

The first is the district's urban poverty rate H^U_{dt}. To see why, let us re-express H^U_{dt} on the basis of the variables in question. Consider that H^U_{dt} depends on urban poverty at *t-1*, on the share of rural-urban migrants at time *t* whose income in the urban sector is below the urban poverty line and on the change in the poverty rate of previous urban dwellers between *t* and *t-1*.[87] Dropping the subscript *d* to save clutter, we have:

$$H^U_t(\pi_t, P^R_{t-1}, \gamma_t, \sigma_t) = \frac{\overbrace{\psi_{t-1}P^U_{t-1}}^{\text{Urban poor at t-1}} + \overbrace{\rho_1(\pi_t)(\gamma_t - \alpha_{t-1}\sigma_t)P^R_{t-1}}^{\text{Non poor rur-urb migrants becoming urban poor between t and t-1}} + \overbrace{\rho_2(\pi_t)\alpha_{t-1}\sigma_t P^R_{t-1}}^{\text{Poor rur-urb migrants becoming urban poor between t and t-1}} + \overbrace{\Delta\psi_t(\pi_t)P^U_{t-1}}^{\text{Change in poverty of existing urban stock between t and t-1}}}{\underbrace{P^U_{t-1} + \gamma P^R_{t-1}}_{\text{Urban population at time t}}} \tag{4.6}$$

where ψ_{t-1} is the urban poverty rate at time *t-1*, ρ_1 and ρ_2 are respectively the share of non-poor rural migrant ($\gamma_t - \alpha_{t-1}\sigma_t$) at time *t* and the share of poor rural migrants $\alpha_{t-1}\sigma_t$ at time *t* who have become urban poor at time *t* (both are a function of urbanisation rate at time t, π_t); $\Delta\psi_t$ is the change in poverty rate (between *t-1* and *t*) of the existing stock of

[86] Note that for ease of exposition in the following discussions on the direct effects of urbanisation on rural poverty we refer only to rural-urban migrants and not to those who live in villages that become urban areas.
[87] For the sake of simplicity we do not consider here rural-to-urban transformation of villages. Adding it would not change the basic argument.

urban population at *t-1*. From this expression it follows that $\rho_1 \leq \rho_2$ and $\partial\rho_1 / \partial\pi_t < 0$, $\partial\rho_2 / \partial\pi_t < 0$. For any values of π_t we can compute the condition for which $H_t^U < H_{t-1}^U$ (i.e. a reduction in the urban poverty rate between *t-1* and *t*) as:

$$z(\sigma, \gamma | \pi_t) = \alpha\sigma(\rho_1 - \rho_2) + \gamma(\psi - \rho_1) > \Delta\psi P_{t-1}^U (P_{t-1}^R)^{-1} \qquad (4.7)$$

with $\partial z / \partial\sigma \leq 0$ (as $\rho_1 \leq \rho_2$) and $\partial z / \partial\gamma \leq 0$ if $\psi \leq \rho_1$.

Equation (4.7) implies that for any given value of urban economic growth at time *t*, urban poverty is more likely to have decreased between *t* and *t-1* the lower the share of rural poor that migrated to the urban areas during this period (σ_t). This is explained by the fact that the probability of poor rural-urban migrants becoming urban poor (after migrating) is higher than the same probability for non-poor rural-urban migrants. On the other hand a smaller rural-urban migrant population will decrease urban poverty only if the incidence of poverty in this population, once it becomes urban, is larger than the pre-existing incidence of poverty in the urban area ($\psi \leq \rho_1$). Condition (7) therefore implies that the evolution of urban poverty over time should capture the evolution of the parameters γ and σ at time *t* for any given value of π_t. This means that at any given time urban poverty should capture the combined effect of economic growth and of the direct effects of urbanisation on rural poverty (the term $h(\sigma_{it}, \gamma_{it})$ in (4.4)).[88]

We also control for the first-round effects of urbanization on rural poverty through the socio-demographic composition of the rural population (i.e. age and literacy). Again, this is an indirect form of control and is probably less effective than the share of urban poor in capturing first-round effects. The rationale behind it relies on the assumption that the income distribution of migrants can be expressed as a function of the migrants' age composition. Other things being equal, poverty incidence tends to be lower among young adults (i.e. 15-34), as they represent the most productive age class. Therefore the higher the share of young adults in the total migrant population (relative to their share in the rural population) the lower the probability that urbanisation will

[88] Following the criticism of Hasan et al. (2006) on the potential bias in Indian urban poverty data at the district level, we use urban poverty at the regional level, which is a Census-based aggregation of a few districts together.

directly reduce rural poverty. Rewriting expression (4.3) (without considering rural areas becoming urban for ease of exposition) we have: $\frac{\sigma}{\lambda}(\lambda_{15-34}) > 1$, with $\partial \frac{\sigma}{\lambda} \big/ \partial \lambda_{15-34} < 0$, where $\lambda_{15\text{-}34}$ is the share of people aged 15-34 in total migrants relative to their share in the rural population. The same argument can be applied to literate migrants. As we do not observe the composition of the migrants' population, we can only control for it indirectly through the composition of the actual rural population. This is based on the plausible assumption that the change in the number of young adults in the rural population is inversely related to the change in their number in the rural-urban migrant population in the same period.

This assumption is supported by the results of regressing the 1981-91 change in the urban population in the 15-34 age group ΔP^{U}_{15-34} on the change in the rural population in the same age group ΔP^{R}_{15-34} (controlling for changes in district's total population and total population in 1981):

$$\begin{array}{lllll} \Delta P^{U}_{15-34} = & -4954 & -1.038\Delta P^{R}_{15-34} & +0.2554\Delta P^{tot} & +0.0123P^{tot}_{t-10} \\ & (2.57) & (29.44) & (38.71) & (11.93) \end{array}$$

N=334 $\quad R^2 = 0.97 \quad$ (robust t-statistics in parenthesis)

The coefficient of ΔP^{R}_{15-34} is not statistically different from -1 indicating that changes in the rural population are reflected in mirror changes in the urban population (through either rural-urban migration or rural-to-urban change in status of villages).

Obviously, the incidence of young adults (as well as literates) in the rural population also directly and positively affects rural income and thus has a direct impact on the poverty rate. Therefore this variable will capture two contrasting effects on rural poverty: a first-order poverty reducing effect and a second-order poverty increasing effect (which should capture part of the direct effect of urbanisation on rural poverty). It should be clear that the ability to control for first round effects of these two types of variables (urban poverty rate and socio-demographic characteristics) is only residual to their direct relationship with rural poverty. Thus they are not likely to fully control for the first round effects of urbanisation on rural poverty. However, to the extent that they can control for at least part of these effects, the direction of change in the urban population coefficient after the inclusion of these variables should provide an idea of the likely intensity of first-round effects.

Aside from the controls of first-round effects, we need to control for any other determinants of rural poverty. We use two variables which should control for the composition of the rural population: the number of people in the age group 15-34, and the proportion of literates in this age group. The latter variable is meant to capture the level of literacy of the most productive part of the population, following the idea that the most powerful influence of education on income and poverty is through its labour market effect. We also include in some specifications the share of rural population which is reported as scheduled castes and scheduled tribes, as this is expected to have an independent (adverse) effect on poverty.

However it is likely that other unobserved factors affect the relationship under scrutiny. We exploit the panel dimension of our dataset to deal with that. First, we include district fixed effects, which absorb any time-invariant component at the district level, such as geographical position, climatic factors, natural resources, etc. Second, we add a whole set of state-year dummies, which control for state-specific time-variant shocks (including economic dynamics and policies). The inclusion of these controls may still not completely account for three other sources of potential bias in the coefficient of interest β_1 (capturing the second-order effects of urbanization on rural poverty in (4.5)).

First, there may be unobserved time varying district-specific shocks which may affect both rural poverty and urban population. For example there may be a localised shock (e.g. the election of an effective district government) which spurs district's economic growth. As economic growth is generally associated with urbanisation, this may foster urbanization while reducing rural poverty at the same time. This omitted variable problem would imply a spurious negative association between the two variables. Data on income per capita at the district level is not available to us. However, as economic growth directly affects urban poverty (as described above) the inclusion of the urban poverty rate in the controls should minimise this problem.

Second, unobserved time varying rural specific shocks may affect urbanisation via increases in agricultural productivity. This view is supported by a long-standing argument in development economics that a country's urbanisation (and industrialisation) process is fuelled by increasing agricultural productivity (e.g. Nurske, 1953). In closed economies an expanding urban population requires increases in productivity of the rural sector in order to be sustained. However, Matsuyama (1992) shows that in open economies this need not be the case, as they may rely on agricultural import for their subsistence (as in the case of the East Asian newly industrialised economies). In our case, districts can safely be considered as small open economies (within India), trading across borders in most agricultural markets. Thus this potential source of bias may not be very relevant in the analysis.[89] In line with this Fafchamps and Shilpi (2003) do not find that agricultural productivity of nearby rural areas is an important determinant of city size in Nepal. To be on the safe side, we also control for a measure of agricultural productivity. The variable is constructed as the sum of the total quantities of 22 different crops produced in a given district, multiplied by the average India-wide price of the respective crop in the same year and divided by the district's rural population. We use an India-wide price instead of district specific prices to minimise both the data gaps (which are several for the latter) and the potential endogeneity of districts' prices to rural poverty. This is in some way an extra control because it may eat up some of the effects of urbanization on rural poverty, which may occur via its effects on agricultural productivity (see channel two above).[90]

Instrumental variable: Finally, following the findings in chapter 3 there may be a problem of reverse causation to the extent that the conditions in the rural sector affect urbanisation. In particular, here we are concerned that rural poverty may drive rural-migration. It could either act as a push factor (i.e. poorer people migrate in search of an escape out of poverty) or, in the presence of high fixed costs of migration, it may act as a restraint to migration. If the former case prevails (i.e. poverty is mainly a push factor), the coefficient β_1 in (5) would have a downward biased; while the opposite is true if the latter effect of poverty on migration dominates. The findings by Ravallion et al (2007) that global rural-urban migration has been associated with large reduction in the number

[89] This argument is not necessarily at odds with the district-level backward linkages channel described above. Urban areas tend to import agricultural products relatively more by surrounding rural areas, but this does not rule out that they can rely on inter-district agricultural trade as well.
[90] Data on agricultural production is not available for all the districts. The inclusion of this variable implies a reduction of the sample to 275 districts.

of rural poor lends some credit to the importance of the former case. Kochar (2004) also provides indirect support to this hypothesis, showing that in India landless households have the highest incidence of rural-urban migrants among rural households.[91]

Regardless of the direction of the bias, we need an additional variable to act as a valid instrument, i.e. it must be correlated with district urban population, but must also be exogenous to poverty-induced rural-urban migration flows. A variable which satisfies both prerequisites is the number of people who migrate to urban areas of the district from states other than the one where the district is located. It is plausible to assume that rural poverty in a given district has no effect on migration decisions in other states, which typically do not share the same rural condition of the district in question. On the other hand, the number of migrants coming to district towns from other states is part of the urban population of the district, thus bearing a positive association with our main explanatory variable.

A concern about the exogeneity of the instrument may arise from the fact that, within a given district, both migration to the cities and rural poverty are likely to be affected by the underlying, unobserved economic trend. However, the first stage of the IV estimation includes all the controls listed in the OLS specification, and particularly the rate of urban poverty, the measure of agricultural productivity, and the interaction of time and States' fixed effects. We argue that these variables would absorb most of the economic trend in the district, thus limiting the potential bias originating from the instrument endogeneity.

Although measurement error is not likely to be a major cause of concern in our analysis, it is worth noticing that the IV estimation may also correct eventual biases arising from errors in the measurement of urban population. This is the case if the measurement error of the instrument and that of the instrumented variable are independent.

4.5. Results

Table 4.2 presents the descriptive statistics for the main variables used in the analysis while Table 4.3 presents the results from regression (4.5) using OLS

[91] His finding emerges in the context of the response of rural schooling decisions to the possibility of employment in urban areas, which tends to be larges amongst landless households.

estimation. Our dataset includes observations of 363 districts for three different time periods: 1983, 1993, and 1999. We run (4.5) applying a two years lag to the measure of urban population and to the other demographic controls for two main reasons. First, in this way we reduce the risk of potential simultaneity bias. Second, the two-year lag allows us to minimise the use of interpolation for obtaining the Census variables (both population and socio-demographic variables), which are recorded in 1981, 1991 and 2001.[92] We also include district and state-year fixed effects in all specifications. Standard errors are robust to heteroscedasticity (using the Huber-White correction) and allow for intra-group correlation within individual observations.[93]

Table 4.2: Descriptive statistics of the main variables, 1981-99

	Obs	Mean	Std. Dev.	Min	Max
Rural poverty (share)	1,170	0.321	0.183	0.004	0.81
Poverty gap index, rural	1,170	0.076	0.061	0	0.315
Rural 15-34 age (share)	1,003	0.247	0.025	0.2	0.326
Rural literates 15-34 age (share in 15-34)	1,003	0.485	0.179	0.107	0.997
Rural poors (abs. nr)	1,000	567,725	485,956	320	4,127,495
Rural population	1,003	1,668,426	982,274	15,078	8,247,888
Scheduled caste (share)	1,001	0.177	0.084	0	0.545
Agr. productivity	793	0.216	0.266	0	3.261
Urb. migr. from other states	1,007	31,098	54,077	0	545,521
Urban population	1,200	436,497	550,895	0	4,526,745
Urban poverty (share)	1,131	0.255	0.178	0	0.701

Sources: see section 4.4

4.5.1. 1981-1999 period

We run a number of different specifications in Table 4.3, testing the robustness of the results to the inclusion of a number of controls and the use of different dependent variables. When controlling only for rural population (as well as for the range of fixed effects described above), the result indicates that the growth of urban population exerts a highly significant poverty reducing effect on rural areas (column 1). This result is robust to the inclusion of socio-demographic controls for the rural population, including the share of scheduled caste, the share of young adults (15-34 age group) in the rural

[92] In any instance the results are not sensitive to the change in the time lag, i.e. applying a 1 and 0 year lags (results available upon request).
[93] Note that the main results are robust to more basic computations of the standard errors as well.

population and the share of literates in the young-adults rural population (column 2).[94] These last two variables are meant to capture a change in the composition of the rural population and therefore should partly absorb the first round effects of urbanization on rural poverty. The inclusion of these controls slightly decreases the urban population coefficient. The signs of the controls are as expected, except for the share of literates: a higher share of young adults decreases poverty, while a higher presence of scheduled caste increases it (although not significantly). This suggests that the direct effect on poverty of the young adult population prevails over their indirect effect which captures the rural-urban migration of young adults. The share of literates has a poverty-increasing, albeit not significant, effect. At a closer inspection, this unexpected effect of literacy is driven by its Post-1993 impact. As shown in column 3, the coefficient of this variable turns negative (but not significant) when we account for the significant poverty increasing impact of literacy in the post-1993 period. In this period a higher incidence of literates in the most productive part of the rural labour force was associated with higher levels of rural poverty. Understanding the rationale of such an unexpected result is beyond the scope of our analysis, but we will suggest a possible reason for this below.

Accounting for this differential impact determines also an increase in the urban population coefficient, as its effect is probably estimated with more precision. This coefficient is slightly above that of column 1, suggesting that rural socio-demographics may be capturing some first-round impact of urbanisation, which in this case increases rural poverty. As discussed above, this would be the case if a high level of urbanization was fuelled by high intra-district migration rates. Considering that young adults are over-represented in the migrant population, and that this is the most productive (and thus least poor) part of the population, there may a positive association between urbanization and poverty via this type of first-round effects. The rest of the direct effects of urbanization on rural poverty should be captured by the inclusion of urban poverty rate as a control. This is significantly and positively correlated with rural poverty (column 4). As urban poverty captures both the effects of district's economic growth (πt) on rural poverty and the direct effects of urbanisation on rural poverty, this suggests that the former are larger than the latter i.e. $\left|\partial H_t^U / \partial \pi_t\right| > \left|(\partial H_t^U / \partial \sigma) + (\partial H_t^U / \partial \gamma)\right|$ in

[94] We tried to include the share of scheduled tribes in rural population as well, but that is never significant in the different specifications we tried. As this variable is systematically less significant than the scheduled caste variable, we only include the latter as a control.

(4.6). The inclusion of urban poverty reduces the absolute size of the urban population coefficient, confirming that the rural poor tend to be over-represented in the migrant population. However this reduction is very mild: the coefficient goes from -0.066 to -0.062 (column 3 to column 4).[95] Following the discussion in the preceding section, we interpret this as a clear indication that most of the effect of urbanization on rural poverty is given by "second-round" mechanisms.

The magnitude of the effects of urban population on rural poverty over the 1981-1999 period is not particularly strong although it is robust. An increase in the district's urban population of 200,000 (a 43% increase from the mean value) reduces on average the poverty rate by 1 to 1.4 percentage points according to the specifications. Given that the average share of rural poverty over the period considered is 32%, this effect ranges between 3.2% and 4.2% of the mean poverty rate.

Results using the poverty gap index as the dependent variable are less robust than those using the poverty rate (columns 5 and 6). Urban population exerts a negative effect on the poverty gap with the other controls keeping the same sign as in the preceding regressions. This result appears to be driven by the effects of urbanisation on those poor who are relatively close to the poverty line. When the rural poverty share is included among the explanatory variables, the urban population has a positive albeit not significant effect on the poverty gap (column 6), which suggests that the poor closer to the poverty line are those who benefit most from urbanisation. This category does not include those poor far behind the poverty line. In the absence of more precise data, we could only speculate about why this may be the case. The effects of urbanisation are not likely to concern the very poor much. For example, the increase in demand for agricultural goods may affect those involved in commercial agriculture, specifically those who own capital and/or certain skills not usually available to the very poor. The same can be said about rural-urban migration: the very poor may not have enough capital to cover the fixed costs of migration. For these reasons urbanisation seems to have a fairly neutral effect on the very poor rural dwellers. Interestingly, the presence of rural dwellers from the scheduled caste is negatively associated with severe poverty.

[95] Note that this reduction is in no way attributable to the slight change in the sample's composition from 363 to 354 districts, as confirmed by running the same regression as in column 3 on the same observations as those of column 4 (results available upon request).

Along with the results from the preceding regressions, this suggests that the scheduled caste population tends to be concentrated among the rural poor close to the poverty line, but not among those in severe poverty.

Table 4.3: The effects of urbanization on rural poverty across Indian districts, 1981-1999

	(1) Rural pov. (share)	(2) Rural pov. (share)	(3) Rural pov. (share)	(4) Rural pov. (share)	(5) Poverty gap	(6) Poverty gap	(7) Rural poor (millions)
Urban pop. (millions)	-0.0616***	-0.0522**	-0.0655***	-0.0615***	-0.0157**	0.00192	-0.1220**
	(0.0220)	(0.0206)	(0.0206)	(0.0218)	(0.00776)	(0.00388)	(0.0517)
Rural pop. (millions)	-0.0126	-0.0192	-0.0110	-0.00758	-0.00193	0.000250	0.9739***
	(0.0163)	(0.0160)	(0.0162)	(0.0149)	(0.00511)	(0.00220)	(0.2193)
Scheduled caste (share)		0.194	0.0686	0.314	-0.0417	-0.132**	0.9605
		(0.284)	(0.278)	(0.299)	(0.116)	(0.0583)	(0.6086)
Rural pop 15-34 age (share)		-2.920***	-3.881***	-4.103***	-1.330***	-0.151	
		(0.770)	(0.825)	(0.826)	(0.271)	(0.120)	
Rural lit 15-34 (% in 15-34)		0.0450	-0.112	-0.122	-0.0203	0.0147	
		(0.179)	(0.172)	(0.167)	(0.0566)	(0.0217)	
Rural lit 15_34 x Post-1993			0.237***	0.215***	0.0807***	0.0189**	
			(0.0680)	(0.0656)	(0.0200)	(0.00821)	
Urban poverty (share)				0.326***		0.287***	
				(0.0616)		(0.00831)	
Rural poverty (share)					0.106***	0.0122	0.3987***
					(0.0210)	(0.00855)	(0.1098)
Rural pop 15-34 age (mln)							-0.2127*
							(0.1166)
Rural literates 15-34 (mln)							-0.1706*
							(0.09710)
Observations	997	996	996	964	964	964	964
No. of districts	363	363	363	354	354	354	354
R-sq. (within)	0.65	0.65	0.66	0.68	0.757	0.949	0.582

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all explanatory variables are lagged two years except for Agricultural Productivity (1 year lag) and urban poverty (contemporaneous).*

We also test for the effects of urbanisation on the *number* of rural poor (column 7), obtaining similar results. For every increase in urban population by 100 people the rural population in poverty decreases by 13 people. The other controls are in numbers rather than in shares (except for scheduled caste). Following the discussion in section 4.4, this represents a different way of controlling for the first round effect of urbanisation on rural poverty. In this way, the urban population variable may capture some of the effects of changes in the remaining rural population (net of the young adult population). The controls maintain the same sign as in the previous regressions, except for the rural population, which is now positive and significant and literates in the 15-34 year group, which is now negative and significant. The former result is expected as, other things being equal, a larger rural population is associated with more rural poor. The latter captures the direct association between literacy and poverty, which is negative. This may differ from the preceding regressions using shares because those may capture second-order effects of literacy on poverty.[96]

4.5.2. 1981-1993 period

We now examine the impact of urbanisation on rural poverty using only the first two time periods available, covering the time interval 1981-1993. This is a robustness check for our results with three time periods, as in this case no interpolation of urban population is needed. It is also an interesting analysis focusing only on the pre-liberalisation period. Overall, the effect of urbanisation on rural poverty is stronger than over the entire period (Table 4.4). The coefficient for the urban population ranges between -0.08 (column 1) and -0.11 (column 3) depending on the specification; this is almost twice as large as the range reported in Table 4.3. An increase in the district's urban population of 200,000 reduces on average the poverty rate by between 1.6 and 2.2% of total rural population. The basic specification without controls (except for the fixed and year effects) confirms the negative relationship between urbanisation and rural poverty, although it is only mildly significant (column 1). The inclusion of socio-demographic controls increases the significance and the size of the coefficient, again confirming that some adverse first-round impacts of urbanisation on rural poverty are taken away by these controls (column 2). Both the share of young adults in the rural population and the share of literates in the young adult population exert a poverty-reducing impact. This supports the hypothesis of a differential impact of literacy on rural poverty over time, i.e. poverty-reducing up to 1993 and then poverty-increasing. The results are robust to the addition of the share of urban poverty (column 3). However, this time the magnitude of the coefficient of urban population increases from 0.099 (column 2) to 0.111 (column 3). This increase suggests that the first-round effects

[96] When we control for the number (instead of the share) of urban poor to better control for first-round effects of urbanisation on rural poverty, the elasticity of reduction in rural poor is slightly lower (results available upon request).

of urbanisation on rural poverty captured by urban poverty may have been poverty-increasing in the eighties. Again this is a very small change, confirming that second-round effects are likely to dominate first-round ones. The impact of urbanisation on the poverty gap index is negative but less significant than for the entire period (column 4), while the impact on severe poverty seems to be neutral again (column 5). Finally, the results also hold when using the number of rural poor as a dependent variable (column 6). Again, the elasticity of poverty reduction is much higher than that considered in the 1981-1999 period.

Table 4.4: The effects of urbanization on rural poverty across Indian districts, 1981-1993, OLS

	(1) Rural pov. (share)	(2) Rural pov. (share)	(3) Rural pov. (share)	(4) Poverty gap	(5) Poverty gap	(6) Rural poor (millions)
Urban pop.	-0.0791	-0.0928*	-0.111**	-0.0265	0.00549	-0.2814**
(millions)	(0.0592)	(0.0553)	(0.0549)	(0.0168)	(0.00809)	(0.1114)
Rural pop.	0.0061	0.0082	0.0047	-0.0028	-0.0015	0.1661***
(millions)	(0.0221)	(0.0316)	(0.0022)	(0.0082)	(0.0066)	(0.0435)
Scheduled caste		0.0691	0.383	-0.00927	-0.120	0.8171
(share)		(0.398)	(0.505)	(0.207)	(0.114)	(1.0103)
Rural pop 15-34		-4.619***	-5.313***	-1.739***	-0.207	
age (share)		(1.306)	(1.408)	(0.473)	(0.224)	
Rur. literates		-0.700***	-0.835***	-0.179**	0.0620	
(share in 15-34)		(0.216)	(0.255)	(0.0845)	(0.0408)	
Urban poverty			0.378***	0.140***	0.0310	0.4839***
(share)			(0.106)	(0.0396)	(0.0233)	(0.1867)
Rural poverty					0.288***	
(share)					(0.0116)	
Rural pop 15-34						-0.1202
age (mln)						(0.2152)
Rural lit. 15_34						-0.4245***
age (mln)						(0.1119)
Observations	682	682	659	659	659	659
No. of districts	363	363	354	354	354	354
R-sq. (within)	0.611	0.640	0.660	0.763	0.940	0.589

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%.*

4.5.3. Further robustness

To control for the possible endogeneity due to the potential effects of agricultural productivity on urbanisation, we add a measure of agricultural productivity to the list of controls. This variable is lagged one year, given that the simultaneity bias should not be an issue in this case (but a contemporaneous specification is not possible due to the lack of data for 1999). The main results reported in Table 4.5 appear to be robust to the inclusion of such a measure. Surprisingly, the urban population coefficient for the entire period increases (column 1). However, this effect is mainly due to the restricted sample for which agricultural data is available. When we run the same regression as in Table 4.3 column 4 with the same sample as in Table 4.5 column 1, the increase in the size of the urban coefficient disappears (column 2). To the extent that part of the poverty-reducing effects of urbanisation may operate through increases in agricultural productivity (see section 4.2 above), the unchanged urbanisation coefficient is a somewhat puzzling result. The key to explain this may be the surprisingly weak (negative) effect of agricultural productivity on rural poverty (column 2). If this is the case, then the effects of urbanisation via productivity increases would be fairly insignificant as well. In fact, when restricting the analysis to the 1981-93 period, the coefficient of agricultural productivity becomes negative (as expected) and the magnitude of the urbanisation impact on rural poverty decreases slightly, although it maintains its significance (column 3 vs. column 4). This suggests that agricultural productivity may have had a different impact on rural poverty in the post-1993 period. Column 5 confirms such a hypothesis, as the post-1993 effect of productivity appears to have been robustly adverse to rural poverty.

Such a surprising finding may be in contradiction with earlier literature on India, which shows the key effect of higher farm yield in poverty reduction only until 1994 (Datt and Ravallion, 1998).[97] Investigating the reasons behind this adverse post-1993 impact is beyond the scope of our analysis, and we only speculate about a possible explanation for it. This may lie in the (negative) effect of agricultural productivity on rural employment in the non-farm tradable sector (e.g. rural industry). Foster and Rosenzweig (2004) find this pattern for Indian villages and explain it through the negative incentives that agricultural productivity growth provides to capital in the non-farm tradable sector through higher wages. To the extent that non-farm growth is especially pro-poor (as rural industry tends to productively employ the main asset of

[97] However, our result may appear to be at odds with recent work by Eswaran et al. (2008), finding that increases in agricultural productivity explain most of the rise in agricultural wages in the 1983-1999 period. The contradiction may be more apparent than real due to substantial methodological differences. First, Eswaran et al. use agricultural wages as an indicator of poverty; second, they perform the analysis on the whole economy without distinguishing between the rural and urban sector; finally, they do not use econometric techniques to estimate the impact of the agricultural productivity on agricultural wages.

poor rural households, i.e. low-skilled labour), this negative effect on non-farm growth may dampen that of agricultural productivity growth on rural poverty. This effect may have been particularly strong in the post-liberalisation period (i.e. post-1991), when labour was freer to move in search for lower-wage locations (see Aghion et al., 2007). Incidentally, the same argument may also help explain the adverse impact of literacy on rural poverty in the nineties. Since literate labour has a higher reservation wage than illiterate labour, a high share of literate labour may have acted as a restraint to investments by the non-farm tradable sector.

We already mentioned that to the extent that rural-urban migration occurs across districts, the identification strategy may not enable us to properly capture the channels linking urbanisation to rural poverty. In order to control for this, we need to construct a variable that measures the weight of rural-urban intra-district migration in the total rural emigrant population. By connecting this variable to the urban population, we may control for the fact that the effects of urbanisation on rural poverty are better identified in those districts with a relatively higher share of internal rural-urban migration in total rural emigrants. However, the data available does not allow us to compute such a share; we instead compute a rough approximation of this measure by dividing intra-district rural-urban migration by rural population. Including the interaction between this variable and the urban population leaves the results unaffected (column 6) with the interaction term bearing an expected but insignificant negative coefficient. We also use a different variable, i.e. the ratio of intra-district rural-urban migrants over the urban immigrants from other districts, obtaining similar (negative and non significant) results (not shown here). The lack of significance of these results may be due to the imprecise measure of the importance of intra-district migration.

We also test for the importance of the backward linkage effects of urbanisation on poverty. Considering that urban agricultural demand affects the district's rural sector more intensely in less spatially integrated markets, we need information on the share of urban demand of perishable products in total urban demand. Since we do not have this information, we instead compute a rough approximation based on agricultural data: the share of land cultivated fruits and vegetables (proxy for perishable goods) in total land cultivated. This measure relies on a number of assumptions, i.e. that a district's supply is a good proxy for urban demand and that fruits and vegetables are the main perishable agricultural goods. The interaction term between this share and the urban population variable has an expected negative coefficient (i.e. the higher the share the more poverty-reducing the urbanisation impact) – column 7. Again, this is not significant probably due to the imprecision of the measure. Also, including this interaction term reduces the explanatory power and the significance of the urbanisation variable. This may be due to the high collinearity between the two variables generated by the small variation of the fruit and vegetable share over time.

Table 4.5: The effects of urbanisation on rural poverty in Indian districts, further robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1981-99		1981-93		1981-99		
	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)
Urban pop. (millions)	-0.0684**	-0.0678**	-0.153**	-0.158**	-0.074***	-0.075***	-0.065*
	(0.027)	(0.026)	(0.063)	(0.065)	(0.027)	(0.027)	(0.039)
Rur. pop. (millions)	-0.0137	-0.00989	-0.0131	0.00411	-0.00946	-0.00992	-0.00392
	(0.019)	(0.018)	(0.026)	(0.024)	(0.019)	(0.019)	(0.021)
Scheduled caste (share)	0.486	0.488	0.738	0.701	0.555	0.540	0.625*
	(0.34)	(0.34)	(0.56)	(0.57)	(0.34)	(0.35)	(0.35)
Rural pop 15-34 age (share)	-4.628***	-4.690***	-5.445***	-5.716***	-5.024***	-5.039***	-4.764***
	(0.97)	(0.99)	(1.47)	(1.54)	(0.98)	(0.98)	(1.02)
Rural lit. 15_34 age (% in 15-34)	-0.0896	-0.0969	-1.067***	-1.004***	-0.135	-0.133	-0.0631
	(0.21)	(0.21)	(0.28)	(0.28)	(0.21)	(0.21)	(0.22)
Rural literates 15_34 x Post-1993	0.215***	0.218***			0.231***	0.233***	0.227***
	(0.074)	(0.074)			(0.075)	(0.075)	(0.077)
Urban poverty (share)	0.327***	0.328***	0.355***	0.380***	0.329***	0.331***	0.371***
	(0.074)	(0.073)	(0.12)	(0.11)	(0.072)	(0.072)	(0.072)
Ln Agricultural productivity	-0.0167		-0.0613**		-0.0274	-0.0268	-0.0260
	(0.020)		(0.030)		(0.020)	(0.020)	(0.021)
Ln Agr. prod. x Post-1993					0.0429***	0.0431***	0.0397**
					(0.015)	(0.015)	(0.016)
Share Internal migrants						-0.285	
						(0.58)	
Urban pop x Share fruits and vegetables							0.0201
							(0.13)
Observations	753	753	519	519	753	753	707
Nr. of districts	275	275	275	275	275	275	253
R-sq. (within)	0.67	0.67	0.65	0.65	0.67	0.67	0.64

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; all explanatory variables are lagged two years except for Agricultural Productivity (1 year lag) and urban poverty (contemporaneous).*

Given that limiting the spatial extent of the effect of urbanization within the border of single districts may be questionable, we run the same specifications of tables 4.3-4.5 adding a spatially lagged urbanization variable, i.e., the average of the urban population of the contiguous districts.[98] We also try to include the spatial lag of total population. These variables however were never significant, while other coefficients were only minimally affected (Table 4.6, first column).

Table 4.6: The effects of urbanisation on rural poverty across Indian districts, further robustness

	(1)	(2)	(3)	(4)
Sample	All	Cities >20k	All	Cities >20k
	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)	Rural pov. (share)
Urban pop. (millions)	-0.0496**	-0.0365	-0.108***	-0.112***
	(0.0222)	(0.0231)	(0.0377)	(0.0408)
Urban pop. of bordering districts (millions)	1.67e-07		2.96e-07	
	(5.79e-07)		(5.97e-07)	
Rural pop. (millions)	-0.0155	-0.00851	-0.0132	-0.00348
	(0.0146)	(0.0145)	(0.0146)	(0.0146)
Urban poverty (share)	0.326***	0.326***	0.322***	0.323***
	(0.0637)	(0.0626)	(0.0629)	(0.0621)
Scheduled caste (share)	0.474	0.372	0.483	0.325
	(0.301)	(0.293)	(0.301)	(0.298)
Rural literates 15_34 age (share in 15-34)	-3.329***	-3.181***	-3.262***	-3.039***
	(0.769)	(0.820)	(0.739)	(0.787)
Rural literates 15_34 age (share in 15-34)	0.0253	-0.118	0.00369	-0.147
	(0.162)	(0.162)	(0.159)	(0.160)
Observations	953	952	914	901
R-squared (within)	0.678	0.682		
Number of districts	343	354	306	305
Method	OLS	OLS	IV	IV

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; urban population is instrumented through the number of urban immigrants from other states.*

Finally, a further bias may be due to small villages upgrading to towns in the census definition. To the extent that these growing villages are systematically located in rural areas where poverty is decreasing (increasing) for reasons independent of urbanisation, we may detect a negative (positive) effect of urban population on poverty share which would be spurious. We therefore re-estimate the models excluding from the

[98] Technically, the variable is equal to Wx, where W is a row-standardized queen contiguity matrix, and x is the vector of urban population of districts.

urban population variable towns with less than 20,000 inhabitants – i.e., the size category which would contain most of the 'upgraded villages'. Results of this regression are extremely similar, although slightly less precise (see Table 4.6, second column). In the last two columns in Table 4.6 we run the same regressions as in the first two but employing IV estimation (using the number of migrants from other states to the urban areas of the district as an instrument). Again neither the spatial lagged variable nor the 'small villages' issue seem to affect the main IV results either (as discussed below).

4.5.4. IV estimation

Although the results are neat, we still need to control for the direction of causality in the relationship between urbanisation and rural poverty. As rural poverty declines, the rural-urban migration rate and thus urbanization may slow down as well and vice-versa. This would provide a source of (downward) bias in the coefficient. Without properly controlling for this potential endogeneity, the coefficient of equation (5) may have a downward biased, which means the estimates in Table 1 may be lower in absolute value than the real ones.[99]

We employ IV estimation to deal with this problem, using the number of migrants from other states to the urban areas of the district as an instrument. The first stage regressions, reported in different specifications in Table 4.7, substantiate the strong correlation of the instrument with the instrumented variable, and F- statistics are well above the confidence threshold of Stock and Yogo (2005) test for weak instruments (Table 4.8-4.9, last row). In analogy with OLS, standard errors in the IV estimations are robust and allow for intra-group correlation at district level.[100]

Results from the second stage regressions confirm the suspect of a downward bias of the OLS parameters, with new estimates being roughly twice as large as the OLS estimation for the period 1981-1999 (Table 4.8). This in turn implies a fairly substantial

[99] This is subject to the caveats that the sign of the bias in a multivariate regression depends also on the correlation with other regressors; and that the direction of the reverse causality may also be the opposite if poverty is a constraint to migration rather than a push factor.

[100] In order to get the covariance matrix of orthogonality conditions of full rank, which in turn allows to calculate clustered s.e., year-state dummies are "partialled out" and their coefficient is not calculated. By the Frisch-Waugh-Lovell theorem, in IV the coefficients for the remaining regressors are the same as those that would be obtained if the variables were not partialled out (Baum et al, 2008).

impact of urbanisation on rural poverty, with the rural poor decreasing by between 2% and 3% of districts' rural populations as the effect of an increase by 200,000 in urban residents (columns 1-3).

Table 4.7: The effects of urbanization on rural poverty across Indian districts, 1983-1999, IV Estimation, first stage

	(1)	(2)	(3)	(4)	(5)	(6)
		1981-99			**1981-93**	
	Urban pop.	Urban pop.	Urban pop.	Urban pop.	Urban pop.	Urban pop.
Urb. migrants from other states	4.248***	4.177***	4.177***	4.095***	3.814***	3.814***
	(0.82)	(0.91)	(0.91)	(0.72)	(0.72)	(0.72)
Rural pop.	0.0558	0.0891**	0.0891**	0.0320	0.0598*	0.0598*
	(0.035)	(0.035)	(0.035)	(0.031)	(0.030)	(0.030)
Scheduled caste (share)	-129121	5459	5459	470989	869828	869828
	(566798)	(742202)	(742202)	(518338)	(786918)	(786918)
Rural pop 15-34 age (share)	3838	-59321	-59321	-1291961	-1125466	-1125466
	(1183220)	(1316907)	(1316907)	(1152661)	(1290325)	(1290325)
Rural literates 15-34	-439969**	-354376	-354376	-141172	-94391	-94391
	(200881)	(241649)	(241649)	(176400)	(237122)	(237122)
Rural literates 15-34 x Post-1993	178712**	195241**	195241**			
	(70735)	(80831)	(80831)			
Urban poverty (share)		-88634*	-88634*		29978	29978
		(52376)	(52376)		(94211)	(94211)
Ln Agr. productivity		64776	64776		53597	53597
		(77841)	(77841)		(72036)	(72036)
Ln Agr. prod. x Post-1993		806.3	806.3			
		(32739)	(32739)			
Observations	996	779	779	682	520	520
Number of districts	363	280	280	363	275	275
R-squared	0.72	0.74	0.74	0.67	0.68	0.68
F-stat	61.83	30.97	30.97	15.17	40.16	40.16

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%.*

The IV analysis confirms the small first-round relative to second-round effects of urbanisation on rural poverty (column 1 to 2). Again, the results are robust when agricultural productivity variables are included as a control (column 3). We also run the IV estimation using the poverty gap as the dependent variable (column 4). The change in the magnitude of the urban population coefficient compared to the OLS specification in Table 4.3 is even bigger, and it maintains its significance. Again, when the share of rural poor is included as a control, the coefficient of urban population becomes insignificant (column 5). This confirms that urbanisation does not have an independent effect on the poverty gap, and thus on the severity of poverty, other than through the effect induced by the decrease in the share of poor in the rural population. The increase in magnitude of the coefficient is confirmed even when using the absolute number of rural poor as a dependent variable (column 6), although in this case the coefficient is only 1.5 larger than in the OLS (cf. Table 4.3, column 7).

Table 4.8: The effects of urbanization on rural poverty across Indian districts, 1981-1999, IV Estimation

	(1) Rural pov. (share)	(2) Rural pov. (share)	(3) Rural pov. (share)	(4) Poverty gap	(5) Poverty gap	(6) Rural poor (millions)
Urban pop. (millions)	-0.112***	-0.117***	-0.139***	-0.0393***	0.00105	-0.1624**
	(0.033)	(0.034)	(0.031)	(0.012)	(0.0052)	(0.0656)
Rural pop. (millions)	-0.00770	-0.00427	0.000761	0.00204	0.00182	1.4754***
	(0.016)	(0.015)	(0.017)	(0.0059)	(0.0025)	(0.2544)
Scheduled caste (share)	0.0646	0.292	0.406	-0.0281	-0.146**	1.0945*
	(0.27)	(0.30)	(0.32)	(0.12)	(0.058)	(0.6497)
Rural pop 15-34 age (share)	-3.845***	-4.057***	-4.808***	-1.573***	-0.172	
	(0.79)	(0.79)	(0.88)	(0.30)	(0.14)	
Rural literates 15_34 age (share in 15-34)	-0.139	-0.153	-0.263	-0.0702	0.00652	
	(0.17)	(0.16)	(0.20)	(0.069)	(0.027)	
Rural literates 15_34 x Post-1993	0.249***	0.230***	0.281***	0.105***	0.0231**	
	(0.067)	(0.064)	(0.070)	(0.022)	(0.0094)	
Urban poverty (share)		0.323***	0.338***	0.116***	0.0173*	0.4001***
		(0.061)	(0.067)	(0.023)	(0.0095)	(0.1124)
Ln Agr. productivity			-0.128	-0.0236	0.0136	-0.4115***
			(0.078)	(0.021)	(0.012)	(0.1462)
Ln Agr. prod. x Post-1993			0.165***	0.0482***	-0.00002	0.3731***
			(0.062)	(0.017)	(0.0085)	(0.1137)
Rural poverty (share)					0.291***	
					(0.0090)	
Rural pop 15-34 age (millions)						-0.3695***
						(0.1036)
Rural literates 15_34 age (millions)						-0.3068***
						(0.1019)
Rural lit. 15_34 age (millions) x post-93						9.128***
						(2.925)
Observations	950	914	753	753	753	753
Number of districts	319	306	255	255	255	255
R-squared	0.04	0.11	0.13	0.14	0.82	0.31
Kleibergen-Paark Wald F statistic	27.089	26.068	21.018	21.018	20.861	20.849

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; urban population is instrumented through the number of urban immigrants from other states.*

We also run the same regressions for the period 1981-93, obtaining similar results (Table 4.9). The coefficient of urban population is magnified by a factor of between 3 and 5 relative to its OLS value (cf. Table 4.4, columns 1-3), although it is estimated fairly imprecisely in the specifications with few control variables (column 1 and 2). The same increase in size is also true for the specification using the poverty gap as a dependent variable (column 4). However the inclusion of the share of rural poor as a control eliminates any effect of the urban population on poverty gap (column 5). This is also the case for the estimation run with the number of rural poor as a dependent variable: the increase of the urban coefficient is four-fold. The robustness checks examined in the previous section, including the spatially lagged variable and the population of towns with more than 20,000 inhabitants, do not affect our results when applied to the IV setting (table 4.6, columns 3 and 4).

Table 4.9: The effects of urbanization on rural poverty across Indian districts, 1981-1993, IV Estimation

	(1) Rural pov. (share)	(2) Rural pov. (share)	(3) Rural pov. (share)	(4) Poverty gap	(5) Poverty gap	(6) Rural poor (millions)
Urban pop. (millions)	-0.268	-0.315	-0.506**	-0.147**	0.00143	-0.8431**
	(0.20)	(0.20)	(0.21)	(0.058)	(0.0157)	(0.3710)
Rural pop. (millions)	0.00030	0.00111	0.00248	0.00047	-0.00256	0.7259
	(0.0024)	(0.0024)	(0.0031)	(0.0011)	(0.00432)	(0.5135)
Scheduled caste (share)	0.174	0.556	0.877	0.133	-0.125	1.1043
	(0.44)	(0.56)	(0.67)	(0.26)	(0.122)	(1.3225)
Rural pop 15-34 age (share)	-4.754***	-5.535***	-5.628***	-1.889***	-0.237	
	(1.31)	(1.41)	(1.52)	(0.55)	(0.259)	
Rural literates 15_34 age (share in 15-34)	-0.738***	-0.867***	-1.073***	-0.257**	0.0582	
	(0.21)	(0.25)	(0.29)	(0.10)	(0.0460)	
Urban poverty (share)		0.390***	0.400***	0.164***	0.0465*	0.5081**
		(0.11)	(0.12)	(0.046)	(0.0260)	(0.2006)
Ln Agr. productivity			-0.0984	-0.0163	0.0126	-0.3782**
			(0.078)	(0.024)	(0.0134)	(0.1494)
Rural poverty (share)					0.294***	
					(0.0132)	
Rural pop 15-34 age (millions)						0.7514
						(2.418)
Rural literates 15_34 age (millions)						-0.497***
						(0.1141)
Observations	636	608	488	488	488	488
Number of districts	318	304	244	244	244	244
R-squared	0.06	0.10	0.04	0.03	0.823	0.306
Kleibergen-Paark Wald F statistic	31.941	32.260	27.910	27.910	20.861	20.939

*All specifications include district and state-year fixed effects. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%; urban population is instrumented through the number of urban immigrants from other states.*

Finally, the substantial downward bias of the OLS estimates implied by the IV results suggests that an increase in poverty may be an important push factor for rural-urban migration. This could indicate that the poverty incidence is higher among migrants than among non-migrants (thus $\sigma > \gamma$). At the same time, our results suggest that first-round effects are quite small, i.e. condition (3) $[\sigma > \gamma + (\varphi - \lambda)]$ does not hold in its strong form. This would imply, consistently with the discussion in section 2, that the poverty incidence is lower in rural areas that are about to become urban than in the other rural areas (thus $\lambda < \varphi$), and interestingly this difference is similar to that of poverty rates between rural-urban migrants and rural non-migrants, i.e. $[(\sigma - \gamma) \approx (\varphi - \lambda)]$. Obviously the evidence provided here is not strong enough to make this more than an interesting speculation. And further research would be necessary to provide more direct empirical testing of such a hypothesis.

4.6. Conclusions

Do the poor in rural areas benefit from population growth of urban areas? And if so, what is the size of the benefits? Answers to these questions could help clarify whether trade-offs exist between urban investment and rural poverty and may help shed new light on the old debate on urban bias in developing countries. Notwithstanding the importance of these questions, little empirical evidence is available to provide adequate answers. We have tried to address this gap, by analysing the effects of urbanization on rural poverty. Using data on Indian districts between 1981 and 1999, we find that urbanization has a significantly poverty reducing effect on surrounding rural areas. Results are robust to the inclusion of a number of controls and to the use of different types of specification. The findings suggest that most of the poverty reducing impact of urbanization occurs through second-round effects rather than through the direct movement of rural poor to urban areas. We resort to IV estimation to test for causality. The results suggest that the effect is causal (from urbanisation to poverty reduction), and that failure to control for causality bias the coefficient of urbanisation downwardly. In our preferred estimations, we find that an increase of urban population by one fifth determines a decrease of between 3 and 6 percentage points in the share of rural poverty. These poverty reducing effects appear to apply mostly to rural poor relatively closer to the poverty line. Although the very poor do not seem to be negatively affected by urbanization, they are not able to reap the benefits of such a growth.

These findings may have a number of potentially important policy implications. First, they may help re-assess the role of public investment in urban areas for poverty reduction. In fact it is a popular tenet that investments in developing countries need to be concentrated in rural areas in order to reduce poverty, as the poor in developing countries are mainly concentrated there (see for instance World Bank, 2008). However, investments in rural areas are often very onerous as substantial resources are needed to reach a population which is scattered around vast territories. To the extent that urbanization may have substantial poverty reducing effects on rural areas, urban investments may become an important complement to rural ones in poverty reduction strategies.

Second, our findings run counter to the popular myth that rural-urban migration may deplete rural areas causing them to fall further behind. The relatively low rate of urbanisation of India itself may also be due to public policies which have not facilitated (and in certain instance even constrained) rural-urban migration (Deshingkar and Start, 2005). At the very least, this chapter questions the appropriateness of this bias against rural-urban migration.

Third, to the extent that the benefits from urbanisation do not spill over to the very poor in rural areas as highlighted by the findings in this chapter, specific actions may be needed to facilitate these rural dwellers to enjoy the benefits of urbanisation. Examples of these may include developing the types of skills useful for an expanding urban sector; or the provision of capital to cover the fixed costs of rural-urban migration.

Although this chapter has not touched upon the issue of urban poverty, rising urban populations may imply that urban poverty could become in the future the main issue in its own right (Ravallion et al., 2007). Further research is needed to assess whether the growth of urban population entails a trade-off between rural and urban poverty reduction.

Chapter 5. Returns to education in Uganda in the nineties: national and regional level analysis

5.1. Introduction

The process of economic integration at the global level has been accompanied by increasing economic inequality both in developed and developing countries. A substantial part of this rise in inequality in developing countries has been accounted for by increasing inter-regional disparities (Kanbur and Venables, 2005b). This stylised fact suggests that regions (and sub-national units more generally) within a country respond differently to changes in market forces. This may be determined by large differences in endowments across space within a country, such as location, factors' composition, but also to differences in local policies and institutions. Given these spatial differences, the recent upsurge of interest in the study of spatial inequality in developing countries is not surprising (e.g. special issue of Journal of Economic Geography, 2005; a number of edited volumes by Kanbur and Venables, 2005). While this focus is important and may explain a substantial part of the recent rise in inequality, it fails to analyse the largest part of economic inequality in developing countries, which appears to be within regions rather than between them. Studies in a number of developing economies confirm that around two third of such inequality is within regions (Demombynes et al., 2003; Yemtsov, 2003, Elbers et al., 2005).

This chapter and the following one represent one of the first attempts to examine the determinants of this within-region inequality using an important measure of labour market inequality, i.e. returns to schooling. Given that access to education and education achievements are unequally distributed across the population (particularly in developing countries), and that education is one of the major determinants of wage earnings, returns to schooling capture a substantial part of a country's economic inequality. In a context of limited geographical labour mobility, such as Uganda in the 1990s, this measure of inequality tends to vary significantly across space. I use this spatial variation to study the determinants of labour market inequalities in a LDC like Uganda.

I apply the analysis to the case of Uganda during the nineties, using different spatial scales as units of analysis (i.e. national and regional level in this chapter and district level in chapter 6). This is a particularly significant context for such a study. Uganda is a LDC with a high poverty incidence, which receives large amounts of aid flows (between 10 and 20% of GDP over the period considered). Investment in education has been identified as a key policy lever to raise the rate of growth and poverty reduction in the country both by the government and by the donor community.[101] However a popular view holds that returns to schooling in poor countries may well fall as educational supply grows unmatched by a proportionate increase in demand. Bennell (2002) argues that in the eighties and nineties the probability of obtaining a formal wage employment has substantially declined in Sub-Saharan Africa. This would imply that returns to education may fall among those who do have a job. The district-level analysis of next chapter helps shed some light on this issue by examining the effects of a number of localised supply and demand-related factors on returns to education, including the effects of an expansion in the supply of education.

Moreover, Uganda undertook profound pro-market reforms during the period of analysis, moving from a fairly closed and planned economy to a liberal open economy in the span of a decade. Interestingly, market reforms and trade liberalisation have been associated with an expanding income wedge between skilled and unskilled workers, as documented by Appleton (2001) and confirmed by the findings in this chapter. The methodology employed in chapter 6 allows testing more explicitly than in previous studies for the effects of these policy changes on returns to education by exploiting the heterogeneous effects across districts. The remainder of this chapter is organised as follows: the next section links the study to the literature; section 5.3 discusses the data and the methodology at different spatial scales; sections 5.4 and 5.5 present the national and regional level results. Chapter 6 then employs the methodology described here to compute returns to education at the district level and tests for the effects of different demand and supply-related factors on such returns.

[101] One of the major steps in this respect has been the introduction of Universal Primary Education (UPE) in 1996.

5.2. Related literature

The analysis in this chapter is directly related to the vast literature initiated by Mincer (1974), which uses micro-level data to compute wage effects, i.e. the change in income earned by a labourer through an additional year of education (all else constant). This wage effect is often referred to as returns to education in the literature. However, the computation of actual (private) returns would require discounting the wage effects by the (private) cost of education. I compute wage effects rather than returns to education proper (for reasons explained below), but for simplicity (and consistent with the literature) I use the two terms interchangeably henceforth. Only a small subset of this literature has focused on Sub-Saharan Africa (SSA), and a handful of studies have provided evidence on SSA for relatively long time spans encompassing the nineties. There is no clear evidence that returns to education have evolved homogeneously in the region in a time of pro-market reforms. Evidence from Ghana between 1987 and 1992 (Canagarajah and Thomas, 1997), from Tanzania (Soderbom et al., 2006) and Uganda (Appleton, 2001) in the nineties suggest the presence of rising returns to education. On the other hand, Krishnan et al. (1998) find that returns in urban Ethiopia remained stable despite labour market reforms in the early 1990s; and Soderbom et al. (2006) find high but falling returns in the Kenyan manufacturing sector over the nineties. Soderbom et al. argue that economic policies are key to account for the differences in levels and dynamics of returns to education across countries. The high returns in Kenya relative to Tanzania at the beginning of the 1990s reflected a policy environment in Kenya more favourable to the liberal operations of market forces than in Tanzania. A decade of reforms in Tanzania brought the country's polices more in line with those of Kenya. As a likely result the earnings profiles (at least in the manufacturing sector) were quite similar in the two countries at the end of the 1990s.

The importance of market reforms in determining changes in labour market inequality has been underscored also by studies on transition countries moving from centrally planned to market oriented economy. A number of studies mainly on China and Eastern European countries show that the freer operation of market forces tends to increase the demand for skills expanding the wedge between skilled and unskilled labour (see Fleisher et al. (2005) for a review; Flabbi et al. (2008) for Eastern European countries; Fleisher and Wang (2005), and Zhang et al. (2005) for China).

This literature has devoted comparatively little attention to the spatial dimension of inequality. Most studies have a national focus and only a few have looked at differences across regions within countries. This is somewhat surprising given the (often large) variation in labour market characteristics across sub-national units. Duranton and Monastiriotis (2002) find that the different evolution of returns to education between regions explain an important fraction of the large and increasing North-South divide in the UK between 1982 and 1997. For example average earnings increased in London relatively more than in the North as returns to education in London grew faster (catching up with those in the other regions) and the share of educated labour increased relatively more quickly in London than elsewhere. However Dickey (2007) shows that most of the increase in inequality in the UK over the last two decades is determined by intra-region rather than inter-regional inequality. This is in line with what studies in Kanbur and Venables (2005a) indicate for developing countries and supports the importance of investigating the determinants of such inequality. Notwithstanding this, the majority of the regional studies are descriptive rather than analytic in nature. Examples of these types include Tokila and Tervo (2007) on Finland and de la Fuente et al. (2003) on Spain.

Taylor (2006) stands out as an exception in this respect as he analyses the determinants of within-regions inequalities in the UK. He uses a measure of inequality based on the residual variance from regional earning equations and finds that trade and technology intensity are the most relevant factors to explain within region inequality in the eighties and nineties.[102] These results are consistent with both the theoretical predictions of the Heckscher-Ohlin (H-O) model for a skilled labour-abundant Northern country and the expected effects of skill-biased technological change (SBTC). He also finds that different education premia across regions are persistent over time. Such variation in inequality across space is likely to be associated with imperfect geographical mobility of labour. In fact in a context of perfect labour mobility, skilled labour would tend to move towards regions where the earnings' skill premium is higher, and unskilled labour would do the opposite, thus equalising labour market inequalities across regions. Limited labour mobility across regions seems to characterise most

[102] Other factors, such as changes in labour market institutions (i.e. level of unionisation) and female labour participation, seem to be less important in explaining inequalities in Taylor's analysis.

countries in the world. For example in a typical year less than 2% of the work force change region in the UK (McCormick, 1997). Limited geographical mobility is usually even more pronounced in developing countries' labour markets, and Uganda is no exception. According to the data used here on average only around 6% of the district's labour force has changed districts between 1992 and 2000.

In light of such limited labour mobility, the paucity of studies on within-regions inequality in developing countries is even more striking. To the best of my knowledge, only one study, by Yang (2005) on China, has analysed within-region labour market inequalities in a developing country context. He examines the determinants of returns to education across urban areas in China before and during the transition from planned to a market economy. His results reveal a large variability across cities, which is partly explained by skilled labour demand and supply related factors, such as the size of the technological and of the public sector, the presence of foreign firms, the level of infrastructure provision. Similarly to Yang, I use a micro approach to analyse how returns to education in Uganda have evolved nationally, regionally (in this chapter) and at district level (in chapter 6) during the 1990s. This is the first analysis of this type for an LDC and in SSA, and should also allow testing for some of the channels mentioned above through which market reforms may have influenced labour market inequalities.

5.3. Data and methods

5.3.1 Data

The analysis is based on data from two Household Surveys carried out in 1992 and 1999/2000 through the World Bank Living Standards Measurement Survey (LSMS). These contain data on about 10,000 households for 1992 and 1999/2000 (2000 henceforth). In particular, after excluding pensioners and students, data is available on about 3,600 and 2,600 wage earners respectively. In cases when more than one of these individuals belong to the same households, their labour market participation may potentially affect the decision of other members of the family to enter the labour market. I control for this by clustering the results at the household level. Data is also available for about 5,600 and 5,700 crop farming enterprises in the two periods. These enterprises are run by self-employed labourers who represent most of the income earning population in Uganda. Despite their importance, I cannot include them in the analysis due to missing data on land used in the production for 1999/2000. The exclusion of land generates a severe bias in the estimation of the production function for self-employed farmers, thus yielding inconsistent estimates for returns to education.[103] Notwithstanding the relevance of self-employment for the dynamics of the Ugandan economy, the analysis of wage employees is important in its own right. First, returns to education in developing countries are usually more sensitive to changing conditions for wage employees than for self-employed, who mainly operate in the informal sector. Second, to the extent that economic growth in developing countries is almost exclusively driven by formal firms (La Porta and Schleifer, 2008), concentrating on the dynamics of the formal sector (most of which is composed by wage employees) is particularly relevant.

Due to the insecure situation in the region bordering with Sudan, the data in 2000 is not available in four districts which were included in the 1992 survey: Kitgum and Gulu (Northern region), Bundibugyo and Kasese (Western region). I exclude these districts from the main analyses in the 1992 sample for comparability reasons. Results including these districts are reported in Appendix 5.1 (Table A5.1), showing very little difference with those regressions run without them.

5.3.2 Empirical specifications

The empirical strategy consists of a two stage analysis similar in spirit to an increasingly large empirical literature (see for instance Guiso et al., 2004; Mattoo et al., 2008), which uses estimated coefficients from first stage regressions as dependent variables in a second stage analysis. I first estimate standard wage regressions (both extended and basic) for employees at national, regional and district levels separately for 1992 and 1999/2000; then I use returns to education from the district-level regressions as dependent variables in the second stage. Here I describe the first stage, which is the focus of this chapter, while the second stage is described in chapter 6. As mentioned above, the analysis does not consider returns to education proper, but rather estimate wage effects. This decision is motivated by two main reasons. First, estimates for

[103] I tried to estimate the production function for crop farming enterprises without land for the 2000 survey obtaining implausible high values for capital and labour relative to the values in 1992. Results are available upon request.

Uganda already exist of both private and social returns to education, as calculated by Appleton (2001) using the same dataset with a full estimation method; second, the main focus of these chapters is to compare wage effects across districts (and over time) within Uganda. Both the direct private cost and the opportunity cost of education are less likely to vary across districts than across countries.

The basic version of the Mincerian wage equation can be written for each period of time as:

$$wage_i = \alpha + \beta S_i + \Gamma X_i + \varepsilon_i \qquad (5.1)$$

where *wage* is the (log of) nominal wage earned by the wage earner, *S* is the number of years of formal education and *X* is a vector of other covariates, which includes also the basic Mincerian controls, i.e. a dummy with the value of 1 if the employee is female, the number of years of experience ($Exp_i = Age_i - 6 - S_i$) and its squared term, and ε is the error term. I also add community-level fixed effects in some of the regressions to control for local labour market conditions. Communities represent the finest level of administrative classification in Uganda. The employees covered in the surveys belong to 820 and 796 communities in 1992 and 2000 respectively (with only a handful of communities being the same in both surveys).

The standard method in (5.1) assumes the same marginal effect across education categories. As it is often the case, the marginal effect may differ between levels of education. In order to allow for such heterogeneous effects, let us re-write (5.1) as an extended wage equation with the educational variable constructed as a spline function with N nodes at selected levels of education (see Moll, 1996):

$$wage_i = \alpha + \sum_{j=1}^{J} \beta_j S_{ij} + \Gamma X_i + \varepsilon_i \qquad (5.2)$$

where S_{ij} is the number of years attended at the *j-th* level. In particular, it seems natural to follow the three educational levels in Uganda to mark the nodes, i.e. primary, secondary and post-secondary. Thus the variable S_{ij} is constructed as follows:

$$S_{i1}=\begin{cases} S_i & if\ 0 \leq S_i \leq 7 \\ 7 & if\ S_i \leq 7 \end{cases} \quad S_{i2}=\begin{cases} 0 & if\ S_i \leq 7 \\ S_i - 7 & if\ 7 < S_i \leq 13 \\ 6 & if\ S_i > 13 \end{cases} \quad S_{i3}=\begin{cases} 0 & if\ S_i < 13 \\ S_i - 13 & if\ S_i \geq 13 \end{cases}$$

This national level analysis is close to that by Appleton (2001) in that it uses the same data and a similar methodology. However, there are a few substantial differences that are worth highlighting. First, this analysis uses more controls (included in the vector X_i) in regressions (5.1) and (5.2) than in Appleton; second, it performs further robustness checks both on the results for 1992 and on the increase in returns to education over time; third, it checks for the robustness of the results to some of the potential sources of bias in the returns to education coefficient. In particular, there are at least two potential sources of bias in this analysis: ability bias and sample selection bias. The former stems from the fact that the choice of education is likely to be determined by individual unobserved abilities which may also have a direct influence on income (i.e. ceteris paribus more able individual tend to be more successful at school and thus tend to pursue higher levels of education). The other potential bias comes from possible self-selection of wage earners. As argued by Soderstrom et al. (2006), labour market participation into the wage sector is an atypical outcome in many Sub-Saharan African countries. Thus wage employees are likely to possess higher than average ability whose effects may be partly captured by the coefficients of the observable wage determinants if it is not properly controlled for.[104] The typical ways to address these biases are through IV estimation (for the ability bias) and through the Heckman (1979) sample selection model. Subject to the availability of data, the analysis below tries to examine the extent to which these biases are problematic in this instance.

However the most important difference with Appleton (2001) is the spatial analysis (at the district and regional levels), which represents the focus of the next chapter. For each year I run separate Mincerian-type regressions at the regional as well as at the district level in order to compute the variation of returns to education across

[104] Soderstrom et al. (2006) use firm level data on employees, while I use household survey on wage earners. The potential self-selection bias is more severe for employees than it is for wage earners (which is a broader group); however, to the extent that the latter are also part of the wage sector, a potential self-selection bias may arise.

space.[105] I adopt the geographical division used by the Uganda Bureau of Statistics. This consists of four regions (Central, Eastern, Western and Northern) and 38 districts for 1992 and 45 districts in 1999/2000 (although data on only 41 of these 45 districts is available in the 2000 survey - see below).[106] The basic equation adopted is a variant of (5.1) to calculate returns to education at different spatial scales (i.e. regions and districts) in each year:

$$wage_i = \sum_{k=1}^{N} \alpha_k d_{ik} + \sum_{k=1}^{N} \beta_k d_{ik} Edu_i + \sum_{k=1}^{N} \mathrm{B}_k D_{ik} W_i + \Pi Z_i + \varepsilon_i \tag{5.3}$$

where N is the number of districts (regions), d_{ik} is a dummy that takes the value of 1 if individual i resides in district (region) k, and the value of zero otherwise; B_k is a column vector of coefficients associated to the interaction between the vector D_{ik} of d_{ik} dummies and that of the Mincerian controls W_i (i.e. gender dummy, experience and its squared term); and Z_i is a vector of the other covariates included in X in the previous equations. Unlike those in vector W, the effects of the covariates in Z are not allowed to vary across areas k. This allows saving degrees of freedom as the estimation in a number of districts relies only on relatively few observations (see discussion in the next chapter). This is also the main reason why I prefer to estimate the β coefficients through a pooled regression as in (5.3) rather than through regressions of the type of (5.1) run separately for each area. Using the latter method and including all the controls in (5.1) would reduce the degrees of freedom available, which may be problematic in those districts with fewer observations available. One possibility could be to run a parsimonious version of (5.1) by district including only the Mincerian controls.[107] This is not ideal however as it may lead to biased coefficients due to potential omitted variable. In any case in the next chapter I test the robustness of the district-level results to using coefficients estimated through this method as well.

I also use the extended method with the education variable constructed as in (5.2) to decompose regional returns to schooling into the different education levels:

[105] Regions are the first layer of administrative division in Uganda, while districts are the third layer (there are sub-regions in between).
[106] A number of districts have been split into two or more districts between 1992 and 1999.
[107] This option would produce the same estimated coefficients as running (5.3) without the controls in Z, although standard errors would be different.

$$wage_i = \sum_{k=1}^{N} \alpha_k d_{ik} + \sum_{k=1}^{N} \sum_{j=1}^{J} d_{ik} \beta_{kj} S_{ij} + \sum_{k=1}^{N} \mathrm{B}_k D_{ik} W_i + \Pi Z_i + \varepsilon_i \qquad (5.4)$$

Chapter 6 will use the β coefficients derived from (5.3), (5.4) and their variants as dependent variables in order to identify the determinants of returns to education across space and over time.

5.4. National level analysis

I first run extended Mincerian wage regressions for wage earners as in (5.2) using the log of annual earnings for 1992 as dependent variable, and then compare the results with those using data from the year 2000 (expressed in constant 1992 prices). The results from these regressions for 1992 are reported in Table 5.1 and reveal that the effects of primary and secondary education are quite similar (an increase of 10% and 9% for each extra year of education respectively). The effects of post-secondary education are twice as large (column 1). All the other variables are highly significant and with the expected sign: experience exerts a positive and diminishing effect and being a woman reduces the wage by 22%.

These results are fairly robust to the inclusion of community fixed effects, which reduces somewhat the coefficients of the education variables, especially primary and post-secondary education (column 2). This suggests that the positive impact of education on wages is partly explained by local conditions, such as labour market characteristics and amenities. The coefficients on the education variables are not strictly comparable to those calculated by Appleton (2001) as he uses age instead of experience and a dummy for university graduation instead of the number of years of post-secondary education. However, the effects of primary and secondary education when using community fixed effects are remarkably in line with those of Appleton (6.5% versus 7% for primary, and 8% in both for secondary), while the coefficient of post-secondary education is lower than in Appleton (13.5% vs. 18%). This difference is likely to be the result of the different definition of the variable. Interestingly, the inclusion of community fixed effects substantially reinforces the anti-female bias (while it has only a minor effect on the coefficient of experience). One possible interpretation of this result is that women may be more economically discriminated within their own community than across communities.

Adding further controls changes little the values of the Mincerian coefficients (columns 3 and 4), but it is worth briefly discussing the impact of these extra variables. Not being married (either as unmarried or in other marital forms, i.e. cohabiting, divorced or widowed) has a negative effect on wage. The causality of this relationship may be questionable, as higher income individuals may be more able to face the cost of marriage. Moreover the relationship may be biased by endogeneity if some unobservable individual characteristics, such as beauty or personality, drive both earnings and the likelihood to marry. However, the fact that even individuals involved in other marital forms (e.g. co-habiting, widowed) earn less than married individuals may lend some support to causality. As expected, urban location increases the wage earned (by 57%). To what extent this is a reflection of higher price levels in urban areas (data to calculate real wages are unavailable) or of higher productivity in cities due to agglomeration economies (with imperfect labour mobility) is matter left for future research. Being in the Central region increases the wage relative to the other regions, while being in the Northern region depresses wages. Eastern and Western regions have similar effects on income.

The results from the basic Mincerian regression (equation 5.1) indicate that the average effect of an extra year of education on employees' wages varies between 7.5% (including community fixed effects and marital controls, column 8) and 10.8% (without community effects and controls, column 5). The inclusion of a full set of controls without community effects (column 6) and the inclusion of community effects without controls (column 7) yield a similar effect of education on wage of around 8.5%. This range is in line with the average coefficient of years of schooling at the world level (10%), calculated by Psacharopoulos (1994), although it is slightly lower than the average for SSA of 13.4%. It is interesting to note that the coefficients of the other variables are almost unchanged relative to the corresponding specifications using the spline-type education variable.

Educational variables are likely to be endogenous to returns to education as unobserved individual ability may influence both the schooling outcome and the level of income (e.g. more able individuals are more likely to achieve higher qualification and earn higher incomes). This endogeneity is fairly well established in the literature and authors have used a number of different instruments (e.g. family background, changes in the schooling systems) to get around this problem.[108] Data availability constrains the possibility of using an adequate instrument in this case; however in line with previous literature (and unlike Appleton 2001, who does not use an instrument for education), I try to use both parents' education to instrument for years of education. The first stage results reported in Table A5.2 (columns 1 and 2) show that these are both powerful predictors of children's education with F-statistics of over 100 and 30 for the specification without and with community effects respectively.

The IV estimation has the effect of almost doubling the coefficient of education in both specifications (with and without community effects – see columns 9-10 respectively). The direction of this difference is consistent with other studies using this type of instruments, although the size of the difference is much larger than in other studies. As argued by Card (1999), the higher IV coefficient is likely to be driven by the influence of family background on the children's income, which suggests an upward bias in the coefficient of education estimated through this instrument. If strong enough such intergenerational links can invalidate the exclusion restrictions of the instruments. The main effects of parents' education on children's income in Uganda seem to be displayed via children's education, as shown by the results of regression (1) with the addition of parents' education as a control (see Table A5.2 in the Appendix). However, some of the effects especially for mother's education is independent of children's education, which may invalidate the exogeneity assumption of the instruments.[109] These results are reported for 1992 but they are similar for 2000 as well: the coefficients and significance of parents' education are substantially reduced after the inclusion of wage earner's education – see columns 4 vs. 3 and 6 vs. 5 (although mother's education remains significant in the specification without community fixed effects – column 4).

This discussion suggests that the large upward bias in the IV coefficient of education may be determined more by errors in measuring parents' education than by

[108] See Card (1999) for a review.

[109] As a matter of fact in the 2SLS regression without community fixed effects in Table 5.1 the Hansen J-statistics fails to reject the null that the instruments are correctly jointly excluded from the second stage at the 10% level of significance (see column 9).

the direct effects of parents' education on children's earnings.[110] An increasingly popular view holds that OLS estimates appear to be remarkably close to the actual returns calculated using samples of twins and family background to control for ability bias (Ashenfelter and Zimmerman, 1997; Card, 1999). Given this little overall bias of the OLS estimates and the substantial difference with IV estimation in the Ugandan data, I will base the bulk of the analysis on OLS rather than IV estimates. Since the eventual bias of the estimated β coefficients via OLS is not expected to differ systematically across districts, this should not be a concern in the district-level analysis.

The above results are based on annual wages from the main occupation, which do not take into account the actual time spent on the main occupation. Therefore the dependent variable may be an imperfect measure of labour productivity. Although the information on the time spent on the main occupation is not available in the 2000 survey, it is useful to check whether the above results for 1992 are robust to using daily and hourly earnings. For that I run the same regressions as in Table 1 using both log daily and log hourly wages as dependent variables.[111] The results, reported in Table A5.3 in Appendix 5.1, are remarkably close to those of Table 5.1, except for the coefficients of primary education, which are between 10% and 28% lower than the correspondent ones using annual income. This difference may be due to measurement error, as there may be a positive correlation between the level of education and the precision in reporting the exact time worked. Overall, these results suggest that the lack of data on time spent in the main activity is not likely to change the results in a significant way.

[110] This hypothesis is supported by the fact that part of the sample did not report information on parents' education and that this information was less precise than for own education (i.e. they reported only the highest education title attained).

[111] Hourly wage is computed as total income from the main occupation earned last year divided by total number of hours devoted to it.

Table 5.1: Determinants of wage for employees in Uganda, 1992 (34 districts)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) 2SLS	(10) 2SLS
Primary	0.099***	0.065***	0.068***	0.056***						
	(0.009)	(0.009)	(0.009)	(0.010)						
Secondary	0.090***	0.083***	0.080***	0.074***						
	(0.012)	(0.011)	(0.011)	(0.011)						
Post-sec	0.193***	0.135***	0.163***	0.132***						
	(0.027)	(0.027)	(0.026)	(0.028)						
Education					0.108***	0.083***	0.087***	0.075***	0.150***	0.121***
					(0.004)	(0.005)	(0.004)	(0.005)	(0.016)	(0.028)
Female	-0.242***	-0.324***	-0.275***	-0.296***	-0.246***	-0.322***	-0.277***	-0.295***	-0.320***	-0.277***
	(0.036)	(0.038)	(0.036)	(0.038)	(0.036)	(0.037)	(0.036)	(0.038)	(0.041)	(0.042)
Experience	0.042***	0.051***	0.034***	0.041***	0.042***	0.050***	0.034***	0.040***	0.051***	0.052***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.008)
Exp squared	-0.0007***	-0.0008***	-0.0006***	-0.0007***	-0.0007***	-0.0008***	-0.0005***	-0.0006***	-0.0007***	-0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Unmarried			-0.299***	-0.258***			-0.285***	-0.252***	-0.076	-0.116
			(0.038)	(0.044)			(0.038)	(0.044)	(0.062)	(0.086)
Other mar.			-0.220***	-0.174***			-0.213***	-0.168***	-0.129**	-0.109*
			(0.050)	(0.052)			(0.051)	(0.053)	(0.061)	(0.065)
Urban			0.556***				0.553***		0.433***	
			(0.031)				(0.031)		(0.042)	
Central			0.306***				0.315***		0.339***	
			(0.038)				(0.038)		(0.042)	
Western			-0.045				-0.035		0.018	
			(0.041)				(0.041)		(0.047)	
Northern			-0.242***				-0.237***		-0.211***	
			(0.054)				(0.055)		(0.058)	
Commun. FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Observations	3286	3286	3286	3286	3286	3286	3286	3286	3025	2861
R-squared	0.251	0.611	0.358	0.618	0.248	0.609	0.354	0.616	0.288	0.176
Nr communities		757		757		757		757		582
1st stage F-stat									100.48	30.44
Hansen J-stat									2.99	0.695

*Dependent variable is log wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%; instruments for education in columns 9 and 10 are father and mother's years of education. F-stat is the F-statistics from the first stage for the excluded instruments; Hansen J-statistic is the test for over-identification.*

In line with the results of Appleton (2001), I also find that the wage effects of education variables substantially increase over the nineties in Uganda. Table 5.2 presents the results of the same regressions as in Table 1 run using 1999/2000 survey data. The size of the proportionate increases diminishes monotonically with the level of education, with the coefficients of primary education experiencing the highest proportionate increases. These range between 62% (OLS with no controls) and 110% (OLS with controls). The increases in the secondary education coefficient range between 64% and 73%, while the effect of post-secondary education increases between 10% and 30%. Overall the education coefficient has increased over the nineties by between 52% and 69%. As in Appleton (2001), the change in coefficients is statistically significant both for primary and secondary education (and for total years of education), while it is not for post-secondary education.[112]

The breakdown of the marginal returns by educational level shows a convex shape in the returns to education (cfr. columns 1-4 in Tables 5.1 and 5.2). In both years the marginal return to post-secondary education is higher than the returns before the secondary level. This convexity is in line with findings on other Sub-Saharan African countries, e.g. Soderbom et al. (2006) for Kenya and Tanzania and Baptist and Teal (2008) for Ghana. However, unlike in Tanzania, the intensity of the convexity seems to have decreased somewhat during the nineties in Uganda. The ratio of post- to pre-secondary returns shrank from a factor of around 2 in 1992 to a factor of around 1.4 in 2000, when the level of returns was significantly lower than in Tanzania (where it exceeded 27%). That is essentially due to a large increase in returns to primary education, which is encouraging for strategies aiming at expanding primary education provision as the one currently pursued by Uganda via UPE. The extent to which increases in the supply of (primary) educated workforce may offset this upward trend of returns to primary schooling is an issue which will be analysed in the next chapter.

[112] I test this by pooling the data for 1992 and 1999/2000 and running the regressions from Table 5.1 with a full set of interactions between the year 2000 and the regressions' variables (results are reported in Table 5.5, columns 1 and 2).

Table 5.2: Determinants of wage for employees in Uganda, 1999/2000 (same districts as in Table 5.1)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) 2SLS	(10) 2SLS
Primary	0.160***	0.122***	0.143***	0.110***						
	(0.010)	(0.013)	(0.010)	(0.012)						
Secondary	0.154***	0.136***	0.135***	0.128***						
	(0.015)	(0.016)	(0.015)	(0.016)						
Post-sec	0.213***	0.176***	0.201***	0.171***						
	(0.034)	(0.039)	(0.034)	(0.040)						
Education					0.164***	0.136***	0.147***	0.127***	0.251***	0.224***
					(0.004)	(0.005)	(0.004)	(0.006)	(0.019)	(0.035)
Female	-0.100**	-0.184***	-0.160***	-0.151***	-0.101**	-0.186***	-0.162***	-0.154***	-0.176***	-0.128**
	(0.042)	(0.046)	(0.043)	(0.045)	(0.042)	(0.046)	(0.042)	(0.045)	(0.053)	(0.064)
Experience	0.055***	0.054***	0.047***	0.043***	0.055***	0.053***	0.047***	0.043***	0.071***	0.062***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.011)
Exp squared	-0.0008***	-0.0007***	-0.0007***	-0.0006***	-0.0008***	-0.0007***	-0.0007***	-0.0006***	-0.0008***	-0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Unmarried			-0.284***	-0.289***			-0.280***	-0.284***	0.067	-0.041
			(0.047)	(0.055)			(0.047)	(0.055)	(0.088)	(0.128)
Other mar.			-0.218***	-0.246***			-0.216***	-0.239***	0.119	0.064
			(0.053)	(0.061)			(0.053)	(0.061)	(0.085)	(0.131)
Urban			0.448***				0.448***		0.277***	
			(0.035)				(0.035)		(0.056)	
Central			0.103**				0.102**		0.147**	
			(0.044)				(0.044)		(0.059)	
Western			-0.082*				-0.081*		-0.024	
			(0.046)				(0.046)		(0.058)	
Northern			-0.173***				-0.176***		-0.083	
			(0.058)				(0.058)		(0.074)	
Commun. FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Observations	2608	2608	2606	2606	2608	2608	2606	2606	1836	1560
R-squared	0.419	0.669	0.467	0.676	0.418	0.668	0.466	0.675	0.378	0.225
Nr communities		796		796		796		796		426
1st stage F-stat									76.81	21.54
Hansen J-stat									2.760	1.103

*Dependent variable is log wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%; instruments for education in columns 9 and 10 are father and mother's years of education. F-stat is the F-statistics from the first stage for the excluded instruments; Hansen J-statistic is the test for over-identification.*

The effect of experience shows a mild increase, while that of marital variables is essentially unchanged. On the other hand, the anti-female bias is substantially reduced over the nineties. A closer inspection of the data suggests that this is mainly due to the drop in wage differentials in the public sector. In fact the results of separate public-private sector wage regressions indicate that the anti-female bias remained almost unchanged in the private sector (from 30.7% to 27.8%) while it decreased from 21.3% to 9.6% in the public sector (results available upon request). This drop in the public sector wage differential is in line with a more merit-based pay system introduced by the public sector reforms in the early nineties (Kagundu and Pavlova, 2007). Finally, location variables suggest that there may have been some spatial convergence over the nineties: the advantage of being in an urban area diminishes, and so does that of being in the Central region relative to the Eastern region and to some extent Western region. The disadvantage of being in the Western region relative to the other regions decreases.[113] These issues are explored in more depth in the regional analysis below.

5.4.1 Robustness checks

As described earlier, in an economy like Uganda employees in the wage sector may not constitute a representative sample of the income earning labour force. Such sample selection may lead to biased OLS estimates and thus bias the comparison of returns over time. The typical way to control for this potential bias is to use a sample selection model along the lines proposed by Heckman (1979). This consists of estimating a wage employment participation equation (i.e. the determinants of the probability of an individual engaged in gainful activities of being a wage employee); and then estimating the wage regression conditional on the estimated wage employment participation. However, due to the different ways in which the questionnaires have been constructed for 1992 and 2000, it is possible to properly identify the broader group of individuals engaged in gainful activities (of which wage employees are a sub-sample) only for 1992. The 2000 survey allows to identify only those individuals receiving income from employment, while the rest of income generating activities (e.g. from self-employment) are attributed to the entire household. This prevents a consistent identification of the larger group of individuals engaged in gainful activities in 2000,

[113] Note that among the location variables only the changes of the coefficients for urban and Central region are statistically significant at standard levels.

which is necessary for the Heckman model. Thus I can estimate the model only for 1992, which is a useful robustness test for the results above. The estimate for returns to education of the two-step Heckman model obtained through maximum likelihood methods is provided in Table 5.3. Mother's and father's education are used as instruments in the first stage both separately and individually to check whether their potential endogeneity may bias the results. To the extent that father's education has an insignificant direct effect on the individual wage (as shown above), the similarity of the results across the different specifications (cf. Table 5.3, columns 1, 3 and 5) is reassuring with respect to the endogeneity concern. The second stage estimates of returns to education are 10-12% larger than OLS estimate (cf. Table 5.1, column 5). This points towards the presence of a small downward bias in OLS estimates, which is consistent with the idea that education raises the wage also by raising the probability of accessing wage employment. The small size of the possible bias of the education coefficient is in line with findings from Kagundu and Pavlova (2007), who use the 2002/03 Ugandan household survey. This is also reassuring for the subsequent spatial analysis in Chapter 6, which because of lack of data needs to use standard OLS estimates of returns to education as dependent variables. Moreover, such a bias is likely to operate in a fairly similar way across districts, thus further attenuating the risk of biased coefficients in the district-level analysis.

Table 5.3: Returns to education from two-stage Heckman model, 1992 (34 districts)

	(1)	(2)	(3)	(4)	(5)	(6)
Instrument	Father's edu		Mother's edu		Both	
	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage
Education	0.120***	0.109***	0.117***	0.106***	0.116***	0.107***
	(0.008)	(0.004)	(0.006)	(0.004)	(0.007)	(0.004)
Female	-0.385***	-0.591***	-0.372***	-0.595***	-0.394***	-0.603***
	(0.055)	(0.029)	(0.050)	(0.029)	(0.053)	(0.029)
Experience	0.047***	-0.004***	0.049***	-0.005***	0.050***	-0.004***
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)
Exp squared	-0.0007***		-0.0008***		-0.0008***	
	(0.0001)		(0.0001)		(0.0001)	
Father's Education	0.101***	0.0497***			0.046***	0.0178
	(0.014)	(0.0114)			(0.016)	(0.0133)
Mother's Education			0.139***	0.0863***	0.117***	0.076***
			(0.015)	(0.0133)	(0.018)	(0.016)
Rho	0.316***		0.258***		0.303***	
	(0.117)		(0.092)		(0.104)	
Observations	12131		12245		11914	
Uncens. obs.	3084		3114		3025	

*Dependent variable is log wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%.*

The potential importance of the differential impact of education on wages between the public and the private sector as well as between the urban and the rural sector has been underscored above. In particular returns to education are usually lower in the public than in the private sector and in the rural than in the urban sectors. Given that the public-private as well as the urban-rural employment composition of the samples have changed between 1992 and 2000 it is worth examining to what extent these compositional changes have affected the changes in returns to education. As shown in Table 5.4, the share of private employees in total has increased in 2000 relative to 1992, while the opposite is true for public employees and other employees (e.g. religious and political workers) maintain roughly the same share. This change is likely to mirror the relative shrinking of the public sector over the period due to the privatisation and the public reform process discussed above. Public wages have increased on average more than private sector wages and this increase adds up to an already higher wage in 1992 (see Table 5.4). On the other hand, the share of urban employees has shrunk in 2000 possibly reflecting an increase in wage employment (relative to self-employment) in rural areas. Both rural and urban wages have grown substantially over the nineties, with the former experiencing a more marked rise.

Table 5.4: Summary statistics for ln(main wage), by employment type and location

Year	Category	Obs	Share	Mean	Std. dev	Min	Max
	Private	1877	57.1%	12.273	1.078	5.704	16.065
	Public	1372	41.8%	12.519	0.951	7.601	15.917
1992	Others	37	1.1%	12.218	1.114	10.309	14.982
	Urban	2113	64%	12.642	0.975	7.313	16.065
	Rural	1173	36%	11.895	0.963	5.704	15.573
	Private	1774	67.4%	12.673	1.154	8.762	16.545
	Public	825	31.3%	13.583	0.902	10.148	16.807
2000	Others	35	1.3%	13.013	1.448	9.133	16.034
	Urban	1369	52%	13.299	1.116	8.762	16.807
	Rural	1265	48%	12.600	1.107	8.762	15.852

Source: LSMS for 1992 and 1999/2000 (based on 34 districts in 1992)

In Tables 5.5 and 5.6 I analyse how returns to education have changed across those sectors during the nineties. This is done by running a wage regression as in (5.2) and (5.1) with pooled data for 1992-2000, adding a series of interaction terms between each variable and a post-1992 dummy. These interactions measure the change in the wage effect of each variable between 1992 and 2000. Table 5.5 reports the results using the extended Mincerian equation and including specifications with and without community fixed effects. The first two columns present the results for the entire sample confirming that returns have increased proportionately more at the primary than at the secondary level (although the absolute variation is the same in the specification with community effects, column 2). Returns to post-secondary schooling rose marginally but not significantly according to the standard levels. The result for the pooled sample confirms the evidence of a reduction in the overall convexity of returns to education in Uganda over the nineties.

Columns 3-6 show the results by educational category for public (columns 3 and 5) and private employees (columns 4 and 6) respectively. There is significant variation: returns in the public sector rose substantially at the primary and post-secondary level while they decreased at the secondary level (albeit this drop is not statistically significant). Those in the private sector steadily rose across educational categories but more modestly than in the public sector. Interestingly, returns to primary education became larger in the public than in the private sector over the nineties reversing the situation of 1992. This may be linked to the process of public sector reforms in the nineties mentioned above but more research is needed to shed light on the issue. The convexity of returns has substantially increased for public sector employees and less so for private employees.

Some variability in the changes as well as in the levels of returns to schooling emerges also between the urban and the rural sectors (columns 7-10). The wage effect of primary education was identical in 1992 when not including community effects (column 7 vs. 8). Their inclusion does not change the coefficient for urban areas (column 9) but substantially reduces it for rural areas (column 10), suggesting that labour market inequality is lower within rural communities than across them (although that is not the case for the wage effects of secondary schooling). The changes in returns are different across areas: returns grew more in urban than rural areas when including community fixed effects (columns 9-10), again in line with smaller wage premia within local rural areas. On the other hand returns to secondary education increased more in rural than urban areas in the specification without community effects (columns 7 and 8). This may suggest the emergence of a demand for a relatively qualified workforce in rural areas over the 1990s.

Table 5.5: Changes in education splines coefficients 1992-2000, splitting the samples

	(1) Pooled	(2) Pooled	(3) Public	(4) Private	(5) Public	(6) Private
Primary	0.068***	0.056***	0.061***	0.074***	0.049**	0.060***
	(0.009)	(0.010)	(0.016)	(0.011)	(0.023)	(0.012)
Secondary	0.080***	0.074***	0.088***	0.096***	0.081***	0.081***
	(0.011)	(0.011)	(0.016)	(0.017)	(0.018)	(0.016)
Post-sec	0.163***	0.132***	0.159***	0.215***	0.136***	0.125**
	(0.026)	(0.028)	(0.030)	(0.051)	(0.034)	(0.057)
Δ Primary	0.075***	0.054***	0.124***	0.041***	0.093*	0.027
	(0.014)	(0.016)	(0.029)	(0.016)	(0.054)	(0.018)
Δ Secondary	0.055***	0.053***	-0.040	0.031	-0.061	0.037
	(0.018)	(0.019)	(0.031)	(0.025)	(0.038)	(0.026)
Δ Post-Sec.	0.038	0.039	0.086	0.056	0.147**	0.085
	(0.042)	(0.048)	(0.059)	(0.070)	(0.070)	(0.082)
Comm FE	NO	YES	NO	NO	YES	YES
Observations	5892	5892	2191	3629	2191	3629
Nr comm..		1553			947	1247
R-squared	0.452	0.671	0.527	0.404	0.787	0.688

	(7) Urban	(8) Rural	(9) Urban	(10) Rural
Primary	0.065***	0.064***	0.063***	0.039**
	(0.011)	(0.014)	(0.012)	(0.017)
Secondary	0.084***	0.071***	0.067***	0.109***
	(0.013)	(0.021)	(0.012)	(0.022)
Post-secondary	0.172***	0.112*	0.153***	0.002
	(0.028)	(0.063)	(0.029)	(0.079)
Δ Primary	0.095***	0.061***	0.065***	0.051**
	(0.020)	(0.020)	(0.021)	(0.024)
Δ Secondary	0.027	0.096***	0.047**	0.041
	(0.023)	(0.032)	(0.022)	(0.038)
Δ Post-Sec.	0.088*	-0.010	0.041	0.127
	(0.049)	(0.088)	(0.055)	(0.107)
Comm FE	NO	NO	YES	YES
Observations	3467	2425	3467	2425
Nr comm..			571	982
R-squared	0.425	0.395	0.592	0.719

*Dependent variable is log wage from main activity. Sample for 1992 is based on 34 districts. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%.; all regressions are run with pooled data for 1992-2000 and include Mincerian controls as well as the other controls as in Tables 5.1 and 5.2 and a series of interaction terms between each variable and a post-1992 dummy. The Δ reports the coefficients of these interactions for the schooling variables.*

Performing the same analysis using the basic earning function with years of education variable (Table 5.6) indicates that the wage premium for an extra year of education in 2000 is between 5.2 and 5.8% larger than in 1992 (columns 2 and 1). This is similar to the changes in primary and secondary coefficients (columns 1-2, Table 5.5). The increase in returns to education is larger for private than for public sector employees (columns 3-6), especially when including community effects (column 5 vs.

6). In the latter case the increase in returns to education for the public sector is not significant, which is the result of rising primary returns and decreasing secondary returns (Table 5.5, column 5). Interestingly, the increase in the education coefficient for both private and public sector employees is lower than that in the pooled sample. This suggests that the changes in returns across sectors (e.g. the wage premium increase of educated private sector over uneducated public sector employees) have played some role in determining the overall increase in the educational coefficient. On the other hand urban and rural employees experienced a very similar increase in returns to education (columns 7-10) and in line with that for the pooled sample.

Table 5.6: Changes in education splines coefficients 1992-2000, splitting the samples

	(1) Pooled	(2) Pooled	(3) Public	(4) Private	(5) Public	(6) Private
Education	0.088***	0.075***	0.097***	0.097***	0.086***	0.075***
	(0.004)	(0.005)	(0.006)	(0.006)	(0.009)	(0.008)
Δ Education	0.058***	0.052***	0.026***	0.037***	0.020	0.038***
	(0.006)	(0.007)	(0.009)	(0.008)	(0.013)	(0.011)
Comm FE	NO	YES	NO	NO	YES	YES
Observations	5892	5892	2191	3629	2191	3629
Nr comm.		1553			947	1247
R-squared	0.449	0.669	0.521	0.399	0.783	0.686

	(7) Urban	(8) Rural	(9) Urban	(10) Rural
Education	0.092***	0.073***	0.079***	0.065***
	(0.005)	(0.007)	(0.006)	(0.009)
Δ Education	0.059***	0.065***	0.054***	0.053***
	(0.007)	(0.009)	(0.009)	(0.012)
Comm FE	NO	NO	YES	YES
Observations	3467	2425	3467	2425
Nr comm.			571	982
R-squared	0.419	0.393	0.589	0.717

*Dependent variable is log wage from main activity. Sample for 1992 is based on 34 districts. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%.; all regressions are run with pooled data for 1992-2000 and include Mincerian controls as well as the other controls as in Tables 5.1 and 5.2 and a series of interaction terms between each variable and a post-1992 dummy. ΔEducation reports the coefficient of this interaction for the number of year of schooling variable.*

Given these results it is worth checking to what extent the change in the private-public sample composition affects the change in the education coefficients over time. In order to do that, I generate 100 random sub-samples from the 2000 samples keeping the same employment (public vs. private) shares as in 1992. I replicate the exercise with the

1992 sample as well. I then calculate the coefficients of the change in the educational variables over time and take the mean of the point estimate across the different iterations. The results (available upon request) suggest that the relative increase in the private sector representation in the composition of the wage employees sample does not seem to have had any significant effect in the rise of returns to education.

5.5. Regional analysis

As discussed above results at the national level may hide large variations in labour market inequality across regions. It is important to analyse such variations, as the within region component explains most of the inequality currently observed in developing countries (Kanbur and Venables, 2005b). Moreover, looking at finer geographical scales than the national level may help shed light on the determinants of the rising returns to education in the nineties. Following the administrative (and statistical) division of Uganda, this sub-national analysis is first carried out for the four regions (i.e. Central, East, North and West), and then for the districts in the next chapter. The latter analysis will also allow for a more formal testing of the determinants of returns to schooling both across space and over time.

The analysis at the regional level supports the idea of fairly heterogeneous returns to education across space in 1992, which have been converging over the nineties (in a period when all regions have experienced rising returns to education). To calculate regional returns to education I adopt the basic as well as the extended Mincerian methods as in (5.3) and (5.4) using regions as the spatial unit of analysis. The coefficients on years of education are reported in Table 5.7. One extra year of education produces an increase in wages of between 8 and 12 per cent in 1992, which is reduced to between 6 and 9 per cent when including community effects. These ranges hold both when including the 4 districts in the 1992 sample which have no data in 2000 (columns 1 and 2) and when excluding them (columns 3 and 4). The education coefficient increases in 2000 to between 15 and 18 per cent (column 5) and between 12 and 14 per cent in 2000 when including community fixed effects (column 6). While the coefficient rises in all regions, the size of these increases varies substantially. The range of increases is smaller when using community effects (last column); in this case, the largest absolute increase (in percentage points) occurs in the Northern region (+7.3 per cent), while the Eastern and Central regions experience the smallest increases (+ 4.5 and +4.9 per cent). However, this ranking is reversed when the rise in return to schooling is obtained without controlling for community effects. In this case, the Eastern region experiences the highest increase (+7.8 per cent), while the Northern region has the lowest (+ 3.3 per cent).[114] This is due to the differential impact of community effects on the coefficients across regions and over time. While in 1992 the inclusion of community effects does not affect the size of return to schooling in the Eastern region, it does in 2000. The opposite is true for the Northern region. This suggests that the spatial nature of labour market inequalities has changed over time: in the Northern region for instance a large share of such inequality was explained at the community level in 1992 (the inclusion of community fixed effects almost halves the education coefficient). On the other hand in 2000 the within community component of inequality was very small. The subsequent district level analysis may help explore the causes of this changing nature of regional inequality further.

Table 5.7: Regional basic Mincerian regressions

	1992		1992		2000		Δ 1992-2000	
	(1)	(2)	(3)	(4)	(5)	(6)	(5)-(3)	(6)-(4)
Central	0.124*** (0.006)	0.091*** (0.007)	0.124*** (0.006)	0.091*** (0.007)	0.179*** (0.007)	0.140*** (0.009)	0.055	0.049
Eastern	0.078*** (0.009)	0.078*** (0.010)	0.078*** (0.009)	0.078*** (0.010)	0.156*** (0.008)	0.123*** (0.012)	0.078	0.045
Western	0.094*** (0.007)	0.079*** (0.008)	0.104*** (0.007)	0.085*** (0.008)	0.164*** (0.007)	0.144*** (0.010)	0.060	0.059
Northern	0.111*** (0.013)	0.062*** (0.011)	0.113*** (0.015)	0.064*** (0.013)	0.146*** (0.012)	0.137*** (0.016)	0.033	0.073
Comm. FE	NO	YES	NO	YES	NO	YES	NO	YES
Nr districts	38	38	34	34	34	34	34	34
Obs.	3585	3585	3286	3286	2608	2608		
R-squared	0.278	0.606	0.292	0.611	0.430	0.670		
Eq. F-stat	7.26**	1.63	6.30**	1.20	2.55*	0.69		

*Dependent variable is log wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%.. All regressions include standard Mincerian covariates (gender, experience and its square term) as well as the other controls as in the previous Tables (marriage variables and urban dummy). Eq. F-stats reports the value of the F-statistics of the test for the equality of the education coefficients across regions.*

[114] Although community effects are capturing a potentially wide range of very local conditions, I tend to prefer the coefficients calculated using these effects over those calculated excluding them. This is precisely because by capturing a range of unobservable local factors, these effects allow to better isolate the impact of the main variables of interest on earnings, and education in particular.

The other standard variables of Mincerian regressions are all highly significant and with the expected sign in all regions both in 1992 and 2000 (see Table A5.4 in Appendix). The anti-female gender bias is relevant in all regions in both 1992 and 2000, but it varies significantly across them.[115] This bias has significantly fallen in all regions during the nineties, except in the Northern region, where it has actually increased. Experience is positive but and with diminishing returns in all regions in both years. Unlike returns to education, the effect of experience on wages has remained fairly stable over time for all regions (between 3 and 7 per cent in both years) but the Northern region where it has increased.

It is difficult to understand the determinants of the levels and changes of returns to schooling across regions without using parametric analysis. However, with four data points one can only speculate about the possible determinants of the cross-regional variation in coefficients, while chapter 6 will perform parametric analysis on such determinants at the district level. Standard economic theory suggests that returns to a factor depend on the relative demand and supply of that factor. Table 5.8 summarises some potential proxies for relative demand and supply of educated labour. To the extent that the private sector pays a higher premium to skilled labour than the public sector (where wages are usually more equally distributed), a higher share of private sector employment may be related to higher returns to education. This seems to be in line with the evidence in 1992, when the ranking of regions according to this share resembles exactly that constructed according to returns to education using community effects (i.e. Central highest, then Western, etc.).

The share of educated labour (with a level of schooling higher than the compulsory primary school) in total employees may proxy for the supply of educated labour. Other things being equal, the higher this supply, the lower the price paid for it should be. Again, this is confirmed by the regional data: the two regions with the highest share of post-primary educated employees (Western and Northern) have also the lowest returns to schooling both in 1992 and 2000.

[115] The only exception is the Northern region in 2000, for which gender bias is insignificantly different from zero according to the standard levels of significance.

The analysis of the share of rural employees in total wage employment provides mixed evidence. Skilled labour may enjoy relatively higher returns in the modern urban sector (where it could exploit skill-biased technology). However, there does not appear to be any clear correlation between the share of rural wage employment and the education coefficient across regions and over time. Incidentally, it is interesting to note that the share of rural employment has substantially increased in all regions over the nineties, despite this was a period of mild urbanisation in Uganda (the urban share of population grew from 11% to 12% between 1990 and 2000).

Some further light on the determinants of changes in returns to education could be shed by the results of the extended method (as in (5.4)). This reveals that returns to all levels of education have increased but with a distinct pattern.[116] The Central region experienced its most significant increase in the post-secondary education coefficient, which doubles between 1992 and 2000. In the Eastern region the largest increase is that of the return to primary education, while in the Western and Northern regions it is returns to secondary education that rose the most. Interestingly, while the convex shape of the earning function with respect to skills holds for all regions in 1992, in 2000 it does so only for the Eastern and Central regions. The shape in the Western and Northern regions is convex up to the secondary school education and becomes concave after that. One possible interpretation of this could be that the demand for highly educated employees (i.e. post-secondary) is low in conflict-torn areas characterised by unsophisticated economic activities.

Table 5.8: Regional characteristics of Ugandan employees, 1992 and 2000

	1992				2000			
	Central	Eastern	Western	Northern	Central	Eastern	Western	Northern
Employees								
Private	65.6%	47.0%	59.8%	45.0%	81.2%	55.6%	65.2%	53.4%
Public	33.6%	51.5%	39.2%	53.3%	18.3%	42.8%	33.0%	44.2%
Other	0.8%	1.5%	1.0%	1.6%	0.5%	1.6%	1.8%	2.4%
Education								
Post-primary (>7)	48.3%	57.5%	41.8%	53.6%	50.0%	61.7%	50.3%	59.8%
Illiterates	11.5%	9.9%	14.0%	16.9%	10.2%	9.2%	9.1%	12.8%
Location								
Rural	34.6%	30.8%	44.4%	32.1%	43.6%	49.5%	56.7%	39.9%
Major town	56.7%	56.9%	51.3%	55.7%	49.3%	45.6%	39.1%	46.3%

Source: LSMS for 1992 and 2000

[116] Complete results are reported in Appendix 5.1 – Table A5.4.

Taken at its face value the evidence is suggestive of possible demand and supply related factors at work in determining returns to education across regions and over time. However, more robust tests for these factors are needed than those allowed by regional data to be able to make any meaningful inference on the effects at work. I turn to such tests in the next chapter using district level data.

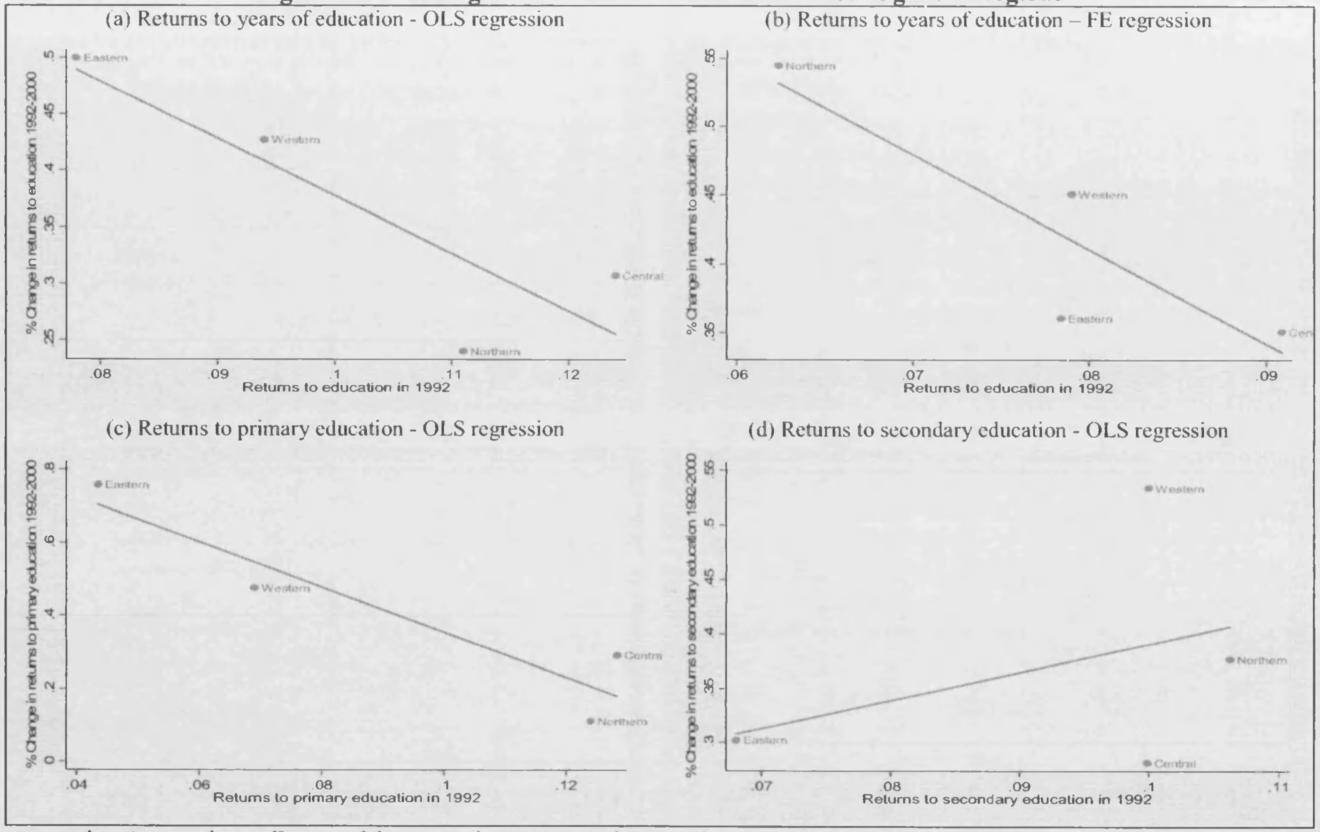
5.5.1 Convergence

How have relative returns across regions evolved over time? The tables above suggest that there appears to be both relative and absolute convergence in returns to education across regions. This idea of convergence is confirmed by an F-test of equality of education coefficients. The test strongly rejects the null (of equality) in 1992 (column 3), while the rejection is much weaker in 2000.[117] Also, the difference in regional coefficients seems to be explained by very local characteristics: when community fixed effects are included the education coefficients are not significantly different from each other (see columns 2, 4 and 6 in Table 5.7).

Figure 5.1 presents some graphical evidence of a standard convergence test. It plots the proportionate change in returns to education $(\beta_{2000} - \beta_{1992})/\beta_{1992}$ against the value of returns in 1992. The figure supports the idea of convergence across regions over time (quadrants a-c). This convergence in returns to education may be consistent with a context of liberalisation and market reforms such as that of Uganda in the nineties, which stimulated trade between regions. In this context movement of factors of production (embedded in goods) may operate as an equalising factor in the relative returns across regions. Regions relatively abundant in skilled labour may have caught up in terms of returns to education as they would specialise in skilled abundant goods (thus raising the relative demand for skilled labour). This finding is in line with the convergence found by Duranton and Monastiriotis (2002) for UK regions. But unlike in the case of the UK, convergence across Ugandan regions applies to returns to years of education but not to returns to experience. In particular the converging factor is returns to primary education while convergence does not seem to occur for secondary (quadrant d) and post-secondary education (not reported here).

[117] It is worth noting that the difference in coefficients among regions in 1992 is almost entirely explained by differences in primary education coefficients, while secondary and post-secondary coefficients are not significantly different from each other.

Figure 5.1: Convergence in returns to education across Ugandan regions



Source: returns to education are the coefficients of the regional regressions whose results are reported in Table 5.7 (OLS is computed without community fixed effects and FE is computed with community FE); returns to primary and secondary schooling are from the regressions whose results are reported in Table A5.4

5.6. Conclusions

The analysis in this chapter provides a robust confirmation that returns to education have increased substantially in Uganda over the nineties. This result is robust to difference specifications, samples, control variables and estimation methods. An additional year of education in 2000 provides on average an extra 5 to 6% increase in the wage premium relative to that in 1992. Given the small share of employees with post-primary education (i.e. 18.3% in the 1992 sample and 20.6% in 2000 had at least a secondary degree), this result is primarily driven by the rise in primary returns. These have doubled in the span of 8 years, experiencing a proportionately larger increase than returns to other schooling levels. This is encouraging for strategies aiming at expanding primary education as the one currently pursued by Uganda. In this light it is important to assess the extent to which increases in the supply of (primary) educated workforce may contrast this upward trend of returns to primary schooling. The district-level analysis in the next chapter helps shed some light on this question.

The pattern of changes in returns to schooling has differed between the public and the private sector, with the latter experiencing more homogeneous increases across educational categories than the former. Overall the skilled-unskilled wage gap has grown slightly more in the private than in the public sector. That is not the case for the rural and urban sectors, which show remarkably similar increases in returns to education. On the other hand some marked variability emerges from the regional analysis. Although returns increase in all regions, the Western and Eastern regions experience larger (absolute) increases when not controlling for community effects, and the Northern and the Western regions when including community effects. These changes seem to be consistent with a process of convergence over time as well as with skills' demand and supply related factors. However with four data point it is difficult to go beyond speculations about these hypotheses. More systematic analysis requires smaller spatial units, which at the same time should not be too small to represent separate labour markets. In the next chapter I will argue that the district is a spatial unit that fulfils these conditions in the case of Uganda.

Appendix 5.1.

Table A5.1: Determinants of wage for employees in Uganda, 1992 (all 38 districts)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) 2SLS	(10) 2SLS
Primary	0.093***	0.060***	0.066***	0.051***						
	(0.009)	(0.009)	(0.009)	(0.009)						
Secondary	0.091***	0.084***	0.078***	0.076***						
	(0.011)	(0.011)	(0.010)	(0.010)						
Post_sec	0.184***	0.126***	0.158***	0.123***						
	(0.026)	(0.026)	(0.025)	(0.027)						
Education					0.104***	0.080***	0.084***	0.073***	0.157***	0.135***
					(0.004)	(0.005)	(0.004)	(0.005)	(0.016)	(0.029)
Female	-0.221***	-0.318***	-0.247***	-0.289***	-0.223***	-0.314***	-0.247***	-0.286***	-0.276***	-0.258***
	(0.034)	(0.035)	(0.035)	(0.036)	(0.034)	(0.035)	(0.034)	(0.036)	(0.039)	(0.040)
Experience	0.041***	0.048***	0.032***	0.038***	0.040***	0.047***	0.032***	0.037***	0.052***	0.053***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.008)
Exp squared	-0.0007***	-0.0007***	-0.0005***	-0.0006***	-0.0007***	-0.0007***	-0.0005***	-0.0006***	-0.0006***	-0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Unmarried			-0.293***	-0.260***			-0.281***	-0.255***	-0.051	-0.084
			(0.036)	(0.041)			(0.037)	(0.042)	(0.060)	(0.084)
Other mar.			-0.200***	-0.170***			-0.192***	-0.163***	-0.103*	-0.099
			(0.048)	(0.048)			(0.048)	(0.049)	(0.058)	(0.061)
Urban			0.547***				0.546***		0.410***	
			(0.030)				(0.030)		(0.041)	
Central			0.302***				0.310***		0.339***	
			(0.038)				(0.038)		(0.042)	
Western			-0.034				-0.025		0.048	
			(0.040)				(0.040)		(0.048)	
Northern			-0.164***				-0.161***		-0.144***	
			(0.048)				(0.048)		(0.052)	
Comm FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Observations	3585	3585	3585	3585	3585	3585	3585	3585	3311	3129
R-squared	0.242	0.605	0.342	0.612	0.239	0.603	0.338	0.610	0.253	0.138
Nr communities		820		820		820		820		627
1st stage F-stat									108.86	30.30
Hansen J-stat									1.438	0.908

*Dependent variable is log wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%; instruments for education in columns 9 and 10 are father and mother's years of education. F-stat is the F-statistics from the first stage for the excluded instruments; Hansen J-statistic is the test for over-identification.*

Table A5.2: The influence of parents' education on children's education and wage, 1992 (34 districts)

Dep. variable	(1) Edu OLS	(2) Edu OLS	(3) Ln(wage) OLS	(4) Ln(wage) OLS	(5) Ln(wage) OLS	(6) Ln(wage) OLS
Father education	0.620***	0.396***	0.073***	0.023	0.040**	0.010
	(0.067)	(0.071)	(0.016)	(0.015)	(0.016)	(0.015)
Mother education	0.368***	0.194**	0.083***	0.053***	0.037**	0.023
	(0.077)	(0.088)	(0.017)	(0.016)	(0.019)	(0.018)
Education				0.081***		0.073***
				(0.004)		(0.004)
Other controls	YES	YES	YES	YES	YES	YES
Comm. FE	NO	YES	NO	NO	YES	YES
Observations	3042	3042	3025	3025	3025	3025
Nr. communities		751			746	746
R-squared	0.365	0.258	0.26	0.35	0.12	0.22
F-stat (mother & father edu)	100.48	30.44	44.42	12.44	10.27	2.04

*Dependent variables are number of years of education (columns 1-2) and log wage from main activity (columns 3-6). Robust standard errors (Huber-White method) in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. F-test tests for the joint significance of mother and father's education.*

Table A5.3: Determinants of daily wage for employees in Uganda, 1992 (34 districts)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Primary	0.089***	0.051***	0.056***	0.040***	
	(0.010)	(0.010)	(0.009)	(0.011)	
Secondary	0.087***	0.086***	0.076***	0.077***	
	(0.013)	(0.012)	(0.012)	(0.012)	
Post-sec	0.186***	0.126***	0.154***	0.119***	
	(0.031)	(0.033)	(0.031)	(0.035)	
Education					0.101***
					(0.004)
Mincerian controls	YES	YES	YES	YES	YES
Other controls	NO	NO	YES	YES	NO
Community FE	NO	YES	NO	YES	NO
Observations	2925	2925	2925	2925	2925
Nr. of communities		703		703	
R-squared	0.225	0.183	0.330	0.200	0.222
	(6) OLS	(7) OLS	(8) OLS	(9) 2SLS	(10) 2SLS
Education	0.078***	0.079***	0.069***	0.154***	0.137***
	(0.006)	(0.004)	(0.006)	(0.018)	(0.033)
Mincerian controls	YES	YES	YES	YES	YES
Other controls	NO	YES	YES	YES	YES
Community FE	YES	NO	YES	NO	YES
Observations	2925	2925	2925	2690	2526
Nr. of communities	703		529		529
R-squared	0.178	0.326	0.193	0.242	0.116
1st stage F-stat			81.33		25.06
Hansen J-stat			3.387		1.027

*Dependent variable is log of daily wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%; instruments for education in columns 9 and 10 are father and mother's years of education. F-stat is the F-statistics from the first stage for the excluded instruments; Hansen J-statistic is the test for over-identification.*

Table A5.4: Regional extended Mincerian regressions (34 districts in 1992)

		1992				2000			
		(1)		(2)		(3)		(4)	
		Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Constant	Western	10.975***	(0.132)	11.345***	(0.077)	11.084***	(0.155)	11.314***	(0.101)
Region effects	Central	0.169	(0.176)			-0.151	(0.206)		
	Eastern	0.689***	(0.201)			-0.058	(0.233)		
	Northern	-0.008	(0.333)			0.197	(0.263)		
Primary	Central	0.128***	(0.014)	0.070***	(0.015)	0.181***	(0.017)	0.133***	(0.020)
	Eastern	0.043**	(0.019)	0.051**	(0.023)	0.181***	(0.024)	0.113***	(0.033)
	Western	0.082***	(0.015)	0.073***	(0.017)	0.132***	(0.019)	0.127***	(0.023)
	Northern	0.127***	(0.030)	0.063**	(0.024)	0.140***	(0.027)	0.086**	(0.034)
Second.	Central	0.100***	(0.018)	0.093***	(0.019)	0.139***	(0.022)	0.104***	(0.026)
	Eastern	0.068***	(0.023)	0.083***	(0.024)	0.097***	(0.030)	0.115***	(0.036)
	Western	0.109***	(0.023)	0.075***	(0.022)	0.215***	(0.029)	0.183***	(0.032)
	Northern	0.097***	(0.034)	0.063*	(0.038)	0.171***	(0.062)	0.208***	(0.070)
Post-sec	Central	0.186***	(0.039)	0.136***	(0.042)	0.313***	(0.053)	0.277***	(0.063)
	Eastern	0.204***	(0.060)	0.140**	(0.067)	0.277***	(0.065)	0.163*	(0.083)
	Western	0.166***	(0.052)	0.160***	(0.053)	0.084	(0.065)	0.065	(0.077)
	Northern	0.110*	(0.066)	0.069	(0.065)	0.084	(0.136)	0.050	(0.170)
Female	Central	-0.311***	(0.055)	-0.373***	(0.065)	-0.098	(0.063)	-0.147**	(0.070)
	Eastern	-0.274***	(0.080)	-0.345***	(0.078)	-0.142	(0.088)	-0.131	(0.102)
	Western	-0.224***	(0.071)	-0.278***	(0.078)	-0.141	(0.091)	-0.299***	(0.101)
	Northern	-0.073	(0.100)	-0.244**	(0.097)	-0.038	(0.117)	-0.139	(0.146)
Exp.	Central	0.055***	(0.006)	0.054***	(0.006)	0.067***	(0.007)	0.061***	(0.008)
	Eastern	0.036***	(0.007)	0.047***	(0.008)	0.057***	(0.010)	0.047***	(0.010)
	Western	0.049***	(0.008)	0.056***	(0.008)	0.047***	(0.010)	0.056***	(0.010)
	Northern	0.019	(0.015)	0.029**	(0.014)	0.032**	(0.013)	0.041**	(0.017)
Exp. sq.	Central	-0.001***	(0.0001)	-0.001***	(0.0001)	-0.001***	(0.0001)	-0.001***	(0.0001)
	Eastern	-0.001***	(0.0001)	-0.001***	(0.0001)	-0.001***	(0.0001)	-0.001***	(0.0002)
	Western	-0.001***	(0.0001)	-0.001***	(0.0001)	-0.001***	(0.0001)	-0.001***	(0.0001)
	Northern	-0.0002	(0.0002)	-0.001**	(0.0002)	-0.000*	(0.0001)	-0.0002*	(0.0001)
Observations		3286		3286		2608		2608	
Comm FE		NO		YES		NO		YES	
Nr of communities				757				796	
Eq. F-test primary		4.97**		0.23		1.71		0.52	
Eq. F-test secondary		0.61		0.23		2.79*		1.64	
R-squared		0.30		0.49		0.43		0.52	

*Dependent variable is log of daily wage from main activity. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%.*

Chapter 6. Why have returns to education increased in Uganda over the 1990s? A district level analysis

6.1. Introduction

The regional analysis in chapter 5 suggests that returns to education are fairly heterogeneous across regions and so are their changes. This chapter uses a finer scale of geographical classification - the district level – to investigate the determinants of these differences across space in a more systematic way.[118] Such analysis is based on the assumption that districts define separate labour markets. I would argue that this is a fair approximation in the case of Uganda as labour mobility between districts is fairly limited: according to the LSMS data inter-district migrants between 1992 and 2000 represent on average only 6% of the total district's labour force. As mentioned above this is a very low figure by international standards.

The idea that districts may represent different labour markets is also supported by the statistically different returns to education across districts, as shown in Table 6.1 below.[119] The F-test strongly rejects the null of equality of returns to education coefficients in 1992 and even more strongly in 2000.[120] These coefficients are the β coefficients in equation (3) and Table 6.1 reports also their standard errors and the number of individuals over which they are computed.[121] Almost all of the coefficients are statistically significant at the standard levels in both years: 29 out of 38 coefficients in 1992 and 33 of 34 in 2000 are significant at the 5% level. The only coefficient which is not significant in any years is

[118] I develop the analysis keeping the districts' configuration of 1992, which is represented in Figure A6.1 in the Appendix (along with regional divisions). The population tends to distribute towards districts in the Central and Eastern regions along borders with Kenya and Tanzania. On the other hand, population density in the Northern region districts is relatively low.

[119] Also returns to other key labour market characteristics such as experience and gender appear to differ substantially across districts (not shown here).

[120] Although there is a tendency towards convergence between 1992 and 2000 in returns to education (both at the regional and at the district level), the stronger rejection of inter-district equality in 2000 may appear somewhat puzzling. In fact it may be explained by the fact that returns to education (as well as to other key labour market characteristics) seem to be estimated more precisely in 2000 than in 1992.

[121] The coefficients are estimated without the use of community fixed effects, which are likely to underestimate returns to education since they net out any effects education may have in encouraging migration to areas of high labour demand or in providing local externalities (Appleton, 2001).

that of Kapchorwa district, which is also the only negative coefficient in both periods. While these negative coefficients may be representative of true negative returns to education in this district, their estimation is very imprecise and the standard intervals of confidence include also positive values. Such imprecise estimates may be due to the very low number of employees over which the returns are computed (37 in 1992 and only 14 in 2000).[122] Because of this I exclude Kapchorwa district from some of the specifications in the subsequent analysis as robustness check.

The importance of performing a district analysis is further confirmed by the sizeable impact of highly localised factors (captured by community fixed effects) on returns to education in the national level regressions above. While using communities as the spatial scale of analysis is ruled out by the very low number of employees per community available in the first stage regressions, the district analysis provides a good compromise in the trade-off between this number and the number of observations (i.e. the number of units) available in the second stage. This analysis also allows one to test for the channels through which market reforms may have influenced labour market inequalities.

There are a number of other reasons why the district is an interesting spatial unit for this empirical work. It represents the most important institutional division of the state in Uganda and the constitution grants it law-making powers in various subjects including health, education and taxation. In particular, districts share responsibilities with the central government for the educational system. This may provide a source of variation in the quality of education across districts, which is likely to apply to the labour force in our data for two reasons: the vast majority of districts in the analysis were already in place when most of the literate labour force in the surveys was educated, and the labour force mobility across districts is limited.[123] A further reason why the district-level analysis is increasingly important in Uganda is that the decentralisation process initiated in 1993 has been transferring to districts more state powers and responsibilities.

[122] These are among the lowest numbers of employees used in districts and over four times below the average number used in each district.

[123] Ninety-five percent of the districts used in the analysis had been in existence since 1971 (just after the Independence).

Table 6.1: Returns to education across Ugandan districts, 1992-2000

District No.		1992			2000		
		β_{1992}	S.E.	Nr obs	β_{2000}	S.E.	Nr obs
1	Kalangala	0.082	0.034	56	0.133	0.023	32
2	Kampala	0.105	0.010	430	0.173	0.013	235
3	Kiboga	0.082	0.026	44	0.158	0.068	23
4	Luwero	0.081	0.019	64	0.193	0.031	74
5	Masaka	0.079	0.015	169	0.159	0.014	128
6	Mpigi	0.146	0.014	237	0.174	0.016	199
7	Mubende	0.141	0.019	73	0.136	0.017	69
8	Mukono	0.070	0.019	161	0.157	0.023	161
9	Rakai	0.208	0.034	58	0.203	0.032	37
11	Iganga	0.072	0.023	93	0.165	0.028	87
12	Jinja	0.130	0.015	242	0.178	0.017	125
13	Kamuli	0.032	0.032	44	0.202	0.023	64
14	Kapchorwa	-0.025	0.019	37	-0.058	0.113	14
15	Kumi	0.048	0.020	49	0.245	0.024	24
16	Mbale	0.072	0.018	126	0.131	0.019	129
17	Pallisa	0.096	0.032	34	0.117	0.037	32
18	Soroti	0.073	0.031	62	0.148	0.032	61
19	Totoro	0.048	0.021	90	0.114	0.019	104
21	Bundibugyo	0.010	0.035	32			0
22	Bushenyi	0.133	0.022	107	0.144	0.017	97
23	Hoima	0.147	0.035	39	0.164	0.023	61
24	Kabale	0.084	0.024	63	0.160	0.023	71
25	Kabarole	0.114	0.013	181	0.182	0.014	115
26	Kasese	0.049	0.022	110			0
27	Kibaale	0.128	0.022	38	0.191	0.036	38
28	Kisoro	0.027	0.041	22	0.184	0.038	41
29	Masindi	0.090	0.029	39	0.162	0.026	45
31	Mbarara	0.098	0.013	240	0.155	0.017	156
32	Rukungiri	0.093	0.032	55	0.140	0.017	58
41	Apac	0.116	0.022	44	0.161	0.031	26
42	Arua	0.061	0.020	92	0.149	0.027	87
43	Gulu	0.072	0.019	93			0
44	Kitgum	0.117	0.029	64			0
45	Kotido	0.175	0.056	65	0.172	0.048	42
46	Lira	0.059	0.023	78	0.113	0.029	72
47	Moroto	0.149	0.030	80	0.115	0.021	29
48	Moyo	0.047	0.025	53	0.180	0.034	45
49	Nebbi	-0.019	0.034	21	0.116	0.034	27
	Mean	0.087	0.025	94	0.153	0.029	77
	S.D.	0.049	0.009	82	0.048	0.018	53
	Signif. (5%)	29/38			33/34		
	Signif. (1%)	24/38			30/34		
	F-stat	15.99			37.70		

Note: βs are estimated through the OLS method in equation (5.3); the F-stat is the test for the joint significance of the β coefficients; Nr obs. refers to the number of observations over which β is estimated.

6.2 Methodology

6.2.1 Data

The data for the dependent variables come from the estimation of regressions (5.3) and (5.4) described in chapter 5. Most of the district-level regressors used in the analysis are reported in the LSMS surveys at either the individual or the household level so that I can use micro information to construct labour market variables for each of the districts. Hence, these district-level factors can be treated as exogenous to individual decisions, but relevant to the determination of returns to education in local labour markets.[124]

I also construct some of the district-level variables from the part of the surveys containing information on the communities of residence of the individuals surveyed. There are 1,216 and 1,086 communities surveyed in 1992 and 2000 respectively, although most of them are not the same over time. The surveys gather various types of information, including on infrastructure, prices, health and education. I average this community data to generate the district-level variables.

6.2.2 Specifications and variables

The basic idea of the empirical analysis is to identify the determinants of labour market inequalities in the Ugandan context using the β coefficients estimated from equations (5.3) and (5.4) as dependent variables. In any market economy skill wage differentials respond to relative supply of and demand for skills. In a world with two factors of production (skilled and unskilled labour), the relative price of skilled labour would decrease in its relative supply and would increase in its relative demand (Katz and Autor, 1999).[125] The basic specification to test these predictions reads as follows:

$$\beta_{kt}^{OLS} = a_k + b_1 \frac{Sup_{kt}^{S}}{Sup_{kt}^{Tot}} + b_2 \frac{Dem_{kt}^{S}}{Dem_{kt}^{Tot}} + d2000 + \varepsilon_{kt} \qquad (6.1)$$

[124] In fact there may be a problem of endogeneity if for instance those district-level variables determine the location of individuals or households with certain characteristics into a specific district. I try to control for this potential source of endogeneity below via instrumental variable estimation.
[125] The use of district fixed effects in the empirical specification ensures that other relatively fixed factors of production, such as land and to some extent capital, are controlled for.

where β_{kt}^{OLS} are the schooling coefficients estimated through (5.3), a_k are district fixed effects, Sup_{kt}^{S} is the supply of skilled labour, Sup_{kt}^{Tot} is the total supply of labour in district k at time t and the same notation holds for labour demand (*Dem*) as well; and *d2000* is a time dummy.

The fact that the dependent variable in (6.1) is estimated rather than observed should not present any difficulties for the regressions aside from a loss of efficiency, unless the sampling error in the dependent variable (i.e. the difference between the true and the estimated value of the dependent variable) is not constant across observations. In this case the regression errors will be heteroscedastic and OLS estimation may generate inconsistent standard errors (Lewis and Linzer, 2005). As the samples in the LSMS are stratified at the district level, the estimated β in (6.1) should not have sampling errors which are systematically different across districts, that is $E(\varepsilon_i \mid \mathrm{Edu}_i, W_i, Z_i) = \hat{\varepsilon}_i$, $\forall\, i$ in (5.3) and (5.4). On the other hand the large cross-district variation in the number of observations over which the β are estimated (see Table 6.1) may represent a source of possible heteroscedasticity in the standard errors in (6.1). Lewis and Linzer (2005) show that the OLS method using White's (1980) heteroscedastic consistent standard errors is generally the most reliable way of estimating regressions with estimated dependent variable in the presence of heteroscedasticity. That is the case except in two instances. First, when the share of the regression residual due to sampling error in the dependent variable is very high (at least 80%) WLS estimation is preferred. Second, when information about the sampling errors in the dependent variable is available and highly reliable, the feasible generalized least squares estimator developed by Lewis and Linzer (2005) is a more efficient option. However neither of these cases is likely to apply here: the eventual sampling error in the dependent variable is unknown, and it is likely to be not very high (if it exists at all) as argued above. Therefore I estimate (6.1) via OLS using White's heteroscedastic consistent standard errors. To be on the safe side I also test the robustness of the results to excluding the district with the lowest number of observations (Kapchorwa), as this may be an important source of the dependent variable's sampling error, and to estimating (6.1) through WLS. As shown below the results are little affected, suggesting that the standard errors in (6.1) are robust to the possible concerns related to the use of an estimated dependent variable.

Table A6.1 in the Appendix describes the construction of the supply and demand variables. The supply of skills is proxied in the analysis through education-based measures. In particular, I use the district's share of wage earners who have completed an education equal or higher than the primary level (*skilled employees*) as the main measure of the relative supply of skills. Completion of primary schooling was still relatively rare in Uganda in the nineties and this is one of the reasons why the government launched the Universal Primary Education initiative in 1997. The unweighted district-wise average share of adults that completed primary school was 32% in 1992 and 34% in 2000. Although these percentages are higher in the case of employees (58% in 1992 and 63% in 2000), they still indicate that primary education is far from universal.[126] To avoid the sensitivity of the results to the choice of the skilled-unskilled threshold I also use the average number of years of formal schooling completed by the wage earners (*employees education*) as an alternative measure of the supply of educated labour. Admittedly, the supply of skills may be endogenous to returns to education; this is the case for instance if higher relative returns in a district act as an incentive for employees to acquire more education or attract people with more education. I use an IV strategy in order to correct for these potential endogeneity biases, as explained below. In the subsequent analysis I also try to correct for the possible bias arising from the skilled migration induced by differential returns to education controlling for the districts' rate of skilled immigration.

The relative demand for skilled labour may depend on a number of different factors. It is widely recognised that greater access to (and use of) technology increases the relative demand for skilled labour, even in low-income developing countries. As argued by Anderson (2005) this might be for a number of reasons. First, adapting to a new technology is a difficult task which requires the use of skilled labour; second, technological progress has often substituted unskilled labour (e.g., automatic assembly lines); third, greater access

[126] This is further confirmed by the unweighted district-wise average of years of education, which was 7.5 in 1992 and 7.8 in 2000 and just above the number of years to complete primary education.

to foreign technology allows developing countries to compete internationally in more skill intensive goods, raising their average skill intensity of production, and thus the relative demand for skilled labour. The data available does not allow the construction of a direct measure of the use and availability of technology in production. Given the data available, I could use a demand-related variable as technological proxy, i.e. the district-wise share of households' purchase of electronic goods in total household income (*electronic purchases*). This is a valid measure to the extent that relatively high shares of electronic goods purchased by households may reflect greater access to and familiarity of usage of modern technology. As firms draw their labour-force predominantly from households in the district where they are located, this measure may proxy for firms' access to and use of modern technology. However, this measure is likely to be correlated with a lot of the same factors which are also related to returns to education, including the level of education, the level of income, eccetera.

Supply-related factors, such as public infrastructure measures, appear to be more exogenous as technological proxies. In particular I use the average distance from each community in a district to the closest telephone booth (*telephone distance*). In the fixed effects specifications employed in the empirical analysis this is effectively close to a measure of district-wise public telephone density. The telephone infrastructure was not developed in the nineties in Uganda (as in most LDCs). The median distance of a community from a telephone boot was 15 Km (and the mean was 27) in 1992 and 12 Km in 2000 (with the mean of 24). This infrastructure can be considered as a good proxy of the technological frontier in such a context. This is confirmed also by the fact that *telephone distance* is an important determinant of consumption of *electronic purchase* itself at the district level (Table 6.2). *Telephone distance* has a negative and significant association with *electronic purchases:* the larger the distance to the telephone (i.e. the lower the telephone density) the lower the consumption of electronic goods (column 1). This result is robust to the inclusion of the average number of years of education in the district (*population edu*) as a control, which has a positive non linear effect (column 2). And it even holds when including the initial share of electronic assets in total durable assets (*electronic assets*) – column 3.

Table 6.2: The relationship between telephone density and electronic purchases

	(1) Elec. puchases	(2) Elec. puchases	(3) Elec. puchases
Telephone distance (x1000)	-0.034***	-0.030***	-0.033***
	(0.006)	(0.008)	(0.008)
population edu (x 1000)		8.461*	6.290*
		(4.448)	(3.472)
population edu squared (x 1000)		-0.858*	-0.715**
		(0.453)	(0.350)
Electronic assets			17.898**
			(8.595)
Observations	72	72	72
R-squared	0.330	0.452	0.546
Number of district	38	38	38

*Dependent variable is the share of electronic purchases in total income. Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%; all regressions include district fixed effects and year effect.*

Another result that emerges from Table 6.2 is that the consumption pattern of the population appears to be influenced by its level of education. In general, the higher this level of education the more sophisticated the consumption of goods and services (i.e. which are more intensive in skills) may become. Leonardi (2008) provides empirical evidence in support of this pattern. As a share of this skill-intensive demand - that of non tradable services - needs to be satisfied locally (e.g. local newspaper, good quality education, doctors) this may raise the relative demand for skilled labour in the district. I test this hypothesis by using the average level of education (*population edu*) of the adult population as a regressor. Note that the use of this variable together with *skilled employees* does not pose a problem of collinearity, as the two are only mildly correlated (correlation coefficient of 0.33). That is because employees represent a relatively minor proportion of the entire adult population, which is mainly composed by self-employed and domestic labourers. In the presence of segmented labour markets these categories are quite distinct from the employees' population. To the extent that the choice of schooling of the general population may be partly determined by returns to education in wage employment, *population edu* may be endogenous in specification (5.1). Again I deal with this problem by instrumenting this variable as illustrated below.

Wages in the public sector are usually less linked to the actual marginal product of labour than those in the private sector. This is very evident in transition economies where public sector wages tend to be more uniformly distributed than private sector ones (see Flabbi et al., 2007, Yang, 2005). Thus I also include the size of the public sector - defined as the percentage of employees employed in state-owned enterprises and local public firms (*public sector*) – as a right-hand side variable in equation (5.1). As highlighted by the findings in chapter 5, returns to education in the public sector, especially at the primary level, experienced a significant increase in 2000 relative to 1992. For this reason I also interact the variable with a post-1992 dummy.

Finally, the urban sector is usually one in which the return to human capital is higher than in the rural sector. That is one of the reasons why other things being equal urban dwellers accumulate more human capital than their rural counterpart (Glaeser and Maré, 2001; Lucas, 2004). Once controlling for the average level of human capital in the workforce, the share of employees located in urban areas (*urban employment*) should then be positively associated with returns to education.

6.2.3 Instrumental variables

The two sets of likely endogeneous variables in (6.1) are the educational level of the workforce (*skilled employees* and *employees education*) and that of the adult's population (*population edu*). As they are both educational variables it would be difficult to find separate instruments for each variable. I use two sets of instruments for both of the endogenous variables: the average distance of each community to the nearest primary school (*primary school distance*) and that to the nearest secondary school (*secondary school distance*). In a context like Uganda with poor transport infrastructure and availability of vehicles, distance may indeed represent an obstacle to school attendance. According to the data in both LSMS surveys over 2% of schooling age population quoted distance from school as the most important reason for never enrolling into school or for dropping out. In 32% of the communities surveyed in 1992 distance from the school was an important (or a very important) reason in the decision of some children not to enrol in one of the three most popular community's primary schools.[127] This is not surprising given that the mean of the *primary school distance* variable in 1992 was 3.7 Km (1.8 the median) with peaks as high as 48 Km.

The major problem with these variables is that they refer to the current period while the workforce and the adult population in general was educated before. However this may not be as big of a problem as it first appears. In fact, like most infrastructure, also schools take time to be built and to become operational. Thus there is likely to be a high degree of persistence over time in any measure based on schooling infrastructure, i.e. the average distance from school in 1992 is likely to be highly correlated to that during the periods when the current adult population was educated. This is indirectly corroborated by the results of the first stage regressions presented in Table 6.2. *Primary school distance* appears to be a negative and significant determinant of skilled employees (column 1). This is the main driver of the relatively high partial R-squared, measuring the explanatory power of the school distance variables, i.e. the power of the instruments. On the other hand *secondary school distance* has a small positive and insignificant effect on education, which is somewhat surprising. This coefficient is in fact hiding a non linear U-shaped relationship between educational variables and *secondary school distance*, which emerges in column 2. When the average distance to the secondary school is high, a further increase in this distance is associated with an increase in the average level of education. A possible explanation may be that a high average distance in the absence of motorised transportation is probably associated with schools providing full boarding to the pupils. This may reduce the drop-out rate relative to a situation where the distance is high but not enough to have boarding schools, determining a higher secondary school completion rate. This non-linear pattern seems not to apply to primary schools probably because the average distance to these schools is generally much lower than that to secondary schools. For instance in 1992 there was only one district with *primary school distance* above 15 Km (i.e. a long enough distance to make a daily commute on foot almost impossible, thus requiring the presence of boarding schools) against five for *secondary school distance*. These results hold for the other educational variables as well (columns 3-4). It is worth noting that distance to secondary school appears to be a more significant determinant of *population edu* than

[127] The question was not posed in the 2000 survey.

primary school distance (as measured by the standardised coefficients and their degree of significance – not shown here).

As the distance data come from slightly different questions in the two surveys, I also add the interaction between each distance variables with a post-1992 dummy.[128] This inclusion does not affect much the other coefficients but raises the predictive power of the instruments (columns 5-7). This is especially through the effect of the secondary school distance interaction term which is positive and significant for the employees' variables and negative and significant for *Population edu.*

Table 6.3: First stage regressions for education variables

	(1) *Skilled employees*	(2) *Skilled employees*	(3) *Populat. edu*	(4) *Empl. education*	(5) *Skilled employees*	(6) *Populat. edu*	(7) *Empl. education*
Primary school distance	-0.009***	-0.008***	-0.014*	-0.067***	-0.007***	-0.016**	-0.065***
	(0.001)	(0.002)	(0.008)	(0.014)	(0.002)	(0.006)	(0.014)
Sec. school distance	0.005	-0.009	-0.064*	-0.126	-0.014	-0.040	-0.173**
	(0.003)	(0.008)	(0.035)	(0.083)	(0.009)	(0.031)	(0.075)
Sec. school distance sq.		0.0003*	0.0017***	0.0031**	0.0004**	0.0011**	0.0045***
		(0.0001)	(0.0006)	(0.0015)	(0.0002)	(0.0005)	(0.0013)
*Primary school dist.*Post-92*					-0.035	-0.127	-0.846**
					(0.038)	(0.096)	(0.398)
*Sec. school dist.*Post-92*					0.015***	-0.062***	0.184***
					(0.005)	(0.020)	(0.046)
Other controls	YES	YES	YES	YES	YES	YES	YES
Observations	72	72	72	72	72	72	72
Nr. of districts	38	38	38	38	38	38	38
R-sq. (within)	0.381	0.452	0.717	0.413	0.498	0.771	0.540
Partial R-sq.	0.283	0.365	0.399	0.370	0.418	0.511	0.510
F-stat.	41.42	54.64	14.63	76.49	53.24	27.50	96.36

*Robust standard errors (Huber-White method) in parentheses; *significant at 10%; ** significant at 5%; *** significant at 1%; all regressions include district fixed effects and year effect; other controls include the other controls in Table 6.4. Partial R-squared is the R-squared of the school distance variables (i.e. excluded instruments in the Tables 6.4-6.6); F-statistics refers to the test for the joint insignificance of the school distance variables.*

Other than being good predictors of a district's level of education, the school distance variables need also to be exogenous in (6.1) in order to be correctly excluded from

[128] In 1992 the questionnaire asked the "distance of the centre of the community to primary school", while in 2000 it asked the distance to the "nearest primary school" as well as to the "most common primary school". I use the former data to construct the variable although the results are very similar using the latter as well.

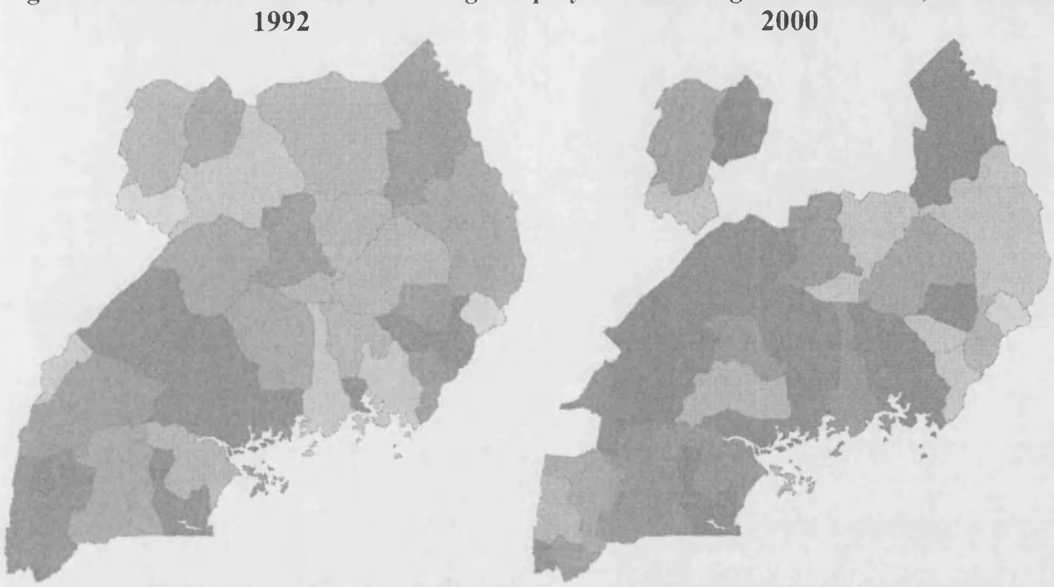
the second stage. In other words distance from schools needs to be unrelated to other determinants of returns to education not included in (6.1). This condition would not hold if for instance some unobserved shocks to the labour market which increased returns to education in a district (e.g. a new manufacturing plant requiring specialised labour) induced also the establishment of new schools in the same district (e.g. to supply perspective specialised labour for the plant). Although possible in principle this type of mechanism is not likely to be at work in a public pre-university schooling system as the Ugandan one (in 2000 only 10% of the primary schools were privately run). As the objective of any public education system should be to provide a public good to the population, this implies that the placement of primary and (possibly) secondary schools is likely to be mainly based on population criteria rather than on the basis of the effective demand for education.[129] The public good nature of schools is also one of the guiding principles of the Universal Primary Education (UPE) programme launched in 1997. The exogeneity of the instruments is confirmed by the data as well, as shown by the results of the 2SLS analyses below: the null hypothesis that the instruments are valid and correctly excluded from the second stage is never rejected at even modest levels of significance.[130]

6.3. Determinants of returns to education

Figure 6.1 graphically represents the evolution of returns to education reported in Table 6.1 on the Ugandan map. Although the increase in returns to education is general across districts, some geographical patterns of increases do emerge: interior districts in the Central and Western regions experience particularly significant rise in returns, while increases in the districts on the Eastern border appear to be more limited. I will test for some of these patterns more formally in the following analysis.

[129] Note that the inclusion of population as a control does not change the results in Table 6.3. I exclude it from the regressions below on the determinants of returns to education as its coefficient is never significant and reduces somewhat the power of the instruments although does not affect the overall results (results available upon request). This little effect of the population variable seems to be due to a similar pattern of population growth across districts with the fixed effects capturing the cross-district variation in population levels.
[130] This hypothesis is tested through the Sargan-Hansen test of overidentifying restrictions. The value of the statistics - reported in the Tables below - is never significant even at the 20% level.

Figure 6.1 Returns to education for wage employees across Ugandan districts, 1992-2000



Note: returns to education are increasing in the intensity of the shade (i.e. the higher the returns, the darker the district)

Table 6.4 presents the results based on equation (6.1). The results largely confirm the expected effects of skills' demand and supply factors. Column 1 runs the fixed effects regression with the endogenous education related regressors. The supply of educated employees – *skilled employees* – depresses returns to education, while the average education of the adult population – *population edu* – has a positive effect on returns. This is consistent with the idea that education does not influence returns to education only through the supply side but also through skills demand channels, which raise the demand for relatively skilled products and services. The other variables are not significant at standard levels, although their sign is consistent with the expectations (negative for *telephone distance* and *public sector* and positive for *urban employment*). In terms of the magnitude of the effects, a 10% increase in the share of primary educated employees (i.e. from the mean of 60% to 70%) is associated with a reduction of 1.5 percentage points in returns (i.e. from the mean of 10.2% to a still healthy 8.7%). On the other hand, a one year increase in the adult population's average education is associated with a 3 percentage points increase in returns. This effect suggests that any population-wide expansion of education, such as the one expected through UPE, would potentially counteract the decrease in returns due to increased supply of skilled employees.

Table 6.4: Determinants of returns to education 1992-2000, demand and supply

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	FE	FE	FE IV	FE IV	FE IV	FE IV	FE IV	FE IV	FE IV	FE IV (WLS)
	β_{OLS}	β_{OLS}	β_{OLS}	$\beta_{OLS}^{PRIMARY}$	$\beta_{OLS}(sep)$	β_{FE}	β_{OLS}	β_{OLS}	β_{OLS}	β_{OLS}
Skilled employees	-0.149*	-0.148**	-0.167***	-0.271*	-0.153**	-0.188***		-0.139**	-0.056**	-0.111*
	(0.088)	(0.070)	(0.058)	(0.163)	(0.077)	(0.063)		(0.061)	(0.026)	(0.059)
Employees edu							-0.018***			
							(0.006)			
population edu	0.030*	0.025	0.045**	0.157***	0.058**	0.045**	0.046***	0.034		0.028
	(0.017)	(0.015)	(0.018)	(0.056)	(0.024)	(0.022)	(0.016)	(0.023)		(0.022)
Telephone distance (x100)	-0.012	-0.011	-0.016	0.046	0.004	-0.091***	-0.023	-0.017	-0.008	-0.015
	(0.016)	(0.020)	(0.018)	(0.043)	(0.025)	(0.027)	(0.019)	(0.019)	(0.023)	(0.019)
urban employment	0.065	0.068	0.067	-0.174	0.041	0.150*	0.072	0.103**	0.102**	0.072
	(0.062)	(0.063)	(0.060)	(0.144)	(0.064)	(0.088)	(0.063)	(0.049)	(0.048)	(0.046)
Public sector	-0.055									
	(0.069)									
Public sector Post-1992*	0.065									
	(0.075)									
Post-1992	0.032	0.063***	0.054***	-0.010	0.046***	0.049**	0.051***	0.065***	0.078***	0.061***
	(0.028)	(0.016)	(0.015)	(0.033)	(0.016)	(0.023)	(0.015)	(0.013)	(0.010)	(0.015)
Observations	72	72	68	68	68	68	68	66	66	68
Nr. of districts	38	38	34	34	34	34	34	33	33	34
R-sq. (within)	0.687	0.676	0.663	0.253	0.615	0.415	0.644	0.746	0.717	0.718
1st stage F-stat			11.86	11.86	11.86	11.86	13.04	13.15	64.43	11.94
Hansen J-stat			0.092	3.246	1.013	0.766	0.130	1.636	1.574	0.844

*Dependent variables are returns to education as estimated through (5.3) and its variants (see main text). Robust standard errors (Huber-White method) in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. F-stat is the statistics for the joint significance of the excluded instruments in the first stage; Hansen J-statistic is the over-identification test. Skilled employees, employees edu and population edu are instrumented through distance to primary school, distance to secondary school, distance to secondary school squared and distance to secondary school x Post-1992 dummy. The 4 districts with a single observation are dropped in the IV estimations. In columns 8 and 9 Kapchorwa district is excluded.*

The share of the public sector in employment – *public sector* – decreases returns, but less so in the second period (measured through the interaction *Public sector* Post-1992*).[131] The insignificance of *public sector* seems to be related to its correlation with *skilled employees*, which captures a substantial share of the *public sector* effect on returns. This is via its (positive) impact on the stock of educated employees, which complements the traditional effect of *public sector* in compressing labour market inequalities.[132] Finally *telephone distance* has a negative effect, i.e. greater access to (and use of) technology raises returns, while *urban employment* has a positive effect on the skilled-unskilled wage premium. The F-test does not reject the hypothesis that the public sector coefficients are jointly equal to zero. Therefore I exclude them in column 2 obtaining very similar coefficients to column 1 except for the time dummy, which doubles in size and becomes highly significant, capturing the general upward trend of returns to education over the nineties. This suggests that its insignificance in column 1 was probably due to the collinearity with the public sector interaction term.

Column 3 presents the results of the IV estimation which deals with the endogeneity of the education related variables. These are instrumented with average distance from primary and secondary school, as well as the squared term of the latter and distance to secondary school interacted with the time dummy. This is the set of instruments that maximises the predictive power of the instruments and follows the results in Table 6.3 (which also employs the same other controls as in Table 6.4). The IV has the effect of raising the education coefficients (cf. columns 3 vs. 2), confirming the hypothesis of endogeneity biasing the coefficients downwards. That is the case if higher returns to education generate incentives to acquire a higher level of education, thus raising both *skilled employees* and *population edu*. It is especially the latter coefficient to be affected by the endogeneity bias: it grows by 80% from column 2 to 3 (versus a 13% growth in the *skilled employees* coefficient). The other variables are little affected although they are estimated more precisely than before (but remain not significant).

[131] Note that without including this interaction term the Public sector coefficient becomes even less significant (not shown here).

[132] I test this effect by taking away skilled employees from the regression in column 1. The coefficient of Public sector doubles and becomes significant at the 14% level (results available upon request).

In column 4 I use returns to primary education estimated through (5.4) ($\beta_{OLS}^{PRIMARY}$) as the dependent variable. This has the effect of increasing both education coefficients substantially, suggesting that a higher level of education impacts on returns to education (via both demand and supply effects) mainly through returns to primary education. The increase in the absolute size of the *pop education* coefficient (350%) is considerably larger than that of the *skilled employees* one (62%, cf. columns 4 vs. 3). This larger increase in *pop education* may be part of the reason why returns to primary have grown proportionately more than the returns to other schooling levels (i.e. returns to schooling have become less convex) during the 1990s, as found in chapter 5. Interestingly, the time dummy becomes insignificant, which implies that all of the increase in returns to primary education over the nineties is explained by the regressors in the model.

One potential concern with these results is that they may be driven by the choice of specification to estimate the β coefficients used as dependent variables in (6.1). In columns 5 and 6 I use β estimated through two different methods as dependent variables, i.e. returns to education estimated via separate regressions for each district ($\beta_{OLS}(sep)$, column 5) and returns to education estimated with community fixed effects (β_{FE}, column 6). In particular the use of β_{FE} poses a great challenge for the identification of the effects on returns. Community fixed effects represent a severe control set given the individual data available: in 2000 for example I estimate the β coefficients on 2,606 individual employees across 796 communities. Despite this challenging identification conditions the coefficients of the education variables in both columns 5 and 6 are not significantly different from those in column 3. In fact in the specification with β_{FE} the coefficients of *telephone distance* and *urban employment* become larger (in absolute size) and significant. This suggests that there are some time invariant community characteristics which purge part of the negative effects of access to technology and of the positive effects of *urban employment* on returns to education. The use of β_{FE} controls for these effects making the technology and urban coefficients significant.

The results are also robust to using *employees edu* as the measure of supply of education (column 7) as well as to excluding the outlier Kapchorwa district from the analysis (column 8). This exclusion reduces the *population edu* coefficient making it significant only at the 15% level and increases the *urban employment* coefficient which becomes significant. When I exclude *population edu* the coefficient of *skilled employees* decreases by more than 50% (column 9). This suggests that failing to control for the (positive) demand side effects of education on returns biases downwards the coefficient of skills supply, which captures part of the demand side effects as well. This is a further confirmation that demand and supply factors appear to operate simultaneously in determining returns to schooling. Finally following Flabbi et al (2008), I also estimate the same regression as in (6.1) through weighted least squares using the inverse of the standard errors of the returns to education coefficients as weights. This should account for the fact that the β are estimated with different levels of precision in (5.3). Estimates obtained through this weighted method change little compared to the unweighted estimation, except for the coefficient of *population edu* which becomes not significant at the standard levels (column 10).

The results in Table 6.4 are robust to the use of different dependent variables and regressors as well as to different estimation methods. However they still do not take into account the possible effects of the drastic economic reforms that the Ugandan government promoted between the first and the second period of analysis. The next section aims to incorporate these effects into the analysis.

6.4. Trade, economic reforms and returns to education

Uganda underwent substantial pro-market reforms during the period considered. With the help of World Bank sponsored Structural Adjustment Programme (SAP), the economy was considerably liberalised starting in 1987 with the liberalisation process intensifying during the nineties. This process included the downsizing of the public sector, the privatisation of state owned enterprises, and measures aimed at lifting constraints to trade both domestically and internationally. These measures along with improvements in transport infrastructure have led to an increase in Uganda's trade both across district and across borders. Such increases in trade are likely to have had a relevant impact on labour market inequalities, and the rest of the empirical analysis is devoted to testing for such impacts. Let us consider international and domestic trade in turn.

6.4.1. International trade liberalisation

International trade liberalisation was an important component of the Ugandan reform process in the 1990s. Table 6.5 shows that between 1994 and 2000 Uganda halved the average tariff rate vis-à-vis the rest of the world and more than halved the one vis-à-vis the other members of the Common Market for Eastern and Southern Africa (COMESA). Maximum tariff rates were reduced dramatically as well. Using actual trade data Rudaheranwa (2005) calculates that the average effective rates of protection due to applied tariffs fell from 35% in 1994 to 18% in 2001. Again, this liberalisation coupled with the improvements in transport infrastructures and reduction in border-post transit times contributed to a large increase in both imports and exports during the nineties. Imports in particular grew by 50% between 1994 and 1999 (see Figure 6.2). And this figure does not likely document the magnitude of the total increase in trade between 1992 and 1999, which is the period under consideration here, as a substantial jump in international trade occurred between 1992 and 1994. Trade with Kenya, which is Uganda's main trading partner, boomed during 1993 and 1994 as political relations stabilised after the two countries nearly went to war in 1992. A liberal and buoyant foreign exchange market and liberal immigration procedures by both countries enacted in those years facilitated the free flow of goods across the border.[133]

Table 6.5: Uganda's import tariff rates between 1994 and 2000

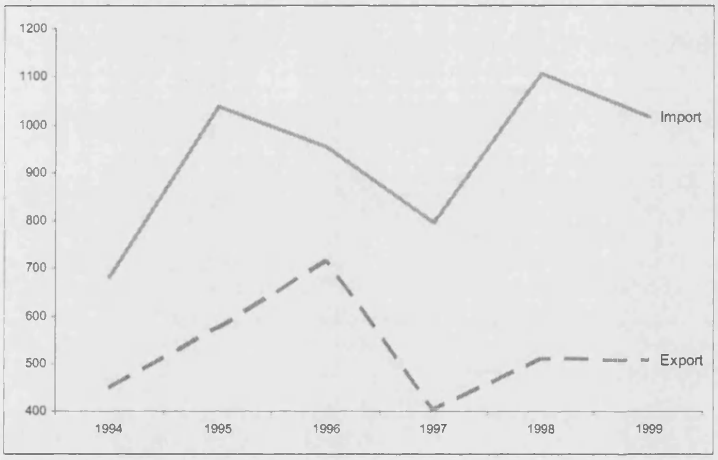
	Average	Std. Dev.	Min	Max
WORLD				
Tariff rate 1994	17.34%	9.09	0	60
Tariff rate 2000	9.36%	5.42	0	15
COMESA				
Tariff rate 1994	9.78%	9.12	0	66
Tariff rate 2000	4.25%	2.11	0	6

Source: UNCTAD Trade Analysis and Information System

[133] Newspapers in 1994 reported that "commercial traffic is so heavy that existing customs personnel are unable to cope with it and hundreds of vehicles loaded with goods spend up to a week awaiting customs clearance" (Ojulu, 1994). And residents of the Uganda-Kenya border town of Busia defined the volume of trade across the border in that year as unprecedented in the past 20 years.

What type of impact on the labour market should one expect from such a rise in international trade in a country like Uganda? The basic (2 goods, 2 factors, 2 countries) H-O framework predicts that trade would increase demand for relatively abundant factor of production. In a simple two-factor model with skilled and unskilled labour, the relatively abundant factor in developing countries would be the latter. Thus it may seem somewhat puzzling that the literature has mainly found increases in the relative demand for skilled labour in developing countries in periods of trade liberalisation.[134]

Figure 6.2: Ugandan exports and imports (US$ million, current prices), 1994-1999



Source: COMTRADE

This may not be a paradox once the assumptions of the basic H-O model are relaxed in two main ways (Anderson, 2005). First, adding natural resources as a further factor of production to the model may change the balance of relative abundance in those developing countries relatively rich in natural resources. Second, if one relaxes the H-O assumption that all countries have equal access to the best available production technology, then greater openness to that technology for countries that did not access it before liberalisation may increase the relative demand for skilled labour even in low-income developing countries. However, these considerations should have limited application to a LDC like Uganda. In fact its abundant factor would be unskilled labour even in a three-factor model. Labour-land ratio is relatively high in Uganda, especially by African standards. In terms of population density it ranks number 49 out of 191 countries of at least 1 million inhabitants; and it ranks

[134] See the review of some of this literature in Anderson (2005)

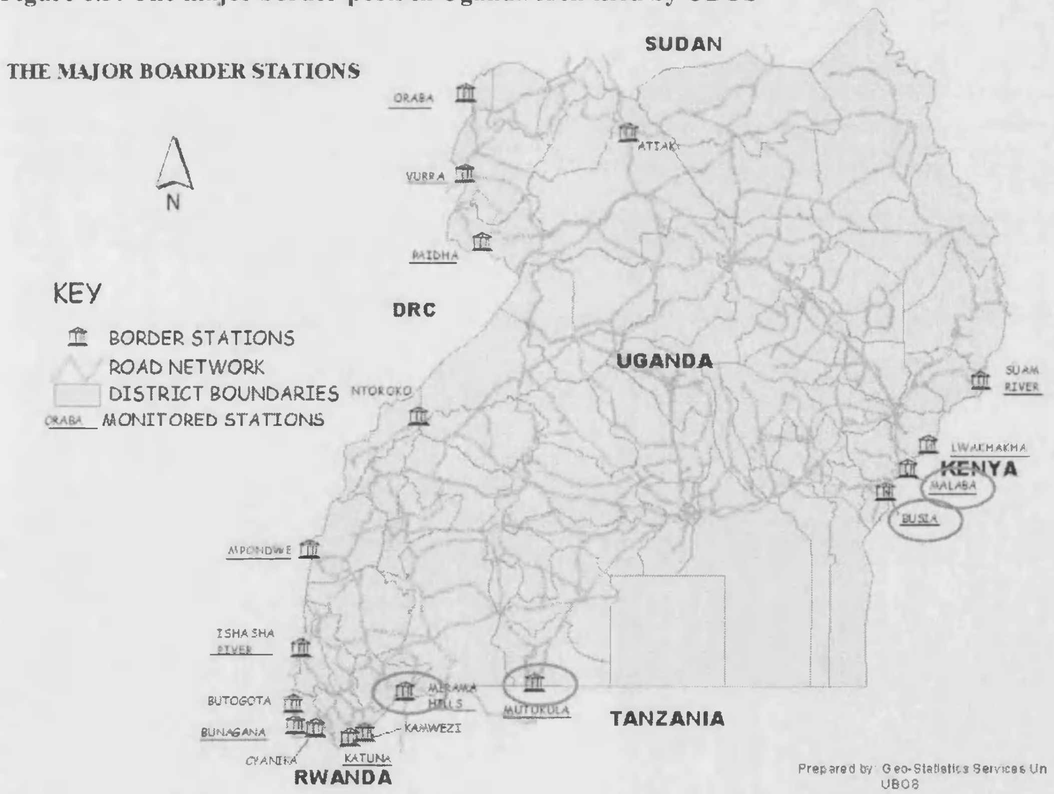
number 7 (out of 52) in Africa. Also, as highlighted by several authors (e.g. Soderbom et al., 2006; Baptist and Teal, 2008) poor Sub-Saharan African countries have limited access to the technological frontier even after trade opening. Given the lack of data it is difficult to identify the effects of trade opening on access to technology in Uganda. The technological proxy used above indicates that access to technology may raise somewhat returns to skilled labour, but it is not possible to assess the extent to which trade opening has facilitated (or not) this access. In any instances controlling for technological factors should allow to separate the pure trade effects from the skilled-biased technological change (SBTC)-type effects. Given this discussion, one would expect that increased international trade reduced labour market inequality in a country like Uganda.

The district-level analysis offers one way to examine the extent to which trade opening and market reforms may have affected the skilled-unskilled wage gap. This is not the first study to look at labour market inequalities at the district level in a period of trade opening. Taylor (2006) uses a relatively small number of regions (10) for a relatively long period of time (15 years) to examine skilled-unskilled wage gap within the UK. In this way identification comes mainly from year-to-year changes *within* regions. As I use a larger number of areas (34) and two time periods, identification in this study comes mainly from relative changes over time *between* districts. In this way this analysis may be able to identify the effects of changes in trade and other economic policies over time with more precision. Perhaps more importantly this chapter makes some effort to develop a strategy that adequately identifies the impact of trade opening on returns to education. In particular, the idea is to compare the changes in returns in districts which are likely to have experienced increases in international trade to different extents.

One natural way to do so would be to use the actual data on cross-district trade as a proxy for a district's exposure to trade. However trade data at the district level is not available, nor does the Ugandan data allows employing another method used in the literature to identify the exposure to trade, i.e. an industry-level analysis using the degree of tariffs' reduction as proxy for the intensity of trade. This strategy has been employed in developing countries' contexts by Attanasio et al. (2003) for Colombia and by Topalova

(2005) for India. In particular, Topalova matches industry composition data at the district level with (exogenous) national level tariff reductions in the same industries. This exercise is not possible in this case as industry classification in the LSMS is not compatible with the tariffs' sectoral classification (from TRAINS), which is based on the Standard International Trade Classification. Instead I exploit the districts' location and that of the main border-posts to identify districts' relative exposure to trade. As a proxy for the intensity of trade I construct a dummy variable *Bpost*$_k$ (named *borderpost*) which takes the value of 1 for districts that host a major border-post or which are close to one (i.e. less than 50 km from its centroid).[135] The identification of the relevant border stations is based on the work of the Uganda Bureau of Statistics (UBOS), which lists the major road border-posts in 2005 (UBOS, 2006). These border-posts are indicated in Figure 6.3.

Figure 6.3: The major border-posts in Uganda identified by UBOS



Source: UBOS (2006)

[135] The main results of the paper do not change when constructing the dummy assigning the value of 1 to the major border-posts and the adjacent districts.

However, not all border-posts were equally operational during the nineties. In particular, those posts bordering Tanzania (to the South and to the East) and Kenya (to the East) were the only ones likely to experience a significant increase in international trade following the liberalisation. This is for three reasons. First, Kenya and Tanzania were among the largest trading partners of Uganda during the nineties (and they still are). Second, the majority of Uganda's international trade transited through Kenya and Tanzania, mainly through the ports of Mombasa and Dar-es-Salaam.[136] Importantly, most of this trade was (and still is) formal thus being actually affected by reductions in tariffs.[137] On the other hand, Uganda's trade with Sudan (to the North), DRC (to the West) and Rwanda (to the South-West) was very limited in the nineties, due to civil unrest in these countries or in the Ugandan districts bordering these countries (in the North and North-West). For these reasons, I consider only the border-posts with Kenya and Tanzania for the construction of the variable. Moreover, among those listed by UBOS I exclude those which were not active for most of the nineties.[138] This leaves four major border stations identified by UBOS (2006) whose name is circled in Figure 6.3: Busia, Malaba, Mutukula and Mirama Hills. I also add Kampala district to this list of road border-posts as it hosts the major airport in Uganda (Entebbe) through which all Uganda's international trade via air transited in the nineties (and still does). Moreover the city of Kampala has also an important border in the railway station for goods transported by train to and from Kampala.

In order to (indirectly) test the appropriateness of the choice of the *borderpost* variable as a proxy for the intensity of trade, let us perform a test using the approach developed by Nicita (2004) for Mexico based on the pass-through literature. The basic idea of this approach is to test the significance of the impact of *borderpost* on changes in prices of traded goods following the liberalisation. Let us take a simplified version of the model developed by Nicita (2004). Prices of imported goods at the district level can be expressed as the product of the international price, the exchange rate, the import tariff and transport costs:

[136] This information is based on personal communication with the UBOS staff.
[137] This naturally applies to imports; however also exports are likely to have been affected as tariffs within COMESA (which Kenya and Tanzania are both part of) were reduced; also there was an increased integration of the East Africa Community (EAC), again which both Kenya and Tanzania are members of.
[138] This identification is based on trade data and information provided directly by UBOS.

$$P_{gkt} = e_t PX^*_{gt}(1+\tau_{gt})TC_{gkt} \quad (6.2)$$

for each traded good g, district k and period t, where the asterisks denote variables expressed in foreign currency. In this framework the international price is given. For each good g, this translates into the empirical specification:

$$P_{kt} = \alpha_k + \rho_1(d_{2000} \times Bpost_k) + \rho_2 d_{2000} + \varepsilon_{dt} \quad (6.3)$$

where $Bpost_k$ is *borderpost* and is an inverse measure of transport costs. As tariffs have been decreasing (and transport connections are improving) over the nineties, the time dummy is effectively an indication of trade liberalisation. I test the specification (6.3) on three different goods (second hand shirt, wheelbarrow and hoe), which were both imported and produced domestically in the nineties, and whose import tariffs were cut between 1994 and 2000. The coefficient of the interaction between the time dummy and the border-post variable $Bpost_k$ should capture the extent to which prices are differently affected across districts according to their proximity to the main border-posts. Following Nicita (2004), ρ_1 is expected to be negative, as the pass-through effect (from international to district-level prices) also depends on transportation costs and local production could become more profitable when transport costs are high. If these are not the major borders through which the goods are imported into Uganda, the interaction effect should not be significant. As Table 6.6 shows (columns 1-4) ρ_1 is negative and significant for all goods (for hoe the coefficient is significant only at the 15% level, column 3). This result is unaffected by the inclusion of a set of controls that may capture co-determinants of prices, i.e. the districts' total population and the share of population in urban areas (column 4). To check whether the negative coefficient is not only picking up the location of those districts on the border, I test whether the same result holds even when using a dummy taking the value of 1 if the district is on the border with another country (*Border*). This variable has an insignificant effect on the price of second hand shirt (column 5), confirming that *borderpost* is identifying the major trading districts of Uganda.

Note that by construction this test is applicable only to imports but not to exports. However the *borderpost* variable is likely to capture the intensity of exports as well, as the exports transit through the same major border-posts as the imports (UBOS, 2006). In any instance the effect of the trade liberalisation process is captured mainly through the changes in imports as those are the direct consequence of the liberalisation. Moreover in 1999 exports were still 10% of GDP while imports represented over 20%. Thus it is mainly the impact of changes in imports that I test through the variable $Bpost_k$.

Table 6.6: The effects of distance from border-posts on changes in prices of a unit of traded goods, 1992-2000

	(1) Shirt	(2) Wheelbarrow	(3) Hoe	(4) Shirt	(5) Shirt
Borderpost	-0.129**	-0.282*	-0.076	-0.129**	
**d2000*	(0.050)	(0.170)	(0.052)	(0.049)	
d_{2000}	-0.027	0.638***	-0.056	-0.056	-0.058
	(0.044)	(0.162)	(0.050)	(0.047)	(0.073)
*Border** d_{2000}					-0.011
					(0.081)
Constant	7.818***	10.079***	8.164***	7.879***	7.819***
Controls	NO	NO	NO	YES	NO
Observations	71	69	72	71	71
Nr. of district	38	38	38	38	38
R-squared	0.201	0.451	0.150	0.215	0.113

*Robust standard errors (Huber-White method) in parentheses; dependent variable is the price (in Ugandan Schelling) of the product in log; * significant at 10%; ** significant at 5%; *** significant at 1%; all regressions include district fixed effects. Controls include the share of population in urban areas and districts' total population.*

I use *borderpost* to test for the effects of international trade on returns to education through an extended version of equation (6.1):

$$\beta_{kt}^{OLS} = a_k + b_1 \frac{Sup_{kt}^{S}}{Sup_{kt}^{Tot}} + b_2 \frac{Dem_{kt}^{S}}{Dem_{kt}^{Tot}} + b_3 Bpost_k d_{2000} + d_{2000} + \varepsilon_{kt} \qquad (6.4)$$

where the interaction between *Bpost* and the post-1992 time dummy d_{2000} identifies the effects of trade on returns to education. This is essentially a difference-in-difference specification, which tests whether changes in returns to education are different in the 'treated districts' (where the treatment is the increase in trade) relative to the control group.

As discussed above, both the simple (based on the 2x2x2 model) and the extended version of the H-O model would predict that $b_3<0$ in the case of Uganda.

6.4.2. Domestic trade integration

While there is data to measure the increase in Uganda's international trade, the rise in inter-district trade is more difficult to document in the absence of data on internal trade. However some pieces of evidence are consistent with a substantial increase in inter-district trade in the nineties. First, the monopoly commodity marketing boards, such as the Produce Marketing Board were dismantled in 1992. These boards *de facto* controlled trade and production by guaranteeing local agricultural producers a minimum price for their crops. Their elimination substantially freed inter-district trade in agriculture (van der Geest, 1999 p. 132). Second, explicit restrictions to cross-district product movements were removed in 1993 (World Bank, 2008). Moreover, the process of domestic trade liberalization included also an attempt to open rural areas to markets through improvement of infrastructure. (DENIVA, 2005). This improvement contributed to a considerable reduction in overland transport costs during the nineties. Rudaheranwa (2005) calculates that the implicit taxation on Ugandan exports relating to transport costs declined from over 31% in 1994 to about 24% in 2003 for 40-foot containerized exports. This was coupled by transport policy reforms, such as the commercialisation of Ugandan Railways in the early 1990s, which improved efficiency in rail transportation and allowed the railway to compete with road transportation (Rudaheranwa, 2005).

Table 6.7: Measures of price variation across Ugandan districts, 1992-2000

	Year	Mean	Std. Dev	Min	Max	range/mean	StDev/mean
Soap	1992	685	58	600	925	0.47	0.09
	2000	818	64	772	1,083	0.38	0.08
Hoe	1992	3,536	518	2,165	4,838	0.76	0.15
	2000	3,277	296	2,981	4,583	0.49	0.09
Wheelbarrow	1992	26,389	9,761	1,500	60,000	2.22	0.37
	2000	42,178	6,717	20,000	54,643	0.82	0.16
Bicycle	1992	94,235	13,745	69,286	131,667	0.66	0.15
	2000	87,632	11,687	76,667	139,375	0.72	0.13

Source: author's elaboration on household survey data

As data on inter-district trade is not available in Uganda, one indirect way to check for the increase in this trade over the nineties is to look at the variation in prices across districts over time. This variation (as measured by the range/mean and the standard deviation/mean ratios) is lower in 2000 than in 1992 for commonly traded goods within Uganda such as those shown in Table 6.7. This finding is consistent with the price-converging effect of increased trade between districts during the nineties.

According to the H-O model, districts relatively endowed with skilled labour are supposed to experience an increase in the relative wage of skilled employees following a rise in inter-district trade. In order to test for this hypothesis, I extend equation (6.4):

$$\beta_{kt}^{OLS} = a_k + b_1 \frac{Sup_{kt}^{S}}{Sup_{kt}^{Tot}} + b_2 \frac{Dem_{kt}^{S}}{Dem_{kt}^{Tot}} + b_3 Bpost_k d_{2000} + b_4 S_{kt,1992} d_{2000} + d_{2000} + \varepsilon_{kt} \qquad (6.5)$$

where $S_{kt,1992}$ is the relative supply of skilled employees as measured by the share of skilled (i.e. at least primary educated) in total employees. The prediction from the H-O model in this case is that $b_4>0$. This test is similar in spirit to that of Michaels (2008), who examines the effect of increased trade between US counties (induced by road infrastructure development) on the skilled-unskilled wage gap.

Along these lines I also examine whether there is a differential effect of increased international trade on returns to education according to the initial level of skills. I add a double interaction term to specification (6.5):

$$\beta_{kt}^{OLS} = a_k + b_1 \frac{Sup_{kt}^{S}}{Sup_{kt}^{Tot}} + b_2 \frac{Dem_{kt}^{S}}{Dem_{kt}^{Tot}} + b_3 Bpost_k d_{2000} + b_4 S_{kt,1992} d_{2000} + \qquad (6.5')$$

$$+ b_5 (S_{kt,1992} \times B_k \times d_{2000}) + d_{2000} + \varepsilon_{kt}$$

with the hypothesis being that $b_5>0$ (i.e. the negative effect of international trade on returns to education is smaller in districts relatively endowed with skilled employees than in the others).

6.4.3 Results

The results of these regressions, reported in Table 6.8, confirm most of the basic theoretical predictions. Districts including a major border-post or located near to one experience a decrease in returns to schooling after the trade opening relative to the other districts, suggesting that international trade has increased the relative demand for unskilled labour in Uganda. Column 1 presents the results of specification (6.4), which adds *borderpost*$*d_{2000}$ to the specification in Table 6.4, column 3. I continue to instrument the education variables with the same set of instruments as in Table 6.4. The *borderpost*$*d_{2000}$ coefficient is negative and highly significant: during the nineties returns to education in districts more exposed to international trade decreased on average by 8 percentage point relative to the others. This is a significant difference considering that the average returns to education were between 13 and 15% in 2000 and they grew during the nineties by around 6 percentage points (see chapter 5). The negative effect of international trade on the skilled-unskilled wage gap is in line with the theoretical prediction from the simple H-O framework as well as from its extensions, as discussed in section 6.4.1. Also, the inclusion of the trade variable increases all the other coefficients, making *telephone distance* and *urban employment* highly significant (cf. Table 6.4, column 3). Therefore failure to consider the impact of trade causes an underestimation of the effects of the co-determinants of returns to education. The inclusion of the *borderpost* variable into specification (6.1) does not substantially affect the time coefficient (d_{2000}), which remains strongly positive (in fact it increases somewhat, cf. column 1, Table 6.8 with column 3, Table 6.4). This suggests that the observed increase in returns to education over the nineties is not likely to have been driven by trade opening, but rather by other factors.

The increase in domestic trade may be a good candidate in this respect, as shown in column 2, which tests specification (6.5). I take care of the endogeneity of the new variable *skilled employees*$_{1992}*d_{2000}$ by adding the interaction between distance to primary school and the post-1992 dummy to the list of instruments.[139] The results indicate that the b_4 coefficient in equation (6.5) is positive and significant as expected. This is consistent with

[139] Note that I also concomitantly take away the distance to secondary school squared term from the list of instruments so as to maximise the instruments power. The results are however similar leaving this instrument as well (results available upon request).

Michaels (2008), who finds that increased inter-county trade induced by the construction of an inter-state highway raised wage inequality in US counties relatively abundant in skilled labour. According to the coefficient in column 2 returns to education in a district with a share of skilled employment 10% larger than the average in 1992 would rise by 2.2 percentage points above the average between 1992 and 2000. The addition of this interaction term (*skilled employees*$_{1992}$*d_{2000}) substantially reduces the *skilled employees* coefficient, which also becomes less significant. This is possibly due to the high correlation between the two variables which may have taken up some of the *skilled employees'* effect. Interestingly, the inclusion of this interaction term makes the time dummy coefficient insignificant. Again it is not clear whether this is the result of the collinearity between the interaction and the time dummy or of the genuine effect of *skilled employees*$_{1992}$*d_{2000} explaining the residual part of the increase in returns over time.

Following specification (6.5') I next test for whether the impact of international trade on returns to schooling differs according to the district's relative level of education. The positive (although not significant) *Skilled empl*$_{1992}$**borderpost**d_{2000} coefficient in column 3 suggests that this may be the case (i.e. within those districts more exposed to trade more skilled districts experience a lower decline in returns relatively to the others). However the result does not hold when I put *skilled employees*$_{1992}$*d_{2000} back in the regression (column 4). This suggests that there is no differential impact of trade across districts on the basis of skills, i.e. b_5 in (6.5') is not significantly different from zero. In fact the addition of this interaction makes the *borderpost**d_{2000} coefficient insignificant probably due to the high collinearity with *Skilled empl*$_{1992}$**borderpost**d_{2000}.

An interesting question is whether any other theory than traditional trade can offer an alternative (or complementary) explanation for the effects of trade reforms on the skilled-unskilled wage gap observed in Uganda. The NEG core-periphery model (Krugman, 1991) is a useful framework to analyse the allocation of activity within a country following a trade shock. It predicts that firms desire to concentrate production near large consumer markets, as this allows them to economise both on transport and fixed production costs. A reduction in trade costs influences the location of economic activity by expanding the set of markets that firms can serve. This gives firms an incentive to move production to regions with relatively good access to foreign markets, such as border areas or port cities. Hanson (1997) shows that following NAFTA production within Mexico relocated towards the Mexico-US border and away from Mexico City.

Table 6.8: The impact of economic reforms on returns to education, 1992-2000

	(1)	(2)	(3)	(4)	(5)	(6)
	FE IV	FE IV	FE IV	FE IV	FE IV	FE IV
	β_{OLS}	β_{OLS}	β_{OLS}	β_{OLS}	β_{OLS}	β_{OLS}
Borderpost d2000*	-0.081***	-0.072***	-0.139***	-0.006		-0.074***
	(0.025)	(0.026)	(0.045)	(0.099)		(0.026)
Skilled employees	-0.302***	-0.188*	-0.289***	-0.185*	-0.081	-0.211**
	(0.080)	(0.108)	(0.082)	(0.107)	(0.098)	(0.097)
Population edu	0.095***	0.082***	0.095***	0.080***	0.048***	0.089***
	(0.024)	(0.030)	(0.024)	(0.029)	(0.017)	(0.028)
Telephone distance (x100)	-0.049**	-0.044***	-0.051***	-0.041**	-0.013	-0.045***
	(0.019)	(0.015)	(0.019)	(0.016)	(0.011)	(0.016)
Urban employment	0.142***	0.168***	0.149***	0.162***	0.116**	0.168***
	(0.054)	(0.051)	(0.056)	(0.054)	(0.055)	(0.051)
Skilled $empl_{1992}$*d_{2000}		0.218*		0.238*	0.273	0.200
		(0.133)		(0.148)	(0.179)	(0.124)
Skilled $empl_{1992}$* *Borderpost**d_{2000}			0.101	-0.112		
			(0.078)	(0.171)		
*Log Distance to Kampala**d_{2000}					0.011**	0.003
					(0.005)	(0.005)
d_{2000}	0.065***	-0.061	0.065***	-0.073	-0.160	-0.065
	(0.011)	(0.074)	(0.010)	(0.084)	(0.120)	(0.085)
Observations	68	68	68	68	68	68
Number of districts	34	34	34	34	34	34
R-squared	0.701	0.739	0.712	0.740	0.672	0.731
1st stage F-stat	6.03	2.36	5.85	2.20	2.17	2.19
Hansen J-stat	0.126	1.188	0.102	1.329	1.481	1.344

*Dependent variable is return to education as estimated through (5.3). Robust standard errors (Huber-White method) in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. F-stat is the statistics for the joint significance of the excluded instruments in the first stage; Hansen J-statistic is the over-identification test. Endogenous variables are skilled employees, population edu and skilled* $employees_{1992}*d_{2000}$. *Skilled employees and population edu are instrumented through distance to primary school, distance to secondary school, distance to secondary school squared and distance to secondary school x Post-1992 dummy (columns 1 and 3). In the specifications with Skilled* $empl_{1992}*d_{2000}$, *distance to secondary school squared is replaced by distance to primary school x Post-1992 dummy in the instrument set (columns 2 and 4-6). The 4 districts with a single observation are dropped in the IV estimations.*

In the case of Uganda the international trade liberalisation was mainly a unilateral liberalisation, as described above. Therefore increased access to markets during the nineties did not happen much along the country's international borders, but it rather occurred within the domestic borders through a reduction in internal trade costs. According to the core-periphery model this should lead to an increased concentration of economic activity near the main market(s). In the case of Uganda Kampala district represents by far the densest area in terms of economic activity in the country. But while the predictions of this framework are clear with respect to the (re-)location of economic activity within a country following a trade shock, they are less clear as far as the districts' skilled-unskilled wage gap is concerned. This will ultimately depend on the skill intensity of the productions the district specialises in. If the domestic goods that Kampala's market demands are relatively skill intensive then a reduction in internal trade barriers may lead to the concentration of relatively skilled economic activity in nearby districts. All else equal, this would generate an increase in the skilled wage premium in those districts and vice-versa in the districts further away. Empirically I try to identify any such effects through an interaction term between the distance (in log) to Kampala and the post-1992 dummy.[140] I include this variable in column 5 while excluding *borderpost*. This term is positive and significant hinting at the opposite situation of the example above, i.e. districts closer to Kampala experience a decline in returns relative to those further away. However re-including *borderpost* in the regression (column 6) wipes away this effect making the distance to Kampala coefficient insignificant (while *borderpost* remains negative and significant). This suggests that the effects of distance to Kampala on returns to education are driven by its role as a major border-post rather than as a major market.

6.4.4 Robustness

Although the main results of the analysis seem to be quite neat, I test their robustness to a variety of checks. Table 6.9 presents the results of these checks. First, I use returns to primary schooling calculated through the extended method as in (5.4) as the dependent variable (column 1). This does not change the sign of the variables, and confirms that returns to primary education are more responsive than returns to education to changes in the regressors. The positive effect of *pop education* is again particularly strong: a 10% increase in the average education of the adult population is associated with an increase by

[140] This distance is calculated with the method of the great circle distance between the city of Kampala and the centroid of the district

1.9 percentage points in returns to primary education. On the other hand the technological proxy and *urban employment* lose significance in this specification. The time dummy remains negative (relative to Table 6.4, column 4) but becomes significant, again possibly due to the collinearity with *Skilled* $empl_{1992}*d_{2000}$, whose coefficient is indeed much larger than in Table 6.8. Most results are robust also to using returns to education estimated via separate regressions for each district ($\beta_{OLS}(sep)$, column 2) and returns to education estimated with community fixed effects (β_{FE}, column 3) as dependent variables. In the former case the only relevant change is that *telephone distance* loses somewhat significance. In the case of β_{FE} the main change is that *skilled* $empl_{1992}*d_{2000}$ becomes insignificant at standard levels (although it remains positive). In column 4 I employ weighted least squares with the inverse of the standard errors of the β coefficients as weights to account for the fact that the β are estimated with different levels of precision. Again, estimates obtained through this method change little compared to the baseline estimation, except for the coefficient of *population edu* which loses some significance. The results are also robust to the exclusion of the outlier district Kapchorwa (column 5).

Until now the results for the effects of international trade on returns to schooling have implicitly relied on the assumption that the 'treatment' (i.e. increased international trade over the 1990s) was exogenous to returns to education as it was determined by districts' location. As all the major border-posts had been in place for some time before the beginning of the liberalisation process, this assumption seems reasonable. But what about if those districts had been selected to host the border-posts on the basis of certain unobserved characteristics (e.g. good infrastructure) which had also an effect on returns to education? The district fixed effects should account for these characteristics as long as they are time invariant. However to the extent that some unobserved shocks (e.g. an improvement in the infrastructure network to facilitate international trade) may affect only those districts hosting a border-post (or close to one) as well as returns to education, this may generate a bias in the $borderpost*d_{2000}$ coefficient. In order to control for this possibility I instrument $borderpost*d_{2000}$ with a dummy identifying all the districts bordering Kenya and Tanzania interacted with the post-1992 dummy. This variable is highly correlated with *borderpost* due to the way the latter has been constructed (i.e. considering only the active border-posts during the 1990s, which were on the borders with Tanzania and Kenya). This is confirmed by the high F-statistics (14.08) and partial R-squared (0.55) of the first stage regression for *borderpost* (not shown here). This instrument is also likely to be exogenous to returns to education, as also confirmed by the Hansen J-test which does not reject the null hypothesis of exogeneity of the instrument set (column 6).

Table 6.9: The effects of economic reforms on returns to education, 1992-2000, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE IV	FE IV	FE IV	FE IV (WLS)	FE IV	FE IV	LIML IV	FE IV
	$\beta_{OLS}^{PRIMARY}$	$\beta_{OLS}(sep)$	β_{FE}	β_{OLS}	β_{OLS}	β_{OLS}	β_{OLS}	β_{OLS}
*Borderpost** d_{2000}	-0.127**	-0.076***	-0.099**	-0.056**	-0.051***	-0.088***	-0.090***	
	(0.052)	(0.024)	(0.047)	(0.024)	(0.017)	(0.025)	(0.027)	
Border d2000*								-0.018
								(0.018)
Skilled employees	-0.072	-0.121	-0.303	-0.093	-0.126	-0.203**	-0.217**	-0.026
	(0.253)	(0.119)	(0.188)	(0.116)	(0.089)	(0.088)	(0.106)	(0.099)
Population edu	0.185***	0.094***	0.097*	0.051	0.059***	0.090***	0.093***	0.036*
	(0.067)	(0.027)	(0.052)	(0.033)	(0.021)	(0.025)	(0.030)	(0.020)
Skilled $empl_{1992}*d_{2000}$	0.927***	0.265*	0.229	0.251*	0.194	0.205*	0.210	0.321*
	(0.333)	(0.161)	(0.221)	(0.130)	(0.132)	(0.126)	(0.152)	(0.176)
Telephone dist. (x100)	0.005	-0.024	-0.130***	-0.038***	-0.036***	-0.049***	-0.051***	-0.016
	(0.042)	(0.017)	(0.041)	(0.014)	(0.012)	(0.014)	(0.016)	(0.012)
Urban employment	0.093	0.152**	0.281***	0.172***	0.175***	0.180***	0.183***	0.124**
	(0.140)	(0.061)	(0.089)	(0.053)	(0.046)	(0.055)	(0.057)	(0.062)
d_{2000}	-0.516***	-0.098	-0.065	-0.070	-0.041	-0.052	-0.055	-0.118
	(0.182)	(0.090)	(0.129)	(0.072)	(0.073)	(0.070)	(0.083)	(0.104)
Borderpost instrumented	NO	NO	NO	NO	NO	YES	YES	NO
Observations	68	68	68	68	66	68	68	68
R-squared	0.468	0.716	0.322	0.762	0.824	0.726	0.713	0.652
Nr. of districts	34	34	34	34	33	34	34	34
1st stage F-stat	2.36	2.36	2.36	1.87	2.67	1.86	1.86	2.08
Hansen J-stat	0.253	1.166	0.331	0.116	0.081	0.917	0.888	1.548

*Dependent variables are return to education as estimated through (5.3) and its variants. Robust standard errors (Huber-White method) in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. F-stat is the statistics for the joint significance of the excluded instruments in the first stage; Hansen J-statistic is the over-identification test. Endogenous variables are skilled employees, population edu, skilled employees*$_{1992}$***d_{2000} *and borderpost**d_{2000} *in columns 2 and 3. Skilled employees, population edu and skilled employees*$_{1992}$***d_{2000} *are instrumented through distance to primary school, distance to secondary school, distance to primary school x Post-1992 dummy and distance to secondary school x Post-1992 dummy. Borderpost**d_{2000} *is instrumented through a dummy for districts bordering Kenya and Tanzania. In columns 5 Kapchorwa district is excluded.*

Instrumenting *borderpost* has the effect of raising slightly its coefficient (cf. Table 6.8, column 2) which remains highly significant, like all the other variables except the time dummy (which continues to be insignificant, column 6). The inclusion of *borderpost* in the endogenous variable list reduces the F-statistics. To control that the results are not influenced by the relatively weak predictive power of the instruments, I use LIML estimation which should be more reliable in case of weak instruments, as discussed in chapter 3. The results are once again unaffected (column 7).

Finally, I check that *borderpost* is not capturing the effects of districts being located on the foreign border of Uganda rather than the actual effects of trade. In order to do that I use a different proxy for the intensity of trade, i.e. a variable identifying the districts on any national borders (i.e. including also those bordering Sudan, DRC and Rwanda) interacted with a post-1992 dummy ($border{*}d_{2000}$). The fact that this variable is not significant (column 8) adds confidence to the claim that *borderpost* identifies those districts more exposed to Uganda's international trade in the 1990s, which instead did not occur much through the country's Northern and Western borders.

6.4.5 Migration

As mentioned above, to the extent that a higher premium for educated labour attract higher skilled employees, in a context of unrestricted labour mobility the results may be biased by reverse causality. Another way to tackle this potential problem other than IV estimation is to control for inter-district migration. Data on this is available only in the 2000 survey, therefore I use proportionate changes in returns to education between 1992 and 2000 as the dependent variable. The regression reads as:

$$\Delta\beta_k^{OLS} = a + b_1\Delta S_k + b_2\Delta D_k + b_3 Bpost_k + b_4 S_{k,1992} + b_5 Mig_{k,92} + \varepsilon_k \qquad (6.6)$$

where the terms in Δ indicate proportionate changes, so that: $\Delta\beta_k^{OLS} = (\beta_{k,2000}^{OLS} - \beta_{k,1992}^{OLS}) / \beta_{k,1992}^{OLS}$; S_k and D_k are standard factors proxying for relative supply and demand of skilled labour, and $Mig_{k,92}$ is the share of skilled migrants (i.e. at least primary educated) in total adult population in district k who immigrated between 1992 and 1999. It is useful to examine this type of analysis graphically. Figure 6.4 plots the change in returns to education between 1992 and 2000 ($\Delta\beta_k^{OLS}$) against the initial level of returns (β_{1992}^{OLS}). The figure suggests that there is some tendency towards convergence in returns to education across districts over the period considered. In quadrant a, the mild negative relation is influenced by Kapchorwa district (number 14), which is the outlier identified above. Once this district is dropped, a convex relationship between proportionate change in returns and the level of returns in 1992 emerges – see quadrant (b). The convergence is apparent also when dropping the other influential district Nebbi (number 49) – quadrant (c). As in the case of regions, the convergence pattern is specific to returns to education, but not necessarily to other labour market characteristics, such as experience (quadrant d). Thus there seem to be some factors at play which tend to equalise returns to education across space over time. This may not be surprising given the increased mobility of goods in a period of market liberalisation.

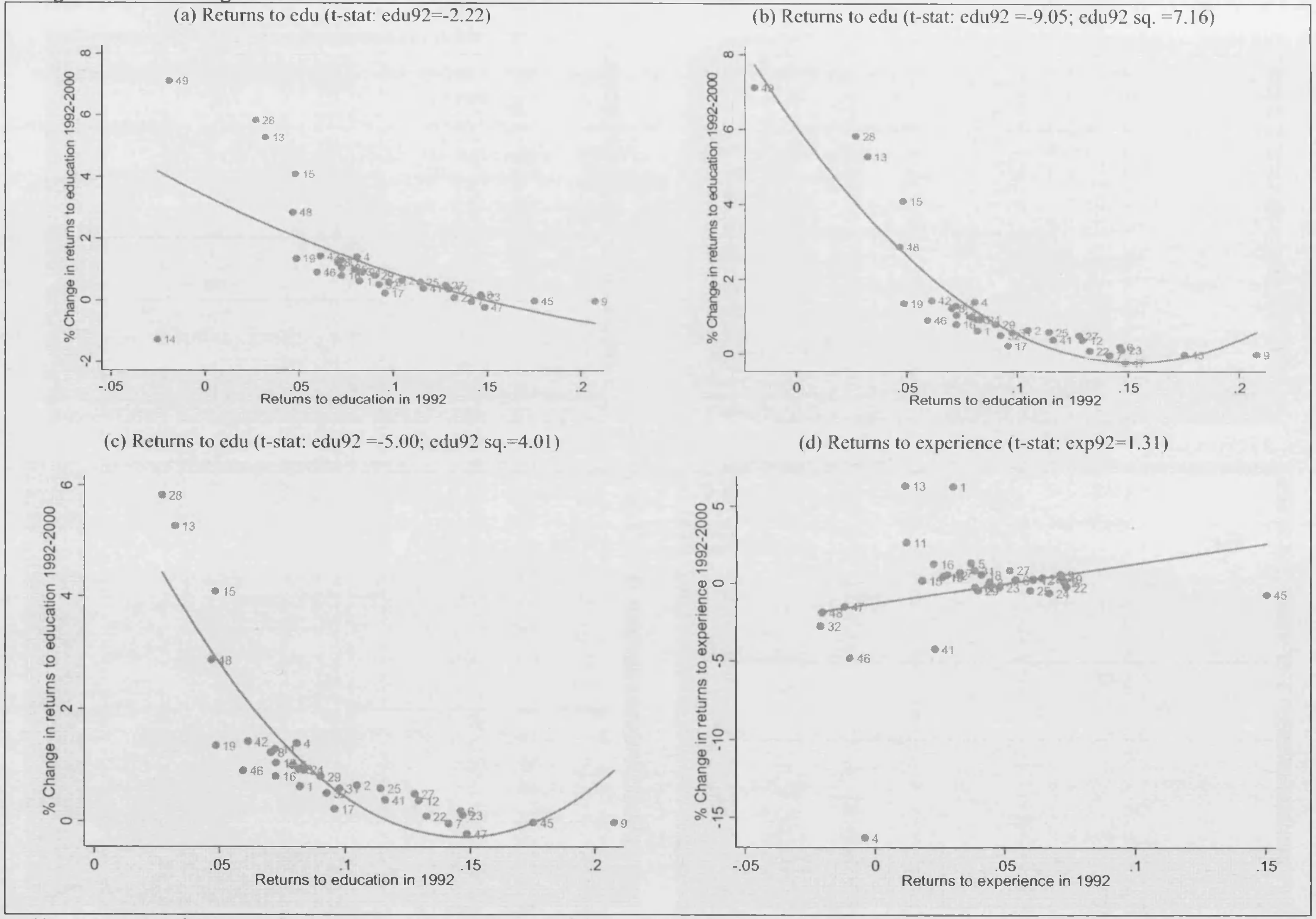
I use IV estimation to implement specification (6.6) employing the same set of instruments as before adapted to the first difference estimation performed here. The results of this analysis are presented in Table 6.10, which confirms the effects of the main variables even when including migration controls. In column 1 I include all the main variables measured in proportionate change (except *Borderpost* which is a time invariant dummy). *Borderpost* exerts a negative effect on proportionate changes in the returns to schooling and so does the change in *Skilled employees.* Also *ΔPopulation edu* and ΔUrban employment maintain their positive and significant effect on returns, while *ΔTelephone distance* exerts a negative and significant effect. The results change little when including the share of skilled employees in 1992, which has a positive but not significant effect on returns to schooling (column 2). This inclusion makes the changes in skills supply insignificant, suggesting that the main effects of the skill supply on returns to education are explained by the initial level of skills rather than by its changes. In line with the results above, relatively highly skilled districts experience a greater rise in returns to education than the other districts. Importantly these results are robust to the inclusion of $Mig_{k,92}$ (column 3), which has a U-shaped relationship with returns to education. The negative sign on the linear term of the migration variable is consistent with the idea that a higher skilled immigration rate depresses returns to schooling as it raises the supply of skills. The inclusion of the migration controls raises substantially the absolute size of the *ΔSkilled employees* and *ΔPopulation edu* coefficients, consistently with the idea that skilled migration occurs towards districts with relatively high returns to education. These results suggest that as the regressions above fail to control for migration they may generate lower bound estimates of the effects of the educational variables on returns to education (even in the context of IV estimation).

Table 6.10: The determinants of changes in returns to education, 1992-2000

	(1)	(2)	(3)
	IV	IV	IV
	$\Delta\beta_{OLS}$	$\Delta\beta_{OLS}$	$\Delta\beta_{OLS}$
Borderpost	-3.706**	-3.131**	-3.570**
	(1.471)	(1.428)	(1.515)
Δ skilled employees	-4.843**	-1.761	-4.286
	(2.174)	(3.427)	(3.171)
Δ population edu	10.904**	7.311	15.304**
	(5.334)	(6.314)	(7.217)
Δ telephone distance	-0.482**	-0.393**	-0.153
	(0.229)	(0.182)	(0.300)
Δ urban employment	4.718**	5.026**	5.253**
	(2.166)	(2.409)	(2.322)
Skilled employees$_{92}$		12.124	3.864
		(11.921)	(9.336)
Mig$_{92}$			-79.800
			(51.227)
Mig$_{92}$ squared			354.779
			(237.627)
Observations	34	34	34
R-squared	0.366	0.388	0.457
1st stage F-stat	3.21	2.30	2.95
Hansen J-stat	1.119	0.047	0.003

*Dependent variable is the proportionate change of return to education as estimated through (5.3). Robust standard errors (Huber-White method) in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. F-stat is the statistics for the joint significance of the excluded instruments in the first stage; Hansen J-statistic is the over-identification test. Endogenous variables are Δskilled employees, Δpopulation edu, skilled employees$_{1992}$. They are instrumented through Δdistance to primary school, Δdistance to secondary school, distance to primary school in 1992 and distance to secondary school in 1992. In column 5 Kapchorwa district is excluded.*

Figure 6.4: Convergence in returns to education across districts over time



Note: returns to education and returns to experience are computed through the OLS estimation in (5.4); district numbers are explained in Table 6.1

6.5. Conclusions

In Uganda returns to education – an important measure of labour market inequality - have substantially increased during the nineties. This was the general trend for the country as a whole, as confirmed in chapter 5, although with substantial variation across regions and districts. By exploiting such variation, this chapter has examined the determinants of returns to education in a within country context. The results confirm the importance of demand and supply (of skills) factors in determining returns to education. In particular, returns are depressed by the supply of skills, and raised by the general level of education, the urbanisation rate, and the access to technology. These effects are robust to using a variety of dependent and independent variables, estimation methods and samples. They also appear to be causal in nature as confirmed by the results of the IV estimation which tackles the endogeneity of the educational variables through a set of instruments based on distance to primary and secondary schools.

The methodology employed in this work has also helped reconcile the widening skilled-unskilled wage gap following trade opening with the standard H-O model predictions of a relative rise in the returns to the (relatively) abundant factor of production, i.e. unskilled labour in the case of Uganda. In fact the analysis – which also controls for other factors influenced by trade integration, such as access to technology - suggests that international trade depresses returns to education in line with the theoretical expectations. On the other hand, the intensification of domestic trade across districts during the 1990s (another by-product of the reforms) appears to have increased returns in those districts relatively endowed with skilled employees. This effect seems to explain a large share of the rise in returns to education over the 1990s.

The findings in these chapters may also help shed some light on why the expansion of the supply of education in Uganda during the 1990s has not depressed returns to schooling and has been associated with substantial reductions of poverty, differently to the experience of other SSA countries (Bennell, 2002; Soderbom et al., 2006). First, although the growth in educational supply of employees has a negative effect on returns to schooling, the rise in education of the entire population increases returns through demand side effects. Second, the wage earning function in Uganda became less convex during the nineties, as returns to primary education appear to have been more responsive to changes in education than returns to other levels of education. In particular the positive effect of the average education of the adult population on returns to primary schooling (demand-side effect of education) is considerably larger than the negative effect of the skilled labour supply (supply-side effect). This differential effect may provide an explanation as to why the convexity of the earning function increased during the 1990s. The expansion of education during the 1990s is likely to have operated relatively more through demand-side than through supply-side effects particularly on returns to primary education. This has raised such returns proportionately more than the returns to other schooling levels (i.e. returns to schooling have become less convex), as found in chapter 5. The expansion of education along with the decrease in convexity may have generated more growth and poverty reduction than if the convexity of the earning function had not changed. This pattern is positive news for the UPE programme in Uganda, although it is not clear whether it will hold regardless of the level of average education of the population. Thus it will be important to continue to monitor the effect of changes in education on returns to schooling in the future.

Moreover, although education expanded rapidly in the 1990s in Uganda, the chaos of the earlier period of economic mismanagement and civil conflict had limited the educational expansion, thus educated labour was scarcer than in other SSA countries. Because of these reasons, the Universal Primary Education policy implemented by the Ugandan government in 1997 has the potential to generate earnings growth despite the associated expansion of the educational supply. It would be important to evaluate its effects as new data comes on stream also including self-employed into the analysis.

Finally, these findings suggest that a rapid convergence process in returns to education has been occurring across districts, which is possibly fuelled by increased movement of goods. This has increased returns in regions and districts relatively more endowed with skilled labour, which initially enjoyed lower returns (due to a larger skills' base). Whether this process will lead to an equalisation of labour market inequalities across districts or will generate higher returns in more skilled districts remains an open question.

Appendix 6.1

Figure A6.1: Ugandan districts in 1992 (by region)

Table A1: Description of district-level variables

Variable's name	Units	Description
Skilled employees	Percentage	District-wise share of wage earners who have completed an education equal or higher than the primary level (i.e. 7 years or more)
Population education	Years	District-wise average number of years of formal schooling of the working age population (i.e. 15 years or above)
Employees education	Years	District-wise average number of years of formal schooling completed by the wage earners
Telephone distance	Km	Average distance from the centre of each community within a district to the closest public telephone booth
Electronic assets	Percentage	District-wise total value of households' electronic assets over total value of durable assets. Electronic goods are defined as "Electronic Equipment e.g TV, Radio, Cassette, etc."
Electronic purchase	Percentage	District's value of purchases by households in the last 12 months of electronic goods over district's total income.
Public sector	Percentage	District-wise number of wage earners employed in state-owned enterprises and local public firms over total district's number of wage earners
Urban employment	Percentage	District-wise number of wage earners located in urban areas over total district's number of wage earners
Primary school distance	Km	Average distance from the centre of each community within a district to the closest primary school
Secondary school distance	Km	Average distance from the centre of each community within a district to the closest secondary school
Borderpost	Dummy	Value 1 if the district hosts a major border-post or is close to one (i.e. less than 50 km from its centroid); 0 otherwise.
Border	Dummy	Value 1 if the district is bordering with a foreign country; 0 otherwise.
Migration$_{92}$	Percentage	District-wise population with primary education or above that migrated into the district after 1992 over total adult population of the district

Source: World Bank LSMS on Uganda for 1992 and 1999/2000.

Concluding remarks

Developing countries have been undergoing important changes in the spatial structures of their economies in recent decades. One of the main messages of this thesis is that without a proper evaluation of these changes our understanding of the development process is incomplete. In the absence of such evaluation both the causes and the consequences of changes in spatial disparities may be ill understood. As these changes are likely to continue as the process of economic development unfolds, bringing a spatial perspective into the analysis of development patterns is a key component to ensure the adequacy of policy prescriptions. Each chapter has tried to do that from a different perspective focussing on relatively under-researched spatial transformations in developing countries.

The first major transformation considered is the evolution of rural-urban disparities in living standards as a country develops economically. Chapter 1 has shown that for Indian states in the Post-Independence period such disparities tend to shrink with economic development for low levels of income per capita and then they start growing again after a certain threshold of income. This pattern seems consistent with the one characterising countries at the very early stages of development (e.g. LDCs) in line with evidence by World Bank (2009, fig. 2a) on a cross-section of countries. This evidence further suggests that there is a tendency for rural-urban disparities to increase in the following transition from low to middle income and then to decrease again after reaching the middle-income stage. As most Indian states in the second half of the 20th century had income per capita typical of the bottom part of today's low income countries, this analysis may be representative of the very early stages of the economic transition. If that is the case, living standards in rural and urban areas may now diverge as Indian states develop and will eventually return to converge when states reach higher levels of income per capita. Two further findings of the chapter may help policy-makers prepare for (and eventually mitigate) these changes. First, land reform has a negative association with rural-urban disparities helping rural areas to catch up relative to urban areas. This is consistent with Besley and Burgess (2000) as well as with the results in chapter 3. Second, rural-urban disparities decrease with the degree of urbanisation again in line with the evidence in chapter 4. Hence to the extent that the relation between urbanisation and rural-urban disparities follows the same path as in the past, urbanisation may dampen the eventual future increase in disparities.

The urban counterpart of this analysis concerns how the growth in wellbeing differs within the urban sector, and in particular between small and large towns. In the absence of direct measures of income for Indian cities, I have used population growth as a proxy of wellbeing for urban areas, as argued by Glaeser et al. (1995). I find a tendency towards convergence in growth rates among Indian towns across all decades of the century. Smaller towns grow faster than larger ones, although this pattern holds until a certain size threshold after which growth rate becomes an increasing function of size. This finding contrasts with the concern in policy circles in India and elsewhere in the developing world that small and medium sized towns have been growing too slowly relative to the large ones. On the other hand very large cities are less subject to this negative size effect on growth than large ones, which may be of some concern especially with respect to the Indian mega-cities.

This concern is common to many developing countries, where the tendency towards urban primacy is more pronounced (Kim, 2007). Chapters 2 and 3 are a contribution in the analysis of the determinants of urban primacy in a developing country experiencing rural-urban migration. They have focused on the possible role of the rural sector in influencing urban primacy. Chapter 2 has developed a two city framework whereby changes in the rural sector (such as agricultural productivity) shift the urban labour supply curve. This shift affects in turn the equilibrium urbanisation level as well as the distribution of urban population across cities. Shocks to the agricultural sector which make the labour supply curve steeper would reduce the urbanisation level, while the effect on the distribution of population across the two cities depends on the shape of the two net wage curves. In particular the decreasing part of the net wage curve for city 1 needs to be flatter than that for city 2 by a certain proportion for a steeper labour supply curve to be associated with an increase in primacy. In other words, congestion costs should kick in more slowly relative to agglomeration forces in the larger than in the smaller city.

The empirical analysis in chapter 3 has applied this framework to the Indian states in the Post-Independence period. The results suggest that the elasticity of rural-urban labour supply (proxied primarily through the agricultural productivity of land) has a positive effect on urban primacy as well as on urbanisation rate. The causality of the relationships (i.e. from elasticity to primacy and from elasticity to urbanisation) is confirmed by the results of the IV estimation, using rainfall levels and land reform legislations as instruments for agricultural variables. These findings are among the first to support the idea that the conditions in the agricultural sector may tilt the balance of population across urban areas. According to the analysis reducing the intensity of the push factors in rural areas may help reducing (or decreasing the rate of growth in) the level of urban primacy. Analytically this strategy would have a similar effect on urbanisation as rural-urban migration restrictions, which also act on the labour supply curve making it steeper. However it would not have the adverse effect on the welfare of potential migrants as restrictions (which prevent migrants to move where their labour is potentially more productive) would do. This may lead to a possible revaluation of the role of agricultural policies as urban de-concentration policies alongside other policies based on the action on pull factors, such as development of poles in remote areas, and political de-congestion policies (e.g. movement of the capital city). At the same time the findings do not seem to contradict the stability of the city-size distribution implicit in the validity of the Zipf's law. In fact the elasticity of urban labour supply increases the dispersion only in the upper tail of the city size distribution (i.e. the gap between the larger and the smaller urban areas in the upper tail of the distribution widens), while it does not appear to affect the city size distribution as a whole.

But the links between rural and urban areas run in the opposite direction as well. Urban growth is likely to affect welfare in rural areas and chapter 4 is concerned with its effects on poverty in surrounding rural areas. Using data on Indian districts between 1981 and 1999, we find that urbanization has a significantly poverty reducing effect. This occurs mostly through second-round effects (i.e. backward linkages, rural non-farm employment, remittances, agricultural productivity, rural land prices and consumer prices) rather than through the direct movement of rural poor to urban areas. IV estimation used to control for reverse causation (such as that illustrated in chapters 2 and 3) suggests that the effect is causal (from urbanisation to poverty reduction), and that failure to control for causality biases the coefficient of urbanisation downwardly. Again these findings may yield implications for a number of ongoing policy debates. First, they may help re-assess the role of public investments in urban areas for poverty reduction. To the extent that urbanization may have poverty reducing effects on rural areas (as studies on rural poverty have often failed to recognise), urban investments may become an important complement to rural ones in poverty reduction strategies. Second, the findings question the appropriateness of the bias against rural-urban migration often justified on the basis that this may deplete rural areas causing them to fall further behind (the 'brain drain'). The relatively low rate of urbanisation of India itself may also be due to public policies which have not facilitated (and in certain instances even constrained) rural-urban migration (Deshingkar and Start, 2005). Third, to the extent that the benefits from urbanisation do not spill over to the very poor in rural areas as highlighted in chapter 4, specific actions may be needed to facilitate these rural dwellers to enjoy the benefits of urbanisation. Examples of these may include the development of the types of skills useful for an expanding urban sector; or the provision of capital to cover the fixed costs of rural-urban migration.

The transformations across space relevant in developing countries' contexts are not only those directly related to income and poverty variables. Inequality is another dimension of welfare which varies markedly both across space and over time, and has potentially important implications on the long-term prospects of developing countries.[141] Analysing inequality within sub-national units is particularly important as this represents the major component of total inequality (larger than the between-units component, as argued in Kanbur and Venables, 2005a and 2005b). Chapter 5 and 6 have examined the variation across sub-national units in Uganda in the 1990s of a popular measure of labour market inequality, i.e. returns to schooling for wage employees. Uganda provides an empirical setting complementary to the Indian one as it is has considerably lower levels of income per capita and a much smaller population and area than India. But similarly to India, Uganda underwent a substantial economic liberalisation process (and experienced sustained economic growth) during the nineties. Chapter 5 has shown that this process has been associated with rising returns to schooling for wage employees at the national level as well as in each of the four regions of the country. This finding holds for the private and public sector employees alike as well as for the urban and rural sectors. This increase in labour market inequality in an unskilled labour abundant country like Uganda during a period of trade liberalisation seems inconsistent with the predictions of traditional trade theory.

[141] See Lopez (2004) for a review of the inequality-growth nexus with particular reference to developing countries.

The district level analysis undertaken in chapter 6 may help reconcile the rising returns to schooling with the predictions of traditional trade theory. This chapter exploits the large variation in returns across districts over time to examine the determinants of returns to education in a within country context. It employs IV estimation to deal with the likely endogeneity of educational variables (i.e. the supply of skilled employees and the average level of education of the adult population), using a set of instruments based on the distance to primary and secondary schools. The results confirm the importance of demand and supply (of skills) factors in determining returns to education. In particular, returns are depressed by the supply of skills, and raised by the general level of education, urbanisation rate, and access to technology. The analysis further suggests that the increase in international imports following the trade opening in the early nineties depressed returns to education in line with the theoretical expectations from the standard H-O model. On the other hand, the intensification of domestic trade across districts during the 1990s (another by-product of the reforms of the nineties) appears to have increased returns in those districts relatively endowed with skilled employees. This effect seems to explain at least some of the rise in returns to education over the 1990s.

The findings in these chapters may also help shed some light on why in Uganda the expansion of the supply of education has not depressed returns to schooling and has been associated with substantial reductions of poverty, differently to the experience of other SSA countries (Bennell, 2002; Soderbom et al., 2006). First, although the rise in educational supply of employees has a negative effect on returns to schooling, the rise in education of the entire population increases returns through demand side effects. Second, the wage earning function in Uganda became less convex during the nineties, as returns to primary education appear to have been more responsive to changes in education than returns to other levels of education. In particular the positive effect of the average education of adult population on returns to primary schooling (demand-side effect of education) is considerably larger than the negative one of the skilled labour supply (supply-side effect). This differential effect may provide an explanation as to why the convexity of the earning function increased during the 1990s. The expansion of education during the 1990s is likely to have operated relatively more through demand-side than supply-side effects on returns to primary education. This has raised returns to primary education proportionately more than the returns the other schooling levels (i.e. returns to schooling have become less convex), as found in chapter 5. The expansion of education along with the decrease in convexity may have generated more growth and poverty reduction than if the convexity of the earning function had not changed. Moreover, although education expanded rapidly in the 1990s in Uganda, the chaos of the earlier period of economic mismanagement and civil conflict had limited the educational expansion, thus educated labour was scarcer than in other SSA countries. Because of these reasons, the Universal Primary Education policy implemented by the Ugandan government in 1997 has the potential to generate earnings growth despite the associated expansion if the educational supply.

The discussion above shows that introducing a spatial perspective into a developmental issue may help shed new light on existing economic analysis. The findings of such an approach could help reshape important policy discussions such as those around decongestion, spatial inequalities and rural-urban migration. Nevertheless I feel that the thesis has perhaps contributed to open as many questions as it has answered. Such questions can provide some directions for much needed further research. In what follows I briefly discuss a number of them that in my view represent particularly fruitful areas of research.

I think push factors should be given a much more prominent role in urbanisation analyses in developing countries. This thesis has shown one way in which some of them may matter, but there is a whole host of other natural and man made push factors whose effects we remain ignorant about. In the light of the increasing concern around climate change, there is still remarkably little knowledge on the likely influence of climatic shocks on urbanisation. Some exceptions exist, such as Barrios et al. (2006), but we lack robust empirical analyses within countries of the effects of climatic variation on rural-urban migration and on the urbanisation rate. The argument applies to policy shocks as well. India itself has experienced a number of important ones, such as the rural banking experiment analysed by Besley and Pande (2005), or the green revolution in the sixties (see Foster and Rosenzweig, 1996 for an analysis of its effects on returns to education). By changing the incentives for rural-urban migration for (some) rural dwellers, these shocks can also be used to look at how push factors eventually change the urbanisation patterns in developing countries. It would be particularly interesting to focus on two distinct aspects of these patterns. Following the findings of the thesis, it would be possible to test for the effects of shocks to the rural sector on urban primacy. A good candidate for such shocks is climatic change in sub-Saharan African countries. Are the worsening climatic conditions in African countries affecting the way in which urban systems are evolving? In particular, could they explain to some extent the lack of second tier towns in sub-Saharan Africa (the 'missing middle' issue)? A different but related issue is whether urbanisation mainly spurred by push factors leads to different types of urban areas than urbanisation mainly driven by pull factors. Again, this distinction may shed light on the determinants of living standards in cities. This type of analysis is data thirsty as it would require data at city level. As such data start to come on stream in developing countries, it is one worth considering.

More generally, it would be important to develop a better understanding of the determinants of city growth in developing countries. While several studies have focused on this question for developed countries (e.g. Glaeser et al., 1995, Glaeser and Shapiro, 2001, Shapiro, 2006, Rappaport, 2007), remarkably little analysis is available for developing countries (with de Mata et al., 2007 on Brazil being an exception). Given the importance of rural-urban migration in fostering city growth in many developing countries but not in developed countries, this lack of analysis is problematic as the findings from one set of countries are unlikely to be applicable to the other.

Although this thesis has not touched upon the issue of urban poverty, rising urban populations may imply that urban poverty could become in the future the main poverty issue in its own right (Ravallion et al., 2007). Rigorous research is needed to assess whether the growth of urban population entails a trade-off between rural and urban poverty reduction. Also, it would be important to understand whether the size of cities which accommodate urban population growth matters in the impact on poverty (e.g. is it better for poverty reduction that urban growth happened in small or large cities? And does it matter at all?).

The findings in this thesis suggest that the level of education affects labour market inequalities both (positively) through demand-side and (negatively) through supply-side channels. This simultaneous effect has important implications for the economic impact of any policies aiming to expand the education level of a country. However it is not clear how the relative demand and supply-side effects will play out when countries reach different average educational levels. As a number of countries in

SSA have embarked in programmes of education expansion it is important to start monitoring the effects of such expansion on returns to education. This will crucially influence the relative success of such programmes in raising incomes and reducing poverty. For example the effects of the Universal Primary Education policy implemented by Uganda in 1997 can start to be evaluated as the first batch of new primary educated children came out of primary school in 2004 and as new household surveys become available (the data of the last one carried out in 2005 was recently released).

In addition, it would be important to analyse the returns to education for the self-employed. This is a key category of workers in developing countries, especially in SSA, as it includes most of the agricultural workers. There is not much evidence on the shape of returns to schooling for this part of the workforce in developing countries although studies on high income countries suggest that they may differ substantially from those for wage employees (van der Sluis, et al., 2004). Having data that allow an analysis of the self-employed would be important for policy evaluation as well. The methodology presented in chapter 6 could be extended to look at whether public policies such as educational policies, trade reforms or infrastructure expansion have a differential impact on self-employed vis-à-vis wage employees.[142]

Finally, proper spatial economic analysis requires data often at a refined scale that is hardly available in developing countries. This can arguably be part of the explanation for the relative neglect of such analysis in the field of development economics so far. Night light data (of the type used in the introduction of the thesis) represents one of the most important advancement in this respect. Using all the necessary caveats and circumspection, this data can be used at very refined spatial scales. Matching it with other data such as socio-economic data from surveys and climatic data can prove an invaluable source of new much needed insights into the spatial distribution of population, economic activity and wellbeing.

[142] This could be done through a triple difference estimation strategy, i.e. including differences between the pre- and post-reform, between treated and non-treated (or 'less treated') populations and differences between self-employed and wage employees.

References

Abril, E.R. and B. Rogaly (2001). Migration and social relations: An annotated bibliography on temporary migration for rural manual work, IIAS/IISG CLARA Working Paper No. 16.

Ades, A.F. and E.L. Glaeser (1995). Trade and circuses: explaining urban giants, *Quarterly Journal of Economics*, 110(1): 195-227.

Aghion P., Burgess R., Redding S. and F. Zilibotti (2007). The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India, *American Economic Review,* forthcoming.

Agnihotri S, R. Palmer-Jones and A. Parikh (2002). Missing women in Indian districts: a quantitative analysis, *Structural Change and Economic Dynamics,* 13(3):285-314.

Alperovich G. (1992). Economic Development and Population Concentration, Economic Development and Cultural Change, 41: 63-74.

Anderson, E. (2005). Openness and Inequality in Developing Countries: A Review of Theory and Recent Evidence, *World Development*, Vol. 33 (7), pp. 1045-63.

Appleton, S. (2001). Education, Incomes and Poverty in Uganda, CREDIT Research Paper 01/22, University of Nottingham.

Arellano, M., and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 277–97.

Ashenfelter, O. and D.J. Zimmerman, (1997). Estimates of the Return to Schooling from Sibling Data: Fathers, Sons, and Brothers , *Review of Economics & Statistics*, Vol. 79, no. 1.

Attanasio, O., Goldberg, P. and N. Pavcnik (2004). Trade reforms and wage inequality in Colombia, *Journal of Development Economics*, 74(2): 331-66.

Au, C-C, and V.J. Henderson (2006a). Are Chinese cities too small?, *Review of Economic Studies* 73, pp. 549–76.

Au, C.-C. and J.V. Henderson (2006b). How immigration restrictions limit agglomeration and productivity in China, *Journal of Development Economics*, 80(2): 350-388.

Bairoch P. (1988). *Cities and Economic Development*, Chicago: Chicago University Press.

Banerjee A. V. and L. Iyer (2006). History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India, *American Economic Review* 95(4): 1190-1213.

Barrios, L. Bertinelli, and E. Strobl (2006). "Climatic Change and Rural-Urban Migration: The Case of Sub-Saharan Africa", *Journal of Urban Economics*, 60.

Barro, R. (1991). Economic growth in a cross-section of countries, *Quarterly Journal of Economics* CVI, pp. 407-44.

Baum, C., M. Schaffer and S. Stillman (2008). IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation, *Statistical Software Components from Boston College Department of Economics.*

Baumol, W.J. (1986). Productivity growth, convergence and welfare: what doe the long-run data show?, *American Economic Review* 76, pp. 1072-85.

Becker, G.S. and K.M. Murphy, (1992). The division of labor, coordination costs, and knowledge, *Quarterly Journal of Economics*, 107(4), pp. 1137–60.

Behrens, K. and A. Polo Bala (2006). Rent-seeking elites, regional income transfers, and urban primacy in sub-Saharan Africa, mimeo, University of Louvain.

Bennell, P. (2002). Hitting the target: doubling primary school enrolments in Sub-Saharan Africa by 2015, *World Development*, Vol. 30, pp. 1179–1194.

Berdegue, J.A., E. Ramirez, T. Reardon, and G. Escobar (2001). Rural nonfarm employment and incomes in Chile, *World Development*, 29(3), pp. 411-425.

Besley T. and R. Burgess (2000). Land Reform, Poverty Reduction, and Growth: Evidence from India, *Quarterly Journal of Economics*, 115(2): 389-430.

Besley T. and R. Burgess (2002). The Political Economy of Government Responsiveness: Theory and Evidence from India, *Quarterly Journal of Economics*, 117(4): 1415-1452.

Besley, T. and R. Burgess (2004), Can labour regulation hinder economic performance? Evidence from India, *Quarterly Journal of Economics* 119(1), 91-134.

Birthal, P. (2007). Linking Smallholder Livestock Producers to Markets: Issues and Approaches, mimeo, Indian Society of Agricultural Economics.

Blundell, R., and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics* 87: 11–143.

Bourguignon, F. and Morrisson, C. (1998). Inequality and Development: The Role of Dualism, *Journal of Development Economics*, Vol. 57(2), pp. 233-257.

Brueckner J.K., 1990. Analyzing Third World urbanization: A model with empirical evidence, *Economic Development and Cultural Change* 38(3): 587–610.

Brückner, M. and A. Ciccone, 2008. Rain and the Democratic Window of Opportunity, CEPR DP6691.

Burgess, R and Pande, R. (2005), Can Rural Banks Reduce Poverty? Evidence from the Indian Social Banking Experiment, *American Economic Review*, 95 (3), 780-795.

Canagarajah, S. and S. Thomas, (1997). Ghana's labor market (1987–92), Policy Research Working Paper 1752, The World Bank.

Card, D., (1999). The causal effect of education on earnings, in Ashenfelter, O. and Card, D., (Eds.), *Handbook of Labor Economics*, vol. 3, North-Holland, Amsterdam.

Cashin P. and R. Sahay, (1996). Internal Migration, Center-State Grants and Economic Growth in the States of India, IMF Working Paper 95/58.

Chandra Mohan Reddy, Y.V.R. (2000), *A Study of Livestock Markets and Marketing of Livestock in Rangareddy District of Andhra Pradesh,* MSc Dissertation, Department of Agricultural Economics, Acharya N.G. Ranga Agricultural University, Hyderabad.

CIA (2003). The World Factbook 2002.

Combes P.P., G. Duranton and H.G. Overman (2005). Agglomeration and the adjustment of the spatial economy, *Papers in Regional Science*, 84(3), pp. 311-49.

da Mata, D., Deichmann, U., Henderson, J.V., Lall, S.V. and H.G. Wang, (2007). Determinants of city growth in Brazil, *Journal of Urban Economics* 62(2), pp. 252-272.

Das Gupta M. (2005). Explaining Asia's Missing Women: A New Look at the Data, *Population and Development Review,* 31(3): 529-535.

Datt, G. (1995). Poverty in India 1951-1992: Trends and Decompositions, Policy Research Department, World Bank, Washington, mimeo.

Datt, G. and M. Ravallion (1996). Why have some Indian states done better than others at reducing rural poverty? World Bank Policy Research Working Paper 1594.

Datt, G. and M. Ravallion (1998), Farm Productivity and Rural Poverty in India, FCND Discussion Paper No. 42, International Food Policy Research Institute, Washington D.C.

Datt, G. and M. Ravallion (2002). Is India's economic growth leaving the poor behind?, *Journal of Economic Perspective* 16, 89–108.

Davis, J. and J.V. Henderson (2003). Evidence on the Political Economy of the Urbanization Process, *Journal of Urban Economics,* 53, 98-125.

De la Fuente, A., R. Domenech and J. Jimeno (2003). Human capital as a factor of growth and employment at the regional level. The case of Spain, final report to the EC.

Deichmann, U. F. Shilpi and R. Vakis, (2008). Spatial specialization and farm-non-farm linkages, Policy Research Working Paper 4611: The World Bank.

Demombynes, G., C. Elbers, J. Lanjouw, P. Lanjouw, J.A. Mistiaen, and B. Özler (2003), Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries, in van der Hoeven R. and A. Shorrocks (eds.), *Growth, Inequality and Poverty,* Oxford University Press, 2003.

DENIVA, (2005). Uganda trade liberalisation and its impact on poverty.

Deshingkar, P. (2005). Seasonal migration: How rural is rural?, ODI Opinions no. 52.

Deshingkar, P. and D. Start (2003). Seasonal Migration for Livelihoods in India: Coping, Accumulation and Exclusion, ODI Working Paper 220.

Dickey, H. (2007). Regional earnings inequality in great britain: evidence from quantile regressions, *Journal of Regional Science,* vol. 47 (4), 775–806.

Dore, R. (1987). *Taking Japan Seriously: A Confucian Perspective on Leading Economic Issues,* Stanford University Press: Stanford.

Durand, J., and D. S., Massey, 1992. Mexican Migration to the US: a Critical Review, *Latin American Research Review,* 27 (2), 3-42.

Duranton, G. (2006) Some foundations for Zipf 's law: Product proliferation and local spillovers, *Regional Science and Urban Economics,* 36, 542–563

Duranton, G. (2008), From cities to productivity growth in developing countries, *Canadian Journal of Economics,* 41(3), 689-736.

Duranton, G. and V. Monastiriotis (2002). Mind the Gaps: The Evolution of Regional Earnings Inequalities in the U.K., 1982–1997, *Journal of Regional Science,* 42(2), 219–256.

Duranton G. and D. Puga, (2004). Micro-foundations of urban agglomeration economies. In: Henderson V and Thisse JF (eds.) *Handbook of Regional and Urban Economics* 4: North-Holland, Amsterdam.

Eaton, J. and Z. Eckstein (1997). Cities and growth: Theory and evidence from France and Japan, *Regional Science and Urban Economics* 27(4-5): 443-474.

Elbers, C., P.cLanjouw, J. Mistiaen, B. Ozler and K. Simler (2005), Are neighbors equal? Estimating local inequality in three developing countries, in R. Kanbur and A. J. Venables (eds), *Spatial Inequality and Development*, Oxford: Oxford University Press.

Ellis, F., (1998). Household strategies and rural livelihood diversification, *Journal of Development Studies*, Vol. 35(1): 1-38.

Eswaran, M. A. Kotwal, B. Ramaswami, and W. Wadhwa (2008), How does poverty decline? Suggestive evidence from India, 1983-1999, mimeo.

Fafchamps, M. and F. Shilpi (2003).The spatial division of labor in Nepal, Journal of *Development Studies* 39(6): 23-66.

Fafchamps, M. and F. Shilpi, (2005). Cities and Specialization: Evidence from South Asia, *Economic Journal*, Vol. 115, April: 477-504.

Ferreira, F., Phillippe G. L. and J.A. Litchfield (2008). The rise and fall of Brazilian inequality: 1981–2004, *Macroeconomic Dynamics*, 12:199-230

Flabbi, L., S. Paternostro and E. Tiongson (2008). Returns to Education in the Economic Transition: A Systematic Assessment Using Comparable Data', *Economics of Education Review*, forthcoming.

Fleisher, B., Sabirianova, K., & X. Wang (2005). Returns to skills and the speed of reforms: Evidence from Central and Eastern Europe, China, and Russia. *Journal of Comparative Economics*, 33, 351–370.

Fleisher, B. M., and X. Wang (2005). Returns to schooling in China under planning and reform. *Journal of Comparative Economics.*

Foster, A and Rosenzweig, M.(1996), Technical change and human capital returns and investments: Evidence from the Green Revolution, *American Economic Review* 86(4), 931-953.

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

Frankema, E. (2006). A Theil decomposition of Latin American income distribution in the 20th Century: Inverting the Kuznets curve?, processed, University of Groningen.

Fujita, M., P. Krugman, P., and A.J. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press: Cambridge and London.

Gabaix X. (1999). Zipf's law for cities: An explanation, *Quarterly Journal of Economics* 114: 738–767.

Glaeser, E.L., J. Kolko, and A. Saiz (2001). Consumer city, *Journal of Economic Geography* 1(1): 27-50.

Glaeser, E.L. and D.C. Maré (2001), Cities and skills, *Journal of Labor Economics*, Vol. 19 (2), pp. 316-42.

Glaeser, E.L., J.A., Scheinkman, and A. Shleifer (1995). Economic growth in a cross-section of cities, *Journal of Monetary Economics*, 36(1), pp. 117-43.

Glaeser E.L. and J. Shapiro (2001). Is There a New Urbanism? The Growth of U.S. Cities in the 1990s., NBER Working Paper No 8357.

Glaeser E.L. and J.M. Shapiro (2003). Urban growth in the 1990s: Is city living back?, *Journal of Regional Science*, 43(1): 139-165.

Government of India, various years. *Statistical Census of India*, 1951, 1961, 1971, 1981, 1991, 2001.

Guiso, L., Sapienza, P. and L. Zingales (2004), Does local financial development matter?, *Quarterly Journal of Economics* Vol. 119, Issue 3, pp. 929 – 969.

Hanson, G. (1997). Increasing returns, trade and the regional structure of wages, *Economic Journal*, 107: 113-33.
Hanson, G. (1997). Increasing returns, trade and the regional structure of wages, *Economic Journal*, 107: 113-33.
Harris J.R. and M.P. Todaro (1970). Migration, unemployment and development: a two-sector analysis, *American Economic Review* 60(1): 126-38.

Hasan, R., D. Mitra and B.P. Ural (2007). Trade Liberalization, Labor-Market Institutions and Poverty Reduction: Evidence from Indian States, *India Policy Forum* 3, pp. 71-122.

Hayashi, F. (2000). *Econometrics*, Princeton: Princeton University Press.

Heckman, J. J. (1979). Sample selection bias as a specification error, *Econometrica*, Vol. 47, pp. 153–161.

Henderson V.J. (1985). *Economic theory and the cities*, New York: Academic Press.

Henderson V.J. (1988). *Urban Development: Theory, Fact, and Illusion*, Oxford University Press.

Henderson V.J. (2004). Urbanization and growth, Brown University, mimeo.

Henderson, V.J., A. Storeygard and D. N. Weil (2009). Measuring economic growth from outer space, Brown University, mimeo.

Henderson V.J. and H.G. Wang (2007). Urbanization and city growth: The role of institutions, *Regional Science and Urban Economics*, 37(3), pp. 283-313.

Hering L. and S. Poncet (2006). Market access and individual wages: Evidence from China, *Review of Economics and Statistics*, forthcoming.

Indian Ministry of Finance (1997), *Economic surveys 1996-97*, New Delhi.

Indian Ministry of Finance (2001), *Economic surveys 2000-01*, New Delhi.

Intergovernmental Panel on Climate Change (2001). *Climate Change 2001: Impacts, Adaptation and Vulnerability*. New York, Cambridge University Press.

Ioannides, Y.M. and H.G. Overman (2003). Zipf's law for cities: an empirical examination, *Regional Science and Urban Economics*, 33: 127–137.

Jacobs, J. (1969). *The Economy of Cities*, Random House: New York.

Jha, V. (2008). Trickle down effects of Inter State Migration in a period of high growth in the Indian economy, mimeo.

Joshi G.V. and N. Lobo (2003). *Rural-urban migration and rural unemployment in India*, Delhi: Mohit Publication.

Journal of Economic Geography (2005), *Special Issue on Spatial Inequality and Development*, 5(1).

Junius, K. (1996). Economic development and industrial agglomeration: An inverted U-curve, Kiel working papers 770, Kiel Institute of World Economics.

Kagundu, P. and O. Pavlova (2007). Gender wage differentials in Uganda: Evidence from the Uganda National Household Survey, Andrew Young School of Policy Studies Research Paper Series No. 07-26.

Kanbur, R. and A. J. Venables (2005a), *Spatial Inequality and Development,* UNU-WIDER Studies in Economic Development, Oxford University Press.

Kanbur, R. and A. J. Venables (2005b), Introduction, in R. Kanbur and A. J. Venables (eds) *Spatial Inequality and Development*, Oxford: Oxford University Press.

Kanbur, R. and X. Zhang (1999). Which Regional Inequality? The Evolution of Rural–Urban and Inland–Coastal Inequality in China from 1983 to 1995, *Journal of Comparative Economics* 27, pp. 686-701.

Katz, L. and A. David (1999). Changes in wage structure and earnings inequality, in Ashenfelter, O. and D. Card, (eds.), *Handbook of Labor Economics*, vol. 3. Elsevier, New York, pp. 1463–1555.

Kijima, Y. and P. Lanjouw (2003). Poverty in India during the 1990s: a regional perspective. World Bank Policy Research Working Paper 3141.

Kim, S. (2007). Spatial Inequality and Economic Development: Theories, Facts, and Policies, Commission on Growth and Development, Working Paper no. 16.

Knight, J., Shi, L. and L. Song (2006), The rural-urban divide and the evolution of political economy in China, in Griffin, K.B., Boyce, J.K., Cullenberg, S. and P.K. Pattanik (eds.), *Human development in the era of globalization*, Edward Elger.

Kochar, A. (2004). Urban influences on rural schooling in India, *Journal of Development Economics* 74, pp. 113– 136

Krishnan, P., T. G. Selassie, and S. Dercon (1998). The Urban Labour Market During Structural Adjustment: Ethiopia 1990–97, CSAE Working Paper 98.9, Centre for the Study of African economies, University of Oxford.

Krugman, P.R. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99, p.483-499.

Krugman P.K. and R. Livas Elizondo (1996). Trade policy and Third World metropolis, *Journal of Development Economics*, Vol. 49: 137-150.

Kuznets, S., (1955). Economic Growth and Income Inequality, *American Economic Review*, 1, Vol. XLV.

La Porta, R. and A. Schleifer, (2008), The unofficial economy and economic development, *Brookings Paper on Economic Activity*, forthcoming.

Lall S.V., H. Selod, and Z. Shalizi, (2006). Rural-urban migration in developing countries: A survey of theoretical predictions and empirical findings. Policy ResearchWorking Paper 3915,World Bank.

Lanjouw, P. and A. Shariff (2002). Rural Nonfarm Employment in India: Access, income, and poverty impact. Working Paper Series no 81. New Delhi: National Council of Applied Economic Research.

Leonardi, M. (2008). Product demand shift and wage inequality, University of Milan, mimeo.

Lewis, J.B. and D.A. Linzer (2005). Estimating regression models in which the dependent variable is based on estimates, *Political Analysis*, 13: 345-64.

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

Lin, S. (2005). International trade, location and wage inequality in China, in R. Kanbur and A. J. Venables (eds) *Spatial Inequality and Development*, Oxford: Oxford University Press.

Lopez, H. (2004). Pro-Poor Growth: A Review of What We Know (and of What We Don't*)*, mimeo. World Bank.

Lu, M. and Z. Chen (2006), Urbanization, Urban-Biased Policies, and Urban-Rural Inequality in China, 1987-2001, *Chinese Economy*, 39(3), pp. 42-63.

Lucas, R.E. (1990), Why doesn't Capital Flow from Rich to Poor Countries? *American Economic Review Papers and Proceedings* 80, 92-96.

Lucas, R.E. (2004), Life Earnings and Rural-Urban Migration, *Journal of Political Economy*, vol. 112, no. S1, pp. S29-S59.

Mattoo, A. Neagu I.C. and C. Ödzen (2008). Brain waste? Educated immigrants in the US labor market, *Journal of Development Economics* 87, pp. 255–269.

Matsuyama, K., (1992). Agricultural Productivity, Comparative Advantage, and Economic Growth," *Journal of Economic Theory* 58, pp. 317-334.

Michaels, G. (2008). The effect of trade on the demand for skill — Evidence from the interstate highway system, *Review of Economics and Statistics,* 90(4), pp. 683-701.

Miguel, E., S. Satyanath and E. Sergenti (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach", *Journal of Political Economy*, 112(4): 725-753.

Mincer, J. (1974). *Schooling, experience and earnings*. New York: National Bureau of Economic Research.

Moll, P.G. (1996). The collapse of primary schooling returns in South Africa 1960-90, *Oxford Bulletin of Economics and Statistics*, 58(1), pp. 185-209.

Moomaw L. and A.M Shatter (1996). Urbanization and Economic Development: A Bias toward Large Cities?, *Journal of Urban Economics,* 40: 13-37.

Nurske, R., (1953). *Problems of capital formation in underdeveloped countries*, Oxford University Press: New York.

Oberai, A.S., Prasad, P.H. and M.G. Sardana (1989). *Determinants and consequences of internal migration in India*, Delhi: Oxford University Press.

Ojulu, E. (1994), Cross border trade booms, African Business, December 1st.

Overman, H.G. and Y.M. Ioannides (2001). Cross-sectional evolution of the US city size distribution, *Journal of Urban Economics* 49: 543–566.

Overman, H. and A. Venables (2005). Cities in the developing world, Report to the

Department of Foreign International Development.

Ozler, B., G. Datt and M. Ravallion, (1996). *A Database on Poverty and Growth in India,* Washington: World Bank.

Park, A. (2007). Rural-urban inequality in China, in Yusuf, S. and T. Saich (eds.), *China urbanizes*, Washington: The World Bank.

Parthasarathy Rao, P., P.S. Birthal, P.K. Joshi and D. Kar (2004), Agricultural Diversification in India and Role of Urbanization, MTID Discussion Paper No. 77, IFPRI.

Planning Commission, (1993). Report on the Expert Group on the Estimation of the Proportion and Number of Poor, Delhi: Government of India.

Plantinga, A.J., Lubowski, R.N. and R.N. Stavins (2002). The effects of potential land development on agricultural land prices, *Journal of Urban Economics* 52, 561-81.

Playne, S., R.V. Solomon, J.W. Bond and A. Wright (2006). *Indian States: A Biographical, Historical, and Administrative Survey*, Asian Educational Services.

Puga D. (1998). Urbanization patters; European versus less developed countries, *Journal of Regional Science*, 38: 231-52.

Quah, D. (1996). Empirics for economic growth and convergence, *European Economic Review*, 40(6), pp. 1353-75.

Ranis, G. and C.H. Fei (1961). A theory of economic development, *American Economic Review* 51: 533–565.

Rappaport, J. (2007). Moving to nice weather, *Regional Science and Urban Economics,* 37 (2007) 375–398.

Ravallion, M. (2002). On the urbanization of poverty, *Journal of Development Economics* 68, 435-442.

Ravallion, M. and S. Chen (2007). China's (uneven) progress against poverty, *Journal of Development Economics,* 82, pp. 1– 42.

Ravallion, M., S. Chen and P. Sangraula (2007). New evidence on the urbanization of global poverty, World Bank Policy Research Working Paper 4199.

Ravallion M. and Q. Wodon (1999). Poor areas, or only poor people? *Journal of Regional Science*, 39(4): 689–711.

Rosen, K. and M. Resnick (1980). The Size Distribution of Cities: An Examination of the Pareto Law and Primacy, *Journal of Urban Economics*, 81: 165-86.

Rozelle, Scott, J. Edward Taylor, and Alan deBrauw (1999). Migration, Remittances, and Agricultural Productivity in China, *American Economic Review*, Vol. 89, pp. 287–91.

Rud, J.P. (2009). Electricity provision and industrial development: Evidence from India, manuscript, London School of Economics.

Rudaheranwa, N. (2005). Trade Policy and Transport Costs in Uganda, CREDIT Research Paper No. 06/09, University of Nottingham.

Safran Foer, J. (2002). *Everyhing is illuminated*, New York: Harper Collins.

Sahn, D., and D. Stifel (2003). Urban-Rural Inequality in Living Standards in Africa, *Journal of African Economies,* 12(1): 564–97.

Satterthwaite, D. (2004). The Under-estimation of Urban Poverty in Low and Middle-Income Nations, IIED Working Paper 14 on Poverty Reduction in Urban Areas, London: IIED.

Satterthwaite, D. (2005). The scale of urban change worldwide 1950-2000 and its underpinnings, Human Settlements Discussion Paper - Urban Change 3, London: IIED.

Satterthwaite, D. (2006). Outside the large cities: The demographic importance of small urban centres and large villages in Africa, Asia and Latin America, Human Settlements Discussion Paper - Urban Change 3, London: IIED.

Schaffer, M. (2007), XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models, *Statistical Software Components from Boston College Department of Economics*

Scott A. and M. Storper (2003). Regions, Globalization, Development, Regional Studies, Vol. 37: 579-593.

Sen A. (1992). Missing women. *British Medical Journal,* 304(6827):587-588.

Shapiro, J.M. (2006). Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital, *Review of Economics and Statistics*, 88(2), 324-335.

Soo, K.T. (2005). Zipf's law for cities: a cross country investigation. *Regional Science and Urban Economics* 35: 239–263.

Söderbom, M., F. J. Teal, A. Wambugu and G. Kahyarara (2006). The dynamics of returns to education in Kenyan and Tanzanian manufacturing, *Oxford Bulletin of Economics and Statistics*, vol. 68 (3), pp. 261-288.

Stark, O. (1980). On the role of urban-rural remittances in rural development, *Journal of Development Studies*, 16(3): 369–74.

Stark, O. and R.E.B. Lucas (1988). Migration, remittances and the family, *Economic Development and Cultural Change*, Vol. 36 (3):465-482.

Stock, J.H. and M. Yogo (2005). Testing for weak instruments in linear IV regression, in D.W.K. Andrews and J.H. Stock (eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, 2005.

Sutton, P.C., C.D. Elvidge and T. Ghosh (2007). Estimation of Gross Domestic Product at Sub-National Scales using Nighttime Satellite Imagery, *International Journal of Ecological Economics & Statistics* 8(S07): 5-21.

Thanh, H.X., D.T., Thu Phuong, N. Thu Huong, and C. Tacoli (2008). Urbanization and rural development in Vietnam's Mekong Delta, IIED Working Paper 14.

Taylor, K. (2006). UK Wage Inequality: An Industry and Regional Perspective, *LABOUR: Review of Labour Economics and Industrial Relations*, 20(1), 91–124.

Tokila, A. and H. Tervu (2007). Are there regional differences in returns to education for entrepreneurs and wage earners?, paper presented at the 47th Congress of the European Regional Association, Paris.

Topalova, P. (2005). Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts, NBER Working Paper No. W11614.

Tostensen, A. (2004). Rural-urban linkages in sub-Saharan Africa: contemporary debates and implications for Kenyan urban workers in the 21st century, Chr. Michelsen Institute working paper 2004:4.

Uganda Bureau of Statistics (2006). The informal cross-border trade survey report, August 2004 to December 2005, Kampala.

United Nations, (2004). *World Urbanization Prospects, 2003 Revision*, UN Population Division, New York.

United Nations, (2005). *World Population Prospects 2004*. UN Population Division, New York.

United Nations, (2006). World Urbanization Prospects: The 2005 Revision Population Database, http://esa.un.org/unpp/, accessed 17 August 2007.

United Nations, (2008). World Urbanization Prospects: The 2007 Revision Population Database, http://esa.un.org/unup/, accessed 23 May 2008.

van der Geest, W. and R. van der Hoeven (1999). Adjustment, Employment & Missing Institutions in Africa - The experience in Eastern and Southern Africa, ILO, Geneva.

Venables, A.J. (1996). Equilibrium locations of vertically linked industries, *International Economic Review*, 37,pp.: 341-359.

Venables A.J. (2004). Spatial disparities in developing countries: cities, regions and international trade, *Journal of Economic Geography*, 5, pp: 3-22.

Venables A.J. (2006). Shifts in economic geography and their causes, CEP Discussion Paper no. 767, London School of Economics.

Vollrath, D. (2007). Land Distribution and International Agricultural Productivity, *American Journal of Agricultural Economics*, 89(1).

Van der Sluis, J., M. van Praag and A. van Witteloostuijn (2004), Comparing the returns to education for entrepreneurs and employees, TI 2004-104/3 Tinbergen Institute Discussion Paper.

von Thünen, J. H., (1966). *Isolated State: an English Edition of Der Isolierte Staat*, Pergamon Press, Oxford.

Wheaton, W.C. and H. Shishido (1981). Urban Concentration, Agglomeration Economies, and the Level of Economic Development, *Economic Development and Cultural Change*, 30:1, p.17

White, H. (1980). A heteroscedastically-consistent covariance matrix estimator and a direct test for the heteroscedasticity, *Econometrica*, 48:817–838.

Williamson, J.G. (1965). Regional inequality and the process of national development, *Economic Development and Cultural Change*, 13:4: 3-84.

Wood, A. (2002). Could Africa be like America?, University of Oxford, mimeo.

World Bank, (2008). *World Development Report 2008: Agriculture for Development*, The World Bank: Washington DC.

World Bank, (2009). *World Development Report 2009: Spatial Disparities and Development Policy*, The World Bank: Washington DC.

Yang, D.T. (2005). Determinants of schooling returns during transition: Evidence from Chinese cities, *Journal of Comparative Economics* 33, 244–264.

Yemtsov, R.G. (2003). Quo Vadis: Regional Inequality and Poverty in Russia, 1992-2000, WIDER Discussion Paper DP2003/67.

Zhang, J., Y. Zhao, A. Park and X. Song (2005). Economic returns to schooling in urban China, 1988 to 2001, *Journal of Comparative Economics* 33, 730–752.