Urbanization, Structural Transformation and Rural-Urban Disparities in China and India*

Viktoria Hnatkovska† and Amartya Lahiri‡

March 2018

Abstract

Over the past three decades India and China have experienced rapid growth and structural transformation. Underneath this similarity however was one significant difference: rural-urban wage gaps declined in India, but widened in China. In both countries, the majority of these wage dynamics are unexplained by worker attributes. We formalize a two-sector-two-location model in which structural transformation and urbanization respond endogenously to productivity shocks. While the structural transformation effect widens the urban-rural wage gap, the urbanization effect reduces it. We attribute the contrasting wage gap dynamics in the two countries to the higher costs of urban relocation for workers in China.

JEL Classification: J6, R2

Keywords: Rural urban disparity, urbanization, structural transformation

*We would like to thank Parikshit Ghosh, Jesse Perla, Frank Portier and seminar participants at Ashoka University, ASU, Bank of Canada, Boston University, Delhi School of Economics, Indian School of Business, ISI-Delhi, Ohio State, Shiv Nader University, UBC, University of Illinois at Urbana Champaigne, the CMSG, the European Econometric Society Meetings and the SHUFE Macro workshop in Shanghai for helpful comments. Thanks also to Ruoying Wang for excellent research assistance. We are also grateful to Jingting Fan for kindly sharing the 2005 Chinese mini census data. Hnatkovska acknowledges research support from Social Sciences and Humanities Research Council of Canada, grant number 12R20471.

†Vancouver School Economics, University of British Columbia, 6000 Iona Drive, Vancouver, BC V6T 1L4, Canada. E-mail address: hnatkovs@mail.ubc.ca.

‡University of British Columbia and CAFRAL. E-mail address: amartya.lahiri@ubc.ca.

1 Introduction

We use individual level micro data from China and India to establish two data facts. First, the median urban-rural wage gap in China *increased* by 23 percentage points between 1988 and 2008. In India, by contrast, the corresponding median wage gap *decreased* by 66 percentage points between 1983 and 2010. Second, in both countries, individual worker attributes accounted for less than half of the observed changes in the urban-rural wage gaps. What can explain these contrasting wage gap dynamics? This paper is an exploration into this question.

A popular explanation for urban-rural wage gaps is based on sorting. As elaborated upon in Young (2013), this explanation operates through workers sorting into urban or rural occupations based on their measured and unmeasured attributes. In as much as unobserved attributes are important for urban occupations, sorting can account for urban-rural wage gaps that are typically observed in the data. In our context, the problem with this explanation is that it has to explain why urban-rural wage gaps widened in China while contracting in India. Put differently, the mechanism has to account for the contrasting dynamic behavior of the gaps rather than generate an urban-rural wage gap at a point in time. Moreover, the explanation has to apply to an environment in which relative sectoral productivities were evolving similarly in the two countries. Indeed we show that worker sorting likely played only a limited role in China and India.

Our explanation for the opposing wage gap dynamics is based on labor misallocation across locations due to the differential costs of relocating labor from rural to urban areas. We develop a two-sector two-location model of an overlapping-generations economy where labor mobility is costly both across sectors and across locations. Productivity growth induces factor movement from agriculture to non-agriculture and from rural to urban locations due to non-homothetic preferences. The rising demand for non-agricultural labor causes the urban-rural wage gap to rise because the non-agricultural sector is predominantly urban. The increasing relocation of labor to urban areas however causes the wage gap to fall.

For the model to match the increase in urban labor in the two countries observed in the data, the required relocation costs in China are significantly greater than the relocation costs in India. The higher relocation costs in China imply that the increase in urban labor supply is not large enough to overcome the rising urban wage induced by higher demand for urban labor. The urban-rural wage gap increases as a consequence. In India on the other hand, the relocation costs are lower which causes the positive urban labor supply effect to more than offset the urban labor demand effect. This causes the wage gap in India to fall. Counterfactual experiments on the model confirm that a key factor behind the widening wage gap in China is the restriction on migration within the country. We find that lowering the implied migration restrictions in China to the corresponding levels in India would generate a sharp 32 percentage point contraction in the wage gap in China along with an additional 9 percentage point decline of the rural share of the workforce between 1988 and 2008.

There is independent evidence in support of the higher labor relocation costs in China. China has an institutional structure called the Hukou system to regulate cross-location worker movement. Workers possessing rural hukous are not eligible for public services like education and health facilities in urban areas. This makes urban relocation expensive. India, by contrast, has no such restriction which makes urban relocation relatively cheaper.

We also test the basic mechanisms formalized in the model using cross-province data in China and cross-state evidence in India. As predicted by our model, we find that for a given urban share of the workforce, places with higher productivity growth have a larger urban-rural wage gap. This is the demand effect of productivity growth that our model (as well as most models of structural transformation) emphasized. On the other hand, we also show that, controlling for productivity growth, locations with greater urban employment shares have smaller urban-rural wage gaps. This is evidence for the supply-side effect of a rising urban workforce that the model's endogenous labor reallocation channel emphasized. We view these results as confirmation for the mechanisms formalized by the model.

Our focus on urban-rural disparities is motivated by three factors. First, there is a long history in the development literature of trying to understand the rural to urban movement of labor during the process of development. This literature is huge but probably the most influential is Harris and Todaro (1970). Second, policymakers in developing countries devote considerable attention and resources to managing the fortunes of vulnerable rural workers.[1] Our work tries to examine how vulnerable rural workers in developing countries truly are and the mechanisms that affect their economic evolution. Third, Young (2013) has recently shown that 40 percent of mean consumption inequality within countries is due to urban-rural inequality. We view Young's finding as suggestive of the importance of understanding the impact of development on rural workers.

Could one examine the dynamics of urban-rural gaps by studying the dynamics of sectoral gaps between non-agriculture and agriculture? This might, at first glance, seem the obvious approach since urban locations are primarily non-agrarian while rural locations are mostly agrarian. As we show in the paper, this is inappropriate both as a description of the data as well as a conceptual framework since it omits the within-location sectoral adjustment of labor, which is a key margin along which economies adjust their structure during development. Consequently, inter-sectoral gaps and inter-location wage gaps do not always move in tandem. As the data for China and India will make clear below, it is important to keep sectors and locations distinct.

We should note that our mechanism for generating structural change relies on a lower

[1] India, for example, introduced one of the largest ever public works programs in 2006 called the Mahatma Gandhi National Rural Employment Guarantee program. This program guarantees 100 days work to all rural workers.

income elasticity of demand for agricultural goods due to non-homotheticity in preferences as formalized in Laitner (2000), Kongsamut, Rebelo, and Xie (2001), Gollin, Parente, and Rogerson (2002), amongst others. An alternative mechanism that has been proposed in the literature (dating back to Baumol (1967)) relies on differential sectoral productivity growth. In particular, Ngai and Pissarides (2007) use a multi-sector model to show that as long as the elasticity of substitution between final goods is less than unity, over time factors would move to the sector with the lowest productivity growth. In both China and India this mechanism leads to a counterfactual implication since productivity growth in non-agriculture was faster than in agriculture in both countries. One could get around this by assuming that the elasticity of substitution between final goods is greater than unity. However, given the lack of precise estimates on this elasticity, it seems heroic to put the entire onus of the explanation on the configuration of a poorly measured parameter. Consequently, we shut down this channel by assuming that the elasticity of substitution between final goods is unity.[2]

Our explanation for wage gap dynamics in China and India contributes to the growing literature focused on misallocation of labor across locations due to migration costs, incomplete markets, and migration risk. Thus, Morten (2016) and Munshi and Rosenzweig (2016) show that the potential loss of village insurance networks constitutes an important cost of migration from rural to urban areas in India. Bryan and Morten (2016) evaluate the contribution of migration costs to the spatial wage differences in Indonesia. Our work contributes to this literature in several respects.

First, the extant literature focuses either on urban-rural disparities and urbanization in developing countries, or on the structural transformation of economies during development. We bring these two strands together by studying the endogenous evolution of structural transformation and urbanization.

Second, the existing literature has primarily focused on explaining spatial wage differentials at a point in time. Our work extends this analysis to add a time-series perspective on spatial wage gaps. Studying the evolution of rural-urban wage gaps over time not only allows us to better identify the factors behind the gaps, but also imposes more discipline on the structural model as we require the implications of such a model to be consistent with both the spatial wage differentials at a point in time and with the dynamic evolution of these gaps in response to aggregate economic developments. Indeed, this time series dimension of the model provides an independent, over-identifying test of the model.

Third, our analysis carefully differentiates the spatial and sectoral gaps in wages and labor

[2] Our work is also related to the factor deepening channel for structural transformation formalized in Acemoglu and Guerrieri (2008). Another possible channel is the skill acquisition cost mechanism proposed by Caselli and Coleman (2001) in their study of regional convergence between the North and South of the USA. In their model a fall in the cost of acquiring skills to work in the non-agricultural sector induces a fall in farm labor supply and leads to an increase in farm wages and relative prices. An overview of this literature can be found in Herrendorf, Rogerson, and Valentinyi (2013a).

allocations. Indeed, in the data we show that both locational and sectoral differences in wages and labor allocations significantly contribute to the overall distribution of wages and workforce in China and India.[3] This is in stark contrast to most of the literature which focuses on the sectoral dimension of the transformation. Examples of such an approach can be found in Herrendorf and Schoellman (2015), Lagakos and Waugh (2012) and Gollin, Lagakos, and Waugh (2014).

Our work is also related to the literature dating back to Kuznets (1955) that found that countries exhibit a wide mix of dynamic patterns on rural-urban inequality during the process of development (see Dudwick, Hull, Katayama, Shilpi, and Simler (2011)). Thus, while Japan between 1955-83 and the United Kingdom between 1871-1955 showed practically no change in regional wage gaps, the Habsburg empire (1756-1910) experienced a sharp divergence of wages between leading and lagging areas. Spain (1860-1975) (studied in Roses and Sanchez-Alonso (2004)) and Sweden (1920-61) (studied in Enflo, Lundh, and Prado (2014)), on the other hand, exhibited fairly sharp regional wage convergence between leading and lagging regions. We add to this literature by examining the experiences of two large developing countries and providing a possible explanation for the diversity of country experiences.

The rest of the paper is organized as follows: the next section presents the data and the main results on rural-urban gaps and their changes over time, as well as the analysis of the extent to which these changes were due to changes in individual characteristics of workers. Section 3 presents our model and examines the role of aggregate shocks in explaining the patterns. In section 4 we present some analytical results while section 5 presents the quantitative results. The last section contains concluding thoughts.

2 Empirical results

Our primary data source for China is the Chinese Household Income Project (CHIP). We use five rounds of the CHIP (1988, 1992, 1995, 2002 and 2008). Since our interest is in determining the trends in wages and determinants of wages such as education, we choose to restrict the sample to individuals in the working age group 16-65 who are identified as working and who report working at least 1900 hours per year. These restrictions leave us with 47,000 to 83,000 individuals per survey round. The data for India comes from successive rounds of the

[3] The distinction between locational and sectoral reallocation of factors shows up both in the data and in the policy initiatives within countries. In both China and India there is evidence of rural workers moving from agriculture into non-agriculture within rural areas. Indeed, a non-trivial share of the structural transformation in these economies occurs through workers switching sectors within the same location. Consequently, one finds a significant share of rural workers engaged in non-agricultural work even though non-agricultural productivity and wages are significantly higher in urban areas. On the policy front, the public works program NREGA in India is a response to a perceived concern that the market mechanism was not effective in generating sufficient urban, non-agricultural employment for workers switching out of agricultural work in rural areas.

Employment & Unemployment surveys of the National Sample Survey (NSS) of households in India. The survey rounds that we include in the study are 1983, 1993-94, 1999-2000, 2004-05, and 2009-10.[4] We restrict the sample to individuals in the working age group 16-65, who are working full time (defined as those who worked at least 2.5 days in the week prior to being sampled), who are not enrolled in any educational institution, and for whom we have both education and occupation information. We further restrict the sample to individuals who belong to male-led households.[5] These restrictions leave us with about 140,000 to 180,000 individuals per survey round. Details on our data are provided in Appendix A.1.

Our primary focus is on real wages. For China, we use annual wage income which is deflated using province-level CPI deflators that differ for rural and urban locations. For India we measure wages as the daily wage/salaried income received for the work done by respondents during the previous week (relative to the survey week), if the reported occupation during that week is the same as worker's usual occupation (one year reference).[6] Wages can be paid in cash or kind, where the latter are evaluated at current retail prices. We convert wages into real terms using state-level poverty lines that differ for rural and urban locations.[7] We express all wages in 1983 rural Maharashtra poverty lines.

We start by computing the wage gaps between urban and rural workers in China and India. The wage gaps are obtained from a regression of (log) wages on age, age squared and a rural dummy. Controls for age are included to account for potential differences in lifecycle stages of urban and rural workers, as is standard in the labor literature (below we also discuss how the inclusion of additional controls affects the wage gaps).[8] Panel (a) of Figure 1 shows the mean and median gaps for China while Panel (b) shows the corresponding gaps for India. Shaded areas represent 95% confidence intervals. The panels present a striking contrast: both the mean and the median urban-rural wage gaps widened in China between 1988 and 2008 while they narrowed in India between 1983 and 2010. Specifically, in China, the mean urban-rural

[4]There is also a survey round for 1987-88, but we did not include it in our analysis as the number of observations for wages in this round falls dramatically relative to the other rounds. This decline is mainly accounted for by the drop in the rural wage observations.

[5]This avoids households with special conditions since male-led households are the norm in India.

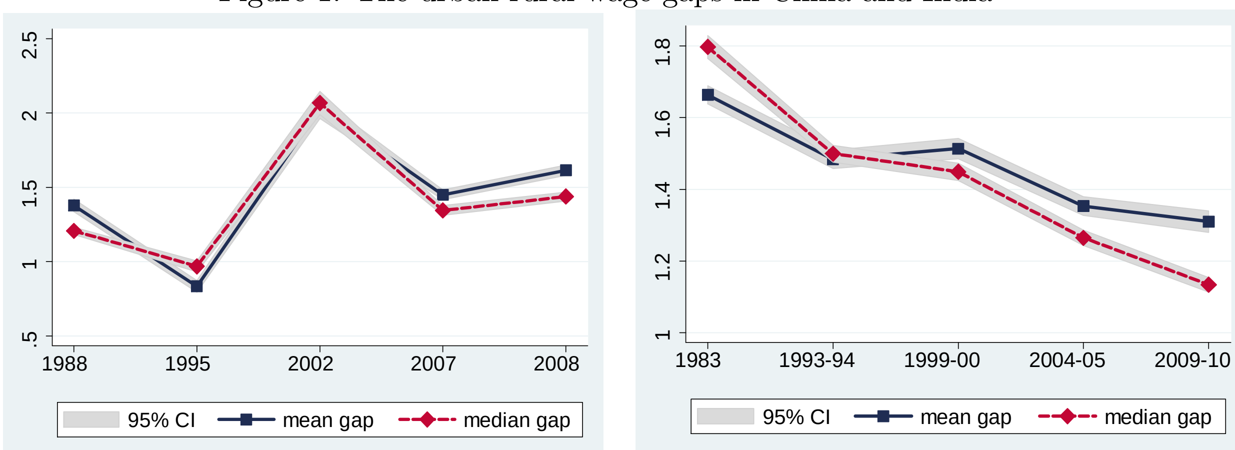
[6]This allows us to reduce the effects of seasonal changes in employment and occupations on wages.

[7]In 2004-05 the Planning Commission of India changed the methodology for estimation of poverty lines. Among other changes, they switched from anchoring the poverty lines to a calorie intake norm towards consumer expenditures more generally. This led to a change in the consumption basket underlying poverty line calculations. To retain comparability across rounds we convert the 2009-10 poverty lines obtained from the Planning Commission under the new methodology to the old basket using a 2004-05 adjustment factor. That factor was obtained from the poverty lines under the old and new methodologies available for the 2004-05 survey year. As a test, we used the same adjustment factor to obtain the implied "old" poverty lines for the 1993-94 survey round for which the two sets of poverty lines are also available from the Planning Commission. We find that the actual old poverty lines and the implied "old" poverty lines are very similar, giving us confidence that our adjustment is valid.

[8]The reported wage gaps are exponents of -1 times the coefficient on the rural dummy. Mean gaps are obtained from the OLS regressions, while median gaps are obtained from the Recentered Influence Function (RIF) regressions (see Firpo, Fortin, and Lemieux (2009) for more details).

wage gap increased from 38% in 1988 to 62% in 2008 – a 24 percentage points rise. In India, on the contrary, the mean urban-rural gap declined from 66% in 1983 to 31% in 2010 – a 35 percentage points decline. The changes in median wages were also pronounced in both countries. The median urban wage premium in China increased from 21% in 1988 to 44% in 2008 – a 23 percentage points rise; while it declined from 80% to 13% in India – a stunning 66 percentage points fall. Thus, we observe divergence in urban and rural wages in China, but a convergence in India over the past 30 years.[9]

Figure 1: The urban-rural wage gaps in China and India



(a) China (b) India

Notes: Panel (a) shows the mean and median urban-rural wage gaps for China, while Panel (b) shows the same wage gaps for India. These are obtained from a regression of (log) wages on a rural dummy, age, and age squared. Shaded areas are 95% confidence intervals.

The urban-rural wage gap could arise due to either the urban-rural gap being large within each sector (Agriculture or Non-agriculture, in our case), or due to between-sector gaps within each location (rural and urban, in our case) being large. Table 1 reports these conditional wage gaps in the two countries. The table highlights an important difference between China and India. In the case of China the major source of the large urban mean wage premium was the high urban-rural wage gap *within each sector,* whereas in India the big contributor was the *between-sector* gap in each location. Moreover, these gaps also evolved differently over time in the two countries. In China, the divergence between rural and urban wages was driven

[9]Note that levels of urban-rural wage gaps that we find are comparable to those reported in other studies that focus on the static differences between rural and urban wages. For instance, Munshi and Rosenzweig (2016) report a real wage gap (PPP-adjusted based on rural consumption) wage gap of 27 percent in India in 2004 (see Table 1 in their paper) and about 8% in China in 2005. While our number for India in 2004 is comparable to theirs, we find higher wage gaps in China. We used the 2005 Chinese mini Census in an attempt to replicate the 8% wage gap reported in Munshi and Rosenzweig (2016), but found that the mini Census only contains monthly income (and not wage) information. The urban-rural gap we get in that data using the sample criteria of Munshi and Rosenzweig (2016) is much larger at 62% which is in line with our income gap measures reported in the robustness Section 2.1 below.

by the divergence of urban-rural gaps within each sector; while in India, the urban-rural wage convergence was primarily due to shrinking sectoral gaps in each location. These patterns emphasize the importance of distinguishing sectors and locations in the analysis.

Table 1: Employment shares and wage gaps

	China			India		
	1988	2008	Δ_t	1983	2010	Δ_t
employment shares:						
L_U	0.26	0.35	**0.09**	0.22	0.30	**0.08**
L_{RA}/L_R	0.79	0.66	**-0.13**	0.78	0.66	**-0.12**
L_{UA}/L_U	0.05	0.03	**-0.02**	0.11	0.07	**-0.04**
wage gaps:						
within A, $\frac{w_{UA}}{w_{RA}}$	1.844	2.778	**0.934**	0.934	1.027	**0.093**
	(0.080)	(0.397)		(0.018)	(0.042)	
within N, $\frac{w_{UN}}{w_{RN}}$	1.289	1.605	**0.316**	1.082	0.994	**-0.088**
	(0.020)	(0.018)		(0.012)	(0.014)	
R between, $\frac{w_{RN}}{w_{RA}}$	1.285	1.272	**-0.013**	1.962	1.679	**-0.283**
	(0.051)	(0.080)		(0.022)	(0.019)	
U between, $\frac{w_{UN}}{w_{UA}}$	0.984	0.751	**-0.233**	2.259	1.709	**-0.550**
	(0.020)	(0.088)		(0.050)	(0.069)	
overall mean, $\frac{w_U}{w_R}$	1.379	1.614	**0.236**	1.664	1.310	**-0.354**
	(0.021)	(0.018)		(0.013)	(0.015)	

Note: Numbers in parenthesis are standard errors.

To illustrate the contribution of the within- and between-sector wage gaps to the overall wage divergence (convergence) in China (India), consider a simple decomposition of the overall urban-rural wage gap at any point in time:

$$\begin{aligned} \frac{w_U}{w_R} &= \frac{w_{UA}\frac{L_{UA}}{L_U} + w_{UN}(1-\frac{L_{UA}}{L_U})}{w_{RA}\frac{L_{RA}}{L_R} + w_{RN}(1-\frac{L_{RA}}{L_R})} \\ &= \frac{\frac{w_{UA}}{w_{RA}}\frac{L_{UA}}{L_U} + \frac{w_{UN}}{w_{RN}}\frac{w_{RN}}{w_{RA}}(1-\frac{L_{UA}}{L_U})}{\frac{L_{RA}}{L_R} + \frac{w_{RN}}{w_{RA}}(1-\frac{L_{RA}}{L_R})}, \end{aligned} \tag{2.1}$$

where the second equality was obtained by dividing both the numerator and denominator by w_{RA}. This decomposition expresses the overall urban-rural wage gap as a function of within- and between-sector wage gaps and sectoral labor shares. We then use it to conduct counterfactual exercises. In particular, we ask: How would the urban-rural wage gap in China have looked like if its urban-rural wage gap within each sector behaved like in India? Similarly, how would the urban-rural wage gap in India have looked like if its between-sector wage gaps in each location behaved like in China? To answer this question for China we substitute its within-sector wage gaps by those in India in both years, while keeping the sectoral labor shares and the between gaps at their corresponding values in China in both years. We find that the resulting mean wage gap in China would have exhibited a decline over time (equal to -14%), instead of the increase observed in the data. For India, we substitute its between-sector wage gaps by those in China in both years, while keeping everything else unchanged. The counterfactual mean wage gap exhibits much more muted convergence (equal to -11% instead of -35%) relative to that observed in the data. This result emphasizes the distinction in the sectoral versus locational drivers of wage gap dynamics in the two countries.

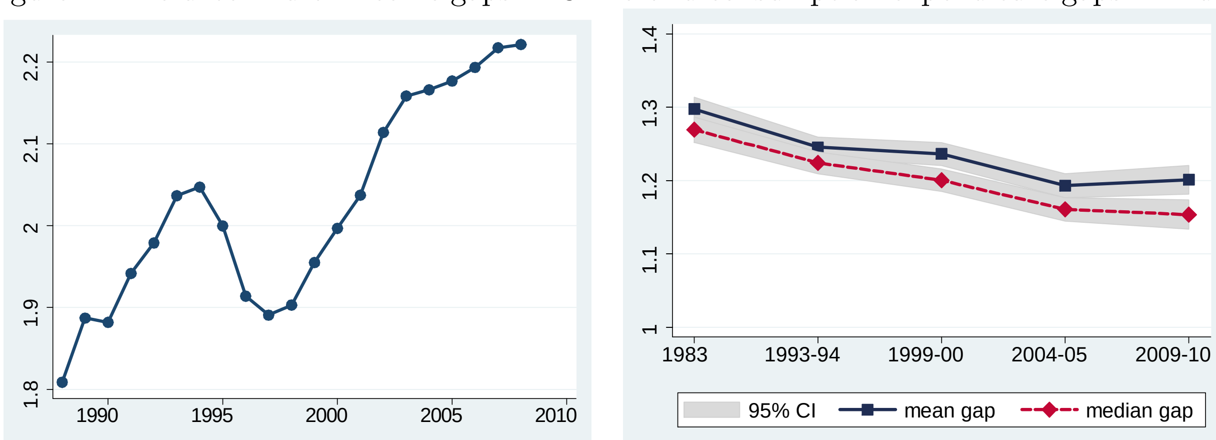
Table 1 highlights another important reasons for our focus on urban-rural gaps as opposed to just inter-sectoral gaps. The top panel of that table reports employment shares of agriculture in each location. It makes it clear that in both China and India, the period since the 1980s has been accompanied by an increase in the share of the rural labor force engaged in non-agricultural activities. Thus, $\frac{L_{RA}}{L_R}$ fell from 0.79 to 0.66 in China while it fell from 0.78 to 0.66 in India during the period under study. In effect, the share of the rural labor force engaged in non-agricultural activities in these two countries rose from just under 1/4 to around 1/3. Clearly, movement of factors from agriculture to non-agriculture is not isomorphic with movement of factors from rural to urban areas.

2.1 Robustness of wage patterns

A potential concern with the wage patterns documented above is that they may not be representative of the overall patterns in urban-rural *income* gaps since a significant proportion of rural workers tend to be self-employed and, therefore, do not report any wage income. We examine robustness of our data findings to this issue by using two supplementary checks. First, we examine the relative share of self-employed workers in the urban and rural labor forces to uncover any systematic trends. Differential trends in the relative proportions of self-employed in urban and rural areas could potentially induce trends in wage gaps simply through a composition effect even if the underlying wage gaps remained unchanged. We find that in India, the share of self-employed in the urban labor force is lower than in rural areas. However, there are no systematically differential trends in these shares over time with the urban share fluctuating around 40 percent and the rural share around 60 percent. In China, the share of self-employed in the urban labor force rose from 1.2 percent to 8.8 percent while the corresponding rural share rose from close to zero in 1988 to 7.4 percent in 2008. Thus, in neither country do we find systematic differences in the dynamics of urban and rural incidence of self-employment.

Second, to check whether the wage patterns carry over to broader measures of income, we also consider family income in China (which consists of wage income and other non-wage sources of income) and household consumption expenditures in India (which is a proxy for family income that includes both wage and self-employed income). Both variables are in real per capita terms.[10] Panel (a) of figure 2 reports the mean gaps in annual family income between urban and rural households in China using the China Statistical Yearbook, while panel (b) shows the mean and median urban-rural gaps in per capita monthly consumption expenditures in India using NSS data. These figures confirm our findings for wages. In China, urban family income, much like wage income, has diverged from rural family income over time with income gaps rising more sharply than wage gaps. In contrast, in India, consumption expenditures in urban and rural families have been converging over time, although the convergence is more muted than the convergence in wages. This is not surprising given that the consumption expenditure gaps are smaller to start with and since consumption habits also tend to adjust slowly over time.[11]

Figure 2: The urban-rural income gaps in China and consumption expenditure gaps in India



(a) family income in China (b) household cons. expenditures in India

Notes: Panel (a) shows the urban-rural per capita family income gaps in China using China Statistical Yearbook data, while panel (b) shows the urban-rural per capita consumption expenditure gaps for India using NSS data. The income gaps are obtained as -1 times the exponents of coefficients on the rural dummy from OLS regression of (log) consumption expenditures on a rural dummy. The consumption gaps are obtained in the same way, (also using RIF regression to get median gaps) except the regressions also include the household size to account for possible scale effects in household consumption.

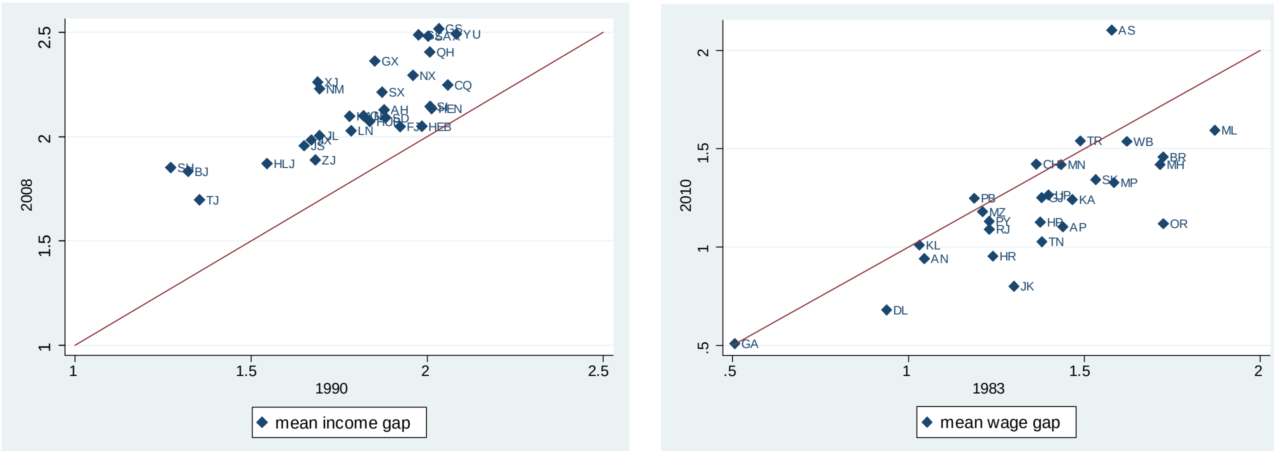
A different concern about the wage and income patterns documented above is that they may be driven by some outlier provinces or states. Panel (a) of Figure 3 plots the scatter of the urban-rural income gaps across provinces in China for 1990 and 2008, while Panel (b) plots the scatter of urban-rural wage gaps for states in India for 1983 and 2010. The key feature to note is that most of the points for China lie above the 45 degree line indicating larger gaps

[10] Family income in China is annual income, while consumption expenditures in India are monthly expenditures.

[11] Another concern regarding the robustness of the wage patterns shown in Figure 1 for China is that the CHIP data is not nationally representative. Using family income data from China Statistical Yearbook allows us to check for this as the Yearbook covers all provinces.

in 2008 relative to 1990. The corresponding scatter of points for Indian states lie primarily below the 45 degree line indicating a narrowing of the wage gap between urban and rural workers between 1983 and 2010. Thus, income divergence in China and wage convergence in India seem to be taking place across-the-board.

Figure 3: The urban-rural wage gaps by province/state in China and India



(a) China (b) India

Notes: Panel (a) shows the urban-rural income gaps for provinces in China for 1990 and 2008. Panel (b) shows the urban-rural wage gaps for India for 1983 and 2009-10 NSS rounds.

We view these results as suggestive of the robustness of the basic fact that the urban-rural wage gap widened in China between 1988 and 2008 while it declined in India between 1983 and 2010.

3 Explaining the trends

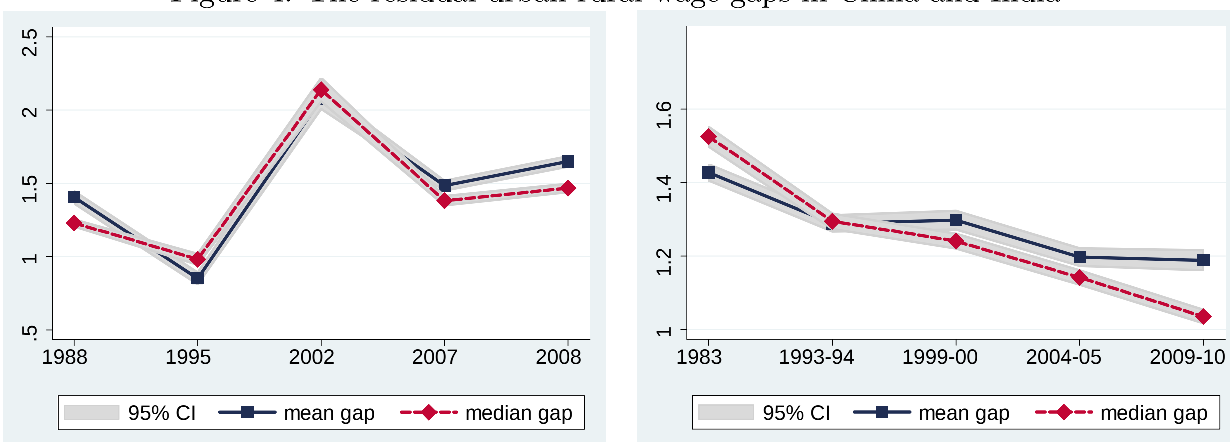
What explains the observed patterns in the urban-rural wage gaps in the two countries? The standard explanations focus on measured attributes in wages such as demographics, education, occupation, etc.. How much of the wage convergence/divergence documented above is driven by a convergence/divergence of measured covariates? We examine this issue in two ways. First, we compute residual wage gaps after accounting for the key demographic characteristics of workers in the two countries in each survey round. Our set of attributes includes age, age squared, gender, and regions of residence.[12] For India we also include a dummy variable for the Scheduled Castes and Scheduled Tribes (SC/STs) who constitute a generally more disadvantaged group and tend to reside more in rural areas. The resulting wage

[12]These regions are constructed based on the similarity of their history and economic activity. In India these regions include North, East, Central, North-East, South and West. In China they include East, Northeast, Central and West.

We do not include occupations in the set of characteristics since they are likely endogenous to wages.

gaps are reported in Figure 4. As is easy to see, the residual mean and median wage gaps are somewhat lower than those reported in Figure 1, but show the trends that are symmetric to the unconditional gaps. For instance, in India, the mean residual wage gap declined from 43% in 1983 to 19% in 2010 – a 24 percentage point decline, which suggests that demographics can explain only about a third of the decline in mean wage gap. In China, the residual mean wage gap increased from 41% in 1988 to 65% in 2008 – a 24 percentage points increase, which is comparable to the rise in unconditional mean wage gap. This suggests that in China individual characteristics had low explanatory power for the observed divergence in wages.

Figure 4: The residual urban-rural wage gaps in China and India



(a) China (b) India

Notes: Panel (a) shows the mean and median urban-rural wage gaps for China, while Panel (b) shows the same wage gaps for India. These are obtained from a regression of (log) wages on a rural dummy, age, and age squared. Shaded areas are 95% confidence intervals.

Second, we use an Oxaca-Blinder-type decomposition to formally decompose the observed *time-series changes* in the mean wage gaps into explained and unexplained components as well as to quantify the contribution of the key individual covariates. The covariates include demographics and education.[13] The details of the decomposition and results are reported in the appendix. We find that these attributes provide only a very partial accounting for the observed wage gap changes in the two countries. In particular, in India at most 23% of the observed wage convergence is due to convergence in individual attributes of urban and rural workers, with education convergence accounting for a third of that. Interestingly, in China we also find that individual attributes of rural and urban workers have been converging, predicting a minor convergence in urban-rural wage gaps. This makes the observed wage divergence in China even more puzzling.[14]

[13]In China we did not include regional dummies into the set of characteristics because several northeastern provinces were not covered by the CHIP survey in 2008. In the decompositions of wage changes between 1988 and 2007 these dummies do not significantly alter the results.

[14]Detailed decomposition results can be found in the Appendix.

Another possible explanation for the wage gap patterns we observe in China and India is workers' sorting on unobserved skill with the resulting wage gaps reflecting difference in the skill intensities of production between rural and urban areas, as suggested by Young (2013). We do not think sorting on skill explains the experience in India and China. First, our empirical results above show that a majority of the change in urban-rural wage gap is left unaccounted for by changing worker characteristics in both China and India over the studied period. To the extent that the included worker characteristics are correlated with her unobserved skills, the latter will not be contributing much to the observed wage gaps dynamics. Second, as Munshi and Rosenzweig (2016) point out, Young's model implies that migration rates in both directions (from rural to urban and from urban to rural areas) should be high when wage gaps are high to produce the appropriate mix of skills in both sectors. But in both countries, both urban and rural out-migration rates are low. Table 2 summarizes gross migration flows in China and India in different years. The top panel (a) is for India and is based on Table 2 from India migration study in Hnatkovska and Lahiri (2015).[15] The rural-to-urban migration flows are just around 2% of Indian labor force across all survey years, while urban-to-rural migration is around 1% of labor force. These are small numbers. They become even smaller if one weights them by the share of migrants who move for a job (as opposed to marriage, due to natural disaster, social problems, displacement, housing based movement, health care, etc.), as this share is about 40% for rural-to-urban migrants, and about 20% for urban-to-rural migrants (see Hnatkovska and Lahiri (2015)). Majority of migration in India takes place between rural areas.

For China we are not aware of studies documenting gross migration flows, so we used 2000 Census to shed some light on this question. Migrants are defined as those individuals whose birth place is different from their place of residence.[16] Unfortunately, Chinese censuses do not contain information about whether the location of birth of an individual was rural or urban, only the rural or urban status of her current residence. Instead one can compute the share of inter-provincial and intra-provincial migrants in each location. The results are reported in panel (b) of Table 2. The largest migration flow is migrants who are in the urban labor force and who were born and moved within the same province. We conjecture that a large share of these are rural-to-urban migrants who have agricultural Hukous but work in non-agriculture. It is harder to infer the urban-to-rural migration flows from these numbers, but they are likely much lower than the rural-to-urban flows.

Overall, the gross migration flows in both China and India are small, especially in comparison to other developing countries of similar level of development (see, for instance, Munshi

[15] The data is from Employment & Unemployment surveys of the National Sample Survey (NSS) of households in India. Only the survey rounds that contain migration information are included.

[16] The same flows are reported in Fan (2015), with the exception that our migration numbers are shares of total labor force in China in 2000.

Table 2: Gross migration flows in China and India

	a. India migrants/labor force, %			
	rural-to-urban	urban-to-urban	rural-to-rural	urban-to-rural
1983	1.98	1.61	5.55	0.79
1999-00	1.96	1.67	5.64	0.93
2007-08	2.10	1.68	5.22	0.74

	b. China migrants/labor force, %			
	rural		urban	
	within prov	across prov	within prov	across prov
2000	3.95	3.95	3.99	2.31

Note: The table reports the ratio of gross migration flows to total labor force. The numbers for India are from the corresponding NSSO rounds, while for China they are from 2000 Census. In India, migrants are identified as individuals who reported that their place of enumeration is different from the last usual residence and who left their last usual place of residence within the previous five years. In China, we define migrants are those whose place of birth is different from their place of residence.

and Rosenzweig (2016) who show that gross urban-rural and rural-urban migration rates were much larger 4.55% and 13.9% in Brazil in 1997). We interpret this evidence as suggesting that worker sorting likely played a limited role in wage gap dynamics in the two countries.

Third, the opposite wage gap dynamics in China and India provide a further challenge to the worker sorting explanation. Both China and India experienced growth in sectoral productivity, with productivity (both labor and TFP) rising faster in non-agriculture than in agriculture (we show this in detail below). Young's model would imply that the wage gaps should be expanding in both countries, which is clearly not the case in India. The opposite wage gap dynamics in the two countries therefore require an explanation that goes beyond worker sorting mechanism. We take on this challenge next.

We argue that aggregate developments during this period have played an important role. Specifically, the period since the 1980s was marked by a sharp increase in the aggregate growth rate, structural transformation of employment and output, and rapid urbanization of the economy in both countries. More precisely, the key aggregate facts are: (i) China and India have both experienced a decline in the share of output and employment in agriculture – the textbook features of structural transformation; (ii) In both China and India, labor productivity was increasing in agriculture and non-agriculture during the last thirty years, with non-agricultural productivity expanding at a much faster pace. However, a key difference was that labor productivity in China grew must faster than in India. Thus, the labor productivity in agriculture increased by only 67 percent in India between 1983 and 2010. In contrast, agricultural labor productivity in China grew by 163 percent between 1990 and 2008. Similarly, the non-agricultural labor productivity rose by 200 percent in India and 338 percent in China during the same periods;[17] (iii) The relative price of non-agriculture declined in both countries: by 23 percent in China and 29 percent in India;[18] (iv) Both countries have become more urban with the urban share of employment rising from 26 to 35 percent in China and 22 to 30 percent in India.[19]

Thus, the patterns of structural transformation, sectoral productivity growth, urbanization and relative price movements were all qualitatively similar in China and India. Our explanation for contrasting wage dynamics in the two countries relies on quantitative differences in sectoral productivity growth and labor reallocation between rural and urban locations. In particular, we argue that the growth in sectoral productivity gives rise to structural transformation of the economy in both countries. However, it also leads to an expansion in locational wage gap in the two countries. Larger wage gaps induce worker reallocation from rural to urban locations, thus raising relative urban labor supply and leading to greater urbanization of both countries. This effect brings rural and urban wages closer together. The relative strength of the two effects determines whether locational wage gaps converge or diverge. The two effects, at the same time, unambiguously lead to structural transformation and urbanization in both countries.

We evaluate this explanation with a calibrated structural model.

3.1 A Structural Explanation

We formalize a simple model with two sectors (agriculture and non-agriculture) and two locations (rural and urban). We begin by presenting the full model. Then we simplify the environment to consider two extreme cases of the model: one with the frictionless labor markets across sectors and locations, and one with extreme frictions prohibiting migration across locations. These special cases allow us to develop the intuition and highlight the minimal model features needed to explain the data. We then return to the full model and quantitatively examine the relative contributions of the identified factors to the observed wage convergence in India and wage divergence in China.

Consider an economy with two locations: rural and urban. Each location produces two goods – an agricultural good and a non-agricultural good. Under our formalization, locations are defined by three key distinguishing characteristics: (a) their productivities in producing the two goods; (b) the amenities they provide for their residents; and (c) the cost of training workers in each location. We elaborate on these below.

[17] When reporting growth rates of labor productivity we used 1990 as the starting year for China instead of 1988 because of discountinuity in the sectoral employment data for China in 1989. We suspect that the definition of employed must have been changed in that year.

[18] It is worth noting that the world relative price of agriculture was actually falling during most of the period since the 1980s, in contrast to the rising relative price of agriculture in China and India.

[19] See Appendix for data sources, computations and more detailed dynamics of these variables.

We assume that goods markets are integrated in this economy so that the price of each good is equalized across locations. However, labor mobility across locations is costly. Hence, factor markets are segmented across locations at any point in time implying that factor prices can also differ across locations. We assume throughout that there is no uncertainty in this economy so that we shall focus on equilibria with perfect foresight.

3.2 Technology

The location-specific technologies for producing the two goods are

$$Y_t^{jA} = A_t^j \left(L_t^{jA} \right)^{\alpha^j}, \quad j = R, U \tag{3.2}$$

$$Y_t^{jN} = N_t^j \left(L_t^{jN} \right)^{\beta^j}, \quad j = R, U \tag{3.3}$$

where $\alpha^j \in (0, 1)$ and $\beta^j \in (0, 1)$. Throughout the paper we shall use R to denote rural and U to denote urban. L^{jA} denotes total employment of labor in the agricultural (A) sector in location $j = R, U$. Similarly, L^{jN} denotes total employment of labor in the non-agricultural (N) sector in location $j = R, U$. Note that underlying these decreasing returns to labor technologies is a fixed factor like land. A^j and N^j denote the total factor productivity in location $j = R, U$ in the agricultural and non-agricultural sectors, respectively. Importantly, we allow the sectoral productivities to be different across locations. Indeed, this is one of the aspects distinguishing locations in the model.[20]

Competitive firms in each location and sector hire labor to maximize profits. Consequently,

$$w_t^{jA} = \alpha^j \frac{Y_t^{jA}}{L_t^{jA}}, \quad j = R, U \tag{3.4}$$

$$w^{jN} = \beta^j \frac{p_t Y_t^{jN}}{L_t^{jN}}, \quad j = R, U \tag{3.5}$$

where w^{jA} denotes the real wage in location j in sector A, while w^{jN} is the real wage in location j in sector N. p is the relative price of the non-agricultural good in terms of the agricultural good, which we treat as the numeraire good throughout. Clearly, profits of firms

[20] Recent work by Hsieh and Klenow (2009) suggests that misallocation of capital across plants in China and India is quite pervasive. The production functions assumed in equations (3.2) and (3.3) preclude discussions of capital misallocation since we do not include capital as an input. Our modeling choice is deliberate. Which way these misallocations might affect urban-rural wage gaps would clearly depend on whether the misallocations are greater in urban or rural locations and in the non-agriculture or agricultural sectors. We neither have that level of disaggregated data to address this issue empirically nor do we have any strong priors on which way the omission might bias our results based on currently exisiting scientific work on the topic.

then are

$$\Pi_t^{jA} = \left(1 - \alpha^j\right) Y_t^{jA} \tag{3.6}$$

$$\Pi_t^{jN} = \left(1 - \beta^j\right) p_t Y_t^{jN} \tag{3.7}$$

These are the returns to the fixed factor.

3.3 Households

Each location is inhabited by overlapping generations of two-period lived individuals. In the first period of life each individual chooses the location where she wants to live next period. Changing locations, however, is costly. Young individuals who choose to change their location have to pay τ units of the agricultural good as a relocation cost. These relocation costs can be financed through borrowing. In the second period of life individuals work in the location they chose when young, have children, repay their debts (if any), consume and then die. Each worker in location $j = R, U$ at date t has 1 kid so that population is constant in this economy over time.[21]

In the second period of their lives, individuals have an endowment of one unit of time which they supply inelastically to the labor market in their location of residence. Labor time supplied to the $A-$sector is directly productive. Labor time supplied to the $N-$sector however requires some sectoral training which entails a cost τ^j units of the agricultural good per unit of labor time. Note that we are allowing the labor training costs to be location specific, since $j = R, U$.

Individuals derive utility from consumption only when old. Hence, lifetime utility of an individual born at date t in location i and who chooses to work and consume in location j at date $t+1$ is

$$V_t^{ij} = u\left(c_{t+1}^{ij}\right) \varepsilon_{t+1}^{j}, \quad u' > 0,\ u'' < 0$$

where

$$c_t^{ij} = \left(c_t^{ijA} - \bar{a}\right)^{\theta} \left(c_t^{ijN} + \bar{n}\right)^{1-\theta}, \quad i = R, U,\ j = R, U.$$

c^{ij} denotes consumption of an individual born in location i and consuming in location j. $\bar{a}$ denotes the minimum consumption level of the agricultural good and $\bar{s}$ is the minimum level of the non-agricultural good that is produced at home.[22]

[21] Our formalization of the migration decision as a dynamic choice is aimed at capturing long-term migration. We could potentially allow within-period, short-term migration as well by having workers change locations after observing the locational productivities for the period and paying for the migration cost out of their current period wage earnings. This would not change the theoretical logic of the model. We choose to abstract from this margin since a large part of these short-term flows tend to reverse themselves within a year as opposed to the long term migration flows that contribute to the growing urbanization of the economy.

[22] This is a standard method of introducing non-homotheticity which makes the income elasticity of demand for the agricultural good less than the corresponding income elasticity of the non-agricultural good.

ε^j is a term reflecting the level of amenities available in location $j = R, U$. It is exogenous to the individual, and identical for all agents in location j. We shall assume that

$$\varepsilon_t^j = \bar{\varepsilon}^j e\left(M_t^j, L_t^{jj}\right), \quad j = R, U \tag{3.8}$$

where M_t^j denotes the number of migrant workers in location j at date t and L_t^{jj} denotes the total number of workers in location j at date t who were also born in location j. $\bar{\varepsilon}^j$ is a location-specific constant scalar. The function $e(.,.)$ captures externalities that could arise from new migrants moving into the location as well as the size of the location. We shall specialize this function as

$$e\left(M_t^j, L_t^{jj}\right) = \left(1 + \frac{M_t^j}{L_t^{jj}}\right)^{\phi}, \qquad j = R, U$$

Note that if $\phi < 0$ then there are negative externalities associated with population growth due to relocation of workers into a location. Note also that e reduces to unity when migration ceases. This reflects the idea that the externalities associated with city growth reflect transitions where population growth exceeds the ability of the location to absorb the new immigrants. In a stationary state, migration ends and the city augments its infrastructure to reflect its new size.

In the following we shall specialize the utility function to the CRRA form:

$$u(c) = \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}},$$

where σ denotes the intertemporal elasticity of substitution.

The budget constraints facing individuals in the two periods of their lives are

$$\text{young: } \tau_t I_t^{ij} = b_{t+1}^i, \qquad i = R, U \tag{3.9}$$

$$\begin{aligned} \text{old} \quad &: \quad c_{t+1}^{ijA} + p_{t+1} c_{t+1}^{ijN} + R_{t+1} b_{t+1}^i I_t^{ij} + \frac{T_{t+1}}{L_{t+1}} = w_{t+1}^{jA} l_{t+1}^{ijA} + \left(w_{t+1}^{jN} - \tau_{t+1}^j\right) l_{t+1}^{ijN} + \frac{\sum_{j=R,U}\left(\Pi_{t+1}^{jA} + \Pi_{t+1}^{jN}\right)}{L_{t+1}}, \\ i, j \quad &= \quad R, U \end{aligned} \tag{3.10}$$

where R_{t+1} denotes the gross interest factor on loans b_{t+1} contracted in period t. I_t^{ij} is an indicator function that takes a value of one if a young individual at time t in location i decides to migrate to location j, and equals 0 otherwise. l^{ij} denotes the labor supplied by an individual born in location i and working in location j. $L_{t+1} = L_{t+1}^R + L_{t+1}^U$ denotes the total population of old at time $t+1$. The last term on the right hand side of equation (3.10) reflects the fact that all firms are owned equally by the old who receive dividends from firms in proportion to their ownership. $\frac{T}{L}$ denotes the common per capita lump sum tax that is imposed by the government on all households.

The first-order-condition dictating the optimal consumption mix of the two goods is given by

$$p_t = \left(\frac{1-\theta}{\theta}\right)\left(\frac{c_t^{ijA}-\bar{a}}{c_t^{ijN}+\bar{n}}\right), \quad i,j=R,U \tag{3.11}$$

Combining equation (3.11) with the budget constraint (3.10) yields the optimal consumption plans of the old in location j at date t as:

$$c_t^{ijA} = \theta\left[\hat{y}_t^{ij} + \left(\frac{1-\theta}{\theta}\right)\bar{a} + p_t\bar{n}\right], \quad j=R,U \tag{3.12}$$

$$p_t c_t^{ijN} = (1-\theta)\left[\hat{y}_t^{ij} - \bar{a} - \left(\frac{\theta}{1-\theta}\right)p_t\bar{n}\right], \quad j=R,U \tag{3.13}$$

where $\hat{y}_t^{ij} = w_t^{jA}l_t^{ijA} + \left(w_t^{jN} - \tau_t^j\right)l_t^{ijN} - R_t b_t^i I_{t-1}^i - \frac{T_t}{L_t} + \frac{\sum_{j=R,U}\left(\Pi_t^{jA}+\Pi_t^{jN}\right)}{L_t}$ denotes the disposable income of a worker in location j at time t who was born in location i at date $t-1$.

The optimal labor supply decision of a worker in location $j=R,U$ at date t dictates that

$$w_t^{jA} = w_t^{jN} - \tau_t^j, \quad j=R,U \tag{3.14}$$

This sectoral indifference condition reflects the fact that the worker can freely allocate their labor time to either the agricultural sector or to the non-agricultural sector in their location at cost τ^j per unit of labor supplied to the non-agricultural sector.

Before proceeding further, it is worth noting that since all individuals supply their one unit of labor time inelastically to the market, and since all individuals in a location j face the same sectoral wages, w^{jA} and w^{jN}, the sectoral labor allocation by all individuals in the same location must be identical, i.e., $l_t^{ijk} = l_t^{jjk} = l_t^{jk}$ for all t and for all $i,j=R,U$ and $k=A,N$.

3.3.1 Location decision

The young at date t get to choose where they want to work next period. If they choose to change locations then they have to pay a fixed cost τ. These costs can be interpreted as representing all costs that impede the conversion of rural workers into urban workers. Reallocation of workers from rural to urban areas can occur through a physical transfer of people (net migration into urban areas from rural locations) and through a spatial transfer of people (due to a reclassification of rural areas into urban areas). Indeed there is evidence of villagers living in urban peripheries pooling their land and capital in order to extract urban rents through land reclassification (see Tiwari, Rao, and Day (2016)). In this event, τ would be interpreted as the collective action cost per capita that has to be incurred in order to reclassify the land. Existing estimates suggest that these two channels have contributed about equally to urbanization in India over the past three decades (Revi (2011)). In China, during 1990s (2000s), about 30% (27%) of population reallocation from rural to urban areas was due to reclassification, while the remaining share was accounted by net migration into urban areas (see Chan and Hu (2003) for the analysis for 1990s and Qin and Zhang (2014) for the 2000s). Since changes in relative aggregate labor supply between rural and urban locations are attributable to both reallocation forces in both countries, we choose to not explicitly distinguish the two channels in our analysis. In what follows we use labor reallocation and migration interchangably.

The location decision of the young at date t is dictated by a comparison of lifetime utility at each location. Recall that V_t^{ij} denotes the lifetime utility of a worker born at date t in location i and working in location j when old. The optimal consumption plans given in equations (3.12) and (3.13) above imply that

$$V_t^{ij} = u\left(\frac{\Gamma}{p_{t+1}^{1-\theta}}\left\{\hat{y}_{t+1}^{ij} - \bar{a} + p_{t+1}\bar{n}\right\}\right)\varepsilon_{t+1}^{j}, \quad i,j = R,U$$

where $\Gamma \equiv \theta^{\theta}(1-\theta)^{1-\theta}$. This expression shows that there are four types of workers at any date depending on where the worker was born and where she chose to locate as a worker.

Since all young individuals can switch location, and there is perfect foresight, for individuals to be indifferent between changing locations or staying on where they were born, their lifetime welfare must be independent of their location choice when young. This implies that

$$V_t^{RR} = V_t^{RU}; \quad V_t^{UR} = V_t^{UU}$$

Given the positive cost of migration and no individual specific amenity effects, in this environment we will either have rural born individuals moving to the urban location or urban individuals moving to the rural location but not simultaneous movement in both directions. To see this note that the rural young will migrate to urban areas only if $V_t^{RU} \geq V_t^{RR}$ while the urban young workers will migrate to rural areas if $V_t^{UR} \geq V_t^{UU}$. For simultaneous migration in both directions we must then have

$$\frac{V_t^{RU}}{V_t^{RR}} \geq 1 \geq \frac{V_t^{UU}}{V_t^{UR}}$$

It is easy to check that with $\tau > 0$ this is generically contradictory. Hence, migration flows will occur in only one direction. Without loss of generality, in the remainder of the paper we shall restrict attention to parameter ranges such that $\frac{V_t^{RU}}{V_t^{RR}} \geq 1 \geq \frac{V_t^{UR}}{V_t^{UU}}$ which implies that only individuals born in rural areas would have an incentive to change locations.

A rural young would be indifferent between switching and not switching locations if and only if $V_t^{RR} = V_t^{RU}$. This gives the migration indifference condition that must be satisfied at all dates[23]

$$w_{t+1}^{UA} - R_{t+1}\tau_t - w_{t+1}^{RA} = \left[\left(\frac{\varepsilon_{t+1}^R}{\varepsilon_{t+1}^U} \right)^{\frac{\sigma}{\sigma-1}} - 1 \right] \left(\hat{y}_{t+1}^{RR} - \bar{a} + p_{t+1}\bar{n} \right). \tag{3.15}$$

The left-hand-side of this expression is the wage increase that the location switch generates for the rural migrant worker tomorrow net of the moving cost. The right-hand-side represents the foregone relative utility from staying on in the rural location. It depends on the relative amenities in the two locations, $\frac{\varepsilon_{t+1}^R}{\varepsilon_{t+1}^U}$, and the disposable income of the rural young who stays on in the rural location (adjusted for the expenditures on minimum level of agricultural and non-agricultural goods). Note that when the amenities of the two locations are identical so that $\varepsilon_{t+1}^R = \varepsilon_{t+1}^U$, the indifference condition reduces to a wage parity in the two locations net of migration cost, $w_{t+1}^{UA} - R_{t+1}\tau_t = w_{t+1}^{RA}$.[24,25]

3.4 Government

A key feature of the model is that the young have to decide on their location decision when they have no source of income. Consequently, this has to be financed through borrowing. We assume that there is a government agency that imposes a lump sum tax T on households and uses these proceeds along with the repayments of past loans by current workers to finance new loans to the young in any period. The budget constraint of the government is

$$T_t + \mu_{t-1} L_{t-1}^R R_t b_t = \mu_t L_t^R b_{t+1}$$

where μ_t denotes the measure of rural young that choose to change location at date t. From equation (3.9) it is clear that $b_t = \tau_t$ for all t since borrowing is needed only to finance the location switching cost. Hence, the government's budget constraint can be written as

$$T_t + \mu_{t-1} L_{t-1}^R R_t \tau_{t-1} = \mu_t L_t^R \tau_t \tag{3.16}$$

The government can either choose the interest rate R_t or adjust the lump sum tax T_t to

[23] The details of all the derivations of the model as well as the proofs below are available in the online Model Appendix that accompanies this paper.

[24] We assume that the initial old generation at date $t = 0$ can freely choose their location so that $w_0^{UA} - w_0^{RA} = \left[\left(\frac{\varepsilon_0^R}{\varepsilon_0^U} \right)^{\frac{\sigma}{\sigma-1}} - 1 \right] \left(\hat{y}_0^{RR} - \bar{a} + p_0\bar{n} \right)$.

[25] Note that from equation (3.15), the urban-rural wage gap would be rising (falling) with $\frac{\varepsilon^R}{\varepsilon^U}$ as $\sigma > (<) 1$. Consequently, the urban-rural wage gap would widen (decline) with migration into urban areas if migration worsens (raises) urban amenities.

ensure that equation (3.16) holds at every t for all values of the other variables. We shall assume that the credit agency lends to the young at a constant interest factor so that $R_t = R$ for all t.[26] In this case T_t becomes endogenous and is given by

$$T_t = \mu_t L_t^R \tau_t - \mu_{t-1} L_{t-1}^R R \tau_{t-1}.$$

We should note that this particular arrangement of the government financing the migration costs is without any loss of generality. It is easy to instead set up the arrangement as a credit cooperative where all workers contribute a lump sum amount T. The cooperative has a balanced budget constraint as in equation (3.16) above wherein it makes new loan disbursements out of the lump sum contributions of the credit cooperative members and loan repayments. The equilibrium outcomes in the two cases would be identical.

3.5 Aggregation

To complete the description of this economy, we now aggregate the individual variables to represent aggregate variables. First, the population dynamics of the two locations and the economy as a whole are given by

$$L_t^R = \left(1 - \mu_{t-1}\right) L_{t-1}^R \tag{3.17}$$

$$L_t^U = L_{t-1}^U + \mu_{t-1} L_{t-1}^R \tag{3.18}$$

$$L_t = L_t^R + L_t^U \tag{3.19}$$

At every date there are three types of workers in this economy – those that were born and work in rural areas; those that were born and work in urban areas; and those that were born in rural areas but changed locations to work in urban areas. Hence, the total sectoral allocation of labor in each location is given by

$$L_t^{Uk} = l_t^{UUk} L_{t-1}^U + l_t^{RUk} \mu_{t-1} L_{t-1}^R, \quad k = A, N$$

$$L_t^{Rk} = l_t^{RRk} \left(1 - \mu_{t-1}\right) L_{t-1}^R, \quad k = A, N$$

Next, we use the individual consumption plans to derive aggregate values of consumption of

[26] Alternatively, the government could finance the switching cost by setting $T_t = 0$ for all t but for the funding agency to adjust the interest rate every period to ensure that $\mu_{t-1} L_{t-1}^R R_t \tau_{t-1} = \mu_t L_t^R \tau_t$. This amounts to the migrant workers being charged an interest rate that is just sufficient to finance location switches by young rural individuals at every date. In this case the interest rate would become endogenous and be given by $R_t = \frac{\mu_t L_t^R \tau_t}{\mu_{t-1} L_{t-1}^R \tau_{t-1}}$. In this scenario, the initial period switches at $t = 0$ would be financed through a one-time lump sum tax $T_0 = \mu_0 L_0^R \tau_0$.

the two goods in the two locations:

$$C_t^{Uk} = c_t^{UUk} L_{t-1}^U + c_t^{RUk} \mu_{t-1} L_{t-1}^R, \qquad k = A, N$$

$$C_t^{Rk} = c_t^{RRk} \left(1 - \mu_{t-1}\right) L_{t-1}^R, \qquad k = A, N$$

Clearly, aggregate consumption of the two goods are $C_t^k = C_t^{Uk} + C_t^{Rk}$ where $k = A, N$.

3.6 Equilibrium

To describe the equilibrium for this economy we first note that at every date all equilibrium allocations must be consistent with market clearing in each sector, i.e.,

$$C_t^A + \tau_t^R L_t^{RN} + \tau_t^U L_t^{UN} + \tau_t \mu_t L_t^R = Y_t^A \tag{3.20}$$

$$C_t^N = Y_t^N \tag{3.21}$$

Define the price and quantity vectors, Ψ and Ω respectively, as

$$\begin{aligned} \Psi_t &= \left\{p_t, w_t^{UA}, w_t^{UN}, w_t^{RA}, w_t^{RN}\right\} \\ \Omega_t &= \left\{c_t^{UUA}, c_t^{UUN}, c_t^{RUA}, c_t^{RUN}, c_t^{RRA}, c_t^{RRN}, l_t^{UUA}, l_t^{UUN}, l_t^{RUA}, l_t^{RUN}, l_t^{RRA}, l_t^{RRN}, \mu_t\right\} \end{aligned}$$

DEFINITION: *The perfect foresight competitive equilibrium for this economy is a time path of the vectors (Ψ_t, Ω_t) such that all young and old individuals, and firms satisfy their optimality conditions, budget constraints are satisfied and all markets clear at all dates for a given path of the productivity vector $\left\{A_t^R, A_t^U, N_t^R, N_t^U\right\}$.*

4 Special Cases

In order to build intuition regarding the model, we now specialize our environment and study two extreme cases: one with frictionless labor market across sectors and locations, and one with prohibitively high reallocation cost which prevents worker reallocation. This exercise also allows us to highlight the minimal features of the model necessary to explain the data facts.

4.1 Frictionless Labor Market

We make five assumptions:

Assumption 1. $\alpha^R = \alpha^U = \beta^R = \beta^U = \beta$.

Assumption 2. $\tau_t^R = \tau_t^U = 0$.
Assumption 3. $\tau_t = 0$.
Assumption 4. $\frac{A_t^R}{A_t^U} \geq \frac{N_t^R}{N_t^U}$.
Assumption 5. $\varepsilon_t^U = \varepsilon_t^R$ for all t.

Assumption 1 implies that production technologies differ across locations and sectors solely due to differences in total factor productivities and nothing else. Assumption 2 sets the training costs of switching to the non-agricultural sector to zero. Assumption 3 makes mobility across locations costless. Assumptions 2 and 3 jointly convert our environment into a model with no frictions in labor allocations, either across sectors or across locations. Assumption 4 implies that rural locations are relatively more productive in producing the agricultural good while urban locations are relatively more productive in producing the non-agricultural good. Assumption 5 says that there are no amenity differences between urban and rural locations. Recalling equation (3.8), this amounts to assuming that $\phi = 0$ and $\bar{\varepsilon}^j = \bar{\varepsilon}$ for all j. Hence, Assumption 5 when combined with Assumption 3 implies that the migration indifference condition for the rural young reduces to $w_t^{UA} = w_t^{RA}$. This greatly simplifies the analytical illustration of the effects of productivity shocks in this economy as the relative urban labor supply function effectively becomes infinitely elastic.

Under Assumptions 1 and 2, the optimality condition for sectoral labor allocation reduces to

$$\frac{L^{UN}}{L^{RN}} = \gamma_t \frac{L^{UA}}{L^{RA}}, \quad \gamma_t \equiv \left(\frac{A_t^R / A_t^U}{N_t^R / N_t^U} \right)^{\frac{1}{1-\beta}} \tag{4.22}$$

Moreover, under Assumption 4 above, $\gamma_t \geq 1$ for all t. Hence, the non-agricultural sector employs relatively more urban labor while the agricultural sector is more intensive in rural labor.[27]

It is important to note that in this model, along paths with a constant population and stationary productivities, the transition to steady state occurs with a maximum lag of one period. Given any initial distribution of the population between rural and urban locations, location adjustments take place at date $t = 0$. From $t = 1$ onwards, there are no further changes in the population distribution between rural and urban locations and the economy remains stationary.

4.1.1 Productivity Shocks

There are four different productivity parameters in the model: A^R, A^U, N^R, N^U. This allows us to examine the effects of both aggregate productivity shocks as well as sectoral productivity shocks on the economy.

[27] Note that our assumption of zero population growth implies that L_t is constant and independent of time.

Aggregate productivity growth We first consider aggregate productivity shocks that raise the levels of A^R, A^U, N^R, N^U equiproportionately at each date. Specifically, suppose the productivity processes are given by

$$A_{t+1}^j = (1+g)\, A_t^j, \quad j = R, U \tag{4.23}$$
$$N_{t+1}^j = (1+g)\, N_t^j, \quad j = R, U \tag{4.24}$$

where $g > 0$ gives the rate of balanced aggregate productivity growth. Hence, $\frac{A_t^U}{A_t^R}, \frac{N_t^U}{N_t^R}, \frac{N_t^U}{A_t^U}$ and $\frac{N_t^R}{N_t^U}$ all remain unchanged even though the levels of all the productivity parameters rise permanently at each t. Model predictions for this case are summarized in the following Proposition:

Proposition 1 *Under Assumptions 1-5 and aggregate productivity growth given by equations (4.23)-(4.24), there is a secular decline in the agricultural employment share of overall labor as well as rural labor and urban labor individually. This structural transformation is accompanied by rising urbanization and a secular fall in the relative price of the agricultural good.*

Proof. See Appendix. ■

Intuitively, this is the standard structural transformation mechanism in models with non-homothetic preferences. The aggregate productivity increase raises the demand for non-agricultural goods more than the demand for the agricultural good. This relative demand shock pushes up the relative price of the non-agricultural good which, in turn, causes a reallocation of workers from agriculture to non-agriculture in both locations. The additional aspect here is that the higher relative price of non-agricultural goods causes a greater incipient rise in the non-agricultural wage in urban locations since they are more productive in producing the non-agricultural good. This results in rising urbanization as young rural individuals relocate to urban locations in order to arbitrage the wage differential.

Sector-biased productivity change We now examine the impact of productivity changes that are biased towards the non-agricultural sector. In particular, suppose the economy is initially in steady state with constant productivities in all sectors given by $A_0^R, A_0^U, N_0^R, N_0^U$. Now suppose that at $t = 0$ news arrives that the productivity process from $t = 1$ will be

$$A_t^j = (1+\varepsilon g)\, A_0^j \text{ for all } t \geq 1, \;\; j = R, U \tag{4.25}$$
$$N_t^j = (1+g)\, N_0^j, \quad j = R, U \tag{4.26}$$

where $\varepsilon < 1$ and $g > 0$. The shock permanently raises $\frac{N_t^j}{A_t^j}$ from $t \geq 1$ for $j = R, U$. Hence, this is a non–agriculture biased productivity change. We collect model predictions for this case in the following Proposition:

Proposition 2 *Under Assumptions 1-5 and N-sector-biased productivity growth given by equations (4.25)-(4.26), there is a decline in the agricultural employment share of overall labor as well as of rural labor and urban labor individually. This structural transformation is accompanied by an increase in the degree of urbanization. The movement in the relative price of the agricultural good, however, is ambiguous.*

Proof. See Appendix. ■

Qualitatively, the predictions for structural transformation are identical to the case of aggregate productivity increase. The main difference in outcomes between the two cases is the impact on the relative goods price. The faster growth in non-agricultural productivity causes an increase in the relative supply of the non-agricultural good. This supply effect provides a counter-weight to the increase in the relative demand for the non-agricultural good. This makes the effect on relative price ambiguous.[28]

4.2 No Migration

The analysis above focused on the extreme case of frictionless labor markets across sectors and locations. We now consider the other extreme case in which frictions associated with location switching are so large that there is no labor relocation across locations, i.e., L^R and L^U are constant over time. Effectively, we let $\tau_t \to \infty$, but continue to maintain Assumptions 1, 2, 4, and 5. In this special case, the location migration indifference condition (equation (3.15)) does not apply. Put differently, the relative supply of urban labor is completely inelastic.

What is the impact of an aggregate productivity increase in this economy? Assume that productivity in both sectors in both locations rises by the same proportion:

$$A_t^j = (1+g)^t A_0^j, \quad j = R, U \tag{4.27}$$

$$N_t^j = (1+g)^t N_0^j, \quad j = R, U \tag{4.28}$$

where A_0^j and N_0^j denote the initial productivity levels in the two sectors for $j = R, U$. The

[28] A different but related experiment would be an unanticipated, permanent increase in sectoral or aggregate productivity starting from a steady state. Thus, suppose all sectoral productivities are constant over time and the economy is in steady state. Now, suppose A^j rises permanently by a factor γ^A and N^j rises by a factor γ^N for $j = R, U$. In this case, the location indifference condition at the initial date will clearly not hold since workers cannot move within that period. Hence, the initial distribution of L would be exogenously given and rural and urban wages would not be equalized at the initial date. It can be shown that such an unanticipated increase in aggregate productivity would induce a structural transformation with rising employment shares of the non-agricultural sector in both locations. However, the urban-rural wage gap would widen and the price of the non-agricultural good would rise in the period of the shock. Analogously, a permanent unanticipated increase in the productivity of sector-N relative to sector-A (an N-sector biased technological improvement), would have similar effects on the structural transformation but have ambiguous effects on the relative price of good N and on the wage gap in the period of the shock. The adjustments in the next period to both shocks would be as described in the propositions above.

model predictions in this case are summarized in the proposition below.

Proposition 3 *Under Assumptions 1, 2, 4 and 5, and aggregate productivity growth given by equations (4.27) and (4.28), the agricultural employment share of both rural labor and urban labor declines. This structural transformation is accompanied by a widening of the urban-rural wage gap and a rise in the relative price of the non-agricultural good.*

Proof. See Appendix. ■

Intuitively, aggregate productivity growth triggers an increase in the relative demand for non-agricultural goods which leads to a rise in the non-agricultural relative price. This raises the wages of non-agricultural workers relative to the wages of agricultural workers. The consequence of this is worker reallocation from agricultural to non-agricultural employment. Since urban locations are predominantly non-agricultural, the urban-rural wage gap widens. This is the standard demand-driven explanation of structural transformation. We refer to this as the "demand effect". Since there is no spatial reallocation in this case, there are no countervailing forces to prevent the widening of the wage gap.

4.3 General insights of model

The propositions above highlight that productivity changes in the model generate two competing effects on prices and allocations. First, from Proposition 3 we have the "demand effect" which raises the relative demand for the non-agricultural good and factors that are used intensively in its production. This causes a rise in the relative price of the non-agricultural good and an increase in the relative urban wage.

The second effect, highlighted by Propositions 1 and 2, arises as a consequence of the first effect. Specifically, widening urban-rural wage gaps trigger worker reallocation from rural to urban areas. Net reallocation into urban areas increases labor supply in those areas. This in turn leads to a higher production of non-agricultural goods, moving the sectoral terms of trade against them. This brings wages of urban workers closer to the wages of rural workers, i.e. the urban-rural wage gap declines. We refer to this effect as the "urbanization effect". Without any labor market frictions, the urban-rural wage gap completely disappears.

When reallocation costs are neither zero nor infinite (the cases analyzed in 4.1 and 3, respectively), but are at the intermediate level, the relative labor supply schedule becomes an upward sloping function of the relative urban wage. Then, an aggregate productivity increase has an additional effect: as productivity rises the relative cost to change locations falls as long as τ does not rise at the same rate as productivity. With the fixed migration cost τ in the model, for the same relative urban wage more people now move to urban areas, i.e., an aggregate productivity shock also causes a rightward shift in relative urban labor supply. For a given relative urban labor demand, this shift of the supply function will cause the relative urban wage to fall.

In general, a rise in productivity would shift both the relative demand and relative supply functions for urban labor. Clearly, the net effect on urban-rural wage gaps will depend on which of these channels dominates. We study the relative strength of these effects in China and India using a calibrated version of the full model.

5 Quantitative Results

We now quantitatively assess the ability of the full model to explain the observed rural-urban wage dynamics along with the aggregate macroeconomic facts. To do so we calibrate the model separately for India and China to match their conditions at the initial dates of our data sample. In particular, for China we use year 1988 as representing its initial steady state, while for India we use year 1983. We then conduct the following experiment. Keeping all the calibrated parameters unchanged, we feed the measured changes in sectoral labor productivities in China during 1988-2008, and in India during 1983-2010 into the model and examine their effects on goods prices, factor prices, and sectoral and spatial factor allocations.

5.1 Calibration for the 1980s

We calibrate the model parameters to match the key moments of wages and employment in the data. More precisely, we choose eleven parameters that minimize the distance between eleven moments in the data in the 1980s and in the model. Our first calibration target is the urban shares of employment, which was equal to 22% in India in 1983, and to 26% in China in 1988.[29] Second, we match the sectoral distribution of the labor force in rural and urban areas summarized in Table 1. This gives us two independent moments to target. Third, we target the four conditional wage gaps also presented in Table 1: the two "within" sector wage gaps and the two "between" sector gaps. Our eighth data target is the output share of agriculture in total GDP. In India in 1983 this was 36%, while it was 17% in China in 1988.

Our last three data targets are moments that characterize consumption expenditures in the two countries: the share of agriculture in total household consumption; the home production share of non-agricultural goods and services; and the minimum consumption level of agricultural goods. In linking the model to the data above, we follow the value-added approach to interpreting a sector.[30] To keep the model internally consistent we define the arguments in the utility function in value added terms as well. We compute agricultural consumption in value

[29] See appendix for more information.

[30] See Herrendorf, Rogerson, and Valentinyi (2013b) for a careful discussion of value added and final expenditure approaches to interpreting the data.

added terms as the agricultural value added, and non-agricultural value-added consumption as non-agricultural value added minus investment.[31] This gives us the share of agricultural value added in total consumption equal as 47% in India and 33% in China. We target the home production share in the consumption of non-agricultural goods and services at 30% in both countries. This number is implied by the time allocation statistics in the US over our sample period, where US households spent on average 12 hours per week in home production and 40 hours per week in market employment. These numbers are also similar to those reported in China since the mid-2000s.[32]

We pin down the minimum consumption level of agricultural goods by following Anand and Prasad (2010) who estimated minimum consumption requirement value to be 50% of food consumption for a sample of six emerging economies, including India. We adjust this number to account for potential differences in minimum food consumption requirements in rural and urban areas. Specifically, we use the estimates of daily calorie needs by age, gender, and physical activity level from the Center for Nutrition Policy and Promotion at the United States Department of Agriculture to compute the necessary adjustments. We assume that rural activities are more strenuous than urban activities and thus require larger calorie intake. Assuming that rural work falls into the "active" activity category, while urban work falls into "sedentary" activity category, we estimate the resulting calorie-intake *premium* in rural areas to be equal to 25%.[33] Therefore, we raise the minimum consumption requirement in rural areas by 11.8%, and reduce it in urban areas by 10.6%, leaving the weighted average of the two areas equal to 50% of food consumption.[34] We use the same numbers for China in 1988.

Our free parameters are the technology parameters α and β, the training costs τ_U and τ_R, migration cost τ, the sectoral productivity levels A^U, N^R, N^U (we normalize $A^R = 1.5$), the agricultural consumption share θ, the minimum agricultural consumption parameter $\bar{a}$

[31] See appendix C.1 for data sources.

[32] See Conference Board.

[33] See http://www.cnpp.usda.gov/sites/default/files/usda_food_patterns/EstimatedCalorieNeedsPerDayTable.pdf

[34] These numbers were obtained by solving the following system of equations for $\bar{a}^R$ and $\bar{a}^U$:

$$\begin{aligned} \bar{a}^R &= (1+\Delta)\bar{a}^U \\ \bar{a} &= s^{RA}\bar{a}^R + (1 - s^{RA}\bar{a}^U), \end{aligned}$$

where Δ is the adjustment factor, $\bar{a}$ is the aggregate minimum agri consumption requirement, and s^{RA} is the rural agri consumption share. We assume $\Delta = 0.25, \bar{a} = 0.5$ and $s^{RA} = 0.47$. The rural agri consumption share, s^{RA}, is approximated as follows.

$$s^{RA} = \frac{MPCE^R * F^R}{MPCE^R * F^R + MPCE^U * F^U},$$

where $MPCE^j, j = R, U$ is the monthly per capita consumption expenditures in location j, while $F^j, j = R, U$ is the food share of total consumption expenditures in location j. All numbers were obtained using 1983 NSS data for India. Specifically, we used the following values: $MPCE^R = 200$, $MPCE^U = 250$, $F^R = 62\%$, $F^U = 55\%$.

and the home production of services parameter $\bar{n}$. These eleven parameters are calibrated to jointly match the eleven data moments described above.

We also need to parameterize the process for amenities in rural and urban locations. In particular, $\frac{\bar{\varepsilon}^R}{\bar{\varepsilon}^U}$ characterizes the relative steady-state level of amenities available in rural and urban locations. Since this ratio is not directly observable and no estimates are available in the existing literature, we make the neutral assumption that rural and urban locations do not differ in terms of amenities they offer to their residents *in the steady state*, i.e. $\frac{\bar{\varepsilon}^R}{\bar{\varepsilon}^U} = 1$.[35],[36]

Parameter ϕ captures the elasticity of available amenities with respect to local population changes, with $\phi < 0$ implying negative externalities associated with population growth due to reallocation into a location. We choose the value for parameter ϕ, such that the model, in response to the observed sectoral productivity changes, reproduces the observed change in the urban employment share over our sample period.[37]

Lastly, we set interest rate to be constant $R_t = R = 1$ for all t in both countries and let T_t be determined endogenously. The resulting parameter values are summarized in Table 3. The resulting fit of the model for the 1980s is summarized in Table 4 for China and Table 5 for India.

A few of our estimates are worth discussing further. First, we estimate the spatial reallocation cost parameter τ to be significantly larger in China than in India. This is not surprising as the data forces this due to the large and persistent urban-rural wage gaps in each sector in China. The proximate data counterpart of the high estimated τ in China would be the Hukou system of household registration which made migration to urban areas very costly for rural households.

Second, the externality parameter ϕ is negative in both countries, but is estimated to be significantly larger (in absolute value) in China than in India. This suggests that congestion externalities associated with reallocation of rural workers into urban areas are larger in China than in India. The higher estimate for China is explained by its relatively low urbanization pace despite the large and persistent urban-rural wage gaps observed there. In India, on the other hand, the rate of urbanization is roughly in line with the incentives provided by the urban-rural wage gaps there. As a result, the estimated ϕ is significantly lower in India.[38]

[35] It is easy to see from the locational indifference condition given by equation (3.15) that for a given steady state distribution of the workforce between the urban and rural locations, there is a downward sloping relationship between the migration cost parameter τ and the relative amenities term $\frac{\bar{\varepsilon}^R}{\bar{\varepsilon}^U}$. Specifically, the lower is $\frac{\bar{\varepsilon}^R}{\bar{\varepsilon}^U}$ the higher must τ be to rationalize the given distribution of workers.

[36] It is worth noting that the parameter ϕ cannot be estimated to target a steady state moment since under our formulation the externality is zero in steady state. Consequently, we estimate ϕ in each country to target the change in their relative urban employment share over the sample period.

[37] It is important to note that we do not impose $\phi < 0$ in our calibration strategy. Rather, we let it be whatever it needs to be in order for the model to match the change in the observed urban employment share during our sample period.

[38] In a related paper Dinkelman and Schulhofer-Wohl (2015) use South African data to show that negative congestion externalities of migration are much higher when land markets are missing. Given the more restricted

	parameter	China	India
fixed parameters			
Intertemporal elasticity of substitution	σ	2	2
Relative steady-state amenities level b/n R and U locations	$\frac{\bar{\varepsilon}^R}{\bar{\varepsilon}^U}$	1	1
Productivity in rural agri	A^R	1.5	1.5
estimated parameters			
Labor weight in A sector	α	0.26	0.08
Labor weight in N sector	β	0.50	0.23
Training cost for U households	τ_U	-0.06	0.23
Training cost for R households	τ_R	0.18	0.19
Migration cost	τ	0.63	0.15
Productivity in rural non-agri	N^R	1.18	0.92
Productivity in urban agri	A^U	0.16	0.07
Productivity in urban non-agri	N^U	1.96	1.10
A consumption share	θ	0.40	0.45
Home production of non-agri goods	$\bar{n}$	0.43	0.41
Minimum agricultural consumption	$\bar{a}$	0.67	0.71
Externalities from migration	ϕ	-0.29	-0.05

Table 3: Model parameters, 1983

Third, our estimated labor share parameters for agriculture are lower than those in non-agriculture for both China and India. This is not unusual for developing countries where the share of land in agriculture is often quite large. Thus, in their study of Indian agriculture, Abler, Tolley, and Kripalani (1994) estimate the labor share in agriculture to be around 60% of the labor share in non-agriculture.[39]

Fourth, our parameter estimates imply a negative training cost for urban households in China. In effect, this implies an educational subsidy for urban households. This is less surprising than one might think. There is a large literature on the redistribution of resources in China towards urban areas through a policy of preferential investment in human capital development. For example, Heckman (2005) finds that the fraction of tuition fees per child in household income is twice as high in rural areas compared to urban areas in China. He also documents that per pupil education expenditure by the state varies systematically (positively) by the wealth of the region. Given that the urban regions are wealthier, it implies that state

land markets in China as compared to India, the Dinkelman and Schulhofer-Wohl (2015) results provide an independent explanation for why the congestion externality paraemter ϕ is higher in China than in India.

[39] We should note that Tombe and Zhu (2015) estimate the sectoral labor shares in agri and non-agri in China to be significantly higher. Part of the reason for the difference is that they use a different model with intermediate inputs which implies that to estimate the factor share of gross output they need to estimate the sectoral value added shares of gross output and the factor's share of value added. For the factor share of value added, Tombe and Zhu (2015) do not use Chinese data. Instead, they use the numbers estimated for the USA by Caselli and Coleman (2001) who found that the labor share was identical in agri and non-agri for the USA. Without getting into the merits of this assumption for developing countries, we would like to highlight that our method imputes these numbers directly for China and India by matching the data moments for the 1980s.

education expenditure is also higher in these regions relative to the poorer rural regions.

With these parameter values the model fits the data in the 1908s quite well. The employment shares are spot on in both China and India. Conditional wage gaps are also matched quite precisely, especially in India. In China the model slightly underpredicts the urban-rural within-agri and urban between sectors wage gaps, and overpredicts the other two wage gaps. The differences, however, are small. In terms of the overall mean urban-rural wage gap – which was not targeted in the calibration – the model slightly underpredicts it in India – 66% in the data, relative to 64% in the model. In China, however, the difference is larger. In general, the overall average urban-rural wage gap in the data will not necessarily add up to the average between and within wage gaps since it is a non-linear function of these gaps (see, for instance, equation (2.1)). In the model, on the other hand, the overall wage gap is computed from the between and within wage gaps. Thus, even if the between and within wage gaps are very similar in the model and in the data, the overall wage gap does not need to be.

We should also note that the model matches exactly the urban to rural reallocation of labor in the data during 2008-1988 in China and 2010-1983 in India. Recall that the parameter ϕ was calibrated to match the change in urbanization share during the respective periods in the two countries, which could be due to differences in natural population growth in rural and urban areas, net migration into urban areas from rural locations and reclassification of rural areas into urban areas. Urban natural population growth is still below rural natural population growth in both China and India. As we argued before the other two channels have both played an important role in the rise of urbanization of both countries. For our focus on locational wage gaps it is the relative aggregate labor supply in the two locations that matters, and since the latter can be attributable to both reallocation forces, we choose to not explicitly distinguish the migration and reclassification channels in our analysis.

One dimension in which the model does not fit the data very well is the shares of agriculture in consumption and output. In both China and India, the model overpredicts these shares. The reason is that any model with a neoclassical production function, will find it difficult to simultaneously account for the very high share of employment in agricultural sector and observed wage gaps. To illustrate this point, consider a bare-bones neoclassical two-sector model with Cobb-Douglas production functions in the agricultural and non-agricultural sectors. When labor markets are competitive and labor is free to move across sectors, such a model would imply that wages, and thus the marginal value products, should be equalized across sectors. The latter can be expressed in terms of agriculture's share of employment ($l_a = \frac{L_a}{L_a+L_n}$) and output ($y_a = \frac{VA_a}{VA_a+VA_n)}$) as

$$\frac{(1-y_a)}{(1-l_a)} = \frac{y_a}{l_a}.$$

That is, the ratio of each sector's share in value added to its share in employment should be the same in the two sectors. The ratio of the left-hand-side to the right-hand-side in the expression above is the agricultural productivity gap in Gollin, Lagakos, and Waugh (2014). Based on the data in Tables 4 and 5, this gap was equal to 7.25 in China in 1988 and 3.06 in India in 1983. In contrast, the wage data in each country gives non-agri to agri wage gaps at much smaller 1.45 in China in 1988 and 2.07 in India in 1983. As a result, there is an internal tension in the model when trying to match the agricultural productivity gap and sectoral wage gap simultaneously. Moreover, since there is no investment or government sector in the model, sectoral consumption and output gaps are very similar. Since we are primarily interested in understanding wage gaps and their changes over time, we choose to assign a higher weight to wage gap moments in our parameter estimation for the 1980s relative to sectoral output and consumption moments.

5.2 Results

How much of the observed dynamics in urban-rural wages in the two countries can be accounted for by the country-specific measured changes in productivity growth in the two sectors? To answer this question we feed the measured sectoral productivity growth in the two countries into the model while keeping all other parameters unchanged. As presented earlier, in India agricultural labor productivity increased by 67% between 1983 and 2010, while non-agricultural labor productivity increased by 200%. The corresponding numbers in China between 1990 and 2008 were 163% and 338%.

Specifically, we use the following dynamic equations for sectoral productivities in each country:

$$A_{t+1}^{j} = A_{t}^{j}\left(1 + g_{t}^{A}\right), \quad j = R, U \tag{5.29}$$

$$N_{t+1}^{j} = N_{t}^{j}\left(1 + g_{t}^{N}\right), \quad j = R, U \tag{5.30}$$

where g_t^A and g_t^N are productivity growth rates in agricultural and non-agricultural sectors, respectively. These sectoral productivities are taken as exogenous by households and firms. They are equal to the growth rates in sectoral labor productivities measured in the data, and are assumed to be the same across locations (we relax this assumption in the next section). The results for China are summarized in Table 4, and for India in Table 5.

Several features stand out. First, changes in productivity lead to an increase in urbanization of the labor force in both countries. In particular, the urban employment share increased by 8 percentage points in India between 1983 and 2010, and by 9 percentage points in China between 1988 and 2008. The model reproduces these urbanization dynamics. This is not surprising since the parameter ϕ was calibrated to precisely match this. The part that is noteworthy is that ϕ is negative for both countries, and that it is larger in absolute value for China.

These two features indicate: (a) absent these negative congestion externalities on amenities, the model would generate greater urbanization in both countries in response to productivity growth; and (b) the higher productivity growth in China implied a higher desired urban labor force growth, which necessitated a larger ϕ in absolute value in China to force the model to match the actual increase in the urban employment share.

Second, the share of the workforce employed in agriculture declines in both rural and urban areas in both countries. Thus, the model successfully replicates the patterns of sectoral transformation observed in the data for China and India.

Table 4: Model and data: China, 1988 versus 2008

	1988		2008	
	data	model	data	model
employment shares:				
L_U	0.26	0.26	0.35	0.35
L_{RA}	0.79	0.79	0.66	0.57
L_{UA}	0.05	0.05	0.03	0.03
wage gaps:				
within A	1.844	1.826	2.778	1.782
within N	1.290	1.317	1.605	1.613
R between	1.285	1.312	1.272	1.088
U between	0.984	0.947	0.751	0.985
overall mean	1.379	1.626	1.614	1.692
aggregates:				
N/A relative price	1.00	1.00	0.779	0.924
A share of Y	0.17	0.65	0.06	0.47
A share of C	0.33	0.62	0.16	0.46

Third, following productivity shocks, the model predicts that the mean wage gap between urban and rural workers should *decline* in India but *increase* in China. This reproduces the pattern in the data though the model somewhat under-predicts the absolute value of the actual changes in both countries. In India, the overall mean wage gap between urban and rural areas falls by 0.20 in response to sectoral productivity growth, which accounts for 57% of the decline in unconditional urban-rural wage gap in the data. In China on the other hand, the model predicts a 7 percentage point increase in the overall urban-rural wage gap which is about 35% of the actual unconditional increase in the data. Given that our interest is in explaining the part of the wage gap that is not accounted for by the observed changes in worker attributes like education and skills, we view these results to be indicative of strong success of the model in explaining the aggregate wage patterns.

We should note that the model broadly reproduces the data pattern of a decline in the inter-sectoral wage gap in urban and rural locations in both China and India. In India this decline in the between-sector wage gap was sufficient to compensate for the relatively stable

Table 5: Model and data: India, 1983 versus 2010

	1983		2010	
	data	model	data	model
employment shares:				
L_U	0.22	0.22	0.30	0.30
L_{RA}	0.78	0.78	0.66	0.59
L_{UA}	0.11	0.11	0.07	0.04
wage gaps:				
within A	0.934	0.937	1.027	1.204
within N	1.082	1.080	0.994	1.211
R between	1.962	1.961	1.679	1.405
U between	2.259	2.260	1.709	1.414
overall mean	1.664	1.642	1.310	1.441
aggregates:				
N/A relative price	1.00	1.00	0.714	0.844
A share of Y	0.36	0.71	0.16	0.58
A share of C	0.47	0.68	0.23	0.56

urban-rural wage gap within each sector, thereby reducing the overall urban-rural wage gap. In China however, the decline in the inter-sectoral wage gaps in each location was insufficient to overcome the expansion in the urban-rural wage gaps within each sector, thus inducing a widening of the overall urban-rural wage gap.

Fourth, the model predicts a decline in the relative price of non-agricultural goods in both countries, consistent with the empirical evidence. The model also predicts a fall in the share of agriculture in output and consumption, with the declines being comparable to those found in the data.

Overall, our results suggest that aggregate factors have played an important role in urban-rural dynamics in India and China in the past 20-30 years. Growth of agricultural productivity and an even faster growth of non-agricultural productivity can account for a large share of the sectoral transformation and relative price dynamics in both countries. The same forces also account for a large part of the observed wage convergence between urban and rural areas in India and predict some divergence in urban-rural wages in China. Furthermore, these factors induce within-sector and between-sector wage adjustments that are consistent with the data.

To understand these results for urban-rural wage gaps, recall that the model relies on two competing effects – the demand effect, which leads to wage divergence; and the urbanization effect which leads to wage convergence. The urbanization effect is stronger in India where the estimated migration costs and negative migration externalities are smaller. In contrast, for China we estimate larger migration costs and higher negative externalities arising from worker reallocation into cities. As a result, the urbanization effect is weaker and the urban-rural wage gap rises over time.

5.2.1 Agglomeration externalities in production

The baseline model that we developed introduced a negative congestion externality of migration on urban amenities. That formalization ignored a second often discussed externality of migration which is its positive effect on aggregate productivity. This positive production externality is typically proposed as an explanation for the concentration of economic activity in locations as well as the growth of cities. This is potentially an important margin for understanding the process of urbanization, so we next explore the role of positive agglomeration externality in production.

We postulate that urban total factor productivity growth gets an additional boost from growth in the urban labor force. Specifically, we assume that

$$1 + g_t^{Uk} = \left(1 + X_t^k\right)\left(1 + g_t^{Rk}\right), \qquad k = A, N \tag{5.31}$$

where

$$1 + X_t^k = \left(\frac{L_t^U}{L_{t-1}^U}\right)^{\phi_k}, \quad k = A, N \tag{5.32}$$

Clearly, as long as there is positive reallocation into urban areas so that $X^k > 0$, productivity *growth* in both sectors is going to be higher in urban areas relative to rural locations.[40]

There are two interesting special cases here. First, for $\phi_A = \phi_N$ the urban production externality is identical across sectors. Second, when $\phi_A = 0$ the externality only affects the non-agricultural sector while agricultural productivities grow at the same rate in urban and rural locations. Notice that since $L_t^U = L_{t-1}^U + M_t$, i.e., the urban population at time t is the sum of the urban population at $t-1$ plus the new migrants at t, equation (5.32) can be written as

$$1 + X_t^k = \left(1 + \frac{M_t}{L_{t-1}^U}\right)^{\phi_k}, \quad k = A, N$$

Reallocation from rural to urban areas now has two effects. On the one hand, negative congestion externalities reduce the level of urban amenities according to equation (3.8). This is unchanged from our baseline case. However, now the process of urban reallocation also raises productivity in urban locations relative to their rural counterparts. Thus, productivity growth becomes urban-biased in this case.

Our identification strategy for the parameters of the model is the same as before: we calibrate the parameters to target the same set of moments in the initial period. We then feed into the model the measured average sectoral productivity growths during the sample period.

[40] Our assumption that urban productivity growth depends on urban population growth is consistent with an environment where productivity growth depends upon ideas that are carried by individuals. Implicitly, the formulation is a stand-in for environments where faster population growth in a location induces greater exchange of ideas and consequently faster TFP growth.

Notice that the average sectoral productivity growths for the country as a whole are given by

$$1 + g_t^k = s_t^{Rk}\left(1 + g_t^{Rk}\right) + \left(1 - s_t^{Rk}\right)\left(1 + g_t^{Uk}\right), \quad k = A, N$$

where s^{Rk} is the fraction of sector-k labor working in rural locations. Hence, the average gross sectoral growth rate is just the weighted average of the corresponding location-specific sectoral growth rates. We measure both g^k and s^{Rk} from the data. We assume that $\phi_A = \phi_N = -\phi$ and calibrate the parameter ϕ to match the net rural-to-urban reallocation of workers during the sample period.[41] This gives a value of $\phi = -0.31$ in China and $\phi = -0.055$ in India. These parameter estimates imply that the observed increase in the urban employment shares induced a 10% boost to urban productivity in China, and 1.7% boost to urban productivity in India.[42]

Table 6 report the changes in the relevant variables predicted by the model under agglomeration externalities as well as those under the baseline case where these production externalities are absent. The table makes clear that a more rapid increase in relative urban productivity stalls the wage convergence in both countries. For instance, in India the resulting wage convergence between urban and rural labor is smaller, with the gap declining by 0.18 compared to 0.20 in the case of shocks that are symmetric across locations. In China, urban and rural wages are diverging by 0.11 instead of 0.07. Crucially, the wage gaps that change the most in both countries are the urban-rural wage gaps within sectors. Since we have assumed that agglomeration effects are symmetric across sectors, the inter-sectoral wage gaps remain relatively unaffected by the introduction of agglomeration externalities.

These results are best understood by noting that the introduction of agglomeration externalities in urban production has two effects. On the one hand, the rural-to-urban migration induces greater urban population growth and density. This raises urban productivity in both sectors which increases the relative wages of urban workers, thereby widening the urban-rural wage gap. On the other hand, a greater incentive for migration, all else equals, drives down the wage gap due to the increase in relative urban labor supply. In our calibration of the two countries, the first effect dominates and the wage gap expands relative to the case where productivity growth is symmetric across locations.

We view these results as indicative of the robustness of our baseline model to allowing for positive externalities of migration through agglomeration effects in production. However,

[41] This restriction forces the amenities congestion externality parameter and the agglomeration productivity externality parameters to be the same in absolute value but of opposite signs. Hence, negative amenities externalities would coincide with positive agglomeration production externalities. This restriction ties our hands in terms of fitting the data but also provides some discipline on the calibration exercise due to the relaitve paucity of independent estimates of these effects.

[42] Note that our specification implies that $1 + g^{Rk} = \frac{1+g^k}{1+(1-s^{Rk})X^k}$, $k = A, N$ which can be used to infer the location-specific productivities by substituting in the measured values for g^k, X^k and s^{Rk} from the data.

they also resonate particularly for China where a number of authors have found evidence of location biased factor allocation. For instance, in China rates of return on capital investment tend to be higher in smaller cities and rural areas, suggesting that urban locations are typically favoured for capital allocation by the government. This raises the productivity of urban workers relative to workers in rural areas (see Jefferson and Singh (1999); Bai, Hsieh, and Qian (2006)). Given that the baseline model was underpredicting the wage divergence in China, the addition of agglomeration effects of migration on urban productivity appears to provide one rationalization for the higher observed wage divergence in the data. It also provides a background rationalization for the evidence in Jefferson and Singh (1999) and Bai, Hsieh, and Qian (2006).

Table 6: Changes under agglomeration production externalities

	Baseline		Agglomeration	
	China	India	China	India
changes in employment shares:				
L_U	0.09	0.08	0.09	0.075
L_{RA}	-0.22	-0.19	-0.19	-0.187
L_{UA}	-0.02	-0.07	-0.02	-0.067
changes in wage gaps:				
within A	-0.043	0.267	0.004	0.289
within N	0.296	0.131	0.332	0.147
R between	-0.224	-0.556	-0.220	-0.551
U between	0.039	-0.846	0.038	-0.849
overall mean	0.066	-0.201	0.111	-0.177
changes in aggregates:				
N/A relative price	-0.076	-0.156	-0.095	-0.072
A share of Y	-0.18	-0.13	-0.17	-0.129
A share of C	-0.16	-0.12	-0.16	-0.119

Note: The results under Baseline report the changes in the variables of interest predicted by the model in response to productivity changes (for India, see Table 5; for China see Table 4). The results under Agglomeration report the results of introducing an urban agglomeration effect on productivity due to migration while keeping overall sectoral productivity growth unchanged as under the baseline case.

5.3 Experiments

The growth experience of China and India are recognized to differ from each other in two key aspects. First, China's growth takeoff was much sharper with growth rates being significantly higher than in India. Second, the role of the state in controlling and directing labor flows across locations through the Chinese household registry system (the Hukou) was significantly greater in China relative to India. The Hukou system effectively raised the cost of labor reallocation from rural to urban locations. How important were these two factors for understanding the differences in the urban-rural dynamics in the two countries? The model allows us to conduct counterfactual experiments to address these questions. We first derive the counterfactual path of urban-rural inequality in India if its growth rate had been like in China. Next, we ask what would happen to urban-rural inequality in China if reallocation costs were reduced to India's levels.

5.3.1 India growing like China

For this experiment we use the model calibrated to India and feed into it the measured sectoral productivity growth for China, keeping all other parameters unchanged. Specifically, we assume that agricultural labor productivity increased by 163% and non-agricultural labor productivity increased by 338% in India. The results are presented in column labelled "Exp1: High Growth, India" of Table 7. The table reports the changes in the relevant indicators as well as the corresponding changes in those variables under the baseline case.

Under the faster sectoral productivity growth, the model predicts a larger reallocation from rural to urban areas, leading to the urban labor share rising to 33% as opposed to 30% in the benchmark model and in the data for India. The labor reallocation from agricultural activities is also larger in this case with the agricultural employment share declining from 0.78 to 0.5 in rural areas and from 0.11 to 0.03 in urban areas. Given the larger urbanization effect, there is a greater urban-rural wage convergence with the overall mean wage gap contracting from 1.642 in 1983 to 1.388 in 2010 implying a 25 percentage point decline as opposed to the 20 percentage point decline in the benchmark case.

Notice that the effect on relative prices now, however, differs from the benchmark result. The relative price of non-agricultural goods *rises* when India grows like China. This is because the "demand effect" is also stronger when India's productivity rises faster. The greater increase in income induces a larger decline in the relative demand for the agricultural good which, in turn, reduces the relative price of agriculture (a rise in p). This effect is now strong enough to more than offset the positive supply effect arising from rural to urban reallocation and consequently causes a fall in the equilibrium relative price of the agricultural good.

5.3.2 China migration costs as in India

In the next experiment we reduce the migration externalities in China to their levels in India. This is equivalent to reducing the spatial reallocation cost. The results of this experiment are presented in the column labelled "Exp2: Low ϕ, China" of Table 7. Not surprisingly, lower costs lead to larger worker reallocation predicted by the model, with the urban employment share rising to 44% in 2008. This reallocation effect is large enough to overturn the urban-

Table 7: Experiments

	Baseline		Exp1: High growth	Exp2: Low ϕ
	China	India	India	China
changes in employment shares:				
L_U	0.09	0.08	0.11	0.18
L_{RA}	-0.22	-0.19	-0.28	-0.18
L_{UA}	-0.02	-0.07	-0.08	-0.02
changes in wage gaps:				
within A	-0.043	0.267	0.344	-0.456
within N	0.296	0.131	0.192	-0.076
R between	-0.224	-0.556	-0.747	-0.230
U between	0.039	-0.846	-1.055	0.035
overall mean	0.066	-0.201	-0.254	-0.323
changes in aggregates:				
N/A relative price	-0.076	-0.156	0.068	-0.163
A share of Y	-0.18	-0.13	-0.18	-0.18
A share of C	-0.16	-0.12	-0.16	-0.16

Note: Experiment1 applies China's sectoral productivity growth to India calibration; Experiment2 uses the migration externality parameter estimated for India ($\kappa = -0.05$) in China's calibration.

rural wage divergence in China. With lower migration externalities the urban-rural wage gap declines by 0.32, i.e. the model predicts wage convergence over time. Moreover, all conditional wage gaps decline.

This experiment suggests that bringing the reallocation costs in China down to their levels in India would produce a significant reduction in wage inequality in China. Put differently, the model suggests that restrictions on labor mobility were a key fact behind the widening urban-rural wage inequality in China during this period.

5.4 Testing the mechanism

The analytical results and the quantitative experiments presented above provide us with several key predictions about the relationship between wage gaps, productivity and urbanization in the model. First, Proposition 3 indicates that an increase in productivity is associated with a widening of the urban-rural wage gap when there is no reallocation across locations. A generalization of this is that conditional on a given size of the urban labor force, an increase in productivity widens the urban-rural wage gap. Second, Propositions 2 and 1, and the quantitative results in Table 7 show that for a given level of productivity, a decrease in reallocation costs is accompanied by greater urbanization and a narrowing of the urban-rural wage gap.

To test the model mechanism we collected state-level data on urban-rural wage (or income) gaps, urban employment and sectoral labor productivity in India and China. Specifically, for India we were able to put together a dataset covering 27 states for year 1983, 2000 and 2010, while for China we collected data on 30 provinces over 1990, 1995, 2002, 2007 and 2008 period. For India we used urban-rural wage gaps from the NSSO dataset, while for China we used urban-rural income gaps from Provincial Statistical Yearbooks. See Appendix for more details on the data sources and computations. We then estimated a regression of wage gaps in India (income gaps in China) on urbanization (as measured by urban employment share) and productivity (as measured by agricultural and non-agricultural labor productivity).

The regression results are presented in Table 8. Consistent with the predictions of the model, we find that urbanization tends to reduce urban-rural wage/income gap in both countries; while agricultural productivity tends to widen the same gaps. The effect of non-agricultural productivity is consistent with the model's predictions in China but goes in the wrong direction in India. Overall, we view these results as independent evidence supportive of the basic mechanisms formalized in the model.

Table 8: Testing model mechanism

	India U-R wage gap	China U-R income gap
Urban employment share	-1.317***	-1.574***
	(0.288)	(0.287)
Agri productivity	0.257**	1.033***
	(0.127)	(0.352)
Non-agri productivity	-1.463***	0.094***
	(0.542)	(0.036)
N	77	142

Note: For India the regressions are at the state level, while for China they are at the provincial level. Sectoral productivity is obtained as a ratio of sectoral output to sectoral employment. Regressions also include a constant (not reported).

6 Conclusion

This paper has studied the experience of China and India over the past thirty years to form a better understanding of the process of structural transformation of countries during the development process. A unique aspect of our work is that we focused on both quantities and prices. In addition, we have examined the process of structural transformation jointly with the process of urbanization. Our data analysis has revealed some interesting contrasts between China and India during this period. While the structural transformation experience of the two economies has been quite similar, the movements in wages have not. Specifically, while urban-rural wage gaps widened in China, they have contracted in India during this period. Interestingly, this has occurred in the backdrop of similar qualitative movements in the relative sectoral prices of goods. This evidence suggests to us that the standard practice of equating the agricultural sector with rural locations and non-agriculture with urban locations is not an innocuous abstraction. Indeed, a significant part of the structural transformation from agriculture to non-agriculture occurs *within* rural locations.

To explain the contrasting trends in the two countries, the paper formalized a two-sector, two-location overlapping generations model of structural transformation. Our model generates structural transformation through non-homothetic preferences and growth in agricultural productivity. We have showed that the model, calibrated to China and India, can generate the opposing movements in the urban-rural wage gaps observed in the two economies. Counterfactual exercises on our baseline model suggest that the restrictions on labor mobility in China from rural to urban areas were a key factor behind the widening urban-rural wage gaps there. An important ancillary result of the model is that it can account for the fact that the relative price of the non-agricultural good declined in both China and India during this period. It is important to note that this data fact is at odds with the rise in this relative price that is implied by standard non-homothetic models of structural transformation.

The key feature of our model that allows it to reproduce the wage and price movements is the endogenous urbanization margin. This introduces an endogenous change in the relative urban labor supply which both tends to reduce the relative price of urban labor (and hence reduces the urban-rural wage gap) and also lowers the relative price of non-agricultural goods which are intensively produced in urban areas. Using cross-province and cross-state data from China and India, respectively, we have shown independent evidence in support of these mechanisms embedded in the model.

We believe that our results suggest that the redistributional and allocational implications of structural transformation cannot be adequately analyzed without explicitly taking into account the accompanying urbanization that is generic to this process. The results also suggest that any analysis of the implications of structural transformation should include an explicit investigation of prices of both factors and goods in order to form a better understanding of the mechanics of the process as well as for devising appropriate analytical structures that best describe them. A larger cross-country study along these lines would appear to be a fruitful avenue for future work.

References

Abler, D., G. Tolley, and G. Kripalani (1994): *Technical change and income distribution in Indian agriculture.* Westview Press, Inc.

Acemoglu, D., and V. Guerrieri (2008): "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 116(3), 467–498.

Anand, R., and E. Prasad (2010): "Optimal Price Indices for Targeting Inflation under Incomplete Markets," IZA Discussion Papers 5137, Institute for the Study of Labor (IZA).

Bai, C.-E., C.-T. Hsieh, and Y. Qian (2006): "The Return to Capital in China," *Brookings Papers on Economic Activity*, 37(2), 61–102.

Baumol, W. J. (1967): "Macroeconomics of unbalanced growth: the anatomy of urban crisis," *The American Economic Review*, 57(3), 415–426.

Bryan, G., and M. Morten (2016): "Economic Development and the Spatial Allocation of Labor: Evidence From Indonesia," Working paper, Stanford University.

Caselli, F., and J. Coleman (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 109(3), 584–616.

Chan, K. W., and Y. Hu (2003): "Urbanization in China in the 1990s: New Definition, Different Series, and Revised Trends," *China Review*, 3(2), 49–71.

Dinkelman, T., and S. Schulhofer-Wohl (2015): "Migration, congestion externalities, and the evaluation of spatial investments," *Journal of Development Economics*, 114(C), 189–202.

Dudwick, N., K. Hull, R. Katayama, F. Shilpi, and K. Simler (2011): *From Farm to Firm: Rural-Urban Transition in Developing Countries.* World Bank, Washington DC.

Enflo, K., C. Lundh, and S. Prado (2014): "The role of migration in regional wage convergence: Evidence from Sweden 1860Ű1940," *Explorations in Economic History*, 52, 93 – 110.

Fan, J. (2015): "Internal Geography, Labor Mobility, and the Distributional Impacts of Trade," Working papers, Pennsylvania State University.

Firpo, S., N. M. Fortin, and T. Lemieux (2009): "Unconditional Quantile Regressions," *Econometrica*, 77(3), 953–973.

FORTIN, N. M. (2006): "Greed, Altruism, and the Gender Wage Gap," Working papers, University of British Columbia.

GOLLIN, D., D. LAGAKOS, AND M. E. WAUGH (2014): "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*, 129(2), 939–993.

GOLLIN, D., S. PARENTE, AND R. ROGERSON (2002): "The Role of Agriculture in Development," *American Economic Review*, 92(2), 160–164.

HARRIS, J. R., AND M. P. TODARO (1970): "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 60(1), 126–142.

HECKMAN, J. J. (2005): "China's Human Capital Investment," *China Economic Review*, 16, 50–70.

HERRENDORF, B., R. ROGERSON, AND A. VALENTINYI (2013a): "Growth and Structural Transformation," *Handbook of Economic Growth, forthcoming.*

——— (2013b): "Two Perspectives on Preferences and Structural Transformation," *American Economic Review, forthcoming.*

HERRENDORF, B., AND T. SCHOELLMAN (2015): "Wages, Human Capital, and the Allocation of Labor across Sectors," Working paper, Arizona State University.

HNATKOVSKA, V., AND A. LAHIRI (2015): "Rural and Urban Migrants in India: 1983-2008," *World Bank Economic Review*, 29, S257–S270.

HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.

JEFFERSON, G. H., AND I. SINGH (eds.) (1999): *Enterprise reform in China : ownership, transition, and performance.* Washington, D.C. : The World Bank.

KONGSAMUT, P., S. REBELO, AND D. XIE (2001): "Beyond Balanced Growth," *Review of Economic Studies*, 68(4), 869–882.

KUZNETS, S. (1955): "Economic Growth and Income Inequality," *American Economic Review*, 65, 1Ű28.

LAGAKOS, D., AND M. WAUGH (2012): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economics Review, forthcoming.*

LAITNER, J. P. (2000): "Structural Change and Economic Growth," *Review of Economic Studies*, 67(3), 545–561.

Morten, M. (2016): "Temporary Migration and Endogenous Risk Sharing in Village India," NBER Working Papers 22159, National Bureau of Economic Research, Inc.

Munshi, K., and M. Rosenzweig (2016): "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, 106(1), 46–98.

Ngai, L. R., and C. A. Pissarides (2007): "Structural Change in a Multisector Model of Growth," *American Economic Review*, 97(1), 429–443.

Qin, B., and Y. Zhang (2014): "Note on urbanization in China: Urban definitions and census data," *China Economic Review*, 30(C), 495–502.

Revi, Aromar, e. a. (2011): *Urban India 2011: Evidence.*

Roses, J. R., and B. Sanchez-Alonso (2004): "Regional wage convergence in Spain 1850-1930," *Explorations in Economic History*, 41(4), 404–425.

Smith, J. P., and F. R. Welch (1989): "Black Economic Progress after Myrdal," *Journal of Economic Literature*, 27(2), 519–64.

Tiwari, P., J. Rao, and J. Day (2016): *Development Paradigms for Urban housing in BRICS countries.* Palgrave Mcmillan, London.

Tombe, T., and X. Zhu (2015): "Trade, Migration and Productivity: A Quantitative Analysis of China," Working Papers tecipa-542, University of Toronto, Department of Economics.

Young, A. (2013): "Inequality, the Urban-Rural Gap and Migration*," *The Quarterly Journal of Economics.*

A Appendix: Not for publication

A.1 Data

A.1.1 India

The National Sample Survey Organization (NSSO), set up by the Government of India, conducts rounds of sample surveys to collect socioeconomic data. Each round is earmarked for particular subject coverage. We use the latest six large quinquennial rounds – 38(Jan-Dec 1983), 43(July 1987-June 1988), 50(July 1993-June 1994), 55(July 1999-June 2000), 61(July 2004-June 2005) and 66(July 2009-June 2010) on Employment and Unemployment (Schedule 10). Rounds 38 and 55 also contain migration particulars of individuals. We complement those rounds with a smaller 64th round(July 2007-June 2008) of the survey since migration information is not available in all other quinquennial survey rounds.

The survey covers the whole country except for a few remote and inaccessible pockets. The NSS follows multi-stage stratified sampling with villages or urban blocks as first stage units (FSU) and households as ultimate stage units. The field work in each round is conducted in several sub-rounds throughout the year so that seasonality is minimized. The sampling frame for the first stage unit is the list of villages (rural sector) or the NSS Urban Frame Survey blocks (urban sector) from the latest available census. The NSSO supplies household level multipliers with the unit record data for each round to help minimize estimation errors on the part of researchers. The coding of the data changes from round to round. We re-coded all changes to make variables uniform and consistent over the time.

In our data work, we only consider individuals that report their 3-digit occupation code and education attainment level. Occupation codes are drawn from the National Classification of Occupation (NCO) – 1968. We use the "usual" occupation code reported by an individual for the usual principal activity over the previous year (relative to the survey year). The dataset does not contain information on the years of schooling for the individuals. Instead it includes information on general education categories given as (i) not literate -01, literate without formal schooling: EGS/ NFEC/ AEC -02, TLC -03, others -04; (ii) literate: below primary -05, primary -06, middle -07, secondary -08, higher secondary -10, diploma/certificate course -11, graduate -12, postgraduate and above -13. We aggregate those into five broader categories: not-literate; literate but below primary; primary; middle; and secondary and above education.

The NSS only reports wages from activities undertaken by an individual over the previous week (relative to the survey week). Household members can undertake more than one activity in the reference week. For each activity we know the "weekly" occupation code, number of days spent working in that activity, and wage received from it. We identify the main activity for the individual as the one in which he spent maximum number of days in a week. If there are more than one activities with equal days worked, we consider the one with paid employment (wage is not zero or missing). Workers sometimes change the occupation due to seasonality or for other reasons. To minimize the effect of transitory occupations, we only consider wages for which the weekly occupation code coincides with usual occupation (one year reference). We calculate the daily wage by dividing total wage paid in that activity over the past week by days spent in that activity.

Lastly, we identify full time workers in our dataset. We assume that an individual is a full time worker if he is employed (based on daily status code) for at least two and half days combined in all activities during the reference week. We drop observations if total number of days worked in the reference week is more than seven.

A.1.2 China

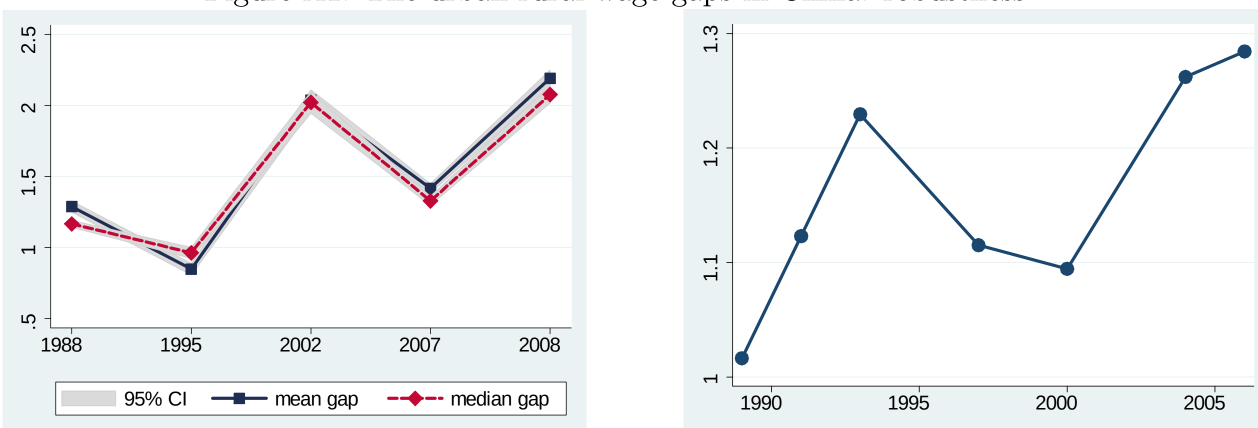
The Chinese Household Income Project (CHIP) is organized by Chinese and international researchers, with the assistance from National Bureau of Statistics (NBS), to study the distribution of personal income in both rural and urban areas in China. There are five waves available: 1988, 1995, 2002, 2007 and 2008. The last two waves were also part of the RUMiC (Rural-Urban Migrants in China) survey project. All waves contain separately a rural and urban survey, on which we base our definition of rural and urban.

Our sample of full-time workers includes observations with working status as employed or self-employed and total annual hour larger than 1900 hours. The status variable categories vary across years. We re-coded it to a consistent 8 categories: Employed, Self-employed, Unemployed, Retired, Homemaker, Disabled, Student/pre-school, and Others. For years with a separate self-employment indicator, we made sure it lines up with status. All years except for 1988 contain hour information. The 1995, 2007, and 2008 waves give hours per week. We first top code the observations larger than 100 to 100 hours, then multiply the hour per week statistics by 50 weeks, assuming two weeks of national holiday in China. In 2002, we divided the annual hour by 50, top code it, and then multiply back 50, to make the numbers consistent with other years. For the 1988 rural survey, we include all employed and self-employed people with non-missing wages. For the 1988 urban survey we dropped temporary workers (about 340 of them) from the sample of employed or self-employed.

A.2 Robustness: China

To further examine the robustness of the wage and income patterns for China, Figure A1 plots the urban-rural wage gaps computed from two other sources: CHIP dataset using family income information (panel (a)), and China Health and Nutrition Survey (panel (b)).[43] Panel (a) shows the mean and median gaps in (per capita) annual family income of urban and rural households in China since 1988, while panel (b) constructs the ratio of mean urban to rural wages over 1989-2006 period. Both plots reveal the same pattern of widening income and wage gaps between urban and rural Chinese workers over the past three decades.

Figure A1: The urban-rural wage gaps in China: robustness



(a) CHIP: mean income gap (b) China Health and Nutrition Survey: mean wage gap

Notes: Panel (a) shows the ratio of mean income in urban areas to mean income in rural areas; and the ratio of median income in urban areas to median income in rural areas together with the 95% confidence intervals using CHIP dataset; Panel (b) reports the ratio of mean wages in urban and rural areas in China Health and Nutrition Survey.

B Oaxaca-Blinder decompositions of wage changes

In this Appendix we present the results of the Oaxaca-Blinder decomposition discussed in Section 3. The results are summarized in Table A1. The top panel reports the change in the measured wage gap in the two countries, as well as how much of that differential is explained by characteristics ("explained"), in particular, by education; and how much is unexplained ("unexplained"). Note, however, that the inter-temporal changes in both the explained and unexplained components may be due to changes in either the attribute gaps or in the returns to those attributes. Thus, we also report the decompositions of these components into explained and unexplained parts. As usual, the explained component is attributable to changes in gaps in the observables, while the unexplained component is due to changes in the returns to these observables over time.[44]

[43] Like the CHIP dataset, the China Health and Nutrition Survey dataset contains individual- and household-level survey data. However, it is significantly smaller than CHIP.

[44] This inter-temporal decomposition of outcome differentials is in the spirit of Smith and Welch (1989) who used such decomposition techniques in their analysis of the change in the black-white wage differential over time.

More precisely, our econometric model for location c and year t is given by

$$y_{ct} = X_{ct}'\beta_{ct} + e_{ct}, \qquad c = 1, 2; \text{ and } t = 1, 2, \tag{A1}$$

where y_{ct} is a vector of log wages while X_{ct} is the matrix of regressors for location c in year t. Here β_{ct} is a coefficient vector, and e_{ct} is the vector of residuals. The differential in expected outcomes between the workers in urban and rural locations in year t is then given by:

$$\Delta y_t^e = \Delta X_t'\tilde{\beta}_t + X_{1t}'(\beta_{1t} - \tilde{\beta}_t) + X_{2t}'(\tilde{\beta}_t - \beta_{2t}),$$

where $\tilde{\beta}_t$ is the vector of coefficients from the model with both groups pooled. The first term above is the explained part while the last two terms give the unexplained parts of the decomposition. Denote E_t to be the explained component of the decomposition, and U_t to be the unexplained part, then

$$\begin{aligned} E_t &= \Delta X_t'\tilde{\beta}_t, \qquad t = 1, 2, \\ U_t &= X_{1t}'(\beta_{1t} - \tilde{\beta}_t) + X_{2t}'(\tilde{\beta}_t - \beta_{2t}), \qquad t = 1, 2. \end{aligned}$$

This is a standard static Oaxaca-Blinder decomposition of differences between groups. We are interested in performing a time-series decomposition of the urban-rural wage gap. For this purpose we employ a two-fold Oaxaca-Blinder procedure where we use coefficients from a pooled regression with a group membership indicator (as in Fortin, 2006) as the reference coefficients. We use the initial year as the base year for the inter-temporal decomposition, so 1983 is the benchmark sample in our analysis of India, and 1988 of China.

The inter-temporal change in the outcome differentials can be written as the sum of changes in the explained, E and unexplained, U components, using the expression above:

$$\Delta y_2^e - \Delta y_1^e = (E_2 - E_1) + (U_2 - U_1) = \Delta E + \Delta U$$

These differentials are reported in Panel (a) of Table A1 below. Note, however, that inter-temporal changes in the explained and unexplained components may be due to changes in either the attribute gaps or in the returns to those attributes. For instance, the inter-temporal decomposition of the explained part, $\Delta E = \Delta X_2'\tilde{\beta}_2 - \Delta X_1'\tilde{\beta}_1$, can be broken down as

$$\Delta E = \Delta X_2' \left(\tilde{\beta}_2 - \tilde{\beta}_1\right) + \left(\Delta X_2' - \Delta X_1'\right)\tilde{\beta}_1,$$

where the first term is the unexplained part of ΔE, while the second term is the explained part of ΔE. This decomposition is presented in Panel (b) of Table A1 below. Similarly, the inter-temporal decomposition of the unexplained part, ΔU can be broken down into an explained and unexplained parts. This decomposition is presented in Panel (c) of Table A1 below.

Our specification of equation (A1) for each location (l = rural, urban) and each year (t =1983 and 2010 for India, or 1988 and 2008 for China) is

$$\log(Wage_{i,l,t}) = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 sex + \beta_4 I_{region} + \beta_5 I_{edu} + \varepsilon,$$

where subscript i is for an individual, I_{region} is a set of regional dummies, consisting of North, East, Central, North-East, South and West in India and East, Northeast, Central and West in China. I_{edu} is a set of education dummies in India, and years of education in China. In India, NSSO data contains many education categories: not-literate; literate but below primary; primary education; middle education; secondary, higher secondary, diploma/certificate course, graduate and above, postgraduate and above. To homogenize these categories across rounds we group them into 5 broader categories: not-literate; literate but below primary; primary; middle; and secondary and above education. In China, the dataset contains education years for each individual, so we include that variable directly into our regressions.

Adding up the explained components of the changes in "explained" and "unexplained" gaps, we find that they have accounted for about 23% ((-0.018-0.037)/-0.236) of the change in the measured gap in India, and -13% ((0.021-0.027)/0.043) in China. The negative number for China suggests that the wage gaps in China should have shrunk as attributes of rural and urban workers have come closer together during 1988-2008 period. Of these, convergence in education was responsible for about a third in India, and almost for the entire change in China. Overall, these results suggest that individual characteristics have played a limited role in driving the time-series dynamics of wage gaps in China and India.

C Aggregate facts

The ongoing process of structural transformation of China and India can be seen through Figures A2 and A3. Figure A2 shows employment shares in agriculture and non-agriculture for China (panel (a)) and India (panel (b)). Figure A3 shows the distribution of output across the agriculture and non-agriculture in the two economies. As is easy to see, agriculture has been releasing workers in both countries, and its share of output has also been declining over time in both India and China. These are the textbook features of structural transformation.

The third aggregate fact of interest is the behavior of sectoral labor productivities in the two countries during this period. We compute sectoral labor productivity as the ratio of real GDP in each sector (agri and non-agri) and their respective employment. Figure A4 shows that in both China and India, labor productivity in both agriculture and non-agriculture was increasing during this period, with non-agricultural productivity expanding at a much faster

Table A1: Oaxaca-Blinder decompositions of wage gaps		
	China	India
	(i)	(ii)
(a) Change over	1988-2008	1983-2010
measured gap	0.043	-0.236
explained	0.129	-0.087
education	0.115	-0.048
unexplained	-0.087	-0.148
(b) Change in explained		
measured gap	0.129	-0.087
explained	0.021	-0.018
education	0.007	0.002
(c) Change in unexplained		
measured gap	-0.087	-0.148
explained	-0.027	-0.037
education	-0.014	-0.020

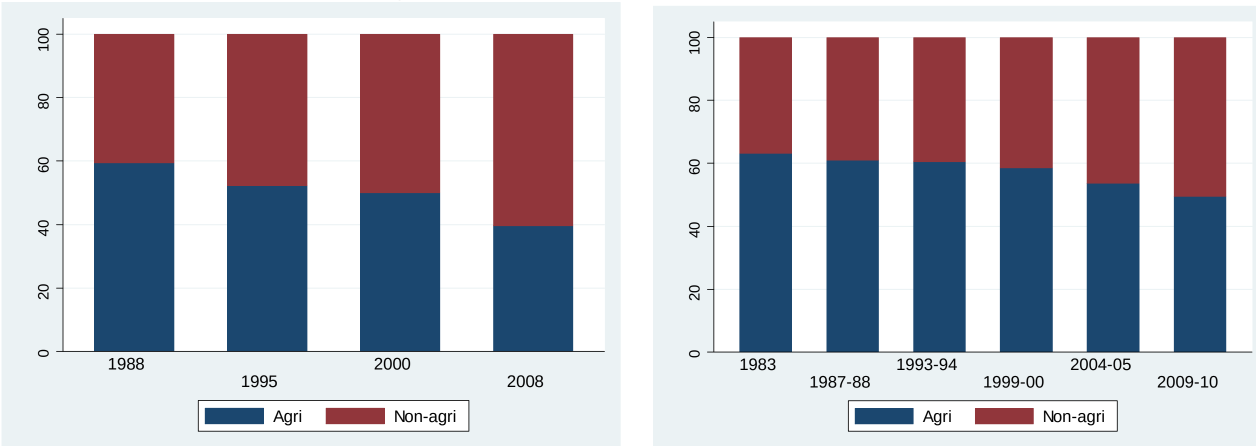
Note: Panels (a) of this table presents decomposition results for the time-series change in the log wage gap between urban and rural workers during 1983-2010 in India (column (i)) and 1988-2008 in China (column (ii)). "Explained" refers to the value of the change in wage gap that is explained by demographic characteristics and education, while "unexplained" refers to the residual change in wage gap. Panel (b) and (c) further decomposes the changes in the explained and unexplained components into the part that is due to changes in characteristics and the residual part. All panels also report the contribution of education to the explained gaps.

pace. While the patterns in the two economies were remarkably similar, a key difference was that labor productivity growth in China was must faster than in India. Thus, the labor productivity in agriculture increased by only 67 percent in India between 1983 and 2010. In contrast, agricultural labor productivity in China grew by 163 percent between 1990 and 2008. The non-agricultural labor productivity rose by 200 percent in India and 338 percent in China during the same periods.[45]

Next, Figure A5 presents the evolution of the relative price of non-agricultural goods (relative to agricultural good) in China and India since the 1980s. The movement in this relative price is very similar in the two countries. The relative price of non-agriculture declined by 23 percent in China and 29 percent in India. It is worth noting that the world relative price of agriculture was actually falling during most of the period since the 1980s, in contrast

[45] When reporting growth rates of labor productivity we used 1990 as the starting year for China instead of 1988 because of discountinuity in the sectoral employment data for China in 1989. We suspect that the definition of employed must have been changed in that year.

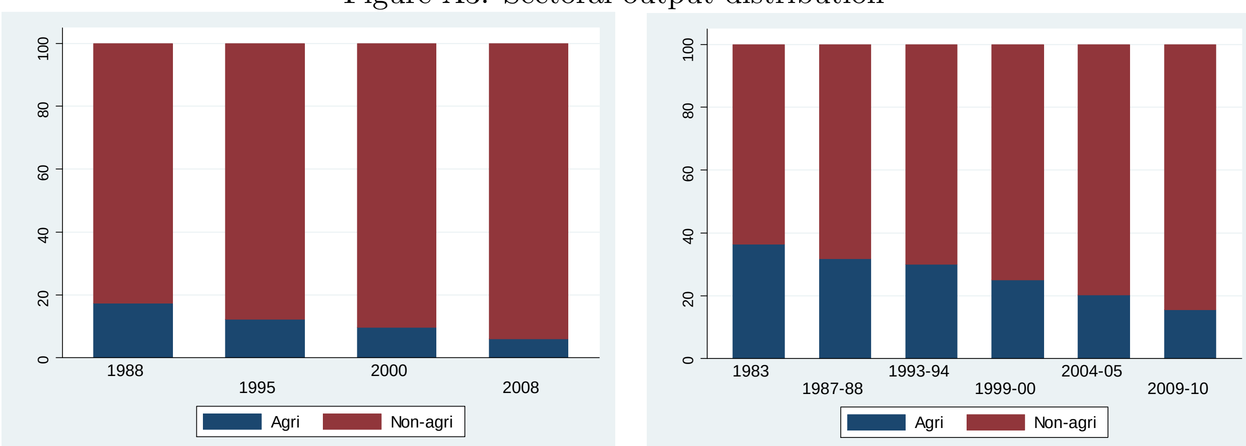
Figure A2: Employment distribution



(a) China: employment shares (b) India: employment shares

Notes: Panel (a) of this Figure presents the distribution of workforce across agricultural and non-agricultural sectors for China while panel (b) presents the employment distribution across the two sectors for India.

Figure A3: Sectoral output distribution



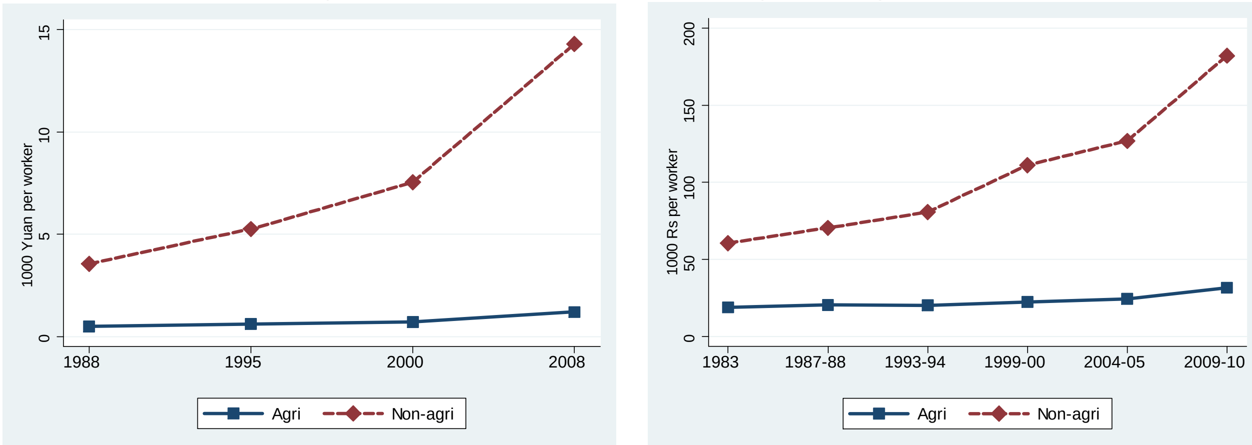
(a) China: output shares (b) India: output shares

Notes: Panel (a) of this Figure presents the distribution of output across agricultural and non-agricultural sectors in China. Panel (b) presents same distribution for India.

to the rising relative price of agriculture in China and India.

The final key aggregate fact relates to urbanization. Figure A6 shows the urban share of both population and employment in China (panel (a)) and India (panel (b)) during their sample periods of 1988-2008 and 1983-2010, respectively. Just as in patterns on relative prices and structural transformation, the urbanization patterns are qualitatively very similar in the two countries with the urban share of employment rising from 26 to 35 percent in China and 22 to 30 percent in India. To obtain this number in India we used the Census of India and NSS data. The Census of India is conducted every 10 years on the first year of each decade.

Figure A4: Sector-biased technological progress



(a) China (b) India

Notes: Panel (a) shows sectoral labor productivity during the 1990-2008 period for China, while panel (b) shows the same for India for the 1983-2010 period.

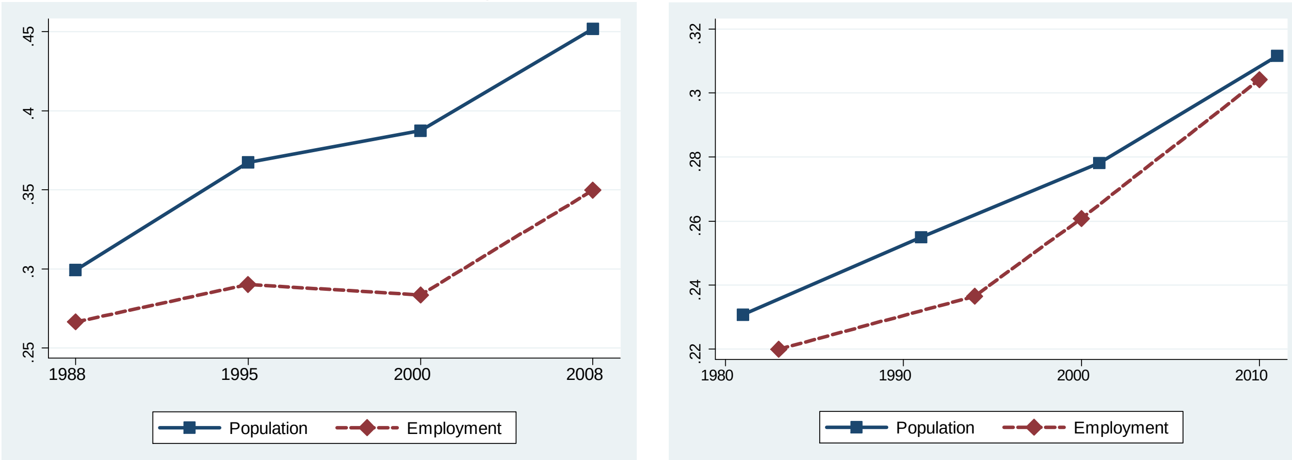
Figure A5: Relative price of non-agriculture

(a) China (b) India

Notes: This figure shows the price of non-agricultural output relative to agricultural output. The relative price in the initial sample year is normalized to 1.

Thus, in 1981 the total population of India was 683.3 million people, of which 525.6 million lived in rural areas and 157.7 million lived in urban areas. To obtain employment numbers we multiply these population figures by the share of working age population in 1983 from the NSS equal to 0.54 in rural areas and 0.59 in urban areas; by the labor force participation rate in 1983 from the NSS equal to 0.66 in rural areas and 0.59 in urban areas; and by the share of employed in the labor force equal to 0.94 in urban areas and 0.96 in rural areas. For China, National Statistical Yearbook reports urban and rural population and employment numbers.

Figure A6: Urban share



(a) China (b) India

Notes: This figures show the urban share of population and employment.

C.1 Consumption moments: Data and calculations

C.1.1 Consumption value added

For India we used sectoral value added from GDP by economic activity data from Statement 10 of National Accounts Statistics provided by the Ministry of Statistics and Programme Implementation (MOSPI) of Government of India. Investment is measured as gross capital formation, and was obtained from Statement 20 of National Accounts Statistics provided by MOSPI. Both value added and investment is in constant 1999-00 prices and can be accessed from
http://mospi.nic.in/Mospi_New/site/India_Statistics.aspx?status=1&menu_id=43.

For China the national level agriculture and non-agriculture employment and GDP was obtained from the National statistics yearbook 2013. GDP is in constant 2004 prices.

C.2 Aggregate and state/provincial data

The series for the relative prices of non-agricultural goods (relative to agricultural good) were obtained using nominal and real output series from the National Accounts Statistics provided by the Ministry of Statistics and Programme Implementation (MOSPI) of Government of India. For China we used the National Statistical Yearbook.

C.2.1 China

For the provincial aggregate data for China, our primary source of data is from the China Compendium of Statistics, which is published in 2009 to celebrate the 60th anniversary of PRC and contains statistics from 1949-2008. Whenever the information needed is missing in the Compendium, we complement it by check the provincial Statistics Yearbooks in various years so that in the end our data could expand the 20 years between 1988-2008.

Sectoral GDP: this information solely comes from the China Compendium of Statistics. The Compendium reports a nominal series of sectoral GDP and GDP index that equals 100 in 1952. There are three sectors: the Primary sector, which includes agriculture, fishing and husbandry; the secondary sector, which includes construction and manufacturing; and the tertiary sector, which is the service sector. To get our own real GDP, we multiply the 1952 level nominal GDP to the GDP index in each year. In this way, we are able to compare GDP across time and provinces. Finally, we define the real GDP in the primary sector as our GDP agriculture, and sum up the real GDP in the secondary and tertiary sectors as our GDP non-agriculture.

Urban (rural) employment: this information is obtained directly from the China Compendium of Statistics.

Urban (rural) population: this information mainly comes from the China Compendium of Statistics, but with missing years for 7 provinces out of the 31 (Hebei, Jilin, Zhejiang, Fujian, Guangdong, Sichuan, and Shaanxi). We supplement the missing data from the provincial yearbooks. In the end, we could get the urban-rural population for all the provinces except for Hebei, Jilin, Guangdong, and Chongqing, which only have agrarian and non-agrarian population (calculated from the hukou status). For these four provinces, we used the non-agrarian population as the number for urban population, and the agrarian population as the rural one.

Urban (rural) per capita income: this information is obtained from the provincial yearbooks in various years. The information is usually under the section for "People's Livelihood". The urban income is reported as per capita disposable income in the urban area, and the rural income is reported as per capita net income in the rural area.

Urban (rural) CPI: this information is obtained from the China Compendium of Statistics. For Beijing and Shanghai, there is only aggregate provincial CPI series. We assigned this series as both urban and rural CPI in these provinces. For Tianjin and Chongqing, there is only urban CPI. So we proxied the rural CPI by the urban CPI in these two provinces. The Shaanxi province doesn't have rural CPI for 1979-1994, we replace that by the aggregate CPI series. We convert all indices to the common base year of 1995. Note that our results on wage or income divergence in China are unlikely to be driven by the mismeasurement of CPI in four provinces above for which separate rural and urban CPI is not available. This is because the divergence of income and wages in China is observed in all provinces, as can be seen from Figure 3 in the main text.

C.2.2 India

For state-level data we rely on the following sources.

Sectoral GDP by state: This information comes from the online database of the India Government Ministry of Statistics and Programme Implementation. The data comes in four set of years: 1980-1996, 1993-2002, 1999-2007, and 2004-2012, in terms of both current and constant prices. We use the constant price series from each dataset, and rescale them so that the base year is the same as the last dataset (the 2004-2012 one), which is 2004. To do so, we first rescale the 1999-2007 data set for each province-sector series, so that the 2004 GDP value is consistent across the two data sets. Then, we rescale the 1993-2002 data set to match the 1999 GDP value in the later data sets, and so on. In this way, each province sector series is at constant 2004 price.

Sectoral employment information by state: from NSS.

D Proposition Proofs[46]

In the following it shall also be useful to follow the notation:

$$k^A = \frac{L^{UA}}{L^{RA}}, \quad k^N = \frac{L^{UN}}{L^{RN}}, \quad k = \frac{L^U}{L^R}, \quad s^A = \frac{L^{RA}}{L^R} \tag{A2}$$

One can use the definitions in (A2) to get

$$s^A = \frac{k - k^N}{k^A - k^N} \tag{A3}$$

Substituting the aggregate solution for C^A into the market clearing condition equation (3.20) and rearranging the result gives

$$p_t = \left(\frac{1-\theta}{\theta}\right)\left[\frac{Y_t^A - \bar{a}L_t - \tau_t^R L_t^{RN} - \tau_t^U L_t^{UN} - \tau_t \mu_t L_t^R}{Y_t^N + \bar{n}L_t}\right] \tag{A4}$$

In addition, the optimality condition for sectoral labor allocations given by equation (3.14) implies that $w_t^{RA} + \tau_t^R = w_t^{RN}$ which implies

$$p_t = \frac{\alpha^R A_t^R \left(L_t^{RA}\right)^{\alpha^R - 1} + \tau_t^R}{\beta^R N_t^R \left(L_t^{RN}\right)^{\beta^R - 1}} \tag{A5}$$

[46] Detailed derivation and expansions of the full model are available in an online appendix.

Combining the two gives the condition

$$\left(\frac{k_t-\gamma_t k_t^A}{k_t^A-k_t}\right)^{\beta-1}=\left(\frac{1-\theta}{\theta}\right)\left[\frac{\left(\frac{k_t-\gamma_t k_t^A}{(1-\gamma_t)k_t^A}\right)^{\beta}\left(1+k_t^A\right)-\frac{\bar{a}}{A_t^R}\left(1+k_t\right)^{\beta}L_t^{1-\beta}}{\left(\frac{k_t^A-k_t}{(1-\gamma_t)k_t^A}\right)^{\beta}\left(1+\gamma_t k_t^A\right)+\frac{\bar{n}}{N_t^R}\left(1+k_t\right)^{\beta}L_t^{1-\beta}}\right] \tag{A6}$$

D.1 Proof of Proposition 1

We start by analyzing how aggregate productivity growth affects equation (A6). It is easy to see that only the right hand side of that equation is affected as $\frac{\bar{a}}{A_t^R}\left(1+k_t\right)^{\beta}L_t^{1-\beta}$ and $\frac{\bar{n}}{N_t^R}\left(1+k_t\right)^{\beta}L_t^{1-\beta}$ both decline along such productivity paths. As a result, for every level of k_t the right hand side rises.

The allocation of labor in each location to the two sectors is given by $\frac{L_t^{RA}}{L_t^R}=s_t^A$ and $\frac{L_t^{UA}}{L_t^U}=\frac{k_t^A s_t^A}{k_t}$. Along paths with aggregate productivity growth, $k_t^A=\left(\frac{A_t^U}{A_t^R}\right)^{\frac{1}{1-\beta}}$ remains unchanged while k_t rises. Moreover, along paths with rising k_t, $s_t^A=\frac{k_t-\left(\frac{N_t^U}{N_t^R}\right)^{\frac{1}{1-\beta}}}{\left(\frac{A_t^U}{A_t^R}\right)^{\frac{1}{1-\beta}}-\left(\frac{N_t^U}{N_t^R}\right)^{\frac{1}{1-\beta}}}$ must decline since $\left(\frac{A_t^U}{A_t^R}\right)^{\frac{1}{1-\beta}}<\left(\frac{N_t^U}{N_t^R}\right)^{\frac{1}{1-\beta}}$. Hence, both $\frac{L_t^{RA}}{L_t^R}$ and $\frac{L_t^{UA}}{L_t^U}$ decline along paths with rising aggregate productivity. This implies that the employment share of agriculture in both locations falls while that of non-agriculture rises as the economy grows.

The overall share of agricultural employment in this economy is given by $\frac{L_t^A}{L_t}=\frac{L_t^{UA}}{L_t^U}\frac{L_t^U}{L_t}+\frac{L_t^{RA}}{L_t^R}\frac{L_t^R}{L_t}$ which can be rewritten as

$$\frac{L_t^A}{L_t}=\frac{s_t^A\left(1+k_t^A\right)}{1+k_t}$$

Since k^A is constant along paths with aggregate productivity growth while s^A falls and k rises, it is clear that $\frac{L_t^A}{L_t}$ must fall. Hence, the economy undergoes a structural transformation along paths with aggregate growth.

To complete a description of the economy along paths with aggregate productivity growth we need to describe the paths of goods and factor prices. Given that there are no costs of switching locations or sectors it is easy to see that factor prices must be equalized across sectors and locations in this special case, i.e.,

$$w_t^{RA}=w_t^{UA};\ w_t^{RA}=w_t^{RN};\ w_t^{UA}=w_t^{UN}\quad \text{for all } t$$

The relative price of the non-agricultural good is given by $p_t=\frac{A_t^R}{N_t^R}\left(\frac{s_t^A}{1-s_t^A}\right)^{\beta-1}$. As we noted above, $\frac{A_t^R}{N_t^R}$ remains unchanged along paths with aggregate productivity increases while s_t^A

declines. Hence, along paths with rising aggregate productivity the relative price of non-agriculture, p_t, must rise, i.e., the agricultural terms of trade worsens.

In the limit as $t \to \infty$, the non-homothetic components in the numerator and denominator of the right hand side of equation (A6) vanish and the economy settles into balanced growth with a limiting stationary degree of urbanization given by

$$\hat{k} = \frac{\left[\hat{\gamma}\left(1+\hat{k}^A\right) - \theta\left(\hat{\gamma}-1\right)\right]\hat{k}^A}{1+\left\{1+\theta\left(\hat{\gamma}-1\right)\right\}\hat{k}^A}$$

where $\hat{\gamma} = \left(\frac{A_0^R/A_0^U}{N_0^R/N_0^U}\right)^{\frac{1}{1-\beta}}$ and $\hat{k}^A = \left(\frac{A_0^U}{A_0^R}\right)^{\frac{1}{1-\beta}}$. It is easy to check that $\hat{k} > 0$. Furthermore, the associated limiting s_t^A is given by

$$\hat{s}^A = \frac{\hat{k} - \left(\frac{N_0^U}{N_0^R}\right)^{\frac{1}{1-\beta}}}{\left(\frac{A_0^U}{A_0^R}\right)^{\frac{1}{1-\beta}} - \left(\frac{N_0^U}{N_0^R}\right)^{\frac{1}{1-\beta}}}$$

D.2 Proof of Proposition 2

Under non–agriculture biased productivity change, we have $\frac{N_t^U}{N_t^R} = \frac{N_0^U}{N_0^R}$ and $\frac{A_t^U}{A_t^R} = \frac{A_0^U}{A_0^R}$. It directly follows that both $k_t^A = \left(\frac{A_t^U}{A_t^R}\right)^{\frac{1}{1-\beta}}$ and $\gamma_t = \left(\frac{A_t^R/A_t^U}{N_t^R/N_t^U}\right)^{\frac{1}{1-\beta}}$ remain unchanged. It is straightforward to verify from equation (A6) that the degree of urbanization in the new steady state must be greater than the level of urbanization in the initial steady state, $k_1 > k_0$. This, in turn, implies that s^A, $\frac{L^{RA}}{L^R}$, $\frac{L^{UA}}{L^U}$ and $\frac{L^A}{L}$ must all decline permanently in response to the shock.

The response of p, the relative price of the non-agricultural good, is however ambiguous since there are two offsetting effects. On the one hand, the productivity process implies that $\frac{A^R}{N^R}$ falls. On the other hand, s^A declines as well. Consequently, the behavior of $p_t = \frac{A_t^R}{N_t^R}\left(\frac{s_t^A}{1-s_t^A}\right)^{\beta-1}$ is ambiguous and depends on the relative strengths of these two opposing effects. As before, there are no sectoral or locational wage differences in this special case since there are no costs of switching.

D.3 Proof of Proposition 3

Equating equations (A4) and (A5), setting $\tau^R = \tau^U = 0$ and $\alpha = \beta$, and rearranging the resulting expression gives

$$\left(\frac{L_t^{RA}}{L^R - L_t^{RA}}\right)^{\beta-1} = \left(\frac{1-\theta}{\theta}\right)\left[\frac{\left(L_t^{RA}\right)^\beta + \frac{A_t^U}{A_t^R}\left(L^U - L_t^{UN}\right)^\beta - \frac{\bar{a}L}{A_t^R}}{\left(L^R - L_t^{RA}\right)^\beta + \frac{N_t^U}{N_t^R}\left(L_t^{UN}\right)^\beta + \frac{\bar{n}L}{N_t^R}}\right] \tag{A7}$$

The structural transformation that is induced by aggregate productivity increases in the case with no migration is easy to see directly from equation (A7). An increase in A_t^R and N_t^R unambiguously increase the right hand side of (A7). Since the left hand side is declining in L_t^{RA} while the right hand side of (A7) is rising in L_t^{RA}, the equation can hold with equality if and only if L_t^{RA} falls. Since L_t^{UN} is declining in L_t^{RA}, L_t^{UA} must fall as well. Hence, agricultural employment declines in both locations. The optimal sectoral labor allocation in rural areas gives $p_t = \frac{A_t^R}{N_t^R}\left(\frac{L_t^{RA}}{L_t^{RN}}\right)^{\beta-1} = \frac{A_0^R}{N_0^R}\left(\frac{L_t^{RA}}{L_t^{RN}}\right)^{\beta-1}$. Since L_t^{RA} falls, p_t must rise, i.e., the relative price of the agricultural good *falls*.

Since there are no inter-sectoral wage gaps within each location under Assumption 2, $\tau^R = \tau^U = 0$, the urban-rural wage gap is given by $\frac{w^{UA}}{w^{RA}}$. From the firm optimality conditions we have $\frac{w_t^{UA}}{w_t^{RA}} = \frac{A_t^U}{A_t^R}\left(\frac{L_t^{UA}}{L_t^{RA}}\right)^{\beta-1} = \frac{A_0^U}{A_0^R}\left(\frac{L_t^{UA}}{L_t^{RA}}\right)^{\beta-1}$. Using the definition $k^A \equiv \frac{L^{UA}}{L^{RA}}$, we can rewrite the wage gap as

$$\frac{w_t^{UA}}{w_t^{RA}} = \frac{A_0^U}{A_0^R}\left(k_t^A\right)^{\beta-1}$$

The effect of the productivity increase on the urban-rural wage gap depends on the response of k^A. If k^A declines then the wage gap widens. From equation (A3) above $s_t^A = \frac{k-k_t^N}{k_t^A-k_t^N} = \frac{\gamma-\frac{k}{k_t^A}}{\gamma-1}$ where $\gamma \equiv \left(\frac{A_0^R/A_0^U}{N_0^R/N_0^U}\right)^{\frac{1}{1-\beta}}$ Clearly, s_t^A is rising in k_t^A. We have seen above that s_t^A declines as productivity rises. Hence, k_t^A must decline as productivity parameters A^j and N^j rise, $j = R, U$. Hence, the urban-rural wage gap widens with rising productivity.

In the limit as $t \to \infty$, the non-homothetic components in the numerator and denominator of the right hand side of equation (A7) vanish and the economy settles into balanced growth with a stationary sectoral labor allocation in each location.