Education and Inequality in India: A Microeconometric Decomposition Analysis

Janneke Pieters [*]

Draft Version, February 2009[†]

Abstract

What effect has educational expansion had on income inequality in India? This study uses household survey data to measure how education affected the distribution of household consumption expenditures in the periods 1987-1993 and 1993-2004. The first results indicate that changes in educational attainment increased inequality in the first period, but decreased inequality in the latter. The returns to education had the opposite effect, namely to reduce inequality in the first period but raise it in the latter. The total direct effect of education was therefore very small in both periods. Changes in fertility, however, contributed to a substantial rise in inequality in both periods.

1 Introduction

Income inequality is increasing in many developing countries. Rising inequality harms the development process as it slows down poverty reduction at given growth rates, and may even reduce growth itself (see e.g.Björnskov (2008)). The average educational attainment is also increasing in developing countries: expansion of education is a key ingredient for development. Unfortunately, there is no clear theoretical prediction of the effect of educational expansion on income distribution (Ram, 1989). Higher average education in developing countries is often accompanied by increased inequality of education, which complicates the relationship. Furthermore, education has indirect distributive effects through, for example, female labor force participation and fertility changes. It should be no surprise that education cannot explain much of cross-country differences in income inequality (as found by De Gregorio and Lee (2002)).

In this paper I investigate the role of education in the rise in inequality in India between 1987 and 2004. Using household survey data, both inequality and educational expansion can be characterized in detail. The question is whether education played an important role in India's rising inequality since the 1990s.

---

[*]University of Groningen, email: janneke.pieters at rug.nl

[†]Work in progress: This draft is based on the first results of the project, and is not complete

In the analysis, changes in educational endowments are separated from changes in the returns to education. Furthermore, the indirect effects of education via fertility and employment status are estimated separately.

Many studies have investigated the distributive effects of education in the distribution of individual earnings. It has been widely documented that the returns to education are rising, as skill-biased technical change and international trade are increasing the demand for skilled workers (Goldberg and Pavcnik, 2007). The rise in returns is generally highest at the higher levels of education. Due to the increasingly convex relationship between education and earnings, educational expansion can increase earnings inequality even if the educational distribution is unchanged.

For India, too, the earnings distribution has been extensively studied, providing evidence of this convexification. Kijima (2006) finds that between 1983 and 1993 the changing distribution of educational attainment was the dominant force increasing wage inequality among male urban workers. Between 1993 and 1999, on the other hand, the rising returns to higher education increased wage inequality most. These findings reflect the reform in Indian economy in the early 1990s. Modern India is characterized by a fast-growing service and soaring demand for skilled workers. The government of India realizes that investment in higher education is necessary not to constrain growth (Government of India, 2006). According to Mazumdar and Sarkar (2008), government policies have been biased towards tertiary education, even though a very large share of the population has less than secondary education.

In the present study, I analyse inequality change in India at the household level in the periods 1987-1993 and 1993-2004. In the household, income from different members and different sources in pooled. The impact of educational expansion will thus consist only partly of changes in the earnings distribution. Moreover, the indirect effects of education through fertility are potentially important (see Ferreira and Leite (2004)).

The results of this study indicate that the returns to education we equalizing in 1987-1993 but disequalizing in 1993-2004. Changes in educational attainment had the opposite effect in both periods, leaving the total direct effect of education very small. Changes in fertility were highly disequalizing in both periods.

2 Inequality in India

The quinquennial surveys of National Sample Survey Organization (NSSO) in India provide the main source of information on inequality. In this study I use the Consumer Expenditure Survey for the years 1987-88, 1993-94, and 2004-05 [1]. Table 1 shows four measures of inequality in Monthly Per Capita Expenditure (henceforth MPCE) in constant prices, and spatially deflated[2]. These figures

---

[1]These years are referred to as 1987, 1993, and 2004, resp.

[2]Deflation is based on the price indexes by Deaton (2003) for 1987-1999 and the official Consumer Price Index for Agricultural Labor and Consumer Price Index for Industrial Workers for 1999-2004. Deaton's indexes are more carefully calculated then the official national price

pertain to the 15 major states plus urban Delhi, and use the household as the unit of analysis.

Table 1: Inequality of Monthly Per Capita Expenditure

Year	Gini	GE(0)	GE(1)	GE(2)
1987	0.338	0.188	0.230	0.540
1993	0.329	0.178	0.224	0.612
2004	0.346	0.197	0.255	0.706

Inequality declined in the period 1987-1993 and increased in the post-reform period 1993-2004. The changes in educational attainment are summarized in table 2, for all adults (age 20 and higher). The share of illiterate adults and adults with education below primary level has declined, especially between 1993 and 2004. The share of adults with middle school, secondary and higher education increased in both periods.

Table 2: Educational Attainment, percentage distribution

Educational Level	1987	1993	2004
1 Illiterate	50.50	44.75	35.87
2 Below primary	10.91	10.91	7.35
3 Primary	12.46	11.00	13.23
4 Middle	9.74	11.61	16.55
5 Secondary	11.21	14.81	18.04
6 Graduate and above	5.17	6.92	8.96
Total	100	100	100

A study by the Asian Development Bank (Asian Development Bank, 2007) shows that the combined effect of changes in the educational attainment of the household head and its returns accounted for more than 60 percent of the increase in the Gini coefficient between 1993 and 2004. In the present study, education of all adult household members will be taken into account, as is explained in the next section.

3 Microeconometric Decomposition of Inequality Change

Bourguignon et al. (2005) and Bourguignon et al. (2008) present a microeconometric decomposition of changes in the household income distribution that is designed to analyse inequality at the household level. The method is a generalization of the decomposition presented by Oaxaca (1973) and Blinder (1973), who separate inequality of mean income between groups into differences in characteristics, differences in returns, and differences in the residual or unobserved

indexes, but are only available for each state's rural and urban areas until 1999.

factors. Besides applying such a decomposition to compare groups, it is possible to compare different points in time. Using micro-economic data, the decomposition can be applied to the entire distribution of income, rather than their means. The difference with DiNardo et al. (1996) and Juhn et al. (1993) is that the level of analysis is the household rather than the individual. The model can thus include labour force participation decisions by all household members and changes in household composition.

To decompose changes in inequality into endowment effects (the effect of changes in the distribution of household characteristics) and price effects (the effect of changes in the returns to these characteristics) distributional counterfactuals are constructed. Let $f^\tau(y)$ be the distribution of MPCE, where $\tau(= t, t')$ denotes the year in which the distribution is observed. X is a vector of household or individual characteristics, including age, education, household size, principal industry of employment, and location (state), and $\chi^\tau(X)$ is the joint distribution of all elements of X in year τ. Denoting $g^\tau(y|X)$, the distribution of income conditional on X, the marginal distribution of MPCE in year τ can be expressed as:

$$f^\tau(y) = \int ... \int g^\tau(y|X)\chi^\tau(X)dX \tag{1}$$

The change in the distribution of MPCE between year t and year t' is thus a function of the change in $g(y|X)$ (the price effect) and the change in $\chi(X)$ (the endowment effect). Regarding the endowment effect, I am principally interested in the educational attainment of household members, the main type of employment of the household, and the number of children.

3.1 Empirical Strategy

The methodology by Bourguignon et al. (2008) allows for the explicit modelling of each household member's labor supply decisions. For India, unfortunately, this is impossible as their is no data which includes both individual employment details and a reliable measure of total household income. The present study is based on the NSSO Consumer Expenditure survey, which includes employment details (status, occupation, industry, etc.) only at the household level.

To measure the price and endowment effects on the distribution of MPCE, four models are estimated. I first estimate MPCE (in constant prices, and spatially deflated):

$$log(MPCE_h) = X_h\beta + \epsilon_i \tag{2}$$

where the vector X_h includes age of the household head and the spouse; gender of the head; education of the head, spouse, and other members; the number of household members by age group (0-15, 16-19, 20-65, and 65+); social group; religion; household type (self-employed, regular labor, or casual labor); principal occupation (professional, administrative, or other); principal industry of employment; state; and survey subround. The equation is estimated separately for rural and urban households, by OLS.

4

Second, I model the choice of educational level, household type, and number of children. There are six different levels of educational attainment,

1. Illiterate

2. Below primary school

3. Primary school

4. Middle school

5. Secondary and higher secondary education

6. Graduate and above

The educational level is estimated for each household member individually, separately for the head, the spouse, and other members, and for rural and urban areas:

$$X_i^E = k \text{ if } Y_i\gamma^k + \eta_i^{Ek} > Max(0, Y_i\gamma^j + \eta_i^{Ej}) \text{ for } j \neq k \qquad (3)$$

where the vector Y_i includes gender, age, social group, religion, and state. I use a multinomial logit model to estimate the probability of choice of each educational level. Similarly, a multinomial logit model is used to estimate the number of children at the household level, separately for rural and urban households:

$$X_h^F = m \text{ if } Y_h\theta^m + \eta_h^{Fm} > Max(0, Y_h\theta^j + \eta_h^{Fj}) \text{ for } j \neq m \qquad (4)$$

where the vector Y_h includes gender of the head, number of adult females, average age of adult females, average education of adult females, social group, religion, and state. The number of children is the number of household members in the age group 0-15, with a maximum of six (i.e., any number of children higher than six is recoded into six). Since the data do not allow us to identify which children belong to which mother, the explanatory variables are averages over all adult females in the household.

Finally, household type is also estimated by multinomial logit, for rural and urban households separately. Household type refers to the principal type of employment, which is only available for the household as a whole. Rural households are classified as self-employed or labor, while urban households are classified as self-employed, regular labor, or casual labor.

$$X_h^T = s \text{ if } Y_h\psi^s + \eta_h^{Ts} > Max(0, Y_h\psi^j + \eta_i^{Tj}) \text{ for } j \neq s \qquad (5)$$

where the vector Y_h includes gender of the head; education of the household head, spouse, and other adult members; number of children; social group; religion; and state.

I estimate educational level, household type, and the number of children to be able to update these characteristics from year t to year t', conditional on the distribution of their respective explanatory variables. That is, I estimate equations 3, 4, and 5 in order to construct the relevant counterfactual distribution of MPCE, but without an elaborate interpretation of the estimated coefficients in γ, θ, and ψ.

3.2 Simulating the Effect of Educational Expansion

To find the endowment effect of education on the distribution of MPCE, first the educational attainment of individuals in year t is 'updated': the probability of each educational level this person would have in year t' is simulated using the multinomial logit coeficients in γ (equation 3) from year t'. The characteristics in Y_i and the unobservable characteristics in η^E are held constant. The simulated educational attainment levels then replace the original values in the vector X_h in equation 2, which changes the distribution of MPCE.

If we want to include the indirect effect of education via fertility changes, we als update the number of children for each household. The number of children is simulated with the multinomial logit coefficients in θ from year t', plus the updated educational attainment in the vector Y_h, holding constant the other explanatory variables in Y_h and the effect of unobservable characteristics in η^F (equation 4). Now, both simulated educational attainment levels and simulated number of children replace the original values in the vector X_h in equation 2, which results in a new distribution of MPCE. Finally, household type can be updated similarly to include the indirect effect of education through household type.

The price effect of education is obtained simply by substituting the coefficients of all education variables in β for year t by those for year t' (equation2). Besides the price effect of education, we also simulate MPCE using all price effects. That is, we subtitute the complete vector β of year t by its counterfactual in year t' in 2. Finally, I also look at the effect of the distribution of unobservable characteristics. This effect is measured by multiplying the residual in year t by the ratio of residual standard deviation of year t' to year t ($\sigma_\epsilon^{t'}/\sigma_\epsilon^t$).

In all simulation excercises, a new value of MPCE is obtained for each household, such that from the simulated distribution any measure of inequality can be calculated. I use the start of the period as base year in all cases: 1987 for the period 1987-93, and 1993 for the period 1993-2004.

4 Results

4.1 Returns to Education

The distribution of educational attainment among adults was discussed in Section 2. When estimating MPCE, educational level dummies were included separately for the household head, spouse, and other members. As figures 1 to 6 show, the returns structure is generally convex and mostly so in 2004. The returns changed relatively little between 1987 and 1993, except for the spouses. The returns to education of the household head declined at all levels between 1993 and 2004. In urban India, the returns to education of the spouse and other members increased between 1993 and 2004, which indicates that the education of the head became relatively less important for the urban households' income. In all years, the correlation between education of the head and spouse is about .60 (slightly higher in urban than in rural areas), while the correlation between

education of the head and other members is only about .16. Including only education of the household head, as in (Asian Development Bank, 2007), may obviously generate the wrong picture of the distributive effects of educational expansion.
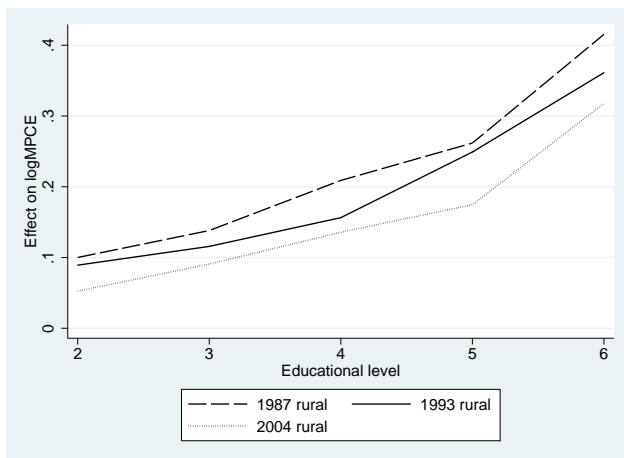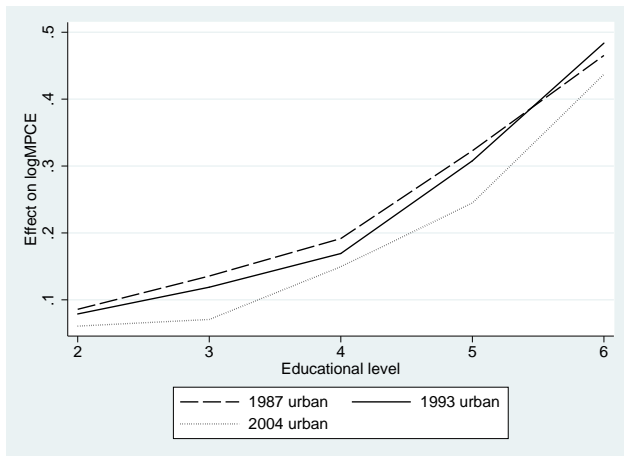

Figure 1: Returns to education head, rural

Figure 2: Returns to education head, urban

For the two periods 1987-1993 and 1993-2004 the endowment effect of education (including the indirect effects through fertility and household type) and the price effect of education on the distribution of MPCE are simulated. The results for the period 1987-1993 are summarized in table 3. Again, inequality is calculated with the household as the unit as analysis, rather than each individual.
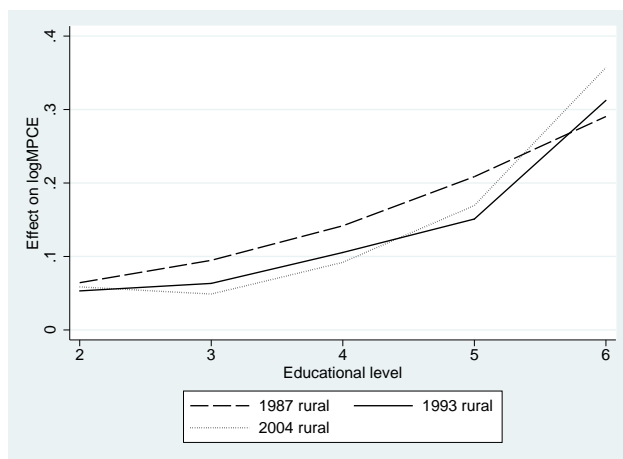
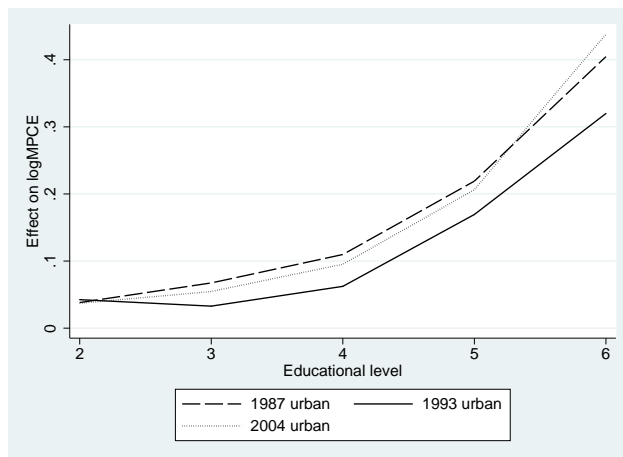Figure 3: Returns to education spouse, rural
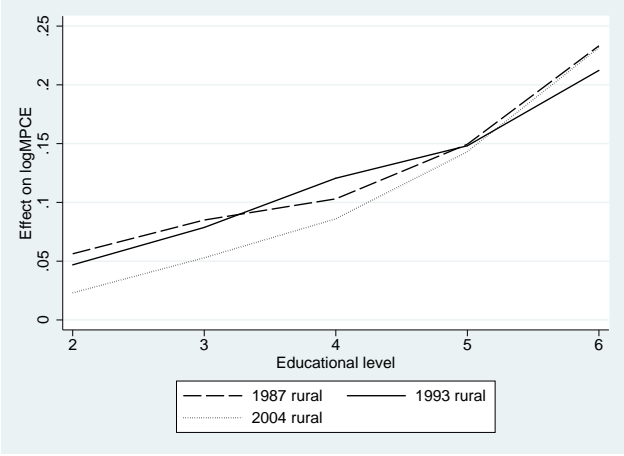

Figure 4: Returns to education spouse, urban
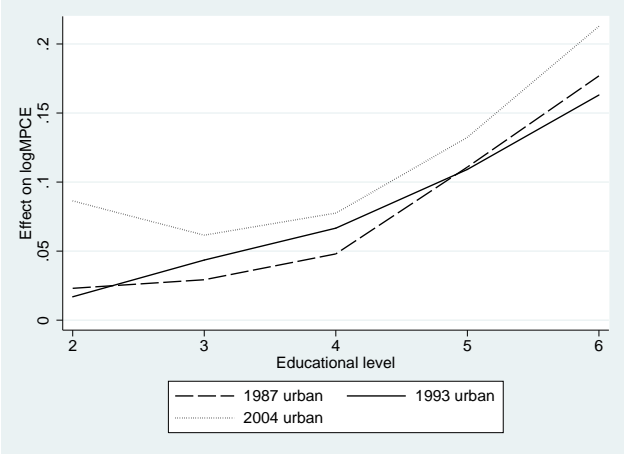
8

Figure 5: Returns to education other, rural

Figure 6: Returns to education other, urban

9

Table 3: Simulated Distribution of MPCE 1987-93: Summary of Results

	Gini	GE(0)	GE(1)	GE(2)
Observed 1987	0.338	0.188	0.230	0.540
Observed 1993	0.329	0.178	0.224	0.612
Unobservable	0.330	0.179	0.216	0.468
Price effect all	0.332	0.181	0.222	0.530
Price effect education	0.333	0.182	0.222	0.517
Total effect education	0.337	0.187	0.227	0.521
Endowment effect education	0.344	0.194	0.236	0.546
+ Indirect, fertility	0.368	0.222	0.270	0.609
+ Indirect, household type	0.368	0.222	0.270	0.615

The first two rows summarize inequality in the beginning and end of the period. All measures of inequality show a decline, except for the GE(2), which is more sensitive to changes at the high end of the distribution. The row 'unobservable' shows the effect of changes in the distribution of unobservables, which reduced inequality by all measures. In this period, the price effect of education was a slight reduction in inequality, while the endowment effect of education was to increase inequality: both effects were similar in magnitude but operated in opposite direction. Cobined, the total direct effect of education was a very slight reduction in inequality. However, when the indirect effect of educational endowment via fertility is included (in the row '+ Indirect, fertility'), we see a large increase in inequality. This suggests that reductions in fertility were greatest at the high end of the educational and MPCE distribution.

Table 4 shows the results of the same simulations, but now for the period 1993-2004. Observed inequality increased more in this period: the Gini coefficient rose from 0.329 to 0.346. Unlike the first period, the price effect of education in the second period was an increase in inequality (in line with the results of Kijima (2006)), and the direct endowment effect of education was a reduction inequality. The combined effect was a small increase in inequality. Again it is clear that the indirect effect of educational endowment via fertility has the largest adverse impact on the MPCE distribution. The subsequent effect via household type (last row) has no further distributive effects.

The clearest way to show the price and endowment effects of education on the distribution of MPCE is by use of growth incidence curves. These depict the percentage growth in MPCE by percentile of the MPCE distribution, and thus include information on the entire distribution. Figure 7 shows the observed change in MPCE distribution for both periods

In Figure 8 we see that the price effects in the first period were equalizing along the entire distribution. Adding the effect of unobservables reduces inequality further. The endowment effects in the first period, in figure 9, are very different from the price effects: growth in MPCE increases along the MPCE distribution. It is also clear that the indirect effect via fertility creates a large income gain compared to the direct endowment effect of education, especially

Table 4: Simulated Distribution of MPCE 1993-2004: Summary of Results

	Gini	GE(0)	GE(1)	GE(2)
Observed 1993	0.329	0.178	0.224	0.612
Observed 2004	0.346	0.197	0.255	0.706
Unobservable	0.326	0.175	0.219	0.596
Price effect all	0.331	0.181	0.227	0.623
Price effect education	0.335	0.185	0.232	0.642
Total direct effect education	0.333	0.183	0.229	0.626
Endowment effect education	0.327	0.176	0.219	0.599
+ Indirect, fertility	0.359	0.214	0.256	0.594
+ Indirect, household type	0.359	0.214	0.256	0.593
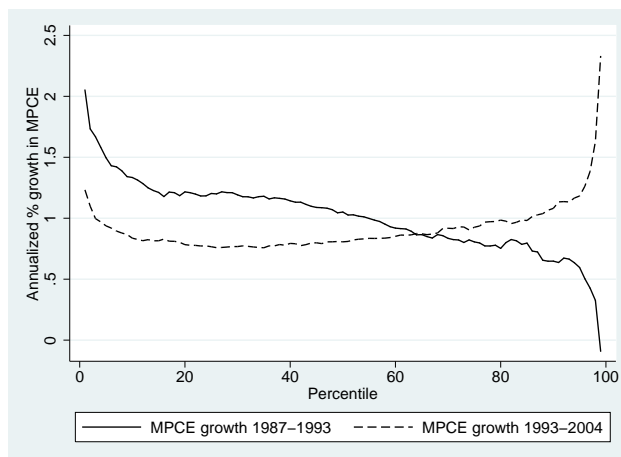
Figure 7: MPCE growth India, 19987-93 and 1993-2004

at the top of the distribution.

The price effect of education in the second period (figure 10) is equalizing in the lower half of the distribution, but very disequalizing in upper half. The endowment effect of education (figure 11) is small, and again, fertility changes led to a substantial increase in inequality: the per capita consumption gains from fertility changes are negative at the bottom of the distribution, and increasing to up to two percent annualy at the top.

5   Conclusions

The first results of this study indicate that in the period 1987-1993 education reduced inequality slightly: changing returns to education reduced inequality, while the endowment effect was exactly opposite. In the period 1993-2004, changing returns to education increased inequality, and this was partly offset by
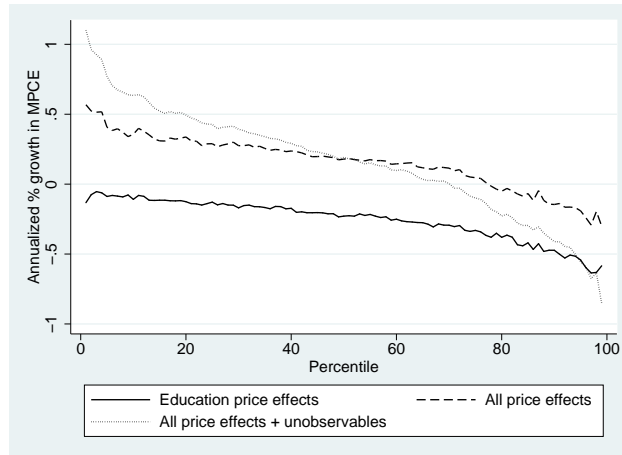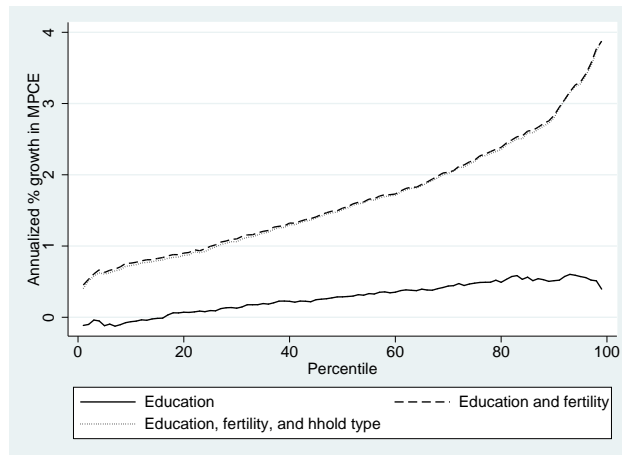
Figure 8: Price Effects, 1987-93

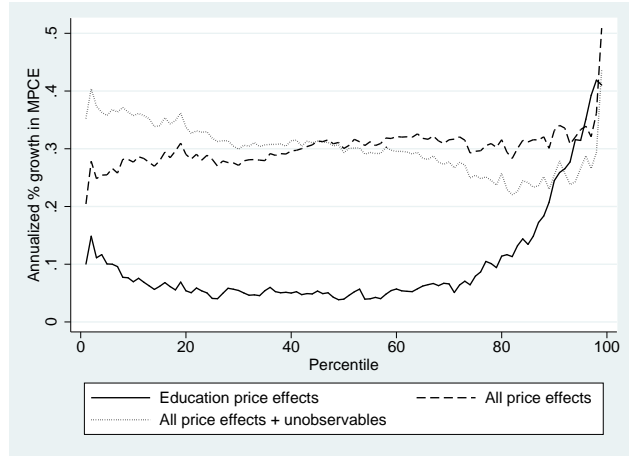Figure 9: Endowment Effect Education, 1987-93
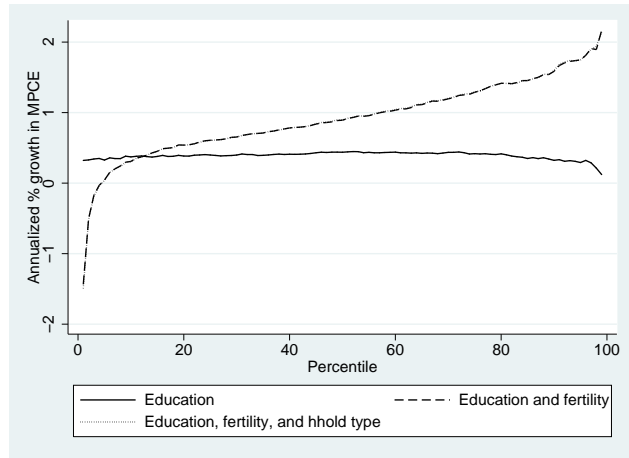
Figure 10: Price Effects, 1993-2004

Figure 11: Endowment Effect Education, 1993-2004

13

the equalizing effect of endowment changes. In both periods, fertility changes had a much greater impact, increasing inequality substantially. These results have to be studied further before drawing any conclusions, so that will be left for the next, complete, version of this paper.

References

Asian Development Bank (2007). *Inequality in Asia*, Chapter 5, pp. 1–97. Asian Development Bank.

Bjӧrnskov, C. (2008, October). The growth-inequality association: Government ideology matters. *Journal of Development Economics 87*(2), 300–308.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources 8*(4), 436–455.

Bourguignon, F., F. H. Ferreira, and P. G. Leite (2008). Beyond oaxaca-blinder: Accounting for differences in household income distributions. *Journal of Economic Inequality 6*, 117–148.

Bourguignon, F., F. H. G. Ferreira, and N. e. Lustig (2005). *The Microeconomics of Income Distribution Dynamics*. New York: World Bank and Oxford University Press.

De Gregorio, J. and J.-W. Lee (2002, September). Education and income inequality: New evidence from cross-country data. *Review of Income and Wealth 48*(3), 395–416.

Deaton, A. (2003, January). Prices and poverty in india, 1987-2000. *Economics and Political Weekly*, 362–368.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996, September). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica 64*(5), 1001–44.

Ferreira, F. H. G. and P. G. Leite (2004). Educational expansion and income distribution: A mircosimulation for ceará. In A. Shorrocks and R. Van der Hoeven (Eds.), *Growth, Inequality, and Poverty: Prospects for Pro-Poor Economic Development*, Chapter 11. New York: Oxford University Press.

Goldberg, P. K. and N. Pavcnik (2007, March). Distributional effects of globalization in developing countries. *Journal of Economic Literature 45*(1), 39–82.

Government of India (2006). Towards faster and more inclusive growth, an approach to the 11th five year plan. Technical report, Planning Commission, New Delhi.

Juhn, C., K. M. Murphy, and B. Pierce (1993). Wage inequality and the rise in return to skill. *Journal of Political Economy 101*, 410–442.

Kijima, Y. (2006, October). Why did wage inequality increase? evidence from urban india 1983-99. *Journal of Development Economics 81*(1), 97–117.

Mazumdar, D. and S. Sarkar (2008). *Globalization, Labor Markets and Inequality in India*. New York: Routeledge.

Oaxaca, R. (1973, October). Male-female wage differentials in urban labor markets. *International Economic Review 14*(3), 693–709.

Ram, R. (1989, April). Can educational expansion reduce income inequality in less-developed countries? *Economics of Education Review 8*(2), 185–195.