The Microeconomics of Changing Income Distribution in Malaysia *

Gary S. Fields and Sergei Soares

Revised Version: February, 2002

I. The Facts and the Questions Asked

The Malaysian income distribution has exhibited major changes over the twenty-four years from 1973 to 1997 for which data are available. Real average per capita income increased by 2.5 times, the absolute poverty rate fell from over fifty percent of the population to under 8 percent, income inequality decreased, and ethnic disparities narrowed (World Bank, n.d.). This record has caused Malaysia to be cited as a successful case of growth with redistribution (Ahuja et al, 1997).

Within this overall period, however, both the growth and the distribution experience were uneven. Economic growth was much slower in the 1984-89 period (just a 1.6% average annual increase in real GDP per capita). Also, income inequality followed two distinct phases. The first phase, from 1973 to 1989, was marked by falling income inequality. This was reversed, however, from 1989 to 1997, during which time income inequality rose. But because the changes in inequality in both periods were modest relative to the magnitude of economic growth, poverty in Malaysia fell continuously, albeit at a slower rate during the slow growth years of the 1980s.

This study uses data from Malaysia's Household Income and Expenditure Surveys to quantify the importance of different factors in accounting for the changes in Malaysia's income distribution between 1984 and 1989 ("Period 1") and between 1989 and 1997 ("Period 2"). These particular years were chosen, because 1997 is the

---

* The authors are very grateful to François Bourguignon, Francisco Ferreira, and Nora Lustig for their support and encouragement at all stages of the project.

most recent available survey, 1984 is the earliest survey comparable to 1997,[1] and 1989 is important for three reasons:

1. Income inequality fell until 1989 and rose thereafter.
2. Economic growth was slow in 1984-89 and fast in 1989-97. and
3. 1989 is the closest year to the beginning of Malaysia's National Development Policy, which placed heightened emphasis on the eradication of hardcore poverty.

The analysis is therefore divided into these two periods, in order to be able to assess the factors responsible for the falling inequality in the first period and the rising inequality in the second. We shall also look at the factors accounting for rising mean income and falling poverty in the two periods. All data are expressed in constant 1997 ringgit.

Two aspects of the income distribution are examined here, and each is measured both non-parametrically and parametrically. The two aspects are location and inequality.

The locational aspect gauges the level of income. The location of any given income distribution is depicted using a quantile function, also called a Pen Parade: $y = F^{-1}(p)$ (i.e., the income amount corresponding to the household at the p'th position in the income distribution). Locational differences are depicted non-parametrically by comparing quantile functions. We also present two summary measures of locational differences: differences in means and differences in poverty headcount ratios.

The inequality aspect tells us how dispersed a given income distribution is. The inequality of any given income distribution is depicted by a Lorenz curve, and non-parametric inequality comparisons may be made by comparing these curves. In

[1] The 1973 Post-Enumeration Survey was not fully comparable with the surveys in later years.

addition to Lorenz curves, two summary measures of inequality are also used: Gini coefficients and Theil indices.

The *location of the distribution,* measured non-parametrically by a quantile function, gives the income amount in real Malaysian ringgit for households at each percentile of the per capita income distribution. These quantile functions are shown in Figure 1. In the 1989-97 period, there was a clear upward movement, which means that at every part of the income distribution, incomes grew. For the 1984-89 period, though, the quantile curve comparisons are much less clear from visual inspection -- this, because per capita among the households in the Household Income and Expenditure Surveys grew at only half-a-percent per year during that period – and so the levels curves in Figure 1 are supplemented by difference curves shown in Figure 2. At each centile of the income distribution, these difference curves display the amount by which that centile's income rose or fell between the base year and the final year. We see in the upper panel of Figure 2 that the differences are all positive, which is another way of establishing that incomes were higher at every position in the income distribution in 1997 than in 1989.[2] This is not the case, however, for the earlier period: although incomes rose throughout most of the income distribution between 1984 and 1989, they fell in the richest three centiles, according to these data.

To supplement these dominance comparisons, we calculated two measures of location -- the mean income and the poverty headcount ratio – for each of the two periods.[3] These measures show that incomes were generally rising in both periods:

[2] This is also called first order stochastic dominance. It implies that for any poverty line, high or low, a smaller percentage of individuals are in poverty in the dominating distribution (in this case, 1997) than in the dominated one.

[3] The poverty line used in this study is half the 1984 median per capita household income.

Location

Location of Actual Distribution of
Per Capita Household Income in
1984 and 1989 (Period 1)

	Mean	Pov Hdct
Actual value, 1984	3,637.76	37.9%
Actual value, 1989	3,752.74	33.7%
Actual change	+3.2%	-11.1%

Location of Actual Distribution of
Per Capita Household Income in
1989 and 1997 (Period 2)

	Mean	Pov Hdct
Actual value, 1989	3,752.74	33.7%
Actual value, 1997	7,070.29	14.4%
Actual change	+88.4%	-57.3%

In summary, although incomes in the 1984-89 period did not become uniformly higher, the two most commonly-used locational measures– the mean income and the poverty headcount ratio (viz., the fraction of people below a fixed real income amount) – show a shift towards higher incomes overall.

The second aspect of the income distribution studied here is the *inequality aspect*. This is measured non-parametrically by a Lorenz curve, which depicts the cumulative percentage of income received by each cumulative percentage of households, ordered from lowest income to highest. The forty-five degree line represents a perfectly equal distribution of income. Therefore, when one Lorenz curve lies closer to the forty-five degree line than another, which is termed "Lorenz dominance," the first income distribution is more equal than the second This means that, as shown by Atkinson (1970) and others, any inequality index obeying the principle of transfers will show lower inequality for the dominant distribution vis-à-vis the dominated one. Figures 3 and 4 depict a dominance relationship in both periods: household income inequality *fell* in Malaysia from 1984 to 1989 and *rose* from 1989 to 1997. As a consequence of Lorenz dominance, two commonly-used inequality indices – the Gini coefficient and the Theil index – also exhibit falling inequality in the first period and rising inequality in the second:

Inequality

Inequality of Actual Distribution
of Per Capita Household Income
in 1984 and 1989 (Period 1)

	Gini	Theil
Actual value, 1984	0.4856	0.4753
Actual value, 1989	0.4610	0.4161
Actual change	-5.1%	-12.5%

Inequality of Actual Distribution
of Per Capita Household Income
in 1989 and 1997 (Period 2)

	Gini	Theil
Actual value, 1989	0.4610	0.4161
Actual value, 1997	0.4993	0.5051
Actual change	+8.3%	+21.4%

These, then, are the basic distributional changes to be explained, about which we ask the following questions:

1. Which factors contributed how much to the increase in household income levels and the fall in absolute poverty in the 1984-89 and 1989-97 periods?
2. Which factors contributed how much to the falling income inequality from 1984 to 1989 and the rising income inequality from 1989 to 1997?

The factors to be examined are changes in households' demographic characteristics, their productive assets, individuals' labor force participation decisions, their opportunities for working in various occupational positions, and the structure of returns to various characteristics in wage employment and self-employment.

Due to lack of data, we have not been able in this study to investigate the ethnic dimension of changing income distribution. This is a question of vital national interest that remains to be explored in the future.

The relative weights of these various factors are quantified using logistic and linear regressions combined in various ways to simulate counterfactual distributions of income. The basis is a two-equation model, the first of which estimates the determinants of occupational position and the second the determinants of earnings conditional on being in a given occupational position. The simulations then involve replacing one year's coefficients or determinants by those from the other year and gauging how different the distributions are. Section II of the paper details the overall methodology. Section III presents the results of the estimation phase. Section IV then gives the simulation methodology and results. Conclusions are summarized in Section V.

II. Methodology

The analysis proceeds by representing the actual income distributions, deriving the simulated income distributions, and comparing the several simulated distributions' ability to fit the actual changes in location and inequality in the 1984-89 and 1989-97 periods.

The actual income distributions

Let $Y_{h\tau}$ represent the income of household h at time τ. Household income is the sum of labor earnings in wage employment, labor earnings in self employment, and other income, summed over all members, all at time τ:

$$Y_{h\tau} = \Sigma_{i\in h} Y_{i\tau}.$$

Household income depends on the demographic make-up of the household, the characteristics of various household members, the productive assets they own, and the returns these productive assets earn in wage employment and in self-employment.[4] This function may be formalized thus:

$Y_{h\tau} = H_\tau (X^D_{h\tau}, X^H_{h\tau}, Y^O_{h\tau}, \Omega_{h\tau}; \beta_\tau, \lambda_\tau)$

where

$Y_{h\tau}$ = Income of household h at time τ,

H_τ = Income-generating function at time τ,

$X^D_{h\tau}$ = Vector of demographic characteristics of household h at time τ,

$X^H_{h\tau}$ = Productive assets owned by household h at time τ,

$Y^O_{h\tau}$ = Other income received by household h at time τ,

$\Omega_{h\tau} \equiv [(\varepsilon_i^w), (\varepsilon_i^{se}), (\eta_i^w), (\eta_i^{se})]$ = Unobserved residuals in the equations determining household members' labor earnings in wage employment (ε_i^w), labor earnings in self-employment (ε_i^{se}), participation in wage employment (η_i^w), and participation in self-employment (η_i^{se}),

$\beta_\tau = (\beta^w_\tau, \beta^{se}_\tau)$ = Regression coefficients in the wage and self-employment equations,

$\lambda_\tau = (\lambda^w_\tau, \lambda^{se}_\tau)$ = Multinomial logit coefficients in the wage employment and self-employment participation equations.

Next, we shall aggregate the observations on each household into an overall economy-wide income distribution. Let X^D_τ, X^H_τ, Y^O_τ, and Ω_τ be vectors denoting the corresponding random variables in the population as a whole. Given the regression coefficients β_τ and the logit coefficients λ_τ, the actual distribution of household incomes at time τ can be written as

$D_\tau = D[X^D_\tau, X^H_\tau, Y^O_\tau, \Omega_\tau; \beta_\tau, \lambda_\tau]$.

[4] Earnings also depend on hours worked in each type of employment, but hours information is not present in the Malaysian data.

Next, the X^D_τ and X^H_τ factors are regrouped into two overlapping sets: those characteristics that enter into the determination of labor earnings (X_τ) and those that enter into the determination of occupational position (Z_τ). Thus, the distribution of household incomes at time τ may be rewritten as

$D_\tau = D[X_\tau, Z_\tau, Y^O_\tau, \Omega_\tau; \beta_\tau, \lambda_\tau]$.

Finally, these relationships are parameterized using two basic equations: (i) a system of occupational position equations, which determines the likelihood that a given person will be a wage employee, self-employed, or a non-earner (either an unpaid family worker or economically inactive), and (ii) an earnings equation, which predicts the individual's earnings within that occupational category. Specifically:

$\text{Prob}(i\text{=self employed})_\tau = \Lambda(\lambda_{j=1\tau} Z_{ij\tau} + \eta_{ij\tau})$

$\text{Prob}(i\text{=wage employee})_\tau = \Lambda(\lambda_{j=2\tau} Z_{ij\tau} + \eta_{ij\tau})$

and

$\ln Y_{ij\tau} = \beta_{j\tau} X_{ij\tau} + \varepsilon_{ij\tau}$,

where

$\Lambda(.)$ is the logistic function,

$\lambda_{j\tau}$ is a set of logit coefficients determining occupational position (wage employee, self-employed, or non-earner) of individual i at time τ,

$Z_{i\tau}$ is a set of determinants of occupational position,

$\eta_{ij\tau}$ are the residuals in the occupation equation,

$\ln Y_{ij\tau}$ is the logarithm of labor income of individual i in occupational position j at time τ,

$\beta_{j\tau}$ is a set of linear regression coefficients,

$X_{ij\tau}$ is a set of determinants of labor income,

and

$\varepsilon_{ij\tau}$ are the residuals in the earnings equation.

The simulated income distributions

To simulate the contributions of groups of explanatory factors to the change in the economy-wide distribution between one year t and another t', the year t' values are substituted in place of the year t values in

$D_\tau = D[X_\tau, Z_\tau, Y^O_\tau, \Omega_\tau; \beta_\tau, \lambda_\tau]$,

holding the other values constant when possible. Five such simulations shall be performed:

i. The "effect of changing the whole reward structure" is defined as the change in the income distribution that would be realized if the year t' values of β are used instead of the year t values:

$B_{tt'} \equiv D[X_t, Z_t, Y^O_t, \Omega_t ; \boldsymbol{\beta_{t'}}, \lambda_t] - D[X_t, Z_t, Y^O_t, \Omega_t ; \boldsymbol{\beta_t}, \lambda_t]$.

ii. Similarly, the "effect of changing the whole occupational position structure" is estimated by using the year t' values of λ rather than the year t values:

$O_{tt'} \equiv D[X_t, Z_t, Y^O_t, \Omega_t ; \beta_t, \boldsymbol{\lambda_{t'}}] - D[X_t, Z_t, Y^O_t, \Omega_t ; \beta_t, \boldsymbol{\lambda_t}]$.

iii. The "effect of changing the whole population structure" is defined as the income distribution that results when the X's and Z's in year t are replaced by those in year t':

$P_{tt'} = D[\mathbf{X_{t'}, Z_{t'}, Y^O_{t'}, \Omega_{t'}}\, ; \beta_t, \lambda_t] - D[\mathbf{X_t, Z_t, Y^O_t, \Omega_t}\, ; \beta_t, \lambda_t]$.[5]

It may be noted that the counterfactual distribution used in describing the change in population structure from t to t' is the same one used in describing the change of both the reward and occupational structures from t' to t. What makes the two expressions different is the reference distribution with which the counterfactual is compared.

iv. We may be interested also in the effect of a change in the returns to the k'th characteristic alone. This is done by replacing β_k in year t by its value in year t', while simultaneously keeping average income constant. Obviously, such a change is meaningful only for analyzing inequality changes.[6]

v. Finally, we may also be interested in the effect of changes in the quantities of a single characteristic such as education. Our way of estimating the population effect of the k'th characteristic alone is to assign to the individual at the p'th position in the education distribution for that gender category in year t the number of years of education at that position in t', holding the Z's and all other X's constant. The contribution of this change to the change in the income distribution between year t and year t' may then be expressed as the population structure effect of the k'th characteristic

$$P_{k,tt'} \equiv D[\mathbf{X_{kt'}}, X_{-t}, Z_t, Y^O_t, \Omega_t; \beta_t, \lambda_t]$$
$$- D[\mathbf{X_{kt}}, X_{-t}, Z_t, Y^O_t, \Omega_t; \beta_t, \lambda_t],$$

[5] A "residual effect" may be defined implicitly by the adding-up requirement that the total change be expressed as the sum of the reward structure effect $B_{tt'}$, the occupational position effect $O_{tt'}$, the population structure effect $P_{tt'}$, and a residual:

$$D_{t'} - D_t = B_{tt'} + O_{tt'} + P_{tt'} + R_{tt'} .$$

This decomposition will not be pursued further here.

[6] Because mean income is kept constant, such a change is of no interest in understanding differences in the location of two years' distributions.

where the changed factor is denoted X_k and the others are denoted X_{-}.

Comparing the five simulations

The final step is to compare the simulations and thereby determine the relative importance for income distribution change of the five simulated factors: the change in the whole reward structure, the change in whole occupational position structure, the change in the whole population structure, the change in the returns to education, and the change in the quantities of education. As described earlier, two aspects of income distribution change are of interest to us: the locational aspect and the inequality aspect.

For each of these aspects of the income distribution, we seek to determine how important different factors are in accounting for the change in income distribution between one year and another. In many cases, we are able to make unambiguous ordinal statements.

We turn now to a more detailed presentation of the implementation of these methods and the empirical results.

III. Estimation Results in the Case of Malaysia

Estimating the determinants of occupational position

At any given time τ, each individual is classified into one and only one occupational position: wage employee, self-employed, or non-earner (which includes those in unpaid family work and the economically inactive).[7] A three-way multinomial logit equation was then run in which the occupational position of individual i at time τ is expressed as a function of the individual's characteristics:

[7] The Malaysian data do not permit multiple classifications.

$$\text{Prob}(i\text{=self employed})_\tau = \Lambda(\lambda_{j=1\tau} Z_{ij\tau} + \eta_{ij\tau})$$

$$\text{Prob}(i\text{=wage employee})_\tau = \Lambda(\lambda_{j=2\tau} Z_{ij\tau} + \eta_{ij\tau}) .$$

In each year, occupational position equations are estimated separately for men and women and for household heads and non-heads.

For heads of households, the $Z_{ij\tau}$ include: (i) The individual's own characteristics: An education spline[8] and an age quartic; and (ii) The average characteristics of family members other than oneself: Their mean education entered as a quadratic, their mean age, also entered as a quadratic, the fraction of them who are female, family size, the household dependency ratio, and rural/urban location. For non-heads of households, the $Z_{ij\tau}$ includes everything that is included for the head, plus the head's own characteristics[9], plus the head's actual occupational position.[10] Since average characteristics of family members other than the head appear here as well, and include the individual whose occupational choice is being estimated, it may have possibly have introduced multicollinearity, but the analysis of the standard errors does not suggest that this happened.

The residuals, $\eta_{ij\tau}$, are interpreted as representing unobserved determinants of occupational choice.

The results of these estimations are reported in Table 1, which consists of four panels -- male heads of household, female heads of household, male non-heads, and female non-heads – in each of the three years. To briefly summarize the results variable by variable:

---

[8] The spline in this case consists of three connected line segments, which allow for the dependent variable to change at one rate for each additional year of primary schooling, at a different rate for each additional year of secondary schooling, and at a third rate for each additional year of higher education.

[9] The Malaysian data set did not include ownership of land or of other productive assets. If it had, we would have included it.

[10] The head's occupational position may be thought of as a proxy for the existence of a family business, thus affecting the occupational choice of other members. As with any proxy, there may be some reporting error.

- Education behaves quite inconsistently. More schooling sometimes increases the likelihood of being a wage employee and sometimes lowers it. On the whole, more schooling decreases the likelihood of being self-employed.
- The age quartic is consistently statistically significant for males and for female non-heads.[11]
- On the whole, the likelihood of being a wage employee is an inverted-U-shaped function of the education of other household members. The effect of others' education on the likelihood of being self-employed shows no consistent pattern or sign.
- A higher dependency ratio, when statistically significant, reduces the likelihood of being a worker of either type.
- Rural residency exhibits inconsistent effects: sometimes positive, sometimes negative, sometimes insignificant.
- For the most part, the head's education and occupational position had no statistically significant impact on the occupational position of non-heads of household. However, one consistently strong relationship was that the head's being self-employed reduces the likelihood that non-heads are either wage employees or self-employed, and therefore increases the likelihood of their being non-earners.
- The percentage of people in the household who are female exhibits no consistent relationship with occupational position.

Overall, these equations explain at most 33% of the variation in occupational position – more typically, about 25%. Because we have not done well in predicting occupational positions from the observed Z's, we expect that the changes in the λ's

[11] Statistical significance of the individual variables implies joint significance of the four variables taken together.

would not explain much of the change in income distribution – a result borne out in Section IV below.

Estimating the earnings conditional on working in occupational position j.

Let ln $Y_{ij\tau}$ denote the log-earnings of individual i if s/he works in sector j at time τ. Mincerian earnings functions are run separately for each sex and occupational position in each year:

$$\ln Y_{ij\tau} = \beta_{j\tau} X_{ij\tau} + \varepsilon_{ij\tau} .$$

$X_{ij\tau}$ includes, for each individual, an education spline, an age quartic, state of residence, and occupation[12]. The estimation method is least squares weighted by survey sampling weights.

The results are reported in Table 2. Again, summarizing briefly, given all of the past work on earnings determination in Malaysia, it is hardly surprising to find that other things equal:

- Education raises earnings.
- The quartic polynomial in age is statistically significant.
- Workers in administrative jobs earn more than professionals while those in other occupations earn less.
- Non-agricultural workers earn more than do agricultural workers.
- Workers in the Kuala Lumpur Federal Territory and in Johor earn more than do workers in other states.

The overall fit of these models is quite good: for male and female wage-earners, half or more of the variance in log-earnings is explained by these variables. For the self-employed, the fit is poorer, which is not surprising given the variability in work hours and in complementary resources among the self-employed.

[12] Industrial sector dummies were not used because of coding changes between 1997 and the other years. The only industrial control that could be trusted was agriculture, whose effect is picked up by the occupation controls.

The next step is to proceed from estimation to simulation, to which we now turn.

IV. Simulating the Role of Various Factors in Explaining the Changing Per-Capita Household Income Distribution in Malaysia

Description of the simulations

The simulations proceed from a change in labor earnings of individuals to a change in the income of the household to a change in the overall household income distribution. In the Malaysian data, labor earnings from wage employment or self-employment are assigned to given individuals. The simulations change these amounts. Any other income received by the individual such as transfer income is assumed to be invariant with respect to any of the simulated changes and is not modeled.

More specifically, the income of household h in a reference year t is the sum of the incomes of each of the household members:

$$Y_{ht} = \Sigma_{i \in h} Y_{it}. \tag{1}$$

Household member i is found in occupational position j (wage employment, self-employment, or non-earner) according to the logit equations

$$\text{Prob}(i=\text{self employed})_\tau = \Lambda(\lambda_{j=1\tau} Z_{ij\tau} + \eta_{ij\tau})$$

and

$$\text{Prob}(i=\text{wage employee})_\tau = \Lambda(\lambda_{j=2\tau} Z_{ij\tau} + \eta_{ij\tau}). \tag{2}$$

Thirdly, the earnings of individual i in occupational position j at time t is a function of a set of income determinants X:

$$\ln Y_{ijt} = \beta_{jt} X_{ijt} + \varepsilon_{ijt} \tag{3}$$

Five simulations are performed by substituting some of the values for a comparison year t' into (1) – (3) in place of the base year (t) values:

Simulation i. "The effect of changing the whole reward structure":

For the reward structure as a whole, change all β's including the constant from their year t values to their year t' values:

$$\ln Y_{ijt}^{i} = \beta_{jt'} X_{ijt} + \varepsilon_{ijt}.$$

Everybody keeps the same occupational category; only the rewards within the category are changed. The residual ε_{ijt} is found by estimating the original wage equation and comparing its prediction to the observed wage. Our interpretation is that the residual represents unobserved determinants of labor income. Consistent with this interpretation, we assign a "price" to these characteristics, which is the variance of the residuals. So whenever we change β_{jt} to $\beta_{jt'}$ we also multiply the residuals by the ratio of their variances in t and t'.

Simulation ii. "The effect of changing the whole occupational position structure": In the occupational position equation, change all λ's from their year t values to their year t' values:

$$Prob(i=\text{self employed})_t^{ii} = \Lambda(\lambda_{j=1t'} Z_{ijt} + \eta_{ijt})$$

$$Prob(i=\text{wage employee})_t^{ii} = \Lambda(\lambda_{j=2t'} Z_{ijt} + \eta_{ijt})$$

The residuals (the η_{ijt}) cannot be uniquely identified from the original estimating equation (2), so they must be assigned. The way this is done is described in the following paragraph.

The residuals in the occupational position equation are not as easily determined as the residuals in the earnings equation. While the former are uniquely identified by the difference between observed and predicted earnings, the latter must be drawn from an inverse hyperbolic secant, which is the distribution consistent with the multinomial logit occupational position equation. However, not any random draw is acceptable, as residuals must be consistent with observed choices. Individual *i*'s choice obeys the following rules:

If $\lambda_{j=1t'} Z_{ijt} + \eta_{ijt} > \lambda_{j=2t'} Z_{ijt} + \eta_{ijt}$ and $\lambda_{j=1t'} Z_{ijt} + \eta_{ijt} > 0$, then individual *i* is self employed.

If $\lambda_{j=2t'} Z_{ijt} + \eta_{ijt} > \lambda_{j=1t'} Z_{ijt} + \eta_{ijt}$ and $\lambda_{j=2t'} Z_{ijt} + \eta_{ijt} > 0$, then individual *i* is a wage earner.

If both $\lambda_{j=2t'} Z_{ijt} + \eta_{ijt} < 0$ and $\lambda_{j=1t'} Z_{ijt} + \eta_{ijt} < 0$, then individual *i* is either inactive or an unpaid family worker.

Note that what changes in these expressions are only the λ's. The Z's, and more significantly, the η_{ijt} residuals, remain the same. This means that the residual drawn must be coherent with the preceding three conditions above to be in accordance with occupational choice observed. An easy way to do this is to draw the residuals randomly from the inverse hyperbolic secant and check to see if they are coherent with observed Z's and choices and estimated λ. For those individuals whose drawn residuals are incoherent, we then redraw them and check again, keeping the drawn residuals for those individuals whose Z's, λ's, and η's were coherent with their observed choices. We keep on redrawing previously incoherent residuals until no more individuals are left with incoherent Z's, λ's, and η's. Generally this takes a few dozen draws.

One final comment is on individuals observed in inactivity but simulated as wage or self employed workers. Since they were observed in inactivity, they have no wage residual associated to them that can be used in the construction of their counterfactual earnings. In this case, new residuals are drawn from normal distributions with zero mean variance equal to the observed variance of the observed residuals of wage or self employed workers.

Simulation iii. "The effect of changing the whole population structure":

For the population structure as a whole, the third simulation changes all X's and all Z's from their year t values to their year t' values:

$$\text{Prob(i=self employed)}^{iii} = \Lambda(\lambda_{j=1t} Z_{ijt'} + \eta_{ijt'})$$

$$\text{Prob(i=wage employee)}^{iii} = \Lambda(\lambda_{j=2t} Z_{ijt'} + \eta_{ijt'})$$

$$\ln Y_{ijt}^{iii} = \beta_{jt} X_{ijt'} + \varepsilon_{ijt'}.$$

(Note: Residuals and other income are considered part of X and Z.) This third simulation puts some people into new occupational categories, and it changes the incomes within occupational categories for others.

Simulation iv. "The effect of a change in the price of education alone":

For the price effect of education alone, we are asking what would happen if the gain in income for an extra year of education were to be changed from the year t to the year t' values, while keeping all other β's constant. This changes only the coefficients on the education spline, using base year values for the coefficients on the other X's, adjusting the constant so that the regression line rotates through the mean:

$$\ln Y_{ijt}^{iv} = \beta_{ed,ijt'} X_{ed,ijt} + \beta_{non\text{-}ed,ijt} X_{non\text{-}ed,ijt} + \varepsilon_{ijt}.$$

By construction, this simulation has no effect on levels, so only its effect on inequality will be looked at.

Simulation v. "The effect of a change in the quantity of education alone":

For the population effect of education alone, the individual at the k'th position in the education distribution for that gender/age category in year t is assigned the number of years of education at that position in t', holding all other X's and the Z's constant:

$$\ln Y_{ijt}^{v} = \beta_{ed,ijt} X_{ed,ijt'} + \beta_{non\text{-}ed,ijt} X_{non\text{-}ed,ijt} + \varepsilon_{ijt}.$$

This is a rank-preserving transformation of the quantity of education each individual possesses.

Earnings within an occupational category are then reestimated with the new, generally-higher years of education substituted in place of the original ones[13].

Assessing the Effects

The relative effects of these five simulations on the location and inequality of the Malaysian income distribution are each assessed both parametrically and non-parametrically. For location, quantile curves are compared as well as specific statistics – the mean income and the poverty rate. For inequality, comparisons are made of Lorenz curves and of two inequality measures – the Gini coefficient and the Theil index.

We look first at the 1984-89 period and then at the 1989-97 period. For each period, there are two sets of simulations: the "A" set takes the earlier year as t and the later year as t', while the "B" set does the reverse.

A criterion is needed for deciding when one effect is more important than another. We shall say that an explanatory factor contributes more to the increase (resp., decrease) in the dependent variable (location as measured by quantile functions, means, and poverty headcounts and inequality as measured by Lorenz curves, Gini coefficients and Theil indexes), the more positive (resp., more negative) is the change in the explanatory factor. In cases where all simulations go part of the way toward explaining an observed change, this ordering rule is simple: the larger the effect, the more important is that explanatory factor. However, some simulated effects may be larger than the observed change and others may be negative, and in these cases, the preceding ranking criterion would say that the most important explanatory factor is the one that goes furthest in the same direction as the observed change, even to the point of overshooting.

[13] Note that in this simulation, the educational endowments change income distributions only through people possessing more years of schooling and earning higher incomes in their pre-existing occupations. The occupational structure is kept fixed.

Results for the 1984-89 Period

Focusing first on the location of the income distribution, as noted earlier, per capita incomes grew by only half a percent a year between 1984 and 1989. Accordingly, the plots of the actual 1984 and 1989 distributions barely diverge, and the simulation graphs in levels can hardly be distinguished from the actual distributions (Figures 5A and 5B). The changes in the actual quantile function between 1984 and 1989, depicted in Figures 6A and 6B, are no more informative: no simulated effect lies everywhere between the zero line and the plus curve, nor does any simulated effect always lie above the actual change. This means that no one factor can be judged to be more important non-parametrically than any other in explaining locational changes between 1984 and 1989. Accordingly, we turn our attention to the two location indices, the mean and the poverty headcount.

The results for these two indices are presented in Tables 3A and 3B. We see that the modest increase in the mean income is accounted for well by the increase in mean education (Simulation v). Indeed, at this time, the population was becoming better-educated (Table 5). According to our estimates, this increase in mean education accounted for 86-87% of the increase in mean income. The changes in the whole population structure, including not only years of education but also everything else, accounted for 145-251% of the increase in the mean. In other words, the actual mean did not increase by as much as the changing population structure would have implied, because other factors were operating to drive the mean downward. The other simulated changes, by contrast, exhibited either small effects (Simulation i) or unstable effects (Simulations ii and iii) on the mean. As for the change in the poverty rate, the change in the population structure accounts almost exactly for the change (100-104%). By contrast, the other factors (the changing reward structure and the changing occupational structure) perform poorly.

Turning now to the inequality aspect of the changing income distribution, Malaysia experienced an unambiguous decrease in inequality from 1984 to 1989, as measured by the Lorenz curves of per capita household income (bottom panels of Figures 3 and 4). To gauge the relative importance of these five different factors, the differences between the actual Lorenz curve and each of the simulated ones are plotted in Figures 7A and 7B.[14] We see that Simulations i and iii fit the actual change most closely, followed by Simulations ii and v, and lastly by Simulation iv. For the B set of simulations, Simulations i and iii also come the closest. Thus, in both sets of simulations, the falling inequality in the 1984-89 period is best accounted for by changes in the reward structure and the population structure, of which a key component was the falling returns to education, particularly higher education (cf. "Higher Education Spline" in Table 2).

A word of explanation is also in order regarding Simulation iv, the results of which appear somewhat paradoxical. During the 1984-89 period, the estimated coefficients of the wage equations shown on Table 2 indicate that the earnings education profile became less steep for all sex and occupation categories. In view of this, why does substituting the 1989 prices of education into the 1984 distribution increase inequality as seen both by Gini and Theil summary measures and also in terms of Lorenz dominance?

The answer, essentially, is the aggregation of the four gender-occupational position groups into families and then into an overall distribution. We examined the Theil and Gini coefficients for each of the four groups separately -- wage earning men, wage earning women, self-employed men, and self-employed women. The summary measures are shown in Table 6. As expected from the regression coefficients, the earnings inequality of Simulation iv within each one of these groups was less than that observed in the original data. Since the average income of each group had been adjusted to remain constant, the increase in inequality could not be attributed to between group inequality.

[14] Because the actual Lorenz curves lie so close together, visual inspection is uninformative, so we have not presented those curves here.

This can be seen by noting that the observed Theil coefficient for the economically active population was 0.483 in 1984, whereas when 1989 education prices are substituted in, it falls to 0.460.

One possibility is that once nonlabor income was added in, some correlation between Simulation iv losers and winners and nonlabor income would lead to the results observed, but this proves not to be the case. Inclusion of nonlabor income, as shown in Table 5, does not change the results: Simulation iv results still decrease the 1984 level of inequality. The Theil index for the economically active population falls from 0.487 to 0.467.

So if Theil indices fall for the economically active population when education prices are substituted into the 1984 distribution, then why do they rise for per capita family income? The reason is that when individuals are aggregated into families, the correlation between simulation losers and winners (remember that the simulation changes only returns to one characteristic – education) and pairing and family size leads to increased per capita family income inequality.

Results for the 1989-97 Period

Between 1989 and 1997, the two principal distributional facts are that the economy became richer at all centiles of the income distribution (see the top panels of Figures 1 and 2) and that the income distribution became unambiguously more unequal (top panels of Figures 3 and 4). Analysis of the same five simulations as in the earlier period suggests the following explanations.

First, as regards the location of the income distribution, the two simulations that perform the best are Simulation iii (changing the whole population structure) and Simulation i (changing the whole reward structure). We can see this in three ways: 1. By comparing the several panels of Figure 8A and 8B; 2. By comparing the five simulated quantile differences to the actual quantile difference (Figures 9A and 9B); and 3. By comparing the five simulated changes in mean and poverty headcount ratio to the actual changes (Tables 4A and 4B). In this period, the educational level of the population was increasing – even more so than in the earlier period (Table 5) – but these increased educational attainments account for only a modest amount of the total change.

If not education, then what explains the large increases in average income from 1989 to 1997? In part, a generalized increase in base income: the constant in the wage equation increases from 1989 to 1997 for all occupational categories, except wage-earning women. For wage-earning men and self-employed women, who together account for around 55% of the population with positive income, the increase in the constant is around 0.5 logarithmic units, meaning an increase in ringgit of about 75%. For self-employed men (around 15% of the those with positive incomes), the log change is 1.5, leading to an increase in ringgit by a factor of almost three.

Another explanation is that there is one dimension of occupational determination that we did not model – the decrease in agricultural occupation. For all years and all modeled occupational categories, there is a very strong negative premium associated with having an agricultural occupation – other things being equal, being in agriculture reduces earnings by around 60%. From 1984 to 1989, the percentage of employed individuals in agriculture remained stable at 32%, but from 1989 to 1997, it fell from by almost half to 17%. This change alone may have increased average incomes by around 10%.

Turning finally to the causes of increased income <u>inequality</u> in Malaysia in the 1989-97 period, Simulation iii exhibits the largest effect in both the A and B set of simulations. From this, we conclude that the increase in inequality in the 1989-97 period is best accounted for by changes in population structure. On the other hand, the factor estimated to be the next largest in the A set of simulations, Simulation iv, works in the opposite direction in the B set. The same aggregation issues discussed in detail for the 1984-89 period probably were at work in the 1989-97 period. Next in importance, the factors that exhibit consistent effects (Simulation i and Simulation v) are also the ones whose effects are small. Interestingly, the very small contribution of education's quantity effect in Malaysia is the exact opposite of what was found in Taiwan. There, the increased equality of years of education was the major factor lowering income inequality (Bourguignon, Fournier, and Gurgand, 1999).

V. Conclusion

This paper set out to answer four questions: Which factors were how important in explaining the rising incomes in Malaysia in 1984-89 and 1989-97, the falling income inequality in 1984-89, and the rising income inequality in 1989-97? Our analysis of the microeconomics of changing income distribution in Malaysia reveals:

1. In the earlier period, the modest increase in mean income and the modest reduction in the poverty headcount ratio are accounted for by the changing population structure.
2. In the earlier period, inequality fell. This is best accounted for by changes in the reward structure and the population structure.
3. In the latter period, mean income rose substantially and the poverty rate fell substantially. The changes in the population structure and the reward structure each make important contributions.
4. For inequality change in the latter period, as for mean income and poverty changes in the same period, these are best explained by changes in the population structure and the reward structure.

REFERENCES

Ahuja, Vinod, Benu Bidani, Francisco Ferreira, and Michael Walton, Everyone's Miracle? Revisiting Poverty and Inequality in East Asia, World Bank, 1997.

Ahuja, Vinod, "Growth with Redistribution?: Inequality and Poverty in Malaysia," World Bank, November, 1997.

Atkinson, Anthony B., "On the Measurement of Inequality," Journal of Economic Theory, 1970.

Bourguignon, François, Francisco Ferreira, and Nora Lustig, "The Microeconomics of Income Distribution Dynamics in East Asia and Latin America," World Bank, April, 1998.

Bourguignon, François, Martin Fournier, and Marc Gurgand, "Fast Development with a Stable Income Distribution: Taiwan, 1979-1994," DELTA, December, 1998.

Economic Planning Unit, Government of Malaysia, The Malaysian Economy in Figures 1998, Kuala Lumpur, 1999.

Milanovic, Branko, "Inequality and Determinants of Earnings in Malaysia, 1989-97," World Bank, December, 1999.

World Bank, "Social Crisis in East Asia: Poverty and Malaysia," Available online at http://www.worldbank.org/poverty/eacrisis/countries/malay/pov1.htm

Table 1A
Occupational Position Equations for Male Heads

	Male Heads 84		Male Heads 89		Male Heads 97	
N	39,056		45,182		30,386	
Log likelihood	-26928.0		-31803.0		-20019.8	
Pseudo R2	0.2432		0.2473		0.2677	
Employee	**Coeficient**	**z value**	**Coeficient**	**z value**	**Coeficient**	**z value**
Primary Education Spline	0.00209	1.85	-0.00172	-1.59	0.00023	0.16
Secondary Education Spline	0.00156	0.75	0.00002	0.01	0.00086	0.47
Higher Education Spline	0.00886	1.00	0.01027	1.60	0.00499	0.73
Age	1.85300	13.55	2.05609	15.10	2.25254	12.44
Age Squared	-0.04661	-10.80	-0.05336	-12.37	-0.05687	-9.84
Age Cubed	0.00044	7.64	0.00053	9.15	0.00054	6.99
Age Quartic	-1.40E-06	-5.12	-1.80E-06	-6.51	-1.76E-06	-4.70
Other household members' mean education	0.0058	2.03	0.0054	2.07	-0.0020	-0.64
Other household members' mean education squared	-0.0001	-2.77	-0.0001	-3.19	0.0000	-0.86
Other household members' mean age	-0.0512	-4.89	-0.0298	-3.12	-0.0282	-2.25
Other household members' mean age squared	0.0005	3.75	0.0002	1.90	0.0001	0.71
Dependency Ratio	-0.0177	-0.11	0.0851	0.55	-0.1769	-0.91
Number of people in household	-0.1067	-10.47	-0.0754	-7.53	-0.0955	-6.44
Percent of members who are female	0.2236	2.17	0.0351	0.38	0.1298	1.11
Rural	0.2097	4.27	0.1781	4.06	0.3048	5.29
Constant	-18.8505	-12.54	-21.2167	-14.16	-23.4937	-11.90
Self Employed	**Coeficient**	**z value**	**Coeficient**	**z value**	**Coeficient**	**z value**
Primary Education Spline	-0.00784	-7.37	-0.01292	-12.96	-0.00685	-5.19
Secondary Education Spline	-0.01653	-7.76	-0.01983	-11.78	-0.01131	-6.08
Higher Education Spline	-0.03074	-3.13	-0.03154	-4.31	-0.02269	-3.11
Age	2.55979	17.94	2.47911	17.70	2.92113	15.55
Age Squared	-0.06916	-15.80	-0.06795	-15.73	-0.07652	-13.38
Age Cubed	0.00075	13.33	0.00075	13.39	0.00080	10.99
Age Quartic	-2.92E-06	-11.22	-2.99E-06	-11.41	-3.00E-06	-8.95
Other household members' mean education	0.0033	1.20	0.0018	0.72	-0.0026	-0.88
Other household members' mean education squared	-0.0001	-2.97	0.0000	-2.05	0.0000	-1.15
Other household members' mean age	-0.0063	-0.64	0.0167	1.86	0.0082	0.69
Other household members' mean age squared	0.0001	1.00	-0.0001	-0.99	-0.0001	-1.14
Dependency Ratio	0.5001	3.16	0.8968	5.95	0.3859	2.01
Number of people in household	-0.0551	-5.75	-0.0317	-3.37	-0.0468	-3.26
Percent of members who are female	0.1324	1.31	0.0128	0.14	0.2271	1.96
Rural	0.6887	14.52	0.7778	18.35	0.9409	16.87
Constant	-28.4668	-17.63	-27.4292	-17.31	-34.1626	-15.86

Table 1B
Occupational Position Equations for Female Heads

	Female Heads 84		Female Heads 89		Female Heads 97	
N	6,092		6,697		4,406	
Log likelihood	-5218.1		-5585.9		-3587.9	
Pseudo R2	0.1721		0.184		0.2055	
Employee						
Primary Education Spline	-0.00973	-5.49	-0.00501	-2.85	-0.00241	-1.11
Secondary Education Spline	0.02970	9.55	0.02027	7.83	0.01852	6.98
Higher Education Spline	0.03715	2.31	0.05905	4.48	0.02561	2.40
Age	0.02401	0.14	0.19955	0.96	0.15186	0.71
Age Squared	0.00534	0.98	0.00155	0.22	0.00277	0.39
Age Cubed	-0.00017	-2.26	-0.00014	-1.44	-0.00015	-1.56
Age Quartic	1.11E-06	3.21	1.11E-06	2.24	1.15E-06	2.39
Other household members' mean education	0.0198	4.87	0.0203	4.96	0.0300	6.60
Other household members' mean education squared	-0.0002	-5.24	-0.0002	-5.36	-0.0002	-6.68
Other household members' mean age	0.0466	3.43	0.0467	3.64	0.0314	2.18
Other household members' mean age squared	-0.0005	-2.83	-0.0006	-3.20	-0.0002	-0.99
Dependency Ratio	-0.1694	-0.92	-0.5716	-3.01	-0.8949	-3.98
Number of people in household	-0.1825	-9.51	-0.1721	-8.79	-0.1742	-7.51
Percent of members who are female	0.2976	2.69	0.2356	2.16	0.3042	2.45
Rural	0.1950	2.60	0.2327	3.15	0.2189	2.46
Constant	-2.2106	-1.19	-4.5755	-2.08	-4.0615	-1.78
Self-emp			**-0.0125563**	**-7.90**	**-0.0065768**	**-3.18**
Primary Education Spline	-0.01473	-8.54	-0.00921	-2.59	-0.00266	-0.79
Secondary Education Spline	-0.01093	-2.11	-0.00769	-0.28	-0.04939	-2.00
Higher Education Spline	-0.00060	-0.02	0.50710	2.29	0.70626	2.07
Age	0.56572	2.29	-0.00850	-1.24	-0.01318	-1.24
Age Squared	-0.01294	-1.65	0.00002	0.21	0.00007	0.51
Age Cubed	0.00012	1.10	0.00000	0.55	0.00000	-0.03
Age Quartic	-4.15E-07	-0.79	1.11E-02	2.99	1.66E-02	3.40
Other household members' mean education	0.0088	2.26	-0.0001	-3.79	-0.0001	-3.75
Other household members' mean education squared	-0.0001	-2.47	0.7374	10.73	0.7776	8.50
Other household members' mean age	-0.0028	-0.23	-0.0001	-0.94	-0.0002	-0.92
Other household members' mean age squared	0.0001	0.77	-0.3764	-2.23	-0.1466	-0.62
Dependency Ratio	0.2932	1.69	-0.1361	-8.39	-0.1269	-5.49
Number of people in household	-0.0879	-5.61	0.2165	2.23	0.4565	3.52
Percent of members who are female	0.0759	0.74	-8.3824	-3.28	-12.5866	-3.21
Rural	0.9015	12.57	0.0127	1.10	0.0165	1.03
Constant	-8.6442	-3.13				

Table 1C
Occupational Position Equations for Male Non-Heads

	Male Non-heads 84		Male Non-heads 89		Male Non-heads 97	
N	30,125		32,968		19,490	
Log likelihood	-19951.4		-23187.6		-11519.0	
Pseudo R2	0.2747		0.2503		0.3348	

Employee	**Coeficient**	**z value**	**Coeficient**	**z value**	**Coeficient**	**z value**
Primary Education Spline	0.0215	15.122	0.0252	17.551	0.0328	14.663
Secondary Education Spline	-0.0239	-22.444	-0.0164	-15.999	-0.0155	-9.623
Higher Education Spline	-0.0325	-8.03	-0.0314	-8.797	-0.0688	-17.592
Age	2.75251	29.475	2.17816	22.931	3.68268	25.588
Age Squared	-0.08989	-22.814	-0.06682	-16.524	-0.12529	-20.111
Age Cubed	0.00123	17.958	0.00086	12.079	0.00181	16.129
Age Quartic	-6.13E-06	-14.79	-4.12E-06	-9.378	-9.61E-06	-13.507
Head's Primary Education Spline	0.0506	2.044	0.0469	2.001	0.1448	3.308
Head's Secondary Education Spline	0.0054	0.217	0.0068	0.29	-0.0471	-1.083
Head's Higher Education Spline	-0.0048	-6.545	-0.0062	-9.174	-0.0077	-8.521
Head Wage Worker	0.1625	3.424	0.1153	2.627	0.2851	4.464
Head Self Employed	-0.7603	-16.817	-0.9299	-22.506	-0.7843	-12.143
Head's Age	-0.0660	-8.736	-0.0895	-12.207	-0.0625	-6.241
Head's Age Squared	0.0006	7.826	0.0008	10.989	0.0006	5.535
Other household members' mean education	-0.00023	-0.081	-0.00059	-0.214	0.00428	1.068
Other household members' mean education square	-0.00002	-0.935	-0.00003	-1.469	-0.00006	-2.457
Other household members' mean age	-0.07876	-7.265	-0.05666	-5.294	-0.07109	-4.847
Other household members' mean age squared	0.00083	4.933	0.00043	2.62	0.00057	2.692
Dependency Ratio	-1.2429	-9.202	-1.2657	-9.436	-1.2411	-6.652
Number of people in household	0.0100	1.706	0.0104	1.762	-0.0077	-0.804
Percent of members who are female	-0.2789	-3.348	-0.3057	-3.922	-0.0820	-0.751
Rural	-0.3841	-11.922	-0.3295	-10.78	-0.2632	-5.875
Constant	-23.9567	-30.015	-19.3282	-23.764	-33.5582	-27.788

Self-Employed	**Coeficient**	**z value**	**Coeficient**	**z value**	**Coeficient**	**z value**
Primary Education Spline	0.0201	10.738	0.0204	11.339	0.0293	9.957
Secondary Education Spline	-0.0302	-17.497	-0.0207	-13.671	-0.0179	-7.706
Higher Education Spline	-0.0687	-5.921	-0.0598	-7.539	-0.0717	-8.789
Age	2.54996	19.639	1.91709	17.629	2.92895	16.84
Age Squared	-0.07722	-15.453	-0.05372	-12.849	-0.08694	-13.245
Age Cubed	0.00099	12.4	0.00063	9.463	0.00108	10.472
Age Quartic	-4.61E-06	-10.35	-2.67E-06	-7.282	-4.84E-06	-8.497
Head's Primary Education Spline	-0.0570	-1.379	0.0649	2.012	-0.0300	-0.409
Head's Secondary Education Spline	0.0550	1.308	-0.0315	-0.963	0.0908	1.234
Head's Higher Education Spline	-0.0038	-2.707	-0.0079	-6.428	-0.0045	-2.767
Head Wage Worker	-0.8832	-10.61	-0.8465	-11.366	-0.7113	-6.497
Head Self Employed	-0.5235	-7.951	-0.5649	-9.792	-0.4950	-5.442
Head's Age	-0.0820	-6.87	-0.0849	-7.542	-0.0275	-1.689
Head's Age Squared	0.0008	6.713	0.0008	7.622	0.0003	2.304
Other household members' mean education	-0.00825	-1.874	-0.01007	-2.53	-0.00808	-1.362
Other household members' mean education square	-0.00001	-0.334	0.00004	1.192	-0.00001	-0.344
Other household members' mean age	-0.08399	-5.505	-0.08254	-5.717	-0.08332	-3.923
Other household members' mean age squared	0.00091	4.46	0.00078	3.947	0.00061	2.18
Dependency Ratio	-1.0785	-4.679	-0.7329	-3.463	-1.2544	-3.882
Number of people in household	0.0356	4.264	0.0114	1.386	-0.0031	-0.215
Percent of members who are female	-0.2687	-1.951	-0.2391	-1.979	-0.1369	-0.776
Rural	-0.0521	-0.975	0.0808	1.663	0.1826	2.551
Constant	-24.6899	-20.094	-19.0942	-18.229	-31.4272	-18.593

Table 1D
Occupational Position Equations for Female Non-Heads

	Female Non-heads 84		Female Non-heads 89		Female Non-heads 97	
N	66,748		76,409		48,010	
Log likelihood	-42365.3		-46667.0		-31708.8	
Pseudo R2	0.1451		0.1603		0.1706	
Employee	**Coeficient**	**z value**	**Coeficient**	**z value**	**Coeficient**	**z value**
Primary Education Spline	-0.0034	-5.15	-0.0004	-0.515	0.0025	2.608
Secondary Education Spline	0.0148	21.578	0.0177	27.184	0.0145	18.966
Higher Education Spline	0.0352	11.57	0.0295	12.001	0.0158	6.955
Age	1.67491	21.484	1.34996	17.853	1.79105	19.422
Age Squared	-0.05940	-17.943	-0.04448	-13.924	-0.06019	-15.47
Age Cubed	0.00088	14.984	0.00062	10.88	0.00085	12.306
Age Quartic	-4.85E-06	-13.01	-3.24E-06	-9.053	-4.49E-06	-10.266
Head's Primary Education Spline	0.0336	1.855	0.0335	1.696	0.0196	0.694
Head's Secondary Education Spline	-0.0430	-2.397	-0.0417	-2.135	-0.0483	-1.734
Head's Higher Education Spline	-0.0008	-1.935	-0.0008	-2.012	-0.0005	-0.961
Head Wage Worker	0.0470	1.327	-0.0060	-0.18	0.1607	3.794
Head Self Employed	-0.9105	-25.628	-0.9937	-29.633	-0.5769	-13.338
Head's Age	0.0036	0.684	-0.0100	-1.911	-0.0277	-4.492
Head's Age Squared	-0.0001	-1.231	0.0001	0.949	0.0002	3.944
Other household members' mean education	0.01834	11.382	0.01828	11.258	0.03353	17.162
Other household members' mean education square	-0.00012	-10.332	-0.00013	-11.859	-0.00022	-17.519
Other household members' mean age	0.08659	13.269	0.08062	13.131	0.09172	13.007
Other household members' mean age squared	-0.00118	-10.882	-0.00102	-10.444	-0.00111	-10.473
Dependency Ratio	-0.6548	-7.317	-0.8010	-8.917	-0.7536	-7.3
Number of people in household	0.0140	3.474	0.0261	6.271	0.0231	4.156
Percent of members who are female	0.3826	7.316	0.3845	7.796	0.3698	6.512
Rural	-0.1427	-6.725	-0.1744	-8.705	-0.2280	-9.238
Constant	-19.0756	-26.816	-16.6686	-23.933	-21.2268	-25.008
Self-Employed	**Coeficient**	**z value**	**Coeficient**	**z value**	**Coeficient**	**z value**
Primary Education Spline	-0.0030	-3.101	-0.0035	-3.728	-0.0014	-1.071
Secondary Education Spline	-0.0087	-5.187	-0.0049	-3.504	0.0009	0.595
Higher Education Spline	-0.0168	-1.271	-0.0129	-1.395	-0.0221	-3.138
Age	0.85015	7.456	0.89822	7.694	1.36133	7.955
Age Squared	-0.02173	-5.244	-0.02240	-5.351	-0.03713	-6.081
Age Cubed	0.00024	3.796	0.00024	3.837	0.00044	4.769
Age Quartic	-1.06E-06	-3.065	-1.06E-06	-3.094	-2.01E-06	-4.006
Head's Primary Education Spline	-0.0162	-0.59	-0.0037	-0.128	0.0466	1.149
Head's Secondary Education Spline	-0.0018	-0.064	-0.0124	-0.43	-0.0637	-1.588
Head's Higher Education Spline	-0.0066	-7.242	-0.0046	-5.618	-0.0038	-3.898
Head Wage Worker	-0.3586	-5.653	-0.3779	-6.381	-0.1490	-1.868
Head Self Employed	-0.6048	-10.202	-0.6650	-12.062	-0.3167	-4.147
Head's Age	-0.0280	-2.918	-0.0124	-1.249	0.0018	0.132
Head's Age Squared	0.0002	2.469	0.0001	1.378	0.0000	0.211
Other household members' mean education	0.01515	5.574	0.01529	5.816	0.00990	3.118
Other household members' mean education square	-0.00011	-3.976	-0.00011	-4.82	-0.00005	-2.135
Other household members' mean age	0.01740	1.936	0.00867	0.978	0.02605	2.264
Other household members' mean age squared	0.00000	0	0.00002	0.192	-0.00024	-1.641
Dependency Ratio	0.5484	3.773	0.3783	2.604	0.6152	3.407
Number of people in household	-0.0427	-5.441	-0.0363	-4.602	-0.0309	-2.708
Percent of members who are female	0.1892	2.08	0.2735	3.173	0.2785	2.631
Rural	0.2726	6.954	0.1193	3.191	0.1100	2.397
Constant	-13.3864	-11.164	-14.7904	-11.845	-20.9209	-11.358

Table 2A
Earnings Functions for Male Wage-Earners

	Wage Men 84		Wage Men 89		Wage Men 97	
N	42,212		43,462		31,037	
R2	0.4954		0.5268		0.5269	
logy	**Coefficients**	**t values**	**Coefficients**	**t values**	**Coefficients**	**t values**
Personal Characteristics						
Primary Education Spline	0.0389	0.0001905	0.0373	0.0001922	0.0361	0.0002724
Secondary Education Spline	0.0624	0.0001717	0.0623	0.0001512	0.0623	0.0001894
Higher Education Spline	0.2316	0.0007179	0.1841	0.0005963	0.1883	0.0006319
Age	0.28416	0.0153997	0.16584	0.0138697	0.18552	0.0210321
Age Squared	-0.00642	0.0005932	-0.00203	0.0005257	-0.00388	0.0008202
Age Cubed	0.00006	0.00000957	0.00000	0.00000835	0.00004	0.0000135
Age Quartic	-2.36E-07	5.47E-08	1.14E-07	4.71E-08	-1.67E-07	7.93E-08
Occupation Type Controls						
Administrative	0.6418	0.0180555	0.6344	0.0156284	0.5910	0.0159054
Clerical	-0.1411	0.013193	-0.1460	0.0116656	-0.2395	0.0136742
Sales	-0.1914	0.0151369	-0.2449	0.0135295	-0.2025	0.0160976
Service	-0.2094	0.0133601	-0.2311	0.0117256	-0.2988	0.0138895
Agriculture	-0.5126	0.013968	-0.5117	0.0124101	-0.5087	0.0157854
Transport	-0.2186	0.0122767	-0.2762	0.0107799	-0.2266	0.0117937
State Controls						
Kedah	-0.3195	0.0131683	-0.2945	0.0120095	-0.3911	0.0146376
Kelantan	-0.3686	0.0144275	-0.3878	0.0131978	-0.5808	0.017606
Malacca	-0.1000	0.0172556	-0.1212	0.0155574	-0.1521	0.0183993
Negeri Sembilan	-0.0254	0.0150409	-0.0507	0.0148235	-0.0977	0.0177568
Pahang	-0.0224	0.0140924	-0.0482	0.0127214	-0.1961	0.0161556
Penang	-0.1685	0.0124166	-0.1309	0.0113615	-0.1667	0.0141379
Perak	-0.0861	0.011218	-0.1569	0.0102817	-0.1796	0.0128203
Perlis	-0.2334	0.0269607	-0.2267	0.0252113	-0.3898	0.0356908
Selangor	0.0687	0.0104173	0.0028	0.0091038	-0.0062	0.0109967
Terengganu	-0.2344	0.0158872	-0.2484	0.0151346	-0.4515	0.0179932
Sabah	0.1780	0.0124087	0.0873	0.0110471	-0.3425	0.0134478
Sarawak	0.1458	0.0128018	0.0996	0.010382	-0.1750	0.0136306
Kuala Lumpur Federal Territory	0.0469	0.0116526	0.0147	0.0107884	0.0968	0.0133574
Labuan Federal Territory	0.3142	0.0408939	0.2288	0.0416833	-0.2285	0.0510746
Rural	-0.1195	0.0060626	-0.0816	0.0054616	-0.1360	0.0069082
Constant	4.8462	0.1421178	5.9796	0.1305984	6.5495	0.1926577

Professional is the omitted category for type of occupation dummies.
Agriculture is the omitted category for industry dummies.
Johor is the omitted category for State dummies.

Table 2B
Earnings Functions for Female Wage-Earners

	Wage Women 84		Wage Women 89		Wage Women 97	
N	19,518		21,361		16,458	
R2	0.4566		0.5206		0.4838	
logy	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Personal Characteristics						
Primary Education Spline	0.0502	0.0003053	0.0466	0.000272	0.0311	0.0003927
Secondary Education Spline	0.0704	0.0003449	0.0682	0.0002629	0.0756	0.0003156
Higher Education Spline	0.2326	0.0011348	0.2123	0.0008219	0.2086	0.0008438
Age	0.19257	0.0244384	0.00263	0.0238335	0.09730	0.0268184
Age Squared	-0.00413	0.0009452	0.00319	0.0009475	-0.00127	0.0010287
Age Cubed	0.00003	0.0000152	-0.00008	0.0000158	0.00000	0.0000165
Age Quartic	-7.19E-08	8.61E-08	5.33E-07	9.37E-08	4.39E-08	9.36E-08
Occupation Type Controls						
Administrative	0.5346	0.0521472	0.5470	0.0376302	0.5078	0.0273037
Clerical	-0.0917	0.0177607	-0.0719	0.0139077	0.0168	0.0154189
Sales	-0.3621	0.0254899	-0.3847	0.0196092	-0.1791	0.0218852
Service	-0.4284	0.0217624	-0.4186	0.0166567	-0.3273	0.0194526
Agriculture	-0.5801	0.023592	-0.5453	0.0194627	-0.5893	0.028392
Transport	-0.4067	0.0203835	-0.3991	0.0155248	-0.1772	0.0173669
State Controls						
Kedah	-0.3575	0.0217858	-0.3202	0.0184167	-0.3000	0.0215843
Kelantan	-0.3967	0.0282192	-0.3777	0.0214635	-0.5953	0.0274921
Malacca	0.0426	0.0277971	-0.0954	0.0215031	-0.0862	0.0263248
Negeri Sembilan	0.0516	0.0242397	-0.0309	0.0214229	-0.0878	0.0257788
Pahang	-0.0408	0.0246622	-0.1224	0.0204205	-0.2123	0.0267125
Penang	-0.0262	0.0199427	0.0375	0.0156935	-0.1145	0.0194869
Perak	-0.1108	0.0185593	-0.2310	0.0152549	-0.1606	0.0191274
Perlis	-0.3763	0.056986	-0.3087	0.0422209	-0.4496	0.0531267
Selangor	0.1340	0.0169423	0.0666	0.0131163	0.0458	0.0160888
Terengganu	-0.3847	0.032457	-0.5049	0.0270794	-0.5387	0.0302347
Sabah	0.1897	0.0231272	0.0788	0.0188336	-0.2162	0.021585
Sarawak	0.0115	0.0241385	-0.0823	0.0171885	-0.2471	0.0215064
Kuala Lumpur Federal Territory	0.1267	0.0187183	0.0483	0.0155756	0.1510	0.018729
Labuan Federal Territory	0.3133	0.0948176	0.0827	0.0842317	-0.1052	0.0972086
Rural	-0.0968	0.010786	-0.0787	0.0082683	-0.1307	0.0106025
Constant	5.7326	0.2234147	7.5171	0.2140876	7.1942	0.2490198

Professional is the omitted category for type of occupation dummies.
Agriculture is the omitted category for industry dummies.
Johor is the omitted category for State dummies.

Table 2C
Earnings Functions for Male Self-Employed

	Self Men 84		Self Men 89		Self Men 97	
N	14,760		18,153		10,296	
R2	0.4024		0.3591		0.4774	
logy	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Personal Characteristics						
Primary Education Spline	0.0457	0.0003598	0.0382	0.0002983	0.0302	0.0004276
Secondary Education Spline	0.0345	0.0005819	0.0081	0.000418	0.0331	0.00049
Higher Education Spline	0.3080	0.0040602	0.2716	0.0028335	0.2269	0.0023625
Age	0.22692	0.0365089	0.18452	0.0298886	0.08582	0.0417171
Age Squared	-0.00442	0.0012065	-0.00312	0.0010029	-0.00010	0.0013409
Age Cubed	0.00003	0.0000168	0.00001	0.0000142	-0.00003	0.0000183
Age Quartic	-5.23E-08	8.43E-08	3.34E-08	7.22E-08	1.97E-07	8.94E-08
Occupation Type Controls						
Administrative	0.7193	0.0755071	0.5252	0.058264	0.6405	0.0609656
Clerical	0.1450	0.1655865	-0.1977	0.1536284	-0.2662	0.1832357
Sales	0.2060	0.0685357	0.1295	0.0514717	0.1104	0.0561268
Service	0.2399	0.0740804	0.2028	0.0574954	0.2022	0.0623777
Agriculture	-0.5708	0.0685861	-0.6016	0.0514621	-0.6660	0.0574319
Transport	-0.0530	0.0698278	-0.1305	0.0527441	-0.1289	0.0578391
State Controls						
Kedah	-0.5506	0.0301611	-0.4238	0.0243937	-0.4177	0.031252
Kelantan	-0.7094	0.0324645	-0.5582	0.0256289	-0.8000	0.032281
Malacca	-0.3005	0.0501995	-0.3475	0.0410869	-0.3181	0.0542258
Negeri Sembilan	-0.0753	0.043619	-0.0391	0.0327049	-0.1309	0.0415467
Pahang	0.1768	0.0333628	0.1363	0.0252698	-0.1363	0.0339798
Penang	-0.4545	0.0356298	-0.3051	0.0304199	-0.2203	0.0403336
Perak	-0.3209	0.028331	-0.3400	0.0232337	-0.1953	0.0309794
Perlis	-0.6057	0.0577062	-0.3310	0.0498519	-0.3992	0.0631633
Selangor	-0.1017	0.0335816	-0.1466	0.0264077	-0.1205	0.0323531
Terengganu	-0.5005	0.0370056	-0.5494	0.0290478	-0.7039	0.0406846
Sabah	-0.1784	0.034745	-0.2077	0.0255397	-0.4927	0.0328248
Sarawak	-0.3089	0.0302265	-0.1176	0.0220151	-0.3287	0.0311621
Kuala Lumpur Federal Territory	0.0784	0.0401453	-0.0700	0.0359767	0.0114	0.0413027
Labuan Federal Territory	0.1450	0.175511	-0.0366	0.1528962	-0.1279	0.1469839
Rural	-0.1983	0.018096	-0.1309	0.0147298	-0.1871	0.0190435
Constant	5.5721	0.3987639	6.1301	0.3205579	7.6345	0.4664547

Professional is the omitted category for type of occupation dummies.
Agriculture is the omitted category for industry dummies.
Johor is the omitted category for State dummies.

Table 2D
Earnings Functions for Female Self-Employed

	Self Women 84		Self Women 89		Self Women 97	
N	4,740		5,238		3,361	
R2	0.2624		0.253		0.2743	
logy	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Personal Characteristics						
Primary Education Spline	0.0271	0.0007504	0.0261	0.0006077	0.0369	0.0008634
Secondary Education Spline	0.0429	0.001525	0.0258	0.0010613	0.0360	0.0010827
Higher Education Spline	0.4955	0.0135178	0.5051	0.0081062	0.4509	0.006585
Age	-0.06235	0.0820761	-0.12307	0.0737428	-0.13394	0.0897577
Age Squared	0.00385	0.0027932	0.00550	0.0024562	0.00543	0.0028281
Age Cubed	-0.00007	0.0000402	-0.00009	0.0000347	-0.00008	0.0000377
Age Quartic	3.85E-07	0.000000207	4.38E-07	0.000000176	3.61E-07	0.00000018
Occupation Type Controls						
Administrative	0.6477	0.253743	0.9640	0.2015068	1.3937	0.1734257
Clerical	0.5204	0.8695833	0.3858	0.4428975	-0.6466	0.479348
Sales	0.5587	0.1116119	0.6119	0.0878289	0.3323	0.1010374
Service	0.3763	0.114728	0.4224	0.0913222	-0.2267	0.1028117
Agriculture	-0.3495	0.1099618	-0.1088	0.0884907	-0.4034	0.1043874
Transport	-0.3703	0.1117319	-0.1466	0.0888045	-0.4564	0.1030985
State Controls						
Kedah	-0.3476	0.0680422	-0.2078	0.0528805	-0.2666	0.0760583
Kelantan	-0.5907	0.058756	-0.3795	0.049004	-0.4668	0.0657163
Malacca	-0.1279	0.0976855	-0.2759	0.0894252	0.0038	0.1140702
Negeri Sembilan	0.1023	0.0909417	0.1481	0.0684099	-0.0169	0.0952677
Pahang	0.0898	0.0796982	0.1735	0.0625389	0.0013	0.0890748
Penang	-0.3133	0.0850171	0.0225	0.0771066	-0.1303	0.0863604
Perak	-0.4626	0.0588165	-0.2652	0.049564	-0.1247	0.0708262
Perlis	-0.7558	0.1743062	-0.1521	0.1119044	-0.1429	0.15133
Selangor	0.0248	0.0774129	0.1454	0.0592302	-0.0927	0.0716125
Terengganu	-0.4968	0.0710387	-0.3404	0.0602639	-0.4435	0.0898252
Sabah	-0.0934	0.0739025	0.1433	0.0620403	-0.1166	0.0797324
Sarawak	-0.2545	0.064145	-0.0310	0.0502364	-0.1980	0.0715687
Kuala Lumpur Federal Territory	0.1785	0.0984459	0.0617	0.0787426	0.2634	0.0877873
Labuan Federal Territory	0.0449	0.4972773	0.3628	0.3519849	-0.0330	0.5882578
Rural	-0.2921	0.0351709	-0.2845	0.029582	-0.2148	0.0381521
Constant	8.0816	0.8626305	8.6192	0.7928124	9.3229	1.02006

Professional is the omitted category for type of occupation dummies.
Agriculture is the omitted category for industry dummies.
Johor is the omitted category for State dummies.

Table 3A.
Distribution of Per Capita Household Income
Substituting 1989 values into 1984 Distribution (Period 1)

	Location		Inequality	
	Mean	Pov Hdct	Gini	Theil
Actual value, 1984	3,637.76	37.9%	0.4856	0.4753
Actual value, 1989	3,752.74	33.7%	0.4610	0.4161
Actual change (as a % of the total change)	3.2%	-11.1%	-5.1%	-12.5%
Simulations substituting the 1989 values into the 1984 distribution:				
i. Change the whole reward structure by substituting 1989 betas into the 1984 distribution	3,650.23	35.9%	0.4710	0.4415
ii. Change the whole occupational structure by substituting 1989 lambdas into the 1984 distribution	3,529.35	38.9%	0.4871	0.4761
iii. Change the whole population structure by substituting 1989 x's and z's into the 1984 distribution	3,804.83	33.7%	0.4708	0.4408
iv. Price effect of education alone: substitute 1989 price of education into the 1984 distribution	n.a.	n.a.	0.5078	0.5152
v. Population effect of education alone: substitute 1989 years of education into the 1984 distribution	3,738.16	36.4%	0.4851	0.4731
Simulated change (as a % of the total change), substituting the 1989 values into the 1984 distribution:				
i. Change the whole reward structure by substituting 1989 betas into the 1984 distribution	10.8%	46.6%	59.5%	57.1%
ii. Change the whole occupational structure by substituting 1989 lambdas into the 1984 distribution	-94.3%	-23.7%	-6.0%	-1.4%
iii. Change the whole population structure by substituting 1989 x's and z's into the 1984 distribution	145.3%	100.1%	60.0%	58.2%
iv. Price effect of education alone: substitute 1989 price of education into the 1984 distribution	n.a.	n.a.	-90.3%	-67.4%
v. Population effect of education alone: substitute 1989 years of education into the 1984 distribution	87.3%	35.2%	2.2%	3.7%

Table 3B.
Distribution of Per Capita Household Income
Substituting 1984 values into 1989 Distribution (Period 1)

	Location		Inequality	
	Mean	Pov Hdct	Gini	Theil
Actual value, 1989	3,752.74	33.7%	0.4610	0.4161
Actual value, 1984	3,637.76	37.9%	0.4856	0.4753
Actual change (as a % of the total change)	-3.1%	12.5%	5.3%	14.2%
Simulations substituting the 1984 values into the 1989 distribution:				
i. Change the whole reward structure by substituting 1984 betas into the 1989 distribution	3,740.57	35.5%	0.4753	0.4445
ii. Change the whole occupational structure by substituting 1984 lambdas into the 1989 distribution	3,739.95	32.8%	0.4577	0.4183
iii. Change the whole population structure by substituting 1984 x's and z's into the 1989 distribution	3,463.74	38.1%	0.4725	0.4429
iv. Price effect of education alone: substitute 1984 price of education into the 1989 distribution	n.a.	n.a.	0.5024	0.5007
v. Population effect of education alone: substitute 1984 years of education into the 1989 distribution	3,654.20	34.8%	0.4593	0.4138
Simulated change (as a % of the total change), substituting the 1984 values into the 1989 distribution:				
i. Change the whole reward structure by substituting 1984 betas into the 1989 distribution	10.6%	44.1%	58.2%	48.0%
ii. Change the whole occupational structure by substituting 1984 lambdas into the 1989 distribution	11.1%	-20.0%	-13.2%	3.7%
iii. Change the whole population structure by substituting 1984 x's and z's into the 1989 distribution	251.3%	104.3%	46.9%	45.3%
iv. Price effect of education alone: substitute 1984 price of education into the 1989 distribution	n.a.	n.a.	168.2%	142.8%
v. Population effect of education alone: substitute 1984 years of education into the 1989 distribution	85.7%	26.6%	-7.0%	-3.8%

Table 4A.
Distribution of Per Capita Household Income
Substituting 1997 values into 1989 Distribution (Period 2)

	Location		Inequality	
	Mean	Pov Hdct	Gini	Theil
Actual value, 1989	3,752.74	33.7%	0.4610	0.4161
Actual value, 1997	7,070.29	14.4%	0.4993	0.5051
Actual change (as a % of the total change)	88.4%	-57.3%	8.3%	21.4%
Simulations substituting the 1997 values into the 1989 distribution:				
i. Change the whole reward structure by substituting 1997 betas into the 1989 distribution	4,768.10	26.0%	0.4802	0.4495
ii. Change the whole occupational structure by substituting 1997 lambdas into the 1989 distribution	3,796.67	32.1%	0.4548	0.3938
iii. Change the whole population structure by substituting 1997 x's and z's into the 1989 distribution	5,252.83	22.4%	0.4963	0.6095
iv. Price effect of education alone: substitute 1997 price of education into the 1989 distribution	n.a.	n.a.	0.4960	0.4836
v. Population effect of education alone: substitute 1997 years of education into the 1989 distribution	3,973.48	31.3%	0.4634	0.4188
Simulated change (as a % of the total change), substituting the 1997 values into the 1989 distribution:				
i. Change the whole reward structure by substituting 1997 betas into the 1989 distribution	983.1%	280.7%	21.9%	43.5%
ii. Change the whole occupational structure by substituting 1997 lambdas into the 1989 distribution	138.2%	136.5%	125.0%	137.6%
iii. Change the whole population structure by substituting 1997 x's and z's into the 1989 distribution	1404.7%	366.7%	-43.4%	-226.4%
iv. Price effect of education alone: substitute 1997 price of education into the 1989 distribution	n.a.	n.a.	-42.1%	-13.9%
v. Population effect of education alone: substitute 1997 years of education into the 1989 distribution	292.0%	156.0%	90.2%	95.3%

Table 4B.
Distribution of Per Capita Household Income
Substituting 1989 values into 1997 Distribution (Period 2)

	Location		Inequality	
	Mean	Pov Hdct	Gini	Theil
Actual value, 1997	7,070.29	14.4%	0.4993	0.5051
Actual value, 1989	3,752.74	33.7%	0.4610	0.4161
Actual change (as a % of the total change)	-46.9%	134.3%	-7.7%	-17.6%
Simulations substituting the 1989 values into the 1997 distribution:				
i. Change the whole reward structure by substituting 1989 betas into the 1997 distribution	5,529.86	19.8%	0.4822	0.4708
ii. Change the whole occupational structure by substituting 1989 lambdas into the 1997 distribution	6,761.98	16.8%	0.5177	0.6778
iii. Change the whole population structure by substituting 1989 x's and z's into the 1997 distribution	4,878.73	24.7%	0.4774	0.4362
iv. Price effect of education alone: substitute 1989 price of education into the 1997 distribution	n.a.	n.a.	0.5258	0.5607
v. Population effect of education alone: substitute 1989 years of education into the 1997 distribution	6,607.18	15.7%	0.4942	0.4962
Simulated change (as a % of the total change), substituting the 1989 values into the 1997 distribution:				
i. Change the whole reward structure by substituting 1989 betas into the 1997 distribution	-52.0%	47.9%	0.7%	0.9%
ii. Change the whole occupational structure by substituting 1989 lambdas into the 1997 distribution	-2617.2%	55.8%	-6.6%	-42.6%
iii. Change the whole population structure by substituting 1989 x's and z's into the 1997 distribution	-979.3%	34.9%	1.7%	8.2%
iv. Price effect of education alone: substitute 1989 price of education into the 1997 distribution	n.a.	n.a.	-8.3%	-18.0%
v. Population effect of education alone: substitute 1989 years of education into the 1997 distribution	-2482.6%	58.5%	-1.8%	-4.4%

Table 5

Rising Educational Attainments in Malaysia, 1984-97.
(Percentage of population aged 14 and above)

School Level	**1984**	**1989**	**1997**
Primary	35.96	33.22	26.78
Junior high	21.18	22.92	21.36
Senior high	17.87	20.61	28.74
University	4.84	6.21	11.19
Religious education only	20.15	17.94	11.93
Total	100	100	100
Average years of schooling, excluding religious education	**1984**	**1989**	**1997**
Men	7.07	7.50	8.61
Women	5.60	6.25	7.60
All	6.32	6.86	8.10
Average years of schooling including religious education	**1984**	**1989**	**1997**
Men	8.03	8.35	9.27
Women	7.80	8.20	9.12
All	7.92	8.28	9.20

Source: Milanovic (1999)

Table 6.
Actual and Simulated Inequality for Disaggregated Gender-Occupational Position Groups.

Theil T for Wage Income by Sex and Activity Category

	Wage Men	Wage Women	Self-Empl.Men	Self-Empl. Women	All Groups
Observed 84	0.369	0.308	0.710	0.645	0.483
89 betas, 84 others	0.348	0.297	0.680	0.638	0.460
Observed 89	0.351	0.295	0.472	0.525	0.409
84 betas, 89 others	0.373	0.306	0.508	0.537	0.433

Gini Coefficient for Wage Income by Sex and Activity Category

	Wage Men	Wage Women	Self-Empl.Men	Self-Empl. Women
Observed 84	0.423	0.414	0.565	0.568
89 betas, 84 otl	0.414	0.408	0.555	0.565
Observed 89	0.415	0.407	0.485	0.520
84 betas, 89 otl	0.425	0.414	0.499	0.524

Theil T for all Income (Wage Income + Other Income)

	Wage Men	Wage Women	Self-Empl.Men	Self-Empl. Women	All Groups
Observed 84	0.388	0.311	0.630	0.550	0.487
89 betas, 84 others	0.370	0.301	0.603	0.545	0.467
Observed 89	0.362	0.300	0.412	0.458	0.405
84 betas, 89 others	0.381	0.310	0.442	0.466	0.427

Gini Coefficient for all Income (Wage Income + Other Income)

	Wage Men	Wage Women	Self-Empl.Men	Self-Empl. Women
Observed 84	0.432	0.415	0.532	0.530
89 betas, 84 otl	0.424	0.409	0.523	0.528
Observed 89	0.422	0.409	0.451	0.491
84 betas, 89 otl	0.431	0.415	0.463	0.494

Figure 1
Changing Quantile Functions

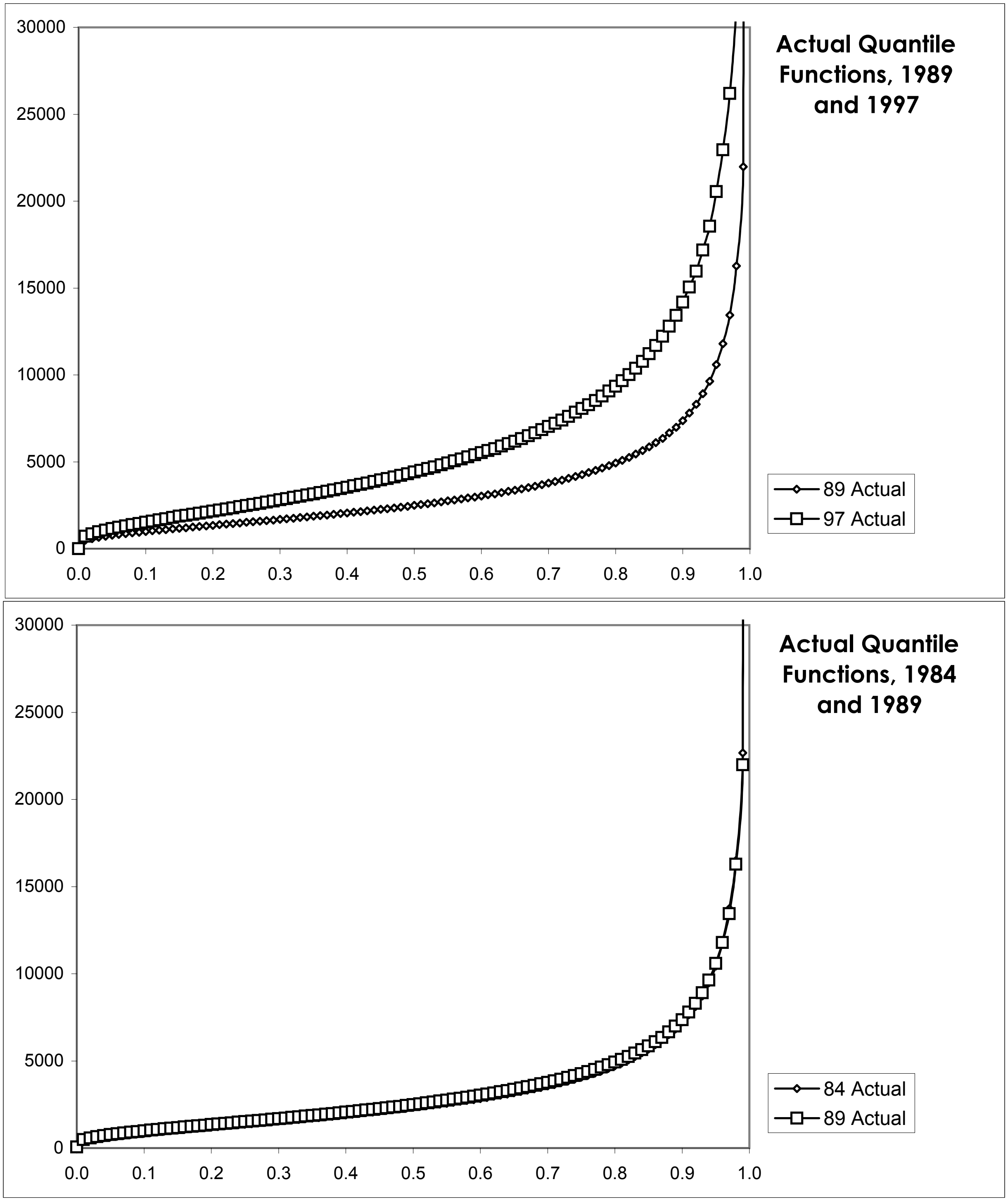
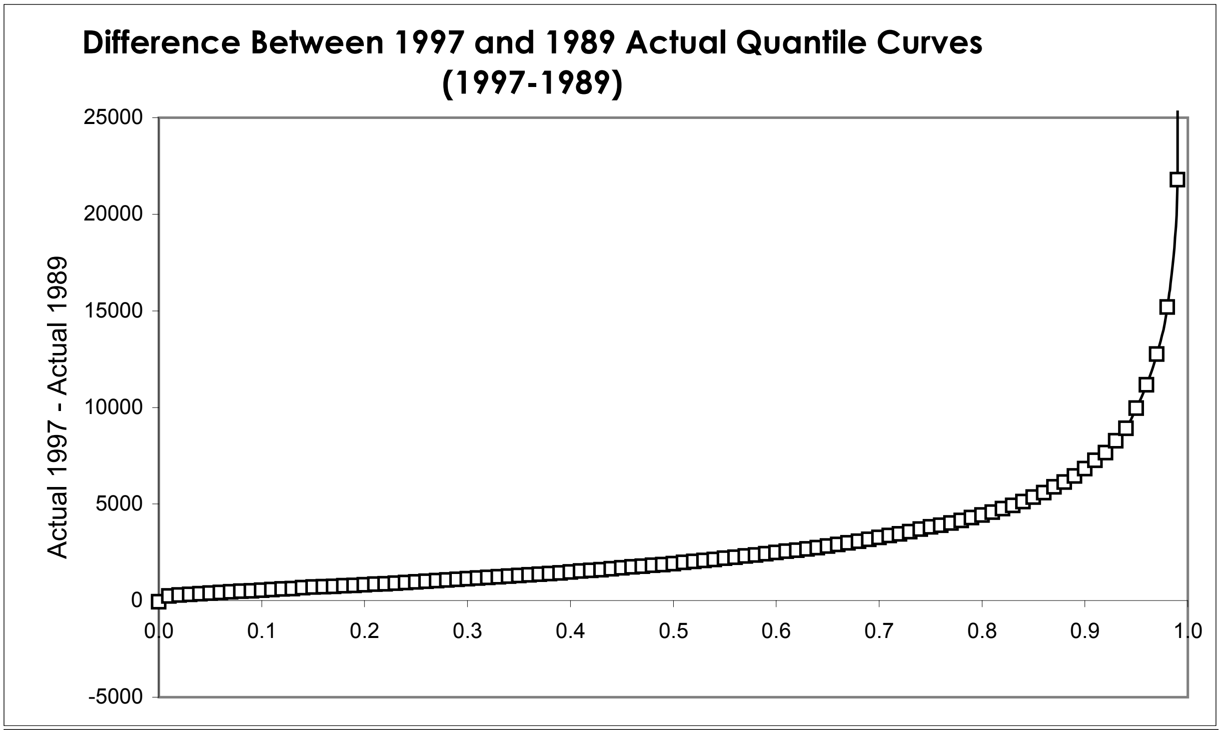


Figure 2
Differences in Quantile Functions



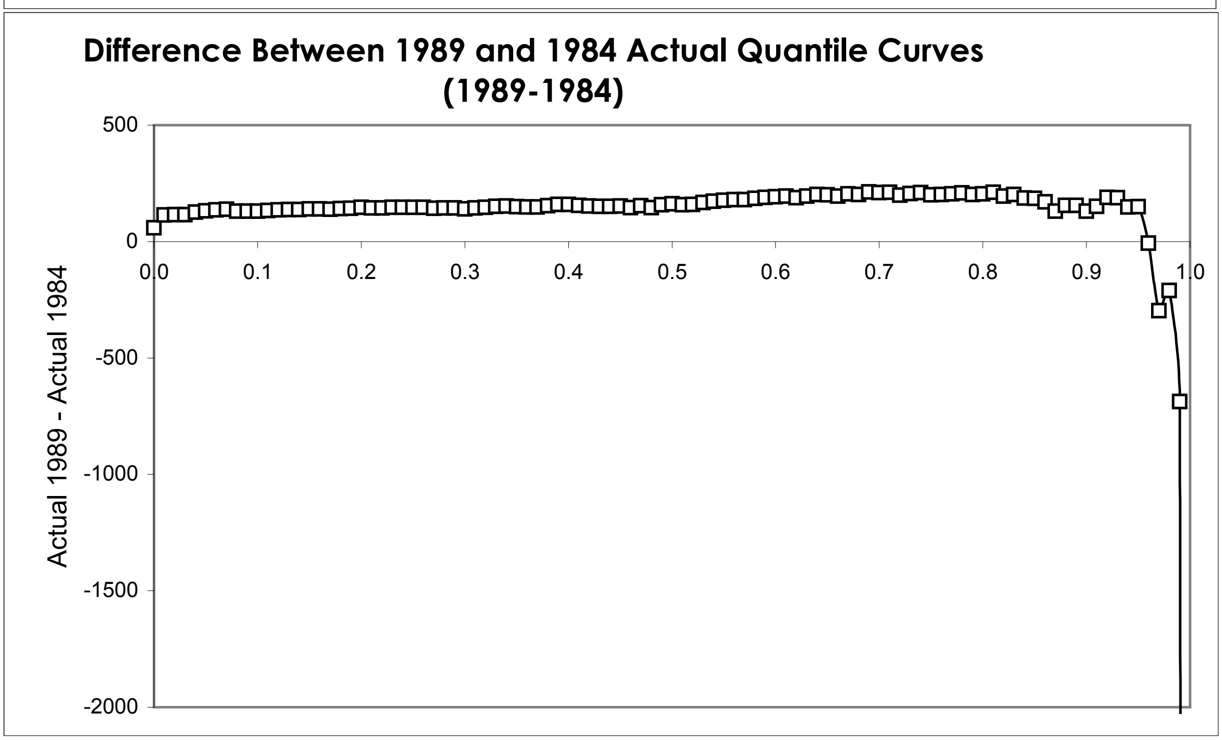
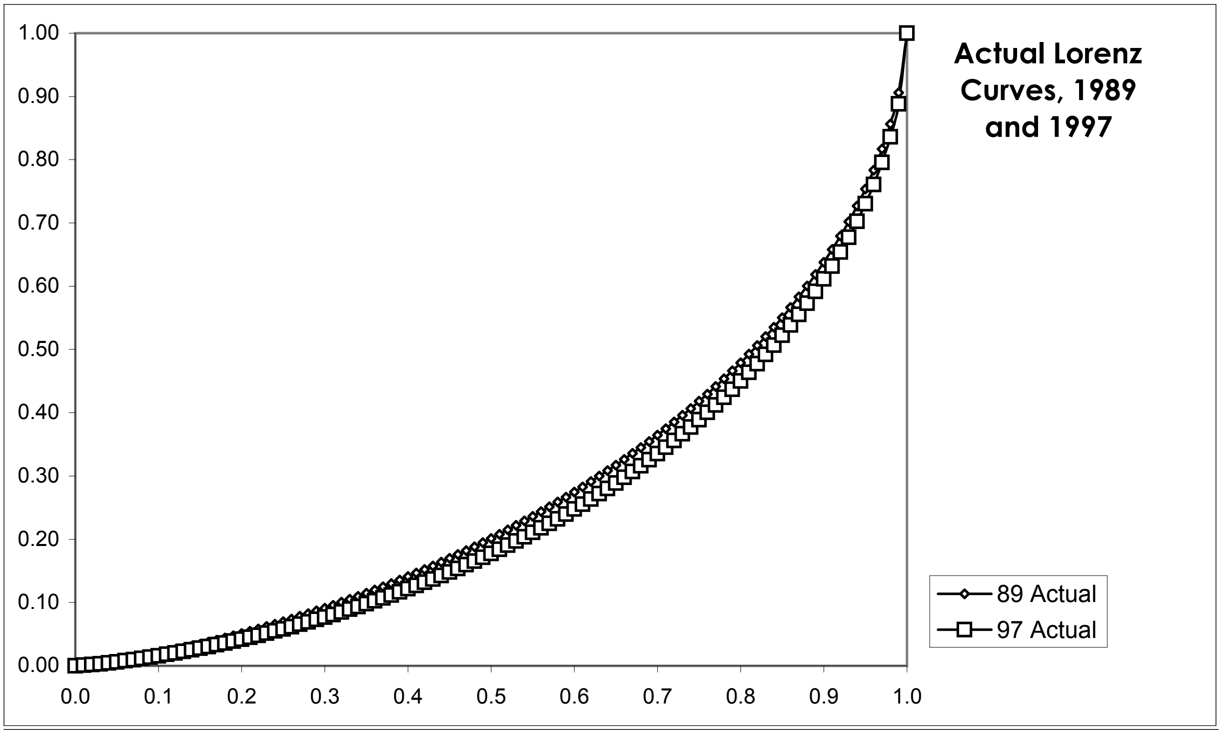


Figure 3
Changing Lorenz Curves



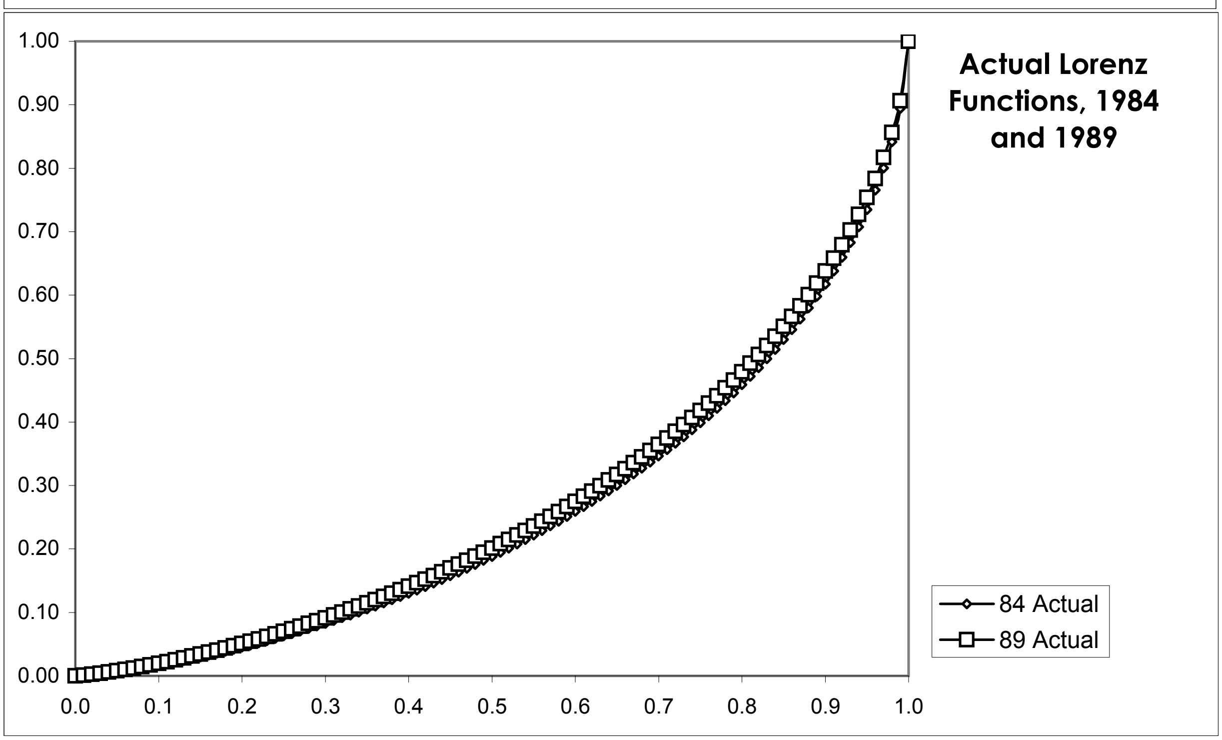


Figure 4
Differences in Lorenz Curves

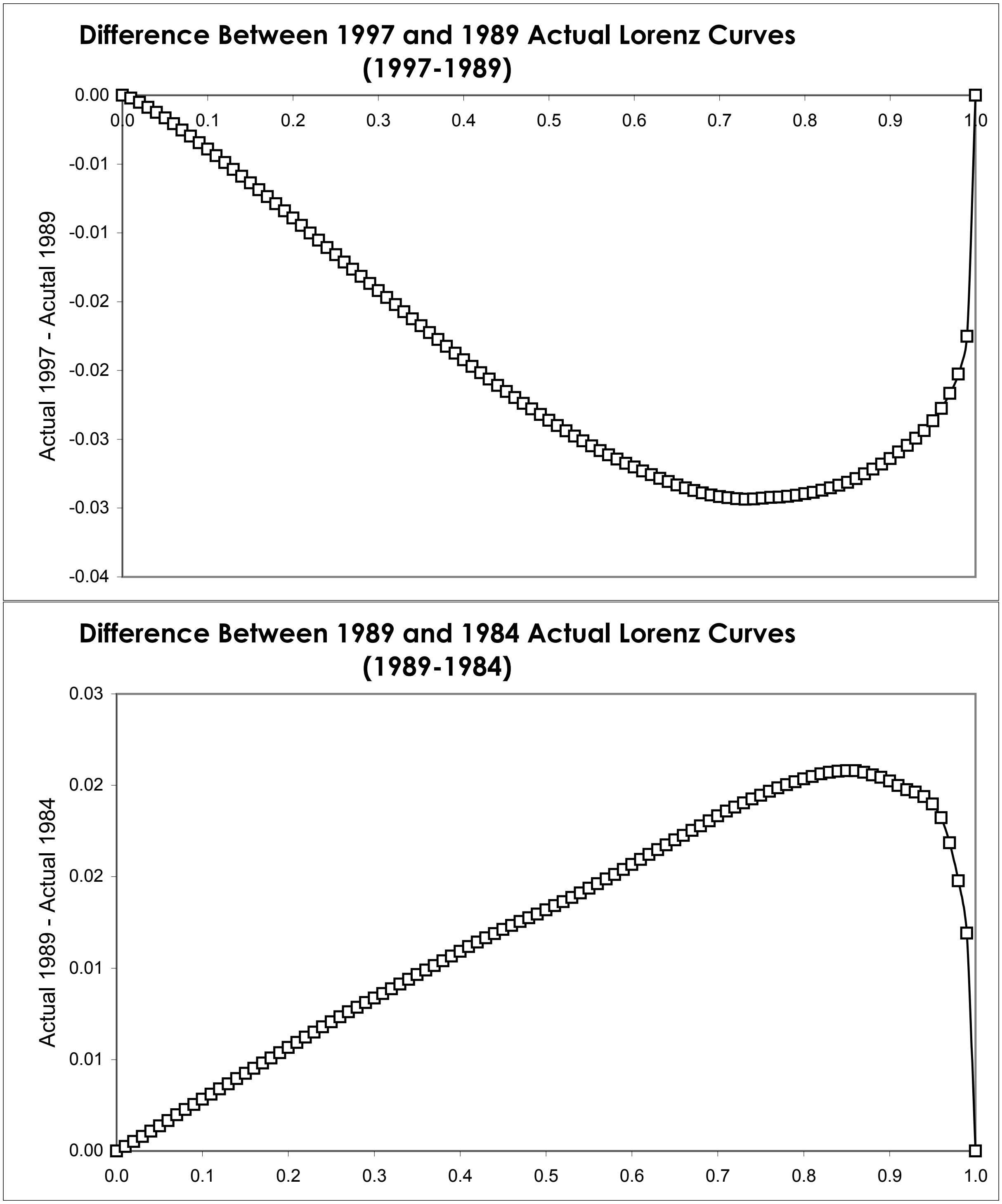
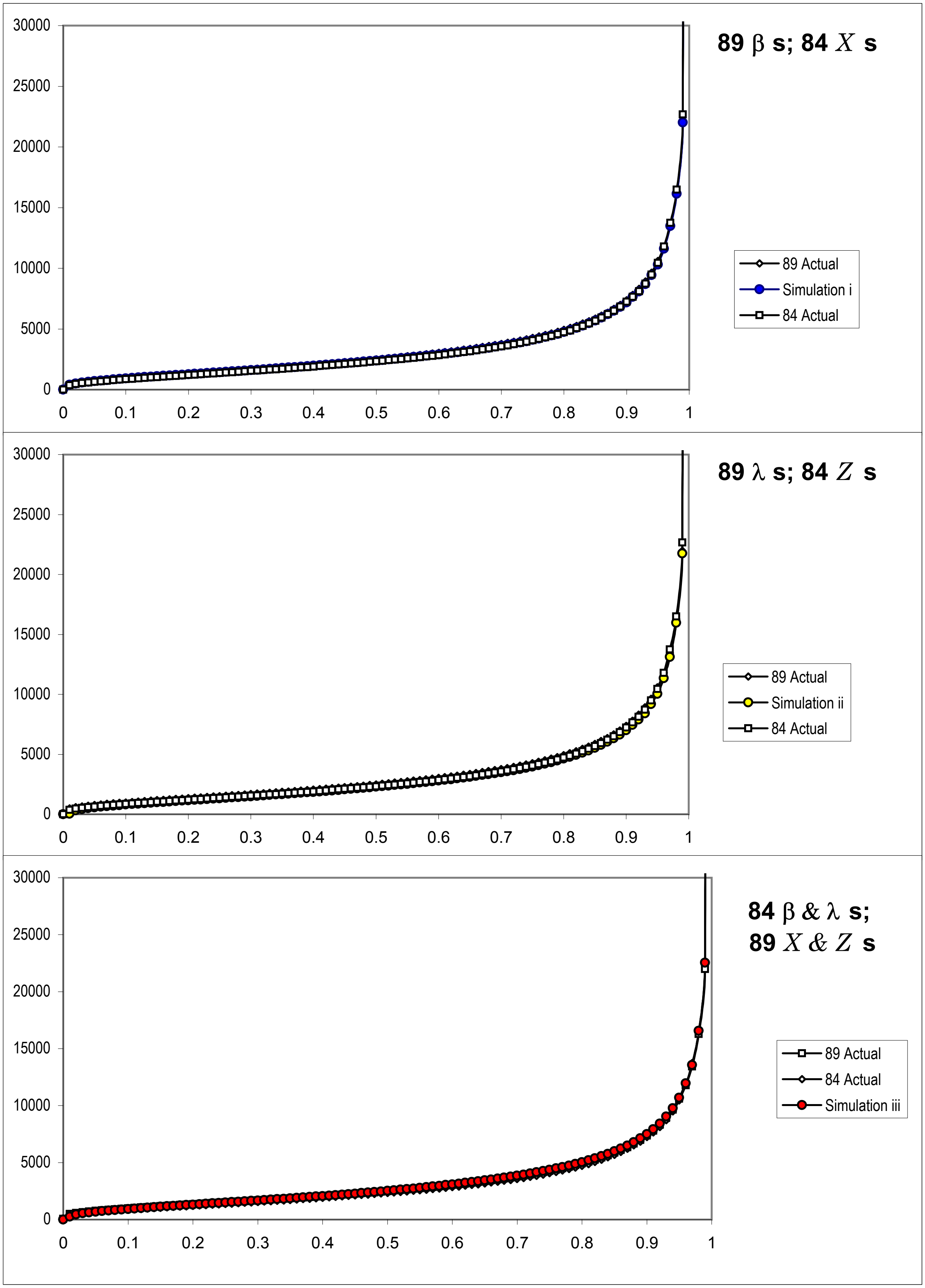
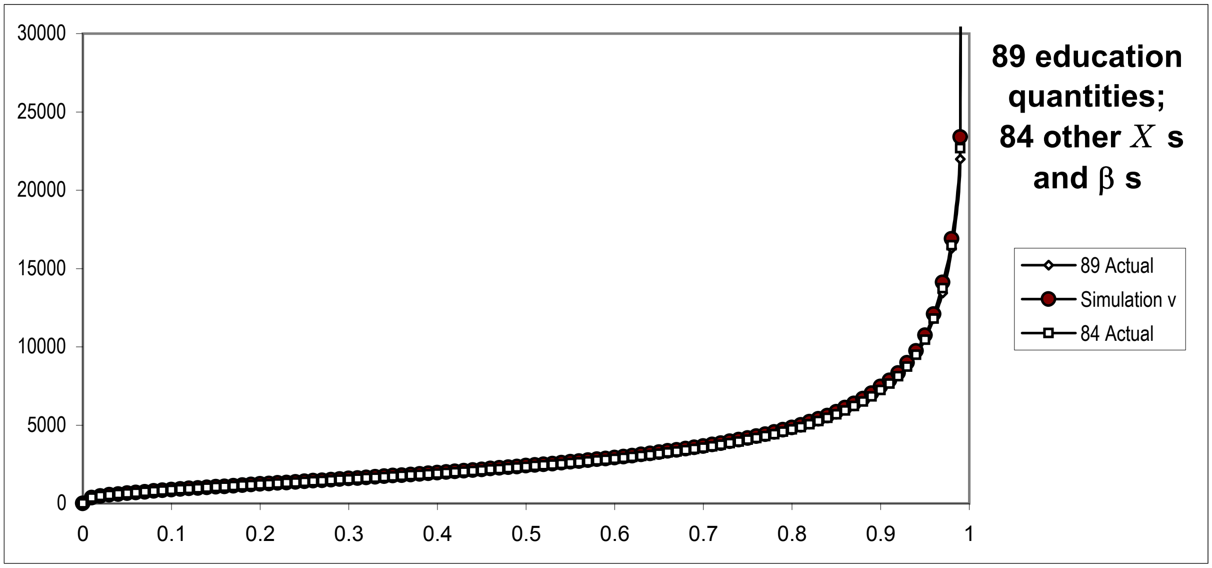


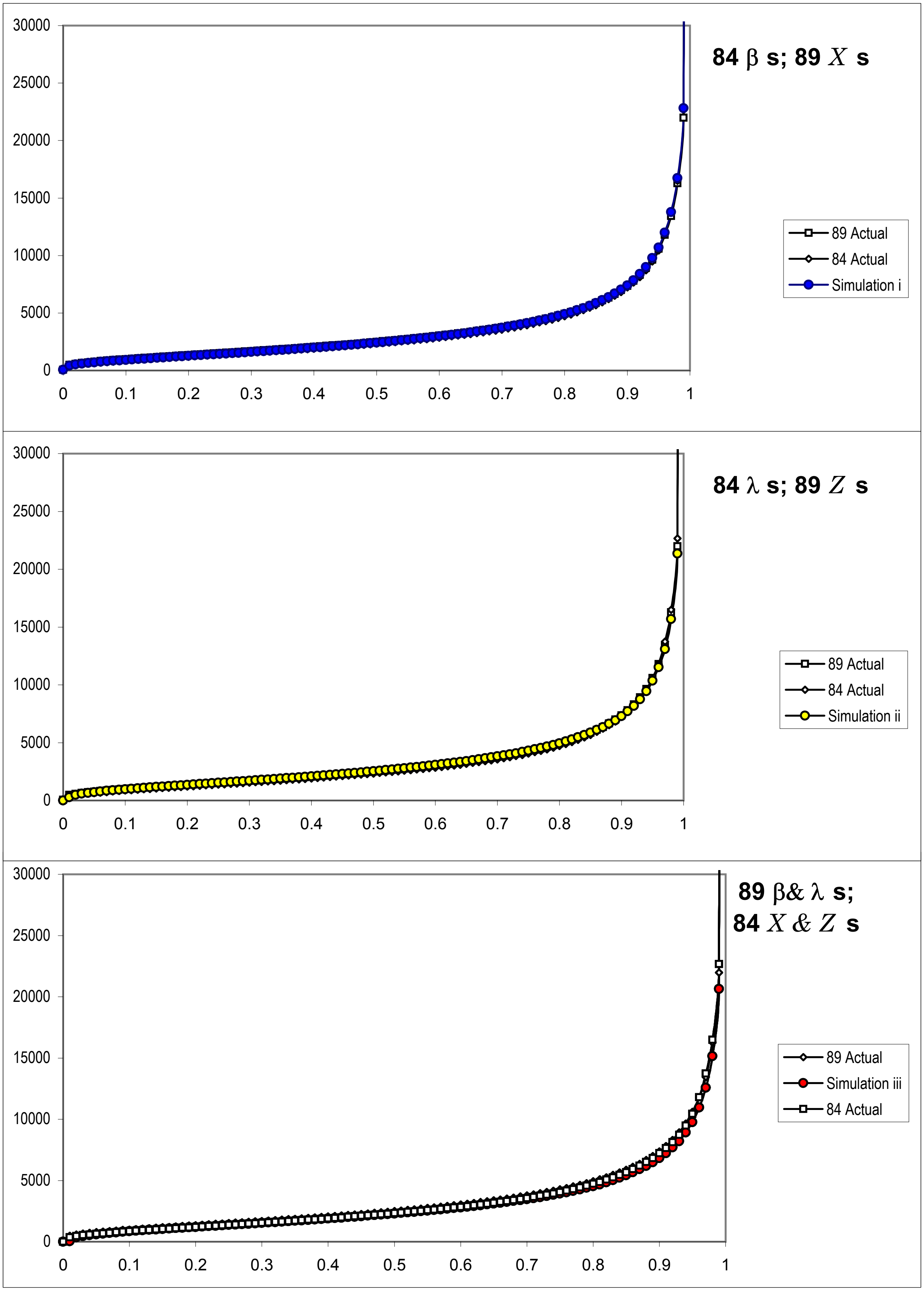
Figure 5A. Household Quantile Curves : 1984 Baseline

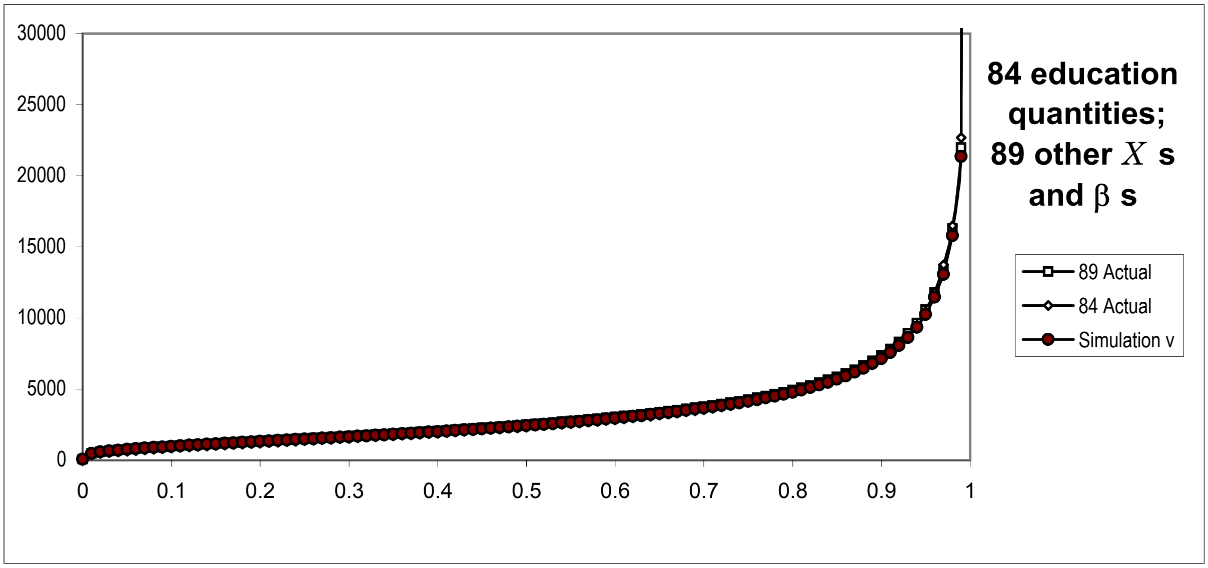




89 education quantities; 84 other X s and β s
89 Actual
Simulation v
84 Actual
30000
25000
20000
15000
10000
5000
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 5B. Household Quantile Curves : 1989 Baseline





30000
25000
20000
15000
10000
5000
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
84 education quantities; 89 other X s and β s
89 Actual
84 Actual
Simulation v

Figure 6A. Quantile Curves: Simulated Values Minus 1984 Actual

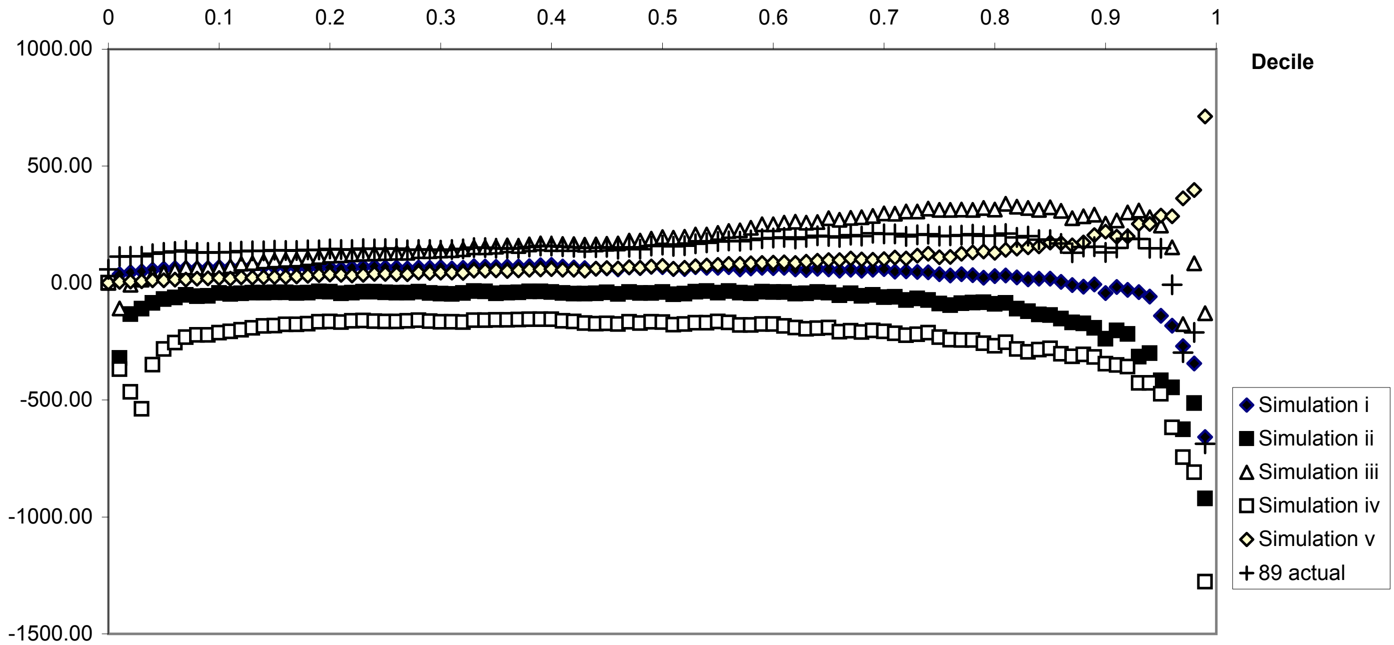


Figure 6B. Quantile Curves: Simulated Values Minus 1989 Actual

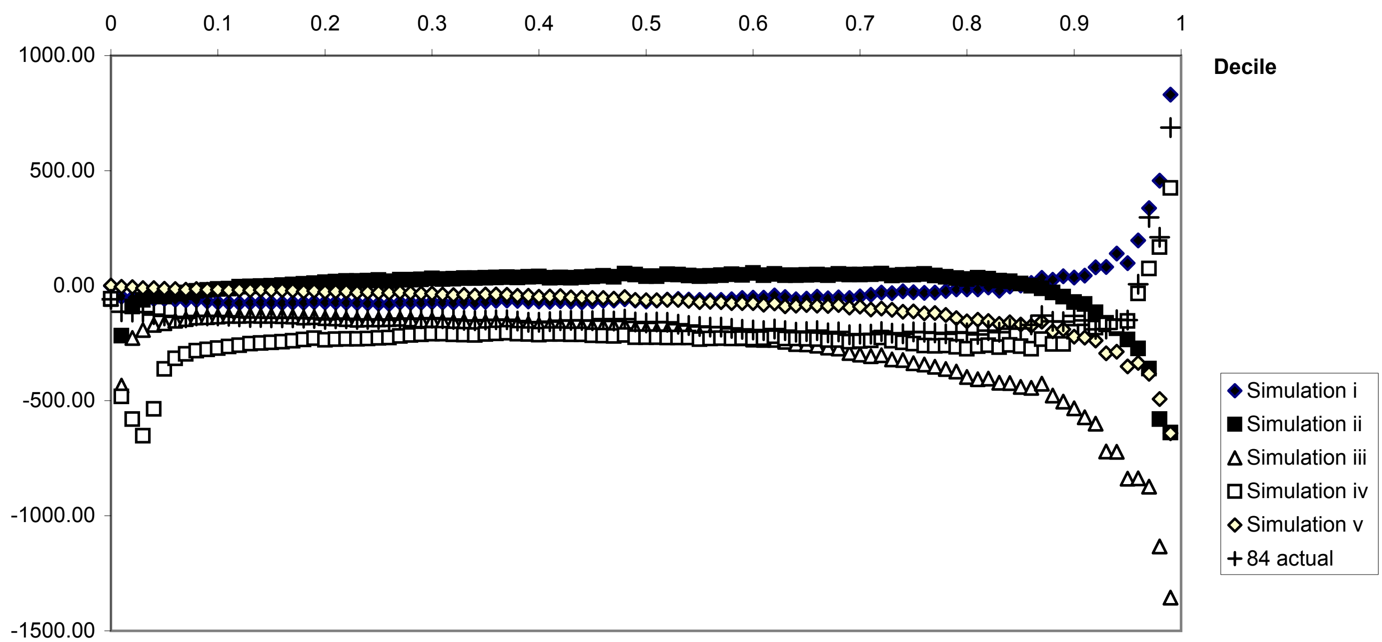


Figure 7A. Lorenz Curves: Simulated Values Minus 1984 Actual

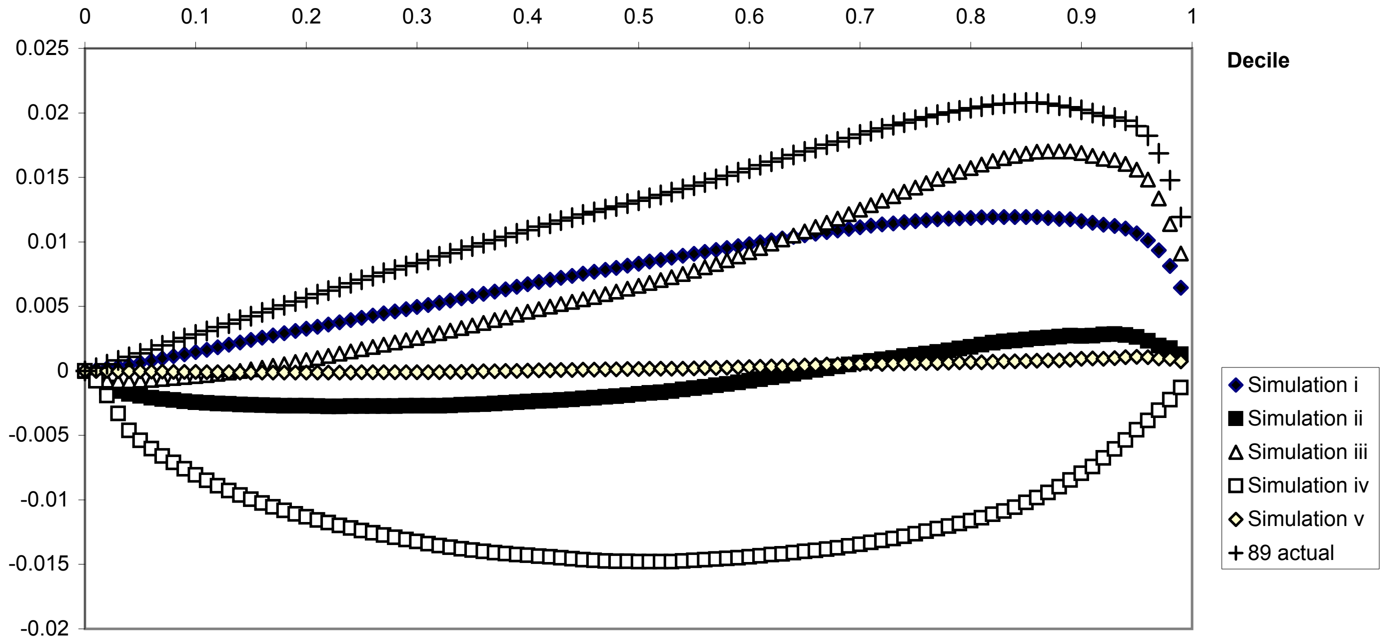


Figure 7B. Lorenz Curves: Simulated Values Minus 1989 Actual

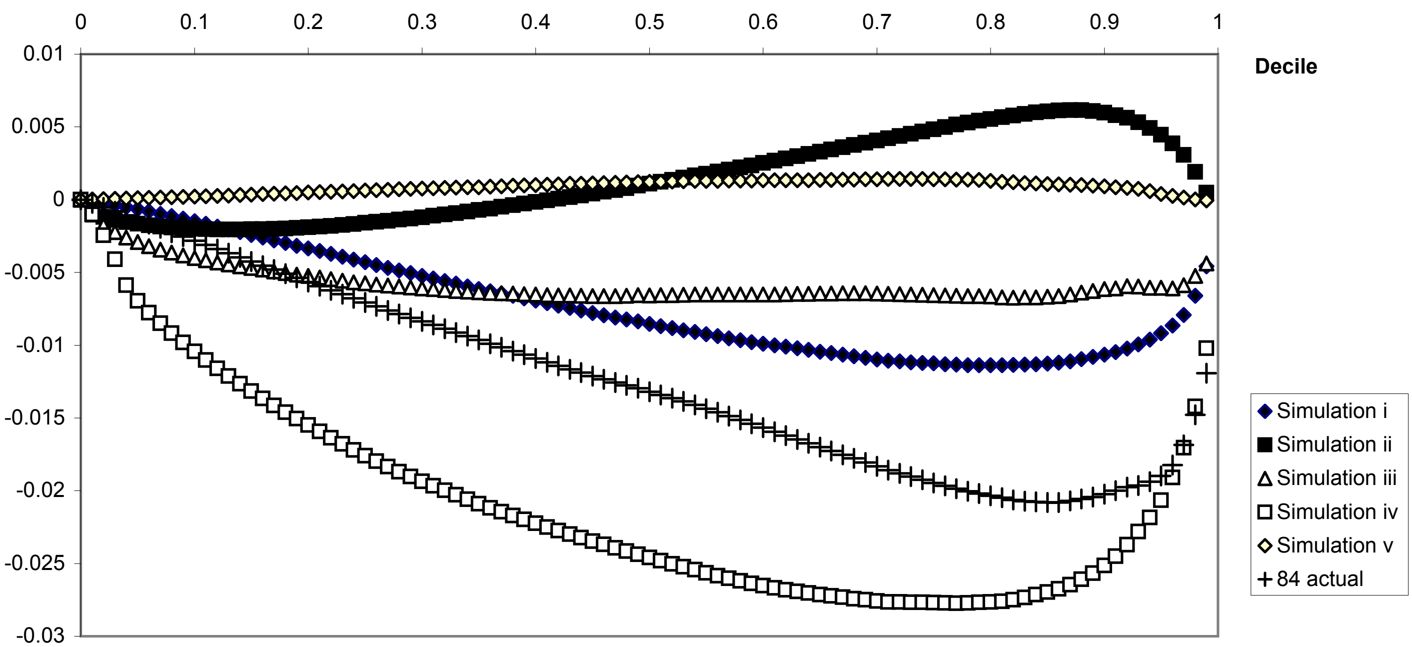
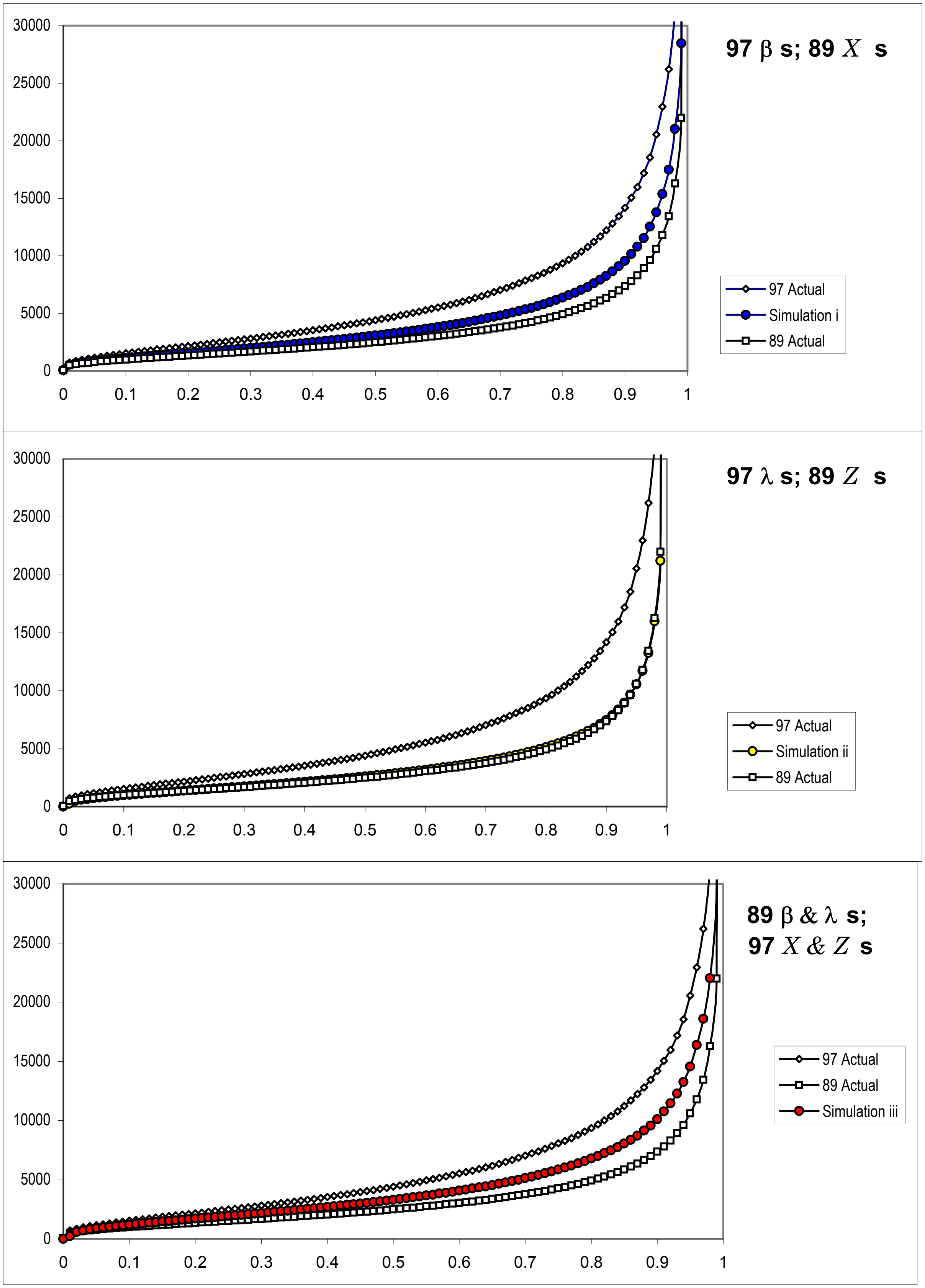
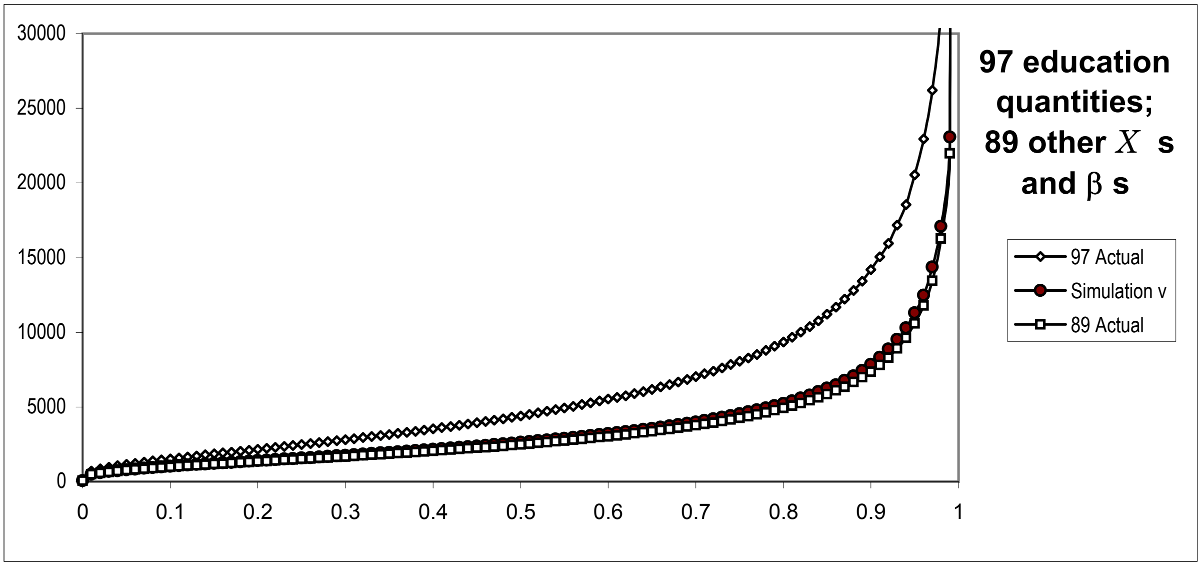


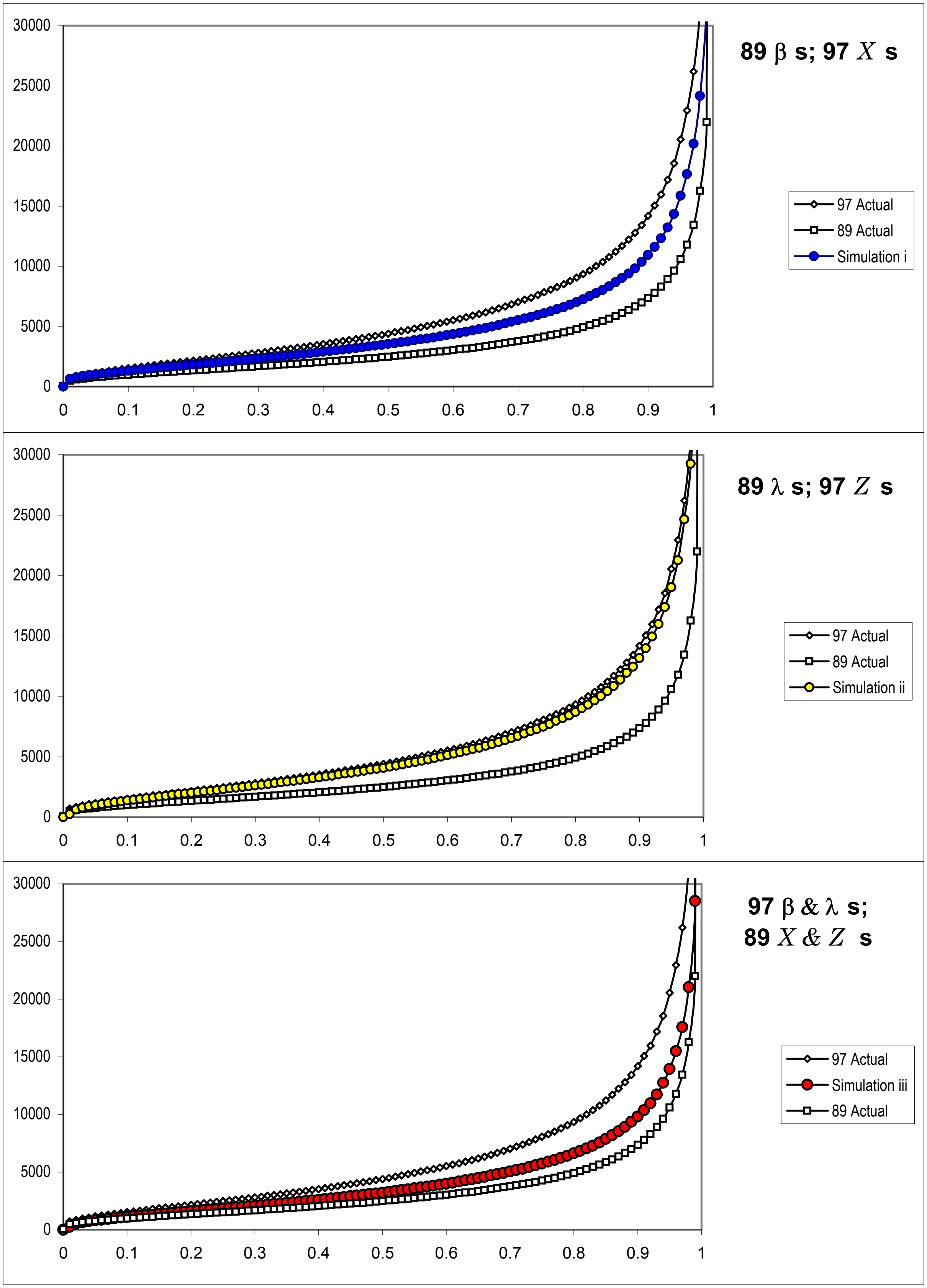
Figure 8A. Household Quantile Curves : 1989 Baseline





97 education quantities; 89 other X s and β s
97 Actual
Simulation v
89 Actual
30000
25000
20000
15000
10000
5000
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 8B. Household Quantile Curves : 1997 Baseline



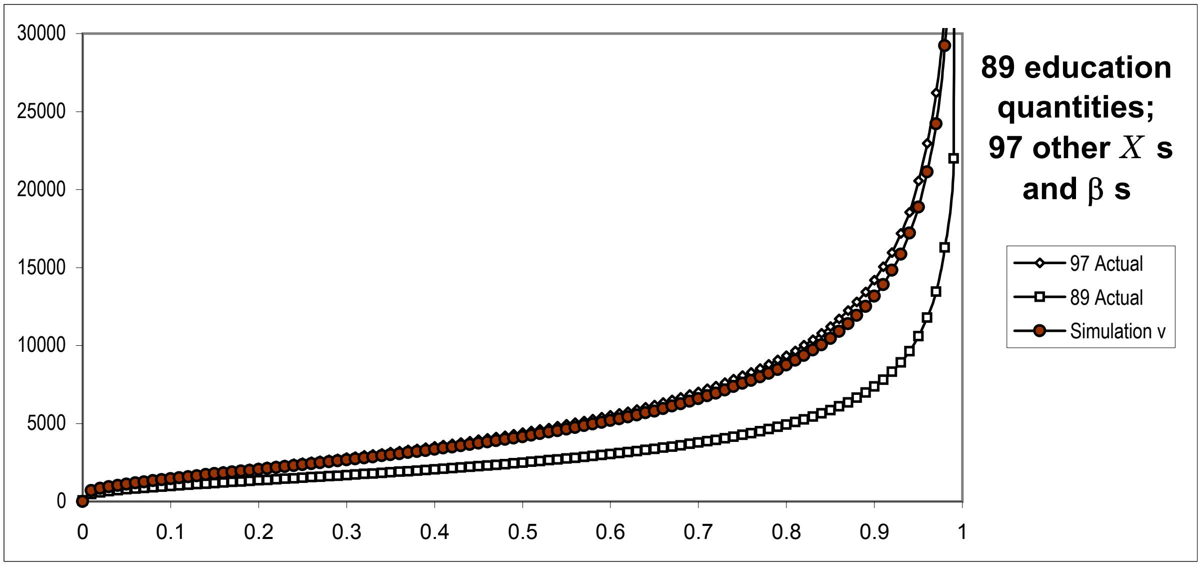

30000
25000
20000
15000
10000
5000
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
89 education quantities; 97 other X s and β s
97 Actual
89 Actual
Simulation v

Figure 9A. Quantile Curves: Simulated Values Minus 1989 Actual

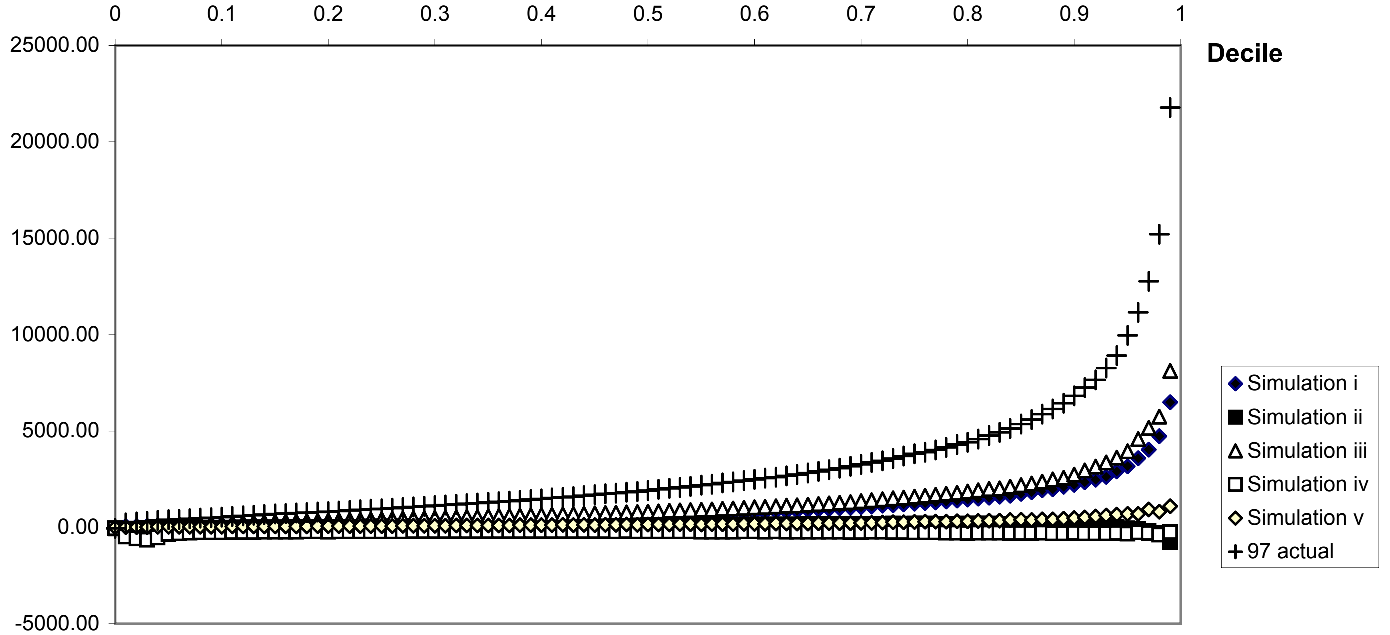


Figure 9B. Quantile Curves: Simulated Values Minus 1997 Actual

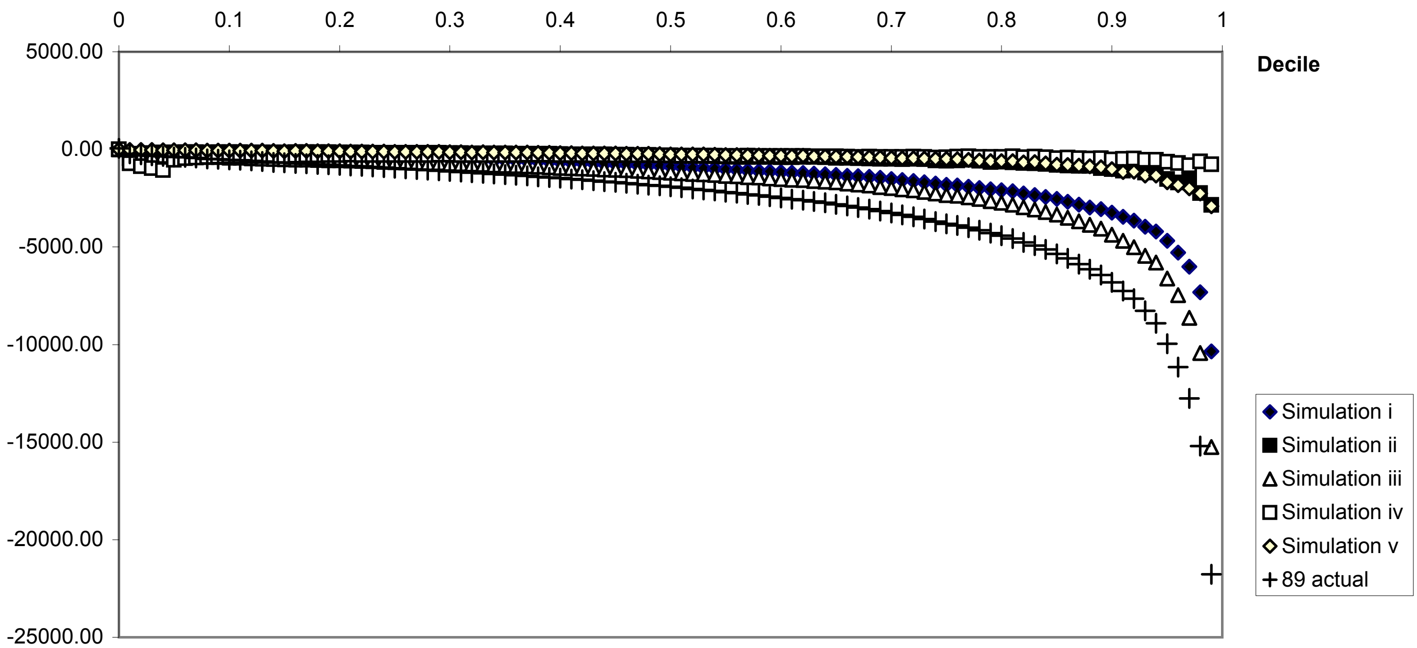


Figure 10A. Lorenz Curves: Simulated Values Minus 1989 Actual

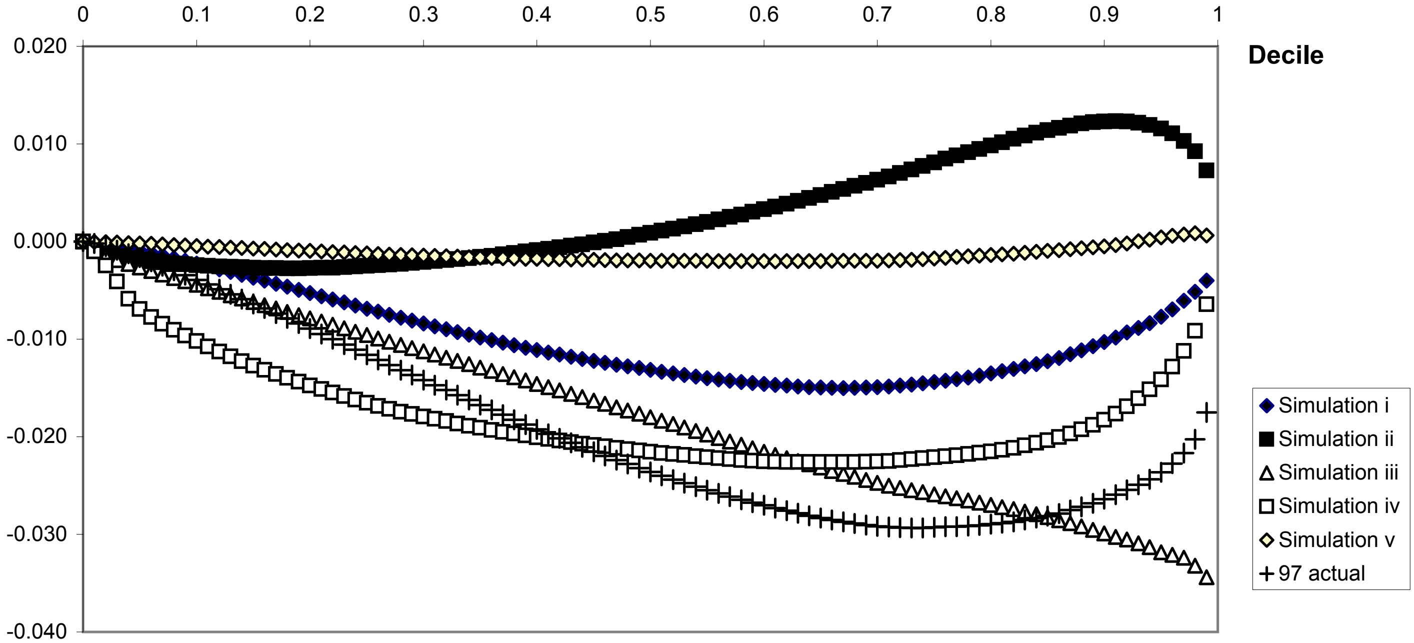


Figure 10B. Lorenz Curves: Simulated Values Minus 1997 Actual