Working Paper Series

# Distributions in motion: Economic growth, inequality, and poverty dynamics

Francisco Ferreira

ECINEQ WP 2010 – 183

ECINEQ 2010-183
October 2010

www.ecineq.org

# Distributions in motion: Economic growth, inequality, and poverty dynamics*

**Francisco Ferreira†**
*The World Bank*

**Abstract**

The joint determination of aggregate economic growth and distributional change has been studied empirically from at least three different perspectives. A macroeconomic approach that relies on cross-country data on poverty, inequality, and growth rates has generated some interesting stylized facts about the correlations between these variables, but has not shed much light on the underlying determinants. "Meso-" and microeconomic approaches have fared somewhat better. The microeconomic approach, in particular, builds on the observation that growth, changes in poverty, and changes in inequality are simply different aggregations of information on the incidence of economic growth along the income distribution. This paper reviews the evolution of attempts to understand the nature of growth incidence curves, from the statistical decompositions associated with generalizations of the Oaxaca-Blinder method, to more recent efforts to generate "economically consistent" counterfactuals, drawing on structural, reduced-form, and computable general equilibrium models.

**Keywords**: Poverty and inequality dynamics; growth incidence curves.
**JEL Classification**: D31, I32.

---

* The author is the World Bank's Deputy Chief Economist for Latin America and the Caribbean. This paper was commissioned as a chapter for the Oxford Handbook of the Economics of Poverty, edited by Philip Jefferson. I am grateful to the editor, to Stefan Dercon, Stephan Klasen and Maria Ana Lugo for comments on and helpful conversations about an earlier version. I am also grateful for comments from seminar participants at the Paris School of Economics, the Katholieke Universiteit Leuven, the Universities of Bari and Göttingen, and the Canazei Winter School on Inequality. I am solely responsible for any remaining errors. The views expressed in this paper are those of the author, and they should not be attributed to the World Bank, its Executive Directors, or the countries they represent.
† **Address of correspondence:** fferreira@worldbank.org.

## 1. Introduction

The dynamic relationship between economic growth, inequality and poverty has always commanded considerable attention among economists. The nexus between growth and development on the one hand, and the distribution "of the whole produce of the earth" on the other has been central to the discipline at least since David Ricardo's (1817) *Principles of Political Economy*.

In the second half of the Twentieth Century, this relationship was approached from a number of different perspectives. Drawing on the seminal work of W. Arthur Lewis (1954) and Simon Kuznets (1955), one strand of the literature has investigated whether income inequality follows a particular dynamic process as economies grow. Lewis saw the process of economic development as largely driven by the transfer of productive resources from backward, low-productivity sectors (such as subsistence agriculture) to high-productivity sectors such as modern plantation agriculture or industry. As labor moved from the poor, but relatively egalitarian backward sectors towards the richer, but more unequal modern sector, Kuznets expected inequality first to rise and then to fall, as the share of workers in the original sector declined. The time-series data on income inequality that was available to Kuznets in the 1950s was extremely restricted. It came from Germany, the United States and the United Kingdom. But in those countries, an "inverted U" pattern for inequality could indeed be identified, thus lending support to the Kuznets hypothesis, which came to dominate a subfield of development economics for many years to come.

In the 1990s, progress in two areas of economic theory – the consequences of asymmetric information and incomplete contracts for how credit and insurance markets function, and endogenous growth theory – combined to generate a fair amount of interest in another approach to the same broad issue. Rather than asking how growth and development affected the distribution of income, a new literature asked whether high initial levels of inequality (or poverty) might be detrimental for future growth. With missing or imperfect credit markets and non-convex production sets, individuals with low wealth might be trapped in low-investment equilibria, leading to divergence between the rich and the poor. Societies starting off with a greater proportion of its population in poverty (or greater inequality given a certain mean) might generate less aggregate investment, or a less advanced occupational structure, and thus lower growth rates.[2] Other stories linking initial income distribution to subsequent growth involved political economy channels. In one strand, higher inequality led the "median voter" to support higher rates of distortionary taxation, leading to lower growth (see, e.g. Alesina and Rodrik, 1994; and Persson and Tabellini, 1994). In another, taxation and public spending were seen instead as potentially efficient mechanisms for correcting market failures, but greater inequality in wealth and political power led to too little public investment (see, e.g., Bénabou, 2000, and Ferreira, 2001).

A considerable expansion in the availability of household survey data, particularly among developing countries, took place in tandem with these theoretical developments. Greater data availability enabled researchers to "test" the predictions of their theories on cross-country data sets that included measures

[2] See, e.g. Galor and Zeira (1993) and Banerjee and Newman (1993).

of economic growth, income inequality and, in some cases, poverty statistics.[3] From the point of view of validating a particular theoretical approach, the resulting empirical literature was not encouraging. As time-series of inequality statistics became available for more countries, evidence of an "inverted U" Kuznets curve appeared to vanish, both when looking within countries over time (see Bruno et al, 1998) and across countries at different levels of GDP per capita (Ravallion and Chen, 1997). Evidence of causality from inequality to subsequent growth fared no better. Some early papers found a negative effect of initial inequality on growth in cross-country regressions (Alesina and Rodrik, 1994; Deininger and Squire, 1998), but later studies cast doubt on the interpretation of that result by finding that, in a panel specification, the effect of lagged inequality on growth appeared to be, if anything, positive (Forbes, 2000).

The literature never converged to a consensus. It was not clear whether the panel specification yielded econometrically superior results (because it eliminated possible endogeneity biases by using lagged variables as instruments), or whether there was something substantively different about recent changes in inequality, vis-à-vis differences in "steady-state" initial levels of inequality. Banerjee and Duflo (2003) argued that cross-country data of the kind used in this literature was not capable of shedding much light on the myriad non-linear processes that might be linking inequality and growth. Voitchovsky (2009) concludes a recent survey by noting that "the inconsistency of reported empirical findings could reflect the gap between the intricacy of the relationship, as expressed in the theoretical literature, and the simple relationships that are commonly estimated." (p.569)

To make matters worse, it seemed for a moment - in the early 2000s - as though even the relationship between growth and poverty – which had always been taken more or less for granted, at least among economists – was being questioned. Using quotes from policy circles, Dollar and Kraay (2002) noted that although "The world economy grew well during the 1990s […], there is intense debate over the extent to which the poor benefit from this growth." (p.195). While some appeared convinced that growth is "a rising tide that lifts all boats", others insisted that, in the 1990s, "the rich were getting richer, and the poor were getting poorer".

In the early 2000s, then, one might have been forgiven for concluding that economists had not made much progress in understanding the relationship(s) between growth and distribution since the days of David Ricardo. Our best shot at a general theory of how structural change and economic development affected the distribution of income – the Kuznets hypothesis – appeared to be rejected by more plentiful data.[4] Similar data appeared to generate contradictory or inconclusive results for tests of alternative theories about the reverse channel of causation: from inequality to growth. And now there was even a

[3] The history of this growth in data availability - and a discussion of how comparable the different data sets are - are topics of considerable interest in and of themselves, but they fall beyond the scope of this paper.

[4] To be fair, there was an alternative view of long-term inequality dynamics, associated with Jan Tinbergen, which proved rather more robust. Tinbergen (1975) saw changes in earnings inequality as largely reflecting the evolution of a "race" between technological progress – which he saw as raising the demand for skills – and the expansion of formal education – which raised the supply of skills. Strangely, this view appears to have been more popular with labor economists than with development economists.

debate on whether growth reduces poverty… Do economists actually know anything about what François Bourguignon (2004) christened "the poverty-growth-inequality triangle"?

This paper reviews some of the recent empirical literature on this triangular relationship, and makes a two-part argument. The first part is that, even if it has not yielded a clear confirmation of any particular grand theory of how growth and distribution are connected, this literature has taught us a few things. I summarize these lessons in terms of three basic "stylized facts", which do seem to be robustly supported by the data.

The second part of the argument is that, while it has been useful in shedding light on those stylized facts, the essentially "macroeconomic" approach followed by this literature is not particularly promising if the objective is to gain a real understanding of the joint determination of growth, inequality and poverty. Average income, poverty and inequality are all aggregate concepts: averages of incomes or income gaps, measured in different ways, and with different weights along the distribution. Their evolution over time – economic growth, changes in poverty and changes in inequality – are all jointly determined by the individual income dynamics in that distribution. Because the three vertices of the triangle are all defined over the distribution of income, they are mechanically related by a statistical identity. But estimating relationships between the three objects – or between two at a time – is unlikely to tell us much about the underlying processes, since it ignores that all three are driven by the interaction of individual behaviors at the microeconomic level.

Of course, like much else in macroeconomics, the estimated relationships between poverty, inequality and growth reveal interesting stylized facts – regularities in the data that are informative of deeper forces. The paper begins by reviewing the useful stylized facts that have arisen from the recent "macroeconomic" or cross-country literature (in Section 2). Section 3 briefly turns to what I term a "mesoeconomic" approach, which uses somewhat more disaggregated data – along the spatial and sectoral dimensions – within individual countries, to investigate the same "triangular" relationship.

In Section 4, I focus on an alternative microeconomic approach, which delves beneath the aggregate measures of poverty, inequality and growth, and considers changes in the entire underlying distribution of incomes. First, each of the three aggregate variables in the poverty-growth-inequality triangle is shown to be expressible as a different aggregation of the information contained in a full description of growth along the income distribution – the *growth incidence curve* (GIC).[5] This allows us to argue that to understand distributional dynamics, one needs to understand the determinants of the GIC. This has generally been attempted by means of counterfactual decompositions of distributional change, which range from the purely statistical, to those with greater economic content. Although that literature has seldom been related to the macro and meso-economic approaches discussed above, and often uses a different language, the paper argues that they are closely related in their essence. A fifth section offers a few concluding remarks, and suggests possible avenues for future research.

One feature that all three approaches reviewed in this paper have in common is their respect for the anonymity (or 'symmetry') axiom of inequality and poverty measurement. This axiom requires that

[5] The GIC was first defined in Ravallion and Chen (2003).

poverty and inequality measures (and therefore, evidently, their differences) be invariant to permutations of the income vector. It implies that, when comparing distributions and their summary statistics, the analyst must ignore the individual dynamics along the distribution: two identical distribution functions $F_t(y) = F_s(y)$, for $t \neq s$ must yield identical inequality and poverty statistics (for a given poverty line), regardless of changes in the identities of the occupants of each percentile *p*.

This is a standard axiom in the analysis of poverty and inequality, and it helps to considerably narrow the scope of the paper. In particular, it means that we do not review work on economic mobility, or on individual or household income dynamics. As panel data has become available in developing countries, a number of recent studies have sought to shed light on the determinants of individual income or consumption dynamics (e.g. Dercon, 2004; Lybbert et al., 2004). These studies have often focused on the role of different asset classes (land, livestock, physical infrastructure) in promoting income growth; on the effects and persistence of various uninsured shocks; and on the evidence for or against poverty traps (e.g. Jalan and Ravallion, 2002). This is an important literature. Because they follow the same individuals over time, such panel-based studies are generally better at identifying the causal effects *of* exogenous shocks (possibly including policy shocks) *on* inequality and poverty than most of the work we review in this paper. At the same time, they typically rely on small samples, which are not representative of entire countries. They are not intended to, and generally cannot, assess the joint determination of aggregate economic growth, and societal poverty and inequality rates, which is our focus here. They are omitted from this review on that basis, but their superiority in terms of causal identification means that there may be much to be gained from deeper thinking about how these different approaches interrelate. We briefly return to this point in the concluding section.

## 2. The macroeconomic approach

What I call the macroeconomic approach to the relationships between poverty, inequality and growth consists of cross-country (panel or single cross-section) analysis of how those three variables or, more often, two at a time, are related to one another. In order to understand this approach, it is useful to start from a very simple accounting identity, which relates the incidence of poverty – known as the headcount index, *H* – both to the mean and to the 'shape' of the distribution of income.

Denote the Lorenz curve for a given distribution as *L(p, π),* where $p = F(y)$, and $\pi$ is a vector of parameters that fully determines the functional form of the Lorenz curve.[6] It is well-known that the derivative of the Lorenz curve at percentile *p* is given by the ratio of the income at that percentile (the quantile function $y = F^{-1}(p)$), to the overall distribution mean:

$$L_p(p, \pi) = \frac{y(p)}{\mu} \tag{1}$$

The headcount is simply the share of the population with incomes no higher than the poverty line, $H = F(z)$. Therefore, evaluating equation (1) at *p = H*, and solving for *H* yields:

[6] *F(y)* is the cumulative distribution of income: it gives the measure of the population with incomes lower than *y*, for every *y*. The Lorenz curve is a related, but different concept. It gives the cumulative share of income that accrues to the poorest *p*% of the population, for every *p*.

$$H = L_p^{-1}(z/\mu\,,\pi) \qquad (2)$$

Equation (2) is the fundamental identity in the "poverty- growth-inequality triangle". For a given poverty line *z*, it relates the incidence of poverty to the mean of the distribution (inversely), and to the parameters of the Lorenz curve. The Lorenz curve is deeply associated with the concept of inequality.[7] Given a constant real poverty line *z*, the poverty headcount is therefore fully determined by the "shape" of the distribution (given by the Lorenz curve) and its central location (given by the mean). Analogous, if somewhat more complicated, identities relate other poverty measures to the distribution's mean and Lorenz curve.[8]

It is straightforward to move from the fundamental identity of the triangle in levels, to its dynamic counterpart. One simply differentiates equation (2) with respect to time. Denoting the time derivative of variable *x* as *dx*, one obtains:

$$\frac{dH}{H} = -\frac{L_{pp}^{-1} z}{L_p^{-1} \mu}\frac{d\mu}{\mu} + \frac{L_{p\pi}^{-1}}{L_p^{-1}} d\pi \qquad (3)$$

Equation (3) relates changes in poverty to the rate of economic growth ($d\mu/\mu$), and to changes in "inequality" or, more precisely, in the Lorenz curve ($d\pi$).[9] It tells us that any change in poverty can be mechanically decomposed into a *growth component* (the first term on the RHS), and a *distribution component* (the second term). Since the Lorenz curve is always increasing and convex by construction (so that $L_p > 0, L_{pp} > 0$ ), equation (3) also tells us that the partial elasticity of poverty with respect to growth ($-\frac{L_{pp}^{-1} z}{L_p^{-1} \mu}$), is always negative: if there is no change in the Lorenz curve, an increase in mean incomes must be accompanied by a decline in poverty. In other words: growth must lead to poverty reduction, unless it is accompanied by substantial distributional change (which may be fairly, albeit somewhat imprecisely, described as "increasing inequality").

The partial inequality elasticity ($\frac{L_{p\pi}^{-1}}{L_p^{-1}}$), on the other hand, cannot be signed a priori: changes in the distribution may increase or reduce poverty, depending on how the Lorenz curve shifts, and on the relative position of the poverty line.

---

[7] The Atkinson (1970) theorem establishes that, given two distributions A and B such that $L^A(p) \leq L^B(p), \forall p$, with the inequality holding strictly at least at one point, then any inequality index that satisfies the Pigou-Dalton transfer principle must take a higher value in distribution A than in B. The (strong) Pigou-Dalton transfer principle states that an inequality measure must rise in response to a transfer of income from a poorer to a richer person. It is regarded as a key axiom for inequality indices.

[8] The most widely-used family of poverty measures is the Foster-Greer-Thorbecke class, $P_\alpha = \int_{-\infty}^{z} \left(\frac{z-y}{z}\right)^\alpha f(y)dy$. The headcount is the FGT measure when $\alpha = 0$. Other commonly-used members of the class take values of $\alpha = 1{,}2$.

[9] We are deliberately glossing over the empirically important question of whether one should measure economic growth as the growth in mean household incomes (from a household survey), or by growth in some concept obtained from national accounts data, such as GDP per capita. While the two should not differ systematically over long periods of time, they have often been found to diverge considerably in the short and medium runs.

This decomposition of changes in poverty into growth and distribution components was first applied empirically by Datt and Ravallion (1992) and Kakwani (1993). Using discrete time intervals, Datt and Ravallion noted that equation (3) could be approximated by writing:

$$P_{t+n} - P_t = \left[P\left(\frac{z}{\mu_{t+n}}, L_t\right) - P\left(\frac{z}{\mu_t}, L_t\right)\right] + \left[P\left(\frac{z}{\mu_t}, L_{t+n}\right) - P\left(\frac{z}{\mu_t}, L_t\right)\right] + R(t, t+n) \qquad (4)$$

Once again, the first term on the RHS of equation (4) corresponds to the growth component, and holds the Lorenz curve constant, while the second term measures the redistribution component, while holding mean income constant. The third term is a residual that arises from the path-dependence of the discrete decomposition, and vanishes if the decomposition is averaged across the two possible orderings. Datt and Ravallion (1992) applied this decomposition to India (1977-1988) and Brazil (1981-1988). Over these periods, poverty fell in India, and stayed broadly constant in Brazil. In rural India, where the decline was of 16%, the growth component accounted for some ten percentage points, and the redistribution component for about six points (the residual was negligible). In Brazil, the very limited growth that did take place during the 1980s contributed towards a reduction in poverty, but this effect was entirely offset by increasing inequality.

This exercise has since been repeated for many countries. While results vary considerably – reflecting cross-country differences both in growth performance and in inequality trends – one may summarize them as follows: when economic growth is high, the negative growth component tends to dominate the decomposition, and leads to falling poverty (although rising inequality can slow this down).[10] When growth is small, as in the Brazilian example above, the distribution component can dominate, and changes in inequality may make the difference between rising and falling poverty.

Besides computing the poverty decomposition for individual countries one at a time, as in equation (4), one could try to estimate equation (3) on a cross-section of countries. Although it holds as an identity for each particular country at each point in time, the partial elasticity terms ( $-\frac{L_{pp}^{-1} z}{L_p^{-1} \mu}$ and $\frac{L_{p\pi}^{-1}}{L_p^{-1}}$ ) obviously differ across countries and periods, as the distributions themselves change. For that reason; because different inequality indices summarize the information in the Lorenz curve somewhat differently; and because of measurement error, it would be possible to estimate equation (3) econometrically, say by regressing changes in poverty on growth and changes in inequality (*I*): $\frac{dH}{H} = \beta \frac{d\mu}{\mu} + \gamma \frac{dI}{I} + \varepsilon$. Such a regression would not have an $R^2$ of one, and estimates of β and $\gamma$ would represent sample estimates of the average values for those elasticities.

This is not, however, what most papers in the macroeconomic approach have done. To avoid the impression of estimating an identity, most studies have omitted either growth or changes in inequality from the RHS of the regression. The vast majority estimate a regression of changes in poverty on economic growth, with or without controls (*X*):

[10] See Ahuja et al (1997) for a set of the decompositions of this sort applied to high-growth East Asian economies prior to the Asian crisis of 1997-98.

$$\frac{dH}{H} = \alpha + \beta \frac{d\mu}{\mu} + X\gamma + \varepsilon \qquad (5)$$

One of the first papers to estimate equation (5) on cross-country data was Ravallion and Chen (1997). Using a compilation of household survey datasets available at the World Bank, these authors assembled a data set with household-survey-based poverty measures and growth rates for 67 countries over the period from 1981 to 1994. With no controls other than a country-specific time trend, and using an international poverty line of US$1 per day at 1985 PPP exchange rates, they found a central estimate for the growth elasticity of poverty of β = -3.1. Excluding the then transitioning economies of Eastern Europe and Central Asia, the regression line went through the origin: the average rate of change in poverty at zero growth in their sample was zero. This reflected the fact that they could find no significant cross-country correlation between changes in inequality and economic growth in their sample.

Although the numbers have changed as the sample expanded, and different specifications have been tried, those two basic results have been confirmed by the more recent literature. Dollar and Kraay (2002) replaced the headcount index with an alternative, slightly unconventional (inverse) measure of poverty: the average income of the poorest 20% of the population. Using a sample of 92 countries over four decades, they could not reject a positive elasticity of 1.0: growth in the incomes of the bottom quintile was, on average, equiproportional to growth in mean incomes. This result, which survived the introduction of a number of controls, once again implied that changes in inequality were not systematically correlated with economic growth. Ravallion (2007) find a slightly lower growth elasticity for the poverty headcount than the earlier Ravallion-Chen study, and a small negative – but statistically insignificant - correlation between changes in the Gini coefficient and proportional changes in mean income, in a sample of 80 countries. Kraay (2006) uses variance decompositions to understand the main drivers of poverty changes in developing countries in the 1980s and 1990s. He finds that changes in distribution account for 30% of the variance in changes in the headcount when all spells are considered, versus 70% for growth in average incomes. When the sample is restricted to long spells, those proportions change to 3% and 97% respectively.

In sum, the macroeconomic approach to the poverty-growth -inequality triangle yields two basic stylized facts:

*(i) Economic growth and changes in inequality are uncorrelated, at least in the sample of countries available to researchers for the two decades from 1980 to 2000.*

*(ii) Poverty generally declines as the economy grows. The longer the growth spells under consideration, the larger the share of the variance in poverty that is accounted for by the growth component.*

Given equation (3), (ii) is clearly implied by (i). The empirical content of the Ravallion and Chen (1997) and Dollar and Kraay (2002) results is that changes in inequality appear to be uncorrelated with growth, at least for this sample – a result that is related to the empirical rejection of the Kuznets hypothesis, documented by Bruno et al. (1998). Equation (3) tells us nothing about that relationship. It decomposes changes in poverty into growth and redistribution components, but says nothing about how inequality and growth should be related. With the apparent demise of the Kuznets curve, and no conclusive result about the possible effects of inequality on subsequent growth, there is no convincing theoretical reason to expect a systematic relationship. And indeed, there does not appear to be one in the data.

This is not to say, however, that inequality does not matter for poverty reduction. In fact, it clearly matters in two ways. The first is that, even if inequality is uncorrelated with economic growth *on average*, there is substantial dispersion around that average. Inequality did rise in many countries, and did fall in many others. In countries with rising inequality, the effect of growth on poverty was dampened or even reversed, while in those where inequality fell, the decline in poverty for a given growth rate was greater. The ensuing heterogeneity in poverty dynamics at any given growth rate was substantial. As an illustration, Ravallion (2007) estimates that the 95% confidence interval around the regression coefficient of growth in (headcount) poverty on growth in survey income or consumption mean implies that a annual growth rate of 2% (roughly the average for developing countries in the 1980s and 1990s) is consistent with poverty reductions ranging from 1% to 7%.

The second way in which inequality matters for poverty reduction is through the dependence of the growth elasticity of poverty reduction on *initial* inequality. Recall that the partial growth elasticity is given by $-\frac{L_{pp}^{-1}z}{L_p^{-1}\mu}$. Given the convexity of the Lorenz curve, it can be shown that if a distribution A is robustly more unequal than a distribution B (in the sense that $L^A(p) \leq L^B(p)$ in the relevant range, then for poverty rates that are not "unusually high", $L_p^{-1A} > L_p^{-1B}$ and $L_{pp}^{-1A} < L_{pp}^{-1B}$. Obviously, this implies that, if the real poverty line and the mean are the same in distributions A and B, then the partial growth elasticity is higher (i.e. its absolute value is lower) in the more unequal distribution. Or, in other words: economic growth always contributes towards poverty reduction but, even if there is no *change* in inequality, its "poverty-reducing power" is less in countries that are *initially* more unequal.

Figure 1, drawn from World Bank (2005), illustrates this point by plotting the partial elasticity of the poverty headcount against the initial Gini coefficient for a sample of 65 countries during 1981-2005, using a US$ 1-a-day poverty line. The result described above is clearly visible in the data: the growth elasticity is strongest among low-inequality countries (with a value of approximately -4.0 for countries with Ginis in the mid-20s) and weakest among high-inequality countries (approaching -1.0 for countries with a Gini index around 60/100). World Bank (2005) shows that the relationship is robust to changes in both the poverty index (e.g. to FGT(2)) and poverty line (e.g. to US$2-per-day). Interestingly, the partial elasticity result carries though to the total empirical elasticity, which also slopes upwards when plotted against initial inequality. See World Bank (2005) and Ravallion (2007). The robustness of this regularity allows us to list it as the third stylized fact that arises from the recent macroeconomic literature on growth, inequality and poverty dynamics:

> *(iii) The (absolute value of the) growth elasticity of poverty reduction falls with inequality. The larger the initial inequality in a given country, the higher the growth rate needed to achieve the same amount of poverty reduction.*

## 3. The "mesoeconomic" approach

Beyond those three stylized facts about country-level correlations, however, the macroeconomic approach to the poverty-growth -inequality triangle has yielded relatively little. During the period of study, inequality was uncorrelated with growth, so the cross-country evidence is limited to pointing out that growth leads to poverty reduction, and that it does so more effectively if initial inequality is lower. But it tells us very little about which country characteristics or policy choices might contribute to faster poverty reduction, through reductions in inequality and a correspondingly higher (total) growth elasticity of poverty reduction. If there were no other way to investigate the relationship between growth and distributional dynamics, the conclusion would presumably have to be: 'since we do not know much about the determinants of distributional change, and since on average growth lowers poverty, let's focus on whatever policies (we believe) yield the highest average growth rates'.[11]

Fortunately, there are other data sets that permit an investigation of this relationship. The closest in nature to the cross-country approach is to look at poverty dynamics within large countries that combine three data characteristics: (i) a sufficiently long time-series of repeated cross-sectional household surveys; (ii) national accounts information on economic growth that is disaggregated both by sectors (e.g. agriculture, industry, and services) and sub-nationally (e.g. by province or state); and (iii) sufficient information on other time-varying poverty determinants (such as inflation rates, patterns of public spending, and the like). With such data, one could adapt equation (5) to a panel of sub-national units, and write:

$$\frac{dH}{H}_{it} = \alpha_i + \sum_{J=P,S,T} \beta_i^J s_{i,t-1}^J \left(\frac{d\mu}{\mu}\right)_{it}^J + X\gamma + \varepsilon \qquad (6)$$

Here, a subscript *i* denotes a sub-national unit (call it a "state"), while *t* denotes a time period. *J* denotes a broad sector of economic activity, such as primary (P), secondary (S) or tertiary (T); and $s_{it}^J$ denotes the share of the sector *J* in the total output of state *i* at time *t*. Naturally, equation (6) is estimated using discrete data on state-level poverty and sector output growth, so it is more commonly written in the form:

$$\Delta lnH_{it} = \alpha_i + \sum_{J=P,S,T} \beta_i^J s_{i,t-1}^J \Delta lnY_{it}^J + \Delta lnX_{t-1}\gamma + \varepsilon \qquad (6')$$

This general specification is designed to shed light on a number of possible "economic determinants" of the distribution component in equation (3). Differences in $\beta_i^J$ across sectors *J* can be informative of whether the pattern of growth matters for poverty reduction. In almost every case where it was tried, researchers have comfortably rejected the null hypothesis that the elasticities are the same across

[11] This is reminiscent of Dollar and Kraay's (2002) conclusion: "In short, existing cross-country evidence – including our own – provides disappointingly little guidance as to what mix of growth-oriented policies might especially benefit the poorest in society. But our evidence does strongly suggest that economic growth and the policies and institutions that support it *on average* benefit the poorest in society as much as anyone else." (p.219, emphasis added).

sectors. For China, Ravallion and Chen (2007) found that the (absolute) elasticity of poverty with respect to agricultural growth was greater than that for any other sector. In India and Brazil, on the other hand, growth in the tertiary (services) sector was more “pro-poor” than growth in either industry or agriculture (Ravallion and Datt, 2002; Ferreira et al. 2010). $\beta_i^J$ can also vary across states (or provinces), and those differences may be driven by various “initial conditions” at the state level. Ravallion and Datt (2002) found that the elasticity of poverty with respect to non-agricultural growth did differ across the states of India, and was higher in states with greater initial farm productivity and literacy rates. Ferreira et al. (2010) found similar results for Brazil, where industrial growth had a greater poverty-reducing effect in states with higher initial levels of human capital (proxied by lower infant mortality rates and higher average years of schooling in the adult population) and of worker empowerment (as measured by initial unionization rates). They also found that the sectoral growth elasticities varied over time, at least in part in response to changes in the policy regime, such as trade liberalization and price stabilization measures.

The time varying controls ($\Delta lnX_{t-1}$), which are usually introduced in lagged differences to alleviate endogeneity concerns, can also be informative. The national inflation rate is generally included as one such control, and it is often significant. Controlling for the composition of growth, state-level fixed effects, and a number of other covariates, lower inflation reduced poverty in both India and Brazil. State-level spending, on the other hand, was effective in reducing poverty in India, but not in Brazil. The growth of *federal* spending on social assistance and social insurance over the study period, however, had a pronounced poverty-reducing effect in Brazil. In fact, one way to interpret the Ferreira et al. (2010) decomposition of poverty changes between 1985 and 2004 is that the full four- to five-point reduction in the headcount over the period can be attributed to expansions in federal redistribution programs, with all other “poverty determinants” – the level and composition of growth, changes in inflation, changes in state-level spending, etc. – essentially canceling one another out.

The budding “mesoeconomic” approach to the study of poverty dynamics has added to the macroeconomic approach in at least two ways. First, it has confirmed at a national level some of the broad findings from the cross-country literature. Two such results are (i) that lower inflation contributes to poverty reduction, after controlling for the growth rate and other covariates; and (ii) that high levels of initial inequality, in assets or incomes, are correlated with lower subsequent poverty reduction.[12] Second, by exploring spatial and temporal variation in the sector composition of economic growth, these papers show that not all growth is the same. Some kinds of growth reduce poverty more than other kinds, and it is not always the same kind. Which types of growth have the greatest impact on poverty depends on the country in question, as well as on spatial differences in economic structure and historical distribution.

[12] On the cross-country association between inflation and poverty, see Easterly and Fischer (2001) and Romer and Romer (1999). On asset inequality and growth see Deininger and Olinto (2000), in addition to the references cited in the introduction.

This of course implies that, for policymakers or researchers interested in any particular country, relying on the three stylized facts in Section 2 is not enough. Individual growth processes *can* be made more or less “pro-poor” and, depending on the objectives[13] of the policy maker, certain trade-offs between higher average growth and higher growth for the poor may become relevant. These results should really come as no surprise, except perhaps to the most hardened representative-agent macroeconomist. “Economic growth” is nothing but an average taken across the proportional output increases across firms (or sectors) and government agencies across the economy. Or, if one approaches it from the income account, it is an average taken across proportional income increases across all households (adjusting from public goods, retained earnings, and measurement error). It should be perfectly plain that policies that allocate resources and opportunities differently across these sectors, firms and households will likely affect both the distribution and the average of future incomes. In fact, perhaps more can be learned from examining how the underlying distribution of incomes changes over time, including in response to shocks and policy changes, than from repeated estimations of how one or more vertices in the triangle related by equation (3) affect the other ones. This microeconomic approach to growth and distributional dynamics is the subject of the next section.

## 4. The microeconomic approach

The final level of work on the link between the dynamics of mean incomes (growth), poverty and inequality investigates this at the microeconomic level. These are studies that do not rely on summary measures of poverty or inequality as their primary data, but are based instead on full distributions of income or consumption expenditures, from representative household (or, in some cases, labor force) surveys. As noted in the introduction, this paper focuses on studies that investigate distributional dynamics under the anonymity axiom, and hence rely on repeated cross-section, rather than panel, data.

The natural starting point for these studies – although the term had not yet been coined when a number of them were written – is the growth incidence curve (GIC). Defined (by Ravallion and Chen, 2003) as the quantile-specific rate of economic growth between two points in time as a function of each percentile *p* in [0, 1], it can be written as follows, in continuous or discrete time:

$$g(p) = \frac{dy(p)}{y(p)} \tag{7}$$

$$g_t(p) = \frac{y_t(p) - y_{t-1}(p)}{y_{t-1}(p)} \tag{7'}$$

The discrete-time version in equation (7’) makes it clear that the GIC is simply the proportional difference between the quantile function at each percentile of the income distribution.

If we define economic growth as the proportional change in the mean of the income distribution, then growth can be expressed as a function of the growth incidence curve. By changing the integration variable, note that: $\mu = \int_{-\infty}^{\infty} y dF(y) = \int_0^1 y(p)dp$. Then it follows that:

[13] His or her “social welfare function”, to use a slightly older language.

$$\frac{d\mu}{\mu} = \int_0^1 \frac{dy(p)}{y(p)} \frac{y(p)}{\mu} dp = \int_0^1 g(p) \frac{y(p)}{\mu} dp \tag{8}$$

Equation (8) simply reminds us of the obvious fact that average income growth is a weighted sum of growth rates along the income distribution, weighted by each individual's initial income level.[14] It is also easy to show that changes in poverty can also be expressed in terms of the GIC, for a large class of poverty measures.[15] Poverty measures that satisfy the symmetry, monotonicity, focus and additive decomposability axioms can be written as:

$$P_t = \int_{-\infty}^{F(z)} \pi(y_t(p), z) dp \tag{9}$$

Where *z* denotes the real poverty line as usual, and $\pi(y_t(p), z)$ denotes the individual poverty function. Well-known examples of this class include the Foster-Greer-Thorbecke family of measures, which is obtained when $\pi(y_t(p), z, \theta) = \left(\frac{z - y_t(p)}{z}\right)^\theta$, and the Watts index, from $\pi(y_t(p), z) = ln \frac{z}{y_t(p)}$. Differentiating (9) with respect to time, while holding *z* constant and using the standard notation in this paper[16], we have:

$$dP_t = \int_{-\infty}^{F(z)} \eta_t(p) g_t(p) dp + \pi(z, z) dF_t(z) \tag{10}$$

Where $\eta_t(p) = \frac{\partial \pi(y_t(p), z)}{\partial y_t(p)} y_t(p)$ and $g(p) = \frac{dy(p)}{y(p)}$.

The first term on the RHS of equation (10) tells us that changes in poverty arise from income changes along the distribution (given by the growth incidence curve), multiplied by how sensitive the particular poverty measure in use is to changes at each point of the distribution (given by the function $\eta_t(p)$). Because of the focus axiom, there is a second term in (10), which captures changes in upper limit of the integral in (9) which might arise as a result of the changes in the distribution of incomes. This term is multiplied by the sensitivity of the poverty measure at the poverty line.

It turns out that there are a number of inequality indices whose changes over time can also be straightforwardly expressed as functions of the growth incidence curve. Consider, for example, the class of population-subgroup decomposable relative inequality measures. These are indices that aggregate functions of relative incomes across the distribution, where 'relative income' refers to an absolute income divided by the mean.[17] Write such a class in general terms as:

$$I_t = G\left[\int_0^1 h\left(\frac{y_t(p)}{\mu}\right) dp\right] \tag{11}$$

---

[14] The fact that the standard measure of economic growth weighs proportional changes in the incomes of the wealthy much more heavily than proportional changes in the incomes of the poor has long been remarked upon. See e.g. Ahluwalia and Chenery (1974) and Klasen (1994).

[15] This analysis follows Kraay (2006).

[16] While this notation is still probably clearest for the paper as a whole, note the possible confusion between two uses of the operator d in (10). When it appears within an integral, it denotes the integrating variable. When it appears alone, it denotes a time-derivative.

[17] Such measures are "relative inequality measures" by construction. They satisfy the scale invariance, rather than the translation invariance, axiom.

Well-known examples of this class of inequality measures include the Atkinson family, when $G\left[\int_0^1 h\left(\frac{y_t(p)}{\mu}\right)dp\right] = 1 - \left[\int_0^1 \left(\frac{y_t(p)}{\mu}\right)^{1-\varepsilon} dp\right]^{\frac{1}{1-\varepsilon}}$, and the Generalized Entropy Class when $G\left[\int_0^1 h\left(\frac{y_t(p)}{\mu}\right)dp\right] = \frac{1}{\theta^2-\theta}\left[\int_0^1 \left(\frac{y_t(p)}{\mu}\right)^{\theta} dp - 1\right]$. Differentiating (11) with respect to time yields:

$$dI_t = G'(\quad) \int_0^1 h'\left(\frac{y_t(p)}{\mu}\right)\frac{\mu}{y_t(p)}\left[g_t(p) - \frac{d\mu}{\mu}\right]dp \qquad (12)$$

Equation (12) is equally intuitive to interpret. At its core are differences between income growth in each percentile, and the growth in mean income, $\left[g_t(p) - \frac{d\mu}{\mu}\right]$. If everybody's income rises in exactly the same proportion, then that term vanishes, and relative inequality remains unchanged. If individual growth rates differ along the distribution, then they will contribute to changes in the aggregate inequality index, in a manner which depends on how sensitive the index is to relative incomes at each particular percentile *p*: $h'\left(\frac{y_t(p)}{\mu}\right)\frac{\mu}{y_t(p)}$.

Equations (8), (10) and (12) indicate that economic growth (at least when measured as the growth in mean household income), changes in poverty and changes in inequality are ultimately just different ways of aggregating the information contained in the growth incidence curve. Each of the three concepts is driven by changes in individual incomes, at the microeconomic level. Because they seek to capture different features of distributional change, they weight those individual changes differently: Economic growth weighs growth in individual incomes by their original income relative to the mean. Poverty measures weigh them according to how sensitive they are to income shortfalls from the poverty lines. And inequality measures weigh them depending on their basic individual "distance" measure.

This fact is more than a simple mathematical curiosity. It highlights the fact that the three "corners" of the poverty – growth – inequality triangle are so deeply related to one another because they are, in fact, variant forms of aggregation of information about the *incidence of growth* on the initial income distribution. Identities such as equation (3), which relates changes in a particular poverty measure (the headcount index) to changes in the mean and the Lorenz curve, arise from this common origin shared by the three concepts.

But if that is the case, then statistical and econometric analysis of the growth incidence curve should prove a sensible way to explore the dynamic relationship between growth, poverty and inequality, under the anonymity axiom. In order to investigate what share of the change in poverty, inequality, or growth, may be associated with a particular economic event (such as a demographic change, a macroeconomic shock, or a structural transformation) one might for instance seek to decompose the GIC into a component that corresponds to the change that can be attributed to the event in question, and a residual component, as follows:

$$g(p) = \frac{y^S(p) - y_{t-1}(p)}{y_{t-1}(p)} + \frac{y_t(p) - y^S(p)}{y_{t-1}(p)} \qquad (13)$$

Where $y^s(p)$ denotes a *counterfactual income distribution*, which would be obtained by the exclusive application of the event to $y_{t-1}(p)$, all else constant. If such a counterfactual could be empirically estimated, then the first term on the RHS of (13) would give the corresponding *counterfactual growth incidence curve*. This term could be substituted into equations (8), (10) or (12), to generate the resulting counterfactual growth, poverty and inequality change arising from the policy. The second term on the RHS of (13) simply measures the residual component of the GIC: changes in income distribution that were not caused by the policy or other object of investigation.

Of course the great challenge is, as usual, that of identification: how does one compute a meaningful causal estimate of $y^s(p)$? In a limited set of cases, such as when estimating the impact of a randomly assigned policy intervention with no diffuse spillovers, the usual experimental methods can be used to construct such a causally interpretable counterfactual. Another (tantalizing) possibility would be to "import" causal estimates (say, of the effect of a weather shock) from panel data analysis, into the construction of a counterfactual GIC. In most other cases – whether one is interested in understanding the effect of endogenous changes in fertility rates on inequality, or the effect of a currency devaluation on poverty – a causal interpretation of (13) is extremely difficult.[18]

That crucial caveat notwithstanding, there is actually a well-established literature on the *nature* (if not causes) of distributional change that basically revolves around estimating statistical decompositions like (13), and using them as suggestive descriptions of the factors that drive poverty and inequality dynamics. Most of the original contributions to this literature – Juhn, Murphy and Pierce (1993), DiNardo, Fortin and Lemieux (1996), Bourguignon, Fournier and Gurgand (2001) – were written before Ravallion and Chen (2003) defined the GIC, and therefore do not use that language. But they all use the "active ingredient" of equation (13), namely the idea of decomposing the overall change in income distribution into two or more steps demarcated by counterfactual income distributions.

### *4.1 Decomposing distributional change using statistical counterfactuals*

The pioneering paper was perhaps Juhn, Murphy and Pierce's (JMP, 1993) attempt to decompose changes in the US wage distribution between 1963 and 1989 into a term attributable to changes in observable worker characteristics (such as educational attainment), a term corresponding to changes in market returns to those characteristics, and another component associated with changes in both the distribution of and the returns to unobservable worker characteristics. These authors extended the classic decomposition of differences in mean earnings between different population subgroups (men and women, blacks and whites) due to Oaxaca (1973) and Blinder (1973), by incorporating changes in the distribution of residuals into the analysis. Like Oaxaca and Blinder, their analysis was based on the standard Mincer earnings equation (for individual *i* at time *t*):

$$y_{it} = X_{it}\beta_t + \varepsilon_{it} \tag{14}$$

[18] For an illuminating discussion of the role of counterfactuals in assessing the causal effects of policies on inequality, including a discussion of the separation between individual heterogeneity and uncertainty, see Cunha and Heckman (2007).

Denoting the cumulative distribution of residuals at time t by $\varphi_t = F_t(\varepsilon_{it}|X_{it})$, they constructed two counterfactual income distributions:

$$y_i^s = X_{i,t=0}\beta_{t=1} + F_{t=0}^{-1}\left(\varphi_{i,t=0}\middle|X_{i,t=0}\right) \quad (15)$$

And

$$y_i^{s\prime} = X_{i,t=0}\beta_{t=1} + F_{t=1}^{-1}\left(\varphi_{i,t=0}\middle|X_{i,t=0}\right) \quad (16)$$

Although JMP do not use this language or notation, their analysis can be written in terms of a decomposition of the growth incidence curve between 1963 (t=0) and 1989 (t=1) as follows:

$$g(p) = \frac{y^s(p)-y_{t=0}(p)}{y_{t=0}(p)} + \frac{y^{s\prime}(p)-y^s(p)}{y_{t=0}(p)} + \frac{y_{t=1}(p)-y^{s\prime}(p)}{y_{t=0}(p)} \quad (17)$$

The first term on the RHS of (17) was interpreted as the "returns component" of the decomposition, since the only differences between $y_{t=0}(p)$ and $y^s(p)$ arise from the change in the β coefficients (see equation 15). The second term on the RHS of (17) was interpreted as the component due to changes in unobserved characteristics, since the only differences between those two quantile functions arise from a rank-preserving transformation of the residuals – i.e. from replacing each residual in the t=0 distribution with the residual with identical rank in the t=1 distribution (see the last term in equation 16). The third term in (17) corresponds to changes in the joint distribution of observed characteristics (X), and was obtained residually.

JMP sought to explain the secular increase in earnings inequality in the US in the period from the mid 1960s to the late 1980s. On the basis of the above decomposition – although without mentioning "growth incidence curves" – they concluded that very little of the increase was accounted for by changes in the distribution of observed worker characteristics, including years of schooling. From around 1979 onwards, a substantial share of the increase in the 90-10 earnings differential could be attributed to increases in the returns to observable characteristics, chiefly formal education. But throughout the entire period, the most important component of the change in inequality was changes in unobserved worker characteristics, and their remuneration. The authors attributed this to increasing returns to unobservable aspects of human capital, such as the ability to adapt to new technologies, which are only imperfectly correlated with formal schooling.

Even though this analysis, only briefly summarized here, can certainly lay no claim to identifying the exact *causes* of changes in the US wage distribution, it was nevertheless an influential piece of statistical evidence in the academic debate. In fact, another pioneering paper in the literature estimating counterfactual income distributions was motivated, at least in part, as a response to some of the findings by Juhn et al. (1993). Dinardo, Fortin and Lemieux (DFL, 1996) sought to investigate whether changes in labor market institutions – in particular the declines in the real value of the minimum wage and in the rate of unionization – had also contributed to rising inequality in the United States. Instead of working with the quantile function $y = F^{-1}(p)$, these authors expressed their decomposition in terms of the primal density function, *f(y)*.

When data is available on both income *y* and a set of covariates *X*, the density function is simply the marginal of the joint distribution $\Gamma(y, X)$. Using the definition of a conditional distribution, DFL write that identity as:

$$f_t(y) = \iiint g_t(y|X)\, \phi_t(X) dX \tag{18}$$

Where *g()* denotes the conditional distribution of *y* on *X*, $\phi_t(X)$ is the joint distribution function of observed covariates (*X*), and $\iiint \ldots dX$ represents the operation of integrating over every element of the vector *X*. Counterfactual density functions can then be generated either by simulating an alternative conditional distribution function $g^s\,(y|X)$, or by constructing the appropriate counterfactual joint distribution of covariates, $\phi^s\,(X)$. The first procedure can be seen as a generalization of importing the estimates of the β parameters from another year, as in equation (15) above. The Mincer equation used by Juhn, Murphy and Pierce (1993) is a representation of the conditional distribution of *y* on *X*, under a log-linear functional form assumption. The more general version of the corresponding counterfactual distribution would have a density given by:

$$f^s(y) = \iiint g^s(y|X)\, \phi_t(X) dX \tag{19}$$

The second procedure – constructing a counterfactual joint distribution of covariates – yields an alternative counterfactual marginal distribution:

$$f^{s'}(y) = \iiint g_t(y|X)\, \phi^s(X) dX \tag{20}$$

How is a counterfactual income distribution, such as those in equations (19) or (20), constructed in practice? Consider first equation (20). If the desired counterfactual is of the sort: "what would the wage distribution be if people had the characteristics of the population at t=1, but were paid according to the conditional distribution at t=0?", then a reasonable approximation would be to use:

$$\phi^s(X) = \phi_{t=0}(X)\psi(X) \tag{21}$$

where $\psi(X) = \frac{\phi(X|t=1)}{\phi(X|t=0)}$ is a re-weighting function, which reweights the sample observed at t=0, with the weights from t=1. Then, in essence, $f^{s'}(y) = \iiint g_{t=0}(y|X)\phi_{t=0}(X)\psi(X)\, dX$ is simply obtained by a kernel density estimation of the density function of *y*, on a sample that has been re-weighted by the ratio of population weights in *t=1* to *t=0*.

DiNardo, Fortin and Lemieux (1996) used this procedure (including some more complex variations on the same basic theme) to investigate whether changes in the composition of the population - say, in terms of its ethnic, educational or demographic make-up – could account for some of the increase in US wage inequality over the period of study. One compositional change which did seem to account for some of the increase in the earnings gap was the decline in the share of the population that was unionized – suggesting that changes in labor market institutions did have a role to play in the period's distributional dynamics. They also found that the persistent decline in the real value of the minimum wage had a negative effect on earnings in the lower tail, particularly for women.

Once a counterfactual density function - like (19) or (20) - has been constructed, the overall change in the distribution of incomes can be decomposed into the part accounted for by these specific changes, and a residual term, as follows:

$$f_{t=1}(y) - f_{t=0}(y) = [f^{s}(y) - f_{t=0}(y)] + [f_{t=1}(y) - f^{s}(y)] \tag{22}$$

The decomposition in (22), which has been termed a generalized Oaxaca-Blinder decomposition, is obviously analogous to the decomposition of GICs in equation (13). The latter can be uniquely obtained from (22) by integrating each density function to obtain cumulative distribution functions, inverting those to obtain the quantile functions, and then dividing the quantile function at t=0.

The basic idea of the generalized Oaxaca-Blinder decompositions is to "break up" the complex and multilayered processes behind distributional change into individual building blocks – e.g. changes in returns, changes in personal characteristics, changes in labor market institutions – so that a sense of their relative importance can be gauged. The counterfactuals can be constructed in different ways: relying on specific functional forms and on importing parameters estimated for one year into another, as in JMP; using non-parametric reweighting methods, as in DFL: or indeed using a combination of the two. The same basic idea has been extended in various directions, including the analysis of changes in the distribution of household incomes (as opposed to the simpler wage distributions studied by JMP and DFL), and the comparison of distributions across countries, rather than time periods.[19]

Although mechanically one could apply a decomposition such as (13) or (22) to any distribution, whether of wages or of household incomes, the choice of the conditional distribution linking *y* to the covariates *X* is likely to differ substantially depending on the nature of the distribution under study. For a distribution of wages or earnings, the Mincerian equation, or some non-parametric equivalent, clearly embodies the relevant features of the conditional distribution. If one were after causal parameters, one might want to correct for sample selection bias into employment, but if the objective is to provide an estimate of the conditional distribution of wages for workers actually employed, even the simple OLS estimate (in the parametric case) might suffice. When *y* denotes household incomes, however, there are other linkages between observed characteristics (X) and final incomes (y). Individual characteristics affect the very composition of households – both through choice of partner and through fertility decisions. They affect occupational decisions within the household – both who works and who doesn't, and the choice of sector and formality status. And they clearly still affect earnings in the usual way, depending on returns in different sectors and occupations.

Bourguignon, Fournier and Gurgand (2001) and Bourguignon, Ferreira and Lustig (BFL, 2005) combine the parametric and semi-parametric approaches described above to the more complex problem of decomposing changes in household income distributions. They note that the vector of household covariates (*X*) includes endogenous variables such as family size and individual occupation, which are

[19] In this paper, we focus on the extension to household income distributions. But see Donald, Green and Paarsch (2000) for a comparison of wage distributions between Canada and the US, using a hazard-function-based estimator of cumulative distributions; and Bourguignon, Ferreira and Leite (2008) for a comparison of household income differences between Brazil and the US.

likely affected by – or correlated with – other elements in the vector, such as individual levels of education, gender and ethnicity. They propose to separate out (as distinct statistical "building blocks") the effects of changes, say, in the distribution of education on the distribution of income through the different mechanisms: changes in family composition (through changes in fertility), changes in occupational structure, and changes in earnings. To do so, they treat the "endogenous covariates" of interest – chiefly fertility and occupational choices – differently from other covariates, and estimate their own conditional distributions on a narrower subset of "exogenous" variables. In other words, they partition the vector $X$ into two sub-vectors $V$ and $W$, where $V$ includes variables such as the number of children in the household, occupation of work and, in some cases, years of schooling. And they replace the overall joint distribution of $X$ with the corresponding product of conditional distributions and a reduced-order joint distribution of $W$. If, for simplicity, $V = (v_1, v_2)'$, then equation (18) can be re-written as:

$$f_t(y) = \iiint g_t(y|V,W)\, h_1(v_1|v_2,W) h_2(v_2|W) \phi_t(W) dV dW \qquad (23)$$

In practice, the case studies in Bourguignon, Ferreira and Lustig (2005) generally use parametric methods to construct counterfactual conditional distributions, *g( )* and *h( )*. The conditional distributions of individual earnings on their covariates are estimated by standard Mincer equations, separately by sector of activity, for initial and terminal years. The conditional distribution of observed occupations (e.g.: unemployment, inactivity, formal employment, informal employment, self-employment) is estimated by means of a discrete-choice model, such as a multinomial logit or probit. In some of the case studies, specifications differ for household heads and other household members, to introduce some measure of intra-household interdependence. Similar discrete choice models are used to estimate the conditional distribution of family size (or, more specifically, the number of children in each household) on observed family characteristics. [20]

After these statistical models of conditional distributions have been estimated for each relevant year in the decomposition (at least one initial and one terminal year), a set of counterfactual income distributions is constructed by importing the relevant set of parameters from one year into another. For example, to simulate the counterfactual distribution corresponding to Brazil's situation in 1976, but with occupational choice parameters from 1996, the coefficients from the occupational choice multinomial logit from 1996 would be imported into 1976, and used to reallocate workers across sectors in the 1976 sample.[21] Those workers who are counterfactually "moved", say from unemployment into informal wage employment, are then counterfactually ascribed the earnings that their characteristics would earn in that sector, given the sector-specific coefficients from a Mincer equation (and a residual drawn from the empirical distribution for the sector).

---

[20] See also Hyslop and Maré (2005) for a "less parametric" application of these decompositions to the distribution of household incomes, which closely follows DiNardo, Fortin and Lemieux (1996).

[21] Reallocating workers on the basis of importing coefficients of a multinomial logit require drawing pseudo-residuals from the appropriate Weibull distribution, subject to the constraint that they be consistent with the originally observed choice. See Bourguignon and Ferreira (2005) for details on this, and the entire BFL methodology.

These and a host of analogous manipulations generate a set of counterfactual income distributions, $f^s(y)$. For each one, counterfactual poverty and inequality measures can be computed, and compared to the initial and final distributions actually observed. As an illustration, the results for just such a set of comparisons – for the distributional change between 1976 and 1996 in Brazil – are presented in Table 1, which is taken from Ferreira and Barros (2005), one of the case studies in Bourguignon, Ferreira and Lustig (2005). The columns of Table 1 list the mean, four inequality, and three poverty measures (for two different poverty lines) for each actual and counterfactual distribution. The first two rows are for the actual distributions observed in 1976 and 1996. Each of the subsequent rows is for a counterfactual distribution, denoted by the Greek letters corresponding to the estimated coefficients that are imported from 1996 into 1976 in each case.

But just as the counterfactual distributions can be used to compute scalar summary measures of inequality and poverty, they can also be used to construct quantile functions, and to compute the GIC decomposition in equation (13). Four such "full distribution" decompositions are depicted, for the same study of Brazil between 1976 and 1996, in Figure 2.

In this figure, the solid dark line is an approximation of the actual growth incidence curve between 1976 and 1996.[22] Each of the other curves, identified by a set of Greek letters, corresponds to the counterfactual GIC obtained from comparing a counterfactual income distribution with the actual 1976 distribution. In each case, the counterfactual income distribution was constructed by importing a set of parameters (denoted by the corresponding Greek letters) from a model estimated in 1996, to the 1976 model, and micro-simulating the resulting changes. So in Figure 2, for example, the line denoted "αs and βs" gives the counterfactual growth incidence attributable only to changes in the returns in all Mincer equations. The next line, which also includes γs, combines those changes in returns with changes in occupational choices (estimated using multinomial logits). The line that also includes $\mu(d)$ further incorporates behavioral changes associated with fertility choices. Finally, the line that includes "$\mu(d), \mu(e)$, αs, βs, and γs", adds the changes in educational attainment, from a multinomial logit that estimates the conditional distribution of years of schooling on a set of exogenous variables such as race, age, gender, and region of residence. Critically, when the GIC represents a *combination* of effects, as in each of these cases, they are made internally consistent. For example, if a person in the 1976 sample is ascribed a new level of education by the simulation arising from importing the $\mu(e)$ parameters, then that new level of education carries through to the person's new occupational choice, decision on fertility, and earnings determination.

In the Brazil case study which I have used to illustrate this approach, the decompositions summarized in Table 1 and in Figure 2 suggest that three main economic forces combine to explain changes in income levels, poverty and inequality – i.e. in Brazil's "poverty-growth-inequality triangle" – between 1976 and 1996. The first is a shift in the occupational structure towards greater unemployment and informality, and towards fewer hours worked within the informal sector by less educated workers. This effect was responsible for the pronounced income losses in the first decile of the distribution, and the

[22] The writing preceded the definition of GICs, and the authors calculated the log differences in incomes for each percentile, which approximates their growth rates.

corresponding increase in bottom-sensitive poverty measures. Contemporaneously, levels of education increased throughout the Brazilian population, leading both to higher endowments of human capital being sold on the labor market, and to marked reductions in desired fertility. These two effects of the educational expansion would have contributed to higher incomes across the distribution. However, in a context of macroeconomic stagnation, the labor markets were unable to absorb the greater levels of education at the going wages, and returns to education fell. This third force essentially cancelled out the gains from the education expansion, leading to an income distribution that was little changed (above the first decile) between 1976 and 1996.

Bourguignon, Ferreira and Lustig (2005) also contain methodologically similar studies for Argentina, Colombia, Indonesia, Malaysia, Mexico and Taiwan (China). Although the decomposition techniques were very similar across these studies, each country case was unique in terms of the real economics underlying distributional change. In Taiwan, a large increase in female labor force participation contributed to a reduction in earnings inequality (because the women entrants had middling wage rates), but an increase in the inequality of household incomes (because they were predominantly married to high-earning men). In Mexico, rising female labor force participation was also a key part of the story, but the pattern was a mirror image of the Taiwanese experience: women entered both at the bottom and at the top of the wage distribution, leading to higher earnings inequality, but contributing to a decline in household income inequality (primarily because of the resulting increase in incomes for the poorest households). Part of the attraction of this version of the GIC decomposition using statistical counterfactuals is its versatility, and the ensuing ability to capture a wealth of different economic forces underpinning distributional change.

But, of course, it also has its limitations. The whole family of GIC decompositions using statistical counterfactuals suffers, for instance, from path dependence: the order in which the decomposition is undertaken affects the size of each individual component. In terms of equation (22), for instance: $[f^s(y) - f_{t=0}(y)] \neq [f^s(y) - f_{t=1}(y)]$ . This is a property that the generalized Oaxaca-Blinder decompositions inherit from their parent. In the original Oaxaca-Blinder decomposition too, the returns component was different if the difference in the β vectors was weighted by the characteristics of blacks (say), than if it was weighted by the characteristics of whites. The problem carries through to all decompositions of the form of (13) or (22). Fortunately, it can be addressed relatively easily, either by showing the results for all different paths of decomposition or, more formally, by taking the appropriate average (which turns out to be a Shapley value) across all of them, as shown by Shorrocks (1999).

A more serious limitation relates to the issue of causality, discussed earlier. The fundamental problem is that counterfactual income distributions are statistical constructs that do not necessarily correspond to a meaningful economic counterfactual. This too can best be understood in the simple terms of the original Oaxaca-Blinder decomposition. We may value a decomposition that tells us that *x*% of the difference in earnings between blacks and whites is associated with differences in characteristics between those two groups, while (100-*x*)% is due to differences in returns to those characteristics. But we understand that the counterfactual used to compute *x*, namely the earnings that blacks (whites) would receive if their characteristics were remunerated with the returns normally associated with whites (blacks), does not correspond to an economic equilibrium. Analogously, most counterfactual income distributions discussed in this sub-section – e.g. the income distribution that would attain in Brazil if the only change since 1976 were the change in returns to schooling observed by 1996 – do not correspond to tenable economic equilibria. Just as in the case of the original Oaxaca-Blinder decompositions, there is some value in decompositions based on counterfactuals that are statistically well-defined, even if they do not correspond to a tenable equilibrium and cannot therefore be interpreted causally.

Nevertheless, some recent research has attempted to construct GIC decompositions where the counterfactual distributions may correspond more closely to a tenable equilibrium, and where the ensuing distributional changes might, therefore, be interpreted as suggestive of causality. The next sub-section briefly reviews three (rather different) approaches to this quest.

### *4.2 Towards GIC decompositions based on economic counterfactuals.*

In order for a counterfactual distribution to correspond to an equilibrium allocation, every outcome in the data-generating process for each individual must be consistent with the equilibrium behaviors of all other agents in the economy. One approach to constructing such a counterfactual, therefore, would be to have a full structural model of behavior for the economic agents involved, and to use such a model to simulate the effects of a particular policy or shock. As an example, consider the model of education, fertility and labor supply estimated by Todd and Wolpin (2006) for the Mexican villages where the *Progresa* conditional cash transfer program was introduced in 1997. It was not the authors' objective to decompose changes in the distribution of income in that rural economy into a component due to the *Progresa* transfers and another (residual) component due to all other changes. Instead, their objective was to use the results of the experimental evaluation of the program – which yielded credible causal estimates of the effect of the transfers on household incomes by comparing outcomes between randomly assigned treatment and control villages – to validate their structural behavioral model.

But if the model succeeded in its objective, so that its predictions of the effects of the transfers on household incomes, accounting for labor supply, enrollment and fertility responses, are correctly validated by the treatment effect estimates from the experimental evaluation, then a distribution of income obtained from simulating the model would be an economically meaningful counterfactual. Under such a "true model", changes in poverty, inequality or growth computed from the corresponding growth incidence curve would indeed be causally – and not just statistically – attributable to the *Progresa* program. [23]

The feature of the Todd and Wolpin (2006) study which lends particular credibility to its model-based construction of an economic counterfactual GIC is the existence of credible estimates of program impact from an experimental evaluation. What happens, though, when one is interested in estimating the effect (on poverty, inequality and growth) of some economy-wide policy that is *not* assigned to specific groups, and cannot be evaluated by experimental or quasi-experimental methods? Examples of such

[23] See also Attanasio, Meghir and Santiago (2004) and Bourguignon, Ferreira and Leite (2003) for alternative models of the effects of conditional cash transfers on household incomes, accounting for behavioral responses.

policies include trade reform, exchange rate devaluations or revaluations, economy-wide labor market reforms, etc.

Growth incidence curve decompositions based on economic counterfactuals have been tried for such "economy-wide" policies as well, although the standards of empirical identification are probably lower than in the case of well-evaluated assigned programs. Let us briefly review two approaches. In the first, exemplified by Ferreira, Leite and Wai-Poi (2010), the treatment of general equilibrium relationships is extremely reduced-form. These authors attempt to estimate the effect of a trade liberalization episode on the distributions of wages and household incomes – and thus on poverty and inequality – in Brazil, between 1988 and 1995.

The authors combine the two-stage regression approach of Goldberg and Pavcnik (2005) with a parametric GIC decomposition in the style of Juhn, Murphy and Pierce (1993). They construct a panel of industries over time, and regress three industry-level dependent variables (wage-premia, skill-premia and employment levels), which they obtain from estimating Stage 1 regressions (24) and (25) below, on (arguably) exogenous trade policy variables (such as changes in tariff rates and industry-specific exchange rates).

$$\ln w_{ij} = X_{ij}\beta + I_{ij} * wp_j + (I_{ij} * S_{ij})sp_j + \varepsilon_{ij} \tag{24}$$

$$\Pr\{j = s\} = P^s(Z_i, \lambda) = \frac{e^{Z_i \lambda_s}}{e^{Z_i \lambda_s} + \sum_{j \neq s} e^{Z_i \lambda_j}} \tag{25}$$

Equation (24) is an individual-level earnings regression, augmented by industry ( $I_{ij}$ ) and skill ($S_{ij}$) dummy variables. The coefficients $wp_j$ and $sp_j$ are correspondingly interpreted as industry and industry-skill wage premia. A panel of such "premia coefficients" (as well as estimates of the constant term $\lambda_0$ in eq. 25) across industries *j* and years *t* is regressed on the trade policy variables, as noted above: these are the Stage 2 regressions. Once they have been estimated, the coefficients from the second-stage regressions are used to predict "trade-mandated" changes in $wp_j$, $sp_j$, and $\lambda_{0j}$ , for actual or counterfactual values of changes in the exogenous trade policy variables. These predicted first-stage coefficients are in turn imputed back into (24) and (25), to generate counterfactual occupation and wage distributions analogous to those in JMP or BFL.

Inequality and poverty statistics can be computed for each of these counterfactual distributions, and compared with the actual (pre- and post-liberalization) distributions, in an attempt to isolate the contribution of the policy to the overall changes. Equation (13) can also be computed for each percentile, showing a full graphical comparison of the actual GIC between 1988 and 1995, and the counterfactual GIC attributed only to the trade-mandated effects of the liberalization. This decomposition is shown in Figure 3, drawn from Ferreira et al. (2010). In this figure, the thick upper line is the actual wage growth incidence curve for Brazil between 1988 and 1995, and the dashed line that matches it closely from the first quintile upwards is the counterfactual GIC obtained from the simulation exercise just described. The line is interpreted in the paper as a lower-bound for the effects of trade liberalization on wage inequality in Brazil. The authors also show that this inequality-reducing effect is almost entirely due to employment reallocation across industries (changes in $\lambda_0$), rather than to changes in wage or skill premia.

The second approach goes into further detail in the modeling of the general equilibrium (or macroeconomic) relationships that link the behavioral responses of different firms and individuals to a particular shock or policy change, and to one another. Rather than relying on a reduced-form relationship between a set of observed exogenous variables (such as changes in tariff rates) and a set of industry-level shifters (as above), studies in this second approach estimate a full macroeconomic (or computable general equilibrium) model. This "macro model" (for short) is used to generate a set of "linkage aggregate variables", such as vectors of wage rates and employment levels, for certain combinations of sectors and types of workers (e.g. wages and employment for high-skilled workers in the informal sector in urban areas).

These variables are then used to connect the macro model to a set of earnings and occupation equations – similar to (24) and (25) above - estimated on household (or labor force) micro-data. A convergence algorithm is used to ensure that the counterfactual distributions of employment choices and wage rates add up to the aggregate simulations from the macro model. At the micro-level, the simulation generates counterfactual occupation and income distributions - and growth incidence curves - much as in the Bourguignon et al. (2005) or the Ferreira et al. (2010) exercises. Behind these counterfactual GICs now lies a full macroeconomic or computable general equilibrium model. To the extent that one trusts the capacity of those models to accurately predict the effects of policies on the general equilibrium of the economy, these GICs are also economically meaningful counterfactuals.[24] That, however, is not a trivial caveat…

## 5. Conclusions

Income distribution dynamics have long been of interest to economists. As the availability of household survey data for developed and (particularly) developing countries expanded in the 1990s, so did their ability to investigate distributional change empirically. The cross-country (or "macroeconomic") literature that sought to exploit international variation in poverty and inequality changes, economic growth and covariates, offered no support for the "grand theories" linking development and distribution. The evidence did not appear particularly supportive of the Kuznets hypothesis, and was inconclusive about the possible effects of initial (or lagged) inequality and poverty on subsequent growth.

That literature did, however, generate three robust stylized facts about growth and distribution in the last couple of decades of the Twentieth Century: (i) there was no statistically significant cross-country correlation between economic growth and changes in inequality; (ii) so economic growth was strongly

[24] See the various chapters in Bourguignon, Bussolo and Pereira da Silva (2008) for a set of examples of this "macro-micro" approach.

and negatively correlated with changes in poverty. However, (iii) the higher a country's initial level of inequality, the higher the growth rate needed to obtain a given amount of poverty reduction.

But economic growth, changes in inequality and changes in poverty are actually just three different aggregations of information about individual income dynamics. They are therefore jointly determined (by the general equilibrium of the economy), and macroeconomic estimates of the reduced-form relationships between them – however useful in identifying empirical regularities – were never likely to shed much light on the fundamental factors underlying distributional change.

Two alternative approaches have been more successful in doing that. The first, which I have called "mesoeconomic", uses sub-national panel data on poverty and on economic growth rates disaggregated by sector, to investigate the role of different growth patterns and initial conditions on poverty reduction. The second approach, which is microeconomic in nature, investigates distributional change at a fully disaggregated level, by decomposing changes in the growth incidence curve. These decompositions have not resolved the identification problems inherent in studying distributional dynamics either. The first crop of studies in this tradition are essentially generalizations of the Oaxaca-Blinder decomposition to a full-distribution, dynamic setting. The counterfactual income distributions on which they rely suffer from the usual problems of equilibrium-inconsistency and path dependence.

Nevertheless, they have succeeded in shedding some light on the nature of distributional change in countries ranging from Indonesia to the United States, in a set of quite informative ways. These studies have employed various different methods, parametric and otherwise. They have generally focused on a set of key factors, including (i) the dynamics of the distribution of educational attainment; (ii) changes in the returns to education (and, less prominently, other covariates); (iii) changes in the structure of occupations, including female labor force and the extent and nature of the informal sector(s); (iv) the links between education, labor force participation, and demographic change; (v) changes in labor market institutions, including unionization and minimum wages. Although the topics are often similar across countries, as the number of studies expands, one interesting result has been just how different each country's specific story is.[25] It appears that the basic pieces of the income distribution dynamics puzzle can be combined in a multiplicity of ways.

A second crop of growth-incidence based studies attempts to get nearer to a causal interpretation of the GIC decompositions, by deriving counterfactual income distributions from models of behavior, or of the general equilibrium of the economy. This paper briefly summarized a few such studies, which were remarkable, if for nothing else, at least for their methodological diversity – ranging from old-fashioned CGEs linked to micro-simulations, to fully structural models of dynamic household behavior.

Despite this great methodological diversity, there are some shared findings and areas of common ground in the multifaceted literature on the poverty-growth-inequality triangle. Growth is good for the poor, and it is particularly good when it is the incomes of the poor that are growing...[26] This is most likely

[25] See, for example, the concluding chapter in Bourguignon, Ferreira and Lustig (2005).
[26] When the incomes of the poor grow faster than those of the non-poor, inequality generally declines, and poverty mechanically falls by more than if growth was uniformly distributed.

to happen when growth takes place in the areas where the poor live, and in the sectors where they work. They are better able to benefit when their initial endowments of human capital, land, and political power are greater. But contemporaneous policy choices can also make a great deal of difference to how the poor share in economic growth. Both market-friendly policies (such as farm-gate price liberalization in China and trade liberalization in Brazil) and state-led redistribution (such as investments in public education in various countries, and well-targeted cash transfer schemes in Brazil and Mexico) have contributed to faster poverty reduction. At the individual country level, in other words, there is a plethora of policy choices that naturally affects the endowments and growth opportunities of people all along the income distribution. These policies naturally affect the incidence and average rate of growth *simultaneously*, and thus jointly determine the evolution of the poverty-growth-inequality triangle.

The microeconomic literature that seeks to empirically describe this joint determination process is still in its infancy, and there is considerable scope for more work on building counterfactual distributions that are consistent with economic equilibria – possibly by striving for a middle ground between the full-scale complexity of structural models of dynamic household behavior and the ad-hoc rigidities of computable general equilibrium models.

Another direction with potentially high research payoffs is to learn from and draw more on the literatures on individual income dynamics and on socioeconomic mobility. As noted in the introduction, all of the literature reviewed in this paper falls under the aegis of the anonymity axiom, and relies essentially on repeated cross-sections. However, a number of concepts highlighted here, including that of the GIC, would remain relevant in a panel data context, subject to interesting adjustments. Grimm (2007) defines a variant of the GIC, termed the "individual growth incidence curve (IGIC)", which follows the same individual over time, and is defined as the income growth rate for each individual as a function of their percentile in the initial distribution. Whereas the Ravallion-Chen GIC is the relevant one for changes in inequality when the symmetry or anonymity axiom is upheld, Grimm's IGIC tells us about individual income trajectories over time, and thus about economic mobility. Scope remains for further interesting work on how the GIC and the IGIC relate, and on whether the IGIC provides as much of a unifying basis for mobility measurement as we have shown that the GIC does for inequality measurement.[27] Similarly, and as indicated earlier, there is probably much to be learned from combining the sort of growth incidence analysis reviewed in this paper with the insights on how various shocks and policy changes affect individual income trajectories, from the more causal literature on individual income dynamics.

[27] On the measurement of various concepts of mobility, see Fields and Ok (1996).

**References**

Ahluwalia, Montek and Hollis Chenery (1974): “The Economic Framework” in H. Chenery et al. (eds.) *Redistribution with Growth*, London: Oxford University Press.

Ahuja, Vinod, Benu Bidani, Francisco Ferreira and Michael Walton (1997): *Everyone’s Miracle? Revisiting Poverty and Inequality in East Asia*, Washington, DC: World Bank.

Alesina, Alberto and Dani Rodrik (1994): “Distributive Politics and Economic Growth”, *Quarterly Journal of Economics*, **CIX** (2): 465-490.

Atkinson, Anthony (1970): “On the Measurement of Inequality”, *Journal of Economic Theory* **2**: 244-263.

Attanasio, Orazio, Costas Meghir and Ana Santiago (2004): “Education choices in Mexico: Using a structural model and a randomized experiment to evaluate Progresa”, University College London IFS Working Paper EWP0404.

Banerjee, Abhijit and Esther Duflo (2003): "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* **8**(3): 267–99.

Banerjee, Abhijit and Andrew Newman (1993): “Occupational Choice and the Process of Development”, *Journal of Political Economy*, **101** (2), pp.274-298.

Bénabou, Roland (2000): "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* **90**(1): 96–129.

Blinder, Alan S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, **8**: 436-455.

Bourguignon, François (2004): “The Poverty-Growth-Inequality Triangle”, Indian Council for Research on International Economic Relations Working Paper #125.

Bourguignon, François, Maurizio Bussolo and Luiz Pereira da Silva (2008): *The impact of macroeconomic policies on poverty and income distribution: Macro-micro evaluation techniques and tools*. Washington, DC: Palgrave Macmillan and the World Bank.

Bourguignon, François, Francisco H. G. Ferreira (2005): “Decomposing changes in the distribution of household incomes: Methodological aspects”, Chapter 2 in Bourguignon, Ferreira and Lustig (2005).

Bourguignon, François, Francisco Ferreira and Phillippe Leite (2003): “Conditional cash transfers, schooling, and child labor: Micro-simulating Brazil’s Bolsa Escola Program”, *World Bank Economic Review* **17** (2): 229-254.

Bourguignon, François, Francisco Ferreira and Phillippe Leite (2008): “Beyond Oaxaca-Blinder: Accounting for differences in household income distributions”, *Journal of Economic Inequality* **6**: 117-148.

Bourguignon, François, Francisco Ferreira, and Nora Lustig (2005): *The Microeconomics of Income Distribution Dynamics*. New York: World Bank and Oxford University Press.

Bourguignon, François, Martin Fournier and Marc Gurgand (2001): "Fast development with a stable income distribution: Taiwan, 1979-94", *Review of Income and Wealth* **47** (2): 139-163.

Bruno, Michael, Martin Ravallion, and Lyn Squire (1998): "Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues." In Vito Tanzi and Ke-young Chu, eds., *Income Distribution and High-quality Growth.* Cambridge, Mass.: MIT Press.

Cunha, Flávio and James Heckman (2008): "A New Framework for the Analysis of Inequality", *Macroeconomic Dynamics* **12** (S2): 315-354.

Datt, Gaurav and Martin Ravallion (1992): "Growth and redistribution components of changes in poverty measures" *Journal of Development Economics* **38**: 275-295.

Deininger, Klaus and Pedro Olinto (2000): "Asset Distribution, Inequality and Growth." World Bank Policy Research Working Paper #2375.

Deininger, Klaus and Lyn Squire (1998): "New Ways of Looking at Old Issues: Inequality and Growth." *Journal of Development Economics* **57**(2): 259–87.

Dercon, Stefan (2004): "Growth and shocks: evidence from rural Ethiopia", *Journal of Development Economics* **74**: 309-329.

DiNardo, John, Nicole Fortin and Thomas Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Approach", *Econometrica*, **64** (5): 1001-1044.

Donald, Stephen G., David A. Green and Harry J. Paarsch (2000): "Differences in wage distributions between Canada and the United States: An application of a flexible estimator of distribution functions in the presence of covariates" *Review of Economic Studies* **67**: 609-633.

Dollar, David and Aart Kraay (2002): "Growth is good for the poor" *Journal of Economic Growth* **7**: 195-225.

Easterly, William and Stanley Fischer (2001): "Inflation and the Poor", *Journal of Money, Credit and Banking*, **33** (2): 160-178.

Ferreira, Francisco H. G. (2001): "Education for the Masses? The interaction between wealth, educational and political inequalities" *Economics of Transition* **9**(2): 533–52.

Ferreira, Francisco H. G. and Ricardo Paes de Barros (2005): "The slippery slope: Explaining the increase in extreme poverty in urban Brazil, 1976-96", Chapter 4 in Bourguignon, Ferreira and Lustig (2005).

Ferreira, Francisco H. G. , Phillippe G. Leite and Martin Ravallion (2010): “Poverty reduction without economic growth? Explaining Brazil’s poverty dynamics, 1985-2004” *Journal of Development Economics* **93**: 20-36.

Ferreira, Francisco H. G., Phillippe G. Leite and Matthew Wai-Poi (2010): “Trade Liberalization, Employment Flows, and Wage Inequality in Brazil”, chapter 8 in M. Nissanke and E. Thorbecke (eds) *The Poor under Globalization in Asia, Latin America and Africa*, Oxford: Oxford University Press.

Fields, Gary and Efe Ok (1996): “The meaning and measurement of income mobility”, *Journal of Economic Theory* **71** (2): 349-377.

Forbes, Kristin J. (2000): “A Reassessment of the Relationship between Inequality and Growth”, *American Economic Review* **90** (4): 869-887.

Foster, James, J. Greer, and Erik Thorbecke (1984): “A class of decomposable poverty measures.” *Econometrica* **52**: 761-65.

Galor, Oded and Joseph Zeira (1993): “Income Distribution and Macroeconomics”, *Review of Economic Studies*, **60**: 35-52.

Goldberg, Penelope and Nine Pavcnik (2005): “Trade Protection and Wages: Evidence from Colombian Trade Reforms”, *Journal of International Economics* **66**: 75–105.

Grimm, Michael (2007): “Removing the anonymity axiom in assessing pro-poor growth”, *Journal of Economic Inequality* **5**: 179-197.

Hyslop, Dean R., David C. Maré (2005): “Understanding New Zealand’s changing income distribution, 1983-1998: A semi-parametric analysis”, *Economica* **72**: 469-495.

Jalan, Jyostna and Martin Ravallion (2002): “Geographic poverty traps? A micro model of consumption growth in rural China”, *Journal of Applied Econometrics* **17**: 329-346.

Juhn, Chinhui, Kevin Murphy and Brooks Pierce (1993): “Wage Inequality and the Rise in Returns to Skill”, *Journal of Political Economy*, **101** (3): 410-442.

Kakwani, Nanak (1993): “Poverty and Economic Growth with Application to Côte D'Ivoire”, *Review of Income and Wealth* **39**: 121-139.

Klasen, Stephan (1994): “Growth and Well-Being: Introducing distribution-weighted growth rates to reevaluate U.S. postwar economic performance”, *Review of Income and Wealth* **40** (3): 251-272.

Kraay, Aart (2006): “When is growth pro-poor? Evidence from a panel of countries” *Journal of Development Economics* **80**: 198-227.

Kuznets, Simon (1955): “Economic growth and income inequality”, *American Economic Review* **65** (1): 1-29.

Lewis, Arthur W. (1954): “Economic development with unlimited supply of labour”, *The Manchester School*, **22**(2): 139-91.

Lybbert, T. J., Barrett, C., Desta, S and Layne Coppock, D. (2004): “Stochastic wealth dynamics and risk management among a poor population” *Economic Journal* **114** (498): 750-777.

Oaxaca, Ronald (1973): "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, **14**: 673-709.

Ravallion, Martin (2007): “Inequality is Bad for the Poor”, in J. Micklewright and S. Jenkins (eds.), *Inequality and Poverty Re-Examined.* Oxford: Oxford University Press.

Ravallion, Martin and Shaohua Chen (1997): “What can new survey data tell us about recent changes in distribution and poverty?” *World Bank Economic Review* **11** (2): 357-382.

Ravallion, Martin and Shaohua Chen (2003): “Measuring pro-poor growth” *Economics Letters* **78**: 93-99.

Ravallion, Martin and Shaohua Chen (2007): “China’s (Uneven) Progress Against Poverty”, .*Journal of Development Economics*, **82** (1): 1-42.

Ravallion, Martin and Gaurav Datt (2002): “Why has economic growth been more pro-poor in some states of India than others?” *Journal of Development Economics* **68**: 381-400.

Ricardo, David (1817): *On The Principles of Political Economy and Taxation.* London: Dent, 1911 edition.

Romer, C.D. and D.H. Romer (1999): “Monetary Policy and the Well-Being of the Poor”, *Federal Reserve Bank of Kansas City Economic Review*, First Quarter, pp.21-49.

Shorrocks, Anthony (1999): “Decomposition procedures for distributional analysis: A unified framework based on the Shapley value” University of Essex, mimeo.

Tinbergen, Jan (1975): *Income Differences: Recent Research*; Amsterdam: North-Holland.

Todd, Petra E. and Kenneth I. Wolpin (2006): “Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioural model of child schooling and fertility”, *American Economic Review* **96** (5): 1384-1417.

Voitchovsky, Sarah (2009): “Inequality and Economic Growth”, Chapter 22 in W. Salverda, B. Nolan and T. Smeeding (eds.) *The Oxford Handbook of Economic Inequality*, Oxford: Oxford University Press.

World Bank (2005): *World Development Report 2006: Equity and Development*, Washington, DC: World Bank and Oxford University Press.

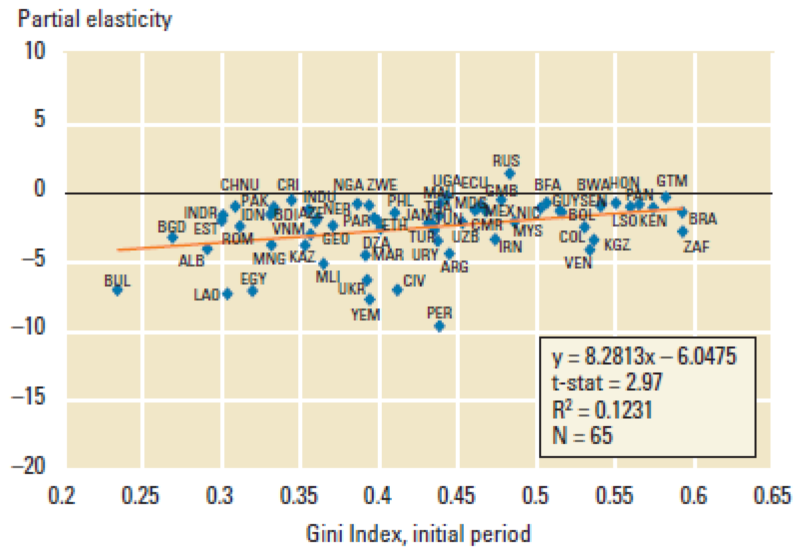
**Table 7: Simulated Poverty and Inequality for 1976, Using 1996 coefficients.**

| | Mean p/c Income | Inequality | | | | Poverty: Z = R$30 / month | | | Poverty: Z = R$ 60 / month | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Gini | E(0) | E(1) | E(2) | P(0) | P(1) | P(2) | P(0) | P(1) | P(2) |
| 1976 observed | 265.101 | 0.595 | 0.648 | 0.760 | 2.657 | 0.0681 | 0.0211 | 0.0105 | 0.2209 | 0.0830 | 0.0428 |
| 1996 observed | 276.460 | 0.591 | 0.586 | 0.694 | 1.523 | 0.0922 | 0.0530 | 0.0434 | 0.2176 | 0.1029 | 0.0703 |
| ***Price Effects*** | | | | | | | | | | | |
| α, β for wage earners | 218.786 | 0.598 | 0.656 | 0.752 | 2.161 | 0.0984 | 0.0304 | 0.0141 | 0.2876 | 0.1129 | 0.0596 |
| α, β for self-employed | 250.446 | 0.597 | 0.658 | 0.770 | 2.787 | 0.0788 | 0.0250 | 0.0121 | 0.2399 | 0.0932 | 0.0490 |
| α, β for both | 204.071 | 0.598 | 0.655 | 0.754 | 2.190 | 0.1114 | 0.0357 | 0.0169 | 0.3084 | 0.1249 | 0.0673 |
| α only, for both | 233.837 | 0.601 | 0.664 | 0.774 | 2.691 | 0.0897 | 0.0275 | 0.0129 | 0.2688 | 0.1040 | 0.0545 |
| All β (but no α) for both | 216.876 | 0.593 | 0.644 | 0.736 | 2.055 | 0.0972 | 0.0303 | 0.0143 | 0.2837 | 0.1114 | 0.0590 |
| Education β for both | 232.830 | 0.593 | 0.639 | 0.759 | 2.691 | 0.0779 | 0.0234 | 0.0110 | 0.2531 | 0.0953 | 0.0488 |
| Experience β for both | 240.618 | 0.600 | 0.664 | 0.771 | 2.694 | 0.0851 | 0.0265 | 0.0125 | 0.2592 | 0.1000 | 0.0525 |
| Gender β for both | 270.259 | 0.595 | 0.649 | 0.751 | 2.590 | 0.0650 | 0.0191 | 0.0090 | 0.2160 | 0.0797 | 0.0404 |
| ***Occupational Choice Effects*** | | | | | | | | | | | |
| γ for both sectors (and both heads + others) | 260.323 | 0.609 | 0.650 | 0.788 | 2.633 | 0.0944 | 0.0451 | 0.0331 | 0.2471 | 0.1082 | 0.0671 |
| γ for both sectors (only for other members) | 265.643 | 0.598 | 0.657 | 0.757 | 2.482 | 0.0721 | 0.0231 | 0.0119 | 0.2274 | 0.0867 | 0.0454 |
| γ, α, β for both sectors | 202.325 | 0.610 | 0.649 | 0.788 | 2.401 | 0.1352 | 0.0597 | 0.0402 | 0.3248 | 0.1466 | 0.0902 |
| ***Demographic Patterns*** | | | | | | | | | | | |
| μd only, for all | 277.028 | 0.574 | 0.585 | 0.704 | 2.432 | 0.0365 | 0.0113 | 0.0063 | 0.1711 | 0.0554 | 0.0264 |
| μd , γ, α, β, for all | 210.995 | 0.587 | 0.577 | 0.727 | 2.177 | 0.0931 | 0.0433 | 0.0321 | 0.2724 | 0.1129 | 0.0677 |
| ***Education Endowment Effects*** | | | | | | | | | | | |
| μe only, for all | 339.753 | 0.594 | 0.650 | 0.740 | 2.485 | 0.0424 | 0.0136 | 0.0073 | 0.1593 | 0.0567 | 0.0287 |
| μd, μe for all | 353.248 | 0.571 | 0.584 | 0.688 | 2.320 | 0.0225 | 0.0078 | 0.0049 | 0.1131 | 0.0359 | 0.0173 |
| μe , μd , γ, α, β, for all | 263.676 | 0.594 | 0.600 | 0.727 | 1.896 | 0.0735 | 0.0374 | 0.0296 | 0.2204 | 0.0913 | 0.0561 |

Source: Based on "Pesquisa Nacional por Amostra de Domicílios" (PNAD) of 1976 and 1996.

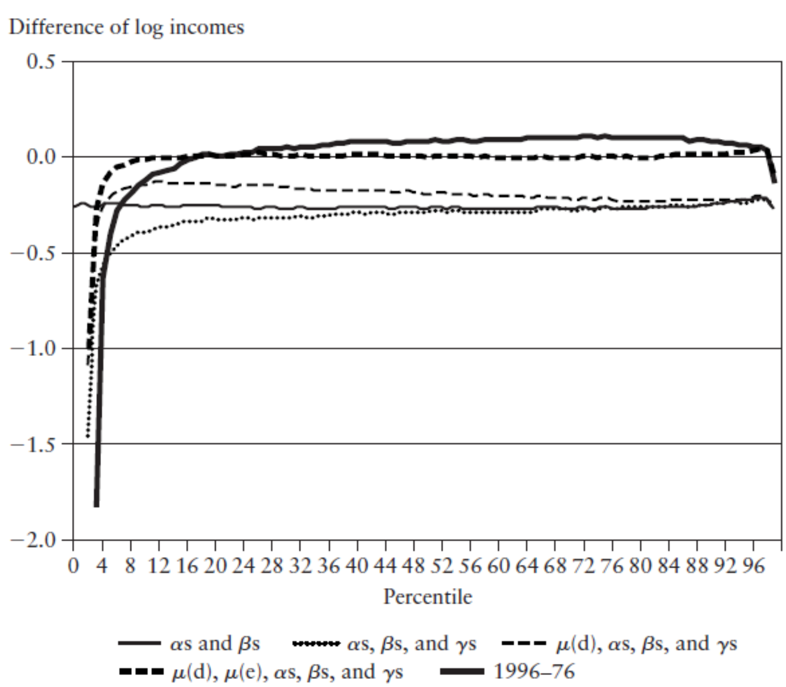
Source: Ferreira and Paes de Barros (2005).

Figure 1: Empirical partial growth elasticities of poverty reduction against initial Gini index: (LDCs in 1981-2004, poverty headcount, z = US$1 a day).




Source: World Bank (2005)

Figure 2: A Generalized Oaxaca-Blinder decomposition of the GIC for Brazil, 1976-1996

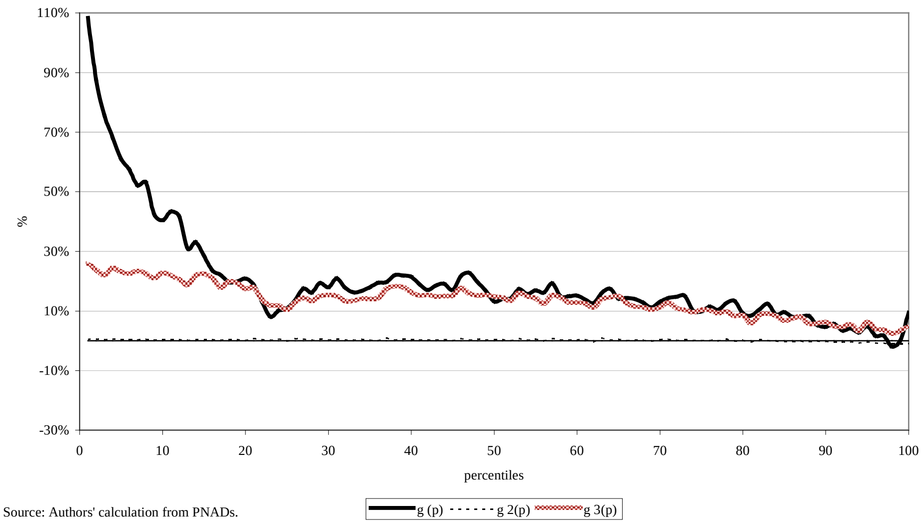



Source: Ferreira and Paes de Barros (2005)

Figure 3:

Observed and counterfactual wage growth incidence curves, 1988-95,

all trade-mandated changes from 2nd stage




Source: Authors' calculation from PNADs.

Source: Ferreira, Leite and Wai-Poi (2010)