Interdistributional Inequality, Stochastic Dominance, and Poverty*

John A. Bishop
East Carolina University, USA

Lester A. Zeager
East Carolina University, USA

Buhong Zheng
University of Colorado, Denver, USA

March 31, 2011

Abstract

Interdistributional inequality comparisons reemerged through the work of Butler and McDonald (1987), Deutsch and Silber (1997, 1999), and Bishop, Chow, and Zeager (2004, 2010), many years after the initial explorations by Gini (1916, 1959). Interdistributional Lorenz curves (ILCs), proposed by Butler and McDonald (1987), are visually attractive. Yet, their relationships with other dominance methods using partial moments of income distributions (e.g., stochastic dominance, beyond the first-degree, and the FGT poverty measures, beyond headcounts) are not clear. Our aim in this paper is to clarify the relationships.

* The authors thank Richard Ericson, Hosong Kim, and participants in the November, 2010, Southern Economic Association meetings in Atlanta, GA, for useful comments and suggestions.

Interdistributional Inequality, Stochastic Dominance, and Poverty

1. Introduction

When researchers began to investigate the relative incomes (wages) between population subgroups (e.g., by race or gender), it was natural for the initial comparisons to focus on the means of the distributions for the subgroups. Some researchers came to realize, however, that other aspects of the distributions should also be taken into consideration. They created more sophisticated methods of comparison for this purpose, including "economic distance" or "economic advantage" (Shorrocks, 1982; Vinod, 1985).

Butler and McDonald (1987) created "Interdistributional Lorenz Curves" (ILCs) to measure the "distance between" black and white income distributions in the United States. These ILCs often look like LCs,[1] but they capture inequality of a different kind. LCs reveal the inequality *within one distribution*; ILCs reveal the inequality *between two distributions*. One form of ILC is reminiscent of Gini's (1916, 1959) proposal for measuring the concentration of each distribution below a reference point in the other. As the roots of the idea suggest, ILCs involve comparisons of corresponding partial moments of two income distributions. Indeed, the comparisons involve what Butler and McDonald call "normalized incomplete moments," created by dividing partial moments by the mean of the distribution, yielding the fraction or proportion of the moment of interest.

[1] ILCs are not necessarily concave to the diagonal line (Butler and McDonald, 1987, 14), and they may even intersect the diagonal, like the concentration curves developed by Kakwani (1980).

Another method that helps researchers to look beyond the means of the income distributions is the FGT class of poverty measures proposed by Foster, Greer, and Thorbecke (1984), which involve comparisons of partial moments for two income distributions. As Foster and Shorrocks (1988) showed, rankings of income distributions by three poverty measures (P_1 : headcount ratio; P_2: per capita income gap; and P_3: distribution sensitive measure) correspond to the rankings of the same distributions by the first three orders of stochastic dominance (FSD, SSD, and TSD). Thus, the relationships between ILCs and FSD, SSD, and TSD, if any, could be extended to the FGT class of poverty measures.

Our investigation reveals that the relationships among these various dominance measures depend on the order of dominance. At the most basic level, ILCs of a certain type (shares of income recipients) and FSD give us equivalent orderings of distributions, which also correspond to the ordering of these distributions by P_1, the headcount ratio. Beyond this basic level, the relationships among the three approaches break down, even when the means of the two distributions are the same. The breakdown in the correspondence can be traced to the normalization used in ILC comparisons.

The next section provides some basic definitions required for the discussion. The following explores relationships between two types of ILCs created by Butler and McDonald (1987) and the orders of stochastic dominance (FSD and SSD). It also briefly characterizes the relationships between the two types of ILCs and the first two classes of FGT poverty measures. The concluding section distills the main results.

2. Definitions

Suppose we have two income distribution functions, *F*(*y*) and *G*(*x*), and wish to compare the extent of inequality *across* the two distributions, rather than the relative inequality *within* the distributions, e.g., using Lorenz curves. For example, we could compare the portions of the distributions that lie below a given reference income. To make this comparison, it is convenient to pool the distributions and use quantile order statistics of the pooled distribution as common reference, or target, incomes. Note here that the proportion of income in the bottom quintile of distribution *F* or *G* considered *separately* – as in Lorenz comparisons – is different from the proportion of income in *F* or *G* that lies below the 20th percentile in the pooled distribution.

Let the range of possible target incomes (t), be $0 \leq t \leq z$, where *z* is the maximum income in the pooled distribution. Then we can express the h^{th} partial moment, corresponding to $y < t$, for the density function *f*(*y*) as

$$I(t;h) = \int_0^t y^h f(y)dy. \quad (1)$$

The corresponding (full) moment is given by $E(y^h) = \lim_{t \to z} I(t;h)$. The fraction of the h^{th} moment of *y*, corresponding to $y \leq t$, is given by

$$\varphi(t;h) = \frac{I(t;h)}{E(y^h)} \quad (2)$$

where $0 \leq \varphi(t;h) \leq 1$. Butler and McDonald (1987) call these expressions *normalized* incomplete (or partial) moments. By plotting combinations of these moments for different income distributions in the unit square (using common values of t), they created the notion of an interdistributional Lorenz curve (ILC).

3. Relations among Dominance Methods

FSD and Partial Moments of the First Type ($h = 0$)

Distribution F first-degree (stochastic) dominates (FSD) distribution G if and only if

$$F(y) \leq G(x) \;\Rightarrow\; \int_0^t f(y)dy \;\leq\; \int_0^t g(x)dx \quad (3)$$

for all t. We illustrate FSD in Figure 1. If $h = 0$, $\varphi_F(t;0) = \; I(t;0) = \; \int_0^t f(y)dy = F(y)$. To construct an ILC for $h = 0$, we plot $\varphi_F(t;0) = \int_0^t f(y)dy = F(y)$ and $\varphi_G(t;0) = \int_0^t g(x)dx = G(x)$ in the unit square using common income targets, as in Figure 2. To create the appearance of a Lorenz curve (lying below the diagonal between 0 and 1), we must put the "disadvantaged" group (with the distribution *G*(*x*)) on the horizontal axis. When equation (3) holds, the $h = 0$ ILC will lie below the diagonal. We summarize this result in the following proposition:

<u>Proposition 1</u>: FSD of G by F implies that $\varphi_F(t;0) < \varphi_G(t;0)$ for all t.

As Foster and Shorrocks (1988) demonstrate, FSD of *G* by *F* also implies that distribution *F* has lower headcount poverty (P_1) than distribution *G*. Hence, dominance of *G* by *F* with respect to the $h = 0$ ILC is equivalent to headcount poverty dominance of *G* by *F*.

SSD and ILCs of the Second Type ($h = 1$)

Distribution F second-degree (stochastic) dominates (SSD) distribution G if and only if

$$\int_0^t F(y)dy \;\leq\; \int_0^t G(x)dx \qquad (4)$$

for all t. When $h = 1$, $\varphi_F(t;1) = \frac{I(t;1)}{E(y)} = \frac{\int_0^t yf(y)dy}{\mu_F}$. Using integration by parts, we can rewrite the numerator of this expression as

$$\int_o^t yf(y)dy = yF(y)\,|_0^t - \int_0^t 1 \cdot F(y)dy = tF(t) - \int_0^t F(y)dy. \quad (5)$$

Similarly, we can use integration by parts to rewrite the denominator, $E(y) = \mu_F$, as

$$\mu_F = \int_0^z yf(y)dy = yF(y)\,|_0^z - \int_0^z 1 \cdot F(y)dy = z \cdot 1 - \int_0^z F(y)dy$$

$$= \int_0^z 1 \cdot dy - \int_0^z F(y)dy = \int_0^z [1 - F(y)]dy. \quad (6)$$

Hence, we can write

$$\varphi_F(t;1) = \frac{F(t)t - \int_0^t F(y)dy}{\int_0^z [1-F(y)]dy}. \quad (7)$$

Therefore, for distribution F in Figure 3, $\varphi_F(t;1) = \frac{Area\ A}{Area\ A+B}$.

From expression (7), we can see that SSD of *G* by *F* does not directly imply anything about the ILC for $h = 1$, due to the presence of the term $F(t)t$ in the numerator and of μ_F in the denominator. To explore matters further, we consider the special case $\mu_F = \mu_G$, and seek to demonstrate the following proposition:

> *<u>Proposition 2</u>: SSD of G by F implies neither $\varphi_F(t;1) < \varphi_G(t;1)$ nor $\varphi_F(t;1) > \varphi_G(t;1)$ for all t, even when $\mu_F = \mu_G$.*

Figure 4 presents two distribution functions, *F*(*y*) and *G*(*x*), drawn such that (1) they intersect only once, and (2) that area J equals area K. Thus, the areas below *F* and *G*, given by $\int_0^z F(y)dy$ and $\int_0^z F(x)dx$, are the same in both cases, which ensures SSD of *G* by *F* in Figure 4. Furthermore, the areas above *F* and *G*, given by $\int_0^z [1 - F(y)]dy$ and $\int_0^z [1 - G(x)]dx$, are the same, which implies that $\mu_y = \mu_x$. Therefore, the denominators of $\varphi_F(t;1)$ and $\varphi_G(t;1)$ in equation (7) are equal.

To evaluate the relative magnitudes of the numerators of $\varphi_F(t;1)$ and $\varphi_G(t;1)$, consider Figure 5, where t $> c$. Here it is clear that $F(t)t > G(t)t$. By the construction of *F* and *G*, $\int_0^t F(y)dy < \int_0^t G(x)dx$ for all t $< z$. Thus, $F(t)t - \int_0^t F(y)dy > G(t)t - \int_0^t G(y)dy$.

Given that $\mu_F = \int_0^t [1 - F(y)]dy = \int_0^t [1 - G(x)]dx = \mu_G$ by assumption, it follows from (7) that $\varphi_F(t;1) > \varphi_G(t;1)$ for $t > c$. We can illustrate this result in Figure 6a by comparing the areas A_F and A_G, corresponding to area A in Figure 3 and in the numerator of (7). In Figure 6a, A_F = J + K + L and A_G = K + M. By inspection, it is clear that J + L > M, which implies that $A_F > A_G$, and thus, that $\varphi_F(t;1) > \varphi_G(t;1)$.

But, we can also find a $t < c$ such that $\varphi_F(t;1) < \varphi_G(t;1)$. For example, consider Figure 6b, where A_F = J + K and A_G = K + L. By inspection, we can see that J > L, which implies that $A_F < A_G$, and thus, that $\varphi_F(t;1) < \varphi_G(t;1)$. Thus, for the distributions F and G in Figure 3, it follows that the $h = 1$ ILC will intersect the diagonal in the unit square, as in Figure 7, even though *F* SSD *G* by construction.

Before leaving this section, we note that Foster and Shorrocks (1988) have shown that SSD of *G* by *F* implies per capita income gap (P_2) dominance of *G* by *F*. Given that SSD does *not* imply ILC dominance for $h = 1$, there is no relationship in general between the second class of FGT poverty measures and ILCs of the second type.

4. Conclusion

We have shown that the step from first- to second-degree stochastic dominance is where the correspondence in orderings of income distributions by ILCs and stochastic dominance, or the class of FGT poverty measures, breaks down. Moreover, we attribute the breakdown to the "normalization" of the incomplete moments (i.e., dividing by the mean of the income distribution) used to plot ILCs.

It is well-known that in second-degree stochastic dominance or per-capita income gap poverty comparisons, a mean-preserving contraction is unambiguously preferred to the original distribution, because it reduces the degree of inequality *within* the altered distribution. In ILC comparisons, mean-preserving contractions in the distribution are not necessarily preferred, because the comparisons are drawn *between* distributions at fixed income levels in the original distribution. With a mean-preserving spread, fractions of persons or incomes in the altered distribution will be smaller at some income cutoffs in the original distribution, but larger at other income cutoffs. Hence, the outcome of the comparison will be ambiguous.

References

Bishop, J. A., K. V. Chow, and L. A. Zeager, (2004) "Lorenz Decomposition and Interdistributional Lorenz Comparisons," *Research on Income Inequality*, 12, 159-177.

Bishop, J. A., K. V. Chow, and L. A. Zeager, (2010) "Visualizing and Testing Convergence Between Two Income Distributions," *Journal of Income Distribution*, 19, 1, 2-19.

Butler, R. J., and J. B. McDonald, (1987) "Interdistributional Income Inequality," *Journal of Business and Economic Statistics*, 5, 13-18.

Deutsch, J., and J. Silber, (1997) "Gini's 'Transvariazione' and the Measurement of Inequality:

Deutsch, J., and J. Silber, (1999) "Inequality Decomposition by Population Subgroup and the Analysis of Interdistributional Inequality," in J. Silber (Ed.), *Handbook of Income Inequality Measurement* (Dordrecht: Kluwer) (pp. 363-397).

Foster, J., J. Greer, and E. Thorbecke (1984) "A Class of Decomposable Poverty Measures," *Econometrica,* 52, 761-776.

Foster, J. and A. F. Shorrocks (1988) "Poverty Orderings," *Econometrica*, 56, 1, 173-177.

Gini, C. (1916) Il Concetto di "Transvariazione" e le sue prime applicazioni," Studi di Economia, Finanza e Statistica, editi del Giornali degli Economisti e Revista de Statistica, Reprinted in Gini, 1959.

Gini, C. (1959) *Memorie de Metadologia Statistica: Volume Secondo – Transvariazione.* Rome: Libreria Goliardica.

Kakwani, N. C. (1980) *Income Inequality and Poverty: Methods of Estimation and Policy Applications*, New York: Oxford University Press.

Shorrocks, A. F. (1983) "Ranking Income Distributions," *Economica*, 50, 3-17.

Vinod, H. D. (1985) "Measurement of Economic Distance Between Blacks and Whites," *Journal of Business and Economic Statistics*, 3, 78-85.

Figure 1:
First-Degree Stochastic Dominance (FSD)

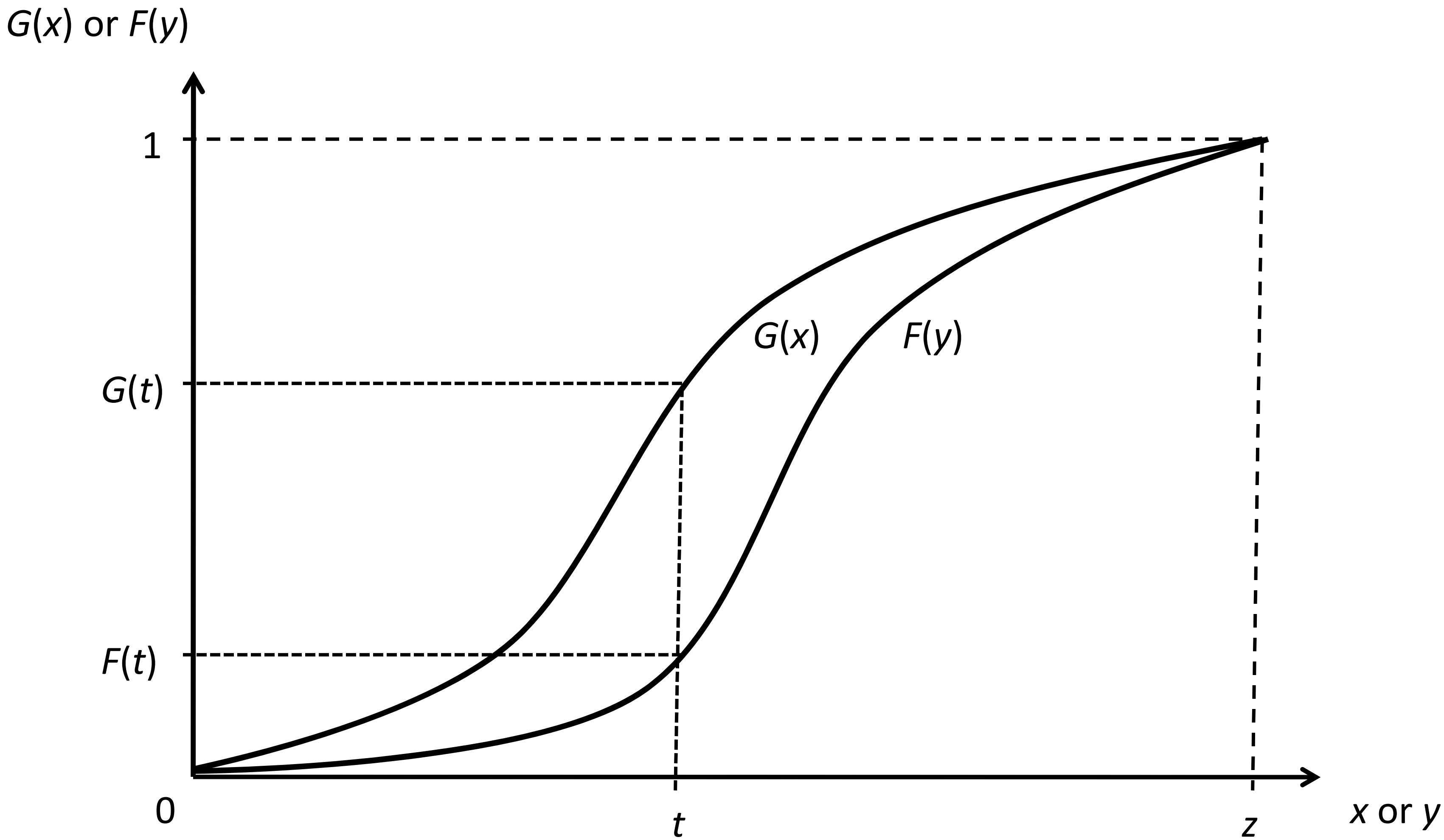

Figure 2
Interdistributional Lorenz Curve for *h*=0

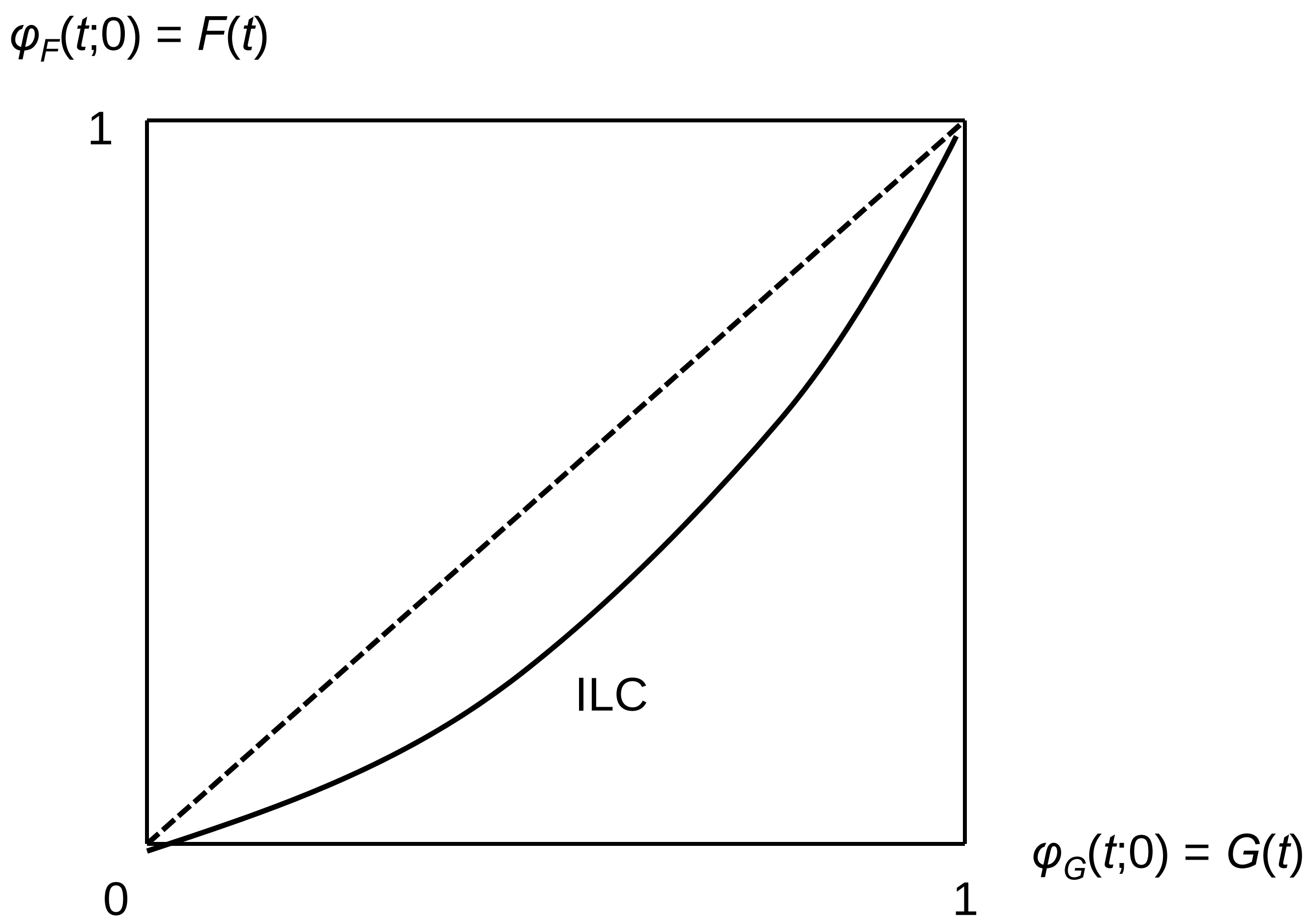

Figure 3:
Relationship Between $\varphi_F(t;1)$ and $F(y)$

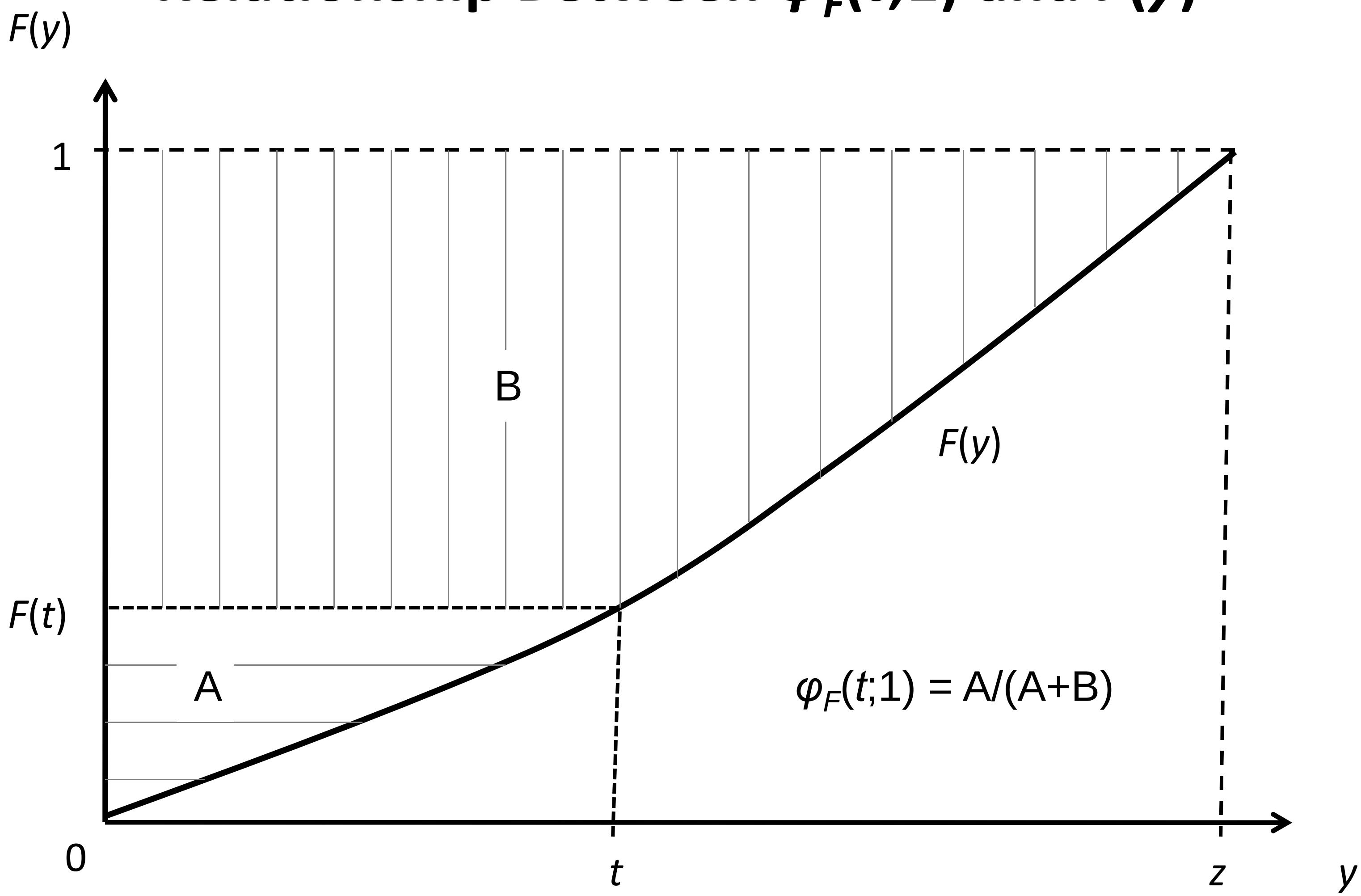

Figure 4:
Distributions with Equal Means

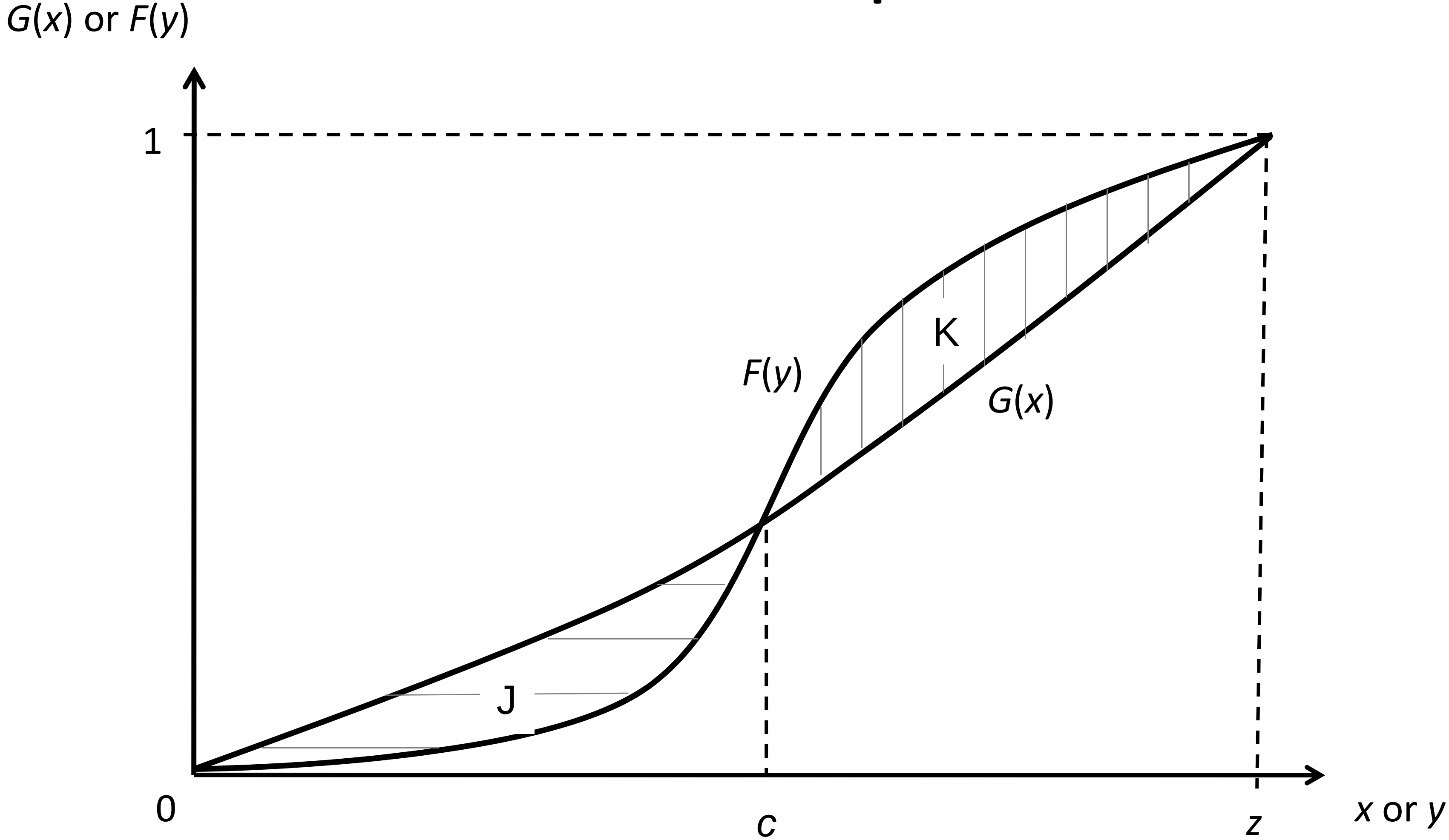

Figure 5:
Distributions with Equal Means

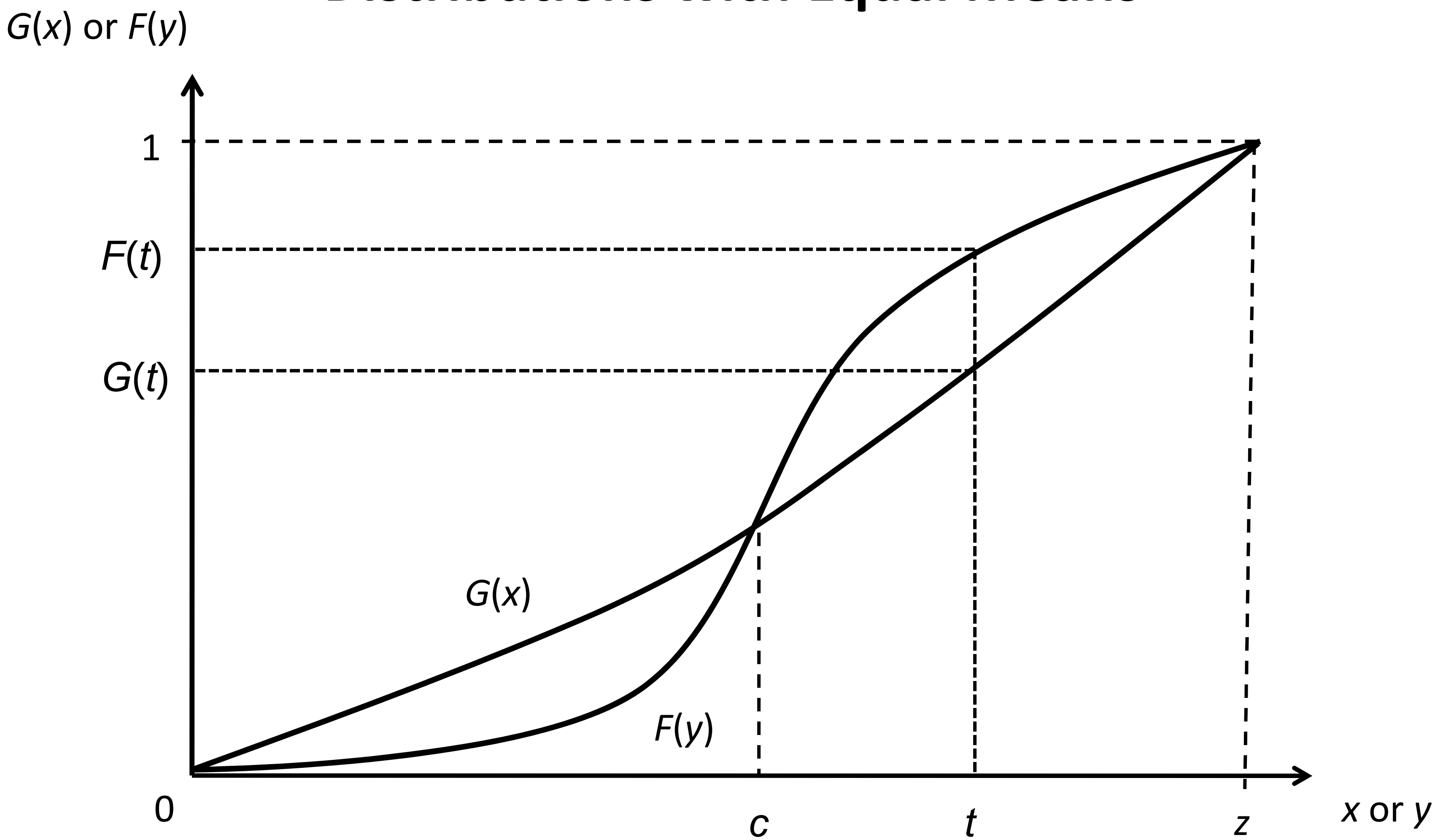

Figure 6a:
$\varphi_F(t;1) > \varphi_G(t;1)$ with Equal Means

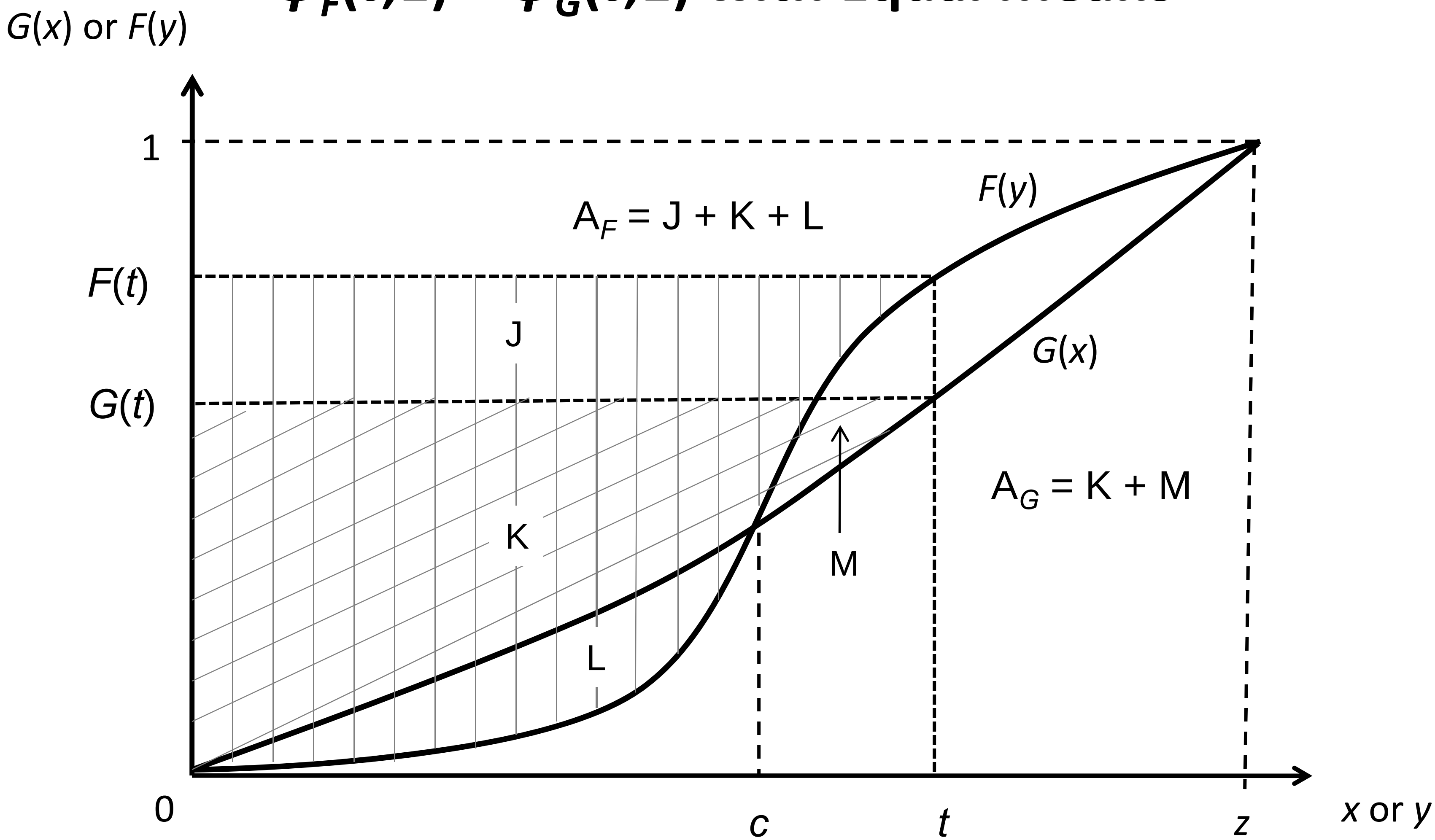

Figure 6b:
$\varphi_F(t;1) < \varphi_G(t;1)$ with Equal Means

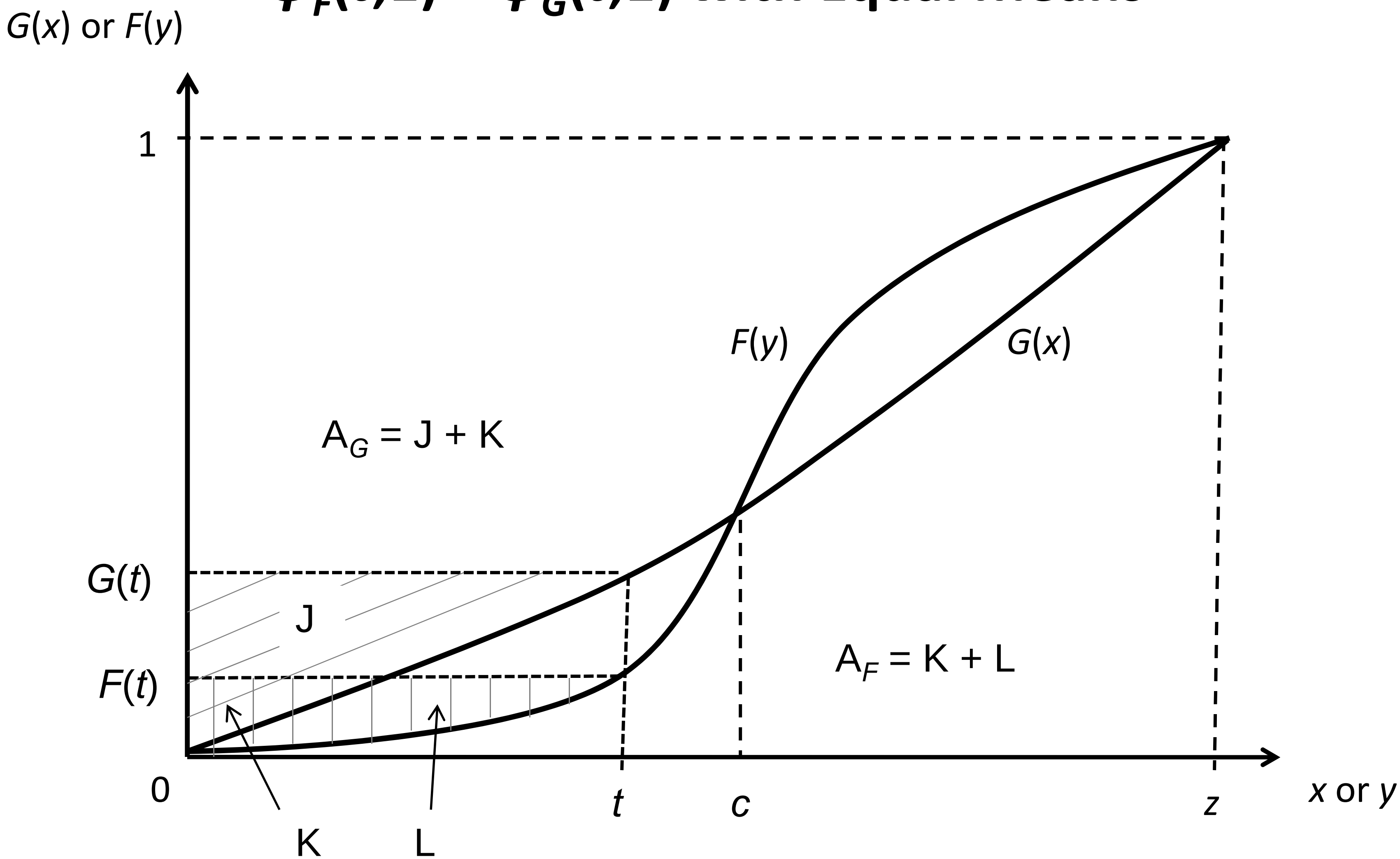

Figure 7
ILC Crossing for *h*=1