Review of Economics & Finance
Submitted on 31/12/2015
Article ID: 1923-7529-2016-02-13-15 Shlomo Yitzhaki

A Potential *Contradiction* Between Economic Theory and Applied Finance

Prof. Emeritus ***Shlomo Yitzhaki***
Department of Economics, Hebrew University, Jerusalem, 91905, ISRAEL
And
Hadassah Academic College, Jerusalem, 91010, ISRAEL
Tel: +972-506-235-240, E-mail: Shlomo.yitzhaki@huji.ac.il

Abstract: One of the basic premises that underlies economic theory in Finance is the assumption of declining marginal utility of income. This assumption imposes risk-aversion on the investors and is necessary requirement to an equilibrium capital markets. A popular method of analyzing empirical evidence among financial analysts is the Ordinary Least Squares Regression. This paper argues that in certain cases involving violation of the linearity assumption by the data, there may be a contradiction between the two approaches. In order to resolve the possibility of a contradiction one has to impose economic theory on the regression. The paper proposes the use of the Gini regression to bridge the gap between economic theory and regression.

Keywords: Stochastic dominance, OLS regression, Gini regression

JEL Classifications: C00, C50, C53, C58

1. Introduction

One of the fundamental principles in Finance is the assumption of risk aversion on the part of investors. This approach can be traced to economic theory, and more specifically to the assumption of declining marginal utility of income. Without this assumption, investors would not reach an equilibrium.

One of the major instruments used in applied Finance is the regression technique. This tool enables identification of correlations among different investment possibilities and facilitates choice of the appropriate combination of investments to minimize exposure to risk while maximizing expected return. The most popular regression technique is the Ordinary Least Squares, hereafter the OLS regression.

I argue that in some practical cases we find a contradiction between the assumptions required by economic theory and those required by the OLS regression, and that this contradiction may lead the user of the regression to arrive at an incorrect decision. To avoid such contradiction, researchers should use alternative regression techniques. All the regressions that are based on L_1 metric are not subject to this contradiction.[1]

[1] Among the regressions not subject to the described contradiction are: Gini, Quantile, and Least Absolute Deviation (LAD) regressions. However, as demonstrated in Yitzhaki and Lambert (2013a), Quantile and LAD regressions are special cases of the Gini regressions.

The paper is constructed as follows: Section 2 presents the role of economic theory in Finance. Section 3 addresses the basic assumptions that may cause the OLS regression to yield incorrect estimates of the regression coefficients, while Section 4 explains the possible contradiction between the OLS and economic theory. Section 5 presents the extended Gini estimators that do not contradict economic theory, in the sense that they impose economic theory on the regression. Section 6 examines the argument that since that we are dealing with equilibrium in capital markets, the rules of stochastic dominance do not apply. On the contrary, it is shown that, at equilibrium, the necessary conditions for dominance remain in place for every investor. Section 7 presents additional statistical arguments that demonstrate the advantages of using the Gini, while Section 8 presents empirical arguments. Section 9 concludes the article.

2. The Principles of Economic Theory

Economic theory is based on optimization of a target function subject to a budget constraint. To determine a local optimum the target function must be concave. At this point, the concept of *Stochastic Dominance* enters, and is applied to all investors. We distinguish between first-degree stochastic dominance (FSD) and second-degree stochastic dominance (SSD). The relevant concept for risk aversion is SSD (see Levy, 2006).

The definitions are as follows:[2]

Stochastic Dominance (SD): Let u(w) be the utility function defined over the wealth of the investor, w. The function u() is unknown to us, but we may assume general properties that apply to investors belonging to specific groups.

FSD (First Degree SD): The first assumption is that u' (w) > 0. This implies that investors prefer more wealth than less wealth.

We tend to assume that investors are risk averse. This assumption leads us to second degree stochastic dominance, SSD.

SSD: Assume that u'() >0; u''()< 0. That is, the marginal utility of wealth is positive but declines with wealth. The target of the investor continues to be expected utility. Given two risky investment opportunities A and B, with cumulative distribution functions F_A (w) and F_B (w) and Absolute Lorenz Curves (ALC), then $E\{U(w_A)\} > E\{U(w_B)\}$ iff $ALC_A(w) \geq ALC_B(w)$ for all w, with at least one w strict inequality. ALC is the *Absolute Lorenz Curve*, which is the familiar Lorenz curve without dividing the vertical axis by the mean. (See Yitzhaki and Schechtman (2013), pp- 76-81).[3]

We may introduce further assumptions regarding the utility function; each additional assumption imposes a further constraint on the assumed class of utility functions. The more assumptions imposed on the utility function, the smaller the efficient set.

In a market with many investors, necessary and sufficient conditions for FSD and SSD rarely exist. The analysis can thus be restricted so as not to contradict SSD. This leads us to an analysis based on necessary conditions for SSD. Unlike necessary and sufficient conditions for SSD that are based on intersections (or non-intersections) of curves, necessary conditions for SSD can be based on parameters. To derive these parameters, it is convenient to start from a decomposition of the

[2] In many practical cases wealth is substituted by income. Hence, we will not distinguish here between wealth and income.

[3] Other presentations of necessary and sufficient conditions for SSD can be found in Levy (2006).

concept "mean preserving spread" into its components.[4] (See Rothschild and Stiglitz (1970, 1971), Hanoch and Levy (1969)).

Mean preserving spread is based on at least two changes in the distribution of wealth: an increase and a decrease of the wealth, such that the mean of the distribution is maintained. The increase (decrease) in the return increases (decreases) the mean. However, the variability measure used may increase or decrease, depending on the shape of the distribution. An increase (decrease) in income increases (decreases) expected utility. The effect of an increase in variability depends on whether the increase occurs due to an increase in higher or lower income. In the former case, the increase in variability tends to mitigate the effect on expected utility because the increase in income has a lower marginal utility than the decrease. The opposite effect occurs if the increase occurs in the lower income.

A necessary condition for SSD comprises two conditions: Let *A*, *B* be two alternative distributions of wealth. Then

$$\mu_A \geq \mu_B \tag{1}$$

$$\mu_A - V_A \geq \mu_B - V_B \tag{2}$$

With at least one strict inequality, μ represents the mean and V the measure used to describe the variability of the distribution of wealth. Inequality (2) insures that the change in the variability measure does not contradict the effect on mean wealth. The Gini Mean Difference, the extended Gini,[5] and the mean Deviation move more slowly than the mean wealth, and can therefore be used to form necessary conditions for SSD. On the other hand, the variance, and consequently the standard deviation, cannot form a necessary condition for SSD unless the types of distribution are restricted.[6]

3. The OLS Regression

Regression is the most popular method used by analysts when dealing with several variables. The basic idea is the following:

The analyst assumes that the expected value of the dependent variable, Y, is a linear function of p independent variables plus a random, statistically independent residual of the independent variables. Formally,

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon \tag{3}$$

Where the β_i (i=0, 1, …, p) are given constants while X_i are random variables, and ε is statistically independent of the Xs. The analyst uses an algorithm that minimizes the variance of the residuals. The estimators attained are called BLUE – i.e., Best Linear Unbiased Estimators. The assumptions leading to (3) may be violated by the data. For example, there is no guarantee that the data form a linear model or that the residuals are statistically independent of the independent variables.

[4] See Rothschild and Stiglitz (1970, 1971) and the comments by Leshno, Levy, and Spector (1997) and Kanniainen (2007).

[5] See Yitzhaki (1982) for the Gini, and Donaldson and Waymark (1983) and Yitzhaki (1983) for the Extended Gini. See Hanoch and Levy (1969) for criticism of the use of the mean-variance criterion in Finance.

[6] See Yitzhaki (1982) and Levy (2006).

Yitzhaki (1996) demonstrated an alternative interpretation of the results of the process of estimation of the regression coefficients. To explain it we need to distinguish between simple and multiple regressions. We start with the simple regression, i.e., with one independent variable.

Assume that the data consist of N observations, (Y_j , X_j), j= 1,…,N. Let us arrange the observations in increasing order of X, the independent variable.[7] Then define the slopes between adjacent observations of X as

$$b_j = (Y_{j+1} - Y_j)/(X_{j+1} - X_j),\ j=1,\ldots, N-1 \qquad (4)$$

The OLS estimator of the regression coefficients is shown to be a weighted average of the slopes defined between adjacent observations of X. Therefore, the OLS estimator in the simple regression case is a weighted average of the slopes defined between adjacent observations of X. Formally,

$$b_{OLS} = \sum_{j=1}^{N-1} w_j\, b_j \qquad (5)$$

We now use (5) to extend the interpretation to the multiple regression case.

Assume that we estimate all possible simple regression coefficients as: b_{ij} i=0,1,…,K , j=1, …, K, with 0 representing the dependent variable. Then, the regression coefficients in the multiple regression case can be interpreted as the solutions of a set of equations with b_{ij} (j=1,…, K) serving as the parameters in each equation, while b_{0i} (i=1,...,K) serve as constants. In the case of an OLS or Gini regression, the set of equations is linear.

The advantage of this interpretation is that regression techniques can be mixed in the same multiple regression, with several equations derived by OLS and others by Gini or the extended Gini.[8] In this way, the effect of the regression technique on the sign and magnitude of the estimates can be evaluated. However, this deviates from the main point of the paper.

4. A Possible Contradiction between OLS Regression and Expected Utility [9]

The weighting scheme of the slopes defined between adjacent observations is determined by the distribution of the independent variable. In the case of OLS and Gini regressions, the weighting scheme is determined by the appropriate Lorenz curve of the independent variable. As shown in Yitzhaki (1996), in the case of the Gini regression the weighting scheme is determined by the Lorenz curve of the independent variable; while for the OLS, the appropriate Lorenz curve is the one in which X instead of F(X) is depicted on the horizontal axis. In both cases, the weighting scheme is determined by the distribution of the independent variable.

To understand the possible contradiction between economic theory, (i. e., SSD), and the conventional method of estimation, (i.e., the OLS regression), let us evaluate the weighting schemes of different distributions of the independent variable: the uniform, the normal, and the lognormal.

This section investigates the relationship between the specific distribution of the independent variable and the weighting scheme of the OLS and Gini estimators. Let Y, X be two continuous

[7] To simplify the presentation we assume no ties in X.

[8] Other regression techniques, like Quantile regression, can be mixed in the same regression. However, the set of equations then ceases to be linear.

9 This section is based on Yitzhaki (1996).

random variables with a density function f(y, x). Let f_X, F_X, μ_X and σ^2_X denote the marginal density, the marginal cumulative distribution, the expected value and the variance of X, respectively. Assume that first and second moments exist. Let $g(x) = E\{Y \mid X=x\}$ be the regression curve, where $g'(x)$ represents its local slope.[10]

Proposition 1: Let $E^*(Y|X) = \alpha + \beta X$ denote the best linear predictor of Y given X. Then β is a weighted average of the slopes of the regression curve:

$$\beta_{OLS} = \int w(x)g'(x)dx \tag{6}$$

where $w(x) > 0$ and $\int w(x)dx = 1$.

The weights are:

$$w(x) = \frac{F(x)}{\sigma_X^2}(\mu_X - E\{X|X \leq x\} \tag{7}$$

Proof: See Yitzhaki (1996) or Yitzhaki and Schechtman (2013).

Since w(x) is a function of the distribution of the independent variable, there are two possible presentations of the weighting scheme. The first presents weight as a function of the independent variable, while the second presents weight as a function of F_X, the cumulative distribution of the independent variable. The second presentation is useful when focusing on the share of the weights assigned to portions of the population. One presentation can be transformed into the other by defining a weighting scheme, $v(p) = w[x^{-1}(p)]$, $(0 \leq p \leq 1)$, where $x^{-1}(p)$ is the inverse of the cumulative distribution. ($F_X(x)$ is a monotonic increasing differentiable function).

To illustrate the effect of the distribution of the independent variable on the weighting scheme let us consider three specific examples: the uniform, the normal, and the log-normal. The first two illustrate interesting cases; the third resembles the distribution of income.

(a) *The Uniform Distribution*. Let X be uniformly distributed between [a, b]. Applying (7), the weight attached to the slope at income x is

$$w(x) = \frac{(b-x)(x-a)}{2(b-a)} \tag{8}$$

This weighting scheme in (8) is symmetric around the median, and the closer the observation to the median, the greater the weight. An interesting feature of this weighting scheme is that its shape remains unchanged regardless of whether it is viewed as a function of x or as a function of F. As a result, the weighting schemes of the OLS and Gini regressions are identical. (Olkin and Yitzhaki (1992).

(b) *The Normal Distribution*. Let X be distributed according to the standard normal. By (5),

$$w(x) = -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} t\, e^{-t^2/2}\, dt = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \tag{9}$$

The weight here is equal to the density of the normal distribution. Hence, each decile of the population receives an equal weight of 10 percent.

(c) *The Log-normal Distribution*. Let X be log-normally distributed, with parameters μ and σ.

[10] Proposition 1 is defined in terms of the population parameters. We may concentrate on the population parameters because the OLS estimators converge to this parameter.

$$w(x) = \frac{\mu_x}{\sigma_x^2}[F_x(x) - F_1(x)] = \frac{F_x(x)}{\sigma_x^2}[\mu_x - E(X \mid X \le x)] \qquad (10)$$

where μ_X, σ^2_X are the expected value and the variance of the log-normal distribution, while $F_1(x) = (1/\mu_X)\,{}_0\int^x t\, dF_X(t)$ is the first moment cumulative distribution. By Theorem 2.6 in Aitchison and Brown (1957, p. 12), the first moment distribution is also log-normal with parameters $\mu + \sigma^2$ and σ^2. Hence, the weight at x is the difference between two cumulative log-normal distributions. Using the usual transformation we can write the weight as:

$$w(x) = \frac{\mu_x}{\sigma_x^2}\left[\Phi\left(\frac{\log(x - \mu)}{\sigma}\right) - \Phi\left(\frac{\log(x - \mu - \sigma^2)}{\sigma}\right)\right] \qquad (11)$$

where Φ**(.)** is the cumulative standard normal. This term can be evaluated numerically.

Table 1 presents the weighting scheme w(x) when X is log-normally distributed, for different values of the parameters of μ and σ. As can be seen, the weighting scheme is not sensitive to μ (note the difference between columns 4 and 5), but is sensitive to σ. For ease of reference, Table 1 also presents the Gini coefficient corresponding to each value of σ. As seen in column 4, if the Gini coefficient is about 0.4 (a typical value for before-tax income), the expected weight of the top decile is around 45 percent; for the highest quintile it is over 60 percent. If wealth were used as an independent variable, a Gini coefficient of 0.55 could be considered typical. In this case, the weight of the top decile may well exceed 60 percent. Experiments with other distributions (the Pareto and the exponential) show that the share of the top decile is never less than 30 percent. While risk aversion requires the attribution of higher weights to the lower deciles of wealth, in the case of asymmetric distributions of the independent variable the OLS tends to impose higher weights on the top deciles. Hence, if the linearity assumption is violated by the data, the OLS may contradict economic theory.

Table 1 The OLS weighting schemes for a log-normal distribution

Parameters								
μ	0.	0.	0.	0.	7.0	0.	0.	0.
σ	0.45	0.6	0.7	0.75	0.75	0.8	1.0	1.2
Gini coef.	0.25	0.33	0.38	0.4	0.4	0.43	0.52	0.60
Column No.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Decile								
1	2.5	2.1	1.0	0.8	0.8	0.7	0.3	0.1
2	3.9	3.4	1.9	1.6	1.6	1.4	0.7	0.3
3	5.0	4.5	2.8	2.4	2.5	2.1	1.1	0.5
4	6.0	5.5	3.7	3.3	3.3	2.9	1.7	0.8
5	7.1	6.6	4.8	4.3	4.4	3.9	2.4	1.3
6	8.8	8.3	6.5	6.0	6.0	5.5	3.6	2.1
7	10.2	9.9	8.3	7.8	7.9	7.3	5.2	3.3
8	12.6	12.5	11.4	11.0	11.1	10.5	8.2	5.7
9	16.1	16.4	16.9	16.7	17.0	16.5	14.4	11.2
10	27.9	30.4	42.8	46.0	45.3	49.3	62.6	74.7

5. The Weighting Scheme of the Gini (and Extended Gini) Regressions

In this section the weighting scheme of the Gini regression and mixed Gini – extended Gini are presented. As in the previous subsection, the presentation is restricted to the same distributions of the independent variable:

The (extended) Gini regression coefficient (EGRC) is defined as:

$$b_{NP}(\nu) = \frac{\text{cov}\,(\,Y, [1 - F_x(X)\,]^{\nu-1})}{\text{cov}(X, [1 - F_x(X)\,]^{\nu-1})}, \qquad \nu > 0, \quad \nu \neq 1 \tag{12}$$

where ν is a parameter determined by the investigator.

The interpretation of ν in the extended Gini in Finance is that it reflects the risk aversion parameter of the investor. If ν equals one, the investor is indifferent to risk. The higher the ν, the more risk-averse is the investor. In the extreme case ($\nu \to \infty$) the investor cares only about the worst result (max-min attitude). The range $0 \leq \nu < 1$ reflects a love of risk, with $\nu \to 0$ meaning a max-max strategy – that is, the investor maximizes the best outcome.

The weighting scheme is determined by the parameter ν and, of course, by the distribution of the independent variable. By determining ν the investigator introduces his/her attitude to risk into the estimation procedure.

All estimators defined in (12) are based on the extended Gini variability index: the denominator is the extended Gini variability index and the numerator is the extended Gini covariance.

Proposition 2: The extended Gini estimators of the regression coefficient have the following properties:

(a) In the population, the parameters are weighted averages of the slopes of the regression curve. That is,

$$\beta_{NP}(\nu) = \int w(x,\nu) g'(x) dx \tag{13}$$

with $w(x, \nu) > 0$ and $\int w(x, \nu) dx = 1$, where

$$w(x, \nu) = \frac{[1 - F_x(x)] - [1 - F_x(x)]^{\nu}}{\int_{-\infty}^{\infty} [[1 - F_x(t)] - [1 - F_x(t)]^{\nu}] dt} \tag{14}$$

(b) In the sample, all estimators are weighted averages of slopes defined by pairs of adjacent observations.

$$b_{NP}(\nu) = \sum_{i=1}^{n-1} w_i b_i \tag{15}$$

where $b_i = \dfrac{y_{i+1} - y_i}{x_{i+1} - x_i}$ $(i = 1, ..., n-1)$; $w_i > 0$, $\Sigma \nu_i = 1$, and

$$w_i = \frac{[n^{\nu-1}(n-i) - (n-i)^{\nu}]\,\Delta x_i}{\sum_{k=1}^{n-1} [\,n^{\nu-1}(n-k) - (n-k)^{\nu}\,]\Delta x_k} \tag{15'}$$

(c) Te estimators $b_{NP}(\nu)$ are ratios of U-statistics. As such, they are consistent estimators of $\beta_N p(\nu)$; for large samples, the distributions of the estimators converge to the normal distribution.

(d) Suppose that $E(Y|X) = \alpha + \beta X$ and $Var(Y|X) = \sigma 2 < \infty$, then all extended Gini estimators are consistent estimators of β.

Proof: Yitzhaki (1996).

Table 2 presents the value assigned at the mid-observation of each decile, while Table 3 presents the weights relative to the first decile for different values of ν. Note, however, that the share of the weight assigned to each decile depends also on the distribution of the independent variable.

Table 2 The weight at the mid-decile for different values of EG parameters

F	1-F	υ=2	υ=3	υ=4	υ=5
0.05	0.95	0.95	0.9025	0.857375	0.814506
0.15	0.85	0.85	0.7225	0.614125	0.522006
0.25	0.75	0.75	0.5625	0.421875	0.316406
0.35	0.65	0.65	0.4225	0.274625	0.178506
0.45	0.55	0.55	0.3025	0.166375	0.091506
0.55	0.45	0.45	0.2025	0.091125	0.041006
0.65	0.35	0.35	0.1225	0.042875	0.015006
0.75	0.25	0.25	0.0625	0.015625	0.003906
0.85	0.15	0.15	0.0225	0.003375	0.000506
0.95	0.05	0.05	0.0025	0.000125	6.25E-06

Table 3 The weight assigned to mid-deciles relative to lowest mid-decile

decile \ v	2	3	4	5
1	1.000000	1.000000	1.000000	1.000000
2	0.894737	0.800554	0.716285	0.640887
3	0.789474	0.623269	0.492054	0.388464
4	0.684211	0.468144	0.320309	0.219159
5	0.578947	0.33518	0.194052	0.112346
6	0.473684	0.224377	0.106284	0.050345
7	0.368421	0.135734	0.050007	0.018424
8	0.263158	0.069252	0.018224	0.004796
9	0.157895	0.024931	0.003936	0.000622
10	0.052632	0.00277	0.000146	7.67E-06

6. The Market's Equilibrium

A possible argument against reliance on rules of stochastic dominance is that in a market with heterogeneous investors the rules of stochastic dominance do not apply. This issue has been handled in Shalit and Yitzhaki (2010) who discussed the role of beta in a market with investors who have different risk aversions. To simplify the presentation we follow Shalit and Yitzhaki (2009) in applying the analysis to two types of investor. Market equilibrium is presented in an Edgeworth's box, as used in Price theory. To imitate consumption commodities, the commodities are defined as goods. This

means that one "commodity" is mean return, while the other "commodity" is the subjective risk-adjusted return of each investor.

Result 1: If the investors have identical preferences toward risk, then the market portfolio will be on the diagonal, (the exact point depends on the inequality in wealth).[11] If, on the other hand, the investors have different attitudes toward risk, then the contract curve will never cross the diagonal, since the less risk averse investor invests more in risky assets. [12] If, for example, the investors are Mean-Variance, or Mean-Gini or Mean-Extended Gini investors, then the expansion curve is linear, such as O_AE or O_BE.

The market price for risk is pp (tangent to the indifference curves that pass at E in Figure 1). The important point from our point of view is that at E, each investor obeys the rules of stochastic dominance, which refutes the argument that since we are dealing with market equilibrium, the rules of stochastic dominance do not necessarily apply.

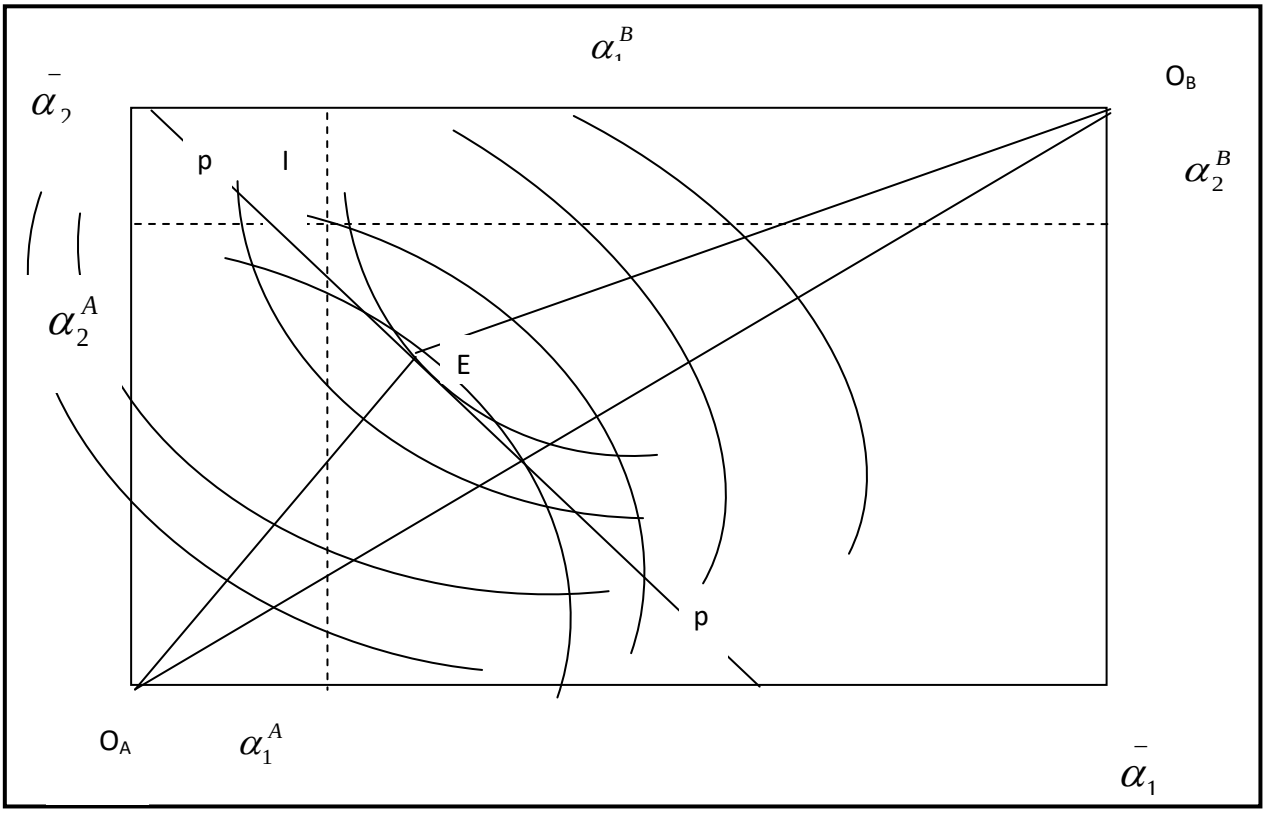

Figure 1 Capital market equilibrium with heterogeneous investors

This price ratio defines a unique equilibrium. Because the equal extended Gini curves are not identical, the optimal expansion paths for type A and type B investors are different. Therefore, the ratio of risky asset optimal portfolio held by each type of investor is different, and no investor will hold the "market portfolio" that is represented by the slope of the diagonal of the Edgeworth box. In other words, the equilibrium is expressed as:

[11] This is the basic assumption in Mean-Variance CAPM.

[12] In Figure 1, if we define the horizontal axis as representing the risk-adjusted return, then investor B is the less risk averse investor.

$$\left.\frac{\alpha_1}{\alpha_2}\right|_A \neq \left.\frac{\alpha_1}{\alpha_2}\right|_B \neq \frac{\bar{\alpha}_1}{\bar{\alpha}_2}$$

This leads us to the second result:

Result 2: Unless risky asset returns are all multivariate-normal, at equilibrium heterogeneous extended Gini investors hold different portfolios of risky assets and no one necessarily holds the market portfolio as expressed by the slope of the diagonal of the Edgeworth box.

The contract curve is the locus of all non-dominated equilibria following various initial endowments. Income distribution is generated by the relative size of the investors' initial endowments. From welfare economics analysis we draw the following two results:

Result 3: As the extended Gini is homogeneous of degree one in asset shares, the contract curve is either identical to the diagonal of the Edgeworth box or lies on one side of the diagonal.[13]

This result implies that once a certain type of investor tends to invest relatively more in one asset, it will continue to do so under all market circumstances. (This conclusion applies to all types of investor). Thus, it is possible to identify and relate types of asset with classes of investor.

Result 4: Expected returns on assets depend directly upon the income distribution across types of investor.

In some sense this result takes us back to traditional microeconomic theory that asserts the significance of income distribution when consumers have different tastes. Yet, this result clearly contradicts the CAPM, which claims that asset returns are determined solely by the demand of a representative investor.

Result 5: Heterogeneous investors who have the same ν will hold an identical portfolio of risky assets.

The mean-extended Gini model has been shown to be richer than the mean variance in that it enables the researcher to construct an infinite number of "capital asset pricing models" for ν homogeneous markets. Shalit and Yitzhaki (1984, 1989) showed that if investors have the same degree of risk aversion, one can estimate capital asset pricing model betas for every ν and then, using the holding of the market portfolio, find the ν that fits the data best. The heterogeneous model with many ν differs considerably from these results as conditions (i.e., point E in Figure 1) establish specific equilibrium relations between asset returns and risk as viewed by all investors in the market.

Finally, one may easily extend the arguments of this section to more than two types of investors. However, the graphical presentation of Figure 1 would be lost.

7. Additional Statistical Arguments

There are several further statistical arguments that justify the use of Gini regression:

(a) Given a linear combination of several variables, $y = \sum_{i=1}^{N} \beta_i x_i + \varepsilon$, one can view the regression as the decomposition of the variability measure of y into the contributions of different independent variables. Decomposition of the Gini Mean Difference, hereafter GMD, includes

[13] This result is derived from the homogeneity property of the isoquants. The contract curve cannot cross the diagonal, as it can only be the diagonal itself or lie on one side of it.

the decomposition of the variance as a special case. In particular, the GMD has two asymmetric correlation coefficients defined between two variables. These correlation coefficients will be equal if the distributions of the two variables differ only by a linear transformation. On the other hand, Shih and Huang (1992) have shown that unless the marginal distributions can differ only in their location and scale parameters, the range of Pearson's correlation can be narrower than [-1, 1]. For example, for lognormal distributions the range of Pearson's correlation is [-.368, 1](De Veaux, 1976). In addition, Denuit and Dhaene (2003) present an example in which Pearson's correlation converges to zero, although the variables are monotonic increasing functions of each other. (See Yitzhaki and Schechtman, 2013, pp-36-37).

(b) The use of transformations:

Econometricians tend to freely apply monotonic transformations to variables. For example, no one would object to the application of a log transformation. As pointed out by De Veaux (1976), this application may change the Pearson's correlation drastically.

(c) The symmetry assumption of the Pearson's correlation:

It is illustrated in Yitzhaki (2015) that the symmetry assumption imposed by cov(x, y) = cov(y, x) may change the sign of a regression coefficient.

(d) In a multiple regression, Yitzhaki and Schechtman (2013, ch-8) interpret the Gini and OLS regressions coefficients as solutions of sets of linear equations with the simple regression coefficients serving as parameters. Applying a monotonic transformation to an independent variable may affect a line and a column of the in an OLS regression, but only a line in a Gini regression.[14] In addition, the cumulative distribution, a component in the covariance in a Gini regression, is not affected by the monotonic transformation. As a result, the Gini regression seems to be more robust with respect to transformations.

(e) In applied finance, the researcher is obliged to use time series:

As pointed out by Wodon and Yitzhaki (2006), the Gini regression is richer than the OLS since the existence of two correlation coefficients implies that looking forward in time may produce a different estimate than looking backward.

(f) Schaffer (2015) has written a STATA program that has the following properties:

(f.1) It can run pure Gini regression or mixed regression. Mixed regression enables the researcher to apply either Gini or OLS or extended Gini to each independent variable. By switching methods one can detect the effect of the regression method on the estimates.

(f.2) The program includes estimation of the LMA curve, which can inform the user whether a monotonic transformation can change the sign of an OLS simple regression coefficient. [15]

[14] To observe this, let us consider the following solution to an OLS regression: $\beta = (X'X)^{-1}X'Y$, then because of the symmetry imposed on the covariance a transformation affects both a line and a column in the inverted matrix, while in a Gini regression only a line is affected.

[15] See also, Yitzhaki and Lambert (2013b), who show that in a multiple-period model, higher variance leads to higher expected return.

8. Empirical Evaluation

The arguments in favor of using the Gini regression in applied finance were of a theoretical nature, and to be effective, one of the assumptions made to obtain the regression estimates must be violated. Therefore, to justify the use of the Gini regression, it is crucial to evaluate the effect of violations when not using the Gini regression on the estimates.

To obtain conservative estimates, we use monthly data and grouped assets that are collected by Ibbotson Associates and can be downloaded from the Internet. The monthly returns we used span the period from January 1926 to mid-2014, which produced a total of 1063 observations. The grouped assets are the following:

1. S&P 500 TR (Total Return) – denoted by LSTK- mean =.0095479; S.E.= .001677.
2. U. S. Small Stock TR – denoted by SSTK – mean= .0129887; S. E. = .0025407.
3. U. S. Long Term corporate Bond – denoted by CBND - mean = .0051175; S. E.= .0006605.
4. U. S. Government Long Term Bonds –denoted by GVBD- mean= .0048086; S. E.= .0007403.
5. U.S. Government Intermediate Bond – denoted by ITGV - mean = .0043875; S.E.= .0003877
6. Short Term Treasury 30 days Bills – denoted by TBIL – mean = .0028575; S.E.= .0000776.
7. Finally we have constructed a portfolio the return of which is the average return in every period.[16]

Table 4a and 4b present the estimates of betas of the different assets and different assumptions of risk aversion by the investor. The Gini and EG (v) risk aversion have an intuitive interpretation: the risk aversion represents an investor who behaves as if he receives the smaller of two independent draws from the distribution of returns, while EG (5) represents an investor who behaves as if he receives the smallest of five independent draws.

Since the Gini and EG (v) have two correlation coefficients defined between two random variables, there are three statistics that substitute for R^2 : GR= $\frac{cov\,(e,F(e))}{cov\,(y,F(y))}$; $\Gamma(y,\hat{y}) = \frac{cov(y,F(\hat{y}))}{cov(y,F(y))}$; $\Gamma(\hat{y},y) = \frac{cov((\hat{y},F(y))}{cov(\hat{y},F(\hat{y}))}$; with $\hat{y}$ being the predicted value of the dependent variable.

Finally, Since R^2 is expressed in square units, to enable comparisons with the Gini, Tables 4a and 4b report the square root of R^2.

Concentrating on Table 4a, we note that different regressions result in different estimates of beta, but what is surprising is that we cannot rank the betas according to risk aversion. To see this, the beta of S&P companies is EG(5) > OLS > Gini; for small companies it is OLS> EG(5) > Gini, while for corporate bonds the ranking is Gini > EG(5) > OLS. Virtually the same pattern emerges in Table 4b: for GVBD we get Gini > EG(5) > OLS ; for ITGV we get Gini > EG(5) > OLS, while for TBIL the ranking is OLS > Gini > EG(5). Note, however, that except for TBIL the rankings are not statistically significant. The financial implications are that because regression curves tend to deviate

[16] It is worth restating that we have created the worst case for the Gini regression. By using aggregate assets we have increased the share of each asset in the portfolio and therefore we have increased the correlation between the assets' returns and the portfolio, while the use of monthly data rather than daily data reduces the fluctuations in the returns. Moreover, the higher the average return on the asset, the higher the bias.

from linearity in no particular order, then although risk averse investors are assumed to stress different portions of the regression curve according to a parameter, the results tend to differ from those expected. An alternative explanation could be that in a market with different types of risk averse investor, the distribution of wealth invested in a market does matter. As in any market in which buyers have different tastes, the income distribution and prices are simultaneously determined.

Table 4a Different betas for different risk aversion parameters: private sector

Regression	Asset	Constant	Beta	$\sqrt{R^2}$	$r(Y, \hat{Y})$	$r(\hat{Y}, Y)$	GR
OLS	LSTK	-.0035 (0008)	1.97 (.031)	.885			
Gini		-.0033 (.0008)	1.95 (.036)		0.88	.879	0.90
EG(5)		-.004 (.0008)	2.045 (.039)		.88	.879	.924
OLS	SSTK	-.007 (.001)	3.01 (.046)	0.895			
Gini		-.006 (.001)	2.86 (.05)		.882	.873	.882
EG(5)		-.006 (.001)	2.92 (.057)		.882	.83	.901
OLS	CBND	.002 (.0006)	.43 (.023)	.243			
Gini		.002 (.0006)	.495 (.027)		.575	.503	.575
EG(5)		.002 (.0006)	.44 (.029)		.575	.503	.511

Table 4b Different betas for different risk aversion parameters: government

Regression	Asset	Constant	Beta	$\sqrt{R^2}$	$r(Y, \hat{Y})$	$r(\hat{Y}, Y)$	GR
OLS	GVBD	.002 (.0007)	.4 (.028)	.165			
Gini		.0017 (.0007)	.48 (.03)		.48	.426	.48
EG(5)		.002 (.0007)	.412 (.034)		.48	.426	.41
OLS	ITGV	.003 (.0003)	.187 (.015)	.132			
Gini		.003 (.0004)	.22 (.167)		.43	.371	.43
EG(5)		.003 (.0004)	.19 (.018)				
OLS	TBIL	.003 (.00008)	.003 (.003)	0.0008			
Gini		.003 (.00008)	.002 (.004)		.02	.04	.02
EG(5)		.0029 (.00008)	.0017 (.0039)		-.02	-035	.03

9. Conclusions

The possible contradiction between economic theory and empirical Finance can be traced to the assumption of the linearity of the regression curve, and to basing the OLS regression on the properties of the variance. To show a contradiction, at least one assumption must be violated by the data. As shown in Yitzhaki (1996) and in Yitzhaki and Schechtman (2013), the algorithm of estimation of the regression coefficient ignores the linearity assumption and the estimated regression coefficient can be presented as a weighted average of slopes of the regression curve economic theory assumes a concave target function, while the OLS uses a quadratic (i.e., convex) function as the target function of estimation. If the distribution of the independent variable happens to be normal, then the OLS is the appropriate estimation method. In such a case each decile of observation receives equal weight, so that the OLS is the most efficient procedure of estimation. However, if the assumption of normality of the distribution of the independent variable is violated by the data, and the distribution is asymmetric to the right, as is typical in the case of income or wealth, then the OLS is not compatible with economic theory.

On the other hand, as shown in Shalit and Yitzhaki (2010), the use of Gini regression is compatible with SSD, which means that the use of Gini regression to estimate the beta coefficients of different investment opportunities does not violate economic theory. Use of the extended Gini regression means imposing the degree of risk aversion on the regression method. This insures that the weighting scheme implied by economic theory is compatible with the weighting scheme used by the analyst.

To summarize the argument: if the linearity assumption is violated and the distribution of the independent variable is positively skewed, then the OLS is not compatible with economic theory, because it assigns a large weight to the sections of the distribution of wealth that the risk-averse investor least cares about.

The empirical section indicates that it is impossible to predict the direction of the bias caused by the use of an OLS regression.

The use of the Gini or extended Gini regression avoids this contradiction. In such a case the analyst imposes a parameter of risk aversion on the income variable. A program that can estimate mixed OLS, Gini and Extended Gini has been written by Mark Schaffer in STATA and can be downloaded from the Internet.

Acknowledgement: I am grateful to a referee and Haim Shalit for helpful comments.

References

[1] Aitchison J., and J. A. C. Brown (1957). *The Lognormal Distribution*. Cambridge: Cambridge University Press.

[2] Donaldson D., and J. A. Waymark (1983). "Ethically Flexible Gini Indices for Income Distribution in the Continuum", *Journal of Economic Theory*, 29(2) : 353-358.

[3] Hanoch, G., and H. Levy (1969). "The Efficiency Analysis of Choices Involving Risk", *The Review of Economic Studies*, 36(3): July, 335-346.

[4] Leshno, M., Levy, H., and Y. Spector (1997). "A Comment on Rothschild and Stiglitz's "Increasing Risk: I. A Definition"', *Journal of Economic Theory*, 77(1): 223-228.

[5] Levy, H. (2006). *Stochastic Dominance: Investment Decision Making under Uncertainty.* (2nd Edition), Springer.

[6] Kanniainen, J. (2007). "Rothschild–Stiglitz's definition of increasing risk and the relationship between volatility and risk premium", *Review of Financial Economics*, 16(4): 363-374.

[7] Olkin I., and S. Yitzhaki (1992). "Gini Regression Analysis", *International Statistical Review*, 60(2): August, 185-196.

[8] Rothschild, M., and J. E. Stiglitz (1970). "Increasing Risk I: A definition", *Journal of Economic Theory*, 2(3): 225-243.

[9] Rothschild, M. and J. E. Stiglitz (1971). "Increasing Risk II: Its Economic Consequences", *Journal of Economic Theory*, 3(1): 66-84.

[10] Schaffer, M E. (2015). *GINIREG: Stata module for Gini Regression*. [Online] Available at https://ideas.repec.org/c/boc/bocode/s457958.html

[11] Shalit, H. and S. Yitzhaki (2009). "Capital market equilibrium with heterogeneous investors", *Quantitative Finance*, 9(6): 757-766.

[12] Shalit, H. and S. Yitzhaki (2010). "How does beta explain stochastic dominance efficiency?", *Review of Quantitative Finance and Accounting*, 35(4): 431-444.

[13] Yitzhaki, S. (1983). "On an Extension of Gini Inequality Index", *International Economic Review*, 24(3): 617-28.

[14] Yitzhaki, S. (1996). "On Using Linear Regression in Welfare Economics", *Journal of Business & Economic Statistics*, 14(4): October, 478-486.

[15] Yitzhaki, S., and E. Schechtman (2013). *The Gini Methodology: A Primer on a Statistical Methodology*. Springer: New York.

[16] Yitzhaki, S., and P. Lambert (2013a). "The Relationship between Absolute Deviation from a Quantile and Gini's Mean Difference", *Metron*, 71(2): 97-104.

[17] Yitzhaki, S. and P. Lambert (2013b). "Is higher variance necessarily bad for investment?", *Review of Quantitative Finance and Accounting*, 43(4): 855-860.