

Transfer Pricing Under Gross Profit Methods: Adjustment for Functions Performed

By
Ednaldo A. Silva

Dr. Ednaldo Silva applies a regression model to estimate the arm's length gross profit margin of a sample of pharmaceutical distributors and evaluates the reliability of this approach.¹

Ah, it is hard to speak of what it was, that savage forest, dense and difficult, which even in recall renews my fear—Dante, Inferno, Canto I, Lines 4-6.

Introduction

The essence of the profit-based methods used in transfer pricing analysis is the functional relationship between economic variables. In practice, the economist must select the best model to explain the behavior of a profit level indicator, such as the gross profit of a given taxpayer under review, as a function of a major predictor variable.

The U.S. regulations specify three transfer pricing methods based on gross profit. The resale price and the cost-plus methods predict an arm's length result by reference to the gross profit of comparable companies. The comparable profits method, when applied using the Berry ratio (gross profit over operating expenses), predicts an arm's length gross profit by reference

to the operating expenses of comparable companies.

In each case, the methods endeavor to predict an arm's length result based on a single scaling variable, either net sales, cost of goods sold, or operating expenses. However, I propose that under general conditions the reliability of each method can be increased by introducing an adjustment for functions performed based on operating expenses as a proxy.² This proposed adjustment is important since the OECD Guidelines and certain tax authorities outside the U.S. appear to have a preference for using gross profit based methods (primarily resale price and cost plus), instead of the comparable profits method ("CPM")³ favored by the Internal Revenue Service ("IRS").⁴ In this article, I use a sample of distributors of pharma-

Ednaldo A. Silva, Ph.D., is an economist with Silva Analytics, Inc., a transfer pricing and valuation firm. Dr. Silva participated in the drafting of the temporary and final Code Sec. 482 regulations.

ceutical products, including cosmetics and perfumes, to illustrate the operating expense adjustment.

Gross Profit Methods

In applying each of the gross profit methods identified above, some economists use a naïve model (without any predictor or explanatory variable) to determine the arm's length amount of gross profit of the taxpayer under review.⁵ In general, the naïve model fails to satisfy the best method rule that taxpayers and tax administrators adopt the most reliable measure according to the arm's length principle.⁶ The naïve model suggests fitting the following equation to determine the arm's length gross profit margin:

$$(1) g_i = \alpha + \epsilon_i$$

where g_i denotes gross profit margin (gross profit over net sales), α denotes a mean or average gross profit margin, ϵ_i denotes a random error associated with the i -th comparable, and $i = 1$ to N comparables. The error term has zero expected value and a constant variance. This indicates that the average error has a null effect in determining the appropriate gross profit margin.

The naïve model asserts that the gross profit margin depends on a single parameter (α), plus a random variation (ϵ_i), but it does not incorporate a predictor variable. Therefore, fitting the naïve model is equivalent to computing an average gross profit margin among a sample of comparables.⁷

A major problem with the naïve model is the absence of any predictor or causal factor explaining the behavior of the gross profit margin. As a result, the variance

of the gross profit margin produced by the naïve model tends to be large compared to any central measure of the data distribution. In this way, the naïve model may produce results that are inconsistent with the regulations because it does not incorporate the effect of functions performed on the gross profit margin.⁸

Adjustment for Operating Expenses

However, in contrast to the underlying theory of the naïve model, gross profit tends to be influenced by variables other than net sales, such as operating expenses. To test the proposition that operating expenses are a basic predictor variable of gross profit, I examined the results of publicly traded U.S. companies primarily engaged in the wholesale distribution of prescription drugs, proprietary drugs, and toiletries, including cosmetics and perfumes, classified in Standard Industrial Classification ("SIC") Code 5122⁹ (see Appendix and Exhibit 2). My selection of industry is arbitrary and is done for the purpose of illustrating the principles involved in the proposed adjustment herein.

Regression Model

After selecting a predictor variable, I examine the relationship between the gross profit margin and the operating expense ratio (excluding depreciation). Therefore, I propose the regression model as a more credible alternative to the naïve model:¹⁰

$$(2) g_i = f(x_i) + \epsilon_i, \text{ for } i = 1 \text{ to } N \text{ comparables.}$$

Above, x_i denotes the operating expense ratio (selling, general, and administrative expenses, except depreciation and amortization, over net sales) of the i -th compa-

table. As a first step in the analysis, I recommend plotting x_i versus g_i , because a scatter diagram provides insight into the functional relationship between two variables, and allows facts and circumstances to dictate the relationship exhibited by the selected comparables. For example, the x_i versus g_i scatter diagram will show the existence of any non-linearity in the particular sample of comparables being examined, or the existence of any extreme observations. Assuming the existence of a single predictor variable, the regression Model 2 can be interpreted as a polynomial of degree 1 (representing a linear model) or of degree 2 (representing a quadratic model).

Arm's Length Range

Reg. §1.482-1(e) (Arm's length range) provides that the arm's length range is determined by applying a single pricing method selected under the best method rule to two or more (*i.e.*, $N \geq 2$) uncontrolled transactions of similar comparability and reliability. Further, Reg. §1.482-1(e)(2)(B) provides that if exact comparables are not found (which, in my experience, represents the general case under ordinary facts and circumstances), the arm's length range will be determined by including the middle 50% of the data.¹¹ Therefore, the arm's length range can be determined by multiplying the standard error of the fitted regression by a known t -scalar derived from a statistical table.¹² As a good approximation, the arm's length range can be computed according to the formula:

$$(3) \hat{g}_i \pm 0.6745 s,$$

where \hat{g}_i denotes the gross profit margin predicted by the fitted re-

Figure 1.
Summary of Regression Results

Model	alpha	beta	gamma	sigma	rho ²	N
1 Naïve	0.2368			0.1724		16
2 Linear	0.0439	0.9984		0.0496	0.923	16
3 Quadratic		1.7206	-1.5001	0.0289	0.972	16

gression for a given level of operating expense ratio, and s denotes the standard error of the fitted regression model.¹³ As shown below, \hat{g}_i can be determined by a series of competing models, including the naïve and regression models.

Empirical Results

A scatter diagram between the paired x_i and g_i for the sample of pharmaceutical distributors examined suggested a quadratic regression in which the intercept is not statistically different from zero (see Exhibit 1). The Greek characters α (alpha) denote the intercept of the regression model, β (beta) denotes the first slope, and γ (gamma) denotes the second slope. Likewise, σ (sigma) denotes the standard error of the fitted model, and ρ (rho) denotes the correlation coefficient between the response and predictor variables. In the case of multiple regression analysis (such as polynomial regression), ρ^2 (or the adjusted R^2) is adjusted for the

number of predictors introduced in the model.

According to the smallest sigma (or minimum error) rule, Model 3 is the best model among the three competing models reviewed. This rule is reflected also in the fact that Model 3 produces the smallest HiLo ratio (an index of reliability), which is the ratio of the upper limit to the lower limit of the predicted gross profit margins produced under the competing models:¹⁴

The statistical results above reflect three-year weighted-average financial information from 16 companies out of the original sample of 19 distributors of pharmaceutical products (see Appendix and Exhibit 2). Three companies were set aside from the analysis because either a given company produced a large standardized regression residual (Andrx Corporation, "ADRX"; and NU Skin Enterprises, "NUS"), or it represented an observation whose predictor variable gives a large influence on the gross profit margin (Scios Inc., "SCIO"). For example, SCIO has a three-year weighted-

average operating expense ratio of 96.6%, and therefore has a large three-year operating loss.

If the three outlying companies are included in the analysis, the shape of the relationship between the gross profit margin and the operating expense ratio does not change (see Exhibit 1, which represents a scatter plot of the full sample). However, the results are much less reliable. For example, the standard error of the linear regression would be 9.88% instead of 4.96%, and the standard error of the quadratic regression would be 7.11% instead of 2.89%, as reported above. Moreover, an analysis of the standardized residuals produced by the fitted regressions does not justify inclusion of the three statistical rascals in the reported results.

Examples

The following two examples illustrate the operating expense adjustment proposed herein.

Example 1: Reg. § 1.482-3(c)(ii)(D) (Sales agent) provides

Figure 2.
Summary of 50% Range of Results When $\alpha = 5\%$

Model	Low \hat{g}	High \hat{g}	HiLo Ratio	Operating profit margin
1 Naïve	0.1205	0.3531	2.93	0.0705
2 Linear	0.0604	0.1273	2.11	0.0104
3 Quadratic	0.0628	0.1018	1.62	0.0128

under the resale price method that if the controlled taxpayer is comparable to a sales agent that does not take title to goods (or otherwise assume inventory risks with respect to ownership of such goods), the commission earned by such sales agent, expressed as a percentage of the uncontrolled sales price of the goods involved, may be used as the comparable gross profit margin. In general, wholesale distributors of large volume (or bulky) products performing limited functions and not assuming inventory risks have low operating expense ratio and, thus, can operate with low gross profit margins.

Assume that a certain controlled taxpayer, A, has an operating expense ratio of 5%, or $x_A = 0.05$. Exhibit 2 shows that seven out of 16 companies have operating expense ratios below or near 5%. Therefore, finding the arm's length range for this low expense ratio taxpayer involves two steps:

■ The first step is to introduce the taxpayer's operating expense ratio into the fitted regression model, and find the corresponding gross profit margin:

$$(4a) \hat{g}_A = 1.7206 (0.05) - 1.5001 (0.05)^2, \text{ or} \\ \hat{g}_A = 0.0823, \text{ or } 8.2\%,$$

which reflects the operations of $N = 16$ distributors of pharmaceutical products operating in the same geographic market.

■ Second, introduce the gross profit margin found in the step above, and the standard error of the fitted regression, into the arm's length range formula (3), and find the limits of the range:

$$(4b) 8.23\% \pm 0.6745 (2.89\%), \\ \text{or } 6.28\% \text{ to } 10.18\%.$$

In comparison, the naïve model produces less reliable results, which are always independent of the operating expense ratio of the taxpayer and, thus, do not reflect the functions performed by the entity under exam. The naïve model produces either of the following pseudo arm's length range:¹⁵

$$(4c) 23.68\% \pm 0.6745 (17.24\%), \\ \text{or} \\ 12.05\% \text{ to } 35.31\% \text{ (Tolerance interval); or}$$

$$(4d) 7.91\% \text{ to } 42.64\% \text{ (Interquartile range); Medium} = 25.28\%.$$

Exhibit 2 shows that virtually all distributors with low expense ratios have gross profit margins in the arm's length range determined above. In contrast, the pseudo arm's length ranges would misrepresent the gross profit margins of the distributors with low expense ratios.

Example 2: In general, companies that perform more functions are expected to earn higher levels of gross profit, and functions performed are basically reflected in operating expenses. Reg. §1.482-5(c)(2)(11) recognize the relationship between gross profit and functions performed, measured by operating expenses: "[R]esources and risks usually are directly related to functions performed ... because differences in functions performed often are reflected in operating expenses, taxpayers performing different functions may have very different gross profit margins but earn similar levels of operating profit [margins]."

Assume that a certain controlled taxpayer, B, performs vari-

ous functions and has an operating expense ratio of 25%, or $x_B = 0.25$. Exhibit 2 shows that only three companies have expense ratios exceeding 25%. Finding the arm's length range for this above-median expense ratio taxpayer involves two steps, as described above:

As the first step, introduce the taxpayer's operating expense ratio into the fitted regression model, and find the corresponding gross profit margin:

$$(5a) \hat{g}_B = 1.7206 (0.25) - 1.5001 (0.25)^2, \text{ or} \\ \hat{g}_B = 0.3364, \text{ or } 33.64\%.$$

Second, introduce the gross profit margin found in the step above, and the standard error of the fitted regression, into the arm's length range formula (3), and find the limits of the range:

$$(5b) 33.64\% \pm 0.6745 (2.89\%), \\ \text{or } 31.69\% \text{ to } 35.59\%.$$

In comparison, the naïve model produces the same pseudo arm's length range as under Example 1, even though taxpayer B has an operating expense ratio that is five times higher than that of taxpayer A. It is not surprising that after obtaining such perverse results from misapplying a statistical model, some analysts resort to the cherry-picking habit of selecting arbitrary comparables (or introducing *ad hoc* adjustments to the gross profit margin of the selected comparables).

Conclusion

Transfer pricing practice often resembles Dante's descent into limbo. Unlike Dante, the econo-

mist does not benefit from the enlightened guidance of Virgil to escape the “shadowed forest” of tax controversy and must rely on analytical skills and intellectual acumen to produce credible results.¹⁷

The regulations require companies (usually through the help of an economist) to establish whether prices or profit margins (or any other form of intercompany consideration, such as royalty rates) are arm’s length. The transfer pricing approach proposed in this article involves extending the traditional or naïve model based on simple averages or interquartile ranges to account for the causal relationship between a response variable and one or more explanatory variables. This means that under the comparable *gross profit* method

(including either the resale price or the cost plus method), and the comparable *operating profit* method, it is important to quantify the relationship between the selected profit level indicator (in this case, the gross profit margin) and the factor(s) that affect its behavior. Since functions performed have a major effect on the gross profit margin, a suitable proxy, operating expenses, has been selected as the most important explanatory variable.

In this article, I offer a regression model to estimate the arm’s length gross profit margin of a taxpayer under review as a function of the operating expense ratio and random forces (as quantified by the error term). Under the sample of pharmaceutical distributors analyzed and several experiments

of other U.S. distributors not included in this paper, the regression model yielded more reliable results than the naïve model. Therefore, I can conclude that the regression approach adopted in this paper is better than the naïve model from the standpoint of economic methodology and statistical practice, as well as in compliance with the reliability concerns expressed in the transfer pricing regulations of the United States and other OECD countries. In this way, following the analytical approach adopted here, the taxpayer is more likely to reduce the uncertainty surrounding speculative transfer pricing documentation and is more likely to avoid the imposition of accuracy-related transfer pricing penalties.

APPENDIX

Search criteria

The following table summarizes the criteria applied in selecting the sample companies from Standard & Poor’s Compustat® database (CD plate 05/31/99).

Criteria	Number of companies passing the criterion	Explanation
SIC = 5122	29	Companies primarily engaged in the wholesale distribution of prescription drugs, druggists’ sundries and toiletries.
FIC = 0	25	The company is incorporated in the U.S. (same geographic market; hence comparable input and output prices).
@and(stko < > 4, exchg < > 5)	25	The company is not a subsidiary and has not undergone a leveraged buyout.
@psum(sale > 0, y96, y98)	19	Active companies for at least the past three years.

Exhibit 1. Gross Profit as a Function of Operating Expenses (% of net sales)

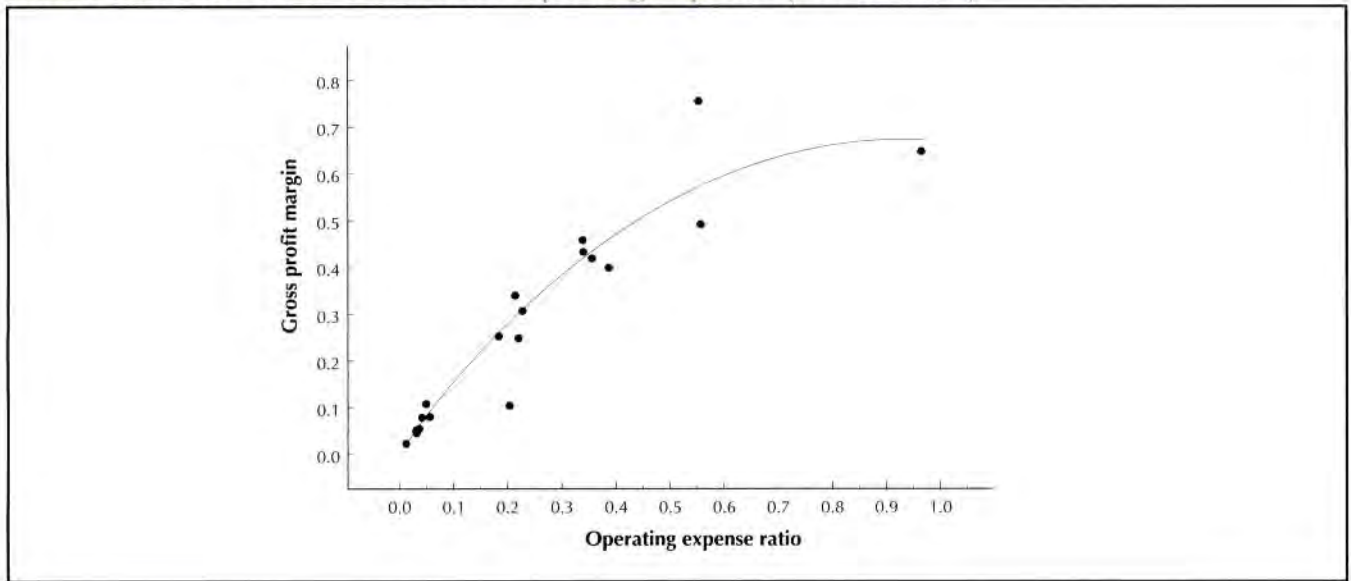


Exhibit 2. Distributors of Pharmaceutical Products, Cosmetics and Perfumes, SIC=5122

@pavg(sale,y96,y98)
 @pwavg(gp/sale,sale,y96,y98)
 @pwavg(xsga/sale,sale,y96,y98)
 (Sales in \$ millions)

Company Name	Ticker	Sales	GPM	Expense ratio
ADVANTAGE MARKETING SYS INC	AMSO	9.870	0.4195	0.3562
AMERISOURCE HEALTH CP -CL A	AAS	7,345.472	0.0508	0.0315
ANDRX CORP	ADRX	161.911	0.1039	0.2039
BARBERS HAIRSTYLING FOR MEN	BBHF	22.275	0.3400	0.2142
BERGEN BRUNSWIG CORP -CL A	BBC	12,908.287	0.0541	0.0370
BINDLEY WESTERN INDS	BDY	6,749.600	0.0224	0.0128
BIOPHARMACEUTICS INC	3BOPM	4.616	0.3991	0.3874
CARDINAL HEALTH INC	CAH	11,916.201	0.0785	0.0424
D&K HEALTHCARE RESOURCES INC	DKWD	536.085	0.0452	0.0326
FUTUREBIOTICS INC	VITK	11.566	0.4917	0.5575
HERBALIFE INTL INC -CL A	HERBA	760.343	0.4579	0.3394
MCKESSON HBOC INC	MCK	21,448.233	0.0798	0.0571
NU SKIN ENTERPRISES -CL A	NUS	827.546	0.7556	0.5536
NUTRITION FOR LIFE INTL INC	NFLI	83.369	0.3067	0.2280
PHARMERICA INC	DOSE	647.502	0.4333	0.3399
PRIORITY HLTHCARE CP -CL B	PHCC	221.618	0.1076	0.0497
SCIOS INC	SCIO	61.789	0.6479	0.9655
SEL-LEB MARKETING INC	SELB	16.057	0.2483	0.2206
SYNCOR INTL CORP/DE	SCOR	398.678	0.2528	0.1844
LOWER QUARTILE			0.0791	0.0461
MEDIAN			0.2528	0.2142
UPPER QUARTILE			0.4264	0.3481

- ¹ The specific companies whose functions are at issue are U.S. distributors of pharmaceutical products, cosmetics, and perfumes. I appreciate the assistance of Elisabetta Linares and the critical comments of Beth Williams and Linwood Smith.
- ² This approach (which may be called an augmented Berry ratio method) is applicable to distributors and service companies. Application of the proposed adjustment to manufacturing companies is complex because of the apparent haphazard allocation of expenses between cost of goods sold and operating expenses. See Reg. § 1.482-5(d)(3) for a definition of operating expenses. The regulations include depreciation in the definition of operating expenses; however, I exclude depreciation to remove the effect of different asset valuation (different vintage of fixed assets and different depreciation schedules) from the results.
- ³ Or the similar transactional net margin method ("TNMM") described in the *OECD Transfer Pricing Guidelines*. See generally Organization for Economic Cooperation and Development, *Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations* (hereinafter "OECD Guidelines").
- ⁴ The avatar found in the OECD Guidelines and repeated, *sans cogito*, by numerous practitioners that the resale price and the cost plus methods are transactional methods and, hence, preferable to "profit based" methods is erroneous. The only transactional (or disaggregated) transfer pricing method is the comparable uncontrolled price method. All aggregated methods, including the resale and the cost plus methods, are profit based methods. Treating a method based on gross profit (instead of operating profit) as not a profit based method, is fatuous.
- ⁵ The culprits cannot derive solace from Reg. § 1.482-3(c)(2)(iii) (Appropriate gross profit): "The appropriate gross profit is computed by multiplying the applicable resale price by the gross profit margin (expressed as a percentage of total revenue derived from sales) earned in comparable uncontrolled transactions." See also Reg. § 1.482-3(d)(2)(ii) (Appropriate gross profit): "The appropriate gross profit is computed by multiplying the controlled taxpayer's cost of producing the transferred property by the gross profit markup, expressed as a percentage of cost, earned in comparable uncontrolled transactions." In my reading, the U.S. transfer pricing regulations and the OECD Guidelines do not prescribe using any particular model to determine the arm's length gross profit margin.
- ⁶ According to Reg. § 1.482-1(c)(2), determining the best transfer pricing method depends on "two primary factors": the comparability between the taxpayer and uncontrolled comparables, and data quality and assumptions used in the analysis.
- ⁷ The expected value of $g_i = \alpha$, because the expected value of $e_i = 0$. In a *chimera* practiced by some analysts, the quartiles of a selected profit level indicator (including gross profit margin, operating profit margin, or rate of return on operating assets, derived from selected comparables) are invariably the best measure according to the arm's length principle. However, median and interquartile range may provide good summary statistics only when the data are skewed or contain outliers (extreme high or low values).
- ⁸ Typically, the potentially comparable company documents reviewed are 10-K reports filed in compliance with SEC regulations. In general, the narrative section of those reports does not provide sufficient information to establish comparability with the taxpayer under review.
- ⁹ I have observed a discernable relationship between the operating expense ratio and the gross profit margin in samples of all U.S. publicly traded distributors of durable goods and non-durable goods distributors.
- ¹⁰ In appraising scientific research, Lakatos suggests that a causal proposition is more credible than a rival proposition when it produces more information than the rival, and some of the additional information can be corroborated by empirical evidence: "[A] theory is 'acceptable' or 'scientific' only if it has corroborated excess empirical content over its predecessor (or rival), that is, only if it leads to the discovery of novel facts." See I. Lakatos, "Falsification and the Methodology of Scientific Research Programs," in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge* (Cambridge University Press, 1970), at 116.
- ¹¹ The drafters of the U.S. transfer pricing regulations overreached by adding the capriccio: "The interquartile range ordinarily provides an acceptable measure of this [arm's length] range"; and then provided the relief: "however a different statistical method may be applied if it provides a more reliable measure." Like a Faustian character, the U.S. regulations represent two souls, and one is forever departing from the other.
- ¹² See D. Lindley and W. Scott, *New Cambridge Statistical Tables* (Second edition) (Cambridge University Press, 1995), Table 10 (Percentage Point of the *t*-Distribution). The multiplier $t = 0.6745$ was obtained from the column corresponding to percentage points at the 25% tail.
- ¹³ In order to limit the proliferation of symbols, except for the distinction between an observed and a predicted gross profit margin, hereinafter I will not distinguish a population parameter from a sample estimate. The exact interval formula for predicting a response variable for a given value of the predictor variable can be found in virtually all statistical books including regression analysis. See, e.g., J. Kmenta, *Elements of Econometrics* (Macmillan, New York, 1986), at 251 (formula 7.53).
- ¹⁴ As a practical matter, given the regression results above, Linear Model 2 may be preferred to Quadratic Model 3 because it is simpler to interpret.
- ¹⁵ There are two kinds of statistical intervals: Confidence intervals to catch an estimate of an unknown parameter of the population, such as its unknown mean or central value, and tolerance intervals to catch a certain proportion of the population. Confidence intervals are narrower because the associated standard errors are scaled by the square root of the respective sample size, so the larger the sample size the smaller the standard error of the parameter estimate. Tolerance intervals are wider because the intent is to catch any individual exhibiting a studied attribute in a selected population, subject to a specified statistical error. See G. Hahn and W. Meeker, *Statistical Intervals* (John Wiley & Sons, New York, 1991). For the purpose of this article, I use tolerance intervals. However, under facts and circumstance when the amount of taxable income subject to a transfer pricing test is large, and thus very sensitive to small deviations in the profit level indicator selected, confidence intervals may be appropriate. The decision to use confidence or tolerance intervals also depends on the characteristics of the sampled comparables, particularly the degree of variability exhibited by the data.
- ¹⁶ These quartiles were produced by Excel® (see Exhibit 2). However, there are several acceptable methods of producing quartiles discussed in M. Frigge, D. Hoaglin, and B. Iglewicz, "Some Implementations of the Boxplot," *American Statistician*, Vol. 43, No. 1 (February 1989).
- ¹⁷ Until Virgil rescued him, Dante's escape from the dark forest was blocked by three wild beasts. Some commentators suggest that the leopard represents lust; the lion, pride; and the wolf, avarice, cupidity, and greed. See A. Mandelbaum, A. Oldcorn, and C. Ross (eds.), *Lectura Dantis: Inferno* (University of California Press, Berkeley, 1998), at 10. *De te fabula narratur*.