# Chapter 2
# The Parametric Approach to Income and Wealth Distributional Analysis

*Essentially, all models are wrong, but some are useful.*
George Edward Pelham Box

**Abstract** This chapter provides a brief overview of the main advantages associated with using parametric models of income/wealth distribution. It also instructs the reader about the numerous statistical models of income/wealth distribution that have been proposed in both the statistical and economic literature for over 100 years since Pareto's breakthrough contribution.

**Keywords** Personal income, wealth and their distributions · Parametric modeling · Literature review

## 2.1 The Idea of a Parametric Model for Income and Wealth Distributions

The analysis of a probability density function is a powerful tool to describe the full distribution of a variable of interest and one which reveals much more information of its several properties compared to standard descriptive statistics, such as the mean, variance, skewness or kurtosis, summarizing each limited properties of the distribution on single values (Cowell and Flachaire 2015). For instance, the analysis of the shape of the income/wealth distribution provides a picture from which at least three important distributional features can be observed simultaneously (Cowell et al. 1996): income/wealth levels and changes in the location of the distribution as a whole; income/wealth inequality and changes in the spread of the distribution; clumping and polarization as well as changes in patterns of clustering at different modes.

In practice, the shape of the density function is unknown and has to be estimated from data on hand. For a long time, the main estimation method was parametric,



F. Clementi and M. Gallegati, *The Distribution of Income and Wealth*,
New Economic Windows, DOI 10.1007/978-3-319-27410-2_2

which requires the choice of a functional form a priori. In the last decades, the availability of powerful computer resources gave rise to various non-parametric methods of density estimation, the most popular being the kernel density estimators which are often used in empirical studies now. Their main attraction consists in the possibility to relax the specific assumptions underlying parametric estimation method, as no distributional form is imposed a priori. However, with small data sets kernel density interpolation might result in imprecise estimates; also, the accuracy of this method typically depends to a large extent on the bandwidth used in the kernel functions.[1] These inaccuracies of model-free density estimation may be reduced using the parametric technique, whose alleged "old-fashionedness"—given its 100-year-plus history as opposed to the relatively young non-parametric methods—has been earnestly defended, among others, by Piketty (2000, p. 462) in his book review of Champernowne and Cowell's (1998) *Economic Inequality and Income Distribution*:

> One criticism that could be addressed to the book is that some readers might find it a little bit old-fashioned. For instance, the authors refer frequently to Pareto coefficients and the way such coefficients can be used to measure income inequality or to fit a Pareto distribution. Many young readers have probably never heard of such techniques, and most of them will probably never use them: at the age of SAS and Stata, one simply needs to push a button on a computer to compute the deciles, centiles and various inequality measures associated with a given income data base. In my view, however, such a criticism would be misplaced: the point is that individual-level databases on incomes have been available only for the past few decades and for the most developed countries, so that these 'old-fashioned' techniques are still very useful. In addition, it seems to me that it is very useful to have a solid background about past economists' work on income distribution (instead of believing that serious research started in 1980 or in 1990, which today's economists frequently tend to assume), and this book provides such a background.

Of course, using parametric models in the analysis of the distribution of income and wealth is not without advantages. For instance, fitting parametric models allows one to represent the entire distribution through a small number of estimated parameters (Brachmann et al. 1996). The estimated parameters may then be used to reconstruct the entire distribution if, for example, income or wealth data released in future are published in grouped form (Hajargasht et al. 2012) or if available microdata are censored to bounds (i.e. bottom and/or top-coded; see Feng et al. 2006, and Burkhauser et al. 2011). Reconstruction of the overall distribution can also be achieved with the help of a reliable parametric model when only empirical estimates of poverty and/or inequality measures are available—such as those published by the

---

[1] In standard kernel method the bandwidth remains constant at all points where the distribution is estimated. This constraint can be particularly onerous when the concentration of data is markedly heterogeneous in the sample. Hence, there would be advantages from using a narrower bandwidth in the dense part (the middle) of the distribution and wider ones in the more sparse tails—as in "adaptive" kernel estimation (Van Kerm 2003)—especially in the cases of heavy-tailed income and wealth distributions. Greater modeling flexibility can also be achieved by means of finite mixture densities, defined as convex combinations of two or more parametric densities. The separate analysis of the components of the mixtures and of the relative importance of these components over time are the main advantages over a non-parametric approach, as they allow capturing the effect of unobserved heterogeneity (Cowell and Flachaire 2015).

World Bank, OECD, Eurostat or other statistical agencies—and direct access to the underlying micro-data is not possible (Graf and Nedyalkova 2014). Furthermore, as maintained by Cowell (2000, p. 145):

> some standard functional forms claim attention, not only for their suitability in modeling some features of many empirical [...] distributions but also because of their role as equilibrium distributions in economic processes.

The parameters of theoretical models often possess also economic interpretation, which allows to gain insights about the causes of the evolution of the distribution over time or interpret the differences between income/wealth distributions across countries. Moreover, once a given parametric model is fitted to a data set, one can straightforwardly compute inequality and other welfare indicators as analytical functions of the parameters of the model. It is also possible to use estimated parameters to perform stochastic dominance testing for inequality and welfare differences between distributions (Kleiber and Kotz 2003) as well as to empirical model the impact of macroeconomic conditions (e.g. GDP growth, unemployment and inflation rates, etc.) on the evolution of the personal income or wealth distribution—see Jäntti and Jenkins (2010) and references therein. Finally, estimated parameters may be made a function of covariates summarizing personal characteristics; this allows distributional shape to vary with population characteristics and provides a route to decomposition analysis of the sources of trends in income/wealth distributions over time or differences between countries (Betti et al. 2008; Biewen and Jenkins 2005; Quintano and D'Agostino 2006).

## 2.2 Brief History of the Models for Studying Income and Wealth Distributions

The interest in finding parametric models for the size distributions of income and wealth has a 100-year-plus history that dates back to the work of Pareto (1895), who was the first to propose a functional form for approximating the observed distributions.[2] The Pareto model—usually referred to as the (strong) "Pareto law"—is a two-parameter distribution with a power-law density on the support $[x_0, +\infty)$; it was found to accurately model high levels of income/wealth, but did a poor job in describing the lower-middle range of the distribution. Based on Pareto's economic foundations, and on the stochastic foundations afterward developed by other authors (e.g. Mandelbrot 1960 and Ord 1975), the Pareto law is now overwhelmingly considered as the parametric model of the rich.[3]

[2] There exists a huge amount of literature on parametric models for the size distributions of income and wealth. Here we limit ourselves to consider the most frequently cited contributions in the area. For the interested reader, a comprehensive survey can be found in Kleiber and Kotz (2003).

[3] In his pioneering contributions at the end of the nineteenth century, Pareto (1896, 1897a, b) suggested two variants of his distribution, occasionally called the three-parameter Pareto distributions. These further Pareto distributions, however, have not been used much in empirical economic studies.

The use of other density functions was later advocated by Gibrat (1931), who provided a theoretical basis (the "law of proportionate effect") for the two-parameter lognormal to be considered as a model for the size distributions of income and wealth. The lognormal distribution was further explored by Aitchison and Brown (1954, 1957) and Lebergott (1959). Another two-parameter statistical model, the gamma, was introduced by Salem and Mount (1974). Later evidence, however, showed that while these two-parameter models fit the data relatively well in the middle range, they tend to exaggerate the skewness and perform poorly in the upper end of the empirical distributions (McDonald and Ransom 1979). Also, they do not allow for intersecting Lorenz curves sometimes observed with income/wealth data (Kleiber 2008b).

Better fits as well as intersecting Lorenz curves were obtained using three-parameter models. Among these, the generalized gamma (Atoda et al. 1988; Esteban 1986; Kloek and van Dijk 1978; Taillie 1981), Singh-Maddala (1976) and Dagum type I (Dagum 1977) distributions are the most popular. These models possess the important property of the "weak Pareto law" (Mandelbrot 1960), that is they converge in distribution to the Pareto model for sufficiently large values of income and wealth. This important convergence property, motivated by the undisputed acceptability of the Pareto distribution as the model of the rich and the extremely rich, is strengthened by their power to describe with similar accuracy the remaining parts of actual distributions, i.e. the lower and middle ranges of income and wealth.

Models with more than three parameters have also been suggested in the size distributions literature for fitting income and wealth data. As a matter of example, the generalized beta distribution of the second kind (hereafter referred to as GB2) is a four-parameter distribution introduced by McDonald (1984) which has not only been very successful in fitting the data, but also includes some of the previously mentioned two- and three-parameter models as special or limiting cases. Other four-parameter models are the generalized beta distribution of the first kind (McDonald 1984), shortly referred to as GB1, and the double Pareto-lognormal distribution (Reed and Jorgensen 2004). Both of these models have proven to work remarkably well in fitting the data, although the former is sometimes outperformed by the GB2 and the three-parameter Dagum type I and Singh-Maddala distributions (Bandourian et al. 2003; Bordley et al. 1996).

McDonald and Xu (1995a, b) also developed the five-parameter generalized beta distribution family, which includes the GB1 and GB2 as special cases—of course along with all of the two- and three-parameter distributions nested within them. In turn, the double Pareto-lognormal distribution was generalized to a five-parameter family of distributions and was called the generalized double Pareto-lognormal distribution (Reed 2007). However, closed-form expressions of the probability density and/or cumulative distribution functions do not always exist for these "super" models, which makes fitting them to data computationally difficult and slow because of the need to resort to numerical techniques (McDonald and Ransom 2008; Reed and Wu 2008). Furthermore, and of considerable importance, one might question about the usefulness and meaning of five-parameter models to approximate income and wealth distributions and could consider them to be just curious theoretical generalizations (Kleiber and Kotz 2003).

Recently, Clementi et al. (2007, 2008, 2009) developed the three-parameter $\kappa$-generalized distribution, a non-Gaussian distribution with power-law tails having its roots in the framework of statistical mechanics. This distribution is a better fit for some countries than the Singh-Maddala, Dagum type I and GB2 models (Clementi et al. 2010, 2012a) and tends to yield better estimates of income inequality even when the goodness-of-fit is inferior to the existing personal income distributions (Okamoto 2012a, b). The $\kappa$-generalized distribution was also successfully used in a three-component mixture model for handling the distinctive features of wealth survey data (Clementi et al. 2012b) and later extended to four-parameter models for the size distributions of income and consumption (Okamoto 2013). Since much of this book is a systematic exposition of all that appeared in the recent literature and was known to us about the $\kappa$-generalized distribution and its extensions, conceptualization and novel empirical research for this "family" of $\kappa$-generalized models will be discussed in depth in subsequent chapters.

http://www.springer.com/978-3-319-27408-9

The Distribution of Income and Wealth
Parametric Modeling with the κ-Generalized Family
Clementi, F.; Gallegati, M.
2016, XVI, 177 p. 24 illus., Hardcover
ISBN: 978-3-319-27408-9