Fundación BBVA

Documentos de Trabajo

1

2010

Francisco J. Goerlich Gisbert

A Simple and Efficient (Parametric Conditional) Test for the Pareto Law

Fundación BBVA

A Simple and Efficient (Parametric Conditional) Test for the Pareto Law

Francisco J. Goerlich Gisbert

UNIVERSITY OF VALENCIA
VALENCIAN ECONOMIC RESEARCH INSTITUTE (Ivie)

Abstract

This working paper presents a simple and locally optimal test statistic for the Pareto law. The test is based on the Lagrange multiplier (LM) principle and can be computed easily once the maximum likelihood estimator of the scale parameter of the Pareto density has been obtained. A Monte Carlo exercise shows the good small sample properties of the test under the null hypothesis of the Pareto law and also its power against some sensible alternatives. Finally, a simple application to urban economics is performed. An appendix presents derivations and proofs.

Key words

LM test, Pareto law, statistical distributions.

Resumen

Este documento de trabajo presenta un contraste sencillo y localmente óptimo de la ley de Pareto. El contraste está basado en el principio de los multiplicadores de Lagrange (ML) y puede ser calculado simplemente al disponer del estimador máximo verosímil del parámetro de escala de la densidad de Pareto. Algunas simulaciones de Monte Carlo muestran las buenas propiedades en muestras finitas del contraste, tanto bajo la hipótesis nula de la ley de Pareto, como su potencia frente a alternativas razonables. Finalmente, se presenta una sencilla aplicación en el campo de la economía urbana. Un apéndice recoge las demostraciones.

Palabras clave

Contrastes ML, ley de Pareto, distribuciones estadísticas.

La serie Documentos de Trabajo tiene como objetivo la rápida difusión de los resultados del trabajo de investigación entre los especialistas de esa área, para promover así el intercambio de ideas y el debate académico. Cualquier comentario sobre sus contenidos será bien recibido y debe hacerse llegar directamente a los autores, cuyos datos de contacto aparecen en la *Nota sobre los autores*.

The Working Papers series is intended to disseminate research findings rapidly among specialists in the field concerned, in order to encourage the exchange of ideas and academic debate. Comments on this paper would be welcome and should be sent direct to the authors at the addresses provided in the About the authors *section.*

Todos los documentos de trabajo están disponibles, de forma gratuita y en formato PDF, en la web de la Fundación BBVA. Si desea una copia impresa, puede solicitarla a través de publicaciones@fbbva.es.

All working papers can be downloaded free of charge in pdf format from the BBVA Foundation website. Print copies can be ordered from publicaciones@fbbva.es.

La serie Documentos de Trabajo, así como información sobre otras publicaciones de la Fundación BBVA, pueden consultarse en: **http://www.fbbva.es**

The Working Papers series, as well as information on other BBVA Foundation publications, can be found at: **http://www.fbbva.es**

A Simple and Efficient (Parametric Conditional) Test for the Pareto Law

EDITA / PUBLISHED BY
Fundación BBVA, 2010
Plaza de San Nicolás, 4. 48005 Bilbao

C O N T E N T S

1. Introduction 5

2. The Pareto Density 7
   2.1. Some results related to the Pareto law 10
   2.2. Maximum likelihood estimation of θ 10

3. A Generalization of the Pareto Distribution 13

4. A Simple and Locally Optimal Test 15

5. An Application to Urban Economics 22

Appendix 24
   Pareto law, Theil indexes and related results 24
   The LM test statistic for the Pareto law 26

References 30

About the Author 33

1. Introduction

GIVEN the recent focus on power laws (Gabaix, 2008), it seems of interest to develop estimation and testing strategies that help us to better discriminate between these laws and alternatives ones. This working paper makes a small contribution to this literature by proposing a simple and locally optimal (parametric conditional) test statistic that can be easily computed once the power exponent of the law has been estimated.

By power law, we mean a distribution for which (at least in the upper tail) the probability of getting values higher than x is (approximated) proportional to $x^{-\theta}$, where θ is a shape parameter known as the power exponent. This law has a long tradition and was stated by Vilfredo Pareto at the end of the XIX century in his studies on income distribution (Pareto, 1896). For this reason, a power law is also known as the Pareto distribution which, in its original form and using the modern tools of distribution theory, can be stated as,

$$\Pr\left[X > x\right] = \left(\frac{\mu}{x}\right)^{\theta} \quad x \geq \mu > 0, \quad \theta > 0. \tag{1.1}$$

From (1.1) it follows that the cumulative distribution function for a Pareto random variable is given by $F_p(x) = 1 - \left(\frac{\mu}{x}\right)^{\theta}$ and the associated density function is $f_p(x) = \frac{\theta\mu^{\theta}}{x^{\theta+1}}$. Although this seems to be a two parameter distribution, in practice μ is assumed to be known, either fixed by statistical design or taken as given by *a priori* information given the problem at hand.

It is worth noting that this is a scale-free distribution. In fact, we can think of a power law as a result of the simple physical principle of scale invariance: if the growth process is the same at all scales, the final distribution should be scale invariant, and if a steady state distribution exists, then a power law emerges (Simon, 1955; Gabaix, 1999). This is essentially Gibrat (1931) law in the economics literature (Sutton, 1997), but the same mechanism has been the focus in physics and the natural sciences (Levy and Solomon, 1996; Sornette and Cont, 1997; Brown and West, 2000; Sornette, 2004). This can explain the extreme propensity for using the Pareto distribution instead of other well known distributions in applied statistics that are capable of displaying a similar shape.

The literature focusing on estimating and testing the power law has taken two different stands. On the one hand, some authors have transformed the size-frequency relation (1.1) into a rank-size rule. This is brought about by considering a random sample of size n ordered as

$$x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(r)} \geq \ldots \geq x_{(n-1)} \geq x_{(n)}. \tag{1.2}$$

Thus, the rank of an element in this sample is given by $r \approx n.\Pr\left[X > x_{(r)}\right]$, which under the power law (1.1) is equal to

$$r \approx n. \left(\frac{\mu}{x_{(r)}}\right)^{\theta}. \tag{1.3}$$

Taking logs in (1.3), we can write $\log r \approx \alpha - \theta \log x_{(r)}$, where α is a function of n, μ and θ. We can measure departures from the Pareto law by departures from this log-log linear relationship between ranks and the ordered sample (1.2). This route has been taken by a number of authors (Rosen and Resnick, 1980; Fan and Casetti, 1994; Gabaix and Ibragimov, 2008), who mainly use ordinary least squares as the estimation method, despite the fact that the ordered sample is neither independent nor identically distributed and therefore it is not the best estimation procedure (Aban and Meerschaert, 2004; Nishiyama and Osada, 2004). Note in passing that, even if the power law holds exactly the rank-size rule represented by the log-log linear equation derived from (1.3), it is just an approximation [1].

On the other hand, some authors have stuck to the size-frequency relation (1.1) and they measure departures from the power law by comparing the Pareto distribution against other distributions, such as the lognormal or the exponential, by means of goodness-of-fit tests. In this vein, Clauset, Shalizi and Newman (2007) advocate the Kolmogorov-Smirnov statistic. However, given the nuisance parameters in constructing the statistic, the p-value for the statistic has to be computed by Monte Carlo integration. As an alternative, section four develops a simple test of the power law (1.1) simply by nesting the Pareto distribution within a broader family. Before, we study some properties of the Pareto density and the generalization considered.

1. A more formal statistical argument uses the probability integral transformation to obtain the expected value if the r^{th} order statistic in (1.2) under the Pareto law and leads to the same approximation.

2. The Pareto Density

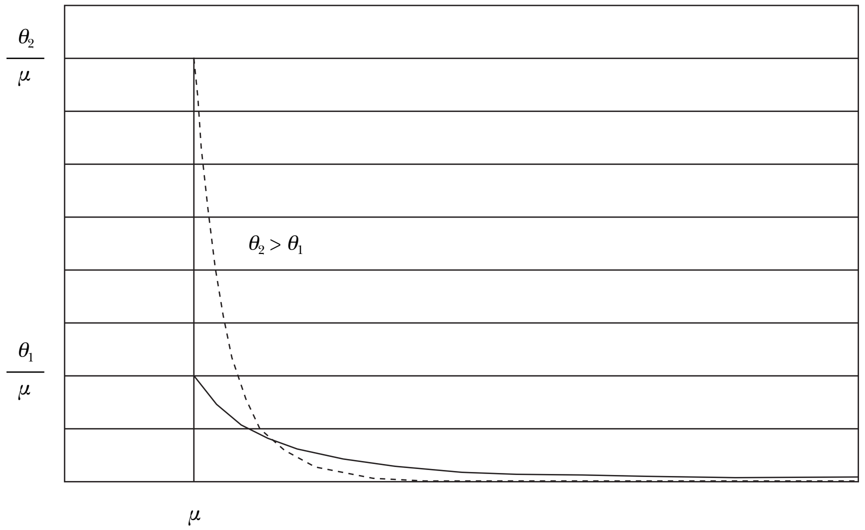
A strictly positive continuous random variable, $x \in R_{++}$, follows a Pareto law if its probability density function *(pdf)* is

$$f_p\,(x) = \frac{\theta}{\mu}\left(\frac{\mu}{x}\right)^{\theta+1} \quad x \geq \mu > 0, \quad \theta > 0. \tag{2.1}$$

For an exhaustive analysis of this distribution see Johnson and Kotz (1970, chapter 19). Together with the lognormal distribution, the Pareto distribution is one of the most frequently used in modeling income and wealth distribution. In fact, it was proposed by Pareto (1896) as a model for the upper tail in these distributions, in particular for values above μ (Cowell, 1977).

From (2.4), graph 2.1 draws the graph for the *pdf* of $x \in R_{++}$, $f_p\,(x) = \frac{\theta\mu^{\theta}}{x^{\theta+1}}$, for the same lower bound, μ, and two different values of θ. We can see that μ is a location parameter, the minimum value for which the density is

GRAPH 2.1: Pareto probability density function



defined, whereas θ is a shape parameter related to the dispersion. Higher values of θ imply more probability mass closer to the lower bound, μ. That is, there is less dispersion (or inequality if we think in terms of income or wealth distributions).

By integrating (2.4), we obtain the cumulative distribution function (*cdf*), (1.1), $F_p(x) = 1 - \left(\frac{\mu}{x}\right)^{\theta}$, so $1 - F_p(x) = \left(\frac{\mu}{x}\right)^{\theta}$ represents the probability of obtaining a value of the variable which is greater than x.

For a Pareto density it is well known that the mean is given by $E(X) = \frac{\theta}{\theta - 1}\mu$, $\theta > 1$, and more generally $E(X^r) = \frac{\theta}{\theta - r}\mu^r$, $\theta > r$.

For $r = 2$ we can easily obtain the variance, $V(X) = \frac{\theta}{(\theta-1)^2(\theta-2)}\mu^2$, $\theta > 2$ and the coefficient of variation, $CV(X) = \frac{1}{\sqrt{\theta.(\theta-2)}}$, $\theta > 2$. Note that these dispersion measures shrink to 0 as $\theta \to \infty$, and that $CV(X)$ depends only on the shape parameter, given its relative nature. On the contrary the main location measure, $E(X)$, depends on both the location and scale parameters, so a change in θ affects the mean of the distribution, $E(X)$[2].

It is of interest to examine the relation between the Pareto law and standard measures of income inequality, the Lorenz (1905) curve and its corresponding Gini (1912) coefficient[3]. By definition, the Lorenz ordinates are

$$\Phi_p(x) = \frac{1}{E(X)}\int_{\mu}^{x} z dF_p(z) = 1 - \left(\frac{\mu}{x}\right)^{\theta - 1}, \qquad (2.2)$$

which given that $\left(\frac{\mu}{x}\right) = (1 - F_p(x))^{1/\theta}$, can be written in terms of the percentiles of the distribution, $p = F_p(x)$, $p \in [0,1]$, as

$$L_p(p) = 1 - (1-p)^{\frac{\theta-1}{\theta}} \qquad \theta > 1. \qquad (2.3)$$

2. In fact $E(X)$ is a decreasing function of θ, $\frac{dE(X)}{d\theta} = -\frac{\mu}{(\theta-1)^2} < 0$. If we want to keep $E(X)$ constant as θ varies, then μ has to change as $d\mu = \frac{\mu}{\theta(\theta-1)}d\theta$.

3. See Cowell (1995, table A.2: 142) for the relation between the Pareto law and other inequality indices commonly used in the income distribution literature.

The corresponding Gini index is twice the area under the line of equality, $L_p(p) = p$, and the Lorenz curve. Hence

$$G(X) = \frac{1}{2.\theta - 1} \qquad \theta \geq 1, \tag{2.4}$$

which again depends only on the shape parameter, θ, and is defined for $\theta \geq 1$. Note that as θ growths $G(X)$ falls and that as $\theta \to 1^+$ $G(X)$ reaches its maximum value of 1. The Zipf (1949) case implies $\theta = 1$, and hence $G(X) = 1$.

In addition, it may seem of interest to develop the corresponding formulas for the Theil (1967) indexes under the Pareto law. By definition, the second Theil index, or the mean logarithmic deviation, is just the logarithm of the ratio of the mean, $E(X)$, to the geometric mean, m_1,

$$T^* = \log \frac{E(X)}{m_1} = \int_{\mathbb{R}_{++}} \log \frac{E(X)}{x} \, dF(x). \tag{2.5}$$

As in the case of Gini this is easier to handle using the inverse of the *cdf*, or the quantile function, $x = F_P^{-1}(p) = \mu\,(1-p)^{-\frac{1}{\theta}}$, so

$$T^* = \log \frac{\theta}{\theta - 1} - \frac{1}{\theta} \qquad \theta > 1, \tag{2.6}$$

which again only depends on the shape parameter θ. Moreover the relation between θ and T^* is monotonic, since

$$\frac{dT^*}{d\theta} = -\frac{1}{\theta^2\,(\theta - 1)} < 0 \qquad \theta > 1. \tag{2.7}$$

By definition, the first Theil index is given by

$$T = \int_{\mathbb{R}_{++}} \frac{x}{E(X)} \log \frac{x}{E(X)} \, dF(x). \tag{2.8}$$

Which in terms of the quantile function, can be written as,

$$T = \log \frac{\theta - 1}{\theta} + \frac{1}{\theta - 1} \qquad \theta > 1, \tag{2.9}$$

which again only depends on the shape parameter θ, and gives us a monotonic relation between θ and T, since

$$\frac{dT}{d\theta} = -\frac{1}{\theta\,(\theta - 1)^2} < 0 \qquad \theta > 1. \tag{2.10}$$

2.1. Some results related to the Pareto law

Result 1. If x follows a Pareto law, then $y_1 = \log \frac{x}{\mu}$ follows an exponential distribution with parameter θ.

Result 2. If x follows a Pareto law, then $y_2 = \frac{\mu}{x}$ follows a power function distribution with parameter θ (a special type of Pearson Type I distribution).

Result 3. If x follows a Pareto law with $\theta = 1$, then $y_2 = \frac{\mu}{x}$ follows a uniform distribution on the interval [0, 1].

2.2. Maximum likelihood estimation of θ

Given a random sample of n observations satisfying the Pareto law, $\{x_i\}_{i=1}^{n}$, the parameters of the distribution can be estimated by a variety of methods (Johnson and Kotz, 1970). Considering μ as known (either by statistical design or *a priori* information) the maximum likelihood estimator (MLE) of θ is simple to obtain, has good known asymptotic properties and its variance can be readily computed (Hill, 1975) [4].

Given $\{x_i\}_{i=1}^{n}$ the logarithm of the likelihood function, $\log L_P$, is given by

$$\log L_P = n \log \theta + n\,\theta \log \mu - (\theta + 1) \sum_{i=1}^{n} \log x_i \tag{2.11}$$

As a function of θ, (2.11) is maximized for

$$\hat{\theta} = \left[\frac{1}{n} \sum_{i=1}^{n} \log \frac{x_i}{\mu}\right]^{-1}. \tag{2.12}$$

4. If μ is unknown (which is usually the case), it will have to be estimated. If the full sample is assumed to come from a Pareto distribution, then the MLE of μ is simply the minimum of the sample, $\hat{\mu} = \min_i \{x_i\}_{i=1}^{n}$ (Quandt, 1966). Otherwise, although methods have been developed to estimate this parameter (Embrechts, Kluppelberg and Mikosch, 1997; Clauset, Shalizi and Newman, 2007), none have gained widespread use since even the statistical literature has not reached a consensus on how to pick the optimal lower cutoff (Beirlant et al., 2004).

which is the MLE of θ. Standard Maximum Likelihood theory (Spanos, 1986) implies that the asymptotic distribution of the MLE is

$$\sqrt{n}\,(\hat{\theta}-\theta) \overset{asy}{\sim} N(0,\theta^2). \tag{2.13}$$

Remark 1. The MLE is very different from the ordinary least squares (OLS) estimator of the log-log linear relationship mentioned in the introduction (Nishiyama and Osada, 2004; Nishiyama, Osada and Sato, 2006); although Aban and Meerschaert (2004) show that the MLE coincides with the Generalized Least Squares (GLS) estimator of this log-log linear relationship taking into account the mean and covariance structure of the extreme order statistics.

Remark 2. Using result 1 and the fact that the exponential distribution belongs to the class of gamma distributions, it is possible in this case to obtain the small sample distribution of $\hat{\theta}$, so we know the exact expectation and variance of the estimator (Johnson and Kotz, 1970, chapter 19.4.4).

Remark 3. We have noticed above that θ can be considered as a measure of inequality, in the sense that higher values of θ imply less dispersion, and in fact lower values of standard relative inequality measures, *CV, G* or the Theil indexes, T and T^*.

All these measures satisfy the Pigou (1912)-Dalton (1920) principle of transfers, so a small transfer from a higher x, say x_i, to a lower x, say x_j, reduces the value of the index (Goerlich and Villar, 2009). However, if we now look at the effect of a Pigou-Dalton transfer on $\hat{\theta}$, for a fixed value of μ, we get a surprising result: Given the concavity of the log function, such a transfer implies an increase in $n^{-1}\sum_{i=1}^{n}\log x_i$, so $\hat{\theta}$ falls [5]. This suggests an increase in inequality! The result casts doubts on using $\hat{\theta}$ as a descriptive measure of inequality.

Remark 4. The MLE of θ, (2.12), has a clear relation with the non-parametric second Theil inequality index for a sample, $\bar{T}^* = \frac{1}{n}\sum_{i=1}^{n}\log\frac{\bar{x}}{x_i}$ [6], where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean. We can write (2.12) as

5. The effect of a small transfer of δ from x_i to x_j, with $x_i > x_j$, on $n^{-1}\Sigma_{i=1}^{n}\log x_i$ is given by $\delta\frac{1}{n}\left[\frac{1}{x_j}-\frac{1}{x_i}\right] > 0$, so $d\hat{\theta} = -\hat{\theta}^2\delta\frac{1}{n}\left[\frac{1}{x_j}-\frac{1}{x_i}\right] < 0$.

6. Note that this is not the MLE of T^*, which by invariance is obtained substituting (2.12) into (2.6).

$$\hat{\theta} = \left[\frac{1}{n} \sum_{i=1}^{n} \log \frac{x_i}{\bar{x}} \cdot \frac{\bar{x}}{\mu} \right]^{-1} = \left[\frac{1}{n} \sum_{i=1}^{n} \log \frac{x_i}{\bar{x}} + \log \frac{\bar{x}}{\mu} \right]^{-1} = \left[- \bar{T}^{*} + \log \frac{\bar{x}}{\mu} \right]^{-1}. \quad (2.14)$$

This is a surprising relation because, contrary to intuition, there is a positive relation between $\hat{\theta}$ and $\bar{T}^{*}$, whereas we have seen before that the relation between θ and T^{*} is a negative one, (2.7).

Given that $\bar{T}^{*}$ satisfies the Pigou-Dalton transfer principle, a progressive transfer reduces $\bar{T}^{*}$, but with $\bar{x}$ and μ unchanged, this transfer also reduces $\hat{\theta}$. In trying to understand this paradoxical result note that, under the Pareto law, *E (X)* depends on both parameters, θ and μ, so that to keep *E (X)* unchanged (as required by the Pigou-Dalton transfer principle) both parameters should move in the same direction. A progressive transfer reduces T^{*} by (2.5), which leads to an increase in θ, (2.7), this in turns needs an increase in μ, so as to keep *E (X)* constant. Therefore, insisting on fixing μ when measuring the effect of a Pigou-Dalton transfer is inconsistent under the Pareto law. In terms of sample estimates, where $\bar{x}$ is not subject to such a dependency, the result is that $\hat{\theta}$ falls as a result of a progressive transfer [7].

7. A similar paradoxical result in relation to welfare under the Pareto law appears in Cowell (1995: 84, footnote 13).

3. A Generalization of the Pareto Distribution

CONSIDER the following member of the Burr (1942) family of distribution functions (Johnson and Kotz, 1970, chapter, 12.4.5)

$$f_B(x) = \frac{\theta}{\sigma}\left(1 + \frac{x-\mu}{\sigma}\right)^{-(\theta+1)} \qquad x \geq \mu > 0, \quad \theta > 0, \quad \sigma > 0. \qquad (3.1)$$

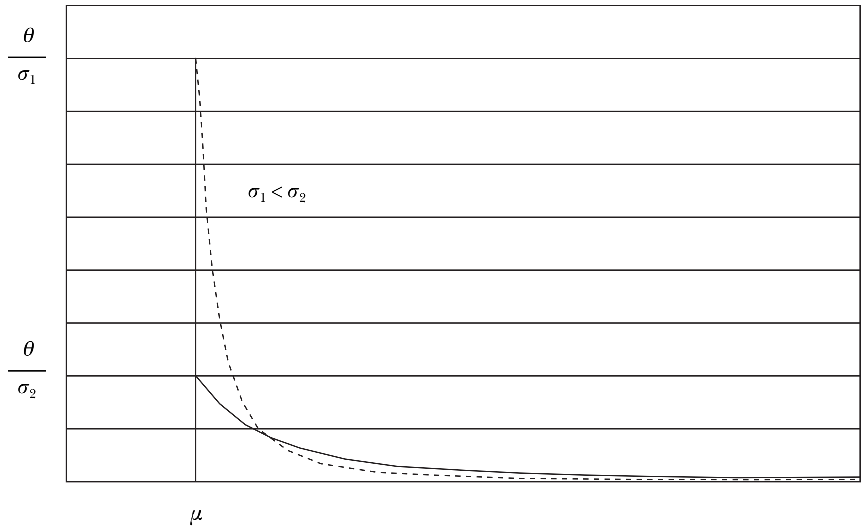
By integrating (3.1) we obtain the cumulative distribution function *(cdf)*, $F_B(x) = 1 - \left(1 + \frac{x-\mu}{\sigma}\right)^{-\theta}$, so $1 - F_B(x) = \left(1 + \frac{x-\mu}{\sigma}\right)^{-\theta}$ represents the probability of obtaining a value of the variable greater than x.

Note that, at the minimum value for which the density is defined, $x = \mu$, $f_p(\mu) = \frac{\theta}{\mu}$ whereas $f_B(\mu) = \frac{\theta}{\sigma}$, therefore σ measures the initial height of the Burr density in relation to the Pareto density. Both densities are *L*-shaped and μ is a location parameter, however in (3.1) we have two shape parameters, θ and σ, and for a given θ, lowering σ has the effect of concentrating the probability mass closer to the lower bound, μ. That is, we get less dispersion as graph 3.1 shows.

For the Burr density (3.1) the mean is given by $E(X) = \mu + \sigma \frac{1}{\theta - 1}$, $\theta > 1$, so it depends on all the parameters, and the variance by $V(X) = \frac{\theta}{(\theta-1)^2(\theta-2)}\sigma^2$, $\theta > 2$, which does not depend on the location parameter μ. Therefore, the coefficient of variation depends on all the parameters of the distribution, which implies that relative inequality measures are, in general, complicated functions of all the parameters [8].

8. Higher order moments of X in this case can be obtained by considering the *standardized* variable $y = \frac{x-\mu}{\sigma}$, and getting the moments for y. Take care to note that μ and σ are not the mean and standard deviation of X. For y, by integrating by parts recursively, we can

GRAPH 3.1: Burr probability density function



It can easily be seen that for $\sigma = \mu$ the *pdf* (3.1) becomes the *pdf* (2.1), so the Burr density collapses to the Pareto density. Hence (3.1) nest (2.1) and a test of the Pareto density can be performed by testing H_0: $\sigma = \mu$ by means of standard methods. In this context a Lagrange multiplier (LM) test is a natural one since the restricted estimator of θ is (2.12), which can be obtained easily without resorting to numerical methods. We develop such a test in the next section.

show that $E(Y^r) = \prod_{j=1}^{r} \frac{r-j+1}{\theta-j}$, $\theta > r$, so for example $E(Y) = \frac{1}{\theta-1}$, $\theta > 1$ and $E(Y^2) = \frac{2}{(\theta-1)(\theta-2)}$, $\theta > 2$. From this general result we can recover the corresponding moments for $x = \mu + \sigma y$.

4. A Simple and Locally Optimal Test

CONSIDER the Burr density (3.1) where, for the same reason as before, μ is assumed to be known, either fixed by statistical design or taken as given by *a priori* information. Therefore, in practice we have only two free parameters to estimate.

Contrary to the Pareto law, while the power exponent in this case can be estimated by a simple formula (either by the log-log linear regression mentioned before or by the method of maximum likelihood), the power exponent from the Burr density is not so easily obtained and requires numerical optimization. First, we do not obtain a log-log linear relationship between ranks and the ordered sample (1.2). Second, given a random sample of observations, $\{x_i\}_{i=1}^{n}$, the log-likelihood function is given by

$$\log L_B = n \log \theta - n \log \sigma - (\theta + 1) \sum_{i=1}^{n} \log \left(1 + \frac{x_i - \mu}{\sigma}\right), \qquad (4.1)$$

which, even for known μ, has to be maximized with respect to θ and σ by numerical methods.

As mentioned before, for $\sigma = \mu$ $f_B(x) = f_p(x)$. Thus, a natural test for the Pareto law in this context is simply a test of the null hypothesis H_0: $\sigma = \mu$. In particular, given that the restricted model can be easily estimated by maximum likelihood, a Lagrange multiplier (LM) test is a natural one because it avoids the need to maximize (4.1), as required by either the Wald or the likelihood ratio approaches. It can also be routinely calculated once the Pareto exponent has been obtained [9].

9. The same approach is used by Urzúa (2000) to develop a test for Zipf (1949) law, which is a particular case of the Pareto law when $\theta = 1$. In his case, there are not nuisance parameters, but the test is a joint one of the Pareto law and a particular value of the exponent. Thus, in case of a rejection, we cannot know if this is due to the particular value of the exponent assumed or to the Pareto law.

The LM test, as developed by Aitchison and Silvey (1958) and Silvey (1959), is given by $\tilde{q}' \tilde{\mathfrak{I}}_n^{-1} \tilde{q}$, where q represents the score vector, $\mathfrak{I}_n$ the Fischer sample information matrix, both from (4.1), and the symbol ˜ means evaluated at the restricted estimates. The restricted estimates are simply $\sigma = \mu$ and $\tilde{\theta} = \hat{\theta} = \left[\frac{1}{n} \sum_{i=1}^{n} \log \frac{x_i}{\mu} \right]^{-1}$, the maximum likelihood estimator (MLE) of the power exponent under the Pareto law.

Using result 2 above, the appendix shows that from (4.1), the first element of $\tilde{q}$ is simply zero and the second element is $n \frac{(\tilde{\theta}+1)}{\mu} \tilde{z}$, where $z = \frac{\theta}{\theta+1} - \frac{1}{n} \sum_{i=1}^{n} \frac{\mu}{x_i}$. Note that z measures the discrepancy between a population and a sample mean, $E(z) = 0$. Given this factor, only the (2,2) element of $\tilde{\mathfrak{I}}_n^{-1}$ matters, which using the distribution of $y = \frac{\mu}{x}$ is $\frac{1}{n} \cdot \frac{\mu^2 (\tilde{\theta}+2) (\tilde{\theta}+1)^2}{\tilde{\theta}}$. Therefore, the LM statistic we are looking for is

$$LM_p = n \cdot \frac{(\tilde{\theta}+2) (\tilde{\theta}+1)^4}{\tilde{\theta}} \cdot \tilde{z}^2 \overset{asy}{\sim} \chi^2 (1) \quad \text{on } H_0 : \sigma = \mu, \qquad (4.2)$$

and can be readily calculated once the Pareto exponent, $\tilde{\theta}$, has been obtained.

Under the null hypothesis, the LM_P statistic is asymptotically distributed as a chi-square variable with one degree of freedom. Table 4.1 presents the critical values from a Monte Carlo simulation for different sample sizes and power exponents for the Pareto law ranging from 1 to 10 [10]. These values accommodate a wide range of dispersion possibilities, since in terms of the Gini index they cover cases from extreme inequality to cases of very low inequality, with values of the Gini around 0.05 for $\theta = 10$.

The evidence shown in table 4.1 presents evidence that the LM_P statistic is well behaved even for small and medium size samples for values of θ below 3, whereas for a higher value of the Pareto exponent we need a slightly larger sample size to avoid size distortions, which are not very large in any case. Regardless, for samples commonly encountered in empirical applications, which are usually carried out with more than 100 observations, the LM_P statistic performs reasonably well for almost any value of θ. Thus, given

---

10. The parameter μ is set to 1 in all simulations. This is irrelevant in this case and in the simulations of the power against the Burr density, but can be of relevance in the simulations of the power against non-nested distributions.

TABLE 4.1: Empirical critical values for LM_P

Significance level	θ	n 50	100	200	500	1,000	∞
10	1	2.66	2.66	2.67	2.68	2.70	2.71
	2	2.56	2.65	2.66	2.69	2.69	2.71
	3	2.45	2.56	2.64	2.67	2.70	2.71
	5	2.33	2.50	2.61	2.64	2.68	2.71
	10	2.17	2.39	2.53	2.64	2.66	2.71
5	1	3.74	3.79	3.79	3.79	3.84	3.84
	2	3.59	3.74	3.78	3.79	3.81	3.84
	3	3.46	3.62	3.72	3.83	3.87	3.84
	5	3.30	3.55	3.70	3.76	3.82	3.84
	10	3.09	3.40	3.60	3.75	3.80	3.84
1	1	6.41	6.58	6.59	6.61	6.57	6.63
	2	6.43	6.65	6.53	6.62	6.65	6.63
	3	6.55	6.47	6.55	6.63	6.74	6.63
	5	6.88	6.81	6.75	6.72	6.78	6.63
	10	7.14	6.97	6.80	6.69	6.69	6.63

Source: Own Monte Carlo simulations using the direct inversion method and 100,000 replications.

the simplicity of the statistic, this can be routinely calculated when the Pareto exponent is estimated by maximum likelihood methods.

However, not only size but also power is important in testing hypothesis. Hence, we first investigate the power of our statistic against the Burr density, from which the LM_P statistic is derived, but second we also investigate the power against relevant and practical alternatives, such as the exponential or the log-normal distribution, which we find more often in applied studies.

Table 4.2 shows percentages of rejections, at different significant levels, when we simulate random samples from a Burr density under the alternative H_1: $\sigma = \mu.(1 + \delta)$ for different values of δ in the range [–0.5, +0.5]. We focus on asymptotic power so the sample sizes in all experiments were set to 1,000 observations and asymptotic critical values were used in rejecting the null of the Pareto distribution. The main points suggested by the inspection of table 4.2 are the following:

TABLE 4.2: Power of the LM_P test against the Burr distribution. Percentage of rejections, n = 1,000 and asymptotic critical values
(percentages)

Significance level	δ	$H_1: \sigma = \mu.(1+\delta)$				
		1θ	2	3	5	10
10	–0.5	100.0	99.8	97.5	80.9	42.9
	–0.4	99.8	96.7	87.4	63.7	32.2
	–0.3	93.7	80.9	65.7	43.3	23.0
	–0.2	64.3	49.5	38.0	25.6	15.7
	–0.1	25.0	20.5	17.2	14.2	11.4
	0.0	10.0	10.0	9.9	10.0	9.7
	0.1	22.5	18.2	15.3	12.5	10.3
	0.2	51.3	39.5	30.2	20.7	12.9
	0.3	77.9	63.8	50.4	32.7	17.5
	0.4	92.8	82.9	69.6	47.4	23.5
	0.5	98.4	93.6	84.1	62.0	31.1
5	–0.5	100.0	99.6	95.4	73.4	33.2
	–0.4	99.4	94.2	81.1	53.4	23.2
	–0.3	89.2	72.5	54.9	33.4	15.5
	–0.2	52.9	38.1	27.9	17.4	9.6
	–0.1	16.1	12.8	10.6	8.3	6.3
	0.0	5.0	5.1	4.9	5.0	4.8
	0.1	13.8	10.5	8.3	6.3	5.0
	0.2	38.2	27.0	19.2	11.8	6.5
	0.3	67.0	50.4	36.6	20.9	9.5
	0.4	87.1	72.9	56.3	33.8	13.6
	0.5	96.4	88.0	74.1	47.9	19.3
1	–0.5	100.0	98.2	88.2	55.7	18.1
	–0.4	97.6	85.2	64.6	34.0	11.0
	–0.3	75.1	52.5	34.4	17.2	6.4
	–0.2	30.5	19.6	13.1	7.2	3.3
	–0.1	5.7	4.3	3.5	2.6	1.7
	0.0	1.0	1.0	1.0	1.0	1.0
	0.1	4.1	2.7	1.8	1.2	0.9
	0.2	17.0	9.8	5.9	2.9	1.2
	0.3	41.6	25.4	14.9	6.2	2.0
	0.4	68.3	47.0	29.3	12.7	3.3
	0.5	86.9	68.7	47.8	21.7	5.4

Source: Own Monte Carlo simulations using the direct inversion method. 100,000 replications and $\mu = 1$.

I) For $\delta = 0$ the simulation shows that the test has appropriate size, which agrees with the results of table 4.1. This is so for any value of θ or the significance level used in the rejection rule.
II) In general, power is moderate for alternatives close to the null, say for $|\delta| < 0.2$, but is substantial as we depart from the null, say for $|\delta| > 0.2$.
III) The previous point is specially true for low and moderate values of the parameter θ, which are the more relevant from a practical point of view.
IV) The power function is asymmetric with respect to δ. Power is always higher for $\delta < 0$, for a given value of θ and significance level used in rejecting the null.
V) Power decreases uniformly as θ increases for a given $\delta \neq 0$. This is, given a significance level and a value of δ, power falls as θ increases. The reduction in power is substantial for high values of θ, say in excess of 10. However, the relevant range for practical purposes in economics is much lower, not in excess of 3 (Gabaix, 2008). Note that this is not true for $\delta = 0$, so size is correct and maintained across significance levels, δ and θ.
VI) As it is sensible, power is lower at the tails, so for a given δ and θ power falls as the significance level used in rejecting the null decreases.

Overall, the results of these power simulations seem promising. For values of δ close to μ the Pareto and the Burr density are similar enough to justify the used of the Pareto distribution given its simplicity and analytical tractability. However, for alternatives far enough from the Pareto the LM_P statistic shows reasonable power.

However even Burr (1942: 223) was unable to find a probability justification for the function (3.1), so even this *pdf* has been the instrumental density in deriving the test statistic it seems worth investigating the power against more useful distributions from a practical point of view, even they are non-nested. We consider in turn two of such densities, the exponential and the log-normal.

First, we simulate the power of the LM_P statistic against an exponential distribution with parameter θ, defined for $x \geq \mu = 1$. This has a similar shape to the Pareto density with the mode at μ and when $\mu = 1$, as in the simulations, the height of the density at the mode is the same for both distributions. Hence, this particular simulation exercise focuses on the power of the LM_P statistic in relation to the curvature of the both densities. As it is well known, the Pareto density falls quicker than the exponential, specially for low values of θ, but the shape of the Pareto approaches to the exponential shape as θ increases. Table 4.3 shows, in panel *a)*, the results of the Monte Carlo simulation. Power

TABLE 4.3: Power of the LM_P test against non-nested distribution. Percentage of rejections using asymptotic critical values
(percentages)

***a)* Exponential distribution**

Significance level	θ 1	2	3	5	10
10	100.0	100.0	100.0	99.9	83.8
5	100.0	100.0	100.0	99.8	71.7
1	100.0	100.0	100.0	97.7	40.7

***b)* Lognormal distribution**

Significance	Percentage of the tail observations included in the estimation of the Pareto exponent				
level	10	5	2.5	1	0.5
10	100.0	99.2	96.4	88.4	81.1
5	99.9	98.0	92.2	79.5	69.5
1	98.7	90.3	75.0	53.5	40.6

Source: Own Monte Carlo simulations using the direct inversion method and 100,000 replications.

is exactly equal to 1 for $\theta \leq 3$ at any significance level, very close to 1 for $\theta = 5$, and only deteriorates slightly for $\theta = 10$, but even in this case it remains much higher than in the corresponding cases for the Burr density, as the inspection of the last column of table 4.2 shows. Hence, we can be sure that the LM_P statistic has very good power against the exponential density [11].

Second, we simulate the power of the LM_P statistic against the log-normal distribution. Because there is no clear relation between the Pareto exponent and the log-normal distribution there are infinitely many ways of simulating a log-normal in trying the measure the power of the test statistic. We began simply by choosing the parameters of log-normal in such a way so as to match two characteristics with a Pareto [12]. If we match the mode and the median of both distributions we found power equal to 1 in all cases con-

11. Note that the particular value of the μ chosen in the simulations, $\mu = 1$, forces the height of the density at the mode to be the same for both distributions, and altering this height cannot reduce power in any case, so we can be sure that the simulation does not bias our result in favour of giving us good power.

12. These results are not shown, but are available from the author upon request.

sidered [13]. Visual inspection of the densities shows that matching these two moments generates a log-normal density in which the height of the density at the mode is too low in relation to the corresponding Pareto density, and also the curvature of the density at the mode is very smooth, so a power equal to 1 is a sensible outcome. Next, we chose parameters in such a way so as to match the mode and the height of the density at the mode for both distributions [14]. Again, power was equal to 1 in all cases considered.

However, these are probably not the interesting cases to consider because visual inspection would reveal quite easily that the reference distribution is not of the Pareto family (Eeckhout, 2004). Given that proportional growth with some friction generates a steady state distribution of the Pareto law in the upper tail (Champernowne, 1953; Gabaix, 1999, 2008) we simulate random samples from a log-normal distribution with both parameters equal to 1, take a given percentage of observations from the upper tail and construct the LM_P statistic only with those observations. The statistic is always constructed with 1,000 observations, so the sample size drawn at each replication is adjusted in relation to the percentage of observations from the upper tail finally used in constructing the LM_P statistic. The results of such an experiment are in panel *b)* of table 4.3, where the percentages of the observations at the tail are 10, 5, 2.5, 1 and 0.5%. Again the results are encouraging, keeping as little as 2.5% of the observations the LM_P statistic discriminates correctly in more than 90% the cases at the 5% significance level or above. Even focusing on the extreme tail observations the power is remarkable.

Overall we see the results of the power analysis quite successful. The LM_P statistic seems more powerful against interesting and practical alternatives than against the Burr distribution, which is the one used to derive the test. Of course this has to be studied more thoroughly, but our limited Monte Carlo exercise signals good power properties of the statistic derived.

---

13. If ν and τ^2 are the (location and shape) parameters of the log-normal distribution (which correspond to the mean and variance of the corresponding normal parent distribution), fixing $\mu = 1$ for the Pareto density, the mode of both distributions is the same when $\nu = \tau^2$. In addition, equating the median implies $\nu = \log 2/\theta$. Hence for a given value of θ we simulate random samples from a log-normal distribution with parameters $\nu = \tau^2 = \log 2/\theta$. Of course in the estimation of the Pareto exponent the simulated values lower than $\mu = 1$ were disregarded, so only values to the right of the mode are included in the estimation.

14. As we have already seen, fixing $\mu = 1$ for the Pareto density and equating the mode of both distributions implies $\nu = \tau^2$ for the lognormal. In addition, equating the height of the densities at the mode implies finding the root of a nonlinear equation, $\tau . e^{\frac{\tau^2}{2}} - \frac{1}{\theta . \sqrt{2\pi}} = 0$, fortunately this equation has only one root. Hence for a given value of θ we solve the nonlinear equation and, given the solution, we simulate random samples from a log-normal distribution imposing $\nu = \tau^2$. As before, in the estimation of the Pareto exponent the simulated values lower than $\mu = 1$ were disregarded, so again only values to the right of the mode are included in the estimation.

5. An Application to Urban Economics

As an example of the use of the LM_P statistic, consider the Spanish urban areas as currently defined by the Spanish Department of Public Works (Ministerio de Fomento, 2000). Not only are most of the population agglomerations with over 10,000 inhabitants included, but also those which fulfil certain other characteristics in terms of economic activity, population trends or mobility and housing. Fixing a lower bound at 10,000 inhabitants and using data from all the census from 1900 up to 2001 we have an increasing number of *cities,* ranging from 157 in 1900 to 294 in 2001 [15].

Table 5.1 offers, in its first panel, the maximum likelihood estimator (MLE) of the Pareto exponent as well as the corresponding LM_P statistic derived in the previous section and its asymptotic p-value [16]. A clear pattern emerges from the estimates of the Pareto exponent, a steadily and monotonic decreasing tendency, from 1.25 in 1900 to 0.86 in 2001. This should be understood as an increasing dispersion or inequality in the distribution of city size, a fact that is confirmed by the results in Goerlich and Mas (2009).

Turning to the question of the suitability of the Pareto distribution as an adequate model for this data, we observe that the LM_P statistic offers little doubts about the validity of the Pareto law for the 20th century. For these years the maximum value of the statistic is reached in 1991, 2.63, with a p-value of about 11%. However at the beginning of the 21st century we observe a clear tendency to depart from the Pareto distribution, since this is clearly rejected for the last census. Using table 4.1, or asymptotic critical values, the p-value for the LM_P statistic in 2001 is significantly lower than 1%. Hence the decreasing tendency in the Pareto exponent is translated into a departure from this distribution. Therefore, in terms of modeling the city

15. It is worth mentioning that that there is no universally accepted definition of city for statistical purposes and that alternative definitions account for many of the empirical disagreement in the literature, as recognized by Gabaix and Ioannides (2004). See Goerlich and Mas (2009) for a much more detailed analysis of this data set.

16. For each year, μ is estimated as the minimum of the corresponding sample, but the same results are obtained if μ is assumed to be known and fixed at the lower bound value, 10,000.

TABLE 5.1: Pareto law and city size distribution

	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Lower bound = 10,000											
Cities	157	168	185	197	214	219	239	265	287	292	294
$\hat{\theta}$	1.247	1.180	1.157	1.096	1.033	0.976	0.967	0.955	0.915	0.888	0.859
LM_P statistic	0.953	0.015	0.004	0.724	1.172	2.037	0.733	0.029	0.368	2.626	7.414
p-value	0.329	0.903	0.948	0.395	0.279	0.154	0.392	0.866	0.544	0.105	0.006
Lower bound = 50,000											
Cities	25	26	28	30	38	45	49	54	64	67	72
$\hat{\theta}$	1.227	1.142	0.974	0.923	0.972	0.998	0.941	0.844	0.839	0.834	0.841
LM_P statistic	0.295	1.452	3.124	2.235	0.508	0.133	0.263	0.713	0.224	0.416	0.228
p-value	0.587	0.228	0.077	0.135	0.476	0.715	0.608	0.398	0.636	0.519	0.633

Source: Own calculations from census data and Ministerio de Fomento (2000). Urban area as city definition. See Goerlich and Mas (2009).

size distribution, alternatives to the Pareto law should be explored for the recent years.

It seems interesting to see what happens if we increase the minimum number of inhabitants for the definition of a city to 50,000. In this case, the number of *cities* falls to a number between 25 in 1900 and 72 in 2001, so we have a really small data set in this case. The Pareto exponent shows a similar decreasing tendency up to 1960, but exhibits a remarkably stability since then, with a value around 0.84. What is more interesting in this case is that, for this restrictive definition of city, the LM_P statistic is unable to reject the null of the Pareto law in any of the years. Only for 1920 we get a value of the statistic of 3.12 with a *p*-value between 5 and 10%, but the overall result is that the Pareto law seems a sensible distribution for this reduced data set. We have not conflicting results.

As a consequence, this probability law can be a sensible approximation for the tail distribution of a given sample, but how to define the tail from a practical point of view is not a trivial question. That is to say, the Pareto law cannot hold, if at all, except for a certain sample size.

Appendix

THIS appendix provides some derivations mentioned in the text and also the manipulation of the likelihood function of the Burr density to obtain the Lagrange multiplier (LM) test.

Pareto law, Theil indexes and related results

Derivation of T^*. To obtain (2.6) from (2.5) we have to solve

$$T^* = \int_0^1 \log \left(\frac{\theta}{\theta - 1} (1-p)^{\frac{1}{\theta}} \right) dp = \log \frac{\theta}{\theta - 1} + \frac{1}{\theta} \int_0^1 \log (1-p)\, dp. \quad \text{(A.1)}[17]$$

Using the fact that $\lim_{z \to 0} z \log z = 0$, we can show (integrating by parts) that $\int_0^1 \log (1-p)\, dp = -[(1-p). \log (1-p) - (1-p)]\big|_0^1 = -1$, which gives us (2.6).

Derivation of *T*. To obtain (2.9) from (2.8) we have to solve

$$T = \int_0^1 \frac{\theta - 1}{\theta} (1-p)^{-\frac{1}{\theta}} \log \left(\frac{\theta - 1}{\theta} (1-p)^{-\frac{1}{\theta}} \right) dp. \quad \text{(A.2)}$$

After some algebra, and using the following results $\int_0^1 (1-p)^{-\frac{1}{\theta}}\, dp = \frac{\theta}{\theta - 1}$, $\lim_{z \to 0} z^{\delta} \log z = 0, \delta > 0$ and $\int_0^1 (1-p)^{-\frac{1}{\theta}} \log (1-p)\, dp = -\frac{\theta^2}{(1-\theta)^2}$ [18], we eventually get (2.9).

17. Note that, the geometric mean for a Pareto random variable is given by $m_1 = \mu . e^{\frac{1}{\theta}}$, since $\log m_1 = \int_0^1 \log \mu\, (1-p)^{-\frac{1}{\theta}}\, dp = \log \mu + \frac{1}{\theta}$.

18. In addition to the previous limit, this integral uses the following general result (Spiegel, 1970: 14.527),

$$\int z^{\delta} \log z dz = \frac{z^{\delta + 1}}{\delta + 1} \left(\log z - \frac{1}{\delta + 1} \right) \qquad \delta \neq 1.$$

Result 1 – Proof. This comes from standard results on the distribution of functions of random variables (Mood, Graybill and Boes, 1974, chapter V). Because of the one-to-one relation between x and y_1, $x = \mu e^{y1}$, the *pdf* of y_1 is given by

$$f(y_1) = \left| \frac{dx}{dy_1} \right| f_p(x(y_1)) = \left| \mu e^{y_1} \right| \cdot \frac{\theta\mu^\theta}{(\mu e^{y_1})^{\theta+1}} = \theta e^{-\theta y_1} \qquad y_1 > 0, \qquad \text{(A.3)}$$

which is the *pdf* of an exponential distribution with parameter θ. Thus, $E(y_1) = \dfrac{1}{\theta}$ and $Var(y_1) = \dfrac{1}{\theta^2}$. ■

Result 2 – Proof. Because of the one-to-one relation between x and y_2, $x = \dfrac{\mu}{y_2}$, the *pdf* of y_2 is given by

$$f(y_2) = \left| \frac{dx}{dy_2} \right| f_p(x(y_2)) = \left| -\frac{\mu}{y_2^2} \right| \cdot \frac{\theta\mu^\theta}{\left(\dfrac{\mu}{y_2}\right)^{\theta+1}} = \theta \cdot y_2^{\theta-1} \qquad 0 < y_2 < 1, \qquad \text{(A.4)}$$

which is the *pdf* of a power function distribution with parameter θ. Thus, $E(y_2) = \dfrac{\theta}{\theta+1}$, $E(y_2^2) = \dfrac{\theta}{\theta+2}$ and in general $E(y_2^r) = \dfrac{\theta}{\theta+r}$, for $r > 0$. ■

Result 3 – Proof. For $\theta = 1$ (A.4) becomes $f(y_2) = 1$, which is the *pdf* of a uniform distribution on [0,1]. Thus, $E(y_2) = ½$, $Var(y_2) = \dfrac{1}{12}$ and $E(y_2^2) = \dfrac{1}{3}$. ■

Remark 2. Given result 1, $z = 2\theta y_1$ follows a χ^2 with 2 degrees of freedom since, by standard arguments, the *pdf* of z is given by

$$f(z) = \left| \frac{dy_1}{dz} \right| f(y_1(z)) = \left| \frac{1}{2\theta} \right| \cdot \theta e^{-\theta \frac{z}{2\theta}} = \frac{1}{2} e^{-\frac{z}{2}} \qquad z > 0, \qquad \text{(A.5)}$$

which is the *pdf* of a χ^2 distribution with 2 degrees of freedom. Thus, $E(z) = 2$ and $Var(z) = 4$.

From this result, given a random sample from the Pareto distribution, and the fact that a sum of independent χ^2 is also χ^2 distributed with degrees of freedom equal to the sum of the degrees of freedom of each term in the sum, we get

$$2\theta \sum_{i=1}^{n} \log \frac{x_i}{\mu} = \frac{2n\theta}{\hat{\theta}} \sim \chi^2 \,(2n). \qquad \text{(A.6)}^{19}$$

Observing that the *pdf* of χ^2 $(2n)$ is $\frac{1}{\Gamma(n)} \cdot \frac{1}{2^n} \cdot x^{n-1} \cdot e^{-\frac{1}{2}x}$, by standard arguments, the *pdf* of $\hat{\theta}$ is

$$f_{\hat{\theta}}\,(x) = \left| -\frac{2n\theta}{x^2} \right| \frac{1}{\Gamma(n)} \cdot \frac{1}{2^n} \cdot \left(\frac{2n\theta}{x} \right)^{n-1} \cdot e^{-\frac{1}{2}\frac{2n\theta}{x}} = \frac{n^n \theta^n}{\Gamma(n) x^{n+1}} \cdot e^{-\frac{n\theta}{x}} \qquad x > 0, \qquad \text{(A.7)}^{20}$$

where $\Gamma(\bullet)$ is the Gamma function.

The LM test statistic for the Pareto law

In the sequel, we obtain the first and second derivatives of log L_B with respect to θ and σ, the information matrix, and next we derive the LM statistic for H_0: $\sigma = \mu$.

The gradient of (4.1) is given by

$$\begin{aligned} \frac{\partial \log L_B}{\partial \theta} &= \frac{n}{\theta} - \sum_{i=1}^{n} \log \left(1 + \frac{x_i - \mu}{\sigma} \right), \\ \frac{\partial \log L_B}{\partial \sigma} &= -\frac{n}{\sigma} + (\theta + 1) \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2 \left(1 + \frac{x_i - \mu}{\sigma} \right)}. \end{aligned} \qquad \text{(A.8)}$$

Note that, by equating these equations to zero we have a system of non-linear equations that have to be solved by numerical methods. Therefore the maximum likelihood estimator (MLE) of the parameters for the Burr density does not have an explicit formula.

The Hessian is given by

19. This is a general result, the sample mean of a random sample from an exponential distribution with parameter θ is distributed as $\frac{1}{2} \cdot \frac{1}{\theta} \cdot \frac{1}{n} \cdot \chi^2\,(2n)$.

20. This is slightly different than the *pdf* shown in Johnson and Kotz (1970, chapter 19.4.4), $f_{\hat{\theta}}\,(x) = \frac{n^{n-1}\,\theta^{n-1}}{\Gamma(n-1)x^n} \cdot e^{-\frac{n\theta}{x}}$, which is due to the fact that they consider the case where μ is estimated as $\hat{\mu} = \min_i \{x_i\}_{i=1}^{n}$, hence the distribution in (A.6) is $\chi^2(2(n-1))$.

$$\frac{\partial \log^2 L_B}{\partial \theta^2} = \frac{n}{\theta^2},$$

$$\frac{\partial \log^2 L_B}{\partial \theta \partial \sigma} = \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2 \left(1 + \frac{x_i - \mu}{\sigma}\right)}, \tag{A.9}$$

$$\frac{\partial \log^2 L_B}{\partial \sigma^2} = \frac{n}{\sigma^2} - (\theta + 1) \sum_{i=1}^{n} \frac{(x_i - \mu)(2\sigma + x_i - \mu)}{\sigma^2 (\sigma + x_i - \mu)^2}.$$

Evaluating (A.8) at the restricted estimate of σ, $\hat{\sigma} = \mu$, we have

$$\left.\frac{\partial \log L_B}{\partial \theta}\right|_{\sigma=\mu} = \frac{n}{\theta} - \sum_{i=1}^{n} \log\left(\frac{x_i}{\mu}\right) = n\left[\frac{1}{\theta} - \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{x_i}{\mu}\right)\right]. \tag{A.10a}$$

Two remarks

I) Given that the restricted estimate of θ is just the Pareto MLE, $\hat{\theta}$ (2.12), it is worth noting that evaluating (A.10a) at $\hat{\theta}$ is just zero, $\left.\frac{\partial \log L_B}{\partial \theta}\right|_{\theta=\hat{\theta},\ \sigma=\mu} = 0.$

II) Defining $z_1 = \frac{1}{\theta} - \frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{x_i}{\mu}\right)$, a function of θ, (A.10a) measures a discrepancy between a population mean and a sample mean, since given result 1 $E(z_1) = 0$. This gives the MLE of θ under the Pareto law a nice moment interpretation.

And

$$\left.\frac{\partial \log L_B}{\partial \theta}\right|_{\hat{\sigma}=\mu} = -\frac{n}{\mu} + (\theta + 1) \sum_{i=1}^{n} \frac{x_i - \mu}{\mu x_i} =$$
$$= \frac{n(\theta + 1)}{\mu}\left[\frac{\theta}{\theta + 1} - \frac{1}{n}\sum_{i=1}^{n} \frac{\mu}{x_i}\right]. \tag{A.10b}$$

One remark

I) Defining $z_2 = \frac{\theta}{\theta + 1} - \frac{1}{n} \sum_{i=1}^{n} \frac{\mu}{x_i}$, a function of θ, (A.10b)

measures a discrepancy between a population mean and a sample mean, since given result 2 $E\,(z_2) = 0$.

Evaluating (A.9) at the restricted estimate of σ, $\hat{\sigma} = \mu$, we have

$$\left.\frac{\partial \log^2 L_B}{\partial \theta \partial \sigma}\right|_{\sigma = \mu} = \sum_{i=1}^{n} \frac{x_i - \mu}{\mu x_i} = \frac{n}{\mu}\left[1 - \frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{x_i}\right], \tag{A.11a}$$

changing its sign and evaluating the expectation,

$$E\left[-\left.\frac{\partial \log^2 L_B}{\partial \theta \partial \sigma}\right|_{\sigma = \mu}\right] = -\frac{n}{\mu}\left[1 - E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{x_i}\right)\right] =$$

$$= -\frac{n}{\mu}\,[1 - E\,(y_2)] = -\frac{n}{\mu}\left[1 - \frac{\theta}{\theta + 1}\right] = -\frac{n}{\mu\,(\theta + 1)}\,. \tag{A.12a}$$

And

$$\left.\frac{\partial \log^2 L_B}{\partial \sigma^2}\right|_{\sigma = \mu} = \frac{n}{\mu^2} - (\theta + 1)\sum_{i=1}^{n}\frac{(x_i - \mu)\,(x_i + \mu)}{\mu^2 x_i^2} =$$

$$= -\frac{n\,(\theta + 1)}{\mu^2}\left[\frac{\theta}{\theta + 1} - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mu^2}{x_i^2}\right)\right], \tag{A.11b}$$

changing its sign and evaluating the expectation,

$$E\left[\left.\frac{\partial \log^2 L_B}{\partial \sigma^2}\right|_{\sigma = \mu}\right] = \frac{n\,(\theta + 1)}{\mu^2}\left[\frac{\theta}{\theta + 1} - E\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mu^2}{x_i^2}\right)\right)\right] =$$

$$= \frac{n\,(\theta + 1)}{\mu^2}\left[\frac{\theta}{\theta + 1} - E\,(y_2^2)\right] = \frac{n\,(\theta + 1)}{\mu^2}\left[\frac{\theta}{\theta + 1} - \frac{\theta}{\theta + 2}\right] = \tag{A.12b}$$

$$= \frac{n\theta}{\mu^2\,(\theta + 2)}\,.$$

Thus, the Fisher sample information, evaluated at $\sigma = \mu$, is given by

$$\Im\,(\theta, \sigma = \mu) = \begin{bmatrix} \dfrac{n}{\theta^2} & -\dfrac{n}{\mu\,(\theta+1)} \\ -\dfrac{n}{\mu\,(\theta+1)} & \dfrac{n\theta}{\mu^2\,(\theta+2)} \end{bmatrix} = \frac{n}{\theta^2} \begin{bmatrix} 1 & -\dfrac{\theta^2}{\mu\,(\theta+1)} \\ -\dfrac{\theta^2}{\mu\,(\theta+1)} & \dfrac{\theta^3}{\mu^2\,(\theta+2)} \end{bmatrix}. \quad \text{(A.13)}$$

And its inverse is

$$\Im^{-1}\,(\theta, \sigma = \mu) = \frac{1}{n} \cdot \frac{\mu^2\,(\theta+2)\,(\theta+1)^2}{\theta} \begin{bmatrix} \dfrac{\theta^3}{\mu^2\,(\theta+2)} & \dfrac{\theta^2}{\mu\,(\theta+1)} \\ \dfrac{\theta^2}{\mu\,(\theta+1)} & 1 \end{bmatrix}. \quad \text{(A.14)}$$

Now we can form the LM statistic. Noting that $\hat{z}_1 = \dfrac{1}{\hat{\theta}} - \dfrac{1}{n} \sum_{i=1}^{n} \log \left(\dfrac{x_i}{\mu}\right) = 0$, the score vector evaluated at the restricted estimate is

$$\tilde{q} = n \begin{bmatrix} \hat{z}_1 \\ \dfrac{(\hat{\theta}+1)}{\mu}\,\hat{z}_2 \end{bmatrix} = n \begin{bmatrix} 0 \\ \dfrac{(\hat{\theta}+1)}{\mu}\,\hat{z}_2 \end{bmatrix}_{2\times 1}, \quad \text{(A.15)}$$

with $\hat{z}_2 = \dfrac{\hat{\theta}}{\hat{\theta}+1} - \dfrac{1}{n} \sum_{i=1}^{n} \dfrac{\mu}{x_i}$. Therefore, only the (2,2) element of (A.15) matters in the LM statistic, $\dfrac{1}{n} \cdot \dfrac{\mu^2\,(\theta+2)\,(\theta+1)^2}{\theta}$. The feasible LM statistic is then given by

$$LM_p = n \,.\, \frac{(\hat{\theta}+2)\,(\hat{\theta}+1)^4}{\hat{\theta}} \,.\, \hat{z}_2^2 \overset{asy}{\sim} \chi^2\,(1) \quad \text{on } H_0 : \sigma = \mu, \quad \text{(A.16)}$$

and can be readily calculated once $\hat{\theta}$ has been obtained.

References

ABAN, I. B. and M. M. MEERSCHAERT (2004): "Generalized Least-Squares Estimators for the Thickness of Heavy Tails", *Journal of Statistical Planning and Inference,* 119, 341-352.

AITCHISON, J. and S. D. SILVEY (1958): "Maximum-Likelihood Estimation of Parameters Subject to Restraints", *The Annals of Mathematical Statistics,* 29 (3), 813-828, September.

BEIRLANT, J., Y. GOEGEBEUR, J. SEGERS and J. TEUGELS (2004): *Statistics of Extremes: Theory and Applications,* England, Wiley.

BROWN, J. H. and G. B. WEST (eds.) (2000): *Scaling in Biology,* Oxford, Oxford University Press.

BURR, I. W. (1942): "Cumulative Frequency Functions", *The Annals of Mathematical Statistics,* 13, (2), 215-232, June.

CHAMPERNOWNE, D. (1953): "A Model of Income Distribution", *The Economic Journal,* 63 (250), 318-351, June.

CLAUSET, A., C. R. SHALIZI and M. E. J.NEWMAN (2007): *Power-Law Distributions in Empirical Data,* e-print, arXiv:0706.1062v1 [physics.data-an] (June 7th). [http://arxiv.org/abs/0706.1062v1].

COWELL, F. A. (1977): *Measuring Inequality* (1st edition), Oxford, Phillip Allan (2nd edition, 1995).

DALTON, H. (1920): "The Measurement of Inequality of Income", *The Economic Journal,* 30, 348-361.

EECKHOUT, J. (2004): "Gibrat's Law for (all) Cities", *The American Economic Review,* 94 (5), 1429-1451, December.

EMBRECHTS, P., P. KLUPPELBERG and C. MIKOSCH (1997): *Modelling Extremal Events for Insurance and Finance,* New York, Springer-Verlag.

FAN, C. C. and E. CASETTI (1994): "The Spatial and Temporal Dynamics of US Regional Income Inequality, 1950-1989", *Annals of Regional Science,* 28, 177-196.

GABAIX, X. (1999): "Zipf's Law for Cities: An Explanation", *The Quarterly Journal of Economics,* 114 (3), 739-767, August.

— (2008): "Power Laws in Economics and Finance", *National Bureau of Economic Research,* Working Paper 14299, September.

— and R. IBRAGIMOV (2008): "A Simple OLS Test of Power Law Behaviour", Working Paper, Harvard University.

— and Y. M. IOANNIDES (2004): "The Evolution of City size Distributions", in J. V. Henderson and J. F. Thisse (eds.), *Handbook of Regional and Urban Economics,* 4, chapter 53, Amsterdam, North-Holland Publishing Company.

GIBRAT, R. (1931): *Les Inégalités économiques. Applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel,* Paris, Libraire du Recueil Sirey.

GINI, C. (1912): "Variabilità e mutabilità, contributo allo studio delle distribuzioni e relazioni statistiche", *Studi Economico-Giuridici dell' Universiti di Cagliari,* 3, part 2, 1-158.

GOERLICH, F. J. and M. MAS (2009): "La distribución empírica del tamaño de las ciudades en España, 1900-2001. ¿Quién verifica la ley de Zipf?", *Revista de Economía Aplicada* [forthcoming].

— and A. VILLAR (2009): *Desigualdad y bienestar social. De la teoría a la práctica,* Bilbao, Foundation BBVA.

HILL, B. M. (1975): "A Simple General Approach to Inference about the Tail of a Distribution", *The Annals of Statistics,* 3 (5), 1163-1174, September.

JOHNSON, N. L. and S. KOTZ (1970): "Distributions in Statistics: Continuous Univariate Distributions", 1, Boston, Houghton Mifflin Company.

LEVY, M. and S. SOLOMON (1996): "Power Laws Are Logarithmic Boltzmann Laws", *International Journal of Modern Physics,* C, 7, (4), 595-601, August. [Available in e-print arXiv:adap-org/9607001v1, physics.data-an, (July 15th) http://arxiv.org/abs/adap-org/9607001v1].

LORENZ, M. C. (1905): "Methods of Measuring the Concentration of Wealth", *Publications of the American Statistical Association,* 9, 209-219.

MINISTERIO DE FOMENTO (2000): *Atlas estadístico de las áreas urbanas de España,* 1st edition, *Centro de Publicaciones,* Madrid, Ministerio de Fomento.

MOOD, A. M., F. A. GRAYBILL and D. C. BOES (1974): "Introduction to the Theory of Statistics", Third Edition, *International Student Edition,* London, McGraw-Hill Book Company.

NISHIYAMA, Y. and S. OSADA (2004): "Statistical Theory of Rank Size Rule Regression under Pareto Distribution", Interfaces for Advanced Economic Analysis, Kyoto University, Discussion Paper 009, 21COE, January.

— S. OSADA and Y. SATO (2006): "OLS-t-Test Revisited in Rank Size Rule Regression", Kyoto Institute of Economic Research, Kyoto University, DEE Discussion Paper 06-3.

PARETO, V. (1896): "Cours d'Economie Politique", Geneva, Switzerland, Droz.

PIGOU, A. C. (1912): "The Economics of Welfare", London, Transaction Publishers.

QUANDT, R. E. (1966): "Old and New Methods of Estimation and the Pareto Distribution", *Metrika,* 10, 55-82.

ROSEN, K. T. and M. RESNICK (1980): "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy", *Journal of Urban Economics,* 8, 165-186.

SILVEY, S. D. (1959): "The Lagrangian Multiplier Test", *The Annals of Mathematical Statistics,* 30 (2), 389-407, June.

SIMON, H. (1955): "On a Class of Skew Distribution Functions", *Biometrika,* 42, 425-440.

SORNETTE, D. (2004): "Critical Phenomena in Natural Sciences", Berlin, New York, Springer.

— and R. CONT (1997): "Convergent Multiplicative Processes Repelled from Zero: Power Laws and Truncated Power Laws", *Journal of Physique I France,* 7, 431-444.

SPANOS, A. (1986): "Statistical Foundations of Econometric Modelling", Cambridge, Cambridge University Press.

SPIEGEL, M. R. (1970): *Manual de fórmulas y tablas matemáticas,* Schaum, New York, McGraw-Hill.

SUTTON, J. (1997): "Gibrat's Legacy", *Journal of Economic Literature,* 35 (1), 40-59, March.

THEIL, H. (1967): *Economics and Information Theory,* Amsterdam, North-Holland.

URZÚA, C. M. (2000): "A Simple and Efficient Test for Zipf's Law", *Economics Letters,* 66, 257-260.

ZIPF, G. (1949): *Human Behavior and the Principle of Least Effort,* Cambridge, MA, Addison-Wesley.

*ABOUT THE AUTHOR**

FRANCISCO J. GOERLICH GISBERT holds a PhD in economics from the University of Valencia and an MSc in economics from the London School of Economics, University of London. He is currently professor at the Department of Economic Analysis of the University of Valencia and senior researcher at the Valencian Economic Research Institute (Ivie). His research fields are applied econometrics, regional economics and income distribution. He has published more than forty articles in national and international specialized journals and has collaborated on more than ten books.
E-mail: francisco.j.goerlich@uv.es

Any comments on the contents of this paper can be addressed to Francisco J. Goerlich Gisbert at francisco.j.goerlich@uv.es.

* The author acknowledges financial support by the Spanish Science and Education Ministry SEC2008-03813/ECON project, and the BBVA Foundation-Valencian Economic Research Institute (Ivie) research program. Results mentioned, but not shown, are available from the author upon request.

Fundación BBVA

D O C U M E N T O S   D E   T R A B A J O

NÚMEROS PUBLICADOS

DT 01/02 *Trampa del desempleo y educación: un análisis de las relaciones entre los efectos desincentivadores de las prestaciones en el Estado del Bienestar y la educación*
Jorge Calero Martínez y Mónica Madrigal Bajo

DT 02/02 *Un instrumento de contratación externa: los vales o cheques. Análisis teórico y evidencias empíricas*
Ivan Planas Miret

DT 03/02 *Financiación capitativa, articulación entre niveles asistenciales y descentralización de las organizaciones sanitarias*
Vicente Ortún-Rubio y Guillem López-Casasnovas

DT 04/02 *La reforma del IRPF y los determinantes de la oferta laboral en la familia española*
Santiago Álvarez García y Juan Prieto Rodríguez

DT 05/02 *The Use of Correspondence Analysis in the Exploration of Health Survey Data*
Michael Greenacre

DT 01/03 *¿Quiénes se beneficiaron de la reforma del IRPF de 1999?*
José Manuel González-Páramo y José Félix Sanz Sanz

DT 02/03 *La imagen ciudadana de la Justicia*
José Juan Toharia Cortés

DT 03/03 *Para medir la calidad de la Justicia (I): Abogados*
Juan José García de la Cruz Herrero

DT 04/03 *Para medir la calidad de la Justicia (II): Procuradores*
Juan José García de la Cruz Herrero

DT 05/03 *Dilación, eficiencia y costes: ¿Cómo ayudar a que la imagen de la Justicia se corresponda mejor con la realidad?*
Santos Pastor Prieto

DT 06/03 *Integración vertical y contratación externa en los servicios generales de los hospitales españoles*
Jaume Puig-Junoy y Pol Pérez Sust

DT 07/03 *Gasto sanitario y envejecimiento de la población en España*
Namkee Ahn, Javier Alonso Meseguer y José A. Herce San Miguel

DT 01/04 *Métodos de solución de problemas de asignación de recursos sanitarios*
Helena Ramalhinho Dias Lourenço y Daniel Serra de la Figuera

DT 01/05 *Licensing of University Inventions: The Role of a Technology Transfer Office*
Inés Macho-Stadler, David Pérez-Castrillo y Reinhilde Veugelers

DT 02/05 *Estimating the Intensity of Price and Non-price Competition in Banking: An Application to the Spanish Case*
Santiago Carbó Valverde, Juan Fernández de Guevara Radoselovics, David Humphrey y Joaquín Maudos Villarroya

DT 03/05 *Sistemas de pensiones y fecundidad. Un enfoque de generaciones solapadas*
Gemma Abío Roig y Concepció Patxot Cardoner

DT 04/05 *Análisis de los factores de exclusión social*
Joan Subirats i Humet (Dir.), Ricard Gomà Carmona y Joaquim Brugué Torruella (Coords.)

DT 05/05 *Riesgos de exclusión social en las Comunidades Autónomas*
Joan Subirats i Humet (Dir.), Ricard Gomà Carmona y Joaquim Brugué Torruella (Coords.)

DT 06/05 *A Dynamic Stochastic Approach to Fisheries Management Assessment: An Application to some European Fisheries*
José M. Da-Rocha Álvarez y María-José Gutiérrez Huerta

DT 07/05 *The New Keynesian Monetary Model: Does it Show the Comovement between Output and Inflation in the U.S. and the Euro Area?*
Ramón María-Dolores Pedrero y Jesús Vázquez Pérez

DT 08/05 *The Relationship between Risk and Expected Return in Europe*
Ángel León Valle, Juan Nave Pineda y Gonzalo Rubio Irigoyen

DT 09/05 *License Allocation and Performance in Telecommunications Markets*
Roberto Burguet Verde

DT 10/05 *Procurement with Downward Sloping Demand: More Simple Economics*
Roberto Burguet Verde

DT 11/05 *Technological and Physical Obsolescence and the Timing of Adoption*
Ramón Caminal Echevarría

DT 01/06 *El efecto de la inmigración en las oportunidades de empleo de los trabajadores nacionales: Evidencia para España*
Raquel Carrasco Perea, Juan Francisco Jimeno Serrano y Ana Carolina Ortega Masagué

DT 02/06 *Inmigración y pensiones: ¿Qué sabemos?*
José Ignacio Conde-Ruiz, Juan Francisco Jimeno Serrano y Guadalupe Valera Blanes

DT 03/06 *A Survey Study of Factors Influencing Risk Taking Behavior in Real World Decisions under Uncertainty*
Manel Baucells Alibés y Cristina Rata

DT 04/06 *Measurement of Social Capital and Growth: An Economic Methodology*
Francisco Pérez García, Lorenzo Serrano Martínez, Vicente Montesinos Santalucía y Juan Fernández de Guevara Radoselovics

DT 05/06 *The Role of ICT in the Spanish Productivity Slowdown*
Matilde Mas Ivars y Javier Quesada Ibáñez

DT 06/06 *Cross-Country Comparisons of Competition and Pricing Power in European Banking*
David Humphrey, Santiago Carbó Valverde, Joaquín Maudos Villarroya y Philip Molyneux

DT 07/06 *The Design of Syndicates in Venture Capital*
Giacinta Cestone, Josh Lerner y Lucy White

DT 08/06 *Efectos de la confianza en la información contable sobre el coste de la deuda*
Belén Gill de Albornoz Noguer y Manuel Illueca Muñoz

DT 09/06 *Relaciones sociales y envejecimiento saludable*
Ángel Otero Puime, María Victoria Zunzunegui Pastor, François Béland, Ángel Rodríguez Laso y María Jesús García de Yébenes y Prous

DT 10/06 *Ciclo económico y convergencia real en la Unión Europea: Análisis de los PIB per cápita en la UE-15*
José Luis Cendejas Bueno, Juan Luis del Hoyo Bernat, Jesús Guillermo Llorente Álvarez, Manuel Monjas Barroso y Carlos Rivero Rodríguez

DT 11/06 *Esperanza de vida en España a lo largo del siglo XX: Las tablas de mortalidad del Instituto Nacional de Estadística*
Francisco José Goerlich Gisbert y Rafael Pinilla Pallejà

DT 12/06 *Convergencia y desigualdad en renta permanente y corriente: Factores determinantes*
Lorenzo Serrano Martínez

DT 13/06 *The Common Agricultural Policy and Farming in Protected Ecosystems: A Policy Analysis Matrix Approach*
Ernest Reig Martínez y Vicent Estruch Guitart

DT 14/06 *Infrastructures and New Technologies as Sources of Spanish Economic Growth*
Matilde Mas Ivars

DT 15/06 *Cumulative Dominance and Heuristic Performance in Binary Multi-Attribute Choice*
Manel Baucells Alibés, Juan Antonio Carrasco López y Robin M. Hogarth

DT 16/06 *Dynamic Mixed Duopoly: A Model Motivated by Linux versus Windows*
Ramon Casadesus-Masanell y Pankaj Ghemawat

DT 01/07 *Social Preferences, Skill Segregation and Wage Dynamics*
Antonio Cabrales Goitia, Antoni Calvó-Armengol y Nicola Pavoni

DT 02/07 *Stochastic Dominance and Cumulative Prospect Theory*
Manel Baucells Alibés y Franz H. Heukamp

DT 03/07 *Agency Revisited*
Ramon Casadesus-Masanell y Daniel F. Spulber

DT 04/07 *Social Capital and Bank Performance: An International Comparison for OECD Countries*
José Manuel Pastor Monsálvez y Emili Tortosa-Ausina

DT 05/07 *Cooperation and Cultural Transmission in a Coordination Game*
Gonzalo Olcina Vauteren y Vicente Calabuig Alcántara

DT 06/07 *The* Extended *Atkinson Family and Changes in Expenditure Distribution: Spain 1973/74 – 2003*
Francisco J. Goerlich Gisbert, María Casilda Lasso de la Vega Martínez y Ana Marta Urrutia Careaga

DT 07/07 *Análisis de la evolución de la dependencia en la tercera edad en España*
David Casado Marín

DT 08/07 *Designing Contracts for University Spin-offs*
Inés Macho-Stadler, David Pérez-Castrillo y Reinhilde Veugelers

DT 09/07 *Regional Differences in Socioeconomic Health Inequalities in Spain*
Pilar García Gómez y Ángel López Nicolás

DT 10/07 *The Evolution of Inequity in Access to Health Care in Spain: 1987-2001*
Pilar García Gómez y Ángel López Nicolás

DT 11/07 *The Economics of Credit Cards, Debit Cards and ATMs: A Survey and Some New Evidence*
Santiago Carbó-Valverde, Nadia Massoud, Francisco Rodríguez-Fernández, Anthony Saunders y Barry Scholnick

DT 12/07 *El impacto comercial de la integración europea, 1950-2000*
Luis Fernando Lanaspa Santolaria, Antonio Montañés Bernal, Marcos Sanso Frago y Fernando Sanz Gracia

DT 13/07 *Proyecciones de demanda de educación en España*
Andrés M. Alonso Fernández, Daniel Peña Sánchez de Rivera y Julio Rodríguez Puerta

DT 14/07 *Aversion to Inequality and Segregating Equilibria*
Antonio Cabrales Goitia y Antoni Calvó-Armengol

DT 15/07 *Corporate Downsizing to Rebuild Team Spirit*
Antonio Cabrales Goitia y Antoni Calvó-Armengol

DT 16/07 *Maternidad sin matrimonio: Nueva vía de formación de familias en España*
Teresa Castro Martín

DT 17/07 *Immigrant Mothers, Spanish Babies: Childbearing Patterns of Foreign Women in Spain*
Marta Roig Vila y Teresa Castro Martín

DT 18/07 *Los procesos de convergencia financiera en Europa y su relación con el ciclo económico*
José Luis Cendejas Bueno, Juan Luis del Hoyo Bernat, Jesús Guillermo Llorente Álvarez, Manuel Monjas Barroso y Carlos Rivero Rodríguez

DT 19/07 *On Capturing Rent from a Non-Renewable Resource International Monopoly: A Dynamic Game Approach*
Santiago J. Rubio Jorge

DT 20/07 *Simulación de políticas impositivas medioambientales: Un modelo de equilibrio general de la economía española*
Antonio Manresa Sánchez y Ferran Sancho Pifarré

DT 21/07 *Causas del crecimiento económico en Argentina (1990-2004): Otro caso de «tiranía de los números»*
Ariel Alberto Coremberg

DT 22/07 *Regional Financial Development and Bank Competition: Effects on Economic Growth*
Juan Fernández de Guevara Radoselovics y Joaquín Maudos Villarroya

DT 23/07 *Política fiscal e instituciones presupuestarias en los países de la reciente ampliación de la Unión Europea*
Carlos Mulas-Granados, Jorge Onrubia Fernández y Javier Salinas Jiménez

DT 24/07 *Measuring International Economic Integration: Theory and Evidence of Globalization*
Iván Arribas Fernández, Francisco Pérez García y Emili Tortosa-Ausina

DT 25/07 *Wage Inequality among Higher Education Graduates: Evidence from Europe*
José García Montalvo

DT 26/07 *Governance of the Knowledge-Intensive Firm*
Vicente Salas Fumás

DT 27/07 *Profit, Productivity and Distribution: Differences Across Organizational Form*
Emili Grifell-Tatjé y C. A. Knox Lovell

DT 28/07 *Identifying Human Capital Externalities: Theory with Applications*
Antonio Ciccone y Giovanni Peri

DT 01/08 *A Multiplicative Human Development Index*
Carmen Herrero Blanco, Ricardo Martínez Rico y Antonio Villar Notario

DT 02/08 *Real Exchange Rate Appreciation in Central and Eastern European Countries: Why the Balassa-Samuelson Effect Does Not Explain the Whole Story*
José García Solanes

DT 03/08 *Can International Environmental Cooperation Be Bought?*
Cristina Fuentes Albero y Santiago J. Rubio Jorge

DT 04/08 *On the Dynamics of Globalization*
Iván Arribas Fernández, Francisco Pérez García y Emili Tortosa-Ausina

DT 05/08 *Los motores de la aglomeración en España: Geografía versus historia*
Francisco J. Goerlich Gisbert y Matilde Mas Ivars

DT 06/08 *Sobre el tamaño de las ciudades en España: Dos reflexiones y una regularidad empírica*
Francisco J. Goerlich Gisbert y Matilde Mas Ivars

DT 07/08 *Managing Waiting Lists in a Fair Way*
Carmen Herrero

DT 08/08 *Estimación del capital social en España: Series temporales por territorios*
Francisco Pérez García, Lorenzo Serrano Martínez
y Juan Fernández de Guevara Radoselovics

DT 09/08 *Estimation of Social Capital in the World: Time Series by Country*
Francisco Pérez García, Lorenzo Serrano Martínez
y Juan Fernández de Guevara Radoselovics

DT 10/08 *Capabilities and Opportunities in Health*
Carmen Herrero y José Luis Pinto Prades

DT 11/08 *Cultural Transmission and the Evolution of Trust and Reciprocity in the Labor Market*
Gonzalo Olcina Vauteren y Vicente Calabuig Alcántara

DT 12/08 *Vertical and Horizontal Separation in the European Railway Sector: Effects on Productivity*
Pedro Cantos Sánchez, José Manuel Pastor Monsálvez y Lorenzo Serrano Martínez

DT 01/09 *How Effective Are Rewards Programs in Promoting Payment Card Usage?: Empirical Evidence*
Santiago Carbó-Valverde y José Manuel Liñares-Zegarra

DT 02/09 *Comparative Advantage Across Goods and Product Quality*
Francisco Alcalá

DT 03/09 *El transporte ferroviario de alta velocidad: Una visión económica*
Javier Campos Méndez, Ginés de Rus Mendoza e Ignacio Barrón de Angoiti

DT 04/09 *¿En qué circunstancias está justificado invertir en líneas de alta velocidad ferroviaria?*
Ginés de Rus Mendoza y Chris Nash

DT 05/09 *Openness and Geographic Neutrality: How Do They Contribute to International Banking Integration?*
Iván Arribas Fernández, Francisco Pérez García y Emili Tortosa-Ausina

DT 06/09 *Firms' Main Market, Human Capital and Wages*
Francisco Alcalá y Pedro J. Hernández

DT 07/09 *Tax Evasion, Technology Shocks and the Cyclicality of Government Revenues*
Judith Panadés Martí

DT 08/09 *Parameterizing Expectations for Incomplete Markets Economies*
Francesc Obiols-Homs

DT 09/09 *Endogenous Financial Intermediation*
Radim Boháček y Hugo Rodríguez Mendizábal

DT 10/09 *The Economic Impact of Migration: Productivity Analysis for Spain and the United Kingdom*
Mari Kangasniemi, Matilde Mas Ivars, Catherine Robinson y Lorenzo Serrano Martínez

DT 11/09 *Time, Quality and Growth*
Francisco Alcalá

DT 12/09 *Competing Technologies for Payments: Automated Teller Machines (ATMs), Point of Sale (POS) Terminals and the Demand for Currency*
Santiago Carbó-Valverde y Francisco Rodríguez-Fernández

DT 13/09 *Education, Utilitarianism and Equality of Opportunity*
Aitor Calo-Blanco y Antonio Villar Notario

DT 14/09 *European Integration and Inequality among Countries: A Lifecycle Income Analysis*
José Manuel Pastor Monsálvez y Lorenzo Serrano Martínez

DT 15/09 *The Determinants of International Financial Integration Revisited: The Role of Networks and Geographic Neutrality*
Iván Arribas Fernández, Francisco Pérez García y Emili Tortosa-Ausina

DT 16/09 *The Distance Puzzle Revisited: A New Interpretation Based on Geographic Neutrality*
Iván Arribas Fernández, Francisco Pérez García y Emili Tortosa-Ausina

Documentos de Trabajo

1

Fundación BBVA

Plaza de San Nicolás, 4
48005 Bilbao
España
Tel.: +34 94 487 52 52
Fax: +34 94 424 46 21

Paseo de Recoletos, 10
28001 Madrid
España
Tel.: +34 91 374 54 00
Fax: +34 91 374 85 22

publicaciones@fbbva.es
www.fbbva.es