An Optimal Shape of Income Tax: Evidence from Zero Income Tax Countries- Paraguay & Uruguay•

Preliminary and incomplete
Comments are welcomed

Arie Beresteanu*
Momi Dahan**
This version: March 2002

Abstract

Most of the existing literature on the optimal shape of income tax has a common result - decreasing marginal tax rates. This result stands in sharp contrast with real world income tax systems that are characterized by increasing marginal tax rates. Diamond (1998) made explicit the factors that affect the optimal shape of income tax rates. A special attention was given to one of the effects: the distribution effect. The main goal of this paper is to empirically explore whether the 'distribution effect' implies rising or declining marginal income tax rates with special interest at high levels of income. We estimate the hourly wage distribution as a proxy for the distribution of skills.

We show that the desired income tax schedule implied by the 'distribution effect' should exhibit increasing marginal tax rates at high levels of income. The analysis is based on data from zero income tax countries – Paraguay and Uruguay. We use a nonparametric estimation technique to avoid using any functional form assumptions on the skill distribution.

Keywords: Optimal income tax; Income distribution; Hazard rate
JEL Classifications: H21; C31

• We wish to thank S. Duryea, R. Funaeres and M. Szekely from the Inter American Development Bank for providing us with the data and helping us to interpret it.
* Department of Economics, Duke University, Box 90097, Durham, NC 27705, USA. E-mail: arie@econ.duke.edu .
** School of Public Policy, The Hebrew University, Jerusalem, Israel. E-mail: momidahan@mscc.huji.ac.il .

1. Introduction

Most of the existing literature on the optimal shape of income tax has a common result - decreasing marginal tax rates.[1] This result stands in sharp contrast with real world income tax systems that are characterized by increasing marginal tax rates. However, since the 1980s many countries, including many Latin American countries, have decided to make income tax systems more flat (less progressive) by reducing the numbers of brackets and tax rates. In other words, the world now is in less contrast with public finance literature.

Diamond (1998) reopened the debate on the optimal shape of income tax by showing an example of optimal increasing marginal tax rates schedule. His paper makes explicit the factors that affect the optimal shape of income tax rates. A special attention is given to one of the effects: the distribution effect. Diamond shows that the degree of progressiveness or regressiveness of the income tax schedule depends on the income distribution.[2] The interesting part in his result is that the distribution of income influences the shape of the income tax schedule not only through income inequality aversion but also through efficiency considerations.

The intuition behind the distribution effect is as follows. On one hand, higher marginal tax rate at a particular income level distorts the decision of individuals at that specific income level causing them to behave in a sub-optimal way. On the other hand, higher marginal tax rate acts as a lump-sum tax on individuals at all higher income levels. Therefore, at high levels of income the decision at the margin is not affected by marginal tax rates in previous brackets. This is purely an efficiency consideration and has clear policy implication.

The main goal of this paper is to empirically explore whether the 'distribution effect' implies rising or declining marginal income tax rates with special interest at high levels of income. We estimate the hourly wage distribution (both before and after controlling for education and experience), as a proxy for the distribution of skills. This allows us to investigate properties of the 'distribution effect'. The exact form of the distribution is important for the degree of progressiveness of the optimal income taxes. We estimate the distribution effect rather than assume its shape and we employ a non-parametric estimation technique in order to avoid assuming a specific functional form.

The distribution of wages before tax was used in previous literature as a proxy for skills A more recent example is Saez, 2001 who also uses the actual distribution in the U.S to estimate the distribution effect on optimal tax rates.[3] However, two factors may introduce bias in estimating the distribution of skills in the presence of, possibly non-linear,

---

[1] See for example simulations done by Mirrlees (1971), Atkinson (1973), Tuomala (1984) and Slemrod *et al* (1994).

[2] More specifically, if w is the income, the inverse hazard rate being f(w)/(1-F(w)) affects the optimal income tax schedule. This will be made clear in Section 2.

[3] Saez (2001) is aware of the bias the actual distribution of wages may has and therefore he chooses to use instead a parametric distribution (Pareto) to derive the optimal income tax rates.

individual income taxes, as it is still common in most countries. First, income taxes affect the incentives to use labor and its mix and as a result the distribution of before-tax-wages is not a good proxy for skills. Second, people may underreport their true income and this bias can vary as a fraction of their income.[4]

Paraguay and Uruguay are ideal cases for this purpose because these countries do not impose income taxes on individuals. To the best of our knowledge, it is the first time that tests of the empirical distribution of wages as a proxy for skills are based on data from zero income tax countries. Using data from zero income tax countries produces cleaner estimates of the distribution of skills.

Paraguay and Uruguay are interesting cases also because they are characterized by distributions with long tails. These two countries have a typical Latin American income inequality level. The Gini coefficient for the distribution of wages is around 0.58, 20 percentage points more than the typical OECD country. The results of this paper can serve as an indication for the implied optimal shape of income tax for countries with high levels of inequality.

The reminder of the paper is as follows. Section 2 sketches the theoretical background on the optimal shape of marginal income tax rates. Section 3 presents a short description of the tax system in Paraguay and Uruguay. Section 4 describes the data. In section 5 we describe the estimates for Paraguay and Uruguay and section 6 concludes.

2. Theoretical Background

The main ingredients of the optimal income tax problem are a) The individual utility function; b) The budget constraints; c) The self-selection constraint; d) The social welfare function; e) The distribution of skills and f) non-observability of skills.
The problem of optimizing a non-linear marginal income tax system consists of maximizing a social utility function taking into account the first order condition at the individual level, the self-selection constraint and the budget constraints both at the individual and economy-wide levels. Solving this problem yields the following expression for the optimum income tax:[5]

$$(2.1) \qquad \frac{\tau}{1-\tau} = \left[\frac{(1+\frac{1}{\varepsilon})}{w}\right]\left[\mathbf{U_C}\right]\left[\frac{\int_{w}^{w_H}[\frac{\gamma}{U_C} - G'(u)]\,\mathbf{fdw}}{\gamma(1-F)}\right]\left[\frac{(1-F)}{f}\right], \qquad \mathbf{f} \equiv \frac{\mathbf{dF}}{\mathbf{dw}}$$

[4] Paraguay and Uruguay use taxes other than individual income tax and transfers that may affect the labor supply and wage distribution. See discussion in Section 3.
[5] A full description of the problem and solution of the optimal income tax may be found in Dahan and Strawczynski (2000).

where τ is the marginal income tax rate, ε is the compensated elasticity of labor supply, w is the level of skill (productivity), Uc is the derivative of individual utility with respect to consumption, G'(u) is the derivative of the social welfare function with respect to individual utility and γ is the shadow price of the government budget constraint.[6] F and f are the distribution and density functions of skills, respectively.

The first term in the right hand side of (2.1) represents the standard efficiency effect. The second term represents income effects. The third term is the inequality aversion effect. If we assume decreasing social marginal utility (G''<0), then as w increases, these three elements will increase and therefore imply higher optimal tax rate.

The last term is the distribution effect. This term is the ratio of the number of individuals above a particular income level, 1-F, to the number of individuals with income at that level, f. Note that a higher marginal tax rate at a low-income level distorts the decision of individuals at this income level. But, on the other hand this new higher marginal tax acts as a lump-sum tax on individuals at higher income levels; since at high levels of income the decision at the margin is not affected by marginal tax rates in previous brackets. The higher (1-F), the higher the quantity of individuals that are paying higher lump sum taxes, and consequently the higher is the optimum marginal tax. On the other hand, the higher is f the lower is the optimal tax at this income level, since more individuals are affected by the distortion.

The intuition becomes even clearer by looking at the uniform distribution case: since f is equal for all income levels, the marginal tax rate declines over all the range - reflecting the fact that as we advance up the income distribution, marginal taxes act as lump-sum for fewer individuals.

Diamond (1998) shows that for individual utility functions with zero income effect (i.e., linear utility of consumption) and constant compensated elasticity of labor supply, the distribution of skills dictates the optimal shape of income tax. In particular, a Pareto distribution implies increasing marginal tax rates for high-income earners. Dahan and Strawczynski (2000) show that this result does not hold for lognormal distribution. In the case of a lognormal distribution of skills, one needs to assume more restrictive utility functions to get increasing marginal tax rates on high-income earners. To get this result, one has to assume, in addition to zero income effect, a logarithmic utility of leisure, as was used by Mirrlees (1971). These assumptions mean that the inverse of the compensated elasticity of labor supply is a linear function of skills. The labor supply of high-skilled individuals is less sensitive than that of low-skilled individuals. However, if we are concerned with high-income individuals it is enough to have this type of behavior among high-skilled individuals only.

---

[6] ε is the compensated elasticity of labor supply only if utility is linear in consumption. In general, the relationship between ε and the compensated elasticity in the additive case, ε^c is:

$$\frac{1}{\varepsilon^c} = \frac{(V_{LL} + w^2 U_{CC})L}{V_L} = \frac{1}{\varepsilon} - \frac{wLU_{CC}}{U_C}$$

Most of the attention in the literature has been devoted to estimation of labor supply elasticities and simulations of various forms of social welfare functions. The distribution effect has captured much less attention. According to the model shown above, clearly the distribution of income is an important factor: since different distributions imply a different ratio for (1-F)/f, it is crucial for the optimum shape of taxes to learn about the distribution of income as a proxy of skills distribution.

The main goal of this paper is to investigate the shape of the optimal income tax implied by the empirical wage distribution, and in particular to explore the properties of the 'distribution effect'. The level of income inequality matters not only because of inequality aversion consideration but also on efficiency grounds. More unequal income distribution has important implications even if we assume out inequality aversion considerations.

3. A short description of the tax systems in Paraguay and Uruguay

Paraguay and Uruguay are chosen for estimating the 'distribution effect' because an income tax is not imposed on individuals in both countries. However, the tax on labor income is far from zero if we take into account the VAT and payroll taxes (social security).

The value-added tax rate is 10 percent in Paraguay. Value-added tax applies to sale of goods and services including imported goods (exports are exempt). In addition, all employees and employers must pay taxes on payrolls. The employer contribution rate is 14% and 2.5 % and employee withholdings are 9% and 0.5% for Instituto de Prevision Social and Banco Nacional de Trabajadores, respectively. Combining these gives us an effective tax rate on labor income of approximately 25 percent.[7]

Likewise, the standard value-added tax rate in Uruguay is 23 percent. The VAT has a very broad scope, as it levied at all stages of production and trading, including retailing, as well as on a wide range of services and is also levied on imports. A reduced rate applies to necessities, such as basic food products, medicines, and interests on loan granted to individuals. There are some exemptions such as farm products, real estate, credits, agricultural machinery, cigarettes and milk. Exports are not subjected to the VAT. The employer contributions for the main social security taxes are 20.5 percent and employee withholdings basic rate is 17.25 (it may increase to 22.25 percents depending on the level of salary). Therefore the labor income tax rate is approximately 36 percent.

The detailed description is intended to make clear that labor income can be substantially different from zero even if the formal income tax rate is zero and that both VAT and payroll taxes are not flat. However, as a first approximation one can consider

[7] The labor tax rate is calculated according to $(\tau_c+\tau)/1+\tau_c$ where τ is payroll tax rate and τ_c is VAT rate. Note that payroll tax rate is 14.1[=(16.5/116.5)+9.5]. The source of tax rates is Coopers and Lybrand (1996).

these taxes as flat rate taxes on labor income in the sense that they are weakly related to income.

Yet, the distribution of wages may be affected by the presence of transfer payments that are financed by these taxes. They may affect the labor supply through the usual income effect and as a result the distribution of wages is not completely clean from government intervention. Nonetheless, the level of transfers as a percentage of GDP is relatively low in Paraguay and (less so) Uruguay.

The relatively low level of taxation, its composition and the low level of transfers are partially explained by t he level of development (income per capita). It is harder to collect taxes, especially income tax, in developing countries because of the flexibility of firms and workers to shift activities to the informal sector. The high level of inequality is also associated (to some degree) with the level of development.

4. The Data

We use data from Paraguay and Uruguay for the year 1995. The sample for Paraguay contains 7663 wage earners; workers (both males and females) that have a non-zero wage in their principal job. The average hourly wage in the principal job is 3,810 (Guaranis) in 1995 where the median is 2,046 (see Table 1). The Gini coefficient is typical for Latin American countries - 0.58. The sample for Uruguay contains 24,463 workers that have a non-zero wage in their principal job. The average hourly wage in the principal job is 20,236 (Uruguayan pesos) in 1995 and the median wage is 13,837. The Gini coefficient for Uruguay is smaller 0.46. Over all, the Gini coefficients in Paraguay and Uruguay are high compared to those of OECD countries. Moreover, we can see that the ratio between the 99% and the 95% quantiles is 2.63 in Paraguay and 2.12 in Uruguay. This indicates that the distribution of wages has a very long and relatively thick tail.

5. Empirical Results

The distribution of hourly wages in the principal job is our first proxy for the distribution of skills both in Paraguay and Uruguay. Figures 1 and 2 describe the hazard rate of wages for Paraguay and Uruguay respectively. The horizontal axis describes levels of hourly wages. Both schedules are decreasing for most of the region. The downward sloping part starts from the 10% quantile of wages in Paraguay but only above median for Uruguay.

The hourly wage of a worker contains both innate and acquired skills. In order to ‘clean’ the skill proxy from acquired skills it is common to use residuals from wage equations as a better proxy for innate skill. We assume that the hourly wage w depends on a vector of acquired (and observed) skills x and on a vector of innate skills ε (unobserved). We also assume these two are separable in the hourly wage equation. We follow the literature on wage equations and assume that the innate and acquired skills are multiplicatively separable:

(5.1) $$w = h(x)g(\varepsilon)$$

In addition we assume that x and ε are statistically independent. Taking logs in equation (4.7) yields

(5.2) $$\log(w) = \tilde{h}(x) + \tilde{g}(\varepsilon)$$

where $\tilde{h}$ and $\tilde{g}$ are the logs of h and g respectively. Therefore, if we assume that x and ε are independent, the residuals from the log-wage equation give

(5.3) $$lres := \log(w) - E(\log(w) \mid x) = \tilde{g}(\varepsilon) - \tilde{\mu}$$

and

(5.4) $$res := \exp(lres) = g(\varepsilon)/\mu$$

where $\tilde{\mu}$ is the mean of $\tilde{g}(\varepsilon)$ over the population. The explanatory variables x (acquired skills) are education and experience (age) for each individual. The regression in (5.4) is estimated non-parametrically and the hazard rate is calculated from the residuals in (5.4). Education here is treated as if it is uncorrelated with innate skills. Education involves both innate skills and a decision to invest in human capital that is not necessarily connected to innate skills. We assume here that education represents only the non-innate skills. Figures 3 and 4 describe the same estimation made for the residuals from log-wage regression. The results are very similar to those with wage levels. The hazard rate is downward sloping in most of the range.

The hazard rate is estimated using a non-parametric approach. However, it is common in the literature to start with a parametric assumption on the distribution of skills (wages) and to estimate the parameters of that distribution from the data. Two of the most commonly used distribution families are lognormal and Pareto distributions. The parameters of the distribution are estimated from the sample using the method of moments or maximum likelihood. The parametric distribution can than be tested (as the null hypothesis) against the empirical distribution.

The Kolmogorov-Smirnov test is almost bounded to reject the null hypothesis given the size of the sample (several thousands observations). In particular Kolmogorov-Smirnov test is not very powerful test against differences in tails of distributions. Tests done on the wages as well as the residuals rejected both distributions from being equal to the empirical distribution with very high significance level[8].

In order to get an idea how well the lognormal and the Pareto distributions fit the data we use the confidence interval constructed for the hazard rates. Once the parameters

[8] Other tests like the Cramer Von-Mises test are available. We plan to elaborate on more parametric families and goodness of fit tests in the next version of this paper.

of both distributions were calculated the hazard rate can be computed. Figures 5 and 6 shows the theoretical hazard rates of the lognormal and Pareto distributions. Both distributions do not fit into the confidence interval in all the range. However it looks like the Pareto distribution is doing a better job than the log normal especially for Uruguay.

6. Conclusions

This paper shows that the desired income tax schedule implied by the ‘distribution effect', which reflects efficiency considerations only, should exhibit increasing marginal tax rates at high levels of income. The analysis is based on data from zero income tax countries – Paraguay and Uruguay – where the hourly wage rates in the principal job is a proxy for the ability of workers. A second proxy for ability is the residuals from a log-wage regression. By using residuals we hoped to clean our skill measure from acquired skills and to be left with only innate skills. In this case we received a decreasing hazard rate (increasing distribution effect).

The estimation method we use is free of functional assumptions. In that sense we feel that our results are strong. The income distribution of South American countries has a very long tail and a careful treatment of the estimation is needed here.

Appendix A: Kernel smoothing of hazard rate estimators

Section 2 describes the role of the 'distribution effect' $\frac{1-F}{f}$ of skills in determining the optimal tax schedule (see equation 2.1). In this appendix we describe the method by which we estimate this function and the properties of this estimator.

The function $\frac{1-F(w)}{f(w)}$ is actually one over the hazard rate $h(x)=\frac{f(x)}{1-F(x)}$. The estimation of the last has received much attention in the statistical and econometric literature. Two pioneering papers by Watson and Leadbetter (1964a, 1964b) introduce two ways to nonparametrically estimate the hazard function. They show that both methods are asymptotically equivalent and thus we employ the one that is computationally less demanding.

The estimator is based on the sample analogue of the hazard rate. Let $\{w_{[i]}\}_{i=1}^{N}$ be the order statistic of the sample. The sample analogue of *f* is *1/n* in each sample point and the sample analogue of *1-F* is (*n-i)/n* in $w_{[i]}$. The ratio is *1/(n-i)*. Convoluting the sample analogue with a kernel density function K(·) yields[9]

(A.1) $$\hat{h}(w)=\sum\frac{1}{n-i+1}K_b(w-w_{[i]}).$$

To overcome problems of low density at the tails we employ an adaptive bandwidth. Bandwidth selection is still an unsettled issue. We rely on a cross validation procedure to choose the bandwidth we employ. We modify the procedure from Müller and Wang (1994) in the following way:

Step 1: Choose a pilot bandwidth *b* and a grid of points $\{\xi_j\}_{j=1}^{m}$ on the support of the sample. Estimate the density *f* using a global bandwidth *b* on the grid $\{\xi_j\}_{j=1}^{m}$ to obtain $\hat{f}_0$. Set $b_{0,j}=C\cdot\hat{f}_0(\xi_j)$ for $j=1..m$ as the initial adaptive bandwidth (*C* is a scalar).

Step 2: Choose a grid of perturbation constants $\{c_i\}_{i=1}^{k}$ and use the bandwidth $\tilde{b}_{i,j}=c_i\cdot b_{0,j}$ to calculate $m_{i,j}=MSE(\xi_j,\tilde{b}_{i,j})$ for $i=1..k$, $j=1..m$.

Step 3: For each *j=1..m* (i.e. for each grid point) choose $b_{1,j}=\arg\min_{\tilde{b}_{i,j}}MSE(\xi_j,\tilde{b}_{i,j})$. That means for each grid point ξ_j chose the best bandwidth from $\{c_i\cdot b_{0,j}\}_{i=1..k}$.

Step 4: Smooth $\{b_{1,j}\}_{j=1..m}$ over the grid $\{\xi_j\}_{j=1}^{m}$. Use the smoothed bandwidth as the adaptive bandwidth to calculate $\hat{h}^{(2)}(x)$ over the estimation grid using (4.4).

---

[9] The sample analogue was changed to *1/(n-i+1)* to avoid division by zero for *i=n*.

This procedure as well as the non-parametric estimation procedure is computationally intensive. To reduce the run time we iterate step 1, 8 times and steps 2-4 additional 8 times. The set of perturbation scalars $\{c_i\}_{i=1}^{k}$ is taken to be (0.85 0.9 0.95 0.99 1.01 1.05 1.10 1.15) allowing the final bandwidth to be between 3 times and 1/3 of the initial bandwidth. The initial scalar, C, was chosen such that on average $C \cdot \hat{f}_0(\xi_j)$ will be equal to Silverman's rule of thumb bandwidth selector.

The method described above yields an estimate of the function $\frac{f}{1-F}$. 95% confidence interval is computed by a bootstrap method in the following way:

Step 1: Choose the number of bootstrap iterations B.
Step 2: Take B random samples with replacement from the data set (bootstrap samples).
Step 3: Calculate $\hat{h}^{(2)}(x)$ using (4.4) for each bootstrap sample.
Step 4: For each estimation point take the 0.025B and 0.975B quantiles of $\hat{h}^{(2)}(x)$.

We choose *B=500*.

REFERENCES

Aitchison, J. and J.A.C. Brown. 1957. *The Lognormal Distribution*, Cambridge University Press, London

Atkinson A. (1973), “How Progressive Should Income Tax Be?” in: M. Parkin and A.R. Nobay, eds., *Essays in modern economics* (Logman, London).

Dahan M. and M. Strawczynski. (2000). “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates: Comment”, *American Economics Review*, 90(3), pp. 681-686.

Diamond P. 1998. “Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax Rates”, *American Economic Review*, March, pp. 83-95.

Emerson, J. D. and D. C. Hoaglin. 1983. “Stem-and-Leaf displays”, In *Understanding Robust and Exploratory Data Analysis*, ed. D.C. Hoaglin, F. Mosteller and J.W. Tukey, 7-30.

Hall P. (1992), *The bootstrap and Edgworth expansions*, Springer series in statistics.

Hardle W. (1990), *Applied Nonparametric Regression*, *Econometric Society Monographs 19*, Cambridge University Press.

Horowitz J.L., and C.F. Manski, (1998), "Censoring of Outcomes and Regressors due to Survey Nonresponse: Identification and Estimation Using Weights and Imputations", *Journal of Econometrics*, Vol. 84, pp. 37-58.

Lydall, H. F. 1968. *The Structure of Earnings*, Oxford University Press.

Mirrlees J. 1971. “An Exploration in the Theory of Optimal Income Taxation”, *Review of Economic Studies*, 38, 135-208.

Müller, H.G., and Wang J.L. (1990), “Locally adaptive hazard smoothing”, *Probability Theory and Related Fields* 85, 523-538.

--------------- (1994), "Hazard Rate Estimation Under Random Censoring with varying Kernels and bandwidths", *Biometrics*, 50, 61-76.

Parzen E. (1962), “On Estimation of a Probability Density Function”, *Annals of Mathematical Statistics*, 74, 105-131.

Saez, E. "Using Elasticities to Derive Optimal Income Tax Rates", Review of Economic Studies. Vol. 68(1), January 2001.

Silverman B. W. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

Slemrod Joel, Shlomo Yitzhaki, Joram Mayshar and Michael Lundholm (1994), “The optimal two-bracket linear income tax”, *Journal of Public Economics*, 53 (2): 269 -290.

Tuomala, Matti .1984. “On The Optimal Income Taxation - Some Further Numerical Results”, *Journal of Public Economics* 23, 351-366.

Watson G.S., and Leadbetter, M.R. (1964a), “Hazard Analysis I”, *Biometrica*, 51, 175-184.
--------------- (1964b), “Hazard analysis II”, *Sankhya*, A, 26, 101-116.

Weiss, Y. 1972. “The Risk Element in Occupational and Educational Choices”, *Journal of Political Economy*, 80, 1203-1213.

1996 International Tax Summaries. Coopers & Lybrand Global Tax Network, John Wiley & Sons, INC.

Table 1: Paraguay -Basic statistics

Moments:

Variable	Mean	Standard deviation	Skewness
Wage	3808.7	13903	46.98
Education	6.96	4.29	0.711
Experience (age)	36.14	14.56	0.583
residual	1.53	3.68	39.04

Quantiles:

Variable	min	1%	5%	10%	25%
Wage	13.84	107.67	323.67	536.67	1099.7
Education	0	0	1	2	4
Experience (age)	7	12	16	19	25
residual	0.013	0.0846	0.233	0.359	0.632

median	75%	90%	95%	99%	max
2046.5	3720.9	7155.6	11628	30565	954080
6	9	12	15	18	18
34	45	57	64	74	92
1.009	1.622	2.702	3.971	9.520	237.8

Gini coefficient for wages: 0.582
Gini coefficient for residuals: 0.487
Number of observations: 7,663.

Table 2: Uruguay -Basic statistics

Moments:

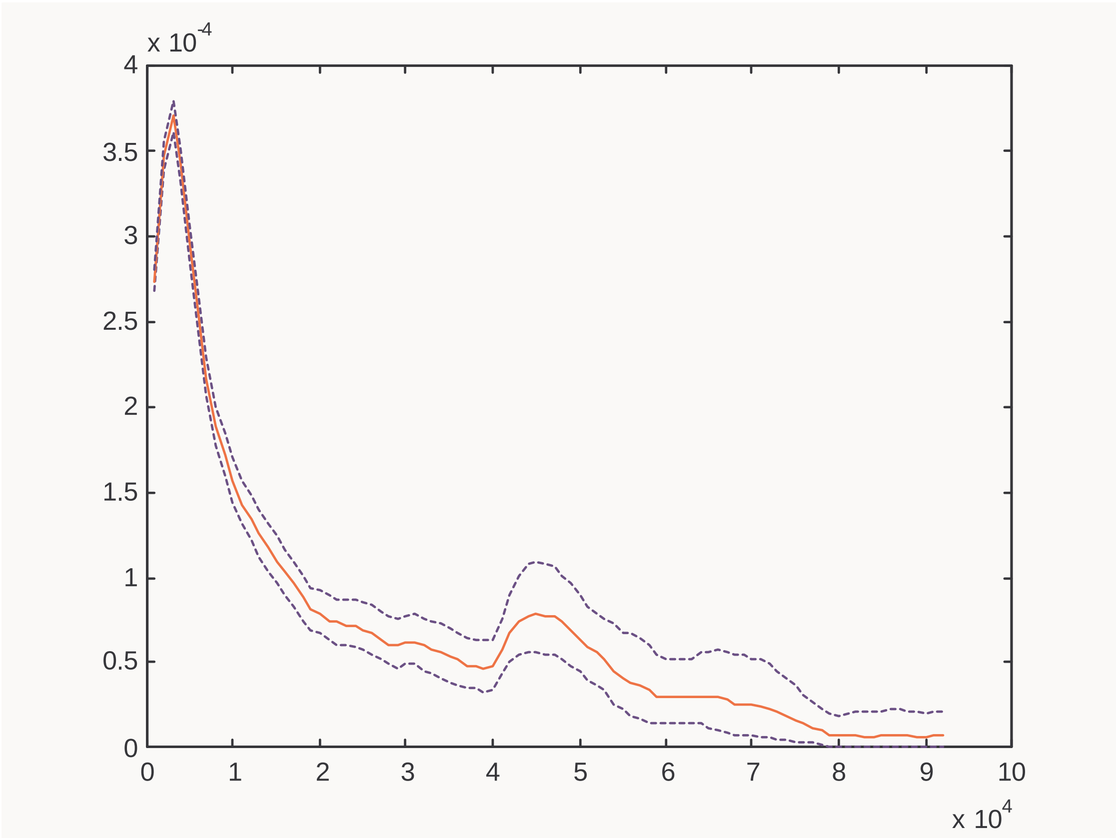
Variable	Mean	Standard deviation	Skewness
Wage	20236	27274	12.08
Education	9.42	4.01	0.185
Experience (age)	39.28	13.93	0.289
residual	1.296	1.398	17.16

Quantiles:

Variable	min	1%	5%	10%	25%	median	75%	90%	95%	99%	max
Wage	242.25	1744.2	3720.9	5160.5	8561.0	13837	23035	38760	56123	119100	1238400
Education	0	1	3	5	6	10	12	15	16	18	20
Experience (age)	14	16	19	21	28	38	50	58	63	70	94
residual	0.0238	0.157	0.317	0.432	0.667	1.012	1.530	2.290	3.018	6.003	84.65

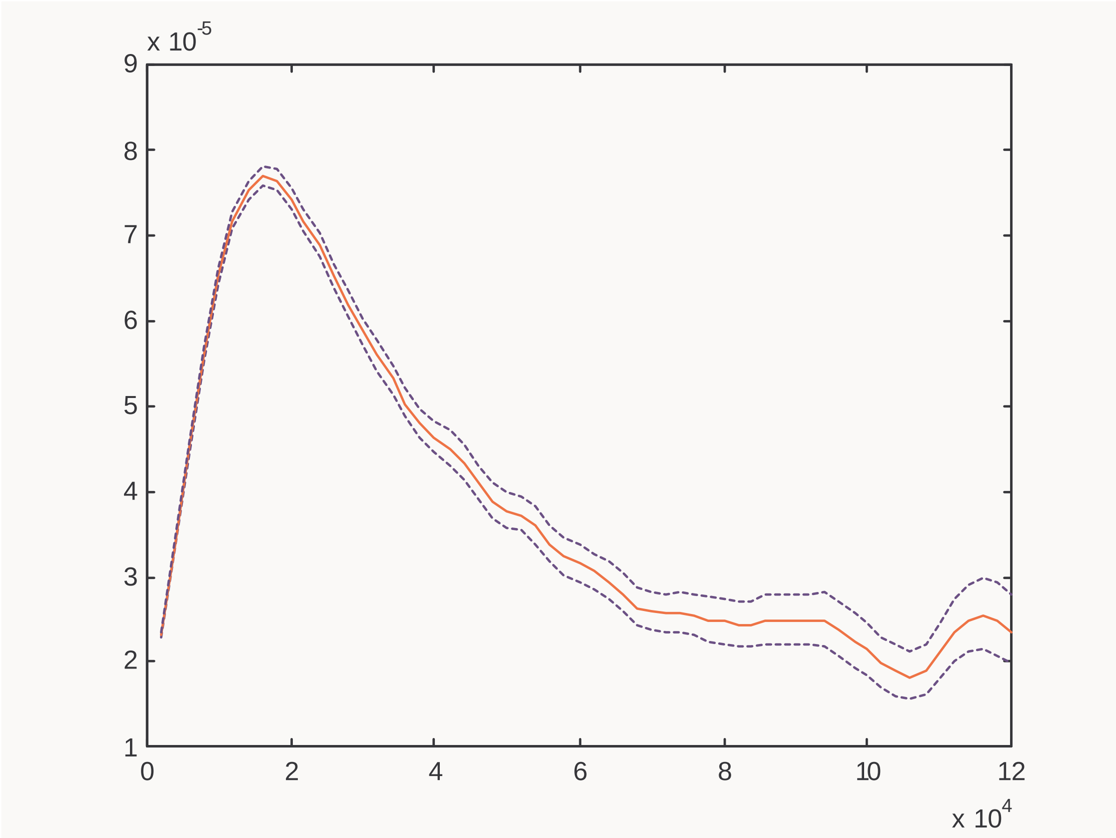
Gini coefficient for wages: 0.460
Gini coefficient for residuals: 0.382
Number of observations: 24,403.

Figure 1
Paraguay- Hazard rate of wages



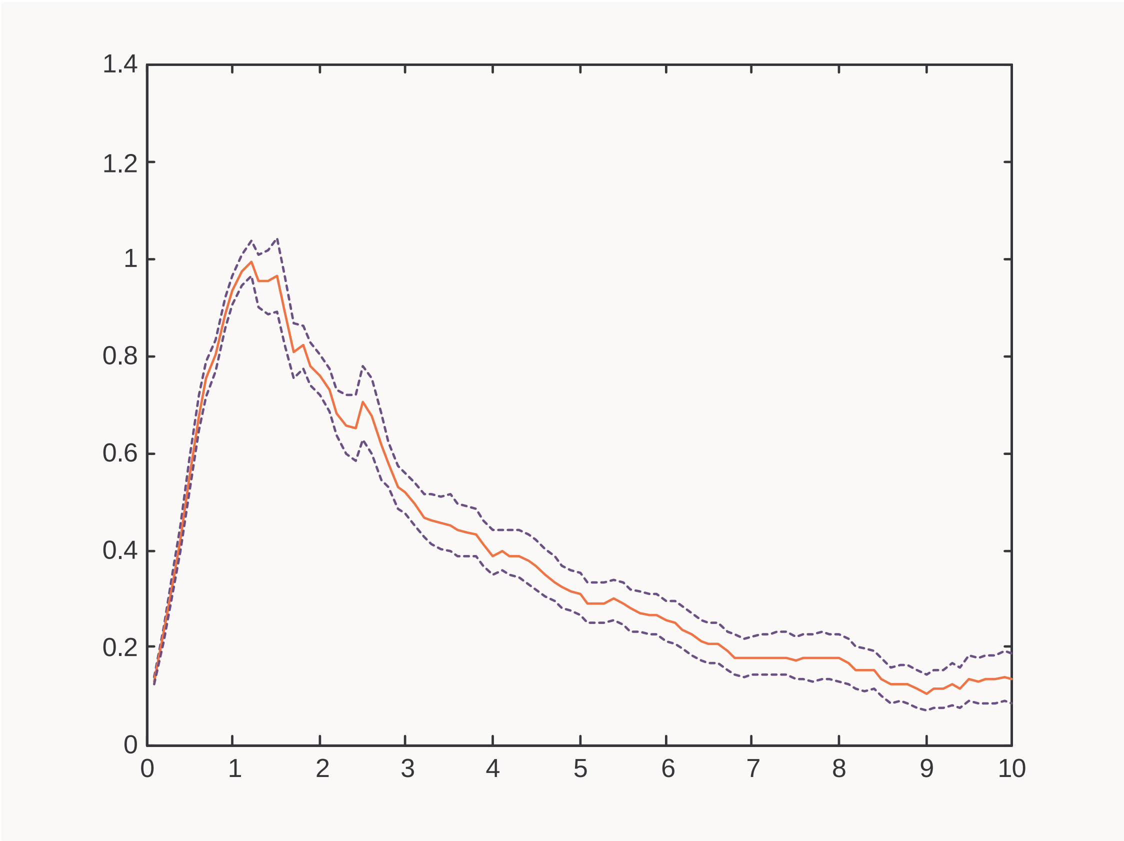
Confidence intervals were constructed using 500 bootstrap iterations.

Figure 2
Uruguay- Hazard rate of wages



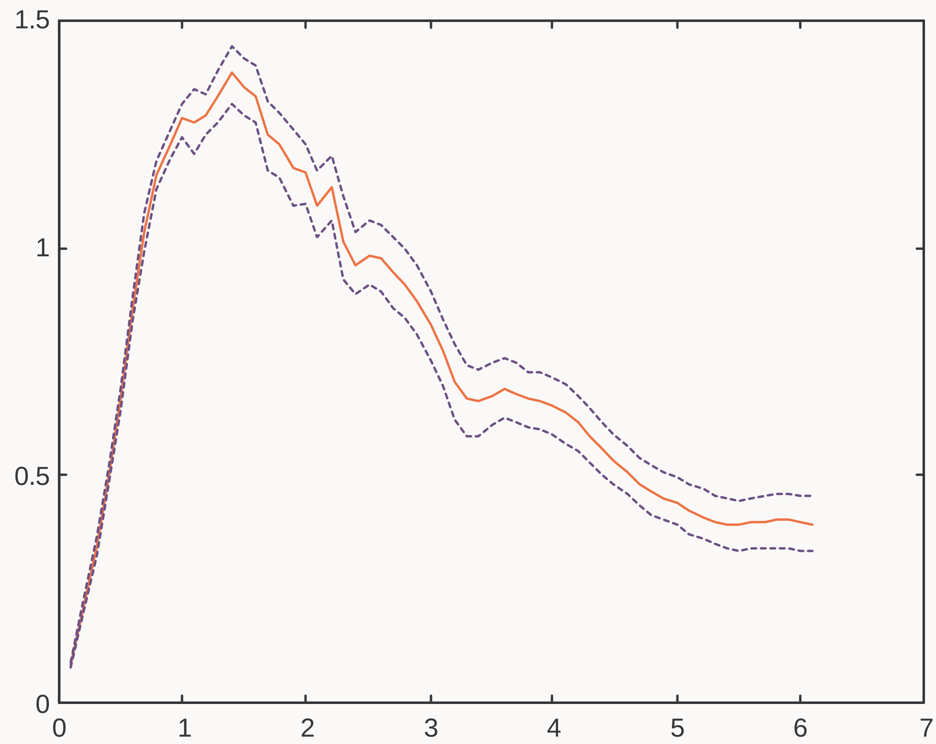
Confidence intervals were constructed using 500 bootstrap iterations.

Figure 3
Paraguay- Hazard rate function of implicit skill measured by the residuals from log-wage regression on education and experience



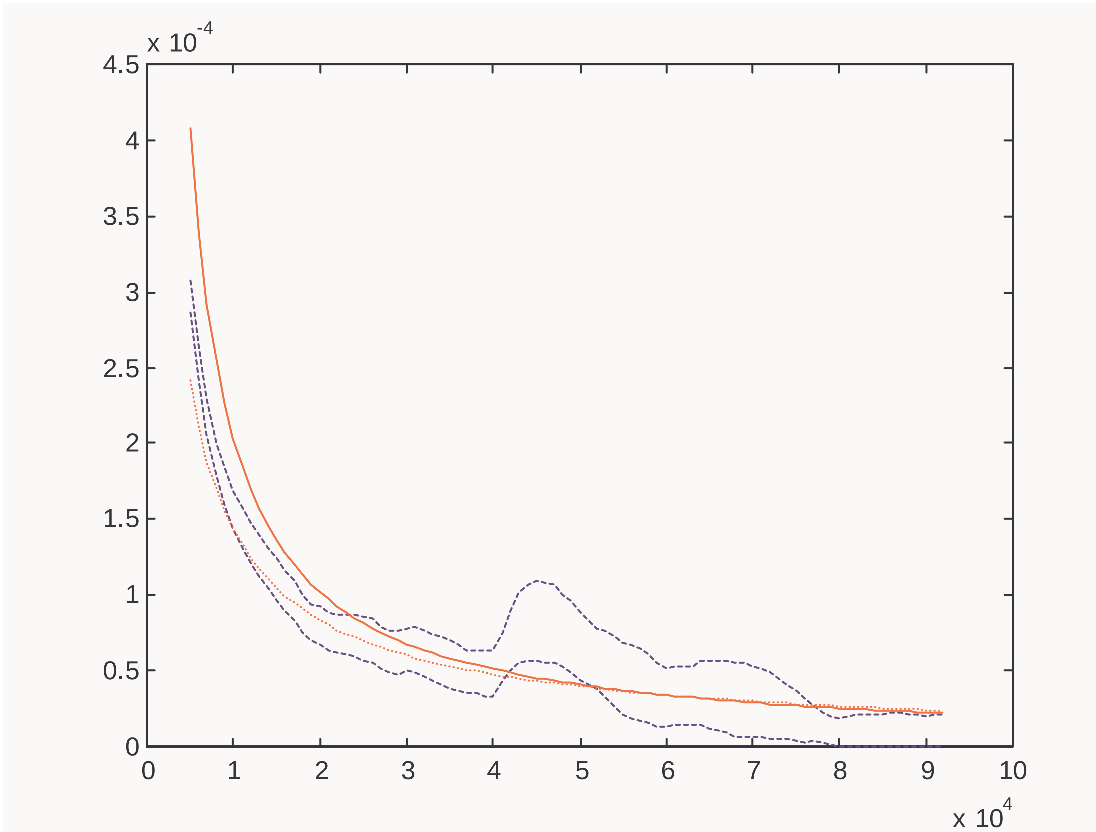
Confidence intervals were constructed using 500 bootstrap iterations.

Figure 4
Uruguay- Hazard rate function of implicit skill measured by the residuals from log-wage regression on education and experience



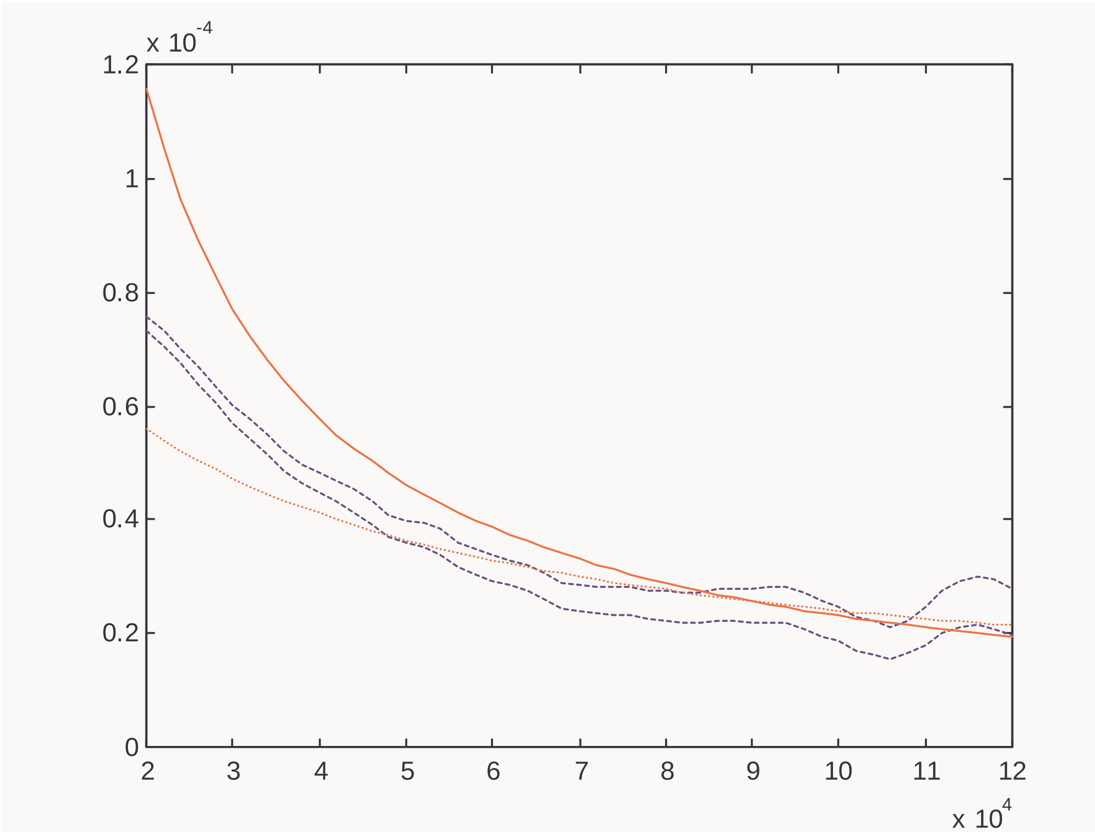
Confidence intervals were constructed using 500 bootstrap iterations.

Figure 5
Paraguay - Goodness of fit for Pareto and lognormal distributions
based on hazard rate for hourly wages



Dotted line - Lognormal
Full line - Pareto

Figure 6
Uruguay - Goodness of fit for Pareto and lognormal distributions
based on hazard rate for hourly wages



Dotted line - Lognormal
Full line - Pareto