Review of Behavioral Economics, 2020, 7: 1–17

Economic Arbitrage and the Econophysics of Income Inequality

Anwar Shaikh[1] and Juan Esteban Jacobo[2]

[1] *Department of Economics, New School for Social Research, New York, NY 10003, USA; shaikh@newschool.edu*

[2] *Departamento de Economía, Universidad Externado de Colombia, Bogotá, Colombia; juan.jacobo@uexternado.edu.co*

ABSTRACT

Yakovenko and his co-authors have established that the bottom 97–99% of individual incomes (labor incomes) follow a near-exponential distribution while the top incomes (property incomes) follow a power law. Initial explanations of these patterns relied on various monetary analogues to the physics principle of energy conservation. Subsequent approaches turned to the stochastic dynamics of economic processes, including those of labor and property income modeled as a drift-diffusion processes. Our paper is in the latter tradition, but our specifications of drift-diffusions are derived from the fundamental economic principle of turbulent arbitrage modeled as a mean-reverting process. This approach is well developed in the domain of interest rate arbitrage as in the case of CIR models. Our contribution is to demonstrate that arbitrage can also explain the observed distributions of wages, rates of return on assets, and property income. In the energy conservation approach, stationary distributions are derived from the assumption of entropy maximization. In both stochastic dynamics approaches, the dynamic paths give rise to stationary distributions that turn out to be entropy maximizing.

Keywords: Economics, arbitrage, econophysics, income distribution, entropy maximization, Fokker–Planck equations

JEL Codes: B12, C10, C18, D31, D33

ISSN 2326-6198; DOI 10.1561/105.00000129

1 Introduction[1]

The analysis of income distribution in capitalist societies goes back to Pareto's (1897) finding that top incomes follow a power law (Silva and Yakovenko, 2005; Yakovenko and Rosser, 2009). But modern evidence indicates that the lower bulk of the income distribution follows different patterns (Dragulescu and Yakovenko, 2001, p. 585).

One way to address these issues is to fit the whole range of incomes by means of single distributions such as the lognormal, Levy, Gamma, Champernowne and others (Dragulescu and Yakovenko, 2001, p. 585). A recent econophysics "two-class" theory of income distribution (EPTC) pioneered by Yakovenko and his co-authors has opened up new ground (Dragulescu and Yakovenko, 2000, 2001; Yakovenko and Rosser, 2009; Banerjee and Yakovenko, 2010). This approach provide substantial empirical evidence that the bottom 97–99% of the overall distribution of personal incomes, which is essentially labor income, is well approximated by an exponential distribution, while top 1–3%, which is essentially property income, is well approximated by a power law. Personal income may encompass both labor and property components, but it is plausible that labor income dominates at lower income levels and the property income at the highest ones. The EPTC approach develops an ingenious method of combining the two distributions to create a parsimonious yet powerful approximation to overall Lorenz curve (Silva and Yakovenko, 2005; Yakovenko and Rosser, 2009). Subsequent research showed that even when labor income data is partitioned by gender or race, the empirical distribution *within* each group is still approximately exponential – despite the fact that Males have persistently higher incomes than Females, and Whites persistently have higher incomes than Non-Whites (Shaikh *et al.*, 2014). In addition, labor incomes by occupation are also approximately exponential (Shaikh, 2016, p. 755). These results make sense if we recognize that the gender and race subsets are samples taken from the overall distribution of individual incomes, while average incomes by occupation are consolidations.

It is important to stress that labor and property incomes *only approximately* follow exponential and Pareto laws, respectively. At an empirical level, the pdf of labor income is better approximated by Gamma or Weibull distributions (Shaikh *et al.*, 2014), while that of property incomes can also be well approximated by the log-normal – as we shall see. Nonetheless, we agree with Yakovenko that the first approximations offered by exponential and Pareto distributions offer great advantages: "Universalities are not easy to uncover, but they form the backbone of regularities in the world around us... Universalities establish the first-order effect, and deviations represent

[1] We wish to thank the anonymous referee for insightful and extremely helpful comments, which we have tried to address here. In particular, the interpretation of the kinetic approach in Section 3 was most illuminating.

the second order effect ... (and) in the first approximation, the distributions are quite similar and universal" (Yakovenko, 2007, p. 2821).

For the portion of individual incomes (r) following a pure exponential distribution, the probability distribution is $P(r) \propto e^{-\left(\frac{r}{\langle r\rangle}\right)}$ where $\langle r\rangle$ represents the mean of the exponential portion. The cumulative probability distribution for incomes *above* r within this same segment is $C(r) = e^{-\left(\frac{r}{\langle r\rangle}\right)}$. Then $\ln C(r) = -\left(\frac{r}{\langle r\rangle}\right)$ is *parameter free* as long as lower incomes in any given year are normalized by their mean. The empirical signature of a normalized exponential distribution is therefore a straight-line plot in $\left(\ln\left(\frac{r}{\langle r\rangle}\right), C(r)\right)$ for all years, and the Gini coefficient (G) is $G = 1/2$ for individual incomes and $G = 3/8$ for two-person family incomes (Dragulescu and Yakovenko, 2001, p. 587). On the other hand, top incomes follow a power law in which the cumulative distribution from above is $C(r) \propto r^{-\alpha}$. The corresponding empirical signature of a power law is a straight-line plot in $(\ln r, \ln C(r))$. The fraction of total income (f) in the power law segment, the Gini $G = (1+f)/2$ and the power law coefficients α ranging from 1.4 to 1.8 all vary with the stock market – consistent with the hypothesis that top incomes are largely property income (Silva and Yakovenko, 2005, p. 307; Yakovenko, 2007, p. 2815).

We utilize US Internal Revenue Service (IRS) data on adjusted gross income of individuals, pre-binned in varying ranges.[2] The main chart in Figure 1 displays the overall distribution in 2011 with income r on the horizontal axis and $C(r)$ on the log-scaled vertical axis. The left inset shows that the log-linear plot of the bottom 97% exhibits an exponential signature, while the right inset displays the log-log plot of the top 3% with its power law signature. Table 1 will show that these patterns hold for every year (Silva and Yakovenko, 2005, p. 305). In what follows, the lower and upper sections are treated as separate populations, so that $C(r)$ begins from 1 in each.

2 Maximum Entropy Approaches

It is a fundamental principle of physics that with a given amount of total energy and a given number of particles interacting in a manner that conserves total energy, the stationary equilibrium distribution of energy corresponding to the maximum-entropy state is Boltzmann-Gibbs. The EPTC approach develops a monetary equivalent of this principle in order to explain labor incomes: a closed economy with a fixed amount of money and a fixed number of agents whose individual monetary transactions conserve total money. Positivity of individual money holding is ensured by assuming that in each interaction individual agents

[2]IRS data for adjusted gross income (AGI) comes in bins of varying size, so we use bin mid-points for income levels.

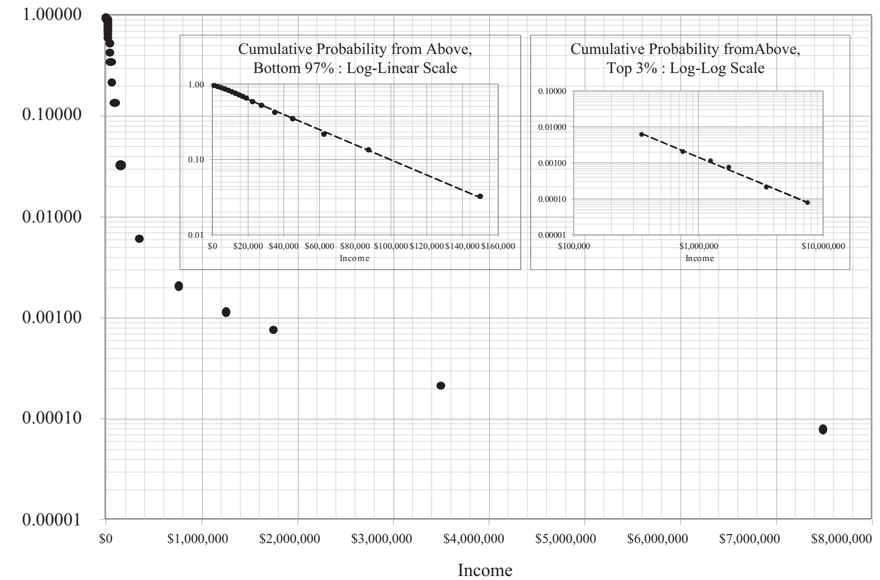

Figure 1: Cumulative Probability from Above, 2011 (Log-Linear Scale).

spends less money than they possess (Yakovenko, 2007, p. 2805). The resulting entropy maximizing stationary distribution is exponential (Dragulescu and Yakovenko, 2000, pp. 723–724). A proposed extension supposes that one agent out of a randomly selected pair is designated as an "agent-firm" (A) who borrows capital K from agent B and uses this to pay the same amount to workers L. The agent-firm then uses these workers to produce a physical product Q which it in turn sells to agents A, B and L at a unit price p such that total sales $pQ > L$. The sales proceeds pay off the loan K with interest, leaving some net profits for the agent-firm. "The net result is a many-body exchange of money that still satisfies the conservation law . . . Our computer simulations show that the stationary probability distribution of money in this model always has the universal Boltzmann-Gibbs form independent of the model parameters".[3] (Dragulescu and Yakovenko, 2000, p. 725). Another proposed extension treats debt as form of money so that assets and liabilities cancel out and "the conservation law is restored" (Yakovenko, 2007, p. 2807). At an empirical level, income flows are used as proxies for individual money and wealth stocks due to the lack of data on the latter, even though "there is no fundamental

[3]Readers will recognize this as a version of money flows in Marx's Schemes of Simple Reproduction. But there, the critical prior issue is the source of the surplus of output value pQ over input value L (Marx, 1967, Chs. XX–XXI).

Table 1

	G′	⟨r⟩	⟨w⟩	f	α		G′	⟨r⟩	⟨w⟩	f	α
2002	0.49	46.39	38.33	0.17	1.45	2010	0.49	56.61	44.36	0.22	1.39
2003	0.49	47.59	38.91	0.18	1.42	2011	0.49	57.61	45.31	0.21	1.40
2004	0.49	51.34	40.62	0.21	1.35	2012	0.49	62.79	46.88	0.25	1.33
2005	0.49	55.24	42.27	0.24	1.31	2013	0.49	61.71	47.62	0.23	1.43
2006	0.49	58.03	43.83	0.25	1.30	2014	0.49	65.75	49.22	0.25	1.40
2007	0.49	60.76	44.94	0.26	1.29	2015	0.48	67.85	50.42	0.26	1.39
2008	0.49	58.01	44.87	0.23	1.35	2016	0.48	68.05	50.96	0.25	1.43
2009	0.49	54.28	43.67	0.20	1.44						

Notes: G' = Gini Coefficient Bottom 97%, $\langle r \rangle$ = Overall Average Income (1000's), $\langle w \rangle$ = Bottom 97% Average Income, f = Top 3% Income Share = $1 - (\langle w \rangle / \langle r \rangle)$, α = Top 3% Power Law Exponent.

reason why the sum of incomes (unlike the sum of money) must be conserved" (Dragulescu and Yakovenko, 2001, p. 588). Exponential distributions are also shown to obtain in diffusion models that only approximately conserve total money or even have declining total money, as well as those in which negative shocks are partitioned out among agents in a manner that keeps individual money stocks positive (Islas-Garcıa *et al.*, 2019). Finally, Ragab develops agent-based simulation models of energy conserving job-competition among workers in which the long term equilibrium is an exponential distribution (Ragab, 2013).

Many models of property incomes and wealth stocks are also proposed. Under the assumption that both money and wealth stocks are separately conserved, various rules of interactions between the two stocks can give rise to a gamma distribution of wealth or a power tail at high wealth. A power tail for high incomes also obtains if agents are assigned random savings rates drawn from a uniform distribution between 0 and 1. If we instead treat individual income as a stochastic process, we get an exponential distribution if income changes are assumed to be independent of the income level (additive diffusion), and a power law if income changes are assumed to be proportional to the income level (multiplicative diffusion) (Yakovenko, 2007, p. 2809).

3 A Kinetic Approach to the Dynamics of Income Distribution

Another strand of the econophysics approach seeks to explain income distribution through "a kinetic approach dealing with the temporal evolution of the probability distribution". The key tool is the Fokker Planck Equation in equation (1), where r is some form of income and $P(r,t)$ is a probability distribution (Silva and Yakovenko, 2005, pp. 308–309).

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial r}\left(AP + \frac{\partial}{\partial r}(BP)\right) \tag{1}$$

Setting r equal to labor income w, assuming Δw is independent of w (additive diffusion) and A and B are positive constants, the stationary solution $\frac{\partial P}{\partial t} = 0$ yields the exponential distribution $P(w) \propto e^{-\frac{w}{\left(\frac{B}{A}\right)}}$. Since a meaningful solution requires $A > 0$, the drift process will push the mean labor income ever downward while the diffusion keeps it aloft. This process requires a reflecting barrier at $w = 0$ (Silva and Yakovenko, 2005, p. 305).

Setting r in equation (1) equal to property income per person π, assuming that $\Delta\pi$ is proportional to π (multiplicative diffusion) and letting $A = a\pi$ and $B = b\pi^2$ for positive constants a, b, the stationary solution $\frac{\partial P}{\partial t} = 0$ yields the power law distribution $P(\pi) \propto \frac{1}{\pi^{\left(1+\frac{a}{b}\right)}}$ (Silva and Yakovenko, 2005, p. 309). Here too, the drift term drives mean property income downward, but the diffusion keeps it in check. Once again, the equivalent dynamic adjustment process requires a reflecting barrier at some appropriately high π.

We now translate the Fokker–Planck formulations into corresponding stochastic differential equations (SDEs) in order to facilitate direct comparisons with the arbitrage approach developed in the next section. Equation (2) depicts the SDE for labor income process and equation (3) for property income. Derivations are in the Appendix.

$$\frac{dw}{dt} = -A + \sqrt{2B}\frac{dW_t}{dt} \tag{2}$$

where A, B are positive constants and W_t is a Weiner process.

$$\frac{d\pi}{dt} = -a\pi + \pi\sqrt{2b}\frac{dW_t}{dt} \tag{3}$$

where a, b are positive constants.

In equation (2), the mean wage will fall and become negative at some point unless we impose a reflecting barrier at $w = 0$. The same applies equation (3), where the drift term drives the mean of property income downward. Once again, the equivalent dynamic adjustment process requires a reflecting barrier, this time at an appropriately large π (Silva and Yakovenko, 2005, pp. 305–309).

It is important to note that actual incomes are combinations of wage and property incomes, with the former dominating at lower ends and the latter at upper ones. On the assumption that the additive and multiplicative diffusion processes are uncorrelated, Banerjee and Yakovenko (2010, p. 11) and Yakovenko and Rosser (2009, p. 19) construct a distribution that smoothly interpolates between the exponential lower end and the power law upper end.

Then the former effectively serves as a reflecting barrier for the latter, which removes the need to explicitly specify a power law barrier. This is an important innovation.

The common characteristic of the preceding models is that mean wage and profit rates fall freely in the absence of diffusion effects. We would argue that this has no real analogue in economic analysis. For instance, in the Malthusian model when population grows faster than the food supply, standards of living fall, mortality rises, and population growth is brought back into line at some subsistence level of real wages. This is a multivariable model, and stochastic factors play no role in determining the convergence to a subsistence real wage. At the most abstract level, all economic theories of competition assume equalized wages, profit rates, interest rates and selling prices of given goods. The ruling principle is arbitrage in such cases: in a given country, labor moves from low wage regions to those with higher wages, capital from low profit rates applications to higher ones, buyers from high priced sellers to low priced ones, etc. (see the next section). In turn, the equalized levels of wages and profit rates, that is, the mean levels, are themselves determined by social-historical factors, including the active agency of the participants, existing technological conditions, and interventions by the state (Shaikh, 2016, Chs. 6, pp. 9–11).

In all cases, these variables are assumed to be equalized *by the drift process alone*. At the highest level of abstraction, each variable therefore appears as a single point. Yet we have seen in the Introduction to this paper that at an empirical level there are specific distributions of wage and property incomes, as well as of profit and interest rates – all of which *change* over time according to various general economic factors. In this paper we focus on wage and property incomes. Our aim is to show that there are general models of arbitrage in which the drift process, taken by itself would produce the single equalization point of economic theory. It is the *interaction* of the drift and diffusion processes that gives rise to the observed distributions. Since mean real wages and property incomes change over time, we normalize all variables by their respective means.

4 An Arbitrage Approach to Income Distribution

Arbitrage is an economic fundamental. The noted economist Irving Fisher illustrates it in the following manner: "[if] ... the cost of transporting wheat from Chicago to New York is, we shall suppose, about 6 cents per bushel ... the equalization of prices will be limited by the cost of transportation. The price in New York can never be above that in Chicago by more than 6 cents per bushel. For similar reasons, the prices in Chicago cannot exceed those in New York by more than the cost of transportation; otherwise the arbitrage merchant would buy wheat in New York and sell it in Chicago. It is by such arbitrage transactions that the prices of the same commodity in different markets seek a common level, just as water flowing from one reservoir to another tends to equalize the levels of the two. The more the costs of transportation are reduced, the more nearly equal will the prices of any commodity in different markets become." (Fisher, 1913, p. 335).

4.1 Labor income

In the case of labor incomes, differences in real wage levels (w) for any given type of labor will induce movement (drift) into high-wage areas, which will increase the labor supply relative to labor demand and hence bring the wage down. Conversely, the outflow from low-wage areas will decrease the labor supply relative to demand and bring the real wage up. This is a process subject to ceaseless fluctuations (Smith, 1973, pp. 58–161, 201–202; Shaikh, 2016, pp. 331–336). It only requires that agents are aware of the differences between high- and low-income regions, as in Fisher's example of wheat prices. If we abstract from costs of transportation, by itself the drift process will equalize wages across areas, that is, bring them exactly to the mean. Since the mean wage will itself be changing over time, we normalize wages by the mean appropriate to each type of model so that wage convergence is always centered around 1. At a more concrete level, we incorporate a diffusion term to accommodate a variety of individual factors affecting each agent's action. This gives us three possible drift-diffusion models for turbulent wage arbitrage. An important requirement is that the wage remain positive (or if we define the wage as the excess over some minimum wage, that the wage not fall below the minimum). Let $\theta =$ the strength of drift, $\sigma =$ the standard deviation of diffusion and $W_t =$ a Wiener process.

$$\frac{dw_t}{dt} = -\theta\left(w_t - 1\right) + \sigma\sqrt{w_t}\frac{dW_t}{dt} \quad \text{[CIR model]} \tag{4}$$

$$\frac{d\left(\ln w_t\right)}{dt} = -\theta\left(w_t - 1\right) + \sigma\frac{dW_t}{dt} \quad \text{[log-linear model]} \tag{5}$$

$$\begin{aligned}\frac{d\left(\ln w_t\right)}{dt} &= -\theta\left(\ln w_t - \ln 1\right) + \sigma\frac{dW_t}{dt} \\ &= -\theta\ln w_t + \sigma\frac{dW_t}{dt} \quad \text{[log-log model]}\end{aligned} \tag{6}$$

The first case is known as the Feller model in the mathematical statistics literature, and as the Cox-Ingersoll-Ross (CIR) model in the finance literature. In the latter domain, it is the seminal model of interest rate analysis, widely used in empirical work and the fount of many theoretical extensions (Dragulescu and Yakovenko, 2002, p. 443; Lamoureux and Witte, 2002, p. 1483). Positivity is ensured in this model by making the variance of noise proportional to the wage (variance $= \sigma^2 w$) so that the standard deviation of the diffusion term

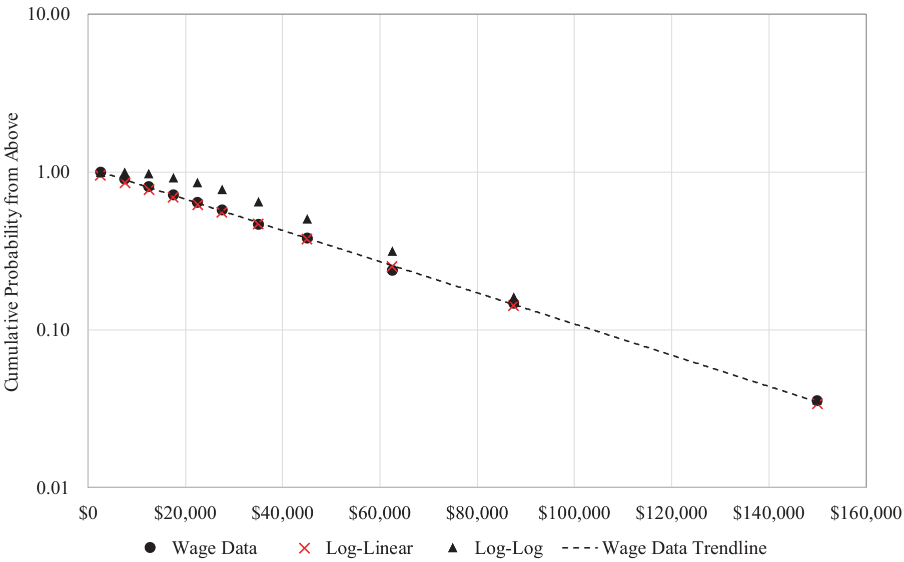

Figure 2: Cumulative Probability from Above, Wage Income Data and Fitted Log-Linear and Log-Log Models, 2011 (Log-Linear Scale).

is $\sigma\sqrt{w}$ (Cox *et al.*, 1985, p. 391). In the second model we consider $d(\ln w)$ on the left-hand side.[4] In this case, the wage remains positive because $\ln(w)$ only exists for $w > 0$, which is the same result as in the CIR model but without any restriction on the variance of the shocks. In the third model we represent mean reversion by log terms on both sides, which implies that the appropriate mean in this case is the geometric mean, not the arithmetic mean as in the first two models. Here too, wages remain positive because of the logarithmic formulation, without any restriction on the variance of the shocks. As previously noted, these arguments apply to the wages of given types of labor, so that partitions of the overall distribution by race and gender, as well as consolidations by occupation, will likely exhibit the same patterns.

The CIR model of equation (4) gives rise to a stationary gamma distribution that encompasses an exponential distribution (Cox *et al.*, 1985, p. 392). The log-linear model of 5 also gives rise to a stationary gamma distribution, which means that we can dispense with the CIR model in our analysis (the log-linear form now being a viable alternative also for interest rate analysis). Finally, the log-log model of equation (6) yields a stationary lognormal distribution. All proofs are in the Appendix. Figure 2 displays the fits of the log-linear and

[4]In stochastic calculus $d(\ln x) = (dx/x) - ((dx)^2/2x)$, so $dx/x \neq d(\ln x)$. If we write equation 6 in terms of (dx/x) then the mean is not equal to 1 except for particular combinations of parameter values.

log-log models with respect to the IRS individual labor income data previously displayed in Figure 1, with the log-linear model being better because its gamma distribution subsumes an exponential.

4.2 Property Income

Property income $\pi \equiv \rho a$ is the product of the stock of financial assets (a) and its rate of return (ρ),[5] both measured relative to their respective means. The rate of return, like labor income, is subject to arbitrage, so we can make use of the form of equations (5) and (6) as models for rates of return. For financial assets per person, it makes sense to posit that when the rate of return of the portfolio is above the mean, the portfolio will increase because its value will have increased and/or because more savings will flow into it. This gives us two possible expressions for the reaction of property income to the rate of return: $d(\ln a_t)/dt = \gamma(\rho - 1) + \sigma dW_t/dt$ or $d(\ln a_t)/dt = \gamma \ln \rho_t + \sigma dW_t/dt$. Because $\pi \equiv \rho a$ is an algebraic identity, we can write $d(\ln \pi_t)/dt = d(\ln \rho_t)/dt + d(\ln a_t)/dt$. Combining these with the two possible adjustment forms for financial assets yields two property income models (see the Appendix for details).

Log-linear Property Income Model

$$\frac{d(\ln \rho_t)}{dt} = -\theta(\rho_t - 1) + \sigma \frac{dW_t}{dt} \tag{7}$$

$$\frac{d(\ln \pi_t)}{dt} = \frac{d(\ln \rho_t)}{dt} + \frac{d(\ln a_t)}{dt} = (-\theta + \gamma)(\rho_t - 1) + \sigma \frac{dW_t}{dt} \tag{8}$$

Log-log Property Income Model

$$\frac{d(\ln \rho_t)}{dt} = -\theta \ln \rho_t + \sigma \frac{dW_t}{dt} \tag{9}$$

$$\frac{d(\ln \pi_t)}{dt} = \frac{d(\ln \rho_t)}{dt} + \frac{d(\ln a_t)}{dt} = (-\theta + \gamma)\ln \rho_t + \sigma \frac{dW_t}{dt} \tag{10}$$

Since property income models only apply to the top end, we restrict the distributions to the upper 5% domain by using a reflecting barrier (Gabaix, 2009, p. 262; Gabaix, 2016, p. 191). As noted, the signature of a power law is a straight-line plot of the log of the cumulative distribution from above versus the log of income. Figure 3 shows that in both models the top 5% of theoretical distributions approximate power laws that provide remarkably good fits to actual top income data, with the log-log being the better of the two since its upper reaches are close to a power law (Mitzenmacher, 2003, p. 229).

[5] The rate of return on a real or financial asset is not the same as the interest rate. Keynes emphasized that new investment is motivated by the difference between the two (Harrod, 1969, pp. 186, 193–194).

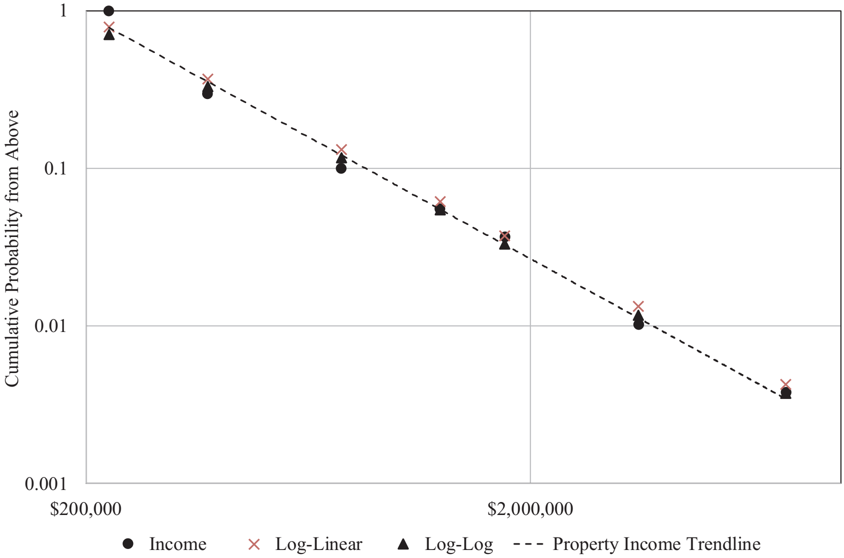

Figure 3: Cumulative Probability from Above, Property Income Data and Fitted Log-Linear and Log-Log Models, 2011 (Log-Log Scale).

Table 1 lists our overall empirical findings for 2002–2016: the empirical Gini coefficient for the exponential portion G' (the expected one being 0.50), the average income of the whole sample $\langle r \rangle$, the average income of the bottom 97% $\langle w \rangle$, the income share f of the top 3% , and the power law coefficient α.

As previously noted, actual income contains both wage and property incomes in varying degrees. Although it is beyond the scope of the present paper, in future work we hope to follow Banerjee and Yakovenko (2010) and Yakovenko and Rosser (2009) in constructing and testing a single distribution that interpolates smoothly between bottom and top distributions.

5 Other Economic Approaches to Income Distribution

We now compare our approach to two other economic ones. Nirei and Souma (2007, pp. 441, 448) argue that U.S. and Japanese income distributions, normalized in each year by the average of the whole distribution, consistently display an exponential pattern for 10^{th}–90^{th} percentile, a power law tail for the top 10^{th} whose exponent fluctuates in the range of 1.3–2.6, falling in bull markets and rising in bear ones. An important contribution of their paper is to show that the power law holds for the individual incomes of 80,000 top Japanese taxpayers. They assume that: labor income follows an additive stochastic process with a *growing* lower bound representing subsistence income (to account for incomes below the 10^{th} percentile); asset returns follows a lognormal distribution with a given mean; asset income is the product of asset size and its rate of return (hence a multiplicative process); and savings arise only from labor income.[6] The resulting Gini coefficient has effects from the exponential section, the power tail and the crossover point, and the power tail exponent is approximately equal one plus the ratio of savings from labor income to asset income. In comparison, we use the fundamental economic notion of mean-reversion to derive the observed distributions of both wages and the rate of return on financial assets, and to derive the top income power law tail on the assumption that each portfolio's relative asset size rises (falls) when its relative rate of return is above (below) the mean. Following (Silva and Yakovenko, 2005) the overall Gini coefficient $G = (1 + f)/2$ depends solely on the share of top income (f) in total income.

Venkatsubramanian (2017) takes the opposite tack: he derives an 'ideal' distribution of income in perfect capitalism. In standard neoclassical economic theory, utility maximizing consumers operate within perfect markets that lead to fully equalized wages for any given type of labor and fully equalized rates of return for any given capital asset. As previously noted, the exact equalization of incomes is the standard assumption in all theories of competition. Venkatasubramanian defines the effective utility from work, which he calls happiness, as utility from pay *minus* disutility from effort *plus* utility arising from a sense of a fair opportunity (49–51). A free market leads to individual maximization of happiness, hence to a fair distribution of income that is a unique Nash equilibrium in a noncooperative population game and an entropy maximizing state. On some further assumptions about the particular determinants of the three components of happiness, he arrives at a central prediction: in a "homogenous population of agents with the same payoff preferences . . . [that is], a classless society – utopia", the ideal income distribution, "the gold standard of fairness", will be a lognormal. In a system with distinct two classes of agents with different payoff corresponding to a bipopulation game, there will be two different fair income distributions, both lognormal, in which "the top lognormal could be easily misidentified as a Pareto distribution" (133–141).

Several issues are worth noting here. First, in the econophysics literature including our own, the term "class" refers to labor income and property

[6] Their equation (2) is $a(t+1) - \gamma(t)a(t) + w(t) - c(t)$ where $a(t+1) =$ the stock of assets in the next period, while $\gamma(t) =$ the rate of return on assets, $\gamma(t)a(t) =$ the flow property income, $w(t) =$ flow of wages and $c(t) =$ and the flow of consumption in this period. This is not stock-flow consistent since the left-hand side is a stock and the right-hand side elements are flows. If instead they had written $\Delta a(t + 1) = \gamma(t)a(t) + w(t) - c(t)$ then this would be consistent: the change in assets (a flow) would equal the flow of savings out of total (property and wage) income flows. But then their claim that savings comes only from labor income would not be true, and the model dynamics might be significantly altered (Nirei and Souma, 2007, p. 449).

income, not payoff preferences. Second, Venkatasubramanian assumes perfect equalization of wages for each type of labor, so that a distribution of labor income can only arise from wage differences among occupations. We argue instead that turbulent arbitrage, a combination of drift and diffusion, gives rise to a near-exponential distribution of wage rates for any given type of labor and hence for the overall distribution. As noted, in our case partitions by race, gender, and occupation are likely to result in similar shapes. Third, in our approach turbulent arbitrage need not be optimal or fair: the diffusion term is a representation of myriad concrete factors that generate satisfaction and disappointment in equal measure. Fourth, turbulent arbitrage, in this case of rates of return, also explains the power-law behavior of property incomes. Finally, as in Venkatasubramanian, our log-log models generate a lognormal distribution for labor income and a *different* one for property income. However, our log-linear wage models generate gamma distributions that fit labor incomes better and are just as good for property incomes, so that in our case there is nothing special about the lognormal.

6 Summary and Conclusions

Yakovenko and his co-authors have established that the bottom 97–99% of individual incomes (labor incomes) follow a near-exponential distribution while the top incomes (property incomes) follow a power law. One branch of econophysics relies on various monetary analogues to the physics principle of energy conservation. The labor income distribution is derived from two-body or multi-body monetary exchanges among agents that conserve total money. Under the assumption of entropy maximization, the corresponding stationary distributions are near exponential. Property income being the product of the stock of wealth and its rate of return, fixed savings rates will generate savings flows proportional to the levels of income. Since savings flows change the stock of wealth, changes in property income will be proportional to level of incomes (multiplicative diffusion) and their entropy maximizing distribution will be a power law.

Another branch of econophysics develops kinetic models of wage and property incomes built around the Fokker–Planck equation. In both cases, the drift process is assumed to push mean incomes downward while the diffusion keeps them aloft. Taken singly, the dynamic adjustment process in each case requires a reflecting barrier to keep the variable from going negative. However, on the assumption that the additive and multiplicative diffusion processes are uncorrelated, it is possible to construct a particular distribution that smoothly interpolates between the exponential lower end and the power law upper end. Then the former effectively serves as a reflecting barrier for the latter, which removes the need to explicitly specify a power law barrier.

From the point of view of economics, the fundamental principle of turbulent arbitrage provides the proper starting point. Differences in levels of real wages for any given type of labor will induce labor movement (drift) into high-wage areas, increase the labor supply relative to its demand and bring wages down. In low wage areas the opposite will occur. With only drift, wages will be equalized so that there will be a single equilibrium wage. At the same time, many concrete factors will influence individual wage paths (diffusion). Therefore, turbulent economic arbitrage, modeled as a mean-reverting drift-diffusion process, gives rise to various near-exponential stationary distributions of wages. In the case of property incomes, drift alone would equalize rates of return as is commonly supposed in economic models, so that it is the interaction between drift and diffusion that gives rise to a distribution. Under the plausible assumption that relative portfolios of financial assets rise in value and/or quantity when their individual rates of return are higher than the mean, the drift and diffusion processes will lead to a stationary power law distribution for property incomes. The log-linear formulation is better for wage models, since it yields a gamma distribution that encompasses an exponential. In turn, the log-log formulation is the better one for property income models, since it can generate a power law for upper incomes. Nothing requires the same formulation for the two domains. Finally, in the maximum entropy models of Section 2, stationary distributions are derived from the assumption of entropy maximization. In the kinetic models of Section 3 and the arbitrage models of Section 4, stochastic adjustment processes with dynamic paths yield stationary distributions that happen to be entropy maximizing. In the first case, entropy maximization is the means, in the second and third cases it is the result.

Appendix

Sources and Methods: Adjusted Gross Income (AGI) bins and number or returns in each bin are from IRS Tables 1.4 for 2002–2016, column 1–2. Our basic data consists of bin midpoints and return frequency. Total AGI is from IRS Table 1 (which turns out to be the same as Total AGI less deficit in Table 1.4), and $\langle r \rangle$ = average AGI = total AGI/total returns. The break points between labor and property income were identified by successively adding points to the lower section plot $\ln C(r)$ vs r until a significant curvature was evident. In all years, the dividing line turned out to be within the bin \$100,000 to \$200,000. Lower and upper sections were subsequently treated as separate populations, so that $C(r)$ begins from 1 in each, and MLE regression was used throughout. The lower section Gini coefficient (G') was calculated using the midpoint method in https://www3.nd.edu/~wbrooks/GiniNotes.pdf. Because the mean income of the lower section cannot be directly calculated from midpoint data, the average income in the bottom section ("income temperature"

$\langle w \rangle$) is estimated through the regression of $\ln C(r)$ vs r, while top income (f) was calculated as $f = 1 - \frac{\langle w \rangle}{\langle r \rangle}$ (Banerjee and Yakovenko, 2010).

Derivations of Equations (1) and (2): The Fokker–Planck equation in equation (1) is equivalent to the Ito stochastic differential equation $dr_t = -A(r_t)\,dt + \sqrt{2B(r_t)}\,dW_t$, where A and B are functions of r_t (Gardiner, 2009, p. 118). Equation (2) is obtained when both $A(r_t)$ and $B(r_t)$ are positive constants. Equation (3) is obtained when $A = ar$ and $B = br^2$, with r equal to the property income π, and a, b are positive constants.

Wage Income Models: The CIR model of equation (4) is known to generate a stationary gamma distribution (Cox *et al.*, 1985). For the log-linear wage model of equation (5) the corresponding Fokker–Planck equation is $\partial_t p(x_t) = \frac{\partial}{\partial x}(J)$ where $J = \theta(\exp(x_t) - 1)\,p(x_t) - \frac{\partial \sigma^2 p(x_t)}{2\partial x_t}$ is the probability current, $x_t = \ln(w_t)$ and w_t is the wage rate. In equilibrium $J = 0$, so $p(x) = \frac{1}{Z_x}\exp(-\frac{2\theta}{\sigma^2}(\exp(x) - x))$, where $\frac{1}{Z_x}$ is the inverse of the partition function. Using a transformation of random variables to derive the stationary distribution of wages, it is obtained that $p(w) = p(\ln(\mathrm{w}))|\frac{d}{dw}\ln(w)| = \frac{1}{Z_x}\exp\left(-\frac{2\theta}{\sigma^2}w\right)w^{\left(\frac{2\theta}{\sigma^2}-1\right)}$, which is a gamma distribution with mean equal to one (Casella and Berger, 2002, p. 50). Finally, for the log-log wage model of equation (6), $J = \theta x_t p(x_t) - \frac{\partial \sigma^2 p(x_t)}{2\partial x_t}$. In equilibrium, $p(x) = \frac{1}{Z_x}\exp\left(-\frac{\theta}{\sigma^2}x^2\right)$ Using a transformation of random variables,

$$p(w) = p(\ln(w))\left|\frac{d}{dw}\ln(w)\right| = \frac{1}{Z_x}\exp(-\frac{\theta}{\sigma^2}\ln^2 w)\left|w^{-1}\right|.$$

For $w \in [0, \infty)$, this corresponds to a log-normal distribution with a geometric mean (which is the median of w) equal to one.

Property Income Models: In the log-linear model, the rate of return (ρ) follows a gamma distribution because equation (7) has the same form as the wage equation (5). In the corresponding financial asset equation (8), the rate of growth of relative financial assets $d\pi_t = (-\theta + \gamma)(\rho_t - 1)\,dt + \sigma\,dW_t$ is the sum of a gamma and a normal distribution. Although we have not been able to derive a stationary solution to the distribution of financial assets (assuming one exists) we note that in the vicinity of equilibrium, $\rho_t \approx 1$ and the growth rate of financial assets behaves as a random growth process as described by Gabaix (2009, p. 3). Imposing a lower reflecting barrier on property income for this type of processes seems to lead to an approximate Power Law as confirmed in Figure 3. In the log-log model of equations (9)–(10), the dynamics of asset returns in equation (9) are similar to those of wages in equation (6) so that the stationary distribution of returns is the log-normal distribution. It follows that $y = \ln(\rho)$ is normally distributed. Since both $y = \ln(\rho)$ and the error terms are normally distributed, the growth rate of assets in equation (10) is a random growth process and imposing a lower reflecting barrier on property income leads to a Power Law (Gabaix, 2009, pp. 261–263; Gabaix, 2016, p. 192).

References

Banerjee, A. and V. M. Yakovenko. 2010. "Universal patterns of inequality". *New Journal of Physics*. 12(3): 1–25. DOI: 10.1088/1367-2630/12/7/075032.

Casella, G. and R. L. Berger. 2002. *Statistical Inference*. CA: Duxbury Pacific Grove.

Cox, J., C. J. E. Ingersoll, and S. A. Ross. 1985. "A theory of the term structure of interest rates". *Econometrica*. 53(1): 385–407. DOI: 10.2307/1911242.

Dragulescu, A. A. and V. M. Yakovenko. 2000. "Statistical Mechanics of Money". *The European Physical Journal B*. 17: 723–729. DOI: 10.1007/s100510070114.

Dragulescu, A. A. and V. M. Yakovenko. 2001. "Evidence for the exponential distribution of income in the USA". *The European Physical Journal B*. 20: 585–589. DOI: 10.1007/PL00011112.

Dragulescu, A. and V. Yakovenko. 2002. "Probability distribution of returns in the heston model with stochastic volatility". *Quantitative Finance*. 0: 443–453. DOI: 10.1080/14697688.2002.0000011.

Fisher, I. 1913. *Elementary Principles of Economics*. London: Macmillan.

Gabaix, X. 2009. "Power laws in economics and finance". *Annual Review of Economics*. 1: 255–293. DOI: 10.1146/annurev.economics.050708.142940.

Gabaix, X. 2016. "Power laws in economics: An introduction". *The Journal of Economic Perspectives*. 30(1): 185–205. DOI: 10.1257/jep.30.1.185.

Gardiner, C. 2009. *Stochastic Methods*. Berlin: Springer.

Harrod, R. F. 1969. *Money*. London: MacMillan.

Islas-Garcıa, J. D. A., A. R. Villagomez-Manrique, M. Castillo-Mussot, and P. G. Soriano-Hernandez. 2019. "Brownian motion, diffusion, entropy and econophysics". *Revista Mexicana de Fısica E*. (65): 1–6. DOI: 10.31349/RevMexFisE.65.1.

Lamoureux, C. G. and H. D. Witte. 2002. "Empirical analysis of the yield curve: The information in the data viewed through the window of Cox, Ingersoll, and Ross". *The Journal of Finance*. 57(2): 1479–1520. DOI: 10.1111/1540-6261.00467.

Marx, K. 1967. *Capital, Vol II*. New York: International Publishers.

Mitzenmacher, M. 2003. "A brief history of generative models for power law and lognormal distributions". *Internet Mathematics*. I(1): 226–251. DOI: 10.1080/15427951.2004.10129088.

Nirei, M. and W. Souma. 2007. "A two factor model of income distribution dynamics". *Review of Income and Wealth*. 53(2): 440–459. DOI: 10.1111/j.1475-4991.2007.00242.x.

Ragab, A. 2013. *Three Essays on the Incomes of the Vast Majority*. PhD Dissertation: New School for Social Research.

Shaikh, A. 2016. *Capitalism: Competition, Conflict, Crises*. New York: Oxford University Press.

Shaikh, A., N. Papanikolaou, and N. Weiner. 2014. "Race, gender and the econophysics of income distribution in the USA". *Physica A*. 415: 54–60. DOI: 10.1016/j.physa.2014.07.043.

Silva, C. A. and V. M. Yakovenko. 2005. "Temporal evolution of the "thermal" and "superthermal" income classes in the USA during 1983–2001". *Europhysics Letters*. 69(1): 304–310. DOI: 10.1209/epl/i2004-10330-3.

Smith, A. 1973. *The Wealth of Nations*. Harmondsworth, Middlesex: Penguin Books.

Venkatsubramanian, V. 2017. *How Much Inequality is Fair? Mathematical Principles of a Moral, Optimal, and Stable Capitalist Society*. New York: Columbia University Press.

Yakovenko, V. 2007. "Statistical mechanics approach to econophysics". *Encyclopedia of Complexity and System Science*. New York: Springer. DOI: 10.1007/978-0-387-30440-3_169.

Yakovenko, V. M. and J. B. Rosser. 2009. "Colloquium: Statistical mechanics of money, wealth, and income". *Reviews of Modern Physics*. (81): 1703–1725. DOI: 10.1103/RevModPhys.81.1703.