Article

Ergodic Inequality

Academic Editor: name
Version March 11, 2016 submitted to Entropy; Typeset by LaTeX using class file mdpi.cls

Abstract: Weak conditions are provided under which society's long-run distribution of wealth is independent of initial asset holdings. The reliance of this result on an objective interpretation of probability, and applications to Nozick's [3] "justice in acquisition" and Piketty's [4] persistent inequalities, are discussed. *JEL* Classification: D63, H1.

Keywords: inequality; ergodicity; libertarianism.

1. Introduction

Nozick's [3] libertarian theory of justice rests on two pillars: *justice in transfer* argues that holdings acquired through voluntary exchange are just, assuming that the parties concerned held legitimate title to the exchanged holdings; *justice in acquisition*, meanwhile, argues that the claim to ownership of a previously unowned resource by an individual is just, provided that it leave nobody else worse off. Both principles have been the subject of controversy (e.g. [1]), but here I focus on justice in acquisition—in particular, its *relevance* rather than its validity. This is particularly important for the application of libertarianism, given the practical impossibility of satisfying this principle. If, as argued recently by Thomas Piketty [4], initial asset holdings have enduring distributive effects, then they are of critical importance in both the theory of justice and the practice of egalitarianism. If, by contrast, the effects of unjust acquisition vanish over time, it offers little cause for concern in implementing a libertarian theory of justice.

In this note, I provide conditions under which justice in acquisition is irrelevant in this way. In particular, it would seem appropriate (and without loss of generality) to model the evolution of property rights as a stochastic process—a sequence of random variables—defined on the space of shares of society's wealth. In the theory of stochastic processes, a process may or may not satisfy the property of "ergodicity," under which every path of the process is representative of the whole; or, in other words, the initial conditions are irrelevant to "long-run" behaviour. By providing conditions under which the evolution of the societal division of wealth is ergodic, I show when unjust acquisition becomes irrelevant.

Of course, the significance of this enterprise is determined by the strength of my conditions. I make three main assumptions, each of which I argue to be "weak." First, I assume that the stochastic process governing wealth shares is a Markov chain. This essentially involves assuming that the division of wealth in the next period is (probabilistically) dependent on the current division of wealth, but not on divisions in previous periods. For this to be appealing, we must simply take a sufficiently long period length; presumably a generation would be ample, for instance. Second, for any given current division of wealth, there is a positive probability that it change in any "direction" in the next period; i.e. there is some chance that any given individual will be slightly better (or worse) off in the next period. This chance could be very small, but it must be positive.[1] Third, small changes in the current division of wealth should not unduly affect its evolution in the next period.

[1] It need not, however, be the same for getting better and worse off, nor for different individuals.

Under these conditions, the long-run division of wealth is independent of its starting point, and unjust acquisition thus becomes irrelevant over time.[2] However, the question then arises of how much time is required to reach—or at least come close to—this "long-run" division. If, for example, we require more periods than there are atoms in the universe, then we may not think the result so interesting. Rates of convergence are, unfortunately, difficult to determine in general. I can establish that ergodic behaviour is approached at a geometric rate, but this rate could still be very slow; between this observation and the lengthened period required for the Markov assumption, unjust acquisition may (or may not) remain relevant for a very long time. Nonetheless, geometric ergodicity establishes that the relevance of unjust acquisition diminishes period-by-period through time. Moreover, for any given approximating neighborhood of the ergodic distribution of wealth, there exists a finite length of time after which the process will always belong to that neighborhood.

Finally, I go on to argue that no such irrelevance result is possible in the absence of objective probabilities. The irrelevance of justice in acquisition in this case is from the modeler's perspective alone; his belief over the eventual distribution of wealth is ergodic, but the determinism of its underlying evolution could render it highly sensitive to its initial conditions forevermore. A libertarian is then forced to wrestle with the ethics of justice in acquisition.

2. The Evolution of the Division of Wealth

Consider a population of N individuals engaged in voluntary exchange through infinite discrete time $t \in \mathbb{Z}_+$. Individual i has a *wealth share* in period t of $x_i^t \in [0,1]$, with $x^t = (x_1^t, x_2^t, \ldots, x_N^t)$ describing the *state* of the process at time t, belonging to the *state space* $X := \{(x_1, x_2, \ldots, x_N) : x_i \in [0,1], \forall i; \sum_{j=1}^{N} x_j^t = 1\}$ of possible divisions of wealth.[3]

Assumption 1. The path of x^t over time is governed by a time-homogeneous Markov chain $\mathbf{\Phi} = \{\Phi^1, \Phi^2, \ldots\}$, taking values in X, and constructed from a set of transition probabilities $P = \{P(x, A), x \in X, A \in \mathcal{B}(X)\}$, where $\mathcal{B}(X)$ is the Borel σ-field on X, $P(\cdot, A)$ is a non-negative measurable function on X for each $A \in \mathcal{B}(X)$, and $P(x, \cdot)$ is a probability measure on $\mathcal{B}(X)$.

As mentioned in the Introduction, this assumption can be made appealing by taking a sufficiently long period length. The next assumption, meanwhile, is the driving force of the analysis.

Assumption 2. For every state $x \in X$, there exists a neighbourhood $\eta(x) \in \mathcal{B}(X)$ such that $P(x, A) > 0$ for all $A \in \mathcal{B}(\eta(x)) = \mathcal{B}(X) \cap \eta(x)$.

Thus, the division of wealth may, at any point, move in any "direction," i.e. there is some chance that any given individual will be slightly better (or worse) off in the next period.

A function h from X to $\mathbb{R}$ is called *lower semicontinuous* if

$$\liminf_{y \to x} h(y) \geq h(x), \qquad x \in X.$$

If $P(\cdot, O)$ is a lower semicontinuous function for any open set $O \in \mathcal{B}(X)$, then $\mathbf{\Phi}$ is called a *(weak) Feller chain*.

Assumption 3. $\mathbf{\Phi}$ is a weak Feller chain.

---

[2] In fact, the conditions are sufficient, but not necessary. I have avoided weakening them in order to retain their simplicity and ease their interpretation.

[3] With some nonzero probability, individual i may die in any given period t, to be replaced by a new "child" inheriting x_i^t. Formally, absent taxation, this is equivalent to having infinitely-lived individuals.

Intuitively, the Feller property requires that, if we change the current division of wealth slightly, the chance of next period's division of wealth shifting in a given way either increases or changes only slightly (in other words, this chance cannot jump dramatically downwards). This more technical assumption is harder to interpret intuitively, but it seems reasonable that small changes in the current division of wealth should not unduly affect its evolution in the next period.

3. Ergodicity

If μ is a signed measure[4] on $\mathcal{B}(X)$, then the *total variation norm* $\|\mu\|$ is

$$\|\mu\| := \sup_{f:|f|\leq 1} |\mu(f)| = \sup_{A\in\mathcal{B}(X)} \mu(A) - \inf_{A\in\mathcal{B}(X)} \mu(A).$$

For the present purposes, the key limit of interest to us is of the form

$$\lim_{t\to\infty} \|P^t(x,\cdot) - \pi\| = 2 \lim_{t\to\infty} \sup_A |P^t(x,A) - \pi(A)| = 0,$$

where π is an *invariant measure* of the process, i.e. a σ-finite measure on $\mathcal{B}(X)$ with the property

$$\pi(A) = \int_X \pi(dx)P(x,A), \qquad A \in \mathcal{B}(X).$$

If this sort of limit holds, then the long-run behavior of the process is described by the invariant measure π, independent of the initial measure from which the process starts. In particular if, for any initial measure λ,

$$\left\| \int \lambda(dx)P^t(x,\cdot) - \pi \right\| \to 0, \qquad t \to \infty,$$

then the process is said to be *ergodic*.

To get to this point, I will require some additional apparatus.[5] $\mathbf{\Phi}$ is called *φ-irreducible* if there exists a measure φ on $\mathcal{B}(X)$ such that, for all $x \in X$, whenever $\varphi(A) > 0$, there exists some $t > 0$, possibly depending on both A and x, such that $P^t(x,A) > 0$. It is called *ψ-irreducible* if it is φ-irreducible for some φ and the measure ψ is a "maximal irreducibility measure," guaranteed to exist by Meyn and Tweedie's [2] Proposition 4.2.2. Letting $\mathcal{B}^+(X) := \{A \in \mathcal{B}(X) : \psi(A) > 0\}$, if $\mathbf{\Phi}$ is ψ-irreducible and every set in $\mathcal{B}^+(X)$ is expected to be visited by $\mathbf{\Phi}$ infinitely often irrespective of the initial state, i.e. $\sum_{t=1}^{\infty} P^t(x,A) = \infty, \forall x \in X, \forall A \in \mathcal{B}(X)$, then $\mathbf{\Phi}$ is called *recurrent*. If it is ψ-irreducible and the probability that every set in $\mathcal{B}^+(X)$ is visited by $\mathbf{\Phi}$ infinitely often is 1 irrespective of the initial state, then $\mathbf{\Phi}$ is called *Harris recurrent*. If it is ψ-irreducible and admits an invariant measure π, then $\mathbf{\Phi}$ is called a *positive* chain. Finally, a set $C \in \mathcal{B}(X)$ is called *ν_m-small* if there exists an $m > 0$ and a non-trivial measure ν_m on $\mathcal{B}(X)$ such that, for all $x \in C$ and all $B \in \mathcal{B}(X)$,

$$P^m(x,B) \geq \nu_m(B).$$

Proposition 4. *$\mathbf{\Phi}$ is ergodic.*

Proof. Under Assumption 2, the process is φ-irreducible for any φ, and hence is trivially ψ-irreducible for any ψ with full support, i.e. any ψ such that $\psi(A) > 0, \forall A \in \mathcal{B}(X)$. Hence, $\mathcal{B}^+(X) = \mathcal{B}(X)$, and the recurrence of $\mathbf{\Phi}$ follows trivially from ψ-irreducibility. Since there are no ψ-null, transient sets, it

---

[4] μ is a *signed measure* on $(X, \mathcal{B}(X))$ if there are two finite measures μ_1 and μ_2 such that for all sets $A \in \mathcal{B}(X)$, $\mu(A) = \mu_1(A) - \mu_2(A)$.

[5] For a more complete account of the following terminology, see Meyn and Tweedie [2].

follows from Meyn and Tweedie's [2] Theorem 9.0.1 that $\mathbf{\Phi}$ is Harris recurrent. Moreover, it follows from their Theorem 10.4.4 that $\mathbf{\Phi}$ has a unique invariant measure π, and is hence a positive chain.

Now, suppose that $C \in \mathcal{B}(X)$ is ν_M-small for some $M \in \mathbb{Z}_+$, with $\nu_M(C) > 0$, and let

$$E_C = \{\, t \geq 1 : \text{ the set } C \text{ is } \nu_t\text{-small, with } \nu_t = \delta_t \nu_M \text{ for some } \delta_t > 0\}.$$

By lower semicontinuity of $P(\cdot, C)$, there exists $\delta > 0$ such that $P(x, C) \geq \delta$ for any $x \in C$, and hence C is also ν_t-small, with $\nu_t = \delta_t \nu_M$ for some $\delta_t > 0$, $t = M+1, M+2, \ldots$, by Meyn and Tweedie's Proposition 5.2.4(i). Thus, the greatest common divisor of the set E_C is 1, i.e. the process is *aperiodic*. The result then follows by Meyn and Tweedie's Theorem 13.3.3. □

Thus, the state x^t of the process is independent of the initial distribution λ after a sufficiently long period of time has passed.

But how long is "a sufficiently long period of time"? This question is difficult to address without significantly stronger assumptions, but a little more can be said in general, once I have introduced some final apparatus. A set $C \in \mathcal{B}(X)$ is called ν_a*-petite* if it satisfies the bound

$$\sum_{t=0}^{\infty} P^t(x, B) a(t) \geq \nu_a(B).$$

for all $x \in C$, $B \in \mathcal{B}(X)$, where ν_a is a non-trivial measure on $\mathcal{B}(X)$ and $a = \{a(t)\}$ is a probability measure on $\mathbb{Z}_+$. Clearly every small set is petite. Lastly, if $\mathbf{\Phi}$ is positive Harris and there exists a constant $r > 1$ such that

$$\sum_{t=1}^{\infty} r^t \| P^t(x, \cdot) - \pi \| \leq \infty,$$

then $\mathbf{\Phi}$ is called *geometrically ergodic*.

Proposition 5. *$\mathbf{\Phi}$ is geometrically ergodic.*

Proof. Since $\mathbf{\Phi}$ is ψ-irreducible with the Feller property, and ψ has full support on X, X is petite by Meyn and Tweedie's [2] Proposition 6.2.8. Condition (iii) of their Theorem 15.0.1 is then trivially satisfied for $V = 1$ and any $b \geq \beta > 0$. This implies that there exist constants $r > 1$, $R < \infty$ such that for any $x \in X$

$$\sum_t r^t \| P^t(x, \cdot) - \pi \| \leq R,$$

establishing the result. □

Thus $P^t(x, \cdot)$ converges to π at a geometric rate.

4. Discussion

My model of the evolution of the division of wealth is deliberately spare, and allows for a wide range of economic activities. For example, it might seem that the individuals in the model do not consume, and that this affects the results: Suppose that at time 0, Alice's land is worth \$100 and yields an income of \$10, whilst Bob's land is worth \$50 and yields an income of \$5, and both need to consume \$6 each period to survive. Then it is plausible that Bob will sell Alice some of his land, and convergence in this case would seem to be to Alice owning all land, by virtue of the fact that she started with the better endowment.[6] But this is not in fact the case, because (assuming neither individual can actually die, or really that both family lines survive in perpetuity)

[6] I thank Michael Allingham for this example.

Assumption 2's small probability of a reversal in the direction of land accumulation causes the process to oscillate between the two extremes of each individual owning all land, given long enough; the resulting long-run distribution over land shares is independent of the starting point. However, such considerations could certainly affect the speed of convergence; if, for instance, land-rich individuals were to consume more than land-poor ones, then convergence would be faster.

More fundamentally, this example highlights a key hidden feature of Assumption 2: In order to make the normative jump from properties of the ergodic distribution to philosophical implications, we have implicitly assumed that the transition probabilities governing the evolution of the process are objective. That is to say, the fundamental forces determining changes in the distribution of wealth from one period to the next are probabilistic. In particular, there is some chance that any given individual will be slightly better (or worse) off in the next period, and that chance is fundamental rather than an expression of the modeler's uncertainty about the true process. This is a highly debatable assumption: If I toss a coin, it is a very good model of the outcome to say that it comes up Heads or Tails with probability 1/2 each, but that does not imply that the model is a true description of the fundamental process at work. If I knew the force and angle at which the coin was propelled, the air resistance and so on, I could theoretically compute a more accurate prediction. Moreover, there is another possible outcome—the coin landing on its side—that I have completely ignored. Indeed, if I had enough information, I could arguably predict the outcome with probability 1. Similarly, probabilities provide a very good model of subjective uncertainty in economic modeling.[7] But if we interpret the model's transition probabilities as nothing more than an expression of the modeler's uncertainty, then only his belief can be taken to be ergodic, and the possibility of a deterministic evolution of the division of wealth that is eternally sensitive to its initial conditions cannot be ruled out.

In the context of the above example, we might think it more reasonable to suppose that there is *zero* probability of leaving the state where Alice owns all of the land, if only because Bob (or his family line) dies out. Or even ignoring the extreme outcome of death—for instance, by replacing it with a large utility loss—the economic forces at work appear too strong to leave any possibility of Alice losing all of her land. This is inconsistent with my Assumption 2.

Might such concerns be alleviated in the presence of economic growth? By studying the evolution of society's *division* of wealth, I have left absolute *levels* of wealth unmodelled. In the above example, for instance, if the income from land grows by 20% each period, Bob's initial $50 of land will be sufficient for his $6 consumption needs within one period. However, for any given growth level, there will clearly exist initial distributions that cannot be so readily escaped—at a minimum, the extreme case where one individual starts off owning all of the land. The acclaimed recent work of Piketty [4] is concerned with less extreme cases where there is nonetheless strong pressure towards unequal distributions, arising from the tendency for returns on assets to exceed growth rates and the resulting difficulty of reducing differences in asset holdings. Such differences must eventually be irrelevant under an objective-probability interpretation of my model, but could have everlasting effects under subjective probabilities.

---

[7] Perhaps the dominant interpretation of probability in economics is subjective [5]; that probability captures an individual's "degree of belief," and hence need not be common across individuals or correspond to any fundamental property of the "real world." However, objective probability does still feature prominently in economic theory, for instance in the expected utility theory of von Neumann and Morgenstern [6].

Conflicts of Interest: The authors declare no conflict of interest.

Bibliography

1. G. A. Cohen. *Self-Ownership, Freedom, and Equality*. Cambridge University Press, Cambridge, 2010.
2. S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993.
3. Robert Nozick. *Anarchy, State and Utopia*. Basic Books, Malden, MA, 1974.
4. T. Piketty. *Capital in the Twenty-First Century*. Harvard University Press, Cambridge, Massachusetts, 2014.
5. L. J. Savage. *The Foundations of Statistics*. John Wiley & Sons, 1954.
6. J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.