Ergodic Inequality

Thomas W. L. Norman

Magdalen College, High Street, Oxford OX1 4AU, UK; thomas.norman@magd.ox.ac.uk; Tel.: +44-1865-276034

Abstract: Weak conditions are provided under which society's long-run distribution of wealth is independent of initial asset holdings.

1. Introduction

Nozick's [1] 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., [2]), 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 Piketty [3], 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, I model the evolution of property rights as a stochastic process on the space of shares of society's wealth. Such 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 adequate, and we can henceforth think in these terms. 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 remain relevant for a very long time. Nonetheless, geometric ergodicity establishes that the relevance of unjust acquisition diminishes period-by-period. 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.

But with finitely lived agents then, does the long-run nature of the analysis doom its own conclusions to irrelevance, particularly with a generation as the period length? If I take all of your money now, for instance, the opportunity many generations hence for your descendants to rob my descendants would be a weak argument for the justice of your current poverty. This is not, however, the domain of justice in acquisition, but rather of Nozick's other libertarian pillar, justice in transfer. For my forced appropriation of your assets is a violation of voluntary exchange, rather than an unjust acquisition of a previously unowned resource. Put differently, the ergodic distribution to which society's division of wealth tends may well be unjust, but this is a failure of the mechanisms of exchange that determine that distribution (i.e., injustice in transfer), rather than the initial societal division of wealth (i.e., injustice in acquisition). It is with conditions for the irrelevance of the latter that I am concerned here.

There is of course a developed literature endogenising the distribution of wealth in a variety of economic models, some of which produce ergodic distributions as their outcome [4–8]. Where these models are rich in their analysis of the details of societal wealth evolution, my model is deliberately sparse, and seeks to provide a simple set of sufficient conditions for ergodicity that are satisfied by these papers. I discuss them, along with other phenomena that can and cannot fit within my framework, in the final section.

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

---

[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.

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 $\boldsymbol{\Phi}$ is called a *(weak) Feller chain*.

Assumption 3 (Downward smoothness)**.** $\boldsymbol{\Phi}$ is a weak Feller chain.

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] $\boldsymbol{\Phi}$ 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 [9] Proposition 4.2.2. Letting $\mathcal{B}^+(X) := \{A \in \mathcal{B}(X) : \psi(A) > 0\}$, if $\boldsymbol{\Phi}$ is *ψ*-irreducible and every set in $\mathcal{B}^+(X)$ is expected to be visited by $\boldsymbol{\Phi}$ 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 $\boldsymbol{\Phi}$ is called *recurrent*. If it is *ψ*-irreducible and the probability that every set in $\mathcal{B}^+(X)$ is visited by $\boldsymbol{\Phi}$ infinitely often is 1 irrespective of the initial state, then $\boldsymbol{\Phi}$ is called *Harris recurrent*. If it is *ψ*-irreducible and admits an invariant measure π, then

---

[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 [9].

Φ 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. Φ *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 Φ follows trivially from ψ-irreducibility. Since there are no ψ-null, transient sets, it follows from Meyn and Tweedie's [9] Theorem 9.0.1 that Φ is Harris recurrent. Moreover, it follows from their Theorem 10.4.4 that Φ 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. $\square$

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 Φ 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 Φ is called *geometrically ergodic*.

Proposition 5. Φ *is geometrically ergodic.*

Proof. Since Φ is ψ-irreducible with the Feller property, and ψ has full support on X, X is petite by Meyn and Tweedie's [9] 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. $\square$

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 sparse, and allows for a wide range of economic activities.

4.1. Consumption and Growth

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 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.

We might, however, 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. Such a "poverty trap" for Bob is inconsistent with my Assumption 2, which is hence unrealistic in certain settings.

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, for instance, the extreme case where one individual starts off owning all of the land. The recent work of Piketty [3] 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.

4.2. Imperfect Capital Markets

With perfect capital markets and identical individuals, everybody has access to the same investment opportunities, hence any persistent wealth inequalities at most reflect initial inequalities, and indeed will vanish when the proportion of income saved is nonincreasing in initial wealth. But with capital market imperfections, investment opportunities vary with inherited wealth and initial wealth inequalities persist in steady state [10]. Nonetheless, a number of papers derive an ergodic distribution of income or wealth. Champernowne [4] shows that, when individuals cannot insure against idiosyncratic equiproportional wealth shocks, there is a unique invariant lognormal distribution of wealth. This result exploits a law of large numbers, possible because of the assumed independence of individual wealth shocks, generating a process consistent with my Assumption 2. Loury [5], by contrast, allows for optimising investments by agents, which are directly related to inherited wealth in the absence of credit markets, so that earnings are correlated across generations of the same family. Diminishing returns to investments and strictly concave utility then imply a contraction mapping in the wealth distribution, which establishes convergence to a unique invariant distribution with persistent inequalities. Banerjee and Newman [6] introduce capital market imperfections resulting from incentive problems in diversifying risky idiosyncratic investments. Higher initial wealth can lead to lower risk-taking due to a tighter constraint on diversifying that risk (to preserve effort incentives), and hence the wealth of the poor can increase faster on average than that of the rich. This again generates a process consistent with my Assumption 2, and yields the existence of a unique ergodic distribution.

---

[6]   I thank Michael Allingham for this example.

Aghion and Bolton [7], meanwhile, endogenise the market rate of return in a model with risk-neutral individuals of limited wealth, whose incentives to exert effort are decreasing in the borrowing required to finance their investments. While the poor in this model initially do not invest, as capital accumulates over time the equilibrium cost of capital decreases, an increasing fraction of the poor can invest and limits are placed on the accumulation of individual wealth. [7] If there is sufficiently rapid capital accumulation then, the distribution of wealth eventually converges to a unique invariant distribution. Otherwise, there exist multiple invariant distributions, as in Picketty [8], where it is possible that both high and low interest rates are self-sustaining; higher interest rates induce a higher fraction of credit-constrained individuals, and hence lower long-run capital accumulation.[8] Each of these outcomes is itself associated with an ergodic distribution, but one determined by the interest rate. This is consistent with my model, and highlights the importance of focusing on the determinants of the ergodic distribution rather than the initial conditions of the process. Ergodicity is not inconsistent with persistent inequality, but merely with its determination by the model's starting point. Put differently, it is Nozick's justice in transfer that matters, not justice in acquisition.

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

References

1. Nozick, R. *Anarchy, State and Utopia*; Basic Books: Malden, MA, USA, 1974.
2. Cohen, G.A. *Self-Ownership, Freedom, and Equality*; Cambridge University Press: Cambridge, UK, 2010.
3. Piketty, T. *Capital in the Twenty-First Century*; Harvard University Press: Cambridge, MA, USA, 2014.
4. Champernowne, D. A model of income distribution. *Econ. J.* **1953**, *63*, 318–351.
5. Loury, G. Intergenerational transfers and the distribution of earnings. *Econometrica* **1981**, *49*, 843–867.
6. Banerjee, A.; Newman, A. Risk-bearing and the theory of income distribution. *Rev. Econ. Stud.* **1991**, *58*, 211–235.
7. Aghion, P.; Bolton, P. A theory of trickle-down growth and development. *Rev. Econ. Stud.* **1997**, *64*, 151–172.
8. Piketty, T. The dynamics of the wealth distribution and the interest rate with credit rationing. *Rev. Econ. Stud.* **1997**, *64*, 173–189.
9. Meyn, S.P.; Tweedie, R.L. *Markov Chains and Stochastic Stability*; Springer-Verlag: London, UK, 1993.
10. Galor, O.; Zeira, J. Income distribution and macroeconomics. *Rev. Econ. Stud.* **1993**, *60*, 35–52.
11. Matsuyama, K. Endogenous inequality. *Rev. Econ. Stud.* **2000**, *67*, 743–759.

---

[7] Matsuyama [11] considers the possible failure of this "trickle-down" mechanism when the rich households continue to borrow as they accumulate wealth, allowing a poverty trap to emerge under certain conditions. This is of course inconsistent with my Assumption 2.

[8] A similar possibility arises in a labour-market context in [6].