Difficulties in Testing for Capital Overaccumulation

Narayana R. Kocherlakota*

November 6, 2023

Abstract

This paper re-considers the question of testing for the presence of Pareto suboptimal capital overaccumulation in overlapping generations economies. The paper allows generation-specific technology shocks to evolve over time according to a stationary Markov chain, and assumes that an econometrician observes a finite sample of aggregate quantities. In this setting, any statistical test of the null hypothesis of capital overaccumulation with size less than one also has zero power against the alternative hypothesis of a dynamically efficient steady state. This result means that the standard assessments of capital overaccumulation based on US aggregate quantity data should be viewed as inconclusive.

*University of Rochester and NBER. This paper originally circulated under the title, "Testing for Capital Overaccumulation is Impossible." I thank two referees for helpful and detailed comments.

1 Introduction

Diamond (1965) proves that it is theoretically possible for a government to undertake a sequence of intergenerational transfers that, relative to a competitive equilibrium outcome, simultaneously improves the welfare of all agents and induces a fall in the capital stock. However, most research in modern macroeconomics abstracts from this form of Pareto sub-optimal *capital overaccumulation*. The basis for this approach is empirical. In his model without risk, Diamond shows that a necessary condition for capital overaccumulation is that the steady-state physical return to investment (net of depreciation) is less than the growth rate. As described in the next section, this requirement has been assessed empirically in US data by a variety of authors, perhaps most notably by Abel, Mankiw, Summers, and Zeckhauser (AMSZ) (1989). It is typically seen as being strongly violated not just on average, but for essentially every annual observation.

In this paper, I re-visit the question of testing for capital overaccumulation. Unlike Diamond (1965), I allow for aggregate risk in the form of intergenerational productivity shocks governed by a stationary Markov chain. I suppose that an econometrician observes a finite sample of aggregate output and investment, which she wishes to use to test the null hypothesis of capital overaccumulation. Her task is considerably simplified by her knowing much about the economy: its trend growth rate, the production function, the rate of depreciation and the set of all possible productivity states. Given her large knowledge base, her testing problem is only that she does not know the transition probability matrix of the productivity Markov chain. Some of the Markov chains imply that there is capital overaccumulation. Others do not.

In this context, I show that any test of the null hypothesis of capital overaccumulation with size less than one has zero power against the alternative hypothesis of a dynamically efficient steady-state.[1] As a result, the US data cannot be seen as conclusively rejecting

[1] I am using a standard definition (Romano, Shaikh, and Wolf (2010, p. 77)) of test size: it is the supremal probability of rejection where the supremum is taken with respect to all models/parameters consistent with the null hypothesis. In the current paper, the supremum is calculated with respect to all Markov chains that

the possibility of a Pareto improving increase in young-to-old transfers that would induce a reduction in capital accumulation.

The intuition behind the result is simple. Suppose that we observe a finite sample such that the economy is always in some state 1 in which the net physical return to capital $(MPK - \delta)$ exceeds the growth rate g. If this state were known to persist forever, then any decrease in capital would reduce available resources and necessarily make some agents worse off. But suppose instead there is some positive probability that the economy could transit to a state 2 in which $(MPK - \delta) < g$. As long as that state is sufficiently persistent, it is then possible to construct a young-to-old transfer increase in *both* states (with a tilt toward state 2) so that that the young agents born in state 1 are made better off. This transfer increase will lead the young agents to cut back on capital accumulation.

I prove two additional results. First, I specifically assess the "spout-or-sink" test implemented by AMSZ (1989). As they demonstrate via its application, this test has positive power. I prove that it has size equal to one.[2] The issue is that AMSZ's test for capital overaccumulation is a delicate one. It does work (as they formally prove) if the physical return to investment is either above the growth rate (implying optimality) or below the growth rate (implying suboptimality) in *all dates and states*. But, as Barbie, Hagedorn, and Kaul (2004) emphasize, this is a non-testable condition. The current paper shows that the AMSZ test may miss a (positive probability) possibility of transiting to a persistently negative $(MPK - \delta - g)$ state, and thereby incorrectly reject the hypothesis of Pareto suboptimality.[3]

The final result concerns the connection between dynamic inefficiency and ex-post Pareto suboptimality in stochastic economies. Dynamic efficiency is a preference-free concept: An allocation is said to be dynamically inefficient if it is possible to reduce capital in a given

are consistent with capital overaccumulation. The term power is also standard: it refers to the probability of rejection if the alternative is true, rather than the null.

[2]The proof is for the class of Markov chain models considered in this paper. However, since size is defined as a supremal probability of rejection, the statement remains valid for any superset of this baseline class of models.

[3]In a similar vein, AMSZ note that, "dynamic efficiency cannot in principle be judged by observing only a particular segment of time."

period, without lowering aggregate consumption in any successor dates and states.[4] The two concepts (ex-post Pareto suboptimality and dynamic inefficiency) are essentially equivalent in economies without risk. But this paper shows that this equivalence is not robust to the addition of aggregate risk. In particular, Section 5.2 demonstrates that Pareto *suboptimal* allocations which exhibit capital overaccumulation may nonetheless be dynamically efficient.[5]

The bulk of the paper describes difficulties in using quantity-based tests of capital overaccumulation. Could one instead use a test based on bond prices/yields? It is true that there are situations in which such a price-based test would be effective. Suppose, for example, that young agents could trade bonds of all maturities, and that r_{long} is the limiting value on the far right of the yield curve. Then, we can use the logic in Kocherlakota (2023) (developed for an endowment economy) to show that that capital is overaccumulated if r_{long} is less than the growth rate[6] g. Intuitively, it is possible to test for capital overaccumulation using bond yields because the *forward-looking* variable r_{long} contains potentially valuable information about low-probability transitions that cannot be observed in a given finite (and intrinsically *backward-looking)* sample of aggregate quantities.

But this argument (that one can test for capital overaccumulation by checking if r_{long} is smaller than g) does not generalize to a world with plausible *financial frictions*. Along these lines, suppose that the agents who engage in physical investment face binding borrowing constraints (or are simply excluded from asset market participation). Then market interest rates may be (much) lower than the shadow interest rates of those agents who have access to the physical investment technology. This wedge between shadow and market rates means that we cannot conclude that capital is overaccumulated just because $r_{long} < g$. In Section 6 of the current paper, I discuss this point further and I illustrate it in an example model in

[4]My use of the term "dynamic (in)efficiency" mirrors Zilcha (1991). Barbie, Hagedorn, and Kaul (2007) instead use Cass (1972)'s term "capital overaccumulation" to refer to what Zilcha (1991) calls "dynamic inefficiency". (Cass' (1972) paper is about the properties of deterministic neoclassical growth models.) In the current paper, the characterization "capital overaccumulation" is explicitly grounded in the welfare criterion of ex-post Pareto optimality.

[5]Bertocchi (1991) reaches a related but distinct conclusion.

[6]In this class of Markov economies, r_{long} is constant. See also Hansen and Scheinkman (2009), Martin and Ross (2019) and Bloise, et al. (2017).

Online Appendix B.

All proofs are in the Appendix.

2 Prior Empirical Implementations of Quantity-Based Tests

In this section, I briefly discuss how prior authors have used observations of aggregate quantities in the US to evaluate the possibility of capital overaccumulation a la Diamond (1965). My main point is that the basic approach takes the form of what can be called an "eyeball" test. Such a test rejects the null hypothesis of capital overaccumulation by showing in a table or graph that all annual observations of capital and output in the US are dynamically efficient (meaning that $(MPK - \delta)$ exceeds the growth rate g in every observation). We shall see in Sections 4 and 5 that the eyeball test (of the null hypothesis of capital overaccumulation) is flawed because it has size equal to one.

2.1 AMSZ (1989)

Abel, Mankiw, Summers, and Zeckhauser (AMSZ) (1989) remains the foundational paper for empirical tests of capital overaccumulation. They propose a variety of methods of checking for dynamic efficiency using US aggregate data. As discussed in more detail in Section 5, these tests can roughly be seen as assessments of whether the investment sector is a "spout or sink". For example, in their Table 1, they compare Profit/GNP to Investment/GNP in the US for the years 1929-85. They see no need for a nuanced parsing of the data. as the former exceeds the latter in every year by at least eight percentage points. In their Table 2, they specialize the analysis to the non-financial sector in the years 1953-85. The story remains the same.

2.2 Blanchard (2019)

Blanchard (2019) undertakes a thorough (re-)assessment of the potential benefits of a further expansion of US debt. In his theoretical work, he allows for a risk-based difference between the return to physical capital and the return to safe assets. He then turns to the data, by saying on p. 1220 that, "In the simulations above ... the welfare effects of an average marginal product far above the growth rate typically dominated the effects of an average safe rate slightly below the growth rate, implying a negative effect of the transfer (or of debt) on welfare. **Such a configuration would seem to be the empirically relevant one**." (emboldening mine).

What is the basis for Blanchard's key last sentence? His Figure 15 on p. 1222 plots the return to physical capital, measured by using the ratio of firms' net operating surplus to the replacement value of their capital stock. The graph shows that this return never falls below eight percent in any year from 1950-2016.[7] Again, Blanchard's conclusion is essentially an application of the eyeball test that Section 4 proves has size equal to one.

2.3 Reis (2021)

In the opening paragraph of a recent working paper, Reis (2021) asserts that "the US data ... strongly suggest the marginal product of capital (m) has stayed relatively constant, well above the growth rate of output." He reaches that conclusion through the examination of two simple graphs (Panels (c) and (d) of Figure 1 on page 2). The first (panel (c)) plots the marginal product of capital, measured as the ratio of net (of depreciation) value added to the non-financial corporate capital stock.[8] That line lies well above the graph of the growth rate

[7] The same graph also depicts the return to physical capital as measured by the ratio of firms' net operating surplus to the financial market value of their capital stock. Blanchard suggests that this variable might be a better measure of the return to capital, given that a growing fraction of earnings is attributable to market power and other rents. Nonetheless, the graph shows that this measure never falls below four percent in any year from 1950-2016.

[8] Geerolf (2018) argues that taking proper account of the labor component of income from unincorporated businesses could weaken, if not overturn, AMSZ's conclusions. However, Reis (2021) explicitly documents that his graphical conclusion about 2000-2021 is robust to this consideration.

of output in every year from 2000-2019. The second (panel (d)) is intended to be a parallel to AMSZ. It plots capital income relative to output, and investment relative to output. It shows that the former variable greatly exceeds the latter in every year from 2000-2019.

Reis (2021) uses these graphical checks to justify his focus on a steady state in which m - the marginal product of capital net of depreciation - always exceeds g. In this sense, he is rejecting the null hypothesis of capital overaccumulation in favor of the alternative that the economy is in a dynamically efficient steady-state. His approach is another application of the eyeball test.

3 Setup

This section describes the class of overlapping generations (OLG) models that is the basis of the analysis, intergenerational transfers within those models, and what is meant by Pareto (sub)optimality.

3.1 Models

This subsection describes the class of OLG models used in the paper.

Time is discrete and is indexed by the natural numbers (so, unlike some prior research in this area, there is an initial date). At date t, a measure 1 agents is born; each member lives for two periods labelled "young" and "old". Every agent has a utility function of the form:

$$ln(c^y) + \beta ln(c^o)$$

over consumption c^y when young and consumption c^o when old. At date 1, there is a measure 1 of initial old agents who live for only one period; they prefer more consumption to less.[9]

The state s_t of the economy is governed by a Markov chain with a finite state space

[9]I use log utility to simplify notation. All of the results in the paper can be generalized to the case of arbitrary power utility functions.

$\{j\}_{j=1}^{J}$ and a $(J \times J)$ transition probability matrix P, where P is restricted to have a unique stationary vector. I denote the stationary vector implied by P as $\pi^{stat}(P)$. The probability of state s_1 equalling j is then given by $\pi_j^{stat}(P)$ (so as to ensure the stationarity of the shock process).

When the Markov state in period t is s_t, each old agent in period $(t+1)$ is endowed with $e_{s_t}^{o}(1+g)^t$ units of consumption, where $e_j^o \geq 0$ for all $j \in \{1, 2, ..., J\}$ and the cross-cohort growth rate g is non-negative.[10] The initial old are endowed with e_{init}^{o} units of consumption.

Each young agent in period t is endowed with $e_{s_t}^{y}(1+g)^t$ units of consumption, where $e_j^y > 0$ for all $j \in \{1, 2, ..., J\}$. A young agent can forego consumption to create capital, which in turn generates consumption for them in the following period. Specifically, if a young agent invests K_t units of consumption at date t, that investment creates:

$$(1-\delta)K_t + A_t K_t^{\alpha}, 0 < \alpha \leq 1, 0 \leq \delta \leq 1 \tag{1}$$

units of consumption in period $(t+1)$. (Note that α may be equal to one.) The initial old agents have no output produced from capital.

The variable A_t in (1) is a technology parameter that is common to all young agents at date t. It fluctuates stochastically around (a scaled version of) the same log-linear trend as do endowments. Specifically:

$$A_t = a_t(1+g)^{t(1-\alpha)}$$

(The log-linear trend is scaled by $(1-\alpha)$ so that output produced by capital grows at the same rate g as endowments.) The technology shock $\{a_t\}_{t=1}^{\infty}$ has J possible realizations $\{\bar{a}_j\}_{j=1}^{J}$ and, when the Markov state in period t is s_t, the realization of $a_t = \bar{a}_{s_t}$.

I close this subsection with three observations about the class of models.

- There is no within-cohort risk within the models, as young agents are fully informed about the realizations of their endowments and investments in the following period

[10] I restrict attention to deterministic growth in order to simplify the relevant notation. The main result (Proposition 4) is readily generalized to the case of stochastic cross-cohort growth.

(when they are old). However, there is cross-cohort risk that is shaped by the evolution of the Markov state.

- Note that capital serves as a person-specific accumulation technology for each young agent. As a result, it plays no role in cross-generational accumulation. Instead, cross-generational output growth is generated by endowment and productivity growth.

- There is no labor in the models. Adding period $(t+1)$ labor as a complementary input to previously accumulated capital in the production of output would greatly complicate the analysis by adding a state variable, as the young agents' period $(t+1)$ incomes would then depend on period t capital choices.

3.2 Transfers

This subsection describes the impact of non-negative transfers from young to old agents. The transfers are history-independent, so that they depend only on the current realization of the Markov state.

Let $\tau \equiv (\tau_1, \tau_2, ..., \tau_J) \in \mathbb{R}^J_+$ be a non-negative J-dimensional vector of state-dependent transfers. (Here, and throughout, I use the letters i and j to represent the corresponding realizations of the Markov state.) At date t, if the Markov state is s_t, each young agent gives up $\tau_{s_t}(1+g)^t$ and each old agent receives $\tau_{s_t}(1+g)^t$.

Given a transfer vector τ, each young agent born in state i can choose a non-negative capital level in response to the capital plan. Hence, a young agent born in state i in period t derives lifetime utility $V^*_{it}(\tau)$, which is defined as:

$$V^*_{it}(\tau) = max_{k\geq 0}\begin{cases} ln(e^y_i(1+g)^t - (1+g)^t k - \tau_i(1+g)^t) \\ +\beta\sum_{j=1}^J P_{ij} ln(e^o_i(1+g)^t + \bar{a}_i(1+g)^t k^\alpha + (1-\delta)(1+g)^t + \tau_j(1+g)^{t+1}), i = 1, .., J \end{cases}$$

By netting out $t * ln(1+g)$ from this expression, we can obtain a time-invariant representation for young agents' welfare:

$$V^*_{it}(\tau) = V_i(\tau; P) + t * ln(1 + g), i = 1, ..., J$$

where:

$$V_i(\tau; P) = max_{k\geq 0}\ ln(e^y_i - k - \tau_i) + \beta \sum_{j=1}^{J} P_{ij} ln(e^o_i + (1-\delta)k + \bar{a}_i k^\alpha + \tau_j(1+g)), i = 1, ..., J \quad (2)$$

The initial old agents receive utility according to their consumptions:

$$e^o_{init} + \tau_i(1 + g), i = 1, ..., J$$

as a function of the initial state.

3.3 Pareto (Sub)optimality

The above formulation of welfare gives rise to a partial ranking of transfer plans.

Definition 1. Let P be a $(J \times J)$ Markov matrix with the unique stationary vector $\pi^{stat}(P)$, and let $S(P) = \{j|\pi^{stat}_j(P) > 0\}$ be the set of states with positive probability according to P. Given the transition probability matrix P, the transfer plan τ' *Pareto dominates* the transfer plan τ if:

$$\tau'_i > \tau_i, i \in S(P)$$

$$V_i(\tau'; P) > V_i(\tau; P), i \in S(P)$$

A transfer plan τ is *Pareto suboptimal* if it is Pareto dominated by another transfer plan τ'. A transfer plan τ is *Pareto optimal* if it is not Pareto dominated by any other transfer plan.

Here, τ' Pareto dominates τ if it makes all agents strictly better off, including the initial old, in all states that have positive probability under P. Note that an agent's welfare is calculated *conditional* on the realization of the state at the time of their entry into the economy. This notion of Pareto optimality is the same as Muench (1977)'s notion of "conditional Pareto optimality" and what Demange and Laroque (1999) term "interim Pareto optimality". It is also the same as the welfare criterion described on page 5 of AMSZ (1989). Because the welfare calculation is done after the determination of the state at the time of an agent's birth, Pareto dominance is not about finding opportunities for cross-generational insurance. In fact, Pareto improvements in transfer plans necessarily *create* incremental future income risk for young agents.

Agents are free to choose their desired levels of capital. Accordingly, a transfer plan τ *induces* a (detrended) capital allocation $(K_1, K_2, ..., K_J)$, where:

$$K_i = argmax_{k \geq 0} ln(e_i^y - k - \tau_i) + \beta \sum_{j=1}^{J} P_{ij} ln(e_i^o + (1-\delta)k + \bar{a}_i k^\alpha + (1+g)\tau_j), i = 1, 2, ..., J$$

The following simple proposition shows that a Pareto dominated transfer plan necessarily induces *overaccumulation* of capital in every positive probability state.

Proposition 1. *Let P be a Markov matrix with the unique stationary vector $\pi^{stat}(P)$, and let $S(P) = \{j | \pi_j^{stat}(P) > 0\}$ be the set of states that have positive probability given P. Suppose that given the transition probability matrix P, a transfer plan τ' Pareto dominates τ. Let K' be the capital allocation induced by τ' and K be the positive capital plan induced by τ. Then $K'_i < K_i$ for all $i \in S$.*

Definitionally, it is possible to improve upon a Pareto suboptimal transfer plan by tilting consumptions toward the old agents. But that tilt leads young agents to cut back on physical investment and to lower levels of capital.[11]

[11] Barbie, Hagedorn, and Kaul (2007, page 572) consider a different class of overlapping generations economies. In this class of models, they show that Proposition 1 is not valid: one allocation may be Pareto superior to another even though capital is lower in some date and state in the latter allocation.

4 Main Result

This section presents the main result. The theoretical theme is that a transfer plan is Pareto suboptimal as long as there is a sufficiently persistent state in which $(MPK - \delta) < g$. The main econometric result is then based on the observation that no *finite* sample can be used to rule out the possibility of such a state.

4.1 Economies without Risk

As mentioned in the introduction, an allocation is said to be dynamically inefficient if it is possible to reduce capital in some date and state without lowering *aggregate* consumption in any successor date and state. This subsection shows that, in a world without risk, this concept is equivalent to Pareto optimality in this class of models.

I first define and characterize dynamic efficiency in a world without risk. (In Section 5.2 and Online Appendix A, I extend this definition to a world with risk.)

Definition 2. Suppose $J = 1$ and that in this single state economy, a transfer plan τ^* induces the capital choice K^*. Let $C^* = (C_t^*)_{t=1}^{\infty}$ be the implied aggregate consumption path:

$$C_1^* = e^y(1+g) + e_{init}^o - K^*(1+g)$$

$$C_t^* = C^*(1+g)^t, t > 1$$

$$C^* = e^y + e^o/(1+g) + K^*(1-\delta)/(1+g) + \bar{a}K^{*\alpha}/(1+g) - K^*$$

Then τ^* is **dynamically inefficient** if there exists a consumption-capital sequence $(C_t', K_t')_{t=1}^{\infty}$

with higher period 1 consumption and no lower aggregate consumption thereafter:

$$C_1' > C_1^*$$
$$C_1' + K_1' = e^y(1+g) + e^o_{init}$$
$$C_t' \geq C_t^*, t \geq 2$$
$$C_t' + K_t' = e^y(1+g)^t + e^o(1+g)^{t-1} + K_{t-1}'(1-\delta) + \bar{a}(1+g)^{(t-1)(1-\alpha)}(K_{t-1}')^{\alpha}, t \geq 2.$$

The transfer plan τ^* is dynamically efficient if it is not dynamically inefficient.

Unlike Pareto (sub)optimality, dynamic (in)efficiency deliberately makes no reference to preferences. Rather, it is about the physical feasibility of a particular kind of allocational perturbation.

The following proposition provides sufficient conditions for dynamic inefficiency and dynamic efficiency in the one-state case:

Proposition 2. *Suppose $J = 1$ and consider a transfer plan τ^* that induces capital K^*. The transfer plan τ^* is dynamically inefficient if:*

$$\alpha\bar{a}K^{*\alpha-1} - \delta < g$$

It is dynamically efficient if:

$$\alpha\bar{a}K^{*\alpha-1} - \delta > g.$$

Proposition 3 shows that, without aggregate risk, the cutoff for dynamic (in)efficiency in Proposition 2 is also the cutoff for Pareto (sub)optimality.

Proposition 3. *Suppose $J = 1$. Consider a transfer plan τ with induced capital level $K_1 > 0$, and define $MPK_1 = \alpha\bar{a}_1K_1^{\alpha-1}$. If $(MPK_1 - \delta) < g$, then τ is Pareto suboptimal. If $(MPK_1 - \delta) > g$, then τ is Pareto optimal.*

If there is no risk, a transfer plan is simultaneously Pareto optimal and dynamically efficient when the induced marginal product of capital, net of depreciation, is no smaller than the growth rate. A transfer plan is simultaneously Pareto suboptimal and dynamically inefficient when the induced marginal product of capital, net of depreciation, is larger than the growth rate. We shall see in Section 5.2, though, that this connection between dynamic efficiency and Pareto optimality disappears once there is aggregate risk.

4.2 An Example

This subsection uses a two-state example to illustrate how the Pareto suboptimality of a transfer plan is shaped by the persistence of dynamic inefficiency.[12]

Suppose that the number of states J is equal to two. Let:

$$\begin{aligned}
\alpha, \delta &= 1 \\
\beta &= 1 \\
\bar{a}_1 &= 1 + \Delta, 1 > \Delta > 0 \\
\bar{a}_2 &= 1 - \Delta \\
e_i^o &= 0, i = 1, 2 \\
e_i^y &= 1, i = 1, 2 \\
g &= 0 \\
e_{init}^o &> 0
\end{aligned}$$

The results in the prior subsection imply that the zero transfer scheme would be Pareto optimal in state 1 if it were the only state and would be Pareto suboptimal in state 2 if it were the only state. We consider which kinds of positive transition probability matrices in the two-state world imply that the transfer scheme $(\tau_1, \tau_2) = (0, 0)$ is Pareto suboptimal.

With a zero transfer scheme, and given the linear technology, the young agents always

[12]Relatedly, Chen and Wen (2017) describe how an economy in which $(MPK - \delta)$ is initially negative may nonetheless be characterized by capital overaccumulation if there is a deterministic transition to a dynamically inefficient steady-state.

choose $K = 0.5$:

$$0.5 = argmax_{k \geq 0} ln(1-k) + \sum_{j=1}^{2} P_{ij} ln(k\bar{a}_i).$$

Hence, the transfer scheme $(0,0)$ is Pareto suboptimal if there exists positive (τ_1, τ_2) such that:

$$ln(0.5 - \tau_1) + P_{11} ln(0.5\bar{a}_1 + \tau_1) + P_{12} ln(0.5\bar{a}_1 + \tau_2) > 2ln(0.5) + ln(\bar{a}_1) \quad (3)$$

$$ln(0.5 - \tau_2) + P_{21} ln(0.5\bar{a}_2 + \tau_1) + P_{22} ln(0.5\bar{a}_2 + \tau_2) > 2ln(0.5) + ln(\bar{a}_2). \quad (4)$$

We know that a small positive vector (τ_1, τ_2) is consistent with (3)-(4) if it satisfies the marginal utility inequalities:

$$-\frac{1}{0.5}\tau_1 + \frac{P_{11}}{0.5\bar{a}_1}\tau_1 + \frac{P_{12}}{0.5\bar{a}_1}\tau_2 > 0 \quad (5)$$

$$-\frac{1}{0.5}\tau_2 + \frac{P_{21}}{0.5\bar{a}_2}\tau_1 + \frac{P_{22}}{0.5\bar{a}_2}\tau_2 > 0. \quad (6)$$

A positive specification of τ_1 reduces the utility of agents in (the dynamically efficient) state 1 because $P_{11}/\bar{a}_1 < 1$. To offset this fall in utility, (5) implies that τ_2/τ_1 must be large enough to satisfy:

$$\frac{\tau_2}{\tau_1} > \frac{\Delta}{P_{12}} + 1. \quad (7)$$

Note that it is possible to find such an interval of small (τ_1, τ_2) pairs for any $P_{12} > 0$.

How does a specification of (τ_1, τ_2) that satisfies (7) affect the welfare of agents in (the dynamically inefficient) state 2? The answer to this question depends on the persistence of state 2, as measured by the transition probability P_{22}. There are three cases.

Case 1 (Pareto suboptimality): the agents in state 2 are made better off with any positive (τ_1, τ_2) (including those that satisfy (5)) if $P_{22} \geq (1-\Delta)$. Note that this cutoff is independent of the persistence P_{11} of state 1.

Case 2 (Pareto suboptimality): even if $P_{22} < (1-\Delta)$, there exists a positive (τ_1, τ_2) that simultaneously makes agents in state 2 better off and satisfies (7) if P_{22} is so large that:

$$P_{22} > (P_{11} - \Delta).$$

Case 3: (Pareto optimality): There does not exist a positive (τ_1, τ_2) that makes agents in both states better off if:

$$P_{22} < (P_{11} - \Delta).$$

To sum up: in this example, the benchmark zero transfer scheme:

- is Pareto suboptimal if the dynamically inefficient state is sufficiently persistent, so that either $P_{22} \geq (1 - \Delta)$ or $(1 - \Delta) > P_{22} > (P_{11} - \Delta)$.
- is Pareto optimal if the dynamically inefficient state is sufficiently transitory relative to the dynamically efficient state:

$$\Delta < (P_{11} - P_{22}).$$

Intuitively, when the dynamically inefficient state is sufficiently persistent, agents in state 2 are made better off by an incremental increase in transfers in both states. Since the probability of transiting to state 2 is positive, agents in state 1 can be made better off by increasing transfers slightly in both states as long as the ratio of increases (state 2 to state 1) is sufficiently large.

4.3 General Theory ...

This subsection generalizes the analysis in the prior subsection to the class of models described in Section 3.

Define $(K_j^{SS})_{j=1}^J$ to be the *steady-state* capitals induced by a tax plan τ as the solution to:

$$\frac{1}{e_j^y - K_j^{SS} - \tau_j} = \beta \frac{(1-\delta) + MPK_j^{SS}}{e_j^o + \tau_j(1+g) + \bar{a}_j (K_j^{SS})^\alpha}, j = 1, .., J$$

where:

$$MPK_j^{SS} = \alpha \bar{a}_j (K_j^{SS})^{\alpha - 1}$$

is the steady-state marginal product of capital in state j. The following proposition considers a transfer plan that induces a *dynamically inefficient* capital allocation as a steady-state in state J (where the index is obviously picked arbitrarily), so that:

$$MPK_J^{SS} - \delta < g$$

The proposition shows that if state J is highly persistent, then - regardless of the structure of the rest of the positive transition probability matrix - τ is Pareto suboptimal.[13]

Proposition 4. *Consider a transfer plan τ that induces the positive steady-state capitals $(K_j^{SS})_{j=1}^J$. Suppose that the steady-state marginal product of capital induced by τ in state J satisfies:*

$$g > MPK_J^{SS} - \delta = \alpha \bar{a}_J (K_J^{SS})^{\alpha - 1} - \delta$$

Consider a sequence of positive transition probability matrices $\{P^m\}_{m=1}^\infty$ which satisfies $lim_{m \to \infty} P_{JJ}^m =$ 1. For an economy with transition probability matrix P^m, where m is sufficiently large, the transfer plan τ is Pareto suboptimal.

The logic behind the proof of this proposition is similar to that illustrated by the example in the prior subsection. The transfer τ_J induces a dynamically inefficient steady-state in state J. Hence, when P_{JJ}^m is near 1 (state J is highly persistent), the welfare of agents in state J is strictly increasing in τ_J. In contrast, an increase in $\tau_i, i < J$, could potentially lower the

[13] In this proposition, and in the remainder of the paper, the superscript on a matrix P represents an index, not an exponent.

welfare of young agents in state i. But, because the transition probability matrix is positive, it is possible to offset this decline by ensuring that the ratio (τ_J/τ_i) is sufficiently large.

4.4 ... And An Important Corollary

Proposition 4 relies on the relevant dynamically inefficient state being sufficiently persistent. The following corollary underscores that this restriction can be satisfied even if the economy spends most of its time in a dynamically efficient state.

Corollary 1. *Consider a transfer plan τ that induces the positive steady-state capitals $(K_j^{SS})_{i=1}^J$. Suppose that the steady-state marginal products of capital induced by τ in states 1 and J satisfy:*

$$\alpha\bar{a}_1(K_1^{SS})^{\alpha-1} - \delta = MPK_1^{SS} - \delta > g > MPK_J^{SS} - \delta = \alpha\bar{a}_J(K_J^{SS})^{\alpha-1} - \delta$$

Consider a sequence $\{P^m\}_{m=1}^{\infty}$ of positive transition probability matrices such that:

$$lim_{m\to\infty} P_{JJ}^m = 1$$
$$lim_{m\to\infty} \frac{(1-P_{11}^m)}{min_{j>1} P_{j1}^m} = 0$$

and let $\bar{\pi}^m = \pi^{stat}(P^m)$ be the stationary probability vector implied by P^m. Let k^m be the capital allocation induced by τ when the transition probability matrix is P^m. Then:

$$lim_{m\to\infty} \bar{\pi}_1^m = 1$$
$$lim_{m\to\infty} k_1^m = K_1^{SS}$$

and, if m is sufficiently large, τ is Pareto suboptimal in an economy with transition probability matrix P^m.

The conditions on $\{P^m\}_{m=1}^{\infty}$ are satisfied by many sequences, including:

$$
\begin{aligned}
P_{11}^m &= (1 - 1/m^2) \\
P_{1j}^m &= \frac{1}{m^2(J-1)}, j > 1 \\
P_{jj}^m &= 1 - \frac{1}{m}, j > 1 \\
P_{j1}^m &= \frac{1}{m(J-1)}, j > 1.
\end{aligned}
$$

The import of the Corollary is that, for large m, the stationary probability of the *dynamically efficient* state 1, in which $(MPK - \delta)$ exceeds g, is near 1. Nonetheless, because the conditional probability P_{JJ}^m is near 1 for large m, it is possible to construct a Pareto improvement, with an associated reduction in capital.

4.5 Main Result

The proceeding result (Corollary 1) suggests that an economy may look dynamically efficient, but still admit a Pareto improving reduction in capital accumulation. This subsection formalizes this observation by proving the main result of the paper: any statistical test of the null hypothesis of capital overaccumulation with size less than one has zero power against the alternative hypothesis of being in a dynamically efficient steady-state.

As in Corollary 1, we consider a version of the economy in which the transfer plan τ induces the positive steady-state capitals $(K_j^{SS})_{i=1}^J$. Suppose too that the steady-state marginal products of capital induced by τ in states 1 and J satisfy:

$$\alpha \bar{a}_1 (K_1^{SS})^{\alpha-1} - \delta = MPK_1^{SS} - \delta > g > MPK_J^{SS} - \delta = \alpha \bar{a}_J (K_J^{SS})^{\alpha-1} - \delta$$

Consider an econometrician who, from period 1 through period T, sees a sample of outputs and capitals:

$$\{a_t K_t^{\alpha}, K_t\}_{t=1}^T$$

It is assumed that the econometrician knows α, g, and $\{\bar{a}_1, ..., \bar{a}_J\}$.

However, the econometrician does not know the matrix P. The set of all possible $(J \times J)$ Markov matrices with a unique stationary probability vector is represented by Ω_J. The econometrician's null hypothesis is that of capital overaccumulation. Accordingly, we define H_0 to be the set of elements of Ω_J such that P implies τ is Pareto suboptimal.

The econometrician uses a statistical *test* of her null hypothesis. A test Υ is an open subset of the set $(\mathbb{R}_+^2)^T$ of possible samples, where the econometrician rejects the null hypothesis if her sample lies in Υ. A test Υ has a *power function* $\Gamma_\Upsilon : \Omega_J \to [0,1]$, where $\Gamma_\Upsilon(P)$ is the probability that the sample lies in Υ if the transition probability matrix is P. The *size* $\rho(\Upsilon)$ of the test is defined as:

$$\rho(\Upsilon) = sup_{P \in H_0} \Gamma_\Upsilon(P).$$

The size of the test is the (supremal) probability of Type I error.

The following proposition is the main impossibility theorem. It considers the Markov matrix P^{**} such that the dynamically efficient state 1 steady-state is known to occur at every date with probability 1, so that:

$$P_{j1}^{**} = 1, j = 1, ..., J.$$

$$P_{ji}^{**} = 0, j = 1, ..., J, i = 2, ..., J.$$

We know from Proposition 3 that P^{**} implies that τ is Pareto optimal, and so P^{**} is not in H_0.

Proposition 5. *Let* $d^* = (K_1^{SS}, \bar{a}_1(K^{SS})^\alpha)_{t=1}^T$ *be the sample with* T *observations of state 1 steady-state capital and output. Consider any test* Υ. *If* $d^* \in \Upsilon$, *then* $\rho(\Upsilon) = 1$. *If* $d^* \notin \Upsilon$, *then:*

$$\Gamma_\Upsilon(P^{**}) = 0$$

where $P^{**} \notin H_0$ *is defined as above.*

The proof is centered on the sample d^* in which the economy stays in the dynamically efficient state 1 for all observed periods. Corollary 1 implies that there is a transition probability matrix $\bar{P}$, with $\bar{P}_{JJ}$ near 1, so that samples close to d^* may have probability arbitrarily close to 1 even though the transfer scheme is Pareto suboptimal. It follows that if d^* leads to rejection, then the size of the test is 1. But under the Pareto optimal transition probability matrix P^{**}, the economy is always in state 1. So, if the sample d^* does not lead to rejection, the test has zero power[14] against P^{**}.

The following corollary to Proposition 5 considers the commonly used eyeball test described in Section 2. It shows that the size of this test is equal to one.

Corollary 2. *Consider a test Υ^* which contains all samples in which the marginal product of capital ($\alpha A_t K_t^{\alpha-1}$) exceeds $(g+\delta)$ in all periods. Then the size of Υ^* is one.*

Corollary 2 is trivial, as the sample d^* in Proposition 5 is an element of Υ^*.

4.6 Failure of a Partial Converse

The analysis in the prior subsection shows that the probability of $(MPK-\delta)$ being larger than g in all periods may be arbitrarily close to 1 even when the economy is Pareto suboptimal. As discussed in Section 2, this kind of dynamically efficient dataset is the empirically relevant case. But suppose hypothetically that $(MPK-\delta)$ were less than g in all periods. Could such a dataset have a high probability even when τ is Pareto optimal? Proposition 3 shows that the answer to this question is no.

To see why, recall that proposition considers a situation in which the transfer plan τ

[14] The logic of Proposition 5 resembles that in the classic papers of Dufour (1997) and Bahadur and Savage (1956). Proposition 5 can also be seen as a time-series illustration of the i.i.d. results of Canay, Santos, and Shaikh (CSS) (2013) and Romano (2004), which establish that a null hypothesis cannot be distinguished statistically from its boundary points when those are defined using the notion of total variation distance. In our Markov chain context, let $\mathcal{F}$ be the Borel sigma-algebra over $\mathbb{R}_+^{2T}$. Define μ^m and μ^{**} to be the probability measures over $\mathcal{F}$ implied by the Markov matrices P^m and P^{**}, for any $m \geq 1$. Define the ***total variation distance*** between the measures μ^m and μ^{**} to be $\psi_{TV}^m = sup_{A\in\mathcal{F}}|\mu^m(A) - \mu^{**}(A)|$. The proof of Proposition 5 establishes that $lim_{m\to\infty}\psi_{TV}^m = 0$. Hence, μ^{**} lies on the boundary of H_0 in the topology induced by the total variation distance. Proposition 5 can then be seen as a parallel to the findings of CSS (2013) and Romano (2004).

induces the positive steady-state capitals $(K_j^{SS})_{i=1}^J$, where the steady-state marginal product of capital induced by τ in state J is dynamically inefficient:

$$g > MPK_J^{SS} - \delta = \alpha \bar{a}_J (K_J^{SS})^{\alpha-1} - \delta.$$

Suppose an econometrician sees a sample of outputs and capitals:

$$\{a_t K_t^\alpha, K_t\}_{t=1}^T.$$

She can deduce the state in each period from this information. Suppose that the Markov transition probability matrix P is such that the probability of the dataset containing T consecutive observations of state J is larger than $(1-\varepsilon)$ for some ε. That can only be true if the transition probability P_{JJ} satisfies:

$$P_{JJ} \geq (1-\varepsilon)^{\frac{1}{T-1}}.$$

Proposition 3 then implies that, for ε positive but sufficiently close to zero, τ is necessarily Pareto suboptimal.

The key to this argument that τ is Pareto suboptimal for P_{JJ} sufficiently close to 1, *regardless of the specification of the rest of the transition probability matrix*. In contrast, Corollary 1 shows that for any $P_{11} < 1$ (where the steady-state in state 1 is dynamically efficient), the rest of the transition probability matrix may be such that τ is Pareto suboptimal.

5 Two Other Results

This section discusses the AMSZ "spout or sink" test and the disconnect between Pareto suboptimality and dynamic efficiency when there is aggregate risk.

5.1 "Spout or Sink" Test

AMSZ (1989) propose a variety of methods to test for dynamic efficiency and Pareto optimality. At the risk of oversimplifying, I focus on their "spout or sink" criterion. This method classifies the economy as being dynamically efficient/Pareto optimal if it invests less than the return to capital and as being dynamically inefficient/Pareto suboptimal if it invests more than the return to capital. Accordingly, consider the difference between the return to last period's investment and current investment. It can be written as:

$$\Delta_{t+1}^{net} \equiv \alpha a_t K_t^{\alpha} + (1-\delta)K_t - K_{t+1} \tag{8}$$

The first two terms capture the "spout" aspect of the investment sector, while the second captures its "sink" aspect.

The following result is essentially a re-statement of Proposition 5 for the AMSZ test. It considers a sequence of positive transition probability matrices $\{P^m\}_{m=1}^{\infty}$ that satisfies the same conditions as in Corollary 1. It shows that for m large, the unconditional probability that Δ_{t+1}^{net} is always positive is near 1, even though τ is Pareto suboptimal.

Proposition 6. *Suppose that the assumptions of Corollary 1 are satisfied. Define:*

$$\Delta_{t+1}^{net} = \alpha a_t K_t^{\alpha} + (1-\delta)K_t - K_{t+1}, t \geq 1$$

to be the investment sector's net contribution to output in period t. Let Pr_m represent the unconditional probability defined by the transition probability matrix P^m. Then, given any sample size T and $\varepsilon > 0$, there exists M_ε such that if $m \geq M_\varepsilon$, τ is Pareto suboptimal given P^m and:

$$Pr_m(T^{-1}\sum_{t=1}^{T} 1_{\Delta_{t+1}^{net}>0} = 1) \geq (1-\varepsilon)$$

where $1_{\Delta_t^{net}>0}$ is an indicator variable that equals one if $\Delta_t^{net} > 0$ and 0 otherwise.

For m large, the economy has a high probability of spending almost all of its time in state

1. The de-trended "spout-sink" $(1+g)^{-t}\Delta^{net}$ defined in (8) is close to $\hat{\Delta}^*(1,1)$ with high probability, where:

$$\bar{\Delta}^*(1,1) \approx MPK_1^{SS} K_1^{ss} + (1-\delta)K_1^{SS} - K_1^{ss}(1+g)$$

This is (possibly large and) positive, because $(MPK_1^{SS} - \delta) > g$. So, with a probability close to 1, the transfer plan passes a strong form of the AMSZ test, as the investment sector is always a spout. Nonetheless, as shown in Corollary 1, a Pareto improvement is possible.

How does the Pareto improvement in Proposition 6 work when m is large? As already noted, the economy spends most of its time in state 1. Since:

$$(MPK_1^{SS} - \delta) > g,$$

the Pareto improving reduction in capital leads, with high probability, to a decline in aggregate consumption. Indeed, with high probability, the realized utility of all agents falls. But, as in the proof of Proposition 3, it is still possible to improve the welfare of agents in state 1 by offering a sufficient increase in transfers in the (low-probability) state J.

5.2 Dynamic Efficiency and Pareto Optimality: A Disconnect

Section 4.1 provides a definition of dynamic inefficiency without risk. Online Appendix A generalizes this definition to allow for risk. Zilcha (1991) collapses the (rather complex) definition in Online Appendix A into a much simpler sufficient condition for dynamic efficiency. In our context, Zilcha's criterion implies that a transfer plan τ is dynamically efficient if it induces a (detrended) capital allocation K with the property:

$$\sum_{i=1}^{J} ln(1-\delta+\alpha\bar{a}_i K_i^{\alpha-1})\pi_i^{stat}(P) > ln(1+g) \tag{9}$$

This is an immediate generalization of the sufficient condition in Proposition 2 (for the case in which $J = 1$).

Proposition 3 showed that, when $J = 1$, any allocation that satisfies (9) is Pareto optimal. However, we can use the analysis from Section 4 to see that Proposition 3 does not generalize to the case in which $J > 1$: with risk, a dynamically efficient transfer plan may be Pareto suboptimal. In particular, suppose that a sequence $\{P^m\}_{m=1}^{\infty}$ of positive transition probability matrices satisfies the conditions of Corollary 1 and that, given P^m, τ induces a capital allocation $(K_1^m, ..., K_J^m)$. Then, for large values of m, the average (logged) MPK is well-approximated by:

$$\begin{aligned} lim_{m\to\infty} \sum_{i=1}^{J} \pi_i^{stat}(P^m)(ln(1-\delta+\alpha\bar{a}_i(K_i^m)^{\alpha-1}) \\ = ln(1-\delta+\alpha\bar{a}_1(K_1^{SS})^{\alpha-1}) \\ > ln(1+g). \end{aligned}$$

For large values of m, the economy spends almost all of its time in state 1. The transfer plan is dynamically efficient, because it satisfies (9), but is also Pareto suboptimal.

6 Bond Prices?

Kocherlakota (2023) considers a class of overlapping generations models without capital in which generational endowments evolve according to a Markov chain around a log-linear trend. He defines r_{long}^{aut} to be the shadow yield in autarky to a very long-term bond (the far right of the yield curve), and shows that it is a deterministic constant if the transition probability matrix is positive. That paper then proves that the autarkic allocation admits a Pareto-improving sequence of young-to-old transfers if $r_{long}^{aut} < ln(1+g)$, where g is the deterministic cross-cohort endowment growth rate.

Can a similar bond price-based test be used in stochastic overlapping generations models to check for the presence of capital overaccumulation? The answer is generally no. The challenge is that, as assumed by many modern macroeconomic models, the agents who are undertaking physical investment may be constrained by binding borrowing limits. As Reis (2021) emphasizes, these models imply that observed market interest rates (including r_{long}) lie below the corresponding shadow discount rates of capital owners, which are the relevant variable for assessing capital overaccumulation. In Online Appendix B, I provide an explicit example of how this disconnect means that capital may not be overacccumulated even if r_{long} is less than g.[15]

7 Conclusion

The message of this paper is simple: the US aggregate quantity data do not reject the null hypothesis that there is capital overaccumulation in the sense of Diamond (1965). As a consequence, it would seem important to perform systematic sensitivity/validation analyses in macroeconomic models to allow for the possibility of as yet unobserved aggregate productivity shocks that engender persistent dynamic inefficiency.

[15] The maximal eigenvalue tests recently proposed by Bloise and Reichlin (2023) are also subject to this same problem. Note too that market power could create a similar gap between r_{long} and $(MPK - \delta)$, as stressed by Ball and Mankiw (2023).

Appendix

This appendix contains all proofs.

Proof of Proposition 1

Since $K_i > 0$ for all i, it satisfies the first order conditions:

$$\frac{1}{e_i^y - K_i - \tau_i} = \beta(1 - \delta + \alpha \bar{a}_i K_i^{\alpha-1}) \sum_{j=1}^{J} P_{ij} \frac{1}{e_i^o + \bar{a}_i K_i^{\alpha} + \tau_j(1+g)}, i = 1, ..., J$$

Since $\tau_i' > \tau_i$ for $i \in S(P)$, we can conclude that:

$$\frac{1}{e_i^y - K_i - \tau_i'} > \beta(1 - \delta + \alpha \bar{a}_i K_i^{\alpha-1}) \sum_{j=1}^{J} P_{ij} \frac{1}{e_i^o + \bar{a}_i K_i^{\alpha} + \tau_j'(1+g)}, i \in S(P)$$

The left-hand side is strictly increasing in K_i and the right-hand side is strictly decreasing in K_i. So, $K_i' < K_i$, for all i in $S(P)$.

Proof of Proposition 2

Suppose $\alpha \bar{a}(K^*)^{\alpha-1} - \delta < g$. We can define an alternative primed allocation of aggregate capital and consumption by lowering capital at all dates:

$$K_t' = (K^* - \varepsilon)(1+g)^t, t \geq 1$$

$$C_1' = e^y(1+g) + e_{init}^o + \varepsilon(1+g)$$

$$C_t' = e^y(1+g)^t + e^o(1+g)^{t-1} + K_{t-1}'(1-\delta) + \bar{a}(1+g)^{(t-1)(1-\alpha)}(K_{t-1}')^{\alpha} - K_t'(1+g)^t, t \geq 2$$

Then:

$$C_t' - C^*(1+g)^t = \varepsilon(1+g)^t - (1-\delta)\varepsilon(1+g)^{t-1} - \bar{a}(1+g)^{t-1}((K^* - \varepsilon)^{\alpha} - (K^*)^{\alpha}).$$

Using the subgradient inequality, the RHS is bounded from below by:

$$\begin{aligned}\varepsilon(1+g)^{t-1}[(1+g)-(1-\delta)-\bar{a}\alpha(K^*)^{\alpha-1}]\\ > \varepsilon(1+g)^{t-1}(g-(\bar{a}\alpha(K^*)^{\alpha-1})\\ > 0.\end{aligned}$$

Hence, τ is dynamically inefficient.

Now let:

$$\alpha\bar{a}(K^*)^{\alpha-1}-\delta < g$$

Suppose by way of contradiction that τ^* is dynamically inefficient. There exists an alternative primed allocation such that:

$$\begin{aligned}&C_1' = C_1^* + \varepsilon(1+g)\\ &K_1' = K^*(1+g) - \varepsilon(1+g)\\ &C_t' = C_t^*, t \geq 2\\ &K_t' = K^*(1+g)^t + (K_{t-1}' - K^*(1+g)^{t-1})(1-\delta) + \bar{a}(1+g)^{(t-1)(1-\alpha)}((K_{t-1}')^\alpha - K^*(1+g)^{(t-1)\alpha}), t \geq 2.\end{aligned}$$

Define $\hat{K}_t' = K_t'(1+g)^{-t}$ and so:

$$\hat{K}_t' = K^* + (\hat{K}_{t-1}' - K^*)(1-\delta)/(1+g) + \bar{a}(\hat{K}_{t-1}^\alpha - K^{*\alpha})/(1+g).$$

or equivalently:

$$(\hat{K}_t' - K^*) = (1+g)^{-1}[(1-\delta)(\hat{K}_{t-1} - K^*) + \bar{a}(\hat{K}_{t-1}^\alpha - K^{*\alpha})]$$

The subgradient inequality implies that:

$$
\begin{aligned}
(\hat{K}_t' - K^*) &\leq (1+g)^{-1}((1-\delta) + \bar{a}\alpha K^{*\alpha-1})(\hat{K}_{t-1} - K^*) \\
&\leq [(1+g)^{-1}((1-\delta) + \bar{a}\alpha K^{*\alpha-1})]^{t-1}(\hat{K}_1 - K^*)
\end{aligned}
$$

But this means that $(\hat{K}_t' - K^*)$ is negative and its absolute value is growing exponentially, which implies that $(K_t')_{t=1}^{\infty}$ is eventually negative. That is a contradiction and so τ^* is dynamically efficient.

Proof of Proposition 3

Case 1: Suppose $\alpha\bar{a}_1 K_1^{\alpha-1} - \delta < g$. Consider $\tau_1' = \tau_1 + \gamma_1, \gamma_1 > 0$. Note that:

$$
\begin{aligned}
&\frac{(-\gamma_1)}{e_1^y - \tau_1 - K_1} + \frac{\beta\gamma_1(1+g)}{e_1^o + \tau_1(1+g) + (1-\delta)K_1 + \bar{a}_1 K_1^{\alpha}} \\
&> \frac{(-\gamma_1)}{e_1^y - \tau_1 - K_1} + \frac{\beta\gamma_1(1-\delta+\alpha\bar{a}_1 K_1^{\alpha-1})}{e_1^o + \tau_1(1+g) + (1-\delta)K_1 + \bar{a}_1 K_1^{\alpha}} \\
&= 0.
\end{aligned}
$$

This implies that for sufficiently small positive γ_1:

$$
\begin{aligned}
&ln(e_1^y - \tau_1 - \gamma_1 - K_1) + \beta ln(e_1^o + \tau_1(1+g) + \gamma_1(1+g) + (1-\delta)K_1 + \bar{a}_1 K_1^{\alpha})) \\
&> ln(e_1^y - \tau_1 - K_1) - \beta ln(e_1^o + \tau_1(1+g) + (1-\delta)K_1 + \bar{a}_1 K_1^{\alpha}
\end{aligned}
$$

which in turn implies that:

$$
V_1(\tau_1 + \gamma_1) > V_1(\tau_1)
$$

Hence, τ_1' Pareto dominates τ_1 and the latter transfer plan is Pareto suboptimal.

Case 2: $\alpha\bar{a}_1 K_1^{\alpha-1} - \delta \geq g$: Suppose τ_1' Pareto dominates τ_1. Then $\tau_1' > \tau_1$ and:

$$ln(e_1^y - \tau_1' - K_1') + \beta ln(e_1^o + \tau_1'(1+g) + (1-\delta)K_1 + \bar{a}_1(K_1')^\alpha)$$
$$> ln(e_1^y - \tau_1 - K_1) + \beta ln(e_1^o + \tau_1(1+g) + (1-\delta)K_1 + \bar{a}_1(K_1)^\alpha)$$

where the K_1's are optimally chosen given the relevant transfer plans. If (c^y, c^o, K) is the allocation induced by τ, then the subgradient inequality implies that:

$$\frac{(\tau_1 - \tau_1' + K_1 - K_1')}{c_1^y} + \beta\frac{((\tau_1' - \tau_1)(1+g) + (1-\delta+\bar{a}_1\alpha K_1^{\alpha-1})(K_1' - K_1))}{c_1^o} > 0$$

Since $\frac{1}{c_1^y} = \frac{\beta(1-\delta+\bar{a}_1\alpha K_1^{\alpha-1})}{c_1^o}$, this inequality can be rewritten as:

$$(\tau_1' - \tau_1)\frac{(1+g)}{(1-\delta+\bar{a}_1\alpha K_1^{\alpha-1})} > (\tau_1' - \tau_1)$$

But this is a contradiction since $\alpha\bar{a}_1 K_1^{\alpha-1} - \delta \geq g$.

Proof of Proposition 4

Suppose τ induces the consumption/capital allocation $(c_i^{m,y}, k_i^m, c_{ij}^{m,o})_{i,j=1}^J$ given the transition probability matrix P^m. This consumption allocation satisfies the first order conditions:

$$\frac{1}{c_i^{m,y}} = \beta(1-\delta+\alpha\bar{a}_i(k_i^m)^{\alpha-1})\sum_{j=1}^J P_{m,ij}\frac{1}{c_{ij}^{m,o}}, i = 1, 2, ..., J$$

Since $lim_{m\to\infty}P_{JJ}^m = 1$, we can conclude that:

$$lim_{m\to\infty}\alpha\bar{a}_J(k_J^m)^{\alpha-1} = MPK_J^{SS}, \quad (10)$$
$$lim_{m\to\infty}c_J^{m,y} = e_J^y - \tau_J - K_J^{SS} \quad (11)$$
$$lim_{m\to\infty}c_{Jj}^{o,y} = e_j^o + \tau_j(1+g) + (1-\delta+\alpha^{-1}MPK_J^{SS})K_J^{SS}, j = 1, 2, ..., J \quad (12)$$

Since $g > MPK_J^{SS} - \delta$, the limits (10) to (12) imply that:

$$\begin{aligned} & lim_{m\to\infty}\left(\frac{\beta P_{JJ}^m(1+g)}{c_{JJ}^{m,o}} - \frac{1}{c_J^{m,y}}\right) \\ > & lim_{m\to\infty}\left(\frac{\beta P_{JJ}^m(1+MPK_J^{SS}-\delta)}{c_{JJ}^{m,o}} - \frac{1}{c_J^{m,y}}\right) \\ = & \frac{\beta(1+MPK_J^{SS}-\delta)}{c_J^{SS,o}} - \frac{1}{c_J^{SS,y}} \\ = & 0 \end{aligned}$$

Hence, there exists M such that:

$$\frac{\beta P_{JJ}^m(1+g)}{c_{JJ}^{m,o}} - \frac{1}{c_J^{m,y}} > 0$$

for all $m > M$. Henceforth, we focus on $m > M$.

A sufficient condition for τ to be Pareto suboptimal in an economy with transition probability matrix P^m is the existence of a positive vector $(\gamma_j^m)_{j=1}^J$ of tax/transfer increments that increases welfare in all states:

$$ln(c_i^{m,y} - \gamma_i^m) + \beta\sum_{j=1}^J P_{ij}^m ln(c_{1j}^{m,o} + \gamma_j^m(1+g)) > ln(c_i^{m,y}) + \beta\sum_{j=1}^J P_{ij}^m ln(c_{ij}^{m,o}), i = 1, ..., J$$

(The condition is only sufficient because the left-hand side could be made even larger by allowing agents to re-optimize with respect to capital.) A small (in a Euclidean norm sense) positive vector $(\gamma_j^m)_{j=1}^J$ improves welfare in all states if it satisfies the marginal utility inequalities:

$$-\frac{1}{c_i^{m,y}}\gamma_i^m + \beta\sum_{j=1}^J P_{ij}^m \frac{\gamma_j^m(1+g)}{c_{ij}^{m,o}} > 0, i = 1, .., J-1 \tag{13}$$

$$-\frac{1}{c_J^{m,y}}\gamma_J^m + \beta\sum_{j=1}^J P_{Jj}^m \frac{\gamma_j^m(1+g)}{c_{Jj}^{m,o}} > 0 \tag{14}$$

Since $m > M$, we know that:

$$-\frac{1}{c_J^{m,y}} + \beta P_{m,JJ}\frac{(1+g)}{c_{JJ}^{m,o}} > 0.$$

It follows that any positive $(\gamma_j^m)_{j=1}^J$ satisfies (14).

Now, for each $m > M$, choose positive κ_m such that:

$$\kappa_m < \frac{\frac{\beta P_{jJ}^m(1+g)}{c_{jJ}^{m,o}}}{\frac{1}{c_j^{m,y}} - \beta(1+g)P_{jj}^m\frac{1}{c_{jj}^{m,o}}} \tag{15}$$

for all j in $\{1, 2, ..., J-1\}$. For each m, define a positive vector $\gamma^m = (\gamma_1^m, ..., \gamma_J^m)$ such that:

$$\frac{\gamma_j^m}{\gamma_J^m} < \kappa_m$$

for all j such that:

$$\frac{1}{c_j^{m,y}} - \beta(1+g)P_{m,jj}\frac{1}{c_{jj}^{m,o}} > 0.$$

Any such positive vector γ^m satisfies (13). (Basically, we are choosing each γ_j^m to be sufficiently small that it generates a small loss in welfare in state j relative to the gain in welfare in state j generated by γ_J^m.)

The ratio restriction (15) can be satisfied by positive $(\gamma_j^m)_{j=1}^J$ that are arbitrarily close to zero in a Euclidean norm sense. Hence, it follows that, for any $m > M$, there is a (set of) positive $(\gamma_1^m, ..., \gamma_J^m)$ sufficiently small such that $(\tau_j^{m\prime})_{j=1}^J = (\tau_j^m + \gamma_j^m)_{j=1}^J$ improves welfare in all states relative to $(\tau_j^m)_{j=1}^J$. This proves the proposition.

Proof of Corollary 1

The hypothesis of the Corollary requires that $lim_{m\to\infty} P_{JJ}^m = 1$. Hence, Proposition 4 implies that, when m is large, τ is Pareto suboptimal in an economy with transition probability matrix P^m.

Since $lim_{m\to\infty}P^m_{J1} = 0$, it follows that $lim_{m\to\infty}min_{j>1}P^m_{j1} = 0$. Hence, $lim_{m\to\infty}P^m_{11} = 1$ and:

$$lim_{m\to\infty}k^m_1 = K^{SS}_1.$$

Finally, the stationary probability vector $(\bar{\pi}^m_1, ..., \bar{\pi}^m_J) = \pi^{stat}(P^m)$ satisfies:

$$\bar{\pi}^m_1(1 - P^m_{11}) = \sum_{j\neq 1}^{J} \bar{\pi}^m_j P^m_{j1} \geq \sum_{j\neq 1}^{J} \bar{\pi}^m_j min_{j>1}P^m_{j1} = (min_{j>1}P^m_{j1})(1 - \bar{\pi}^m_1)$$

Then:

$$\begin{aligned}\bar{\pi}^m_1 &\geq \frac{min_{j>1}P^m_{j1}}{1 - P^m_{11} + min_{j>1}P^m_{j1}} \\ &= \frac{1}{\frac{(1-P^m_{11})}{min_{j>1}P^m_{j1}} + 1}.\end{aligned}$$

It follows that $lim_{m\to\infty}\bar{\pi}^m_1 = 1$.

Proof of Proposition 5

In this subsection, I prove Proposition 5 under the conditions set forth in Section 4.5.

Consider a sequence of positive transition probability matrices $\{\bar{P}^m\}_{m=1}^{\infty}$ which satisfies the properties delineated in Corollary 1:

$$lim_{m\to\infty}\bar{P}^m_{JJ} = 1$$
$$lim_{m\to\infty}\frac{(1 - \bar{P}^m_{11})}{min_{j>1}\bar{P}^m_{j1}} = 0$$

From Corollary 1, as m converges to infinity, the stationary probability vectors $\pi^{stat}(P^m)$ converge to a limit which puts all weight on state 1. From Proposition 4, there exists M such that τ is Pareto suboptimal given $\bar{P}^m$ for any $m \geq M$.

Suppose $d^* \in \Upsilon$. Let K^m be the capital allocation induced by τ given $\bar{P}^m$. The sequence

$\{K_1^m\}_{m=1}^{\infty}$ of state 1 capitals satisfies the first order condition:

$$\frac{1}{e_1^y - \tau_1 - K_1^m} = \beta(\bar{a}_1(K_1^m)^{\alpha-1} + (1-\delta)) \sum_{j=1}^{J} \bar{P}_{1j}^m \frac{1}{e_1^o + \tau_j + (1-\delta)K_1^m + \bar{a}_1(K_1^m)^\alpha}$$

and so converges to the state 1 steady-state:

$$lim_{m\to\infty} K_1^m = K_1^{SS}.$$

Let d^m be the sample consisting of T replications of $(\bar{a}_1(K_1^m)^\alpha, (K_1^m)^\alpha)$. Since Υ is open, and $lim_{m\to\infty} d^m = d^*$, there exists $M^* \geq M$ such that $d^m \in \Upsilon$ for all $m \geq M^*$. It follows that:

$$\begin{aligned} &sup_{P \in H_0} \Gamma_\Upsilon(P) \\ &\geq sup_{m \geq M^*} \Gamma_\Upsilon(\bar{P}^m) \\ &\geq sup_{m \geq M^*} \pi_1^{stat}(P^m)(P_{11}^m)^{T-1} \\ &= 1. \end{aligned}$$

The size of the test is equal to one. Now suppose $d^* \notin \Upsilon$. Then:

$$\Gamma_\Upsilon(P^{**}) = 0$$

where $P^{**} \notin H_0$ is defined immediately before the proposition. The proposition follows.

Proof of Proposition 6

Let $(K_1^m, ..., K_J^m)$ be the capital allocation induced by τ when the transition probability matrix is P^m. The variable $\hat{\Delta}_{t+1}^{net} = (1+g)^{-t} \Delta_{t+1}^{net}$ is a function of the lagged and current Markov states (s_t, s_{t+1}) :

$$\hat{\Delta}_{t+1}^{net} = \hat{\Delta}_m(s_t, s_{t+1}) \equiv \alpha \bar{a}_{s_t}(K_{s_t}^m)^\alpha + (1-\delta)K_{s_t}^m - K_{s_{t+1}}^m(1+g)$$

Given the assumptions in Corollary 1, there exists M such that:

$$\hat{\Delta}_m(1,1) > 0$$

for all $m > M$. Hence, for $m > M$, given any sample size T, the unconditional probability of the economy always being in a spout is bounded from below by the probability of the state always being 1:

$$Pr_m(T^{-1}\sum_{t=1}^{T} 1_{\Delta^{net}_{t+1}>0} = 1) \geq (\pi_1^{stat}(P^m))(P_{11}^m)^{T-1}.$$

Since the right-hand-side is arbitrarily close to 1 for large values of m, the proposition follows.

References

[1] Abel, Andrew, N. Gregory Mankiw, Lawrence Summers, and Richard Zeckhauser (1989): Assessing Dynamic Efficiency: Theory and Evidence, Review of Economic Studies 56, 1-20.

[2] Bahadur, R. R., and Leonard J. Savage (1956): The Non-Existence of Certain Statistical Procedures in Non-Parametric Problems, Annals of Mathematical Statistics 25, 1115-1122.

[3] Ball, Lawrence, and Mankiw, N. Gregory (2023): Market Power in Neoclassical Growth Models, Review of Economic Studies 90, 572-596.

[4] Barbie, Martin, Marcus Hagedorn, and Ashok Kaul (2004): Assessing Aggregate Tests of Efficiency for Dynamic Economies, Topics in Macroeconomics 4, Article 16.

[5] Barbie, Martin, Marcus Hagedorn, and Ashok Kaul (2007): On the Interaction Between Risk-Sharing and Capital Accumulation in a Stochastic OLG Model with Production, Journal of Economic Theory 137, 568-579.

[6] Bertocchi, Graziella (1991): Bubbles and Inefficiencies, Economic Letters 35, 117–122.

[7] Blanchard, Olivier (2019): Public Debt and Low Interest Rates, American Economic Review 109, 1197-1229.

[8] Bloise, Gaetano, Herakles Polemarchakis, and Yiannis Vailakis (2017): Sovereign Debt and Incentives to Default with Uninsurable Risks, Theoretical Economics 12, 1121-1154.

[9] Bloise, Gaetano, and Pietro Reichlin (2023): Low Safe Interest Rates: A Case for Dynamic Inefficiency?, SSRN 4244152.

[10] Canay, Ivan, Andres Santos, and Azeem Shaikh (2013): On the Testability of Identification in Some Nonparametric Models with Endogeneity, Econometrica 81, 2535-2559.

[11] Cass, David (1972): On Capital overaccumulation in the Aggregative, Neoclassical Model of Economic Growth: A Complete Characterization," Journal of Economic Theory 4, 200-223.

[12] Chen, Kaiji, and Yi Wen (2017): The Great Housing Boom of China, American Economic Journal: Macroeconomics 9, 73-114.

[13] Demange, Gabrielle, and Guy Laroque (1999): Social Security and Demographic Shocks, Econometrica 67, 527-542.

[14] Diamond, Peter (1965): National Debt in a Neoclassical Growth Model, American Economic Review 55, 1126-1150.

[15] Dufour, Jean-Marie (1997): Some Impossibility Theorems with Applications to Structural and Dynamic Models, Econometrica 65, 1365-1387.

[16] Hansen, Lars, and Jose Scheinkman (2009): Long-Term Risk: An Operator Approach, Econometrica 77, 177-234.

[17] Geerolf, Francois (2018): Reassessing Dynamic Efficiency, UCLA working paper.

[18] Kocherlakota, Narayana (2023): Infinite Debt Rollover in Stochastic Economies, Econometrica 91, 1629-1658.

[19] Martin, Ian, and Stephen Ross (2019): Notes on the Yield Curve, Journal of Financial Economics 134, 689-702.

[20] Muench, Thomas (1977): Optimality, the Interaction of Spot and Futures Markets, and the Nonneutrality of Money in the Lucas Model, Journal of Economic Theory 15, 325-344.

[21] Reis, Ricardo (2021): The Constraint on Public Debt When r < g But g < m, BIS working paper 939.

[22] Romano, Joseph (2004): On Non-Parametric Testing, the Uniform Behavior of the t-Test, and Related Problems, Scandinavian Journal of Statistics 31, 567-584.

[23] Romano, Joseph, Azeem Shaikh, and Michael Wolf (2010): Hypothesis Testing in Econometrics, Annual Reviews in Economics 2, 75-104.

[24] Zilcha, Itzhak (1991): Characterizing Efficiency in Stochastic Overlapping Generations Models, Journal of Economic Theory 55, 1-16.