When regulation kills growth

Summary for Davos

Javier Milei
Demian Reidel

January 21, 2026

Abstract

Policy debates routinely treat increasing returns to scale (IRS) as a standing warrant for structural intervention: declining average cost is read as natural monopoly, and natural monopoly is read as regulate-by-default. That conclusion is imported from toy partial-equilibrium environments that assume away endogenous labor supply, capital accumulation, and intertemporal feasibility. This paper restores those margins in a one-sector Ramsey–CES benchmark with an explicit returns-to-scale parameter. The steady-state problem becomes globally tractable: the Euler equation pins down an iso–user-cost locus $F_K(K, L) = D$, and stationary interior allocations reduce to zeros of a one-dimensional diagnostic along that locus. In the fold regime (IRS with strong complementarity), the Euler restriction implies a minimum feasible labor scale $L_{\min}$ and a folded feasibility set, generating steady-state multiplicity and poverty-trap dynamics driven by scale viability rather than monopoly pricing. The policy implication is direct: IRS are not a monopoly theorem, and state interventions that fragment scale or raise wedges can increase $L_{\min}$ and manufacture stagnation.

Keywords: increasing returns to scale; endogenous labor supply; nonconcave growth; multiple equilibria; poverty traps; antitrust
JEL: E13; E32; O41; L40

1 Introduction

Increasing returns to scale (IRS) are widely viewed as empirically relevant—from fixed costs and network effects to complementarities in production and organization—yet they are routinely set aside in benchmark macroeconomics. The standard reason is not that IRS are implausible, but that they are analytically inconvenient: once aggregate technology is nonconcave, familiar existence and uniqueness arguments break, steady states need not be unique, and the usual "one steady state, one set of comparative statics" logic can fail. Constant or decreasing returns (CRS/DRS) are therefore often imposed as a tractability device.

That tractability choice has had an unintended consequence for policy debate. In the absence of a disciplined general-equilibrium treatment of IRS, it is common to import a partial-equilibrium cost-curve story and treat it as a general policy conclusion. The resulting reflex—*IRS implies natural monopoly, therefore the state should intervene by default*—is not a theorem. It is an artifact of toy environments that assume away the very margins that determine whether scale is attainable, sustainable, and welfare-improving in the long run.

This paper develops a tractable IRS framework with endogenous labor supply and uses it to clarify what IRS do (and do not) imply for antitrust-style policy. The central message is blunt: *IRS are not a license for structural intervention.* In the environment we study, the primary implication of IRS is a *minimum viable scale* and potential *steady-state multiplicity*, including poverty-trap dynamics. Policies that fragment scale or raise wedges can move an economy toward (or even below) that minimum-viability boundary. In such a setting, "doing something" is not innocuous—it can be permanently destructive.

1.1 Why IRS is usually avoided

The benchmark growth model with CRS/DRS is attractive because it produces a globally concave planning problem (or an equivalent competitive equilibrium), which in turn delivers a clean characterization: a unique interior steady state (under standard assumptions), stable dynamics, and well-behaved comparative statics. By contrast, IRS correspond to technologies that are homogeneous of degree $\theta > 1$, which generically destroys global concavity. Once this happens, multiple stationary allocations and threshold dynamics become natural objects rather than pathologies.

The fact that IRS complicate the mathematics has encouraged an unfortunate conceptual shortcut: treating IRS as if they were synonymous with monopoly power. But IRS are a technological property; monopoly is a market-structure outcome. Conflating the two becomes especially misleading once labor supply and capital accumulation are endogenous. In a dynamic general-equilibrium setting, the key question is not "does a large producer exist?" but "can the economy reach and sustain the scale required to make high productivity feasible?"

1.2 The policy shortcut relies on toy partial-equilibrium models

A common policy argument proceeds as follows: IRS imply declining average costs; declining average costs imply subadditivity; subadditivity implies natural monopoly; therefore scale is presumptively harmful and the state should regulate or break up large-scale production. This chain of reasoning is rooted in a *toy partial-equilibrium* framing in which (i) factor supplies are fixed, (ii) intertemporal tradeoffs are absent, and (iii) feasibility constraints over scale are not central objects. With those margins stripped out, the only remaining issue is pricing, so "natural monopoly" appears to settle the welfare question by construction.

The problem is that the toy conclusion is not robust once the omitted margins are restored. When labor supply is endogenous, scale is chosen; when capital accumulates, today's scale affects tomorrow's productivity and feasibility; and when steady-state optimality must satisfy an Euler equation, the required marginal product of capital is pinned down by fundamentals, not by regulatory preference. In this environment, IRS do not mechanically translate into a monopoly distortion that calls for structural remedies. Instead, IRS reshape the *feasible set* and can generate a minimum-viability threshold. Treating size as presumptively suspect then risks the worst kind of policy error: remedies that reduce effective scale can push the economy toward the wrong side of a feasibility boundary.

1.3 Preview of the mechanism and results

We study a one-sector Ramsey economy with endogenous labor supply and a CES technology that is homogeneous of degree θ, with IRS corresponding to $\theta > 1$. Complementarity is captured by the CES parameter ξ; the empirically relevant case for our results is strong complementarity, $\xi < 0$ (elasticity of substitution below one).

The analysis turns on two simple steps that make the IRS case tractable. First, in steady state the Euler equation pins down a *user-cost wedge*

$$D \equiv \beta^{-1} - (1 - \delta) > 0,$$

and imposes the restriction

$$F_K(K, L) = D.$$

This defines an *iso–user-cost locus* in (K, L) space. Second, homogeneity allows a ratio representation with $r \equiv K/L$, so that along this locus the steady-state system collapses to a *one-dimensional diagnostic* $G(r)$: interior steady states correspond to roots of $G(r)$ subject to feasibility.

The main economic content comes from the "fold regime" defined by IRS and complementarity, $\theta > 1$ and $\xi < 0$. In that regime, the reduced marginal product of capital $\varphi_K(r)$ is hump-shaped in the capital–labor ratio. As a consequence, the Euler restriction implies a *minimum feasible labor scale*:

$$L_{\min} = \left(\frac{D}{\varphi_K^{\max}} \right)^{1/(\theta-1)},$$

where $\varphi_K^{\max}$ is the peak value of the reduced marginal product schedule. Below $L_{\min}$, the steady-state Euler requirement is infeasible: no capital–labor composition can deliver the required marginal product of capital.

This geometry produces three headline implications.

1. **Minimum viable scale.** The economy has an endogenous minimum labor scale $L_{\min}$ required to satisfy intertemporal optimality. Policy cannot repeal this constraint; it can only move the

economy relative to it.

2. **Multiplicity and poverty traps.** In the fold regime, the iso–user-cost locus has two branches for $L > L_{\min}$, and the diagnostic condition can intersect these branches multiple times. The resulting steady-state multiplicity supports poverty-trap dynamics: low scale depresses marginal products and discourages labor; low labor prevents reaching scale.
3. **The antitrust fallacy.** IRS are not a monopoly theorem. The central object is feasibility and coordination over scale, not a presumption of market-power abuse. Moreover, structural interventions that fragment scale or raise wedges can reduce $\varphi_K^{\max}$ or raise effective D, thereby increasing $L_{\min}$ and shrinking (or eliminating) the high-scale equilibrium. In an IRS economy, default intervention is therefore not "safe"—it can manufacture the very stagnation it claims to prevent.

1.4 Roadmap

Section 2 presents the model: preferences with endogenous labor supply, a CES technology with explicit returns-to-scale and complementarity parameters, and the equilibrium conditions, including the user-cost wedge D. Section 3 develops the steady-state reduction: the Euler restriction defines the iso–user-cost locus, and homogeneity yields a one-dimensional diagnostic whose roots correspond to interior steady states. Section 4 analyzes the fold regime ($\theta > 1, \xi < 0$), deriving the minimum feasible labor scale $L_{\min}$, the two-branch geometry, and the resulting multiplicity and poverty-trap mechanism. Section 5 draws the policy implications: IRS do not justify antitrust by default, and scale-fragmenting intervention can backfire by pushing the economy toward (or below) minimum viable scale. Section 6 concludes.

2 Model

This section lays out a minimal one-sector Ramsey economy with endogenous labor supply and a CES technology that allows for increasing returns to scale. We formulate the problem in terms of a representative planner. This keeps the analysis focused on feasibility and intertemporal optimality—the objects that matter for scale and long-run outcomes—without importing any monopoly pricing assumptions.

2.1 Preferences

Time is discrete, $t = 0, 1, 2, \ldots$. A representative household values consumption C_t and leisure $1 - L_t$, where labor satisfies $L_t \in (0, 1)$. Preferences are

$$\sum_{t=0}^{\infty} \beta^t \Big[\eta \log C_t + (1-\eta)\log(1-L_t)\Big], \qquad \beta \in (0,1),\ \eta \in (0,1). \tag{2.1}$$

The logarithmic specification provides two boundary disciplines that are useful later: marginal utility of consumption diverges as $C_t \downarrow 0$, and marginal disutility of labor diverges as $L_t \uparrow 1$ (i.e., as leisure vanishes). These properties prevent corner "solutions by decree" and make feasibility constraints economically binding rather than cosmetic.

2.2 Technology

Output is produced using capital $K_t \geq 0$ and labor $L_t \in (0,1)$ according to a CES technology with an explicit returns-to-scale parameter:

$$Y_t = F(K_t, L_t) \;\equiv\; A\Big(\omega K_t^{\xi} + (1-\omega)L_t^{\xi}\Big)^{\theta/\xi}, \qquad A > 0,\ \omega \in (0,1),\ \xi \in \mathbb{R},\ \theta > 0. \tag{2.2}$$

Returns to scale are governed by θ: CRS corresponds to $\theta = 1$, DRS to $\theta < 1$, and IRS to $\theta > 1$. The parameter ξ governs substitution; when $\xi < 0$, the elasticity of substitution is below one, capturing strong complementarity between capital and labor.

Let

$$Z(K,L) \equiv \omega K^{\xi} + (1-\omega)L^{\xi}.$$

Then $F(K,L) = AZ(K,L)^{\theta/\xi}$, and the marginal products are

$$F_K(K,L) = A\theta\,\omega\, Z(K,L)^{\theta/\xi-1}\, K^{\xi-1}, \tag{2.3}$$

$$F_L(K,L) = A\theta\,(1-\omega)\, Z(K,L)^{\theta/\xi-1}\, L^{\xi-1}. \tag{2.4}$$

By construction, F is homogeneous of degree θ, so Euler's theorem implies

$$F_K(K,L)\,K + F_L(K,L)\,L = \theta\, F(K,L). \tag{2.5}$$

Under IRS ($\theta > 1$), factor payments at marginal products do not generally exhaust output; this accounting fact is separate from, and not the source of, the scale-feasibility mechanism emphasized below.

2.3 Feasibility and equilibrium conditions

Capital depreciates at rate $\delta \in (0,1]$, with given initial stock $K_0 > 0$. The resource constraint is

$$C_t + K_{t+1} = (1-\delta)K_t + F(K_t, L_t), \qquad t \geq 0, \tag{2.6}$$

with $C_t \geq 0$ and $K_{t+1} \geq 0$.

A planner chooses $\{C_t, L_t, K_{t+1}\}_{t\geq 0}$ to maximize (2.1) subject to (2.6). For interior allocations, the first-order conditions can be written in the familiar intratemporal and intertemporal forms. The labor-leisure condition is

$$\frac{1-\eta}{\eta}\frac{C_t}{1-L_t} = F_L(K_t, L_t), \tag{2.7}$$

and the Euler equation for capital is

$$\frac{1}{C_t} = \beta\,\frac{1}{C_{t+1}}\Big(F_K(K_{t+1}, L_{t+1}) + 1 - \delta\Big). \tag{2.8}$$

We impose the standard transversality condition on optimal plans.

A central object for the steady-state analysis is the user-cost wedge

$$D \;\equiv\; \beta^{-1} - (1-\delta) \;>\; 0. \tag{2.9}$$

In any stationary interior allocation (C, K, L), (2.8) collapses to

$$F_K(K, L) = D. \tag{2.10}$$

This restriction is not a modeling convenience: it is the intertemporal optimality requirement that ties long-run feasibility to fundamentals (β, δ), and it will anchor the diagnostic characterization in Section 3.

3 Steady-State Characterization via a One-Dimensional Diagnostic

This section provides a global characterization of *stationary interior allocations* in the Ramsey–CES environment. The key step is to organize the steady-state system around a single measurable wedge—the user cost D—and to collapse the steady-state problem to zeros of a one-dimensional diagnostic evaluated along the iso–user-cost locus. This reduction is precisely what toy partial-equilibrium arguments miss: in general equilibrium, long-run feasibility must satisfy the Euler restriction $F_K = D$, which is pinned down by fundamentals and cannot be legislated away.

3.1 Steady-state system

A *stationary interior allocation* is a triple (C, K, L) with

$$C_t = C, \qquad K_t = K, \qquad L_t = L, \qquad K_{t+1} = K,$$

satisfying $C > 0$, $K > 0$, and $L \in (0, 1)$.

In steady state, the resource constraint (2.6) becomes

$$C = F(K, L) - \delta K. \tag{3.1}$$

The intratemporal optimality condition (2.7) becomes

$$\frac{1-\eta}{\eta}\frac{C}{1-L} = F_L(K, L). \tag{3.2}$$

Finally, the Euler equation (2.8) collapses to a restriction on the marginal product of capital, derived next.

3.2 Euler restriction as an iso–user-cost locus

Recall the user-cost wedge defined in (2.9),

$$D \equiv \beta^{-1} - (1-\delta) > 0.$$

In any stationary interior allocation, (2.8) implies

$$F_K(K, L) = D. \tag{3.3}$$

Equation (3.3) defines an *iso–user-cost locus* in (K, L) space:

$$\mathcal{I}(D) \;\equiv\; \{(K, L) \in (0, \infty) \times (0, 1) :\; F_K(K, L) = D\}.$$

In our application, this locus is the backbone object. It pins down which (K, L) pairs are even candidates for a steady state: if (K, L) does not satisfy $F_K = D$, it cannot be stationary regardless of any policy narrative about "desired" scale.

3.3 Ratio representation

Define the capital–labor ratio

$$r \equiv \frac{K}{L}, \tag{3.4}$$

and the associated CES aggregator in ratio form

$$Z(r) \equiv \omega r^{\xi} + 1 - \omega. \tag{3.5}$$

Using $K = rL$, the production function (2.2) can be written as

$$F(K, L) = L^{\theta}\, y(r), \qquad y(r) \equiv A\, Z(r)^{\theta/\xi}. \tag{3.6}$$

Similarly, the marginal products (2.3)–(2.4) admit the scale–ratio decomposition

$$F_K(K, L) = L^{\theta-1}\, \varphi_K(r), \qquad F_L(K, L) = L^{\theta-1}\, \varphi_L(r), \tag{3.7}$$

where

$$\varphi_K(r) \equiv A\theta\omega\, Z(r)^{\theta/\xi-1}\, r^{\xi-1}, \tag{3.8}$$
$$\varphi_L(r) \equiv A\theta(1-\omega)\, Z(r)^{\theta/\xi-1}. \tag{3.9}$$

This representation cleanly separates *scale* (powers of L) from *composition* (functions of r). It is the device that makes the global steady-state problem tractable.

3.4 Parametrization along $F_K = D$

Combining (3.3) with (3.7) yields

$$L^{\theta-1}\varphi_K(r) = D. \tag{3.10}$$

Throughout the paper we focus on IRS, $\theta > 1$, in which case (3.10) can be solved for labor as a function of r:

$$L(r) = \left(\frac{D}{\varphi_K(r)}\right)^{1/(\theta-1)}. \tag{3.11}$$

Associated capital is then

$$K(r) = r\,L(r). \tag{3.12}$$

Substituting into (3.1), consumption along the iso–F_K locus is

$$C(r) = F(K(r), L(r)) - \delta K(r) = L(r)^{\theta}\, y(r) - \delta\, r\, L(r). \tag{3.13}$$

Equations (4.13)–(4.14) provide a single-valued parametrization of the candidate steady-state set, even when the mapping from L to r is multi-valued (which is exactly what will happen in the fold regime).

3.5 Diagnostic function

Define the *steady-state diagnostic* along the iso–F_K locus by

$$G(r) \;\equiv\; \frac{1-\eta}{\eta}\frac{C(r)}{1-L(r)} - F_L(K(r), L(r)) \;=\; \frac{1-\eta}{\eta}\frac{C(r)}{1-L(r)} - L(r)^{\theta-1}\varphi_L(r), \tag{3.14}$$

where the second equality uses (3.7) and (4.6).

Proposition 3.1 (Steady states as zeros of a one-dimensional diagnostic)**.** *Suppose $\theta > 1$ and let D be defined by* (2.9)*. A triple (C, K, L) with $C > 0$, $K > 0$, $L \in (0, 1)$ is a stationary interior allocation satisfying the steady-state first-order conditions* (3.1)–(3.3) *if and only if there exists $r > 0$ such that*

1. *$L(r) \in (0, 1)$ and $C(r) > 0$ as defined in* (4.13) *and* (4.14)*;*
2. *$G(r) = 0$ as defined in* (3.14)*;*
3. *$(K, L, C) = (K(r), L(r), C(r))$ with $K(r)$ given by* (3.12)*.*

Proof. ($\Rightarrow$) Let (C, K, L) be a stationary interior allocation satisfying (3.1)–(3.3). Define $r \equiv K/L > 0$. By (3.3) and the ratio form (3.7), we have $L^{\theta-1}\varphi_K(r) = D$, which implies $L = L(r)$ as in (4.13), and hence $K = K(r)$ as in (3.12). By the resource identity (3.1), $C = C(r)$ as in (4.14). Finally, the intratemporal condition (3.2) is equivalent to $G(r) = 0$ by (3.14). The feasibility restrictions $C > 0$ and $L \in (0, 1)$ translate directly into item (1).

($\Leftarrow$) Conversely, suppose there exists $r > 0$ such that $L(r) \in (0, 1)$, $C(r) > 0$, and $G(r) = 0$. Set $(K, L, C) = (K(r), L(r), C(r))$. By construction, (3.10) holds, hence $F_K(K, L) = D$, i.e. (3.3). The definition of $C(r)$ implies (3.1). Finally, $G(r) = 0$ is exactly the intratemporal condition (3.2). Thus (C, K, L) is a stationary interior allocation satisfying the steady-state first-order conditions. □

Proposition 3.1 is the main tractability result used throughout the paper: rather than solving a three-equation nonlinear steady-state system in (C, K, L), we can trace the iso–F_K locus via (4.13)–(3.12) and locate stationary interior allocations by finding zeros of the scalar function $G(r)$ subject to the feasibility filters $L(r) \in (0, 1)$ and $C(r) > 0$.

4 The Fold Regime: Minimum Viable Scale, Multiplicity, Poverty Traps

This section analyzes the steady-state geometry implied by increasing returns to scale and strong complementarity. The analysis is entirely driven by feasibility and intertemporal optimality: in steady state, the Euler equation imposes a required marginal product of capital D, and under IRS the mapping from scale to marginal products becomes nonlinear. The core result is a *fold* in the iso–user-cost locus $F_K(K, L) = D$, which implies an endogenous minimum viable labor scale $L_{\min}$ and generically permits multiple interior steady states. The economic content is straightforward: below a minimum scale, the steady state is not merely undesirable—it is *infeasible*.

Throughout, we focus on interior stationary allocations that satisfy the planner's first-order conditions (Section 2); with IRS the problem is generally nonconcave, so these conditions need not guarantee global optimality. Nevertheless, they provide the correct object for understanding whether high-scale outcomes are feasible at all, and how thresholds arise.

4.1 Fold regime definition

We define the *fold regime* by two restrictions:

$$\theta > 1 \qquad \text{and} \qquad \xi < 0, \tag{4.1}$$

with $A > 0$ and $\omega \in (0, 1)$ as in (2.2). The first inequality imposes increasing returns to scale. The second imposes strong complementarity (elasticity of substitution below one), which makes capital and labor jointly necessary for high marginal products.

The fold regime matters because the steady-state Euler restriction pins down the required marginal product of capital,

$$F_K(K, L) = D \equiv \beta^{-1} - (1 - \delta),$$

and under (4.1) the marginal product of capital is not monotone in the capital–labor ratio. The result is a minimum scale requirement and a two-branch iso–F_K geometry.

4.2 Hump-shaped reduced MPK

Introduce the capital–labor ratio $r \equiv K/L$ and define

$$Z(r) \equiv \omega r^{\xi} + (1 - \omega). \tag{4.2}$$

By homogeneity of degree θ, production and marginal products can be written in scale–ratio form:

$$F(K, L) = L^{\theta} y(r), \qquad F_K(K, L) = L^{\theta-1}\varphi_K(r), \qquad F_L(K, L) = L^{\theta-1}\varphi_L(r), \tag{4.3}$$

where

$$y(r) \equiv A\,Z(r)^{\theta/\xi}, \tag{4.4}$$

$$\varphi_K(r) \equiv A\theta\omega\,Z(r)^{\theta/\xi-1}\,r^{\xi-1}, \tag{4.5}$$

$$\varphi_L(r) \equiv A\theta(1-\omega)\,Z(r)^{\theta/\xi-1}. \tag{4.6}$$

The function $\varphi_K(r)$ is the *reduced* (scale-free) marginal product of capital: scale enters F_K only through the factor $L^{\theta-1}$.

Lemma 4.1 (Boundary behavior of reduced MPK)**.** *Under* (4.1)*,* $\varphi_K(r) \to 0$ *as* $r \downarrow 0$ *and as* $r \uparrow \infty$*.*

Proof. As $r \downarrow 0$ with $\xi < 0$, we have $r^\xi \to \infty$ and hence $Z(r) \sim \omega r^\xi$. Using (4.5),

$$\varphi_K(r) \sim A\theta\omega\big(\omega r^\xi\big)^{\theta/\xi-1} r^{\xi-1} = A\theta\omega\,\omega^{\theta/\xi-1}\,r^{\xi(\theta/\xi-1)}r^{\xi-1}.$$

Since $\xi(\theta/\xi-1) = \theta-\xi$, the exponent on r is $(\theta-\xi)+(\xi-1) = \theta-1 > 0$. Thus $\varphi_K(r) \to 0$ as $r \downarrow 0$.

As $r \uparrow \infty$ with $\xi < 0$, $r^\xi \to 0$ and $Z(r) \to 1-\omega \in (0,1)$. Then

$$\varphi_K(r) \sim A\theta\omega(1-\omega)^{\theta/\xi-1}r^{\xi-1},$$

and since $\xi - 1 < 0$, $\varphi_K(r) \to 0$ as $r \uparrow \infty$. □

Lemma 4.1 implies that φ_K must attain an interior maximum. In fact, the maximizer is unique and admits a closed form.

Lemma 4.2 (Unique maximizer of reduced MPK)**.** *Under* (4.1)*,* $\varphi_K(r)$ *has a unique maximizer* $r^\star \in (0,\infty)$ *given by*

$$(r^\star)^\xi = \frac{(1-\xi)(1-\omega)}{\omega(\theta-1)}. \tag{4.7}$$

Define $\varphi_K^{\max} \equiv \varphi_K(r^\star)$*.*

Proof. Differentiate $\log\varphi_K(r)$ using (4.5) and (4.2). For $r > 0$,

$$\frac{d}{dr}\log\varphi_K(r) = \Big(\frac{\theta}{\xi}-1\Big)\frac{Z'(r)}{Z(r)} + \frac{\xi-1}{r}, \qquad Z'(r) = \omega\xi r^{\xi-1}.$$

Setting the derivative to zero and simplifying yields

$$\omega(\theta-1)r^\xi = (1-\xi)(1-\omega),$$

which is (4.7). Because $r \mapsto r^\xi$ is strictly monotone for $\xi \neq 0$ (strictly decreasing when $\xi < 0$), (4.7) has a unique solution. By Lemma 4.1, $\varphi_K(r) \to 0$ at both boundaries, so the unique stationary point is the unique global maximizer. □

4.3 Minimum feasible labor scale

In any stationary interior allocation, the Euler restriction (2.10) requires

$$F_K(K, L) = L^{\theta-1}\varphi_K\left(\frac{K}{L}\right) = D. \tag{4.8}$$

Fix any labor scale $L \in (0, 1)$. The left-hand side is maximized over the ratio r at $r^\star$, yielding

$$\max_{r>0} F_K(rL, L) = L^{\theta-1}\varphi_K^{\max}. \tag{4.9}$$

Since (4.8) must hold exactly in steady state, feasibility requires D not exceed the maximal attainable MPK at scale L.

Proposition 4.3 (Minimum viable labor scale)**.** *Suppose* (4.1) *holds and define* $\varphi_K^{\max}$ *as in Lemma 4.2. Any stationary interior allocation satisfying the Euler restriction* (2.10) *must satisfy*

$$L \geq L_{\min} \;\equiv\; \left(\frac{D}{\varphi_K^{\max}}\right)^{1/(\theta-1)}. \tag{4.10}$$

If $L_{\min} \geq 1$*, then no stationary interior allocation can satisfy* (2.10)*. If* $L_{\min} < 1$*, then the Euler restriction is feasible only for labor scales* $L \in [L_{\min}, 1)$*.*

Proof. From (4.9), for any (K, L),

$$F_K(K, L) = L^{\theta-1}\varphi_K(K/L) \leq L^{\theta-1}\varphi_K^{\max}.$$

If $F_K(K, L) = D$, then necessarily $D \leq L^{\theta-1}\varphi_K^{\max}$, which is equivalent to $L \geq (D/\varphi_K^{\max})^{1/(\theta-1)}$. The remaining statements follow immediately. □

Equation (4.10) is the key feasibility object. Below $L_{\min}$, the required steady-state MPK cannot be achieved by *any* rearrangement of the capital–labor ratio. This is not a market failure and not a pricing issue. It is a hard constraint implied by intertemporal optimality and the technology.

4.4 Two-branch structure and fold geometry

Assume $L_{\min} < 1$ so that the Euler restriction can be satisfied for some interior labor scales. Consider the iso–user-cost locus defined by (4.8). Using (4.3), the locus solves

$$\varphi_K(r) = \frac{D}{L^{\theta-1}}. \tag{4.11}$$

For a fixed L, the right-hand side is a constant. Because φ_K is continuous, hump-shaped, and attains its unique maximum $\varphi_K^{\max}$ at $r^\star$, the equation (4.11) has a two-solution structure whenever the constant lies strictly below $\varphi_K^{\max}$.

Proposition 4.4 (Fold and two branches)**.** *Maintain* (4.1) *and suppose* $L_{\min} < 1$*. Then:*

1. *For $L = L_{\min}$, equation (4.11) has the unique solution $r = r^\star$.*
2. *For any $L \in (L_{\min}, 1)$, equation (4.11) has exactly two solutions*

$$r_-(L) \; < \; r^\star \; < \; r_+(L),$$

with $r_-(L)$ on the "left" side of the hump and $r_+(L)$ on the "right" side.

Consequently, the iso–user-cost locus $F_K(K, L) = D$ is a folded curve in (K, L) space, with fold point

$$(K_{\min}, L_{\min}) \equiv (r^\star L_{\min},\, L_{\min}). \tag{4.12}$$

Proof. For $L = L_{\min}$, we have $D/L^{\theta-1} = \varphi_K^{\max}$ by (4.10), so (4.11) holds if and only if $\varphi_K(r) = \varphi_K^{\max}$, which occurs uniquely at $r^\star$ (Lemma 4.2).

For $L \in (L_{\min}, 1)$, $D/L^{\theta-1} \in (0, \varphi_K^{\max})$. Since φ_K is continuous, strictly increasing on $(0, r^\star)$ and strictly decreasing on $(r^\star, \infty)$ (implied by the uniqueness of the maximizer and boundary limits), the level set $\varphi_K(r) = D/L^{\theta-1}$ intersects each side exactly once, yielding exactly two solutions $r_-(L) < r^\star < r_+(L)$. The fold point follows from $K = rL$. □

Economically, Proposition 4.4 says that once scale exceeds the minimum viable level $L_{\min}$, there are two distinct capital–labor *compositions* consistent with the same required MPK. One branch corresponds to "too little capital per unit labor" ($r_-(L)$), the other to "too much" ($r_+(L)$). The fold arises because under complementarity, MPK collapses at both extremes: either capital is starved by labor, or labor is starved by capital.

4.5 Multiplicity of steady states

Section 3 shows that steady states can be characterized by a scalar diagnostic $G(r)$ evaluated along the iso–user-cost locus. The fold geometry implies that the domain of r consistent with interior labor is typically a compact interval containing $r^\star$, split into two branches. This alone does not mechanically force multiplicity, but it makes it the generic outcome: the intratemporal condition can intersect the folded locus on one branch, the other branch, or both. When it intersects both, the economy has multiple interior steady states.

To make this transparent, define labor along the locus by solving (4.8) for L:

$$L(r) \equiv \left(\frac{D}{\varphi_K(r)} \right)^{1/(\theta-1)}. \tag{4.13}$$

By Lemma 4.2, $L(r)$ is well-defined for $r > 0$ and attains its minimum at $r^\star$, with $L(r^\star) = L_{\min}$. Moreover, under $D < \varphi_K^{\max}$ there exist two values $r_1 < r^\star < r_2$ solving $\varphi_K(r) = D$, and hence $L(r_1) = L(r_2) = 1$. The interior-feasible r-domain corresponds to $r \in [r_1, r_2]$ (subject also to $C(r) > 0$).

Let consumption along the locus be

$$C(r) \equiv F(K(r), L(r)) - \delta K(r), \qquad K(r) \equiv rL(r). \tag{4.14}$$

Define the diagnostic

$$G(r) \equiv \frac{1-\eta}{\eta}\frac{C(r)}{1-L(r)} - F_L(K(r), L(r)). \tag{4.15}$$

Steady states correspond to roots of $G(r) = 0$ with $L(r) \in (0,1)$ and $C(r) > 0$.

The next observation provides the basic sign behavior that drives multiplicity.

Lemma 4.5 (Diagnostic explodes as $L(r) \uparrow 1$)**.** *Suppose $C(r) > 0$ and $L(r) \uparrow 1$ along the iso–user-cost locus. Then $G(r) \to +\infty$.*

Proof. As $L(r) \uparrow 1$, the term $\frac{C(r)}{1-L(r)}$ diverges to $+\infty$ whenever $C(r)$ remains strictly positive. Meanwhile $F_L(K(r), L(r))$ is finite for interior $K(r)$ and $L(r) \le 1$. Hence $G(r) \to +\infty$. □

Lemma 4.5 says that near the upper labor boundary, the intratemporal condition (2.7) cannot hold unless the economy is exactly at a point where the MRS and MPL match; otherwise the MRS term dominates. The fold then makes it easy for G to cross zero separately on each branch.

Proposition 4.6 (Two steady states in the fold regime)**.** *Maintain* (4.1)*, assume $L_{\min} < 1$, and suppose $C(r) > 0$ for all $r \in [r_1, r_2]$ where $L(r) \le 1$. If*

$$G(r^\star) < 0, \tag{4.16}$$

then there exist at least two interior steady states: one with $r \in (r_1, r^\star)$ and one with $r \in (r^\star, r_2)$.

Proof. Under the stated conditions, G is continuous on each branch $(r_1, r^\star]$ and $[r^\star, r_2)$. By Lemma 4.5, $\lim_{r \downarrow r_1} G(r) = +\infty$ and $\lim_{r \uparrow r_2} G(r) = +\infty$ because $L(r) \uparrow 1$ at r_1 and r_2. If $G(r^\star) < 0$, then by the intermediate value theorem there exists at least one root of G on $(r_1, r^\star)$ and at least one root on $(r^\star, r_2)$. □

Condition (4.16) is economically mild: at the minimum-viable scale, MPL is typically high relative to the household's MRS (the economy "wants to expand"), so the intratemporal equality fails on the low-scale boundary and must be restored at higher labor on each branch. The essential point is that the fold creates two distinct feasible compositions consistent with the same required MPK, so the labor-supply condition can hold at more than one point.

4.6 Poverty trap mechanism

The fold regime provides a direct poverty-trap logic driven by feasibility, not by market power. Proposition 4.3 establishes a minimum viable labor scale: if labor is below $L_{\min}$, the steady-state Euler requirement $F_K = D$ cannot be satisfied. When labor is just above $L_{\min}$, the economy operates near the fold point where the iso–user-cost locus is most fragile: small reductions in effective scale or small increases in the required MPK can render high-productivity allocations unattainable.

The feedback mechanism is immediate. Near the low-scale region:

1. With IRS, output and marginal products are low at small scale because scale itself is a productivity amplifier.

2. Low marginal products reduce the return to working (and to accumulating capital), depressing labor supply and investment incentives.
3. Lower labor supply reduces scale further, pushing the economy back toward the minimum-viability boundary.

This is the poverty trap: *low scale sustains low scale.* In a nonconcave environment, steady states can have distinct basins of attraction. A temporary shock (or a policy wedge) that reduces effective scale can move the economy across the boundary separating the basin of a high-scale steady state from that of a low-scale steady state. Once the economy falls into the low-scale basin, recovery is not guaranteed by "more competition" or "more enforcement"; the constraint is technological and intertemporal.

Section 5 uses this logic to make the policy point explicit. The object that matters is not firm size as such; it is the economy's distance from $L_{\min}$. Any intervention that reduces the attainable peak $\varphi_K^{\max}$ or raises the effective user cost D pushes $L_{\min}$ upward. In the fold regime, that is exactly how well-intentioned structural remedies can manufacture stagnation.

5 Policy: The Antitrust Fallacy and the Case for Restraint

The analysis in Sections 3–4 is intentionally austere: a standard Ramsey environment, endogenous labor supply, and a CES technology that allows increasing returns to scale. There are no markups, no strategic interaction, no entry barriers, and no monopoly pricing assumptions. Yet the model delivers a minimum-viable-scale constraint, steady-state multiplicity, and poverty-trap dynamics in the fold regime. The policy lesson is therefore not subtle: *the usual inference "IRS $\Rightarrow$ antitrust intervention" is not a logical implication of general-equilibrium theory.*

This section explains why the policy reflex is analytically misplaced and why, in the fold regime, structural intervention is not merely unnecessary but potentially harmful.

5.1 IRS is not a monopoly theorem

A recurring policy argument begins with a partial-equilibrium observation: with IRS, average costs can fall with output, so a single producer can undercut rivals and the market "tends" toward natural monopoly. The leap is then made from natural monopoly to a presumption of regulation or structural breakup. This chain conflates a technological property (IRS) with a market-structure conclusion (monopoly), and it relies on a framing that assumes away the margins that discipline scale in general equilibrium.

In our environment, IRS do *not* operate through monopoly pricing. They operate through *feasibility* and *intertemporal optimality.* In steady state, the Euler equation pins down a required marginal product of capital, $F_K(K, L) = D$, where D is determined by patience and depreciation. That condition is agnostic about industrial organization. It is a hard restriction on what stationary allocations can exist in the first place.

Once labor is endogenous, scale is a choice and a constraint: low labor implies low scale; low scale implies low marginal products; low marginal products further discourage labor. In the fold regime, this interaction produces a minimum feasible labor scale and multiple long-run allocations. None of this requires market power. It follows from the equilibrium conditions of a neoclassical economy with IRS and complementarity.

The implication for antitrust is immediate. If a policy conclusion requires monopoly distortions, it must be derived from assumptions about conduct, exclusion, or pricing power. It cannot be claimed as a free corollary of IRS. Treating "large scale" as presumptively suspect is not economic analysis; it is a presumption imposed on the model from outside.

5.2 The policy-relevant statistic is $L_{\min}$

The fold regime supplies a sharp object that organizes policy thinking: the *minimum feasible labor scale* $L_{\min}$. Section 4 shows that in the fold regime ($\theta > 1$ and $\xi < 0$) the reduced marginal product schedule in the capital–labor ratio is hump-shaped, so the Euler restriction implies a minimum labor requirement:

$$L_{\min} = \left(\frac{D}{\varphi_K^{\max}} \right)^{1/(\theta-1)}. \tag{5.1}$$

Here $D \equiv \beta^{-1} - (1-\delta)$ is the user-cost wedge and $\varphi_K^{\max}$ is the peak value of the reduced marginal product of capital.

Equation (5.1) is the correct lens for policy in this class of models. It says:

- There is an endogenous *minimum viability boundary.* If $L < L_{\min}$, the Euler condition is infeasible: no composition of inputs can deliver the required marginal product of capital.
- Policy cannot legislate feasibility. It can only move the economy *relative* to that boundary by shifting D or the technology objects embedded in $\varphi_K^{\max}$.

In an IRS economy with complementarity, the central risk is not "too much scale" but *insufficient scale.* The welfare-relevant question is whether policy pushes the economy away from, or toward, its minimum viable scale.

5.3 How intervention can backfire

The usual structural remedies advocated under the banner of "anti-monopoly" tend to do one of two things in practice: (i) fragment productive scale and integration, or (ii) raise wedges and uncertainty through compliance burdens, restrictions, or discretionary enforcement. In the present model, these map cleanly into the determinants of $L_{\min}$.

Fragmentation lowers $\varphi_K^{\max}$. Many interventions that aim to prevent "bigness" directly attack the mechanisms that generate high productive scale: integration, coordination, network complementarities, and fixed-cost spreading. In the model, these effects appear as reductions in effective productivity or complementarity—and hence as a reduction in $\varphi_K^{\max}$. From (5.1), a lower $\varphi_K^{\max}$ raises $L_{\min}$ mechanically. Raising the minimum viable scale makes the high-scale allocation harder to reach and easier to lose.

Wedges raise D. Policies that increase the effective user cost of capital—through taxation of returns, regulatory risk, higher effective depreciation (e.g., shorter planning horizons), or policy uncertainty—map into a higher wedge D. Again, (5.1) implies that higher D raises $L_{\min}$. In the fold regime, where feasibility is nonconvex, this is not a second-order effect. Increasing $L_{\min}$ can shrink the basin of attraction of the high-scale steady state and, for sufficiently large shifts, remove it altogether.

Nonlinearity makes "small" interventions not small. The fold geometry is a warning label for policymakers. Near a minimum-viability boundary, marginal changes in wedges or in effective complementarity can produce discontinuous changes in long-run outcomes. This is precisely the environment in which broad structural intervention is most dangerous: it can move the economy across a threshold and permanently lock it into the low-scale branch. The interventionist claim that "we can always fix it later" is therefore not credible in a model where feasibility itself is nonlinear.

The model thus flips the usual presumption. Under IRS with complementarity, the main policy hazard is not that scale emerges; it is that policy destroys the scale required for high productivity and then mistakes the resulting stagnation for a reason to intervene further.

5.4 Presumption of restraint

The analysis supports a presumption of restraint in competition policy when the justification rests primarily on scale economies.

Target conduct, not scale. If there is exclusionary behavior, collusion, or coercive restrictions on entry, these are distinct claims that require distinct evidence. They are not implied by IRS. Using IRS as a blanket rationale for structural remedies confuses technology with conduct and treats size as guilty by definition.

Do not destroy complementarities. In the fold regime, complementarities are the point: they are what make scale productive, and they are what generate the minimum viability boundary. Remedies that forcibly fragment production, limit integration, or otherwise suppress coordination are precisely the interventions most likely to raise $L_{\min}$ and deepen poverty traps.

Respect feasibility. The Euler restriction ties steady-state feasibility to D; policy cannot vote $F_K = D$ out of existence. If an intervention raises wedges or reduces effective productivity, it makes the feasibility problem harder. When the economy is near its minimum viable scale, the right default is not aggressive structural engineering. It is humility: do not impose policies whose predictable effect is to push the economy toward (or below) the fold.

In short, the correct posture in an IRS economy is not "intervene because scale exists." It is *refuse the interventionist shortcut.* Without a demonstrated conduct-based distortion, the burden of proof lies with the state. Structural intervention that treats scale as inherently suspect is not only analytically unjustified; in the fold regime it can be actively harmful.

6 Conclusion

This paper makes a simple point that is routinely obscured by tractability conventions and policy reflexes: once aggregate increasing returns to scale (IRS) are taken seriously in a dynamic general-equilibrium setting, the relevant implications are about *feasibility and scale coordination*, not an automatic presumption of monopoly distortions and structural regulation. The widespread policy syllogism—IRS $\Rightarrow$ "natural monopoly" $\Rightarrow$ default intervention—rests on toy partial-equilibrium logic that strips out endogenous labor supply, accumulation, and the intertemporal optimality constraints that govern long-run allocations.

Our framework restores these missing margins in a stripped-down Ramsey benchmark with endogenous labor and a CES technology that embeds an explicit returns-to-scale parameter. The analysis is organized by two objects. First, the steady-state Euler equation pins down a user-cost wedge D and imposes the restriction $F_K(K, L) = D$, which defines an iso–user-cost locus in (K, L) space. Second, homogeneity allows a ratio representation in terms of $r = K/L$, so that the steady-state problem collapses globally to zeros of a one-dimensional diagnostic $G(r)$ evaluated along that iso–user-cost locus. This reduction yields a tractable way to characterize stationary interior allocations even when the underlying problem is nonconcave under IRS.

The main economic content arises in the fold regime, defined by IRS and sufficiently strong complementarity ($\theta > 1$ and $\xi < 0$). In this regime the reduced marginal product of capital is hump-shaped in the capital–labor ratio, implying a *minimum feasible labor scale* $L_{\min}$: below $L_{\min}$, no composition of inputs can satisfy the Euler requirement $F_K = D$. Above $L_{\min}$, the iso–user-cost locus generically has two branches, and the diagnostic condition can intersect these branches more than once. The result is steady-state multiplicity and a transparent poverty-trap mechanism: low scale depresses marginal products and wages, discouraging labor supply; low labor supply prevents reaching scale. Importantly, this mechanism does not rely on markups, monopoly pricing, or exclusionary conduct. It is a feasibility-and-coordination phenomenon generated by IRS in an otherwise standard neoclassical environment.

The policy implication is therefore sharp. IRS are not a monopoly theorem, and they do not justify antitrust-style structural remedies by default. In a fold economy, interventions that fragment scale or raise wedges are not neutral "corrections"; they can move the economy toward (or across) minimum-viability boundaries. In particular, policies that reduce the peak attainable reduced marginal product $\varphi_K^{\max}$ or raise the effective user cost D increase $L_{\min}$, shrinking (and potentially eliminating) the high-scale steady state and deepening the poverty trap. The presumption should therefore run in favor of restraint: the burden of proof lies with the state to show conduct-based harms that outweigh the first-order risk of destroying viable scale. Targeting specific exclusionary conduct is conceptually distinct from punishing scale itself; the latter is precisely the error that the toy partial-equilibrium narrative encourages.

Finally, a scope note is warranted. Under IRS, the planner problem is typically nonconcave, so stationary allocations characterized by first-order conditions need not be globally optimal, and dynamic selection can depend on initial conditions and thresholds. Our results deliver a disciplined, global characterization of stationary interior allocations and a clear minimum-scale mechanism. A full welfare ranking and selection analysis is a separate question—but the central warning for policy does not require it: when feasibility is folded by IRS, indiscriminate structural intervention is not "safe," and it can permanently lock economies into low-scale outcomes.