# Notes on the Increasing Returns to Scale Problem and Romer (1990)
# Reflections on Lectures 5 and 6 of ECON 605

FRD*

February 11, 2026

## Introduction

In Lectures 5 and 6 of ECON 605 we discussed a simplified version of the Romer (1990) endogenous growth model, in which the growth rate of the economy depends on the endogenous creation of *productivity*—or, equivalently, of new ideas or varieties.[1] As argued in those lectures, one of the key challenges in this class of models is the so-called *Increasing Returns to Scale Problem* (IRTS Problem, hereafter). The IRTS Problem captures the tension that arises in competitive markets when the creation of new ideas collides with the fact that no one has incentives to pay for them—thus, no one produces them.

In these notes, I argue that this IRTS Problem is, in essence, a public goods problem: ideas are *non-rival*. The "standing on the shoulders of giants" effect implies that firms cannot choose the optimal level of ideas to employ, as existing ideas can be used freely. As a result, $A_t$ has no marginal cost, making it impossible to attach a price to its use. I will also discuss more broadly how imperfect competition in the intermediate goods sector helps resolve this problem, explore whether alternative mechanisms could achieve the same, and leave open questions that came up while I was working through this material.

## The Research Firm

The setup in Romer (1990) features a perfectly competitive research firm that takes the current level of ideas (or varieties) $A(t)$ as given, receives an exogenous price $P_A(t)$[2], and pays a wage $W(t)$ per unit of scientific labor $S(t)$.

*All (very probable) mistakes are my own and *only* my own. Not even of my friend ChatGPT that helped a lot!

[1] Using varieties as a synonym for ideas (and even for productivity) can be seen as a simplification. Still, it is a useful one if we think that the creation of a new variety always involves some creative or innovative process.

[2] For now, think of this as the firm facing a perfectly elastic demand curve. I will return to the implications of this assumption later.

At each point in time $t$, the firm chooses the amount of labor $S(t)$ that maximizes its instantaneous profits:

$$\max_{S(t)} \; P_A(t)\,\dot{A}(t) - W(t)S(t), \tag{0.1}$$

where $\dot{A}(t)$ is determined by the idea production function,

$$\dot{A}(t) = A(t)\,S(t). \tag{0.2}$$

Implicitly, we assume that the research firm immediately receives the infinitely-lived patent for producing the new variety $j$ and sells it at the market price $P_A(t)$ to potential intermediate goods producers.

The source of increasing returns to scale (IRTS) lies in the fact that the idea production function in (0.2) is homogeneous of degree two. This means that if we multiply each production factor by a constant $\lambda$, output increases more than proportionally. Formally:

$$\dot{A}(t) = F(A,S) = A(t)\,S(t) \quad \Rightarrow \quad F(\lambda A, \lambda S) = \lambda^2 A(t)\,S(t).$$

## Understanding the IRTS Problem

Now that we have identified where the IRTS appear, we can discuss their origin and implications. Looking closely at the maximization problem in (0.1), we see that the firm does not choose $A_t$ but only $S_t$. This implies that $A_t$ is indeed an input in the production function—otherwise we would not have IRTS—but one that the firm cannot choose how much to use.

This feature is crucial. The reason is that ideas are *nonrival goods*: once an idea has been created, it can be used freely by any firm at zero marginal cost.[3] If I, as a firm, can use at no cost all the available ideas in the world, and if these ideas raise my productivity, I will naturally want to use all of them.

But this is precisely the essence of the problem. If every firm can use all existing ideas *at no cost*, then what do I gain in exchange for producing a new one? This is the core intuition behind the IRTS Problem: the model must include an additional mechanism to make idea creation profitable.

**Question 1.** *In the lecture notes it is mentioned that firms cannot pay a fixed cost since they earn zero profits. Is that the whole story?*

No. In a perfectly competitive market, we know that variable (non-extraordinary) profits are zero. Thus, if the firm faces a fixed cost, it will incur losses—unless, in the long run, enough firms exit for profits (including fixed costs) to return to zero. But this situation can occur in any perfectly competitive industry, regardless of whether returns to scale are increasing or

[3] In more realistic settings, there might be frictions—for instance, a firm in Latin America may find it harder to adopt a technology developed in Europe—but the baseline assumption in Romer is full diffusion.

constant.

The problem here runs deeper. Since the research firm has no cost of acquiring the stock of ideas $A_t$, its marginal cost includes only the wage $W_t$. Consequently, the competitive price will equal this marginal cost, which is "too low" to compensate for the contribution of $A_t$. In equilibrium, price equals marginal cost and variable profits vanish once labor costs are paid. This is, in essence, what Euler's theorem for homogeneous functions tells us.

**Question 2.** *Okay, but if the firm can access all existing ideas basically for free, why should it include them as a cost?*

Good question. The short answer is: because these same ideas are the key resource that generates value for the research firm. If firms do not pay a cost for using the existing stock of ideas, then when the flow of new ideas becomes part of that stock, no one will receive income for having produced it. In this framework, intermediate goods firms do not pay for the ideas—they simply use them.

It is a vicious cycle: if I, as a researcher, do not have to pay for the ideas I use, then nobody will pay me for the ideas I create. The result is that the price of ideas collapses to zero, and the incentive to innovate disappears. This is the heart of the public-goods dimension of the IRTS problem.

**Question 3.** *But if the core is a public–goods problem, why do we call it the IRTS Problem rather than simply a public–goods problem?*

For me, this is the most difficult question. The two labels describe **the same friction from different angles**, but they are not interchangeable in general.

From a public–goods lens (micro), as discussed above, ideas are *nonrival*: once created, everyone can use them at zero marginal cost. Efficient use then requires a price of use equal to zero, which creates a *funding problem* for their creation. This is the classic public–goods tension.

From an IRTS lens (macro–GE), when nonrival ideas enter the production side, they typically *induce increasing returns at the aggregate level* (e.g., $Y = AL$ or $\dot{A} = AS$). Those increasing returns break the convexity/CRTS assumptions behind competitive general equilibrium: with linear prices that pay factors their marginal products, the product is exhausted and no rents remain to finance idea creation (see Euler's theorem reference in Lecture Note 5). In short: *nonrivalry* $\Rightarrow$ *aggregate IRTS* $\Rightarrow$ *the competitive decentralization fails*. This is the sense in which the literature calls it the *IRTS Problem*.

So, in Romer's baseline, calling it the *IRTS Problem* stresses that *nonrivalry of ideas manifests as aggregate increasing returns*, and those IRTS are what clash with competitive pricing and convexity. It *is* a public–goods problem at its core, but the term *IRTS Problem* flags the macro–GE consequences that force us to depart from perfect competition and/or introduce policy to restore incentives. In other words, nonrivalry is the friction and IRTS is its macro face.

**Question 4.** *Then every public good generates IRTS in production?*

No. A public good (nonrival and nonexcludable) does not necessarily induce increasing returns to scale (IRTS) in production. The confusion arises because in Romer's work, ideas do enter technology in such a way that they multiply rival inputs (or the law of the movement of knowledge), and this generates aggregate IRTS. But this is a particular functional form, not a universal property of "the public."

**Question 5.** *So, if we change how the production function looks, can we solve the problem?*

Yes—at least partially. If we modify the functional form of the production technology so that $A$ enters additively rather than multiplicatively, or if we "dilute" its effect by the size of the economy (for example, through $A/N$), the aggregate returns to scale can become constant rather than increasing.

In that case, the model no longer exhibits strong scale effects, and a competitive equilibrium can exist without the need for monopolistic rents. However, this transformation does not eliminate the *public-goods problem*: ideas remain nonrival, and their efficient use (at zero marginal cost) still conflicts with the need to finance their creation.

**Question 6.** *So we can have IRTS in a production function but not a public–goods problem? What would this produce in our model? Would we still have issues paying costs?*

Yes. Increasing returns to scale (IRTS) need not come from a public good. Think of large setup plant (large fixed costs, e.g. most natural monopolies), learning-by-doing internal to the firm or engineering economies that make average cost fall with scale. In all these cases, there are IRTS *without* a nonrival input like ideas.

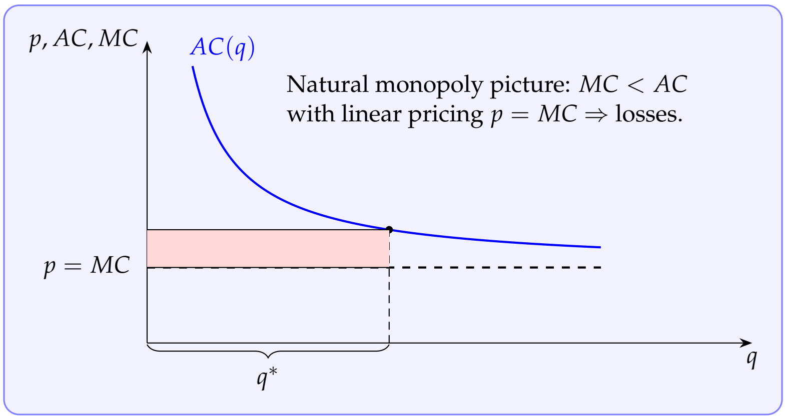
With IRTS, we will have decreasing average cost ($AC' \leq 0$) at least to a certain point. Decreasing average costs means that marginal costs ($MC$) runs below this $AC$ and then the competitive rule $p = \text{MC}$ implies $p < \text{AC}$, and then *losses* under perfect competition. This is the *natural monopoly / nonconvex cost* problem. It differs from Romer's public–goods tension: here the issue is *covering average cost*, not financing a nonrival input whose efficient price of use is zero. But you can see the simmilarities.

In a setup like this inputs are rival, there is *no* public–goods dimension: we don't require patents to finance a nonrival stock. The policy question is *how to fund fixed costs* (access fees, subsidies, franchise/licensing), not how to reconcile efficient free use with funding creation.

**Question 7.** *Is the problem always the same whenever there are IRTS?*

The common IRTS case already breaks $p = MC$ because $MC < AC$, so the firm loses money at competitive prices. Still, the good has a positive user price and you can, in principle, fund $F$ with two–part tariffs, average–cost pricing, Ramsey pricing, or franchises. It is more a local problem, i.e. how to fund the part of the costs that you are not being able to fund with the price.

With nonrival ideas it's tougher. Efficient use pushes the *user* price of the idea to (essentially) zero. Then there is nothing to collect per unit of use to pay for creation. So while both cases feature IRTS and trouble under linear pricing, the ideas case removes the very margin (a positive user price) that could contribute to funding. That is the sharper funding gap.




At $p = MC$, revenue $= p\,q^*$ and cost $= AC(q^*)\,q^*$, so losses are $\left[AC(q^*) - p\right] q^*$. In the ideas case the user price of the nonrival input is near zero, so this funding margin disappears.

**Question 8.** *Why did Romer not use a different production function and make the problem easier to solve?*

I think there are two main reasons, although I must admit that I am not completely sure. First, from a realistic point of view, we want $A$ to be a public good. That said, productivity should evolve exponentially, for two reasons.

1. The change in productivity should be, in some sense, exponential—otherwise we would not be able to explain post–Industrial Revolution growth. If $Y(t) = A(t)\,\bar{y}$, then $\dot{Y}(t) = \dot{A}(t)$, and we need $A(t)$ to grow exponentially to match sustained economic growth.

2. This second point is less clear, but we also need productivity to exhibit increasing returns to scale. In other words, why would we adopt a new technology if it leaves us in exactly the same position as before? We adopt new technologies only if they make us marginally more productive than the previous ones. This mechanism is, in fact, at the core of increasing returns to scale and of Romer's logic for using this type of production function.

## How Imperfect Competition Solves the Problem

Now that the origin of the tension is clear, the last step is to see how the model gets around it. The Romer logic is very direct: someone has to pay for the public good. Who benefits from ideas in the production side? Intermediate-goods firms. Who should pay? The same firms that use the ideas. For that to happen, we need two ingredients working together: (i) rents on the intermediate side so that there is actual money to pay for ideas, and (ii) a legal mechanism that makes the new idea temporarily exclusive so that someone can sell it. This is exactly why the model combines monopolistic competition in intermediates with patents.

On the pricing side, intermediate producers face CES demand from the final-good aggregator and a linear-in-labor technology. The static problem delivers a constant markup over marginal cost:

$$p_j(t) = \frac{\sigma}{\sigma - 1} W(t), \qquad \sigma > 1,$$

so each active variety earns a positive flow of operating profits $\pi_j(t)$. This is the same as having extraordinary variable profits. On the asset side, there is free entry into the market for designs. The price of a new design equals the present discounted value of those future profits,

$$P_A(t) \; = \; \int_t^{\infty} e^{-\int_t^{\tau} r(s)\, ds} \, \pi_j(\tau) \, d\tau,$$

and here is the key: A competitive research firm uses this price *as given* to choose scientific labor $S(t)$ to maximize profit as in the sense of (0.1). This price is higher than the price they would receive in the absence of one of the two ingredients mentioned before: monopolistic competition or patents. With $S(t) > 0$, the zero-profit condition for research pins down the knife-edge relation[4]

$$P_A(t) \, A(t) = W(t),$$

so the research firm breaks even each instant: it sells new ideas at $P_A$, pays wages, and earns zero economic profits given free entry. Intuitively, the rents are created in the intermediate sector through the markup, capitalized into the patent price, and then handed to the research sector to finance idea creation.

This decentralization solves the basic funding problem: we created a pool of rents without charging a per-use price for ideas. The final-good firm still buys intermediates competitively; ideas are used at zero marginal cost; and yet research is financed through the value of patents. That is the core appeal of imperfect competition here.

Two important caveats remain. First, the patent price internalizes producer surplus in intermediates, but not the consumer surplus that extra varieties generate in the final-good aggregator. As a result, the private return to innovation falls short of the social return (a classic externality problem where consumers are being benefited for something they are not paying!). Second, the research technology exhibits the shoulders effect: a higher $A$ today raises the productivity of future research. That dynamic externality is also not internalized by private innovators. Put differently, the markup-patent mechanism fixes funding but not efficiency: the equilibrium growth rate is typically below the planner's benchmark.

**Remark** (About symmetry and profits)**.** In the baseline, all intermediate producers are identical and set the same markup, so each earns a strictly positive flow of operating profits. There is no contradiction with free entry: entry occurs into the patent market, and the design price equals the discounted stream of those profits. After paying that price up front, an entrant earns the flow of profits over time. In present-value terms, there are no abnormal returns. Introducing firm heterogeneity would simply make profits and patent values differ across varieties; the logic of rents, patent pricing, and research zero-profit would remain the same.

[4] *Problem to the reader*: think the two *intuitive* reasons why $S(t)[P_A(t)\, A(t) - W(t)]$ has an interior solution.

## Conclusion

Ideas are nonrival which means that once created, anyone can use them at (essentially) zero marginal cost, so efficient allocation pushes toward a zero user price. Under competitive, linear pricing, rival inputs exhaust revenue, leaving no residual to finance discovery. That creates a funding gap: unless some rents (or public funds) are channeled to research, the private value of new designs drifts toward zero and innovation slows. This is the IRTS problem—not a complaint about competition per se, but a reminder that nonrival inputs don't fit neatly in convex, price-taking environments. Any workable decentralization must both keep use cheap and redirect some surplus to creation. The real task is striking that balance—via patents, markups, subsidies, or prizes—without overshooting on static distortions. That tension is the core of the problem.