Talent as Operator, Not Stock: The Curvature of Misallocation

Hyun Yoo*

June 26, 2026

Abstract

The misallocation-of-talent literature delivers one disciplined number: the output recovered by moving talent to where it belongs. That number is linear in the mass of reallocated talent, and it is linear for a single reason that is never argued, only assumed—talent is entered into the technology as a first-order multiplicative stock, a scalar that scales an input, rather than as an operator on the map by which human capital accumulates. The decision is equivalent to setting a cross-partial to zero, $\phi(T) \equiv \partial^2 F/\partial T\,\partial C \equiv 0$, where C is the investment intensity of the environment a person enters. We show that this is a substantive, sign-determining choice rather than a harmless normalization; that independent quasi-experiments holding talent fixed and shifting investment exogenously measure $\phi \neq 0$, and measure it positive, before discarding it as unexplained heterogeneity; and that restoring it decomposes the reallocation gain into a linear-stock term, an own-curvature term governed by F_{CC}, and an interaction term equal to $\phi(\bar{T})\ \mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$, so that the bias of the received number is signed by $\mathrm{sign}(\phi) \cdot \mathrm{sign}(\mathrm{Cov}_{\mathrm{moved}}(T, \kappa))$ and by nothing else. The entire reconstruction reduces to one estimable cross-partial, of which the received account is precisely the hypothesis that it is zero.

Keywords: talent allocation; dynamic complementarity; misallocation; human-capital accumulation; skill formation; innovation; reallocation accounting.

JEL codes: J24; O15; O31; O33; I26; D61; E24.

1 Introduction

Few quantities in modern macroeconomics are cited with the confidence reserved for the gains from putting talent where it belongs. Hsieh et al. (2019) attribute between 20 and 40 percent of the growth in U.S. output per worker from 1960 to 2010 to the improved allocation of talent across occupations—the dismantling of the barriers that had kept able women and Black Americans out of the professions in which their productivity was highest. Brouillette (2025) estimates that lifting the gender barriers to invention would raise U.S. income per person by 14.2 percent in the long run, against a mere 1.5 percent from a 30 percent subsidy to research and development. Akcigit et al. (2025) conclude, from a structural model of talent-driven endogenous growth calibrated to Danish administrative data, that subsidies to education can dominate research-and-development tax credits as engines of long-run growth. These are not marginal estimates. They are among the most-cited talent numbers in the discipline, and each licenses a policy conclusion of the first order: that the returns to repairing the allocation of human talent rival or exceed the returns to the

*I thank seminar participants for helpful comments. All errors are my own.

instruments—R&D incentives, subsidies to physical capital—that growth policy has traditionally favored.

Each of these numbers shares a property that has gone unremarked because it has never been stated: each is a *linear* functional of the mass of talent reallocated. Double the number of able people who are freed to flow to their best use and you double the headline gain; the marginal able person freed is priced at the same value as the first. This linearity is not a finding. It is not recovered from the data, and it is not the conclusion of any theorem. It is implied, silently and exactly, by a single modeling decision: *where the scalar that indexes a person's talent is permitted to enter the technology.* In each of these frameworks talent enters as a first-order multiplicative factor—a scalar T_i (or its model-specific cousin) that scales an input and is then carried, undeformed, through the aggregator. A factor that merely multiplies an input contributes the same proportional increment wherever it is applied, and so the gain from moving it is proportional to how much of it is moved. The linearity of the prize is the shadow of a functional-form choice about the placement of a scalar.

The aim of this paper is to locate that choice exactly, to show that it is substantive rather than cosmetic, and to derive the consequence of relaxing it. The decision can be read off a single object. Let human capital accumulate according to a map $H_{i,t+1} = F(H_{i,t}, C_{i,t}; T_i)$, in which $C_{i,t}$ is the intensity of the investment—course load, mentoring, the productivity of the surrounding research environment—that individual i receives, and T_i indexes the individual. The placement of talent as a multiplicative scale factor is exactly the restriction that the cross-partial

$$\phi(T) \equiv \frac{\partial^2 F}{\partial T_i \, \partial C_{i,t}}$$

vanishes everywhere. When this cross-partial is zero, talent's contribution is independent of how richly the person is invested in, the marginal value of moving a person is a constant across destinations, and the headline gain is linear in mass. When it is not zero, none of these is true. The entire edifice of linear talent accounting rests on the hypothesis $\phi \equiv 0$, a hypothesis that is nowhere written down as an assumption because it is built into the technology through the place the type is allowed to occupy.

That the choice is consequential is suggested by a pattern visible in evidence the allocation literature does not use. Take a population in which talent has already been measured, hold that measured talent fixed, and shift the investment a person receives by an exogenous intervention—random assignment to an enriched course, a discontinuity in eligibility for an accelerated track. Realized output then moves; the investment matters at fixed talent. More tellingly, it moves *more* for the abler. The return to the same exogenous dose of investment rises with the talent of the recipient, so that the effect of the intervention and the level of talent do not merely add—they interact, with a sign. A later section establishes this carefully and names the designs that show it. For now it is enough to register the tension: the interaction that the allocation models set to zero by construction is precisely the object that experiments built to measure investment effects find, find significant, and find positive.

Our contribution is deliberately narrow, and saying so plainly is part of the argument. We do not invent the object that does the work. The notion that ability and investment are complementary in the formation of skill—that a richer environment raises the return to talent, and more talent raises the return to a richer environment—is the technology of skill formation studied by Cunha and Heckman (2007) and estimated by Cunha et al. (2010); we take their cross-partial as established and borrow it intact. Nor do we contest that the allocation of talent matters for growth; that is the achievement of the literature we engage. What we do is carry an existing object across a boundary it has not crossed. The complementarity that has been measured *within* a person's life cycle—when and how much to invest in a child of given endowment—we install in the *cross-sectional* allocation

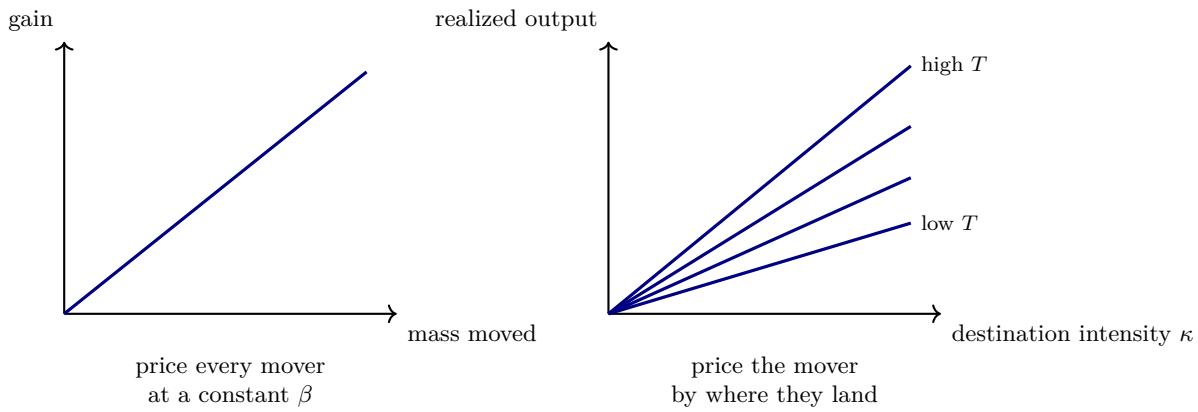


Figure 1: Two pricings of a reallocation. *Left:* when talent multiplies an input, every reallocated person carries the same constant marginal value β and the aggregate gain is linear in the mass moved, independent of destination. *Right:* when talent modulates the return to investment, the dose-response of realized output to a destination's investment intensity κ is steeper for the abler, the curves fan apart with talent, and the value of a reallocation depends on where talent lands. The figure is conceptual and carries no data.

problem, where it has a consequence no one has drawn: it makes the reallocation gain nonlinear, and it makes the bias of the linear number carry a determinate sign. The whole construction collapses to a single estimable cross-partial that may, as an empirical matter, be zero. A reader inclined to suspect that a contribution resting on one number is thin should hold that thought against the symmetry of the situation: the account we engage also rests on one number, the same cross-partial, fixed at zero by assumption rather than by measurement.

Because the claim is narrow, its boundaries should be stated at the outset rather than discovered later. We do not dispute the reallocation accounting of Hsieh et al. (2019); the wedges they recover from wage and occupational-share data are not in question, and the linear term we isolate is exactly their term. We do not claim to measure the talent of those never exposed to the relevant environments; our entire argument lives in an interaction estimated from people who were exposed, and is silent on latent ability in the unobserved. And we do not claim that restoring the cross-partial reverses the ranking of education against R&D in Akcigit et al. (2025); that ranking depends on objects their model fixes and ours does not re-solve, and we leave it genuinely open. What we claim is one thing: that the linearity of the canonical numbers is a choice, that the choice has a sign-bearing consequence, and that the consequence is estimable.

Figure 1 fixes the contrast that organizes the paper. On the left is the picture implicit in linear talent accounting: every person moved to a better use is priced at the same constant marginal value, so the gain is a straight line in the mass reallocated, indifferent to where each person lands. On the right is the picture this paper argues is the correct one: the value of moving a person depends on where that person lands, because the return to a destination's investment intensity is steeper for the abler, and the dose-response curves at different talent levels fan apart. The whole of what follows is the work of deciding, formally and empirically, which panel is right—and of showing that the question reduces to the sign of one cross-partial.

Before that choice can be indicted, it must be stated at its strongest—and on its own terms it is one of the genuine achievements of modern macroeconomics.

2 The Linear-Stock Apparatus at Its Strongest

The case against linear talent accounting cannot be made against a caricature of it. The frameworks at issue are among the cleanest in applied macroeconomics: their assumptions are explicit, their counterfactuals are point-identified from a small number of moments, and their conclusions are falsifiable. This section sets them out in their strongest form, on their own terms, before any of those terms is contested. The argument of the paper is not that the machine is incoherent. It is that the one assumption which makes its central output linear is a substantive choice, and that case is worth making only if the machine is first granted everything else.

The Roy template: talent as a multiplicative stock

The canonical statement is Hsieh et al. (2019). Each individual draws, for every occupation, an idiosyncratic talent in the tradition of Roy (1951), with the vector of draws governed by a multivariate Fréchet distribution whose parameters are assumed identical across demographic groups; whatever differences emerge across groups in occupational outcomes are therefore attributed not to talent but to the frictions and preferences that distort sorting. Human capital is produced once, through a schooling technology in which the time invested enters with an occupation-specific exponent ϕ^{HHJK}, and the talent draw appears *separately*, as a multiplicative scalar in net earnings rather than as an argument that interacts with the schooling input. The architecture is deliberate and powerful: because talent multiplies rather than reshapes the human-capital map, the aggregate is a clean integral over the talent distribution, the wedges that distort sorting are point-identified from wage gaps and occupational shares, and the counterfactual that removes the wedges is a transparent re-integration.

The closed-form intuition that results is the template the rest of this paper engages. To first order, the output gained by removing the distortions is proportional to the covariance, across groups and occupations, between occupational productivity and the distortion wedge—a misallocation is costly in proportion to how systematically the able are kept from the occupations in which they are most productive—and the aggregate gain rises with the *number* of people in the previously distorted group. The structure of the result is exactly "talent as a multiplicative stock implies a gain linear in the mass reallocated." It is worth being precise about why no curvature on the investment margin can arise here. Talent and the schooling input enter the human-capital map multiplicatively and separably, so the second derivative of human capital with respect to talent and the schooling input is simply the product of the two marginal effects—a positive constant, carrying no information beyond what the two marginals already carry. There is no margin along which more talent changes the *return* to investing more; the only complementarity is the mechanical one that any product exhibits, and it does nothing to bend the reallocation gain away from linearity.

Brouillette: the intensive margin without an accumulation stage

Brouillette (2025) sharpens the template by removing accumulation altogether. Inventive talent is the product $z_i = z_i^s \, z_i^n$ of a signal observed at birth and a noise term revealed only on entry into a scientific career, both drawn from heavy-tailed distributions; there is no life-cycle production function for inventive capability, because career entry is modeled as a one-shot revelation rather than a process of formation. Talent enters in exactly two places. It is the individual's supply of research labor, and it scales the input to the idea-production technology

$$\dot{M}_t = M_t^{\phi^{\text{spill}}} \, R_t,$$

in which the stock of ideas M_t grows with aggregate research effort R_t and the spillover exponent is calibrated to $\phi^{\text{spill}} = -2.1$, capturing that ideas become harder to find as the stock grows.

The calibration is disciplined and its anchors are worth stating, both because the paper returns to them and because they show how little freedom the framework leaves. The returns-to-research parameter is $\gamma = 1/3$; the tail index of the birth signal is $\theta_s = 1.26$; the elasticity of substitution across varieties is $\sigma = 1.97$; and the spillover is $\phi^{\text{spill}} = -2.1$, with $\gamma = 1/[(\sigma - 1)(1 - \phi^{\text{spill}})]$. From wage and participation data the model infers the distortions facing women in science: an exposure distortion $\tau^E = 0.79$, a labor-market distortion $\tau^L = 0.033$, and a capital-market distortion $\tau^C = 0.07$. Removing them raises aggregate research labor by roughly 49 percent over fifty years, and because output inherits the returns-to-scale exponent, $1.49^{1/3} \approx 1.142$ recovers the headline 14.2 percent gain in income per person. The estimate is robust in a band from 2 to 27 percent as γ sweeps the range $[0.05, 0.6]$ spanned by the growth literature, and falls to 8.7 percent when the tail index is set to $\theta_s = 2$. Decisively for what follows, the gain is almost entirely an *intensive*-margin phenomenon: the share of women among inventors barely changes in the counterfactual, and the income gain comes from better inventors rather than more of them—the freed Marie Curies are valuable because of how able they are, not because of how many they are. This framework complements Einiö et al. (2022) on the social determinants of the direction of innovation.

Tapping into talent: the richest member, and its binary gate

The most elaborate member of the family is Akcigit et al. (2025). Here talent is heterogeneous and growth is endogenous; financial frictions impede investment in human capital, career choice is endogenous, and the whole is calibrated to Danish administrative data linking childhood ability to adult outcomes. Idea production is a function of an individual's talent, the number of researchers, and the stock of automation, and aggregate growth inherits a power of average talent. The framework is rich enough to rank policies: it concludes that subsidies to education can dominate R&D tax credits, because education widens the pool of those who reach the research frontier. What is decisive for the present argument is the form education takes. Education is modeled as a binary gate—a fixed number of doctoral slots that an individual either passes through or does not—rather than as a continuous quantity of investment whose return talent might modulate. Because the investment margin is a discrete admission rather than a dose, there is no continuous input for talent to interact with, and the policy counterfactuals remain linear in the mass of talent admitted.

Three shared properties

Across these three frameworks the same structure recurs, and it is convenient to name it as a triple that the rest of the paper will invoke.

(P1) Talent is a level—T_i, or z_i, or a Fréchet draw—that multiplies an input.

(P2) Human-capital accumulation is either absent, as in the one-shot revelation of career entry, or separable from talent, so that the form of the accumulation map is the same at every level of talent.

(P3) The reallocation gain is $\Delta Y = \beta \cdot \Delta(\text{mass moved})$ for a constant marginal value β.

Table 1 lays out how each framework realizes the triple. The columns record where talent is permitted to enter the technology, the aggregator that carries it to output, the moments that identify the exercise, the headline number it delivers, and its status against (P1)–(P3). The table

Table 1: The linear-stock apparatus. Three frameworks, one structure: in each, talent is a level that multiplies an input (P1), accumulation is absent or separable from talent (P2), and the reallocation gain is linear in the mass moved (P3).

Framework	Where talent enters	Host aggregator	Identifying moments	Headline number	(P1)/(P2)/(P3)
Hsieh et al. (2019)	Multiplicative scalar in net earnings; separate from schooling input $s^{\phi^{\text{HHJK}}}$	Roy sorting across occupations	Wage gaps, occupational shares by group	20–40% of U.S. growth, 1960–2010	yes / yes / yes
Brouillette (2025)	Research-labor supply and idea production $\dot{M} = M^{\phi^{\text{spill}}} R$; no accumulation stage	Idea-stock growth, returns $\gamma = 1/3$	Inventor shares, inferred τ^E=0.79, τ^L, τ^C	14.2% income per person	yes / yes / yes
Akcigit et al. (2025)	Output scale, with education a binary PhD gate of fixed slots	Talent-heterogeneity endogenous growth	Danish ability–income link, career choice	Education subsidy $\succ$ R&D credit	yes / yes / yes

is the reader's reference card for everything that follows; the indictment of the next sections is, in the end, an indictment of the third column read against the last.

It bears emphasizing what this section does not do. It does not allege that the apparatus is internally inconsistent, that its identification is loose, or that its conclusions are unfalsifiable. The opposite is true on every count: the apparatus is coherent, elegant, point-identified from a disciplined set of moments, and falsifiable. The paper attacks none of this. The achievement of the linear-stock literature is real, and the critique that follows is not a critique of its rigor.

The question is therefore not whether this machine is coherent—it is—but whether the one assumption that makes (P3) linear, the separability buried in (P2), is a harmless normalization or a substantive, sign-determining choice.

3 The Undeclared Identifying Assumption

The linearity of the canonical numbers rests on an assumption that is never written down as one. This section locates it precisely, shows that it is a choice rather than a normalization, and draws out the consequence that the rest of the paper measures. The argument is still conducted in the language of the frameworks of Section 2; nothing in it presumes the conclusion. It establishes only that a single object governs whether the headline numbers are right, and that the received account has fixed that object by construction rather than by evidence.

The general map and its cross-partial

Begin with the most general statement of how capability forms. Let an individual's human capital evolve according to

$$H_{i,t+1} = F(H_{i,t},\, C_{i,t};\, T_i)\,, \tag{1}$$

where $C_{i,t}$ is the intensity of investment applied to i at t—the richness of the learning environment, the load of the curriculum, the productivity of the surrounding research setting—and T_i indexes

the individual. The frameworks of Section 2 are all special cases of (1): they differ in the aggregator that carries H to output and in how C is measured, but each chooses a particular place for T_i to enter F.

The object that governs the placement is the cross-partial

$$\phi(T) \equiv \frac{\partial^2 F}{\partial T_i \, \partial C_{i,t}} = F_{CT}.$$

Read in words, ϕ is the degree to which more talent raises the marginal return to investment—or, equivalently and by the symmetry of the second derivative, the degree to which a richer learning environment raises the marginal return to talent. It is the formal content of the idea that ability and opportunity reinforce one another: where $\phi > 0$, investing in an abler person yields more than investing in a less able one, and an abler person gains more from a given environment than a less able one would. It is the same cross-partial that the technology of skill formation puts at the center of human development, here written for an arbitrary accumulation map.

The chain of equivalences

The placement of talent as a multiplicative stock is exactly the restriction $\phi \equiv 0$, and that restriction propagates through the apparatus along a chain of equivalences that ends at the linearity of the headline:

$$\phi \equiv 0 \iff F \text{ separable in } T \text{ and } C \iff \text{talent's contribution independent of investment intensity}$$

$$\iff \beta \text{ constant across destinations} \iff \text{exact re-integration of the gain} \iff \text{(P3)}.$$

The first equivalence is a fact about functions: a twice-differentiable map has a vanishing cross-partial in two arguments if and only if it is separable in them, additively or, after a transformation, multiplicatively. The second restates separability in economic terms. The third is its allocative consequence: if talent contributes the same increment regardless of where a person is placed, the marginal value β of moving a person is a constant, independent of the investment intensity of the destination. The fourth follows because a constant marginal value can be integrated over the reallocated mass without reference to who lands where, and the fifth is just the resulting linearity. Every link is reversible. The linear-stock apparatus is neither more nor less than the hypothesis that the cross-partial in (1) is identically zero.

Why the assumption goes undeclared

What makes this assumption distinctive is that it is never declared. No paper in the literature writes "assume $\phi = 0$" and defends it; the restriction enters silently, through the seemingly innocent decision to let the type appear as a scalar that multiplies an input. The decision looks like notation. It is in fact identification. A type that merely multiplies an input is a stock of human capital by another name—a quantity of efficiency units that scales what is applied to it—and a separable scale factor cannot represent curvature at all, because separability sets the cross-partial to zero by construction. To place talent as a multiplicative scalar is therefore not to remain agnostic about the complementarity between ability and investment; it is to assert that the complementarity is exactly zero, while appearing to assert nothing.

The deeper error this conceals is a confusion of two distinct things: the *existence* of talent and a high *stock* of accumulated human capital. To say that a person is talented is to say that a characteristic of theirs modulates how capability accumulates—that the map F responds to them

differently. To say that a person has a large stock of human capital is to say that the level of H is high. These are not the same statement, and the multiplicative-stock placement collapses the first into the second: it represents the talented person as one who carries more efficiency units into the aggregator, which is precisely a statement about the level of the stock, not about the shape of the map. A framework built this way cannot distinguish a person whose talent bends the accumulation curve from a person who has simply accumulated more, because it has thrown away the only object—the cross-partial—that the distinction lives in. The category error is committed not as a claim but as a choice of where to write a letter.

The consequence, before any data

The consequence can be stated exactly, and it can be stated before a single estimate is brought to bear. Suppose $\phi \neq 0$. Then β is not a constant: the return to moving a person depends on the investment intensity of the destination, because the marginal value of investment is itself a function of the talent being invested in. The linear arithmetic, which prices every mover at one constant value, then omits a term—a curvature term that the next sections make explicit—and that term does not average to zero. It carries a sign, fixed by the sign of ϕ and the way talent sorts across destinations. The bias of the headline number is therefore *systematic*: a directional over- or under-statement of the true gain, not mean-zero measurement noise that careful estimation would wash out. A literature that has set $\phi = 0$ has not merely simplified; it has committed to a particular direction of error whose magnitude it cannot see, because the object that would reveal it has been assumed away. Proposition 1 makes the term and its sign precise.

There is a standing reply to all of this, and it should be put on the record now even though its refutation comes later. The reply is that $\phi = 0$ is a harmless normalization, that any interaction between talent and investment is a second-order detail, and that whatever heterogeneity it might generate is absorbed by the fixed effects that allocation exercises routinely include. The reply has the form of a methodological commonplace, and it would be persuasive if the interaction were in fact small and incidental. Whether it is small is an empirical question, and the reply assumes the answer.

That defense has a fatal feature: the interaction it dismisses as second-order noise is exactly what the field's own quasi-experiments measure—and then discard.

4 The Anomaly: When Talent Is Held Fixed

That interaction is not hypothetical. It is measured—signed, significant, and then set aside as unexplained heterogeneity—in the very populations and institutions the allocation literature treats as its own. This section assembles that evidence. The argument requires a precise target. Call a design a *counterexample to* $\phi \equiv 0$ if it exhibits identical talent, exposed to different accumulation inputs, yielding different realized output, where the difference is neither a level shift in talent alone nor a level shift in the input alone but their interaction. A level shift in talent moves output even when the input is held at zero; a level shift in the input moves output identically across talent levels. Only the third object—a difference in the *return* to the input that itself depends on the talent level—is the cross-partial F_{CT}. A design isolates ϕ when it varies the input exogenously, holds the talent level fixed and measured, and recovers a return to the input that rises (or falls) monotonically with that talent level. The designs below do exactly this, in populations and institutions the allocation literature treats as its own.

4.1 An experiment that writes the cross-partial into its design

The cleanest instance is the global randomized trial of Agarwal and Gaule (2026). The design is built, almost as if to specification, to vary the accumulation input while measuring talent independently of it. Six hundred and twenty Olympiad-nominated students from forty-four countries, spread across four cohorts, were assigned by matched-quadruplet randomization either to an eighteen-week advanced-combinatorics course or to equivalent self-study with the same materials and no live instruction or feedback. The talent level is a continuous, pre-determined object: each student's baseline selection-test score, recorded before any course exposure. The accumulation input is the live instruction the course supplies, and engagement with it.

Three readings of the experiment stack into a single conclusion. First, access alone—the intention-to-treat effect of being offered the course—raises learning by 0.165 SD. This is the effect of moving a student into a richer accumulation environment without conditioning on how much of it they take up. Second, because uptake is imperfect, assignment can be used as an instrument for engagement; scaling the access effect by compliance yields a return to full engagement on the order of 0.6 to 0.7 SD. The figure sometimes quoted as "0.66 SD" is not a single estimated coefficient but the product of a per-month local effect and the course's duration, a derived quantity that should be read as such; the per-month effect itself is estimated as essentially constant across the engagement spell, so the engagement margin carries no detectable curvature of its own. We return to the significance of that flatness when we separate the two second-order terms; for now it matters only that the engagement margin is locally linear.

The third reading is decisive for the cross-partial, and it is not a compliance story at all. Splitting the sample at the median of baseline ability, the treatment effect is 0.254 SD for students above the median and 0.006 SD—indistinguishable from zero—for students below it, with the equality of the two rejected at $p = 0.097$. The variable generating this gap is the continuous baseline selection score: a talent *level*, measured before exposure. A return to the accumulation input that is an order of magnitude larger for higher-ability students, where ability is fixed in advance and the input is randomized, is $\partial^2 H / \partial T \, \partial C > 0$ written into an experiment. It is the rate operator ϕ, recovered with the sign positive and the interacting variable beyond dispute.

That this is genuinely an ability-by-input interaction, and not a relabeled selection effect, is reinforced by how little of engagement the design can predict from anything but assignment. Regressed on the full baseline characteristic set, engagement has an R^2 of 0.086—under ten percent. Who takes up the input is, to first order, not forecastable from observables; the assignment instrument supplies variation in the input that is close to orthogonal to the level. The interaction is therefore read off the experiment's exogenous variation in the dose, interacted with a talent level it cannot itself manufacture.

4.2 The same latent talent, realized along different paths

If the trial shows that the return to an input rises with talent, a second pair of designs shows the complementary fact: the same latent talent, routed through different accumulation paths, realizes as different excellence. Bell et al. (2019), tracking who becomes an inventor in America, find that early measured ability accounts for strikingly little of the gaps that matter. Third-grade math scores explain 31.2% of the income gap in the propensity to invent, 2.4% of the gender gap, and—extrapolating the grade-by-grade fit back to birth—only about 5.7% of the income gap at birth. The lion's share of who invents, and crucially *in which direction*, is determined not by the level of childhood ability but by directed exposure. The probability of inventing in a given technology class jumps to roughly one in a thousand for children whose own environment placed them adjacent

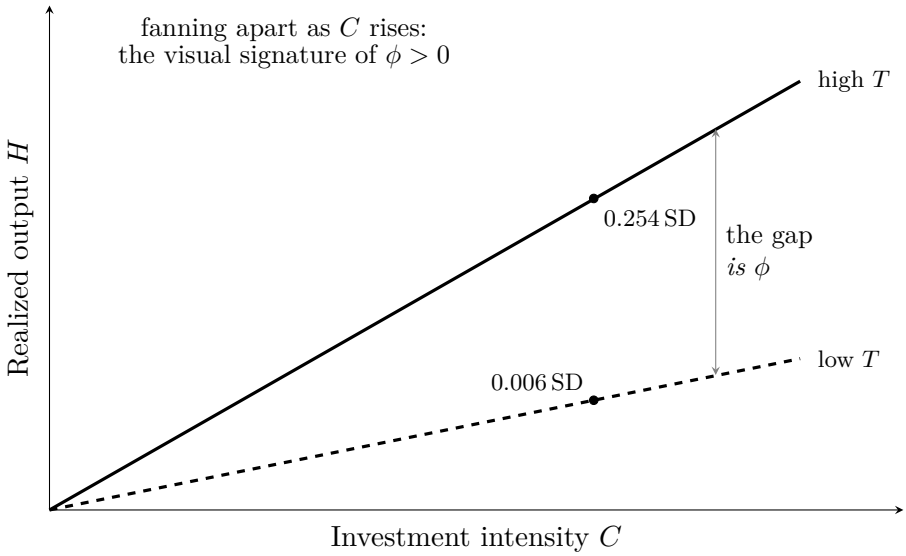


Figure 2: Dose-response of realized output on the accumulation input, drawn at two talent levels. Under $\phi > 0$ the curves fan apart as investment intensity rises: the same increment of input buys more for the abler. The marked points overlay the experimental instance of Agarwal and Gaule (2026)—a treatment effect of 0.254 SD above the baseline-ability median against 0.006 SD below it, the ability split read off a pre-determined talent level.

to that class, against under two in ten thousand one step removed—a more-than-fivefold gap, significant at $p < 0.01$. Ability does not vanish; it is held latent and then routed.

Card et al. (2024) make the routing visible at a sharp threshold. Their regression-discontinuity design exploits a gifted-eligibility cutoff at an IQ of 116, with a donut excluding the immediate neighborhood of the threshold to guard against manipulation. Crossing the threshold has *zero* effect on later test scores—math, reading, and college-admission tests alike are unmoved for both genders. Yet for disadvantaged boys it raises college entry by 28 percentage points, lifting the entry rate from 46% to 74% at the preferred eight-point bandwidth; for girls the effect is statistically indistinguishable from zero. The authors' own production function for the outcome is $y = \min[C, N]$, a perfect-complements form: realized excellence is the smaller of the accumulation input the institution supplies and a second, binding input the student brings. This is the strongest possible complementarity, and it makes the point in its starkest form—identical measured talent, a different accumulation path, a different realized outcome, with the difference carried entirely by the interaction of the two inputs rather than by either level alone.

Figure 2 renders the common structure of these designs. Plot realized output against investment intensity at two talent levels: under $\phi \equiv 0$ the two curves are parallel, the return to the input identical at every ability; under $\phi > 0$ they fan apart, the same increment of input buying more for the abler. The above- and below-median split of the trial sits on the figure as the empirical instance—two points on two diverging curves, the vertical gap between them the cross-partial itself.

4.3 Why the anomaly binds this literature in particular

It would be possible to dismiss all of this as evidence about gifted children, remote from the aggregate accounting of talent allocation. That dismissal fails on a point of adjacency. These designs are run on the very populations and institutions the allocation models invoke: gendered entry into science, the Olympiad pipelines that feed the inventive frontier, the gifted tracks through

which high-ability children are or are not routed. The misallocation accounting prices the gain from moving such talent to where it belongs, and prices it as linear in the mass moved. The interaction those models set to zero by construction is not an exotic object measured in some unrelated setting; it is measured, signed, and statistically significant in precisely the data that would discipline the allocation counterfactual. The heterogeneity discarded as unexplained noise is, in these designs, not noise and not unexplained: it is the effect, and it is signed positive.

One of these designs, however, is weaker than the others, and it is the one that most tempts an easy reading. Saying so plainly is not a retreat—it is what fixes the exact identifying assumption the reconstruction will have to carry.

5 What the Anomaly Identifies

The weaker design is the one whose headline is the most rhetorically convenient, and it is best to surrender that convenience at once. The accelerated-mathematics study of Card and Giuliano (2024) produces a result that looks, on first reading, like a clean talent-by-investment interaction; it is not, and seeing exactly why it is not is what sharpens the identification the rest of the paper relies on.

5.1 A doubling that is read off a group indicator

The study is a fuzzy regression discontinuity on entry into an accelerated middle-school mathematics track. The running variable is the fifth-grade state mathematics score; the cutoff sits at roughly the eightieth percentile; crossing it raises the probability of track entry by about 52 percentage points, and the students are followed for thirteen years. The result that draws the eye is sharply gendered. Marginally eligible girls roughly double their probability of completing a STEM degree, from 6% to 11%, and gain further in business and economics, so that the combined probability of a STEM-or-business/economics degree rises by 9.1 percentage points ($t = 2.5$). Boys at the same score are centered on zero across these outcomes.

It is tempting to read the gendered gap as a cross-partial: hold the score fixed at the cutoff, vary access, and the return is large for one group and null for the other. But the variable actually *varied* to generate the differential is not the talent level. The design fixes the score at the threshold and reads the heterogeneity off a group indicator—gender. What it identifies, with admirable cleanliness, is a group-by-access interaction. To the structural literature this is not a curvature object at all: a gender-by-access effect is precisely what a gender fixed effect absorbs, with no implication whatever for ϕ. A model in which talent enters as a separable scalar can accommodate the entire result by letting the level differ by gender, leaving the cross-partial at zero.

The result therefore bears on ϕ only through a bridging assumption, which it is more honest to name than to leave implicit. It informs the cross-partial if, and only if, the gender gap proxies a talent-relevant accumulation characteristic—some gendered difference in prior accumulation that shifts which input is binding at the margin—rather than a preference shift orthogonal to accumulated human capital. The two readings are observationally equivalent within this design. Under the first, access relaxes a binding accumulation constraint that the abler-but-under-accumulated group disproportionately faces, and the differential return is a shadow of the curvature. Under the second, girls simply value the STEM path more once admitted, and the differential is taste, not technology. Nothing internal to the discontinuity distinguishes them.

5.2 The gifted-education discontinuity, placed identically

The same discipline applies to the gifted-education discontinuity discussed in Section 4. Its threshold holds *cognitive* talent fixed at an IQ of 116; the input the institution moves is the non-cognitive margin N in the complementarity $\min[C, N]$; and the heterogeneity is again read across gender, with disadvantaged boys gaining 28 percentage points in college entry and girls essentially nothing. So this design is not direct evidence on ϕ as a function of cognitive talent either. It is evidence that the return to a non-cognitive investment depends on the gap between the two inputs, identified through a gender proxy for that gap, with the perfect-complements form an *assumed* functional specification rather than an estimated second derivative.

Yet the same study contains the most direct in-paper evidence that the operative margin is the binding-input gap and not gender as such. When the sample is split by baseline enjoyment of school—a proxy for the non-cognitive input the institution supplies—it is the *lower*-enjoyment students of *both* genders who benefit. The effect tracks the input gap, not the group label: where the institution's input is most scarce relative to the student's, the return is largest, regardless of sex. This is exactly the pattern the bridging assumption posits. It does not convert the discontinuity into a clean estimate of the cross-partial, but it shows that the gender split through which the effect is read is standing in for the binding-input gap, which is the talent-relevant accumulation characteristic the assumption requires.

5.3 The caveat the lead design must also carry

Honesty about the corroborating designs forces a matching caution about the lead one. The cross-partial of Section 4 is read off the interaction of baseline ability with engagement. If engagement is itself partly an expression of talent—if the persistence that elite assessments are designed to select on is a component of T rather than a separable input C—then the ability-by-engagement interaction conflates the cross-partial with a talent self-productivity term. Random assignment purges *who* engages: it severs the selection of high-ability students into high engagement, which is why the instrumented dose is close to orthogonal to the level. What randomization cannot purge is whether the *content* of engagement is itself talent. For this reason the engagement return—the 0.6 to 0.7 SD figure—is best carried as a bound rather than a point estimate of the curvature.

The clean cross-partial evidence is therefore not the engagement return but the ability split itself: 0.254 SD above the baseline-ability median against 0.006 SD below it. Here the interacting variable is unambiguously a pre-measured talent level, the dose is the instrumented access, and the conflation does not arise, because the level is fixed by the baseline test and the input is supplied by the randomized offer. The split point-identifies the sign of the cross-partial; the engagement return only bounds its magnitude.

5.4 The hierarchy this leaves

Conceding all of this does not weaken the case; it narrows it to its defensible core and, in doing so, strengthens it. Three tiers of evidence emerge, ordered by cleanliness. The ability split of Agarwal and Gaule (2026) point-identifies the *sign* of the cross-partial—bounding its magnitude—with that sign positive and the interacting variable a pre-determined talent level. The two regression discontinuities, Card and Giuliano (2024) and Card et al. (2024), corroborate the curvature only under the stated assumption that the group gap proxies a talent-relevant accumulation characteristic. And the directed-exposure results—Bell et al. (2019), together with the evidence from Nobel laureates of Novosad et al. (2026) that scientists disproportionately succeed in the fields their own origins routed them toward—speak to a different object entirely, the direction along which accumulation

is routed rather than its talent-modulated magnitude. That object is the subject of the section that follows.

Having fixed exactly what the data identify, and exactly what they do not, we can name the object that represents it—an object the separable stock cannot contain by construction.

6 Talent as the Curvature Operator

The evidence of the preceding two sections forces a definition of talent that the separable stock cannot accommodate. If the return to an accumulation input rises with a characteristic measured before that input is applied, then the characteristic does not multiply the input—it bends the map by which the input is converted into capability. Talent is not a stock that scales one undifferentiated input but an *operator* on the accumulation map; and it carries two formally distinct signatures, of which only one prices the reallocation.

6.1 The accumulation map and its cross-partial

Definition 1 (Talent)**.** *Talent* is the characteristic T_i that modulates the accumulation map—the rate at which an accumulation input is converted into realized human capital—measured prior to and independently of that input. It is an existence predicate, not a quantity of accumulated capability: talent is the curvature it induces, not a high value of the stock.

The qualifier "measured prior to and independently of the input" is not decorative. A characteristic measured *after* exposure—a wage, a patent, a post-program grade—is already a realized stock that embeds the input, and interacting it with the input would manufacture a cross-partial mechanically, committing on the empirical side the very category error the modeling side commits when it equates the existence of talent with a high stock of human capital. The measurement guardrail that enforces the definition is the subject of Section 8; here it is enough that talent is defined by what it does to the map, on a signal collected before the map is traversed.

Write the accumulation map as

$$H_{i,t+1} = F(H_{i,t}, C_{i,t}; T_i), \tag{2}$$

where $H_{i,t}$ is human capital carried into period t, $C_{i,t} \geq 0$ is the accumulation input applied at t, and T_i enters as a parameter of the map rather than as a factor multiplying an argument of it. The object that distinguishes an operator from a stock is the cross-partial of the map in the input and the parameter,

$$\phi(T_i) \equiv \frac{\partial^2 F}{\partial T_i \, \partial C_{i,t}} = F_{CT}, \tag{3}$$

the degree to which more talent raises the marginal return to investment or, read the other way, the degree to which a richer accumulation environment raises the marginal return to talent. We refer to a destination d by its investment intensity, the average input it supplies,

$$\kappa_d \equiv \mathbb{E}[C_{i,t} \mid \text{dest} = d],$$

so that moving talent across destinations is a movement along the input axis of (2).

6.2 The rate operator

The first signature of the operator is the cross-partial $\phi(T)$ of (3) itself: talent modulates the return to the *magnitude* of investment. This is the object the experimental ability split of Section 4 measures—a return to the randomized dose that is an order of magnitude larger above the baseline-ability median than below it—and the object the gifted-education complementarity exhibits, where realized excellence is governed by the interaction of the institution's input with a binding input the student brings. It is the only object on which the central proposition rests. When $\phi(T) > 0$ the marginal value of moving talent depends on the investment intensity of the destination it is moved *to*, and the linear arithmetic that prices every mover at a constant marginal value omits a term that carries a sign. Everything the reallocation accounting gets wrong, it gets wrong through this single second-order object.

6.3 The transfer structure

The second signature is distinct in kind and must be held rigorously apart from the first. Let the accumulation input be a vector over K field-directions, $C = (C^{(1)}, \ldots, C^{(K)})$, so that exposure has a direction as well as a magnitude. Define the *transfer kernel*

$$\Psi_T(k, k_i^\star) \equiv \frac{\partial H_{i,t+1}}{\partial C^{(k)}},$$

the directional gradient of realized capability in the direction of exposure k, peaked where the exposure direction k matches the individual's productive direction $k_i^\star$. Talent enters Ψ_T only as an index identifying which direction is productive for a given person; it is a first-order object, a gradient over directions, not a second derivative in talent. This is the structure the directed-exposure evidence identifies: the more-than-fivefold jump in the probability of inventing in a technology class when childhood exposure is adjacent to it (Bell et al., 2019), and the tendency of laureates to win in the fields their own backgrounds routed them toward (Novosad et al., 2026). Accumulation has a direction, and exposure that matches it transfers more.

6.4 Why the two signatures must not be conflated

The discipline the rest of the paper depends on is that Ψ_T is *not* a component of ϕ. The cross-partial ϕ differentiates the return to investment *with respect to talent*; the kernel Ψ_T is a gradient over directions *at fixed talent*. They answer different questions and are identified by different variation. The directed-exposure designs vary the direction of exposure while neutralizing the talent margin—through a class fixed effect, a cross-class difference, a gender-split regression, or a field-permutation null—so by construction they exhibit no $\partial\Psi_T/\partial T$. What they identify is the routing of accumulation, the kernel, not its talent-modulated magnitude.

There is, in principle, a third object: the genuine transfer cross-partial $\partial^2 H/\partial T\,\partial C^{(k)}$, measuring how talent modulates the *directional* return. Identifying it would require varying talent at fixed direction, which the present designs do not supply. We name it only to set it aside: it is a separate, currently unidentified object, flagged for future work and claimed nowhere in what follows. The paper's results rest on ϕ alone, and the kernel enters only as corroboration that accumulation is talent-modulated at all.

It follows that the directed-exposure designs cannot identify ϕ, and the reason is structural rather than incidental. Each estimator neutralizes the talent margin as a matter of design: a class fixed effect differences it out, a cross-class comparison holds it fixed, a gender split reads heterogeneity off a group rather than a level, a permutation null destroys exactly the talent-direction

match. What survives each of these is the routing of accumulation. To recover ϕ one must do the opposite—vary talent and investment *jointly*, as the randomized ability split does—and these designs are built to do the reverse.

6.5 The linear stock as a special case

With both signatures defined, the apparatus of the earlier sections is exactly located. The linear-stock literature is the conjunction

$$\{\, \phi \equiv 0 \quad \text{and} \quad \Psi_T \text{ direction-free} \,\}:$$

the cross-partial vanishes, so talent does not modulate the return to the magnitude of investment, and the kernel is flat across directions, so accumulation has no direction to match. Under this conjunction the map F is separable in talent and the input, realized capability is a stock that scales one undifferentiated input, and the reallocation gain is linear in the mass moved. The received account is not a different theory of talent; it is this corner of the present one. To restore ϕ is to leave the corner—and to inherit an obligation to derive what leaving it does to the headline.

With the operator named and its two signatures held apart, we can derive its consequence for that headline—a derivation that must keep apart a different pair, the two second-order terms, ϕ and F_{CC}, that separability silently fuses. The decomposition of Section 7, and the central result Proposition 1, carry that derivation; the re-pricing of Section 10 turns it into a curve.

7 The Curvature Term and the Sign of the Bias

The operator changes the algebra of reallocation. If talent merely scaled an input, the gain from moving a unit of talent to where it belongs would be a constant times the mass moved, and the headline numbers of Section 2 would be exactly what they claim to be. Once talent enters the accumulation technology as $\phi(T) = F_{CT}$, the gain acquires a second-order structure whose sign is determined by two observable objects. This section derives that structure, isolates the one term that carries ϕ from the one that does not, and shows that the bias of the linear-stock number is signed — not by an appeal to magnitudes, but by the product of two signs.

7.1 The allocation problem

Let talent be distributed across the population according to $G(T)$, with support in $\mathbb{R}_{>0}$. An *allocation map* $a : T \mapsto d$ sends each individual of talent T to a destination d, and each destination is summarized by its investment intensity $\kappa_d \equiv \mathbb{E}[C \mid d]$, the average accumulation input that the environment delivers. Through the accumulation technology $H_{t+1} = F(H_t, C; T)$, an individual of talent T allocated to a destination of intensity κ produces $F(H_0, \kappa; T)$ units of human capital. Aggregating over the population,

$$\mathcal{H}(a) = \int F\big(H_0, \kappa_{a(T)}; T\big)\, dG(T), \qquad Y = \mathcal{Y}\big(\mathcal{H}(a)\big), \tag{4}$$

with output a strictly increasing function $\mathcal{Y}$ of aggregate human capital. A reallocation policy is a change in the map, $a_0 \to a_1$; the object of interest is the induced change $\Delta\mathcal{H} = \mathcal{H}(a_1) - \mathcal{H}(a_0)$ and, through $\mathcal{Y}$, the change in output. The linear-stock literature evaluates this change by pricing every individual who moves at one constant marginal value. The question is what that pricing omits.

7.2 The decomposition

Consider a marginal mass of talent reallocated from a low-intensity environment κ_0 to a high-intensity environment κ_1, with each mover's investment allocated in proportion to the local intensity, so that the input applied to a mover traverses the interval $[\kappa_0, \kappa_1]$. Write the technology at fixed initial human capital as $F(\kappa; T) \equiv F(H_t, \kappa; T)$, treating F_C, F_{CC}, F_{CT} as partial derivatives in the investment argument, and anchor the expansion at $(\bar{\kappa}, \bar{T})$, the mass-weighted center of the move.

Proposition 1 (The curvature decomposition)**.** *Let F be twice continuously differentiable in (C, T). Then, per unit mass reallocated from intensity κ_0 to intensity κ_1, the change in aggregate human capital admits the decomposition*

$$d\mathcal{H} = \underbrace{F_C(\bar{\kappa}; \bar{T})\,(\kappa_1 - \kappa_0)}_{\text{linear-stock}} + \underbrace{\tfrac{1}{2}\,F_{CC}(\bar{\kappa}; \bar{T})\,(\kappa_1 - \kappa_0)^2}_{\text{own-curvature, } \propto F_{CC}} + \underbrace{\phi(\bar{T})\ \mathrm{Cov}_{\text{moved}}(T, \kappa)}_{\text{interaction, } \propto \phi}, \tag{5}$$

where $\mathrm{Cov}_{\text{moved}}(T, \kappa)$ is the covariance, taken over the reallocated subpopulation, between a mover's talent T and the intensity κ of the destination at which that mover lands.

The first term is the entire content of the received account: a constant marginal return $F_C(\bar{\kappa}; \bar{T})$ multiplied by the size of the move. The second and third terms are what separability discards. They are second-order, but they are not noise: each carries a definite sign, so each shifts the headline in a known direction rather than averaging out.

7.3 Proof at the level of the argument

The full derivation, the global version, and the closed-form special cases are collected in Section A; the structure of the argument is short enough to give here, because every subsequent claim of the paper depends on which approximation is made and where.

The per-mover gain is exact before any expansion. Holding initial human capital fixed and applying the fundamental theorem of calculus to the investment argument,

$$\Delta F(T) = F(\kappa_1; T) - F(\kappa_0; T) = \int_{\kappa_0}^{\kappa_1} F_C(\kappa; T)\, d\kappa, \tag{6}$$

which requires only that F be C^1 in its investment argument along the path. Equation (6) is the gain to one mover of talent T, written as the integral of the marginal return to investment along the path the mover's input traverses. No functional form has yet been imposed.

A single approximation enters, and enters exactly once: a first-order joint Taylor expansion of the *integrand* F_C in (κ, T) about the anchor $(\bar{\kappa}, \bar{T})$,

$$F_C(\kappa; T) \approx F_C(\bar{\kappa}; \bar{T}) + F_{CC}(\bar{\kappa}; \bar{T})\,(\kappa - \bar{\kappa}) + F_{CT}(\bar{\kappa}; \bar{T})\,(T - \bar{T}), \tag{7}$$

which discards third-order terms (F_{CCC}, F_{CTT}, F_{CCT}); these are reinstated in the global treatment of Section A. Substituting (7) into (6) and integrating term by term, with the anchor set at $\bar{\kappa} = \kappa_0$, yields a linear-stock piece $F_C(\bar{\kappa}; \bar{T})(\kappa_1 - \kappa_0)$, an own-curvature piece $\frac{1}{2}\,F_{CC}(\bar{\kappa}; \bar{T})(\kappa_1 - \kappa_0)^2$, and a per-mover interaction piece $\phi(\bar{T})\,(T - \bar{T})(\kappa_1 - \kappa_0)$.

The decomposition is completed by aggregating over the movers. A mover of talent T does not land at a common destination; she lands at the intensity $\kappa(T)$ that the reallocation actually delivers to her, displacing $\kappa(T) - \bar{\kappa}$. Aggregating the per-mover interaction against the mover distribution G_{moved},

$$\int \phi(\bar{T})\,(T - \bar{T})\big(\kappa(T) - \bar{\kappa}\big)\, dG_{\text{moved}}(T) = \phi(\bar{T})\ \mathrm{Cov}_{\text{moved}}(T, \kappa), \tag{8}$$

because the first moment of the centered talent over the movers is zero. This is the step that gives the third term of (5) its economic content: a *homogeneous* reallocation, in which every mover is sent to the same destination κ_1, makes the aggregate interaction vanish, since then $\kappa(T) - \bar{\kappa}$ is constant and $\mathbb{E}_{\text{moved}}[T - \bar{T}] = 0$. What survives is the *sorting*: the term is nonzero precisely when higher-talent movers land in systematically higher- (or lower-) intensity environments than lower-talent movers.[1]

One feature of this derivation is load-bearing and deserves to be stated as the methodological discipline it is. The own-curvature term in (5) contains no ϕ whatsoever: it is $\frac{1}{2} F_{CC}(\kappa_1 - \kappa_0)^2$, the curvature of the dose-response in investment *at fixed talent*. Had we held F_C constant in κ along the path — the natural shortcut — we would have set $F_{CC} = 0$ and silently re-imposed part of the very separability the paper places under scrutiny. Diminishing returns to investment are real and are present even under the separable null $\phi \equiv 0$; they belong in the accounting, and they must be calibrated on their own, from the engagement dose-response curvature, rather than folded into the talent margin. Keeping F_{CC} explicit and apart from ϕ is what allows the interaction term to be read as a clean test of complementarity rather than a residual that absorbs any nonlinearity.

7.4 What the decomposition implies

Three consequences follow by inspection of (5).

Corollary 1 (Direction of the error)**.** *If $\phi > 0$ and $\text{Cov}_{\text{moved}}(T, \kappa) > 0$, the interaction term is positive: the linear-stock account* understates *the gain from moving talent up the investment gradient and* overstates *the gain from moving it down.*

Corollary 2 (Sign of the bias)**.** *The sign of the interaction component of the linear-stock bias is $\text{sign}(\phi) \cdot \text{sign}(\text{Cov}_{\text{moved}}(T, \kappa))$, and nothing else.*

Corollary 3 (When the linear number is exact)**.** *The gain reduces to the linear-stock term alone if and only if $\phi \equiv 0$ and $F_{CC} \equiv 0$. Under $\phi \equiv 0$ alone, a pure own-curvature correction $\frac{1}{2} F_{CC}(\kappa_1 - \kappa_0)^2$ remains.*

Corollary 2 is the analytic payoff of the paper. The bias of the most cited numbers in the misallocation literature is not a quantity to be bounded by guesswork; it is the product of two signs, the curvature of the accumulation map and the sorting of talent across destinations, each of which can be estimated. Corollary 3 locates the received account precisely: it is not the special case $\phi \equiv 0$, but the strictly smaller special case in which the accumulation map is linear in investment as well as separable in talent. The distinction matters, because a finding that investment has diminishing returns ($F_{CC} \neq 0$) revises the headline without vindicating the operator; only the talent margin does that.

7.5 Scope of the proposition

Proposition 1 is local and second-order. It signs the *marginal* bias, around the anchor $(\bar{\kappa}, \bar{T})$, and the single approximation (7) that delivers it is exact only in the limit of small moves. For large discrete reallocations the third-order terms it discards can matter, and the sign of the interaction

[1] The interaction term has two equivalent expressions that differ only in units. Writing $\tilde{\kappa}$ for the displacement normalized by the move length, so that $\text{Cov}_{\text{moved}}(T, \kappa) = (\kappa_1 - \kappa_0)\,\text{Cov}_{\text{moved}}(T, \tilde{\kappa})$, the interaction is $\phi(\bar{T})\ \text{Cov}_{\text{moved}}(T, \kappa)$ in raw intensity and equivalently $\phi(\bar{T})\,(\kappa_1 - \kappa_0)\ \text{Cov}_{\text{moved}}(T, \tilde{\kappa})$ in normalized displacement. The two are identical; the raw form is the more transparent, since it expresses the interaction directly as a covariance of two measured quantities. All sign statements below are invariant to the choice, because $\kappa_1 - \kappa_0 > 0$ for an upward move.

can reverse. The mechanism is transparent in (6): the per-mover gain integrates F_C along a path, and if $\phi(T)$ changes sign over the talent range traversed — complementarity at moderate talent, substitutability in the extreme tail — a reallocation that straddles the sign change can carry an aggregate interaction opposite to the one the locally evaluated $\phi(\bar{T})$ predicts. Section A states the exact gain as a double integral, gives the sign conditions under which the local conclusion survives globally, identifies the three channels through which it can reverse, and works a numerical example in which a locally positive ϕ produces a globally negative interaction. The proposition is the analytic anchor of the argument, not the whole of it: where the moves are large, the sign is an empirical question that the global version, applied to measured curvature and measured sorting, answers.

7.6 Where the constant return nests

A word is owed on how the linear-stock term relates to the constant β of the received accounting, because the two price different things. The linear-stock term in (5) prices a change in a mover's *investment intensity*: it is the marginal return to the input, multiplied by the change in the input. The constant β of Hsieh, Hurst, Jones, and Klenow, or of Brouillette, prices *people*: it is a return per unit of reallocated mass. The two coincide only through an explicit mapping from occupations to investment intensities, and only under separability, where the marginal value of moving a person is independent of where the person lands.

For a host that already contains an investment stage — as in the human-capital production function of Hsieh, Hurst, Jones, and Klenow, where schooling and experience are inputs — the reconstruction is a genuine re-pricing of an object the model already carries. For a host with no accumulation stage at all — as in Brouillette, where research labor is the talent draw itself and capability is revealed, not built — the reconstruction does not nest the original by setting $\phi = 0$, because there is no investment margin to switch off. It instead inserts an accumulation stage between the talent draw and realized capability, and recovers the original number as the degenerate, no-curvature limit of that stage. The honest description in that case is a re-specification that adds a stage, not a re-pricing of the same object; Section 10 and Section C carry the construction through and show that the inserted stage reproduces the original headline exactly at $\phi = 0$, so the enrichment is disciplined rather than free.

7.7 The economic reading

The decomposition acquires its force from the structure of the reallocations the literature actually studies. Frictions do not keep talent out of high-intensity environments at random; they keep *under-exposed* high-talent groups out — the able child in a district without an accelerated track, the woman inventor without access to a research-intensive lab. Removing such frictions moves talent *up* the investment gradient, so the freed movers land where intensity is higher than where they began, and the higher-talent among them tend to reach the highest-intensity destinations. That is exactly the configuration $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) > 0$. Under complementarity, $\phi > 0$, Corollary 1 then says the received literature understates the gain from removing these frictions — and understates it in a concave, not a linear, way, because the return bends with the intensity of the destination rather than scaling with the count of movers.

This decomposition is worth deriving only if its load-bearing object, the cross-partial ϕ, can be measured without circularity. That is a live danger, not a formality: the wrong talent proxy — one that already embeds the very investment whose interaction with talent we mean to detect — makes ϕ mechanically nonzero and the entire test self-confirming. The next section builds the

measurement so that this cannot happen.

8 Measuring the Cross-Partial Without Circularity

A decomposition is only as good as the object it turns on, and the object that Proposition 1 turns on is exposed to a danger the linear-stock literature never had to face. Because separability holds talent and investment apart by construction, a structural modeler who sets $\phi \equiv 0$ can never measure something that is not there; the price of that safety is precisely the bias the decomposition signs. The reconstruction pays a different price. Once ϕ is the load-bearing coefficient, the wrong proxy for talent does not merely add noise—it can manufacture $\phi \neq 0$ out of nothing, because a "talent" measure that is itself realized human capital already contains the investment whose return it is supposed to modulate. The task of this section is therefore not estimation in the ordinary sense but the construction of a measurement protocol under which ϕ is identified as the talent×investment cross-partial and not as an artifact of how talent was measured. The argument is built so that every channel through which circularity could enter is named, signed, and either closed or bounded before a single coefficient is reported.

Four estimands

The protocol delivers four objects, kept rigorously distinct; their formal statements are deferred to Section B. The primary object is the rate cross-partial $\phi(T) = \partial^2 F/\partial T\,\partial C$, the degree to which talent raises the marginal return to investment intensity. The second is the own-curvature $F_{CC} = \partial^2 F/\partial C^2$, the diminishing return to investment at *fixed* talent; it is alive even under the separable null and must be calibrated separately so that it is never mistaken for ϕ. The third is the transfer kernel Ψ_T, the directional return to exposure of Section 6; it corroborates that accumulation is talent-modulated but does not price the reallocation and does not enter the estimating equation for ϕ. The fourth is the re-priced reallocation-gain functional—not a number but a curve in the destination intensity κ—formed by plugging the first three estimates into the decomposition; it is the secondary deliverable taken up in Section 10, and it is the only one of the four that inherits the sampling error of all the others.

The primary estimating equation

Let T_i be a continuous, pre-determined ability score and let C_i be the accumulation input, entering only through its instrumented, assignment-induced component $\hat{C}_i$. The relationship of interest is

$$H_{i,t+1} = \beta_0 + \beta_1 T_i + \beta_2 \hat{C}_i + \beta_3 \left(T_i \times \hat{C}_i\right) + \mathbf{x}_i'\gamma + u_i, \tag{9}$$

in which β_3 is the sample analogue of ϕ: it measures how the return to a unit of investment shifts with the talent of the recipient. The variable that interacts with the input is the ability score *itself*, a level, so a nonzero β_3 is a strict statement about $\partial^2 H/\partial T\,\partial C$ rather than about any group margin. Because engagement is chosen and therefore endogenous, both regressors that contain it—C_i and $T_i \times C_i$—are treated as endogenous and instrumented by $\{Z_i,\ T_i \times Z_i\}$, where Z_i is random assignment and T_i serves as its own instrument by virtue of being exogenous and pre-determined. With two endogenous regressors the conventional first-stage F is uninformative about weak identification of the interaction specifically; we therefore report the Sanderson and Windmeijer (2016) first-stage F for the interaction equation, the statistic that conditions out the other endogenous regressor and tests whether assignment moves engagement *differentially across the ability range*—the relevance condition on which β_3 actually rests.

The guardrail

The protocol's heart is a single restriction on what may serve as T_i. It is stated as an assumption because every result downstream is conditional on it.

Assumption 1 (Exposure-free talent)**.** The talent measure T_i is a signal collected before and independently of the accumulation input C_i: a uniformly administered test taken prior to any program, within-family or sibling variation, or a regression-discontinuity running variable held locally fixed at a cutoff. It is never a realized human-capital output—not a wage fixed effect, not a granted patent, not a post-program grade.

The reason is exact and not merely prudential. If the proxy for talent is itself an output that already embeds C, then $T_i \times C_i$ is in part $C_{\text{past}} \times C_{\text{now}}$, an input×input interaction that loads on F_{CC} and makes the estimated cross-partial mechanically nonzero whether or not talent modulates anything. This is the empirical twin of the modeling category error of Section 3: the structural literature mistakes the existence of talent for a high stock of accumulated human capital on the page, and an undisciplined empiricist commits the same confusion in the data. Assumption 1 is the axis along which the entire approach is positioned: a design earns the right to claim ϕ exactly insofar as its talent measure is collected upstream of the investment whose return it interacts with.

Separating the operator from diminishing returns

A rejection of $\beta_3 = 0$ vindicates the operator only if it cannot be explained by curvature in investment alone. To enforce this, F_{CC} is absorbed before the talent margin is tested. Within narrow talent bins we first estimate the dose-response curvature $\partial^2 H/\partial C^2$ at fixed talent. We then test the talent margin only after absorbing a flexible, common-in-talent function $s(\hat{C})$ that soaks up any such curvature:

$$H_{i,t+1} = \beta_0 + \beta_1 T_i + s(\hat{C}_i) + \beta_3 \big[(T_i - \bar{T}) \times \hat{C}_i\big] + \mathbf{x}_i'\gamma + u_i. \tag{10}$$

The operator test is $H_0 : \beta_3 = 0$ conditional on $s(\cdot)$ saturated, and centering talent at its mean $\bar{T}$ orthogonalizes the interaction from the main investment curvature so the two are not numerically entangled. The discrimination this enforces is the discipline of the whole paper: a nonlinearity in investment alone is $F_{CC} \neq 0$ with $\phi = 0$, and it does *not* vindicate the operator. Only the talent-margin interaction does.

Whether the available variation can support this separation is itself an empirical question, and answering it honestly bounds the claim. The global randomized trial of Agarwal and Gaule (2026)—620 students, with the above- versus below-median ability split significant at $p = 0.097$—supplies a single clean interaction margin but may not span both a fixed-talent investment margin and a fixed-investment talent margin with the precision the test demands. Where it does not, the test is under-powered net of F_{CC}, and the population register described below is required rather than optional. The protocol therefore reports the minimum detectable ϕ from the ability-split interaction after spline-absorbing $s(\hat{C})$ at $n = 620$, so the reader can see exactly when one trial suffices and when it does not.

Measurement error in the interaction

Suppose the proxy departs from true talent as

$$\widetilde{T}_i = T_i + L_a(\text{env}_i) + \epsilon_i, \tag{11}$$

with L_a an institution-embedded component correlated with investment and ϵ_i classical. The behavior of the estimated interaction then splits into three cases, each carrying a known sign; the leading-order plims are derived in Section B.

First, with classical centered ϵ alone, the interaction is the least reliable regressor: by the product-reliability result of Busemeyer and Jones (1983), the reliability of $T \cdot C$ is approximately the product of the separate reliabilities, so the interaction coefficient is *attenuated toward zero*. This case is benign—a rejection of $\phi = 0$ survives it—and the protocol concedes the benchmark openly rather than overstating the threat.

Second, with the environment-correlated component L_a, the term $L_a C$ is non-classical and contaminates the estimate with sign $\text{sign}(\text{Cov}(L_a, C)) \cdot \text{sign}(F_C)$. Because the marginal return to investment is positive and institution-embedded talent components covary positively with investment intensity, the generic contamination is positive: it biases the estimate *away* from zero and *toward* the operator conclusion. This is the dangerous, self-fulfilling case, and it is exactly the case in which attenuation toward zero fails to protect the inference.

Third, heteroskedastic measurement-error variance $\partial\sigma_\epsilon^2/\partial C \neq 0$ adds a term signed by that derivative, manufacturing a spurious interaction even at $\phi = 0$ when reliability drifts with investment. A separate, easily overlooked channel operates whenever talent is not centered: with $\mathbb{E}[T] \neq 0$, error in the *level* smears into the interaction coefficient, so a genuine main effect can masquerade as a cross-partial (Abel, 2018; Garber and Klepper, 1980; Griliches and Hausman, 1986), and the interaction-attenuation logic can even reverse (Muff and Keller, 2015). The test is therefore valid only after centering, after instrumenting $T \times C$ with $T \times Z$, and after confirming that the discontinuity-local estimate, where the proxy is fixed at a cutoff, survives. Collecting the channels, contamination drives the estimate away from zero exactly when

$$\text{Cov}(L_a, C) \cdot F_C - \frac{\partial \sigma_\epsilon^2}{\partial C} \cdot \mathbb{E}[T] > 0, \tag{12}$$

with $\mathbb{E}[T]$ the uncentered level. When this expression is positive the critique can be self-fulfilling; when it is negative the contamination is conservative and a rejection of $\phi = 0$ is, if anything, understated.

Signing the contamination

The expression in (12) would be a hedge if its sign were unknown. It is not: the protocol estimates that sign from data, converting a concession into a deliverable. The instrument is a *proxy ladder*—the same cross-partial estimated off measures ordered by exposure-freeness, from sibling or conscription-test variation, to a uniformly administered test, to a discontinuity-local running variable, and finally to a post-program grade used *for diagnosis only*. A monotone drift in the estimate as the proxy becomes less exposure-free signs $\text{Cov}(L_a, C)$; the post-program-grade rung is excluded from every reported estimate or bound, entering solely as the contaminated anchor that calibrates the contamination slope. In parallel, $\text{Cov}(L_a, C)$ is estimated directly by projecting the proxy on environment instruments at a cutoff where true talent is fixed; a reliability-by-investment regression signs $\partial\sigma_\epsilon^2/\partial C$; and the result is reported as a bound in the spirit of Black et al. (2000) rather than as a point with an assumed-away error. Where a structural route is available, the latent-factor and nonclassical-deconvolution machinery of Cunha et al. (2010) and Hu and Schennach (2008), applied to two or more pre-program ability indicators, estimates the interaction on the latent factor and dissolves product-reliability attenuation by construction; Lewbel (2012) supplies heteroskedasticity-based instruments where no external one exists; and the supermodularity inequalities of Athey and Stern (1998) test complementarity directly.

Data, and what survives

Three quasi-experimental panels supply the in-hand variation: the Florida district records behind the acceleration and gifted-education discontinuities, the 620-student global trial behind the ability split, and the patent–tax–neighborhood linkage behind the exposure results. None of these reaches beyond the upper tail of the talent distribution, and the macro claim concerns the whole of it. The protocol therefore seeks a population-representative register pairing an early, exposure-free ability measure—conscription cognitive tests or uniformly administered early tests in sibling and adoptee registries—with later output, so that ϕ can be estimated off the tail rather than within it.

This yields a deliberately two-tier deliverable, stated plainly so the reader is never misled about reach. With data in hand, the protocol delivers a tail-local bound on ϕ together with its qualitative sign, after the absorption of F_{CC} and the centering, instrumenting, and discontinuity-local safeguards; an interim pooled identification across the trial and the two discontinuities, under the group-proxies-talent assumption of Section 5, buys power while the register is pending. With the register, it delivers the population-wide ϕ and hence the re-priced macro curve of Section 10. If the register never arrives, what survives is exactly the tail-local bound and the sign result—not the re-priced macro headline. Saying so is not a retreat but the boundary of the claim. Once ϕ is measured under this guardrail, the entire critique collapses to a single coefficient—and that coefficient may be zero.

9 The Decisive Test

The reconstruction of Section 6 and the decomposition of Proposition 1 have a property that critiques of established results rarely possess: they reduce to a single number that can come back wrong. The entire apparatus—talent as operator, the signed bias, the re-priced curve—rests on one coefficient, the response of realized human capital to the interaction of pre-determined talent with an exogenously shifted investment, after the within-talent dose-response curvature F_{CC} has been absorbed. If that coefficient is a precise zero, talent and investment are separable in the cross-partial sense, the reallocation gain carries no operator correction, and the linear-stock arithmetic stands exactly as the literature wrote it. The purpose of this section is to stake that bet in the open, before any estimate is in hand, so the argument is a hypothesis rather than an unfalsifiable reframing.

The test is the estimable coefficient ϕ itself: the coefficient on (pre-determined talent level $\times$ exogenously shifted input) in the absorption specification (10), identified under the guardrail of Assumption 1. Its measurement is decisive precisely because it is one number with a sharp null. The received account is not adjacent to the hypothesis $\phi = 0$; it *is* that hypothesis. To accept the linear-stock literature's arithmetic is to assert that talent multiplies an input without bending the return to it, and that assertion is the statement $\phi = 0$ stated in the data rather than assumed on the page.

Everything depends on isolating the right nonlinearity. A curved dose-response in investment alone is $F_{CC} \neq 0$ with $\phi = 0$, and it does not vindicate the operator: diminishing returns to investment are present even under the separable null and price no reallocation bias. The test must therefore detect the *talent-margin* interaction specifically, which is why F_{CC} is absorbed by the saturated common-in-talent function $s(\hat{C})$ before ϕ is read off, and why centering at $\bar{T}$ orthogonalizes the interaction from the main investment curvature. A finding that realized output is concave in investment is consistent with the linear stock; only a finding that the abler extract a higher return from the *same* exogenous dose discriminates against it.

The decisive instance is the instrumented-engagement-by-continuous-baseline-ability interaction

of Agarwal and Gaule (2026), where the interacting variable is unambiguously a pre-measured talent level and the dose is randomized access; the above-median treatment effect of 0.254 SD against 0.006 SD below median ($p = 0.097$) is that interaction in reduced form, awaiting the absorption and instrumenting that turn it into a clean ϕ. Under the assumption of Section 5 that a group gap proxies a talent-relevant accumulation characteristic, the access effect across the talent-group margin in Card and Giuliano (2024)—marginally eligible girls roughly doubling their STEM-degree probability from 6% to 11% where same-score boys are centered on zero—and the gifted-label effect across the binding prior input in Card et al. (2024)—a 28 percentage-point college-entry effect for disadvantaged boys with zero effect on test scores—corroborate the same margin. None of these is asked to carry the identification alone; the trial's ability split leads, and the rest corroborate.

The implication that makes the test falsifiable beyond its own coefficient is that the marginal reallocation return is heterogeneous in the destination intensity κ, with its sign pinned by $\text{sign}(\phi)$. Under the operator, moving talent up the investment gradient earns a return that bends with where talent lands; under the null, that return is flat in κ. A flat profile is not a weak result—it is the linear stock vindicated, and the cleanest concession the argument can make. If the test returns zero, the paper is wrong in the way a falsifiable paper is permitted to be wrong: the operator correction vanishes, F_{CC} may still warrant a separate non-operator adjustment, and the canonical number needs no ϕ-re-pricing. The strength of the position is that this outcome is named in advance and welcomed as informative.

Supposing the test does not return zero, the deliverable is not a corrected number but a corrected object—a curve.

10 From a Number to a Curve

The argument to this point has been about what the reallocation arithmetic omits. It is time to put that omission to work on a concrete number. The sharpest recent statement of the lost-talent counterfactual is Brouillette (2025), who shows that removing the barriers that have kept women out of invention raises aggregate research labor by about 49% over fifty years and, through an aggregator with increasing returns ($\gamma = \frac{1}{3}$), a long-run income gain of $1.49^{1/3} - 1 \approx 14.2\%$ – an order of magnitude larger than the 1.5% a thirty-percent R&D subsidy would buy. The result is exact and, on its own terms, clean. It is also, on its own terms, a single number. The purpose of this section is to show what that number becomes once talent is allowed to act on the map by which inventive capability accumulates rather than merely to scale a fixed input.

The transformation is concrete: the headline ceases to be a scalar and becomes a curve indexed by where the freed talent lands. If the women now able to enter research arrive in high-intensity environments, 14.2% is a lower bound; if access without engagement leaves the accumulation input low, it is an upper bound. The single number was never wrong; it was the value of a function read off at one point, with the rest of the function suppressed by the assumption that talent is a stock.

10.1 The inserted stage, in the model's own objects

Brouillette (2025) carries no life-cycle of inventive productivity: talent enters the technology as the research-labor identity $\ell_i^R = z_i$, with $z_i = z_i^s z_i^n$ the product of a birth signal and an entry-stage noise draw, each Pareto. There is no stage at which an environment's investment in a person shapes the capability that person brings to the frontier. We insert exactly one such stage, and nothing else. Between STEM entry and realized research-labor productivity we place a realized inventive capability

$$H_i = F(z_i, C_i), \qquad F_{CT} = \phi > 0, \tag{13}$$

where C_i is the investment intensity of the research environment the now-larger pool of women inventors enters, and the rest of Brouillette (2025)'s structure – the savings problem, the idea law of motion, the overlapping-generations transition, the calibration targets – is left untouched. The talent measure feeding (13) is the pre-investment birth signal, observed strictly prior to any research-environment investment, so the object (13) interacts is talent as it exists before the environment acts on it.

Two functional forms make the cross-partial transparent and bracket the economically relevant range. The first is Cobb–Douglas,

$$H = z\,C^{b}, \qquad \phi = \frac{\partial^2 H}{\partial z\,\partial C} = b\,C^{\,b-1} > 0 \iff b > 0, \qquad F_{CC} = b(b-1)\,z\,C^{\,b-2} < 0 \iff b < 1,$$

which sets the cross-partial to zero precisely at $b = 0$ and there reduces to $H = z$, the original research-labor identity; because $H = z\,C^{b}$ remains a power law in z for any destination-constant C, the Pareto-tail algebra that Brouillette (2025) relies on survives intact, re-weighted by a factor C^{b}. The second is constant-elasticity,

$$H = \left(\omega z^{\rho} + (1-\omega)C^{\rho}\right)^{1/\rho}, \qquad \operatorname{sign}\phi = \operatorname{sign}(1-\rho),$$

which spans complementarity ($\rho < 1$), the separable knife-edge ($\rho = 1$), and substitutability ($\rho > 1$), and which admits as its Leontief limit $\rho \to -\infty$ the perfect-complement production $\min[C, N]$ that Card et al. (2024) estimate for gifted placement among disadvantaged students – a destination environment in which an able student's realized capability is capped by the slot available to receive it. In both forms the received linear-stock specification is recovered as a single interior parameter value; the operator claim is that this value is not the one the data select.

A point of notation must be settled to avoid a genuine collision. Brouillette (2025) uses the symbol ϕ for the knowledge-spillover exponent in the idea law of motion $\dot{M} = M^{\phi^{\text{spill}}}R$, calibrated to $\phi^{\text{spill}} = -2.1$. That object governs the decreasing returns to the existing stock of ideas in the production of new ones; it is a different object from the talent–investment cross-partial in (13). Throughout we write the spillover exponent ϕ^{spill} and reserve ϕ for the cross-partial.

10.2 Why this re-prices the channel rather than replacing it

The natural objection is that adding a stage replaces the original model rather than re-pricing its result. The objection is answered by the place the new stage acts. Brouillette (2025)'s central finding is that the gains come almost entirely from the intensive margin – better inventors, not more of them – because the female share of inventors barely moves while the capability behind each unit of research labor rises. In the enriched model that intensive margin is precisely an average of realized capability H_i, and under $\phi > 0$ that average is convex in the destination intensity κ: the very margin the original result leans on is the margin the cross-partial bends. The inserted stage does not introduce a new channel and credit it with the gain; it supplies the missing curvature of the channel the original already identified as load-bearing.

The honest content of the nesting must be stated exactly, because it is not the usual kind. Setting $\phi = 0$ (equivalently $b = 0$, or $\rho = 1$) returns the headline to 14.2% exactly, as the arithmetic of Section C shows for every landing intensity. But this is a recovery of the original's *number*, not a structural nesting of an existing object: the original has no investment stage for the operator to switch off, so the enrichment inserts a stage and recovers 14.2% as that stage's degenerate, no-curvature limit. We claim no more than this, and no less; the discipline that keeps the move legitimate is that the inserted stage acts on the original's own intensive margin and that all of its

calibration targets are held fixed, as we make precise below. The distinction between recovering a number and nesting an object is the same one that separates a re-specification from a re-pricing, and we keep it visible rather than concealing it (cf. Section 7).

10.3 The aggregator re-weight and the headline curve

The full arithmetic is in Section C; the result is short. Aggregate research input ceases to be the talent mass and becomes the capability mass,

$$R^{\text{op}} = \int z_i \, C_i^b \, di,$$

so the only thing that changes downstream is the content of R, re-weighted by C_i^b; the aggregator $c \propto R^\gamma L/N$ with $\gamma = \frac{1}{3}$, the idea law of motion, and the overlapping-generations transition are unchanged. Writing the destination re-weight $\Lambda(\kappa_d, b) = (\kappa_d/\kappa_0)^b$ relative to incumbent intensity κ_0, and applying it to the 49% labor rise, the gain *conditional on a representative freed mover landing at destination intensity* κ_d is

$$\boxed{\text{gain}(\kappa_d, b) = \left[1 + 0.49\,(\kappa_d/\kappa_0)^b\right]^{1/3} - 1, \qquad \gamma = \tfrac{1}{3}.} \tag{14}$$

At $b = 0$ (or $\rho = 1$), $\Lambda = 1$ for every κ_d and (14) returns $1.49^{1/3} - 1 = 0.142$ for every destination: the flat linear-stock baseline is the no-curvature reading of the curve.

One distinction governs how (14) may be used, and conflating the two objects it separates is the exact error this paper indicts in the linear literature. Equation (14) is a per-destination *conditional* object: it prices a representative mover at a given κ_d in order to display the mechanism, and it varies with the *mean* destination intensity. It is not the aggregate gain. The aggregate re-priced gain integrates over the distribution of landing intensities, and by Proposition 1 its bias relative to the linear number is governed not by the mean destination but by the interaction $\phi \cdot \text{Cov}_{\text{moved}}(T, \kappa)$ – the covariance, over movers, between talent and where it lands. A mean-preserving spread of κ with $\text{Cov}_{\text{moved}}(T, \kappa) = 0$ leaves the aggregate first-order unchanged; positive sorting, high-talent freed inventors arriving in high-intensity laboratories, is what moves it. We therefore report the two objects side by side: the conditional curve to show how the headline bends with the level of the destination, and the $\phi \cdot \text{Cov}_{\text{moved}}(T, \kappa)$ term to instantiate the proposition.

Figure 3 draws the conditional curve. Each line fixes a value of the complementarity b and traces (14) as the destination ranges from half to double incumbent intensity. All lines pass through the pivot $(1, 14.2\%)$ – a representative mover landing at incumbent intensity is priced identically regardless of b – and the flat reference at 14.2% is the $b = 0$ linear-stock line. The band between the most pessimistic and the most optimistic complementarity is the range the single number conceals once curvature is admitted.

The shape of the dependence is itself informative, and worth isolating from the level. Figure 4 fixes the destination and instead traces the gain against the complementarity b, for a low, an incumbent, and a high landing intensity. The incumbent line is flat at 14.2% – when a mover lands exactly at κ_0, the re-weight $\Lambda = 1$ for every b, so complementarity cannot move the conditional headline. Away from the pivot the lines fan: a high destination rewards complementarity, a low destination penalizes it, and the two open symmetrically around the pivot as b grows. The figure makes vivid that the operator matters only to the extent that landing intensity departs from the incumbent norm – which is exactly the content of the interaction term.

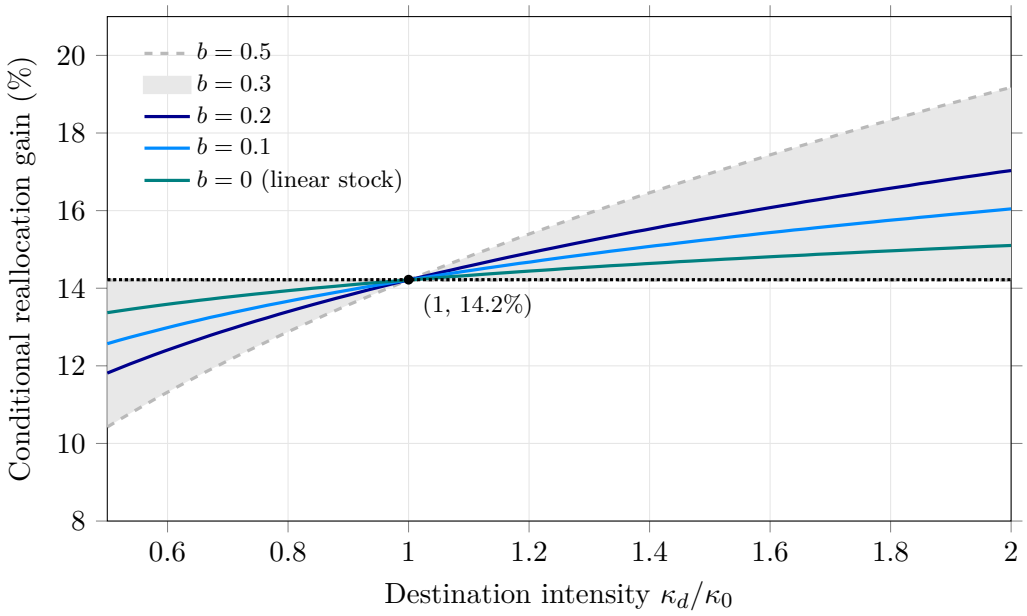


Figure 3: The headline as a curve. Each line evaluates the per-destination conditional gain (14) for a fixed complementarity b as the representative freed mover's destination intensity ranges from half to double incumbent intensity κ_0. All curves pass through the pivot $(1, 14.2\%)$; the dotted line is the flat linear-stock baseline ($b = 0$); the shaded band spans the no-complementarity case and $b = 0.5$. Values are illustrative pending estimation of b.

10.4 Separating the operator from the increasing-returns amplifier

A curve that widens could do so for two reasons that must not be confused. One is the operator channel proper: the cross-partial re-weights the capability mass R^{op}. The other is the aggregator's increasing returns: with $\gamma = \frac{1}{3} < 1$, any perturbation to R is raised to a concave power, which compresses spreads, while the underlying re-weight itself is what generates them. To keep the attribution clean, Figure 5 redraws the $b = 0.2$ conditional curve under the calibrated $\gamma = \frac{1}{3}$ alongside the same re-weight fed through a constant-returns aggregator, $\gamma = 1$, in which the gain is linear in R. The vertical distance between the two is the work the increasing-returns lever does; the slope common to both is the operator channel. The spread the operator opens is not an artifact of the aggregator, and the aggregator's amplification is not by itself evidence of an operator.

The same separation must be enforced for the bias decomposition, where a third quantity enters. The reallocation correction has three pieces – the linear-stock term, the interaction term $\propto \phi \cdot \text{Cov}_{\text{moved}}(T, \kappa)$, and the own-curvature term governed by F_{CC} – and only the second is the operator. The own-curvature term is diminishing returns to investment at fixed talent; it is present even under the separable null and must never be folded into the operator correction. Figure 6 displays the three at a representative upward move, with illustrative magnitudes consistent with a modest complementarity $b \approx 0.2$ and positive sorting: the linear-stock term carries most of the gain, the interaction adds to it under $\phi > 0$ and $\text{Cov}_{\text{moved}}(T, \kappa) > 0$, and the own-curvature term subtracts a smaller amount because $b < 1$. Reading the figure, the operator's contribution is the middle bar alone.

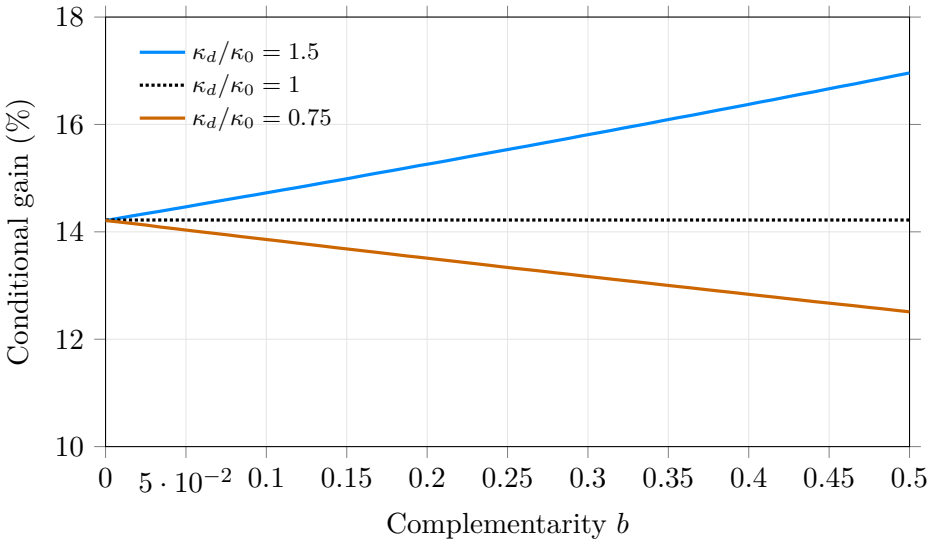


Figure 4: Conditional gain against complementarity b, at fixed destination intensities. At the incumbent norm $\kappa_d = \kappa_0$ the gain is flat at 14.2% for every b; above and below it the lines fan with b, since complementarity bites only when landing intensity departs from the norm. Illustrative pending estimation.

10.5 Re-calibration, with no free parameters

The re-pricing earns its credibility by spending no calibration freedom. The complementarity b (equivalently ρ) is to be estimated from an instrumented talent–investment interaction, not chosen to fit; the own-curvature F_{CC} is calibrated separately and carried as its own term; the sorting $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$ is swept across its full range, negative as well as positive; and every other calibration target in Brouillette (2025), together with the three identifying moments, is held fixed, so the operator receives no recalibration of the host model. What the exercise then reports is the new conditional curve, its concavity in κ, the signed interaction component of the bias, and the F_{CC} correction – each as a distinct object. The protocol that delivers them, including the precise reason the available quasi-experimental coefficients identify ϕ but not F_{CC}, is set out in Section C.

The illustrative magnitudes, labeled illustrative pending estimation, are these. Under a modest complementarity $b \approx 0.2$, the conditional headline widens from 14.2% to roughly 12.6%–16.0% as the destination ranges from half to double incumbent intensity; under a stronger $b \approx 0.5$, to roughly 10.4%–19.2%. The central case is the upward branch, toward 16–19%, because the sorting we argue for – frictions having kept high-talent women out of the highest-intensity environments, so that lifting them lets the ablest land where investment is deepest – is $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) > 0$; the downward branch is the contingency in which access without engagement leaves the input low. The direction is pinned, by Proposition 1, to $\mathrm{sign}(b) \cdot \mathrm{sign}(\mathrm{Cov}_{\mathrm{moved}}(T, \kappa))$. This sweep is a distinct axis from Brouillette (2025)'s own 2%–27% robustness band over the returns-to-scale parameter: that band varies γ holding the specification linear, while this one varies where talent lands and how complementary the map is, holding $\gamma = \frac{1}{3}$; because the operator works through the same increasing-returns lever, the two compound. Where the human-capital moments of the broader growth accounting permit, the same insertion re-prices the 20–40% allocation result of Hsieh et al. (2019) in the identical way, turning its headline into a curve indexed by the sorting of reallocated talent.

A curve this contingent on empirical inputs owes the reader an explicit account of where it

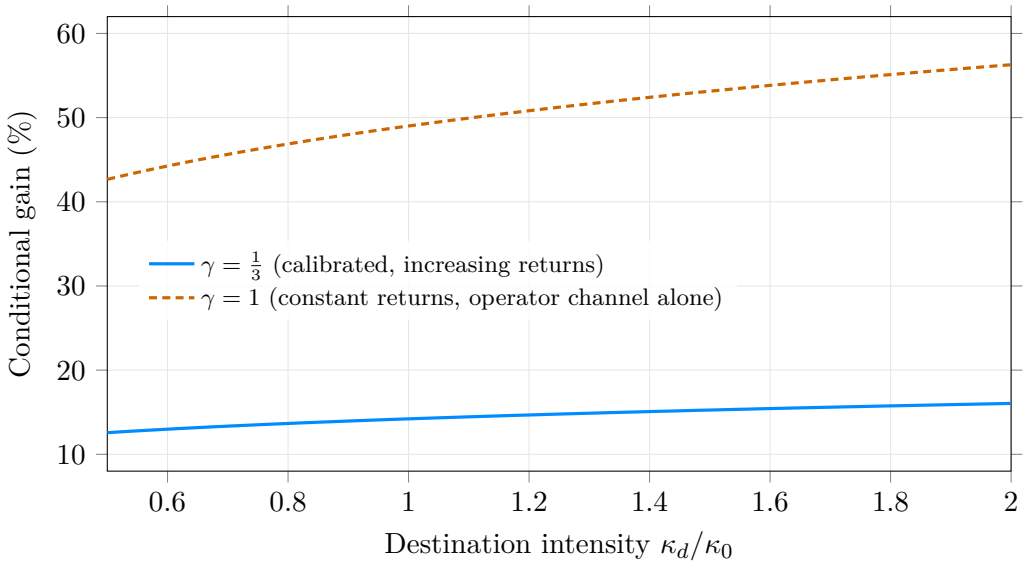


Figure 5: Operator channel versus increasing-returns amplifier. Both lines carry the same complementarity $b = 0.2$; the solid line uses the calibrated aggregator $\gamma = \frac{1}{3}$, the dashed line a constant-returns aggregator $\gamma = 1$ in which the gain is linear in re-weighted research input. The gap isolates how much of the widening is the increasing-returns lever raising any perturbation to a power, as distinct from the operator re-weight that generates the perturbation. Illustrative pending estimation.

could break. We give that account next – not as a retreat, but as the boundary of the claim, drawn alongside the frontier the claim bridges.

11 Scope, Limits, and the Frontier

The argument of this paper sits at the meeting point of three mature literatures, and its claim is made precise by saying exactly what each of them already owns and what is added here. The territory is bordered on three sides; the contribution is the bridge across them. The purpose of this section is to show that no single neighbor owns that bridge, and then to draw the boundary of the claim on every side without softening it.

What each neighbor owns, and what this paper adds

The cross-partial $\phi(T) = \partial^2 F/\partial T\,\partial C$ is not a new object. Cunha et al. (2010) and, before them, Cunha and Heckman (2007) established the technology of skill formation in which capability accumulates with self-productivity and dynamic complementarity, and they built the measurement apparatus by which the cross-partial of a latent-skill production function can be identified from noisy indicators; Heckman and Mosso (2014) synthesize the resulting account of human development. That literature owns the cross-partial, and it owns the discipline of estimating it. But it deploys the object almost exclusively for a within-person, life-cycle question — when and how much to invest in a child of a given endowment — and never carries it across into a cross-sectional allocation problem. What is new here is precisely that transport: the same cross-partial, embedded not in an individual's intertemporal investment problem but in the integral that aggregates realized capability across a reallocated population, where it acquires a consequence the within-person setting cannot see — a signed bias in the measured gain from reallocation. The measurement is

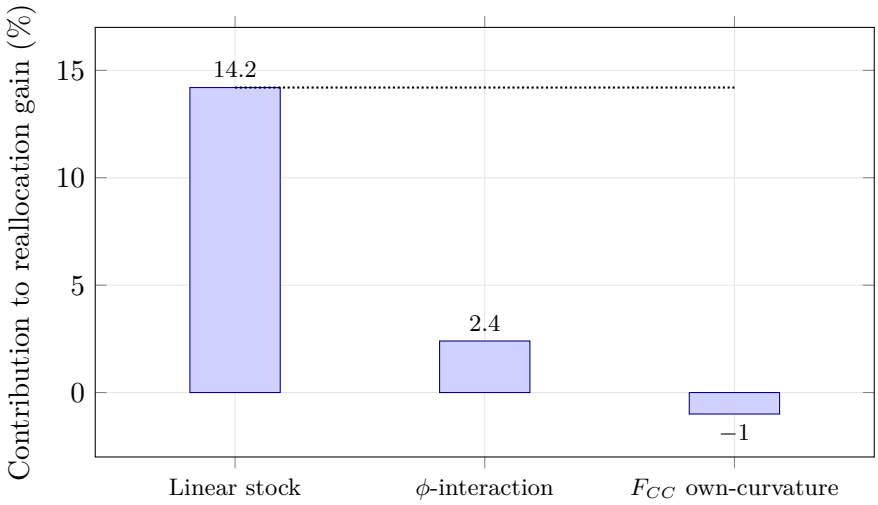


Figure 6: Decomposition of the reallocation correction at a representative upward move. The linear-stock term is the received 14.2%; the ϕ-interaction term is the operator correction, positive here under $\phi > 0$ and positive sorting $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) > 0$; the F_{CC} own-curvature term is diminishing returns to investment, present even under the separable null and reported separately so it is never mistaken for the operator. Bar heights are illustrative pending estimation; the dotted line marks the linear baseline.

theirs; the allocation consequence is the paper's.

The reallocation accounting itself belongs to Hsieh et al. (2019). Their finding that the movement of women and Black men into high-skill occupations accounts for a substantial share of U.S. growth over 1960–2010 is the canonical talent-misallocation number, and this paper does not contest the accounting that produces it. It signs the interaction component of the bias that their linear, multiplicative treatment of talent induces — a component their framework cannot represent, because talent enters as a Fréchet scalar multiplying an input rather than as an operator on the map by which capability is built. The same structural choice runs through the wider misallocation tradition: the spatial frictions of Hsieh and Moretti (2019), the selection-driven cross-country gaps of Lagakos and Waugh (2013), the financial frictions of Buera et al. (2011), and the general-equilibrium second-order corrections of Baqaee and Farhi (2020) all price reallocated talent at a first-order multiplicative value, so that the gain is linear in mass and is summarized, to first order, by a covariance of productivity with a wedge. The closest member of this tradition is Bryan and Morten (2019), who endogenize *which* talent moves and thereby capture a direction-of-reallocation object akin to the transfer structure of Section 6; but talent there remains a multiplicative stock, and the analysis stops short of any cross-partial between talent and the intensity of investment it meets at its destination.

The endogenous-growth treatment of talent is owned most fully by Akcigit et al. (2025), whose model of talent-heterogeneous innovation with financial frictions to human-capital investment delivers the ranking in which education subsidies can dominate research-and-development tax credits. This paper adds the curvature their model lacks: education there is a discrete gate with a fixed number of slots rather than a continuous investment that talent complements, so their counterfactuals remain linear in talent mass. Whether restoring the cross-partial overturns their ranking is a question this paper deliberately leaves open. A flip is plausible in principle — their idea-production aggregator represents talent as a fixed-elasticity scale and cannot price the convex return to deepening investment on infra-marginal high-talent students, and education policy plausibly moves talent up the investment gradient — but the direction cannot be asserted without re-solving their model

with the binary gate replaced by a continuous investment that talent complements, re-deriving the career-choice cutoff and the growth equation, and signing the sorting covariance in their own calibration. Until that is done the ranking is undetermined in either direction, and nothing here predicts its reversal. The broader talent-and-growth literature shares the same multiplicative form: Jones (1995) and Bloom et al. (2020) treat research input as a scalar feeding an idea-production function, Çelik (2023) studies the misallocation of innovative talent with talent as a type, and Einiö et al. (2022) owns the *direction* margin — which fields talent is pushed toward — rather than the curvature of the return to investing in a person.

The nearest rival is Prato (2025), and it must be named as such. Prato's model of the global race for talent contains a genuine accumulation cross-partial: talent evolves by learning-by-meeting, with higher-talent individuals learning relatively more from meeting a high-innovation peer, and this complementarity is fed into an H-1B allocation counterfactual that yields substantial long-run growth effects. It is the only existing paper that carries an accumulation cross-partial into an allocation counterfactual, and it falls short of the object studied here in exactly three respects. First, its second argument is *peer knowledge* — an equilibrium meeting draw — not own, policy-chosen investment, so its cross-partial governs sorting and migration, the question of which talent ends up where, rather than the return to investing in a given person; the operator $\phi(T) = \partial^2 F / \partial T \, \partial C$ requires C to be a controllable input, which a meeting draw is not. Second, conditional on talent, output in that model remains multiplicative: the curvature lives only in the law of motion for peer learning, not in the technology that maps capability into product. Third, the gain is never decomposed into a linear-stock term, an own-curvature term, and an interaction term, and no bias on a linear number is signed. The distance is therefore one argument wide — the substitution of own investment for the peer draw, together with the decomposition that signs the bias — and that argument is the contribution.

Within the body of quasi-experimental evidence, Brouillette's lost Marie Curies (Brouillette, 2025) is not a rival but a host: its counterfactual is re-priced in Section 10 by inserting an accumulation stage between talent and realized inventive capability, with the fourteen-percent headline recovered exactly in the no-curvature limit. The remaining quasi-experiments are read as reduced-form evidence for the two signatures of the operator. The global randomized trial of Agarwal and Gaule (2026) supplies the cleanest evidence on the rate operator, in the form of a treatment effect that is far larger for above-median than for below-median baseline ability; the directed-exposure findings of Bell et al. (2019) and Novosad et al. (2026) supply evidence on the transfer kernel, the direction along which exposure is productive. The acceleration and gifted-education results of Card and Giuliano (2024) and Card et al. (2024) corroborate the rate channel under the stated reading that a group gap proxies a talent-relevant accumulation characteristic rather than a preference shift orthogonal to accumulated capability.

Two further papers mark the far end of the spectrum, where the cross-partial is set to zero outright. Bennedsen et al. (2025) measure talent as a high-school grade-point rank — a fixed scalar set once and carried forward — and Amanzadeh et al. (2024) measure it as a time-invariant worker wage fixed effect; in both, talent is a stock, and $\phi \equiv 0$ by construction. Amanzadeh et al. (2024) comes closer than any other stock-based paper to the rate operator, because its location-dependent returns — a year of U.S. labor-market experience worth on the order of fifty-nine to two hundred and four percent more than a year of home-market experience — are exactly an environment-dependent accumulation rate. What keeps it on the near side of the boundary is that those returns are assumed uniform across individuals: the individual-specific return is not estimated, as the authors themselves acknowledge, so the design cannot separate a talent-by-environment complementarity from a pure repricing of fixed traits. This paper claims the cross-partial these papers leave unmeasured, not their wage results.

It is worth saying plainly that the complementarity literature is systematically richer than the allocation models admit. The most explicit talent-by-investment complementarity design in this body of work is Calaway (2026), whose static Cobb–Douglas specification of realized talent in latent exceptional ability and mentor inputs finds that ordinary teachers more than double the number of identified top students — a complementarity between a talent endowment and an investment input, estimated rather than assumed, even if cast in a static rather than a dynamic form. Jin et al. (2025) document a talent-by-talent supermodularity in production, native–immigrant founding teams outperforming both pure types; and Cohodes et al. (2022), Fry and Glennon (2025), and Ainsworth et al. (2025) bear on the investment and direction channels through which capability is built and identified. The skill-formation literature furnishes the closest methodological precedents for what is required here: Aizer and Cunha (2012) document parental reinforcement — the tendency to invest more in higher-endowed children precisely because of complementarity — which is the household analogue of the positive sorting on which the sign of the bias turns; and Houmark et al. (2024) estimate a cross-partial by interacting a genetic endowment measured prior to investment with birth order, the design most closely aligned with the measurement guardrail of Section 8. The attendant cautions are equally instructive: Agostinelli and Wiswall (2025) show that re-normalizing skills each period can silently impose separability and assume the curvature away, and Muff and Keller (2015) show that interaction coefficients can be biased *away* from zero when covariates carry heteroskedastic measurement error — both of which discipline the measurement strategy rather than threaten the claim. Taken together, the evidence that talent and investment interact is abundant; what is missing is its transport into the allocation integral. The territory is bordered by the measurement apparatus of Cunha, Heckman, and Schennach, by the reallocation accounting of Hsieh, Hurst, Jones, and Klenow, and by the talent-growth models of Akcigit, Pearce, and Prato and the global-talent-race model of Prato—and no single neighbor owns the bridge between them.

The boundary of the claim

Each of the load-bearing limits of the argument marks the edge of a well-bounded claim rather than a retreat from it, and stating each precisely is what keeps the claim falsifiable.

The cross-partial is identified, in the evidence assembled here, among Olympiad-nominated and gifted-cutoff populations — the upper reaches of the ability distribution. Extending the estimated curvature to the broad reallocation the macro accounting contemplates is therefore an assumption, not a result, and the honest deliverable of Section 10 is for this reason a *bounded curve* rather than a single corrected headline. Bounding the curve by the range of ϕ does not close the gap if the entire range is itself estimated at the tail. The route off the tail is the population-register design of Section 8 — an early, exposure-free ability measure drawn from administrative records and merged to later output, capable of estimating ϕ across the distribution rather than only at its top. That register is not an optional refinement; it is what the population-wide claim requires.

Proposition 1 signs the bias of the linear number, but it signs only the *marginal* bias. For large discrete reallocations the higher-order terms of the path integral that underlies the proposition cease to be negligible, and the sign of the interaction can reverse if the cross-partial itself changes sign across the talent range — complementarity at moderate talent giving way to substitutability in the extreme tail, so that a reallocation straddling the sign change carries an aggregate interaction opposite to the locally evaluated cross-partial. The global, non-local version of the result, the exact conditions under which the marginal sign survives, and a worked numerical instance in which it reverses are developed in Section A. The proposition is the analytic anchor of the argument, not the whole of it.

The direction of the bias depends on the sign of the sorting covariance $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$ between

the talent of movers and the investment intensity of where they land, and that covariance is an equilibrium object, not a theorem. It is positive under the natural reading that frictions disproportionately keep high-talent members of under-exposed groups *out* of high-intensity destinations, so that removing those frictions moves talent up the gradient; but it could be negative if freed talent is instead absorbed into low-intensity mass employment. Rather than a weakness, this is the contribution stated in operational terms: there is a sign to estimate, and it is estimable, where the linear literature did not recognize that a sign existed at all. The discipline this imposes on the rhetoric is exact. Absent a measured covariance, the honest claim is that the linear number is biased by a signable amount of as-yet-undetermined direction; it is the curve of Section 10, fed by a measured covariance, that converts a signable direction into a signed one.

Finally, the paper does not resolve whether talent exists in the never-exposed, and this is a virtue of its design rather than a gap in it. Like the inventor-exposure literature, the analysis cannot observe the latent talent of those never given the opportunity to realize it. But it deliberately does not need to. The entire claim lives in the interaction, not the level: *given* that a measured characteristic modulates the accumulation map — which the test of Section 9 adjudicates — the reallocation arithmetic is biased with sign $\text{sign}(\phi) \cdot \text{sign}(\text{Cov}_{\text{moved}}(T, \kappa))$, and that bias is signable from observed movers alone. It is exactly this restriction of the claim to the interaction that frees the result from any assertion about latent talent in the unexposed.

What remains, having drawn the boundary on every side, is to state in one line what the whole apparatus reduces to.

12 Conclusion

The distinction at the center of this paper can be stated in a single breath. A stock scales one undifferentiated input: more talent is more of the same thing, and its contribution to output is read off a constant marginal price. An operator does something different in kind — it bends the curve along which capability accumulates, raising the return to investing in a person in proportion to how much talent that person carries, and the bending *is* the cross-partial $\phi(T) = \partial^2 F / \partial T \, \partial C$. The misallocation-of-talent literature delivers its disciplined headline — the output recovered by moving talent to where it belongs — as a linear functional of reallocated talent mass, and it does so for exactly one reason that is never argued, only built into the technology: talent is entered as a multiplicative stock, which is to set the cross-partial to zero by construction.

Once the cross-partial is restored, the gain from reallocation no longer reduces to that single number. It separates into the linear-stock term the received literature already prices, an own-curvature term governed by F_{CC} — the diminishing return to investment at fixed talent, alive even under the separable null — and an interaction term equal to $\phi(\bar{T}) \cdot \text{Cov}_{\text{moved}}(T, \kappa)$. The two second-order terms are distinct objects and must be kept distinct: a nonlinearity in investment alone is $F_{CC} \neq 0$ with $\phi = 0$, and it does not vindicate the operator. It is the interaction term, and it alone, that carries the operator's signature, and through it the bias of the linear number acquires a definite sign,

$$\text{sign}\left(\text{bias of the linear-stock counterfactual}\right) \;=\; \text{sign}(\phi) \cdot \text{sign}(\text{Cov}_{\text{moved}}(T, \kappa)),$$

and nothing else. Misallocation arithmetic that ignores the curvature is not merely imprecise; it is biased, and biased in a direction that named, estimable quantities determine. When talent sorts up the investment gradient and the cross-partial is positive, the linear accounting *understates* the gain from moving talent toward richer environments and overstates the gain from moving it the other way — and it does so in a concave, not a linear, way.

The reach of the critique is therefore also the measure of its discipline. The entire reconstruction reduces to a single coefficient. The whole received literature is reframed as the null hypothesis $\phi = 0$, the special case in which talent and investment are separable and the linear-stock arithmetic is exactly correct. The constant-elasticity rendering of the accumulation technology makes this concrete and testable: there $\operatorname{sign}\phi = \operatorname{sign}(1-\rho)$, so "talent is an operator, not a stock" is the falsifiable claim $\rho < 1$, talent and investment are complements, and the separability the literature assumes is the measure-zero knife-edge $\rho = 1$, at which talent and investment become perfect substitutes, the elasticity of substitution diverges, and the technology collapses to the separable additive form $\omega T + (1-\omega)C$ in which talent no longer bends the return to investment. A claim that turns on whether a structural elasticity sits at a single point, rather than on either side of it, is a claim that can be wrong, and the paper stakes itself on that.

What remains is an empirical program whose shape the theory fixes. It calls for a talent measure collected before and independently of investment — exposure-free by construction, drawn from administrative registers rather than from any realized output that already embeds the investment whose return is in question — so that the cross-partial can be estimated off the tail and across the distribution. It calls for the global numerical re-solution of the host counterfactuals with the accumulation stage inserted and every other target held fixed, turning the illustrative band of Section 10 into a measured curve. And it calls, above all, for the measured sorting covariance that converts the direction of that curve from signable to signed: the curve already tells us there is a sign; the covariance tells us which one.

The cleanest statement the paper can make is also its terminal one. If the talent-margin interaction is a precise zero once talent is measured separably from investment and the own-curvature surface F_{CC} is absorbed — if instrumented investment does not interact with pre-determined ability — then talent and investment are separable in the cross-partial sense, the reallocation gain carries no operator correction, and the linear-stock arithmetic is vindicated. The operator critique is, in that event, empirically void. This is not a hedge but the form of the contribution: the received account of the misallocation of talent is exactly the hypothesis that the curvature of the accumulation map is zero, and this paper makes that hypothesis something that can be put to the test.

A The Formal Core

This appendix carries the mathematics of Section 7 to full rigor. It states the setup precisely, derives the three-term decomposition from first principles, gives the exact global version with its sign theorem and the channels through which the local sign can reverse, and works the two closed-form parametric families that make "operator, not stock" a parametric and testable claim.

A.1 Setup and conventions

The accumulation technology is $H_{t+1} = F(H_t, C; T)$, with the rate operator defined as the cross-partial in the talent and investment arguments, $\phi(T) \equiv F_{CT}$. A destination d is summarized by its investment intensity $\kappa_d \equiv \mathbb{E}[C \mid d]$. A marginal mass of talent T is reallocated from a low-intensity environment κ_0 to a high-intensity environment κ_1, with investment allocated in proportion to local intensity, so that the input applied to a mover traverses the interval $[\kappa_0, \kappa_1]$ in the second argument of F. Initial human capital H_t is held fixed during the move, so we suppress it and write $F(\kappa; T) \equiv F(H_t, \kappa; T)$, with F_C, F_{CC}, F_{CT} understood as partial derivatives in the investment argument. The *linear-stock counterfactual* prices every mover at a single constant marginal value, $F_C(\bar{\kappa}; \bar{T})$, evaluated at the mass-weighted anchor $(\bar{\kappa}, \bar{T})$ of the move. Throughout, G_{moved} denotes the talent distribution restricted to the reallocated subpopulation.

A.2 Local derivation and the three-term decomposition

Step 1 — the per-mover gain is an exact path integral. Holding H_t fixed and applying the fundamental theorem of calculus to the investment argument,

$$\Delta F(T) = F(\kappa_1; T) - F(\kappa_0; T) = \int_{\kappa_0}^{\kappa_1} F_C(\kappa; T)\, d\kappa. \tag{15}$$

This requires only that F be C^1 in its investment argument on $[\kappa_0, \kappa_1]$. No expansion has been made; (15) is exact and global, and it reproduces (6) of Section 7.

Step 2 — the single approximation. The only approximation in the local proposition is a first-order joint Taylor expansion of the integrand F_C in (κ, T) about $(\bar{\kappa}, \bar{T})$:

$$F_C(\kappa; T) \approx F_C(\bar{\kappa}; \bar{T}) + F_{CC}(\bar{\kappa}; \bar{T})\,(\kappa - \bar{\kappa}) + F_{CT}(\bar{\kappa}; \bar{T})\,(T - \bar{T}). \tag{16}$$

The expansion discards third-order terms in F — F_{CCC}, F_{CTT}, and F_{CCT} — which are retained exactly in Section A.3. It is used once, inside the integral (15), and nowhere else; in particular the path integral itself is never approximated.

Step 3 — integrate the three pieces. Substituting (16) into (15) and integrating, with the anchor set at $\bar{\kappa} = \kappa_0$ (any interior anchor changes only the third-order remainder),

$$\int_{\kappa_0}^{\kappa_1} F_C(\bar{\kappa}; \bar{T})\, d\kappa = F_C(\bar{\kappa}; \bar{T})\,(\kappa_1 - \kappa_0) \qquad \text{(linear-stock)}, \tag{17}$$

$$\int_{\kappa_0}^{\kappa_1} F_{CC}(\bar{\kappa}; \bar{T})\,(\kappa - \kappa_0)\, d\kappa = \tfrac{1}{2}\, F_{CC}(\bar{\kappa}; \bar{T})\,(\kappa_1 - \kappa_0)^2 \qquad \text{(own-curvature)}, \tag{18}$$

$$\int_{\kappa_0}^{\kappa_1} F_{CT}(\bar{\kappa}; \bar{T})\,(T - \bar{T})\, d\kappa = \phi(\bar{T})\,(T - \bar{T})\,(\kappa_1 - \kappa_0) \qquad \text{(interaction, per mover)}. \tag{19}$$

The own-curvature term (18) contains no ϕ. It is investment-by-investment curvature — the degree to which the marginal return to investment diminishes at *fixed* talent — and it is present whenever investment has diminishing returns, including under the separable null $\phi \equiv 0$. Carrying it explicitly, rather than tacitly setting $F_{CC} = 0$ by holding F_C constant along the path, is the decisive technical discipline of the decomposition: it prevents the interaction term from absorbing ordinary diminishing returns and being misread as evidence of complementarity.

Step 4 — the covariance collapse. A mover of talent T lands not at a common destination but at the intensity $\kappa(T)$ that the reallocation delivers to her, displacing $\kappa(T) - \bar{\kappa}$; the per-mover displacement is therefore $\kappa(T) - \bar{\kappa}$, not the common $\kappa_1 - \kappa_0$ that appears in (19). Aggregating (19) against G_{moved}, and using $\mathbb{E}_{\text{moved}}[T - \bar{T}] = 0$,

$$\int \phi(\bar{T})\,(T - \bar{T})\big(\kappa(T) - \bar{\kappa}\big)\, dG_{\text{moved}}(T) = \phi(\bar{T})\ \mathrm{Cov}_{\text{moved}}(T, \kappa). \tag{20}$$

A homogeneous reallocation, sending every mover to the same κ_1, makes the aggregate interaction vanish: then $\kappa(T) - \bar{\kappa}$ is a constant pulled outside the integral, and the centered talent integrates to zero. The interaction is thus an artifact neither of the size of the move nor of its average destination, but of the *sorting* of talent across destinations within the move.

The decomposition. Collecting (17), (18), and (20), the change in aggregate human capital per unit mass reallocated is

$$\boxed{d\mathcal{H} = \underbrace{F_C(\bar{\kappa};\bar{T})\,(\kappa_1-\kappa_0)}_{\text{linear-stock}} + \underbrace{\tfrac{1}{2}\,F_{CC}(\bar{\kappa};\bar{T})\,(\kappa_1-\kappa_0)^2}_{\text{own-curvature}} + \underbrace{\phi(\bar{T})\ \mathrm{Cov}_{\text{moved}}(T,\kappa)}_{\text{interaction}}} \tag{5'}$$

which is (5) of Proposition 1.

Units of the interaction term. The interaction admits two equivalent forms that differ only in the units of the displacement. With $\tilde{\kappa}$ the displacement normalized by the move length, so that $\mathrm{Cov}_{\text{moved}}(T,\kappa) = (\kappa_1-\kappa_0)\ \mathrm{Cov}_{\text{moved}}(T,\tilde{\kappa})$, the interaction is $\phi(\bar{T})\ \mathrm{Cov}_{\text{moved}}(T,\kappa)$ in raw intensity and $\phi(\bar{T})\,(\kappa_1-\kappa_0)\ \mathrm{Cov}_{\text{moved}}(T,\tilde{\kappa})$ in normalized displacement. The two expressions are identical. The raw form is preferred in the text because it presents the interaction directly as a covariance between two measured quantities — a mover's talent and the intensity of her destination — and because all sign statements are invariant to the convention: $\kappa_1-\kappa_0>0$ for an upward move, and $\operatorname{sign}\mathrm{Cov}_{\text{moved}}(T,\kappa) = \operatorname{sign}\mathrm{Cov}_{\text{moved}}(T,\tilde{\kappa})$.

Corollaries. By inspection of (5): if $\phi>0$ and $\mathrm{Cov}_{\text{moved}}(T,\kappa)>0$ the interaction is positive, so the linear-stock account understates upward moves and overstates downward ones (Corollary 1); the sign of the interaction component of the bias is $\operatorname{sign}(\phi)\cdot\operatorname{sign}(\mathrm{Cov}_{\text{moved}}(T,\kappa))$ (Corollary 2); and the gain reduces to the linear-stock term if and only if $\phi\equiv 0$ and $F_{CC}\equiv 0$, a pure own-curvature correction surviving under $\phi\equiv 0$ alone (Corollary 3).

A.3 The global version

The local proposition signs the marginal bias. For large discrete reallocations the approximation (16) is dropped and the gain is written exactly. Aggregating the exact path integral (15) over the movers, each traversing her own path $[\kappa_0(T),\kappa_1(T)]$,

$$\Delta\mathcal{H} = \int_T \int_{\kappa_0(T)}^{\kappa_1(T)} F_C(\kappa;T)\,d\kappa\,dG_{\text{moved}}(T). \tag{21}$$

To isolate the part of (21) that vanishes under separability, write the marginal return as its value at the anchor talent plus the talent integral of the cross-partial, $F_C(\kappa;T) = F_C(\kappa;\bar{T}) + \int_{\bar{T}}^{T}\phi(\kappa;\tau)\,d\tau$. The first piece is independent of the talent of the mover and contributes the own (talent-free) part of the gain; the second is the *exact interaction component*,

$$\mathcal{I}_{\text{exact}} = \int_T \int_{\kappa_0(T)}^{\kappa_1(T)} \left(\int_{\bar{T}}^{T} \phi(\kappa;\tau)\,d\tau\right) d\kappa\,dG_{\text{moved}}(T). \tag{22}$$

Equation (22) is exact: it retains all higher-order own-curvature, all curvature of ϕ in talent, and all dependence of ϕ on the investment level, inside $\phi(\kappa;\tau)$ and the path limits. It reduces to $\phi(\bar{T})\ \mathrm{Cov}_{\text{moved}}(T,\kappa)$ when ϕ is constant over the traversed support and the paths are short, recovering (5).

Theorem 1 (Global sign of the interaction)**.** *Suppose that along every mover's path $\kappa\in[\kappa_0(T),\kappa_1(T)]$ and for every τ between $\bar{T}$ and T the cross-partial $\phi(\kappa;\tau)$ has constant sign $\sigma\in\{+,-\}$ (sign-definiteness over the traversed support), and that movers sort positively, $T\geq\bar{T}\iff\kappa_1(T)\geq\kappa_0(T)$. Then $\mathcal{I}_{\text{exact}}$ has sign σ for upward reallocations.*

Table 2: An illustrative reversal: local complementarity, global substitution.

Group	T	mass	move $\kappa_0 \to \kappa_1$	$\phi(T) = 3 - 2T$
L (bulk)	1.2	0.95	$1.0 \to 2.0$	$+0.6$
H (tail)	3.2	0.05	$1.0 \to 4.0$	-3.4

Proof. Fix a mover of talent T. Under sign-definiteness, the inner talent integral $\int_{\bar{T}}^{T} \phi(\kappa;\tau)\, d\tau$ has sign $\sigma \cdot \mathrm{sign}(T-\bar{T})$, because integrating a function of constant sign σ from $\bar{T}$ to T yields a quantity of sign σ when $T > \bar{T}$ and of sign $-\sigma$ when $T < \bar{T}$. The middle integral over κ runs from $\kappa_0(T)$ to $\kappa_1(T)$; under positive sorting this interval is oriented upward ($\kappa_1(T) \geq \kappa_0(T)$) exactly when $T \geq \bar{T}$ and downward exactly when $T < \bar{T}$, so the κ-integral preserves the sign of the integrand for $T > \bar{T}$ and flips it for $T < \bar{T}$. The two sign flips for $T < \bar{T}$ cancel, so the integrand of the outer integral has sign σ for every mover, whether above or below the anchor talent. The integral of a function of constant sign σ against the probability measure G_{moved} has sign σ. □

Restricting attention to the interaction component, the linear-stock counterfactual understates the gain whenever $\mathcal{I}_{\text{exact}} > 0$, for which Theorem 1 gives a sufficient condition: ϕ sign-positive over the traversed support, and positive sorting holding all along the path rather than merely at the anchor. The hypotheses of the theorem name exactly what must hold for the local conclusion of Corollary 2 to survive globally, and they expose three channels through which it can fail.

(R1) The cross-partial changes sign across the support. If ϕ is complementary at moderate talent but turns substitutable in the extreme tail, $\int_{\bar{T}}^{T} \phi\, d\tau$ ceases to be sign-definite. A reallocation whose paths straddle the sign change can carry an aggregate interaction opposite to the sign of the locally evaluated $\phi(\bar{T})$.

(R2) The sorting covariance flips for large moves. Even with $\phi > 0$ everywhere, if freed high-talent movers are absorbed into low-intensity mass employment for large reallocations, the path covariance turns negative and the global interaction reverses.

(R3) Higher-order own-curvature dominates. If $|F_{CCC}|$ or $|F_{CCT}|$ is large relative to ϕ, the third-order remainder discarded in (16) swamps the interaction, and the measured bias is governed by own-curvature rather than by the talent margin — a further reason to calibrate F_{CC} separately.

A numerical reversal. The first channel can be exhibited in closed form, with own-curvature switched off so that the interaction is isolated. Let

$$F(\kappa;T) = \nu\, T + \kappa\, s(T), \qquad s(T) = -(T-1.5)^2. \tag{23}$$

Then $F_C = s(T)$, $F_{CC} \equiv 0$, and $\phi(T) = s'(T) = 3 - 2T$, so $\phi > 0 \iff T < 1.5$. Consider two groups of movers, both reallocated upward, in Table 2.

The mass-weighted anchor talent is $\bar{T} = 0.95(1.2) + 0.05(3.2) = 1.30 < 1.5$, so the locally evaluated cross-partial is $\phi(\bar{T}) = 3 - 2(1.30) = +0.40 > 0$, and the sorting is positive (the higher-talent group makes the longer upward move). The *local* prediction of Corollary 1 is therefore that the linear-stock counterfactual understates the gain. The *exact* interaction is computed from (22); with $F_{CC} \equiv 0$ the marginal return $F_C = s(T)$ does not depend on κ, so the inner κ-integral collapses

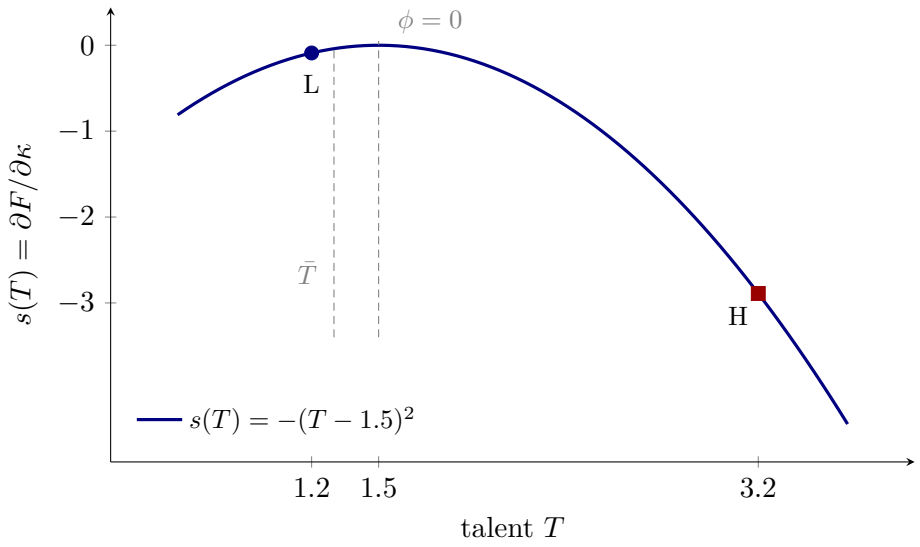


Figure 7: The reversal of Table 2. The marginal return to investment $s(T)$ peaks at $T = 1.5$, where the cross-partial $\phi = s'(T)$ crosses zero. The bulk group L sits on the rising arm ($\phi > 0$) and the tail group H on the falling arm ($\phi < 0$). The mass-weighted anchor $\bar{T} = 1.30$ lies below 1.5, so the locally evaluated $\phi(\bar{T}) = +0.40$ predicts understatement, yet the exact interaction (24) is negative: the long upward moves of the high-talent tail, deep in the $\phi < 0$ region, dominate. Values illustrative.

to the move length $\Delta\kappa_g$, and the exact interaction reduces to $\sum_g m_g \, \Delta\kappa_g \, [\, s(T_g) - s(\bar{T}) \,]$. Using $s(1.2) = -0.09$, $s(3.2) = -2.89$, and $s(1.30) = -0.04$,

$$\mathcal{I}_{\text{exact}} = 0.95\,(1)\,(-0.09 + 0.04) + 0.05\,(3)\,(-2.89 + 0.04) = -0.0475 - 0.4275 = -0.475. \tag{24}$$

The exact interaction is *negative*: the linear-stock counterfactual in fact overstates the gain. The five-percent-mass high-talent movers sit deep in the region where $\phi < 0$, and their long upward paths dominate the interaction integral, overturning the sign that the locally evaluated $\phi(\bar{T}) = +0.40$ would predict. The reversal is robust to the size of the tail: for tail mass in $[0.05, 0.15]$ the anchor stays below 1.5, so $\phi(\bar{T}) > 0$ throughout while $\mathcal{I}_{\text{exact}}$ remains negative. The figures here are illustrative; they demonstrate the structure of channel (R1), not an empirical magnitude. Figure 7 plots the construction.

A.4 Closed-form parametric special cases

Two parametric families show that the decomposition is not an abstraction. The first has a globally signed cross-partial and so cannot reverse through channel (R1); the second makes "operator, not stock" a single parametric hypothesis.

(A) Additive Cobb–Douglas. Let

$$F = H + A\,T^a C^b, \qquad A > 0,\ a, b > 0. \tag{25}$$

The relevant derivatives in the investment argument are

$$F_C = A\,b\,T^a C^{\,b-1}, \qquad F_{CC} = A\,b(b-1)\,T^a C^{\,b-2}, \qquad \phi = F_{CT} = A\,a\,b\,T^{\,a-1} C^{\,b-1}.$$

Hence $\phi > 0$ for all $a, b > 0$: the cross-partial is globally positive and channel (R1) cannot arise within this family. The exact per-mover gain is $\Delta F(T) = A\,T^a(\kappa_1^b - \kappa_0^b)$. Evaluating each term of Proposition 1 at the anchor $\bar{\kappa} = \kappa_0$,

$$\begin{aligned} \text{linear-stock} &= A\,b\,\bar{T}^a\kappa_0^{b-1}(\kappa_1 - \kappa_0), \\ \text{own-curvature} &= \tfrac{1}{2}\,A\,b(b-1)\,\bar{T}^a\kappa_0^{b-2}(\kappa_1 - \kappa_0)^2 \quad (\text{negative iff } b < 1), \\ \text{interaction} &= A\,a\,b\,\bar{T}^{\,a-1}\kappa_0^{b-1}\,\mathrm{Cov}_{\text{moved}}(T, \kappa) > 0 \quad (\text{under positive sorting}). \end{aligned}$$

The own-curvature term is negative precisely when investment exhibits diminishing returns ($b < 1$), confirming that it is a genuine concavity correction independent of the talent margin. Taking the ratio of the interaction to the linear-stock term, the interaction equals the fraction

$$\frac{a\ \mathrm{Cov}_{\text{moved}}(T, \kappa)}{\bar{T}\,(\kappa_1 - \kappa_0)} \tag{26}$$

of the linear-stock term. For $a = 1$ and a one-standard-deviation sorted move, this fraction is first-order in the data, not a negligible correction: the operator revises the headline by an amount of the same order as the headline itself.

(B) CES. Let

$$H = \left(\omega\,T^\rho + (1-\omega)\,C^\rho\right)^{1/\rho}, \qquad \omega \in (0,1). \tag{27}$$

Writing $\Omega \equiv \omega\,T^\rho + (1-\omega)\,C^\rho$, the first derivative in investment is $F_C = (1-\omega)\,C^{\rho-1}\,\Omega^{1/\rho-1}$, and differentiating in talent,

$$\phi = F_{CT} = \omega(1-\omega)(1-\rho)\,T^{\rho-1}C^{\rho-1}\,\Omega^{1/\rho-2}. \tag{28}$$

Every factor in (28) is positive except $(1-\rho)$, so

$$\boxed{\operatorname{sign}\phi = \operatorname{sign}(1-\rho)} \tag{29}$$

talent and investment are complements for $\rho < 1$, the cross-partial vanishes at $\rho = 1$, and they are substitutes for $\rho > 1$. The elasticity of substitution is $1/(1-\rho)$. The linear-stock literature is exactly the knife-edge $\rho = 1$, at which (27) collapses to the separable additive form $H = \omega\,T + (1-\omega)\,C$ and talent is a separable scale on one undifferentiated input. The central claim of the paper is therefore parametric and testable: "talent is an operator, not a stock" is the hypothesis $\rho < 1$, and separability is the measure-zero knife-edge $\rho = 1$. At $T = C = \omega = \frac{1}{2}$, (28) gives

$$\phi = +1.5,\ +1.0,\ +0.5,\ 0,\ -0.25,\ -0.5 \qquad \text{for} \qquad \rho = -2,\ -1,\ 0,\ 1,\ 1.5,\ 2,$$

with the value $+0.5$ at $\rho \to 0$ matching the Cobb–Douglas cross-partial $\frac{1}{4}T^{-1/2}C^{-1/2}$. The Leontief limit $\rho \to -\infty$ recovers the perfect-complement form $H = \min[T, C]$, the strongest complementarity— structurally the same Leontief shape as the gifted-education production function $\min[C, N]$ of Card et al. (2024), in which the student-brought input N occupies the position of the talent argument here. With a single global ρ the cross-partial is sign-definite, so the local sign theorem of Theorem 1 holds globally; a CES reversal of the local result requires ρ itself to vary across the talent support, the parametric face of channel (R1).

B The Empirical Strategy

This appendix develops the estimation of Section 8 and the decision rule of Section 9 in full. It states the four estimands and their identifying restrictions, derives the leading-order behavior of the interaction coefficient under measurement error in the talent proxy, formalizes the guardrail that distinguishes the operator cross-partial from an input×input curvature, and states the decisive test as a decision rule on a single coefficient.

B.1 The four estimands

(a) The rate cross-partial $\phi(T)$. The primary object is estimated, in the lead design, from

$$H_{i,t+1} = \beta_0 + \beta_1 T_i + \beta_2 \hat{C}_i + \beta_3 \left(T_i \times \hat{C}_i\right) + \mathbf{x}_i'\gamma + u_i, \tag{30}$$

with T_i a continuous, pre-determined ability score and $\hat{C}_i$ the assignment-induced component of engagement. The two endogenous regressors $\{C_i,\ T_i \times C_i\}$ are instrumented by $\{Z_i,\ T_i \times Z_i\}$ through the first stage

$$C_i = \pi_0 + \pi_1 Z_i + \pi_2 T_i + \pi_3 \left(T_i \times Z_i\right) + \mathbf{x}_i'\rho + v_i, \tag{31}$$

in which T_i, exogenous and pre-determined, is its own instrument. Then β_3 is the sample analogue of ϕ. Identification rests on three restrictions. *Exclusion:* $Z_i \perp u_i \mid \mathbf{x}_i$ and $(T_i \times Z_i) \perp u_i \mid \mathbf{x}_i$—assignment affects H only through the engagement it induces and does not shift the talent-margin return except through that engagement, an encouragement-only and not a content-bearing design. *Relevance:* $\pi_3 \neq 0$—assignment must move engagement *differentially across the ability range*; we report the Sanderson and Windmeijer (2016) first-stage F for the interaction equation specifically, conditioning out the other endogenous regressor. *Pre-determination:* T_i is realized before any C and is not a function of the C-history. For discontinuity hosts $Z_i = \mathbf{1}[r_i \geq c]$; with T_i the running variable held fixed at c the interaction is read off a group marker and is corroborating, while with T_i a second orthogonal ability axis it is the cross-partial proper. The trial of Agarwal and Gaule (2026) is primary precisely because its interacting variable is a continuous talent level. To trace curvature in talent—and to probe the sign-change region where local and global conclusions can diverge—we allow $\phi(T) = \phi_0 + \phi_1 T$ via an added $T_i^2 \times \hat{C}_i$ term.

(b) The own-curvature F_{CC}. Within narrow talent bins b the dose-response is estimated flexibly,

$$H_{i,t+1} = \sum_b \mathbf{1}[T_i \in b]\left[\alpha_b + g_b(\hat{C}_i)\right] + \mathbf{x}_i'\gamma + e_i,$$

with g_b a quadratic, restricted cubic spline, or local polynomial, so that F_{CC} at bin b is $\partial^2 g_b / \partial \hat{C}^2$. The operator test then uses the absorption specification

$$H_{i,t+1} = \beta_0 + \beta_1 T_i + s(\hat{C}_i) + \beta_3\left[(T_i - \bar{T}) \times \hat{C}_i\right] + \mathbf{x}_i'\gamma + u_i, \tag{32}$$

testing $H_0 : \beta_3 = 0$ conditional on the common-in-talent surface $s(\cdot)$ saturated; each term in $s(\hat{C})$ is instrumented as in (31), and centering at $\bar{T}$ orthogonalizes the interaction from the main investment curvature. A nonzero F_{CC} with $\beta_3 = 0$ is diminishing returns to investment, not the operator.

(c) The transfer kernel $\Psi_T(k, k^\star)$. Over (individual × candidate-direction) cells with distance d_{ik} between exposure direction k and productive direction $k^\star$,

$$\mathbb{P}(\text{invent in } k)_i = \lambda_0 + \lambda_1 \mathbf{1}[d_{ik} = 0] + h(d_{ik}) + \mathbf{x}_i'\gamma + \varepsilon_{ik}, \tag{33}$$

with Ψ_T concentrated at $k = k^\star$ identified by the $d = 0$ versus $d = 1$ jump λ_1, holding the smooth proximity profile $h(\cdot)$ fixed, under the condition that the specific direction is as-good-as-random given overall exposure intensity. The parent–child field-match permutation of Novosad et al. (2026) supplies the placebo: $\lambda_1 \to 0$ under random reassignment of directions. The kernel corroborates that accumulation is talent-modulated; it does not identify ϕ and does not enter the re-pricing.

(d) The re-priced reallocation-gain functional. From Proposition 1 we plug $\widehat{\phi}$ from (30) or (32), $\widehat{F_{CC}}$ from the within-bin dose-response, and the separately measured sorting covariance $\widehat{\mathrm{Cov}_{\text{moved}}}(T, \kappa)$—taken from realized assignment in the host calibration, not from the experiment—into the decomposition, reporting the gain as a curve in κ bracketed by the range of $\widehat{\phi}$. The signed interaction component of the bias, $\mathrm{sign}(\widehat{\phi}) \cdot \mathrm{sign}(\widehat{\mathrm{Cov}_{\text{moved}}}(T, \kappa))$, is reported separately from the F_{CC} correction so the operator channel is never conflated with diminishing returns. Curve standard errors follow from the delta method or a bootstrap that propagates $\widehat{\phi}$, $\widehat{F_{CC}}$, and $\widehat{\mathrm{Cov}_{\text{moved}}}(T, \kappa)$ jointly.

B.2 Measurement error in the interaction

Let the proxy be

$$\widetilde{T}_i = T_i + L_a(\text{env}_i) + \epsilon_i,$$

with L_a an institution-embedded component correlated with investment and ϵ_i classical, and let the true model be $H = \beta_0 + \beta_1 T + \beta_2 C + \phi\,(T \times C) + u$. We regress H on $(\widetilde{T}, C, \widetilde{T} \times C)$ and trace the leading-order behavior of the interaction coefficient.

(i) Classical centered error. With ϵ classical and the regressors centered and jointly normal, $\widetilde{T}C = TC + \epsilon C$, so the interaction is the least reliable regressor. By the product-reliability result of Busemeyer and Jones (1983), $\mathrm{rel}(T{\cdot}C) \approx \mathrm{rel}(T) \cdot \mathrm{rel}(C)$, whence

$$\text{plim } \hat{\beta}_3 = \phi \cdot \mathrm{rel}(T{\cdot}C) < \phi \quad \text{in magnitude,}$$

an attenuation *toward zero*. This is benign: a rejection of $\phi = 0$ survives it, and the protocol concedes the benchmark rather than overstating it.

(ii) Environment-correlated component. With $L_a \neq 0$, $\widetilde{T}C = TC + L_aC + \epsilon C$, and the L_aC term is non-classical. To leading order

$$\text{plim } \hat{\beta}_3 \approx \phi \cdot (\text{attenuation} < 1) + \underbrace{\frac{\mathrm{Cov}(L_aC,\ H \mid C, \widetilde{T})}{\mathrm{Var}(\widetilde{T}C \mid C, \widetilde{T})}}_{\text{contamination, signed}},$$

with contamination sign $\mathrm{sign}(\mathrm{Cov}(L_a, C)) \cdot \mathrm{sign}(F_C)$. Since $F_C > 0$ and institution-embedded talent components are positively correlated with investment intensity, the generic contamination is *positive*, biasing $\hat{\beta}_3$ away from zero and toward the operator conclusion. Attenuation toward zero fails; this is the self-fulfilling, dangerous case.

(iii) Heteroskedastic error. With $\mathrm{Var}(\epsilon \mid C) = \sigma_\epsilon^2(C)$, a term signed by $\partial\sigma_\epsilon^2/\partial C$ is added: a reliability gap that rises with investment manufactures a spurious interaction of sign $\mathrm{sign}(\partial\sigma_\epsilon^2/\partial C)$ even at $\phi = 0$.

The non-centering channel. When talent is not centered, $\mathbb{E}[T] \neq 0$ gives $\mathrm{Cov}(\widetilde{T}, \widetilde{T}C) \neq 0$, so error in the *level* smears across the level and interaction coefficients: a genuine main effect can masquerade as ϕ, and an irrelevant $T \cdot C$ acquires a nonzero coefficient through correlation with the mismeasured level (Abel, 2018; Garber and Klepper, 1980; Griliches and Hausman, 1986). Under heteroskedastic covariate error the interaction attenuation can even reverse (Muff and Keller, 2015). The test is therefore valid only after centering, after instrumenting $T{\cdot}C$ with $T{\cdot}Z$, and after confirming that the discontinuity-local estimate survives.

The dangerous condition. Collecting the channels, $\hat{\beta}_3$ is biased away from zero exactly when

$$\boxed{\mathrm{Cov}(L_a, C) \cdot F_C - \frac{\partial\sigma_\epsilon^2}{\partial C} \cdot \mathbb{E}[T] > 0} \tag{34}$$

holding $F_C > 0$, with $\mathbb{E}[T]$ the uncentered level, since the non-centering channel is precisely the $\mathbb{E}[T] \neq 0$ effect. When the expression is positive the critique can be self-fulfilling; when negative the contamination is conservative and a rejection of $\phi = 0$ is a fortiori credible.

Signing the contamination. The sign of (34) is estimated rather than assumed. (1) A *proxy ladder* estimates $\hat{\phi}$ off measures ordered by exposure-freeness—sibling or conscription test, then uniformly administered test, then discontinuity-local running variable, then post-program grade. A monotone rise as the proxy becomes less exposure-free signs $\mathrm{Cov}(L_a, C) > 0$, the dangerous direction, and the exposure-free estimate is the conservative bound; a flat profile indicates L_a immaterial. The post-program-grade rung is excluded by construction from every reported estimate or bound of ϕ—it enters only as the contaminated anchor that calibrates the contamination slope, and any interval that includes it is non-compliant with the guardrail. (2) $\mathrm{Cov}(L_a, C)$ is estimated directly by projecting the proxy on environment instruments orthogonal to true talent at a cutoff where T is fixed; the fitted component is $\widehat{L_a}$. (3) A regression of squared proxy residuals at fixed T on C signs $\partial\sigma_\epsilon^2/\partial C$. (4) These plug into (34), and $\hat{\phi}$ is reported as a bound with an empirically established contamination direction, in the spirit of Black et al. (2000). *Structural alternatives.* The latent-factor and nonclassical-deconvolution machinery of Cunha et al. (2010) and Hu and Schennach (2008), applied to two or more pre-program ability indicators, estimates the interaction on the latent factor and dissolves product-reliability attenuation by construction; Lewbel (2012) supplies heteroskedasticity-based instruments where no external one exists; and the supermodularity inequalities of Athey and Stern (1998) test complementarity directly, with the caveat that several such procedures can share a bias.

B.3 The guardrail, formalized

The coefficient $\hat{\beta}_3$ estimates ϕ—and not an F_{CC}-type input×input interaction—if and only if the proxy satisfies both conditions. *Timing:* T_i is realized before any C in the estimated relationship. *Orthogonality:* $\widetilde{T}_i$ is not a function of the C-path, $\partial\widetilde{T}_i/\partial C_{i,s} = 0$ for all s, equivalently $L_a \perp \{C_{i,s}\}$. If either fails, $\widetilde{T} \times C$ is partly $C_{\text{past}} \times C_{\text{now}}$, an input×input interaction loading on F_{CC}, and the critique becomes circular—the empirical twin of equating realized H with talent. Four falsification tests guard the assumption. (1) *Cross-proxy exposure-freeness:* stability of $\hat{\phi}$ across the ladder.

(2) *Post-treatment placebo:* a deliberately post-treatment "talent" proxy must yield a mechanically inflated $\hat{\beta}_3$, confirming the test detects circularity and quantifying what the guardrail removes. (3) *Discontinuity-fixed-proxy check:* the genuine interaction is identified off an orthogonal ability axis, and if it vanishes when the proxy is fixed at the cutoff it was a level or sorting artifact. (4) *Field-match placebo:* the randomized direction-match null of (33) for Ψ_T.

B.4 The decision rule

Estimate the coefficient on (pre-determined talent level × instrumented input) after absorbing the common-in-talent surface $s(\hat{C})$ in (32) and after the centering, instrumenting, and discontinuity-local safeguards above. The decisive instance is the instrumented-engagement-by-continuous-baseline-ability interaction of Agarwal and Gaule (2026). A nonlinearity in investment alone—$F_{CC} \neq 0$ with $\phi = 0$—does not vindicate the operator. If the talent-margin interaction is a precise zero, separability holds, the reallocation gain carries no operator correction, and the canonical number needs no ϕ-re-pricing, though it may still carry a separate, non-operator F_{CC} adjustment. The paper stands or falls on $\phi \neq 0$ specifically—one estimable coefficient, of which the received account is the hypothesis that it is zero. The consequences of rejecting that hypothesis are worked out in Section 10, and the boundary of the resulting claim is drawn in Section 11.

C The Re-Pricing Protocol

This appendix develops the re-pricing of Section 10 in full. It states the inserted accumulation stage in both functional forms, maps it into the host aggregator and shows the exact recovery of the headline number, specifies the calibration and the grid over which the curve is computed, works the illustrative back-of-envelope, and states the measurement guardrail that keeps the exercise free of circularity. Every numerical value here is illustrative pending estimation; it is used to display direction and rough magnitude, never as a result.

C.1 The inserted accumulation stage

We replace the research-labor identity $\ell_i^R = z_i$ of Brouillette (2025) – in which realized research labor is talent itself – with an inserted accumulation stage,

$$\ell_i^R = H_i = F(z_i, C_i),$$

sitting between STEM entry (signal z_i^s, entry noise z_i^n, total $z_i = z_i^s z_i^n$) and research-labor productivity. Parents continue to carry the existing time factor, so the labor supplied is $\ell_i^R = H_i\, \alpha^{1+\delta}$ with $\alpha^{1+\delta}$ exactly as in the host model. The original is recovered when $F(z, C) = z$ for all C: that is, when the inserted stage is the identity in talent and is independent of investment. Two parameterizations are carried throughout.

Cobb–Douglas. $H = z\, C^b$. The cross-partial, own-curvature, and marginal return are

$$\phi = \frac{\partial^2 H}{\partial z\, \partial C} = b\, C^{\,b-1} > 0 \;\; \forall b > 0, \qquad F_{CC} = b(b-1)\, z\, C^{\,b-2} < 0 \;\; \text{iff} \;\; b < 1, \qquad F_C = b\, z\, C^{\,b-1}.$$

The null is $b \to 0$. Two properties matter for the host model. First, ϕ and F_{CC} are distinct functions of b: the cross-partial is positive for every $b > 0$ while the own-curvature is negative for every $b < 1$, so the operator (ϕ) and diminishing returns (F_{CC}) are present together but carry opposite signs

and must be tracked separately. Second, with C a destination-level constant, $H = z\,C^b$ is a power law in z, so the Pareto-tail closed forms of Brouillette (2025) survive: the intensive-margin and idea-production aggregates carry through unchanged up to a multiplicative C^b re-weight.

Constant elasticity of substitution. $H = \left(\omega z^\rho + (1-\omega)C^\rho\right)^{1/\rho}$, with $\omega \in (0,1)$. Writing $\Omega \equiv \omega z^\rho + (1-\omega)C^\rho$,

$$\phi = F_{CT} = \omega(1-\omega)(1-\rho)\, z^{\rho-1} C^{\rho-1}\, \Omega^{1/\rho-2} \implies \operatorname{sign}\phi = \operatorname{sign}(1-\rho).$$

Thus $\phi > 0$ for $\rho < 1$ (complements), $\phi = 0$ at $\rho = 1$ (perfect substitutes, the separable null), and $\phi < 0$ for $\rho > 1$. The Leontief limit $\rho \to -\infty$ delivers the perfect-complement form $\min[C, N]$ that Card et al. (2024) estimate, in which an able student's realized capability is capped by the slot that receives it. The elasticity of substitution is $1/(1-\rho)$, and the linear-stock specification is the measure-zero knife-edge $\rho = 1$ at which $H = \omega z + (1-\omega)C$.

The destination object. Define

$$\kappa_d \equiv \mathbb{E}[C_i \mid d],$$

the average investment intensity a freed woman inventor receives at destination d once the exposure barrier is lifted – in the host calibration, $\tau^E = 0.79$. The sorting object is the across-mover covariance $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) = \mathrm{Cov}_{\mathrm{moved}}(z, \kappa)$ between pre-investment talent and landing intensity.

C.2 Mapping into the aggregator, and exact recovery of 14.2%

Aggregate research input is the only host object whose content changes. It moves from the talent mass to the capability mass,

$$R_t = \int z_i\, di \;\rightsquigarrow\; R_t^{\mathrm{op}} = \int z_i\, C_i^b\, di,$$

and everything downstream is left untouched: the consumption aggregator $c \propto R^\gamma L/N$ with $\gamma = \frac{1}{3}$, the idea law of motion $\dot{M} = M^{\phi^{\mathrm{spill}}} R$ with the knowledge-spillover exponent $\phi^{\mathrm{spill}} = -2.1$, and the overlapping-generations transition. Only the integrand of R is re-weighted by C_i^b.

The three terms, in the host model's own objects. Split the population into incumbents at normalized intensity κ_0 and the freed mass Δ that enters under the counterfactual. Expanding κ^b around κ_0 inside (13) reproduces exactly the three components of Proposition 1:

$$\underbrace{\kappa_0^b \int_\Delta z_i\, di}_{\text{linear-stock}} + \underbrace{b\,\kappa_0^{b-1}\left[\bar{z}_\Delta(\bar{\kappa}_\Delta - \kappa_0) + \mathrm{Cov}_{\mathrm{moved}}(z,\kappa)\right]}_{\text{interaction} \,\propto\, \phi\cdot\mathrm{Cov}_{\mathrm{moved}}(T,\kappa)} + \underbrace{\tfrac{1}{2}b(b-1)\,\kappa_0^{b-2}\int_\Delta z_i(\kappa-\kappa_0)^2\, di}_{\text{own-curvature} \,\propto\, F_{CC}},$$

where $\bar{z}_\Delta$ and $\bar{\kappa}_\Delta$ are the freed mass's mean talent and mean landing intensity. The linear-stock term is the talent mass moved priced at the incumbent margin; the interaction term carries the cross-partial and is the operator correction; the own-curvature term carries $\frac{1}{2}b(b-1)$ and is diminishing returns, present even at the separable null and never to be folded into the operator.

The conditional headline. With destination re-weight $\Lambda(\kappa_d, b) \equiv (\kappa_d/\kappa_0)^b$ and the 49% labor rise treated as the freed-women inflow, the gain conditional on a representative freed mover landing at κ_d is

$$\frac{R_1^{\mathrm{op}}}{R_0^{\mathrm{op}}} = 1 + 0.49\,\Lambda(\kappa_d, b), \qquad \boxed{\mathrm{gain}(\kappa_d, b) = \left[1 + 0.49\,(\kappa_d/\kappa_0)^b\right]^{1/3} - 1.} \tag{35}$$

What (35) is, and is not. The boxed object is a per-destination *conditional* curve. It prices a representative mover at a chosen κ_d, deliberately, so the figures of Section 10 can show how the headline bends with where talent lands. It is not the aggregate gain, and the two must not be conflated – conflating a mean destination with a sorting is precisely the error the linear literature commits. The *aggregate* re-priced gain integrates over the distribution of landing intensities, and by Proposition 1 its bias relative to the linear number is governed by $\phi \cdot \mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$, not by the mean destination κ_d. A mean-preserving spread of κ with $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) = 0$ leaves the aggregate first-order unchanged; positive sorting is what moves it. The two are reported side by side: the conditional curve to display the mechanism, the $\phi \cdot \mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$ object to instantiate the proposition.

Exact recovery. At the null $b = 0$ (equivalently $\rho = 1$), $\Lambda(\kappa_d, 0) = 1$ for every κ_d, and (35) gives

$$\text{gain} = \big[1 + 0.49\big]^{1/3} - 1 = 1.49^{1/3} - 1 = 0.142,$$

that is 14.2% exactly, for every landing intensity. The inserted stage's degenerate, no-curvature limit reproduces the host model's number. We state the status of this recovery plainly. It is a recovery of the *number*, not a structural nesting of an existing object, because the host model has no investment stage C for an operator to switch off. The honest answer to the charge that the result is replaced rather than re-priced is therefore twofold: the host model lacks an investment stage, so the enrichment does not nest it by setting $\phi = 0$ but inserts the stage and recovers 14.2% as its no-curvature limit; and the inserted stage acts on the host model's own intensive margin – the average realized capability that its "better, not more, inventors" result already identifies as load-bearing – so the enrichment disciplines, rather than supplants, the channel the original leans on (cf. Section 7).

C.3 Calibration, solve, and figures

Identifying ϕ, and not from a raw ratio. The complementarity b (or ρ, with the null $\rho = 1$) is recovered from the instrumented talent–investment interaction coefficient, after instrumenting the investment input, and not from any raw ratio of intention-to-treat effects. The global mathematics-talent experiment of Agarwal and Gaule (2026) is instructive here precisely because it shows why the raw ratio will not do. That trial finds an access effect of 0.254 standard deviations among above-median-ability students against 0.006 among below-median, with the difference significant only at $p = 0.097$. The ratio of these two numbers is a reduced-form composite of the cross-partial and a differential first stage – higher-ability students also engage more, as the same study's engagement regression shows – so it is an upper bound on the talent gradient, not a clean estimate of ϕ. The interaction must be instrumented before b can be read from it.

Why F_{CC} is not identified by the same experiment, and is carried separately. The own-curvature F_{CC} is not identified by the gap between the access intention-to-treat effect of 0.165 standard deviations and the 0.6–0.7 standard-deviation return implied for a fully engaged student in Agarwal and Gaule (2026). That gap is an intention-to-treat-versus-local-average-treatment-effect compliance gap over a per-month engagement effect the study estimates as constant, i.e. linear in engagement; the derived 0.66 standard-deviation figure is the product of a per-month effect of about 0.147 and roughly 4.5 months, not a separately estimated curvature. Along the engagement margin, therefore, the local own-curvature is approximately zero, $F_{CC} \approx 0$, and the compliance gap

carries no information about F_{CC}. The own-curvature must instead be recovered from a within-talent-bin dose-response or imported from an external estimate, and in either case is carried as its own term, never inferred from the cross-partial.

Sweeping the sorting. The covariance $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$ is swept over its full range, including negative values. The default is positive: frictions that kept high-talent women out of the highest-intensity environments imply that lifting them lets the ablest land where investment is deepest, so $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) > 0$. The negative contingency is the case in which access without engagement leaves the investment input low. We compute $\bar{\kappa}_\Delta/\kappa_0 \in \{0.5, 0.75, 1, 1.25, 1.5, 2\}$.

Grid. The curve is computed over $b \in \{0, 0.1, 0.2, 0.3, 0.5\}$ (equivalently $\rho \in \{1, 0.5, 0, -1, -\infty\}$); $\kappa_d/\kappa_0 \in \{0.5, 0.75, 1, 1.25, 1.5, 2\}$; and $\mathrm{sign}\,\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) \in \{-, 0, +\}$. Every calibration target of Brouillette (2025) and all three identifying moments are held fixed, so the operator receives no free recalibration of the host model.

Solve. Two routes. The first is the closed-form rule of thumb of (35), which prices the Cobb–Douglas conditional curve directly. The second, in prospect, is a full re-solution of the host model's balanced-growth path with $\ell_i^R = H_i(z_i, C_i)$ substituted into research labor and the intensive-margin constants, re-imposing the same three identifying moments under the CES form so that, again, no calibration freedom is spent. The figures of Section 10 – the conditional curve Figure 3, its dependence on b Figure 4, the operator-versus-aggregator comparison Figure 5, and the three-term decomposition Figure 6 – are produced from this grid.

C.4 Illustrative back-of-envelope

Take Cobb–Douglas, $\gamma = \frac{1}{3}$, baseline factor 1.49, and a conservative $b \approx 0.2$. The conditional gain (35) at the two ends of the intensity range and at the pivot is

$$\begin{aligned}
\kappa_d/\kappa_0 = 2: &\quad \Lambda = 2^{0.2} = 1.149, &&\text{gain} = [1 + 0.49 \cdot 1.149]^{1/3} - 1 = 16.0\%;\\
\kappa_d/\kappa_0 = 0.5: &\quad \Lambda = 0.5^{0.2} = 0.871, &&\text{gain} = [1 + 0.49 \cdot 0.871]^{1/3} - 1 = 12.6\%;\\
\kappa_d/\kappa_0 = 1: &\quad \Lambda = 1, &&\text{gain} = 14.2\% \text{ exactly.}
\end{aligned}$$

So a representative freed mover's conditional gain ranges from about 12.6% (landing at half incumbent intensity) to about 16.0% (double), with the flat 14.2% at the pivot. A stronger complementarity $b \approx 0.5$ widens the conditional range to about 10.4%–19.2%. This sweep moves the *level* κ_d; the *aggregate* re-priced gain instead moves with $\phi \cdot \mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$, so the band displays the mechanism while the aggregate instantiates the proposition. The central case is the upward branch – 14.2% a lower bound pushed toward 16–19% – because the argued sorting is $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa) > 0$; the downward branch is the contingency that Agarwal and Gaule (2026) flags, access without engagement. The direction is pinned by $\mathrm{sign}(b) \cdot \mathrm{sign}(\mathrm{Cov}_{\mathrm{moved}}(T, \kappa))$, the content of Proposition 1. This is a distinct sensitivity axis from the host model's own 2%–27% band over the returns-to-scale parameter: that band varies γ holding the specification linear, while this one varies where talent lands and how complementary the map is, holding $\gamma = \frac{1}{3}$. Because the operator works through the same increasing-returns lever, the two axes compound: a high-γ world magnifies the operator spread as well.

C.5 Guardrail in the re-pricing

The exercise is free of circularity only if both arguments of the inserted stage are measured before the stage acts. On the talent side, the measure feeding $F(z, C)$ must be the pre-investment birth signal z_i^s – or, in an experimental setting, a baseline selection score – and never a realized-capability output. The host model's own timing makes this clean: z_i^s is observed at birth, strictly prior to any research-environment investment C_i, so interacting z_i^s, rather than realized research labor, with destination intensity respects the guardrail by construction.

The symmetric requirement is on the destination side, and it is equally load-bearing. The intensity κ_d must be a pre-realization property of the environment – course intensity, mentoring availability, laboratory or environment productivity – and never the mover's own realized engagement or output. Using realized engagement as κ would make $\mathrm{Cov}_{\mathrm{moved}}(T, \kappa)$ co-move mechanically with the interaction term, re-introducing on the destination side exactly the circularity the talent-side guardrail rules out. The two conditions together are what allow the bias to be signed from observed movers alone, the discipline carried through Section 9 and bounded in Section 11.

References

Andrew B. Abel. Classical measurement error with several regressors. Working paper, The Wharton School, University of Pennsylvania, 2018. Originally May 2017, revised May 2019.

Ruchir Agarwal and Patrick Gaule. Developing math talent worldwide: Evidence from a global rct. IZA Discussion Paper 18381, IZA – Institute of Labor Economics (IZA@LISER Network), Bonn / Esch-sur-Alzette, February 2026. ISSN: 2365-9793. Pre-registered AEA RCT Registry AEARCTR-0011178.

Francesco Agostinelli and Matthew Wiswall. Estimating the technology of children's skill formation. *Journal of Political Economy*, 133(3):846–887, 2025. doi: 10.1086/733396.

Nicholas J. Ainsworth, Aaron J. Ainsworth, Christopher Cleveland, Leah R. Clark, Quentin Brummet, Emily K. Penner, Jacob Hibel, Andrew Saultz, Michelle Spiegel, Paul Hanselman, and Andrew Penner. Gifted identification across the distribution of family income. CES Working Paper CES 25-73, U.S. Census Bureau, Center for Economic Studies, Washington, DC, December 2025.

Anna Aizer and Flavio Cunha. The production of human capital: Endowments, investments and fertility. NBER Working Paper 18429, National Bureau of Economic Research, September 2012.

Ufuk Akcigit, Jeremy G. Pearce, and Marta Prato. Tapping into talent: Coupling education and innovation policies for economic growth. *The Review of Economic Studies*, 92(2):696–736, 2025.

Naser Amanzadeh, Amir Kermani, and Timothy McQuade. Return migration and human capital flows. NBER Working Paper 32352, National Bureau of Economic Research, Cambridge, MA, April 2024. URL http://www.nber.org/papers/w32352. JEL No. F22, J61, O15.

Susan Athey and Scott Stern. An empirical framework for testing theories about complementarity in organizational design. NBER Working Paper 6600, National Bureau of Economic Research, Cambridge, MA, 1998.

David Rezza Baqaee and Emmanuel Farhi. Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics*, 135(1):105–163, 2020. doi: 10.1093/qje/qjz030.

Alexander M. Bell, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. Who becomes an inventor in america? the importance of exposure to innovation. *The Quarterly Journal of Economics*, 134(2):647–713, 2019. doi: 10.1093/qje/qjy028. Earlier version circulated as NBER Working Paper No. 24062 (November 2017, revised January 2019).

Morten Bennedsen, Francisco Pérez-González, Maria Schlier, Margarita Tsoutsoura, and Daniel Wolfenzon. The missing talent in public sector. Study Paper 238, The ROCKWOOL Foundation Research Unit, Copenhagen, Denmark, February 2025. Working paper dated September 12, 2024.

Dan A. Black, Mark C. Berger, and Frank A. Scott. Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association*, 95(451):739–748, 2000. doi: 10.1080/01621459.2000.10474262.

Nicholas Bloom, Charles I. Jones, John Van Reenen, and Michael Webb. Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144, 2020.

Jean-Félix Brouillette. The lost marie curies and foregone economic growth. Working paper, HEC Montréal, August 2025.

Gharad Bryan and Melanie Morten. The aggregate productivity effects of internal migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):2229–2268, 2019. doi: 10.1086/701810.

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002, 2011. doi: 10.1257/aer.101.5.1964.

Jerome R. Busemeyer and Lawrence E. Jones. Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 93(3):549–562, 1983. doi: 10.1037/0033-2909.93.3.549.

Ian Calaway. Early mentors for exceptional students. Job Market Paper, DePaul University. April 29, 2026. Supported by the Global Math Talent Fellowship through the NBER and the Agency Fund, 2026.

David Card and Laura Giuliano. Targeted acceleration in middle school math: Impacts on college entry, degree completion, and stem. Nber working paper, National Bureau of Economic Research, Cambridge, MA, October 2024.

David Card, Eric Chyn, and Laura Giuliano. Can gifted education help higher-ability boys from disadvantaged backgrounds? NBER Working Paper 33282, National Bureau of Economic Research, Cambridge, MA, December 2024. URL http://www.nber.org/papers/w33282. Revised May 2026.

Murat Alp Çelik. Does the cream always rise to the top? the misallocation of talent in innovation. *Journal of Monetary Economics*, 133:105–128, 2023.

Sarah R. Cohodes, Helen Ho, and Silvia C. Robles. STEM summer programs for underrepresented youth increase STEM degrees. NBER Working Paper 30227, National Bureau of Economic Research, Cambridge, MA, July 2022. URL http://www.nber.org/papers/w30227. Revised September 2024.

Flavio Cunha and James J. Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, May 2007. doi: 10.1257/aer.97.2.31.

Flavio Cunha, James J. Heckman, and Susanne M. Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, May 2010. doi: 10.3982/ECTA6551.

Elias Einiö, Josh Feng, and Xavier Jaravel. Social push and the direction of innovation. CEP Discussion Paper 1861, Centre for Economic Performance, London School of Economics, July 2022.

Caroline Fry and Britta Glennon. In good company: Coethnic advisors and career paths of immigrant ph.d. students. NBER Working Paper 33782, National Bureau of Economic Research, Cambridge, MA, May 2025. URL http://www.nber.org/papers/w33782. JEL No. F22, I23, J24, O31.

Steven Garber and Steven Klepper. Extending the classical normal errors-in-variables model. *Econometrica*, 48(6):1541–1546, 1980. doi: 10.2307/1912825.

Zvi Griliches and Jerry A. Hausman. Errors in variables in panel data. *Journal of Econometrics*, 31(1):93–118, 1986. doi: 10.1016/0304-4076(86)90058-8.

James J. Heckman and Stefano Mosso. The economics of human development and social mobility. *Annual Review of Economics*, 6:689–733, 2014. doi: 10.1146/annurev-economics-080213-040753.

Mikkel Aagaard Houmark, Victor Ronda, and Michael Rosholm. The nurture of nature and the nature of nurture: How genes and investments interact in the formation of skills. *American Economic Review*, 114(2):385–425, February 2024. doi: 10.1257/aer.20220456.

Chang-Tai Hsieh and Enrico Moretti. Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39, 2019. doi: 10.1257/mac.20170388.

Chang-Tai Hsieh, Erik Hurst, Charles I. Jones, and Peter J. Klenow. The allocation of talent and U.S. economic growth. *Econometrica*, 87(5):1439–1474, 2019. doi: 10.3982/ECTA11427.

Yingyao Hu and Susanne M. Schennach. Instrumental variable treatment of nonclassical measurement error models. *Econometrica*, 76(1):195–216, 2008. doi: 10.1111/j.0012-9682.2008.00823.x.

Zhao Jin, Amir Kermani, and Timothy McQuade. Native-immigrant entrepreneurial synergies. NBER Working Paper 33804, National Bureau of Economic Research, Cambridge, MA, May 2025. URL http://www.nber.org/papers/w33804. JEL F22, G24, L26, M13, O32.

Charles I. Jones. R&d-based models of economic growth. *Journal of Political Economy*, 103(4): 759–784, 1995.

David Lagakos and Michael E. Waugh. Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–980, 2013. doi: 10.1257/aer.103.2.948.

Arthur Lewbel. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1):67–80, 2012. doi: 10.1080/07350015.2012.643126.

Stefanie Muff and Lukas F. Keller. Reverse attenuation in interaction terms due to covariate measurement error. *Biometrical Journal*, 57(6):1068–1083, 2015. doi: 10.1002/bimj.201400157.

Paul Novosad, Sam Asher, Catriona Farquharson, and Eni Iljazi. Access to opportunity in the sciences: Evidence from the nobel laureates. Working paper, Dartmouth College, Imperial College London, Harvard University, and University of Pennsylvania, May 2026.

Marta Prato. The global race for talent: Brain drain, knowledge transfer, and growth. *The Quarterly Journal of Economics*, 140(1):165–238, 2025.

A. D. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146, 1951. doi: 10.1093/oxfordjournals.oep.a041827.

Eleanor Sanderson and Frank Windmeijer. A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221, 2016. doi: 10.1016/j.jeconom.2015.06.004.