The Shapley Singularity: Contribution, Intelligence, and the Compression of Economic Value

PIATRA . INSTITUTE

June 2026

Abstract

A firm of 100 units of revenue hires 10 people and ends the year at 120, and every available theory of value answers the question of who made the 20 badly: labor-value accounting hands it to an undifferentiated substance called labor, marginal-product accounting books it to whoever was measured last, and bonus arithmetic divides it by headcount. The question is a credit-assignment problem, the same problem that reinforcement learning has faced under that name for sixty years, and cooperative game theory solved its allocative core in 1953: the Shapley value, the unique division of a jointly produced surplus consistent with efficiency, symmetry, and indifference to everything except contribution. This paper asks what political economy looks like when that solution is taken seriously at the moment machine intelligence becomes a production factor. A minimal firm, solved exactly, shows naive growth attribution overstating the new hires' contribution more than fourfold and isolates a residue of the surplus, the three-way synergy, that belongs to no participant alone. A task-continuum production game over intelligence, compute, energy, and human labor, solved exactly along the whole automation path, shows total output multiplying while labor's share of attributed value collapses from all of it to an eighth, and it shows two structural facts the singularity discourse misses, both consequences of modeling the stack factors as perfect complements: the Shapley values of intelligence and its essential complements are forced equal by symmetry, so attribution cannot see which member of the stack is the clever one, and intelligence's share of value is capped at one-third by two essential complements, at $1/(k+1)$ by k, for any finite capability whatever. We define the Shapley singularity as the limit in which intelligence's share of attributed value approaches one across frontier production games, and within the perfect-complement model the computation shows it is unreachable by capability growth alone: it is reached only by absorbing or shedding complements, which makes the singularity of attribution an ownership event rather than an intelligence event. The qualification is load-bearing: with imperfect substitutability between the stack factors the equalities loosen and the ceiling moves, and a capability path that erodes a complement's essentiality is a genuine route to a higher ceiling that the model excludes by assumption, so the ownership thesis holds within the perfect-complement structure rather than unconditionally. A final pair of games locates the last human wage in verification, the best-attributed factor on the frontier until machine output certifies itself, after which labor's share falls threefold in one step, and exhibits the gap between the Shapley point and the core, between what each factor contributes and what its owner can stably extract, which is where politics lives. The frame is bounded: a Shapley value is an audit under a chosen value function, never a desert claim, and the choices of what counts as value and what follows from the audit remain political acts that no attribution mathematics can perform.

1. The Twenty Units

A company produces 100 units of revenue, hires 10 people, and a year later produces 120. Someone must now decide what the 10 contributed, because raises, bonuses, equity, and the next hiring round all depend on the answer. The arithmetic that gets used in practice divides the growth by the headcount, 2 units a head, and the assumptions buried in that division are visible the moment they are written down: that none of the growth would have happened without the new people, that each of them contributed the same amount, that the incumbent team's productivity did not move, that demand, capital, brand, and timing contributed nothing, and that the 10 were independent producers rather than complementary parts of a system. Every one of those assumptions is usually false, and the falseness is structural rather than accidental: a salesperson produces little until engineering ships, a senior engineer may produce nothing measurable while unblocking five others, and the coordinator who removes friction appears in no ledger at all. Individual contribution inside joint production is not directly observable even when the joint product is known to the unit, and the institutions that pretend otherwise are pricing a fiction.

The problem is old, and it is not owned by any one tradition. Marxian political economy needs an answer when it speaks of the surplus that labor produces and capital appropriates, and its critics have always pressed the point that "labor" names a heterogeneous, interacting collection rather than a homogeneous substance. Meritocratic capitalism needs an answer when it defends pay as the reflection of contribution, and it is no closer to having one, because market wages reflect bargaining position, scarcity, and institutional rules as much as any marginal product. Machine learning needs the answer operationally: a reinforcement learner that wins a game must distribute the credit for the win across the thousands of decisions that produced it, a problem named credit assignment at the field's founding (Minsky, 1961) and placed at its center ever since (Sutton and Barto, 2018). The firm's question and the learner's question are the same question, how to divide the consequence of a joint history among its contributors, and the mathematics that answers one answers the other.

This paper takes the cooperative-game answer, the Shapley value, and asks what happens to political economy when the answer is applied at the moment machine intelligence enters production as a factor rather than a tool. The thesis has three parts, and each is exhibited in a model small enough to solve exactly rather than argued by gesture. The first part is corrective: in any production with complementarity, naive attribution is wrong by large factors in identifiable directions, and a portion of every joint surplus belongs to no participant alone. The second part is the heart: as the automatable share of the task space grows, attributed value migrates from labor to the intelligence stack, but the migration has a ceiling fixed by the structure of complements, and the ceiling is indifferent to how capable the intelligence becomes. The third part is political: attribution and appropriation are different objects, a point and a zone, and the distance between them, which the Shapley value makes visible and the market does not, is the territory in which the distributional fights of the coming decades will be fought.

Throughout, the restraint is the one good formal work imposes: the games are calibrated to no firm and no economy, and their numbers, the fourfold overstatement, the one-third ceiling, the threefold cliff, are facts about precisely specified structures, offered because those structures are the simplest in which the claimed phenomena can be realized at all. What survives transfer to the world is the shape of the results, the asymmetries and the ceilings and the orderings, and not the decimals.

2. From Labor Value to Contribution Value

Classical political economy began with a celebration of synergy and never built the instrument to divide it. The pin factory that opens Smith's account produces tens of thousands of pins from ten men whose solitary outputs would have been a few dozen each, and the difference, the overwhelming bulk of the product, is attributable to the division of labor itself, to the arrangement rather than to any man (Smith, 1776). Smith drew from this the wealth of nations; he did not ask, because his problem was aggregate, what share of the day's pins the wire-drawer made. Ricardo put distribution at the center, declaring the division of the produce among classes the principal problem of political economy, and answered it with class-level laws of rent, profit, and wages rather than with any account of individual contribution (Ricardo, 1817). Marx made the gap constitutive: surplus value is what labor produces beyond the reproduction cost of labor power, and the capitalist appropriates it through ownership, and this is a structural relation between classes, deliberately not a metering of who added what (Marx, 1867). The strength of that move is that it asks about power and property rather than bookkeeping; its weakness is that "labor" enters as a single substance, and a theory whose central quantity is jointly produced by heterogeneous, complementary contributors cannot say, even to its own satisfaction, how the surplus decomposes.

The marginalists believed they had the metering. In Clark's hands the claim became moral: under competition each factor is paid its marginal product, and so, under natural law, each gets what it creates (Clark, 1899). The trouble was identified before the doctrine hardened. Wicksteed showed that paying every factor its marginal product exhausts the product exactly only under constant returns to scale, by Euler's theorem, and the interesting economies, the ones with fixed costs, increasing returns, and complementarity, are where the books fail to close (Wicksteed, 1894). The failure is now textbook microeconomics in another costume: where production is team production, the joint output is not a sum of separable individual outputs, the marginal products do not add up, and metering individual contribution is the firm's defining difficulty (Alchian and Demsetz, 1972). Holmström sharpened it into an impossibility: no sharing rule that exactly distributes the team's output among its members can make every member internalize their full marginal effect, so the firm cannot, even in principle, pay everyone their marginal product out of the product (Holmström, 1982). Complementarity is not a corner case being waved at here; the economics of modern manufacturing is a portrait of supermodular systems in which the return to each activity rises with every other, so that the firm is a web of interactions rather than a list of contributions (Milgrom and Roberts, 1990). Both grand traditions, the labor theory and the marginal-product theory, founder on the same rock: joint production with synergy.

The mathematics that handles the rock exists and is 70 years old. What a coalition is worth was formalized by von Neumann and Morgenstern (1944) as the characteristic function $v(S)$, and Shapley (1953) proved that four innocuous requirements, that the whole surplus be distributed, that interchangeable players receive the same, that a player who adds nothing receive nothing, and that attribution add across games, pick out exactly one division: each player's contribution averaged over every order in which the coalition could have assembled. Young (1988) later showed the same point can be carried by marginality alone, that if attribution depends only on a player's marginal contributions, symmetry and efficiency force the Shapley value, which is as close as allocation theory comes to a uniqueness theorem for the concept of individual contribution. Harsanyi (1963) supplied the complementary ledger: every coalitional game decomposes uniquely into dividends, pure interaction terms attached to each subset, so that the value of any coalition is the sum of the dividends of its parts, and the Shapley value is what each player collects when every dividend is split equally among its members. Synergy, in this accounting, stops being noise around individual contributions and becomes an entry in its own right, with its own line in the ledger.

Political economy noticed, and the noticing is a story of approach and refusal. The analytical Marxists of the September Group brought these very tools to this doctrine (Elster, 1985). Roemer rebuilt exploitation theory on coalitional counterfactuals: a group is exploited if it would do better withdrawing from the economy with its per-capita share of assets, a characteristic-function construction in all but name, and the result relocated exploitation from the labor process to the distribution of property (Roemer, 1982). Cohen, pressing on the labor theory's foundations, worked the two-player case explicitly: land that yields 1 unit alone and 10 with labor, labor that yields nothing alone, and the Shapley division gives land 5.5 and labor 4.5, an arithmetic embarrassment for any claim that labor, being the active principle, deserves the whole (Cohen, 1995). Reiff, surveying compensation theory, considered the Shapley value as the repair for the adding-up problem and rejected it on desert grounds, asking why anyone is owed their average marginal contribution across hypothetical orderings rather than their actual contribution in the actual history (Reiff, 2013). The other side's discomfort was formalized by von Platz (2021) as a dilemma: meritocratic defenses of capitalism need a notion of contribution that both justifies the actual capital-labour split and is morally plausible, and no candidate does both. And the tradition's center rejected the whole idiom; Wood's attack on rational-choice Marxism, static, individualist, ahistorical, spoke for the majority view that class struggle does not factor through a characteristic function (Wood, 1989).

So the tool was examined by the one tradition that most needed it and shelved, and the reason is not mathematical. A serious Shapley accounting is doctrinally dangerous in both directions at once. It will assign positive shares to capital goods, to coordination, sometimes to management, because in the counterfactual ledger they are not dummies, and that dissolves the clean claim that labor creates all value. It will equally expose gaps between what owners extract and any computable contribution, pricing rent, monopoly position, inherited assets, and captured luck as what they are, and that dissolves the clean claim that the market already pays contribution. A tool that audits everyone recruits no one. Marx offers a place in history; an attribution engine offers a coefficient. There are Marxists and there are marginalists, and there are no Shapleyans, and the asymmetry says nothing about the mathematics and everything about what ideologies are for. The value itself, meanwhile, went into the infrastructure: it allocates joint costs inside firms (Shubik, 1962), measures voting power in committees (Shapley and Shubik, 1954), and sits, with the matching theory begun in Gale and Shapley (1962) and honored alongside it with a Nobel six decades later, in the plumbing of market design. The pattern is exact: Marx became a flag, Shapley became plumbing, and an economy about to be reorganized by machine intelligence is about to need its plumbing made visible.

3. The Shapley Frame

A production game is a pair (N, v): a finite set of factors N, which may be people, teams, machines, datasets, energy contracts, or institutions, and a characteristic function v assigning to every coalition $S \subseteq N$ the value the coalition would produce operating together, with $v(\emptyset) = 0$ (von Neumann and Morgenstern, 1944). The Shapley value of factor i is its marginal contribution averaged over every order in which the grand coalition can assemble,

$$\phi_i(v) \;=\; \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!}\,\big[\,v(S \cup \{i\}) - v(S)\,\big],$$

and it is the unique allocation that exhausts $v(N)$, treats interchangeable factors identically, gives nothing to a factor that adds nothing anywhere, and adds across independent games (Shapley, 1953). The adding-up problem that broke marginal-productivity ethics is closed by construction: the shares sum to the product whatever the returns to scale, because the averaging spreads the increasing-returns surplus over the orders instead of double-counting it at the margin.

The Harsanyi decomposition is the same object read as a ledger of interactions. Every game has a unique expansion

$$v(S) \;=\; \sum_{\emptyset \neq T \subseteq S} d(T), \qquad\qquad \phi_i(v) \;=\; \sum_{S \ni i} \frac{d(S)}{|S|},$$

in which $d(T)$, the dividend of coalition T, is the part of T's value that exists over and above everything its proper sub-coalitions account for (Harsanyi, 1963). A standalone contribution is a dividend on a singleton; a pure two-party synergy is a dividend on a pair; and the dividend on the grand coalition is the residue that exists only because everyone showed up. The Shapley value then has a one-sentence statement: every dividend is split equally among the factors it needs. The split of the synergies is the substantive act, and the equal split is what symmetry buys; everything contested in attribution lives in the dividends of size two and up.

Two disciplines must be imported before this machinery touches the world. The first is causal: $v(S)$ is a counterfactual quantity, what the coalition would produce, an interventional expectation in the sense of the do-operator rather than an observed conditional (Pearl, 2009). Estimating attribution from correlations invites the standard catastrophe, the strong engineer assigned to failing projects who is scored as causing failure, and the recent attribution literature has had to rediscover this inside machine learning, where Shapley explanations computed without causal structure mislead in the ways confounding always misleads (Heskes, Sijben, Bucur, and Claassen, 2020). The second discipline is to remember what the number is. A Shapley value is the unique attribution consistent with the axioms under a specified v; it is an audit, and it carries no theory of desert, no claim about need, risk, history, or right. The politics enters twice, in choosing what v counts as value and in deciding what follows from the audit, and a frame that hides either entrance is propaganda with a theorem in it.

What recommends the frame is that engineering keeps rediscovering it under selection pressure. The multi-agent learning literature, needing to train swarms whose members share one global reward, converged on difference rewards: score agent i by $G(z) - G(z_{-i})$, the system's outcome against the counterfactual outcome with i's action replaced by a default (Wolpert and Tumer, 2001). The counterfactual multi-agent policy gradient does the same with a learned critic, marginalizing out one agent's action while holding the others fixed (Foerster, Farquhar, Afouras, Nardelli, and Whiteson, 2018). These are marginal contributions to coalitions, the integrand of the Shapley formula, derived not from fairness axioms but from the engineering fact that agents trained on raw global reward free-ride and agents trained on purely local metrics destroy the whole. The firm's two failure modes, the all-hands bonus and the siloed KPI, are the same two failure modes, and the repair converges on the same mathematics. SHAP, which industrialized Shapley attribution over input features and made ϕ one of the most-computed quantities in applied statistics (Lundberg and Lee, 2017), completes the picture: the value is already the working theory of contribution everywhere except in the theory of contribution's home discipline.

4. The Firm, Recomputed

The opening puzzle can now be solved rather than lamented. Take three factors: O, the incumbent organization, the existing team with its product, processes, and assets, treated as one bloc; N, the 10 new hires as a bloc; and M, the favorable market, the demand wave the firm happened to ride that year. The characteristic function encodes the counterfactuals a serious post-mortem would actually debate. The incumbents alone were the firm as it was: $v(\{O\}) = 100$. Hires without a firm produce nothing, demand without a firm sells nothing, and the two together still produce nothing: every coalition without O is worth zero. The old firm riding the wave alone would have grown to $v(\{O, M\}) = 110$; the enlarged firm in a flat market would have reached $v(\{O, N\}) = 106$; and the observed year is $v(\{O, N, M\}) = 120$. The numbers are an illustration with the realistic shape, most of the counterfactual growth attaches to conditions and to the enlarged firm's interaction with them, and any reader's preferred numbers can be run through the identical machinery.

The exact Shapley division of the 20 units of growth gives the incumbent organization 9.33, the 10 hires 4.33, and the demand wave 6.33. Per hire that is 0.43 of a unit against the naive 2 units a head, an overstatement of 4.6 times, and the error is not pedantry: it is the difference between a bonus pool that prices contribution and one that prices recency, and in the other direction it is the standard managerial error of crediting the wave to the surfers already on payroll. The Harsanyi ledger says where the 20 actually live. The hires' standalone dividend is zero, as is demand's; 6 units are the pure interaction of incumbents with hires, 10 are the pure interaction of incumbents with demand, and 4 units exist only in the three-way coincidence of an enlarged firm meeting a rising market. Those 4 units are the irreducibly social residue of the year: they are attributable to no factor alone and to no pair, and every accounting that forces them onto someone's line, the founder's, the new team's, the market's, is making a political allocation while calling it measurement. The Shapley value's treatment, splitting each dividend equally among the factors it requires, 3 and 3, 5 and 5, 4/3 each, is the disciplined version of the act, and Figure 1 lays the two decompositions side by side.

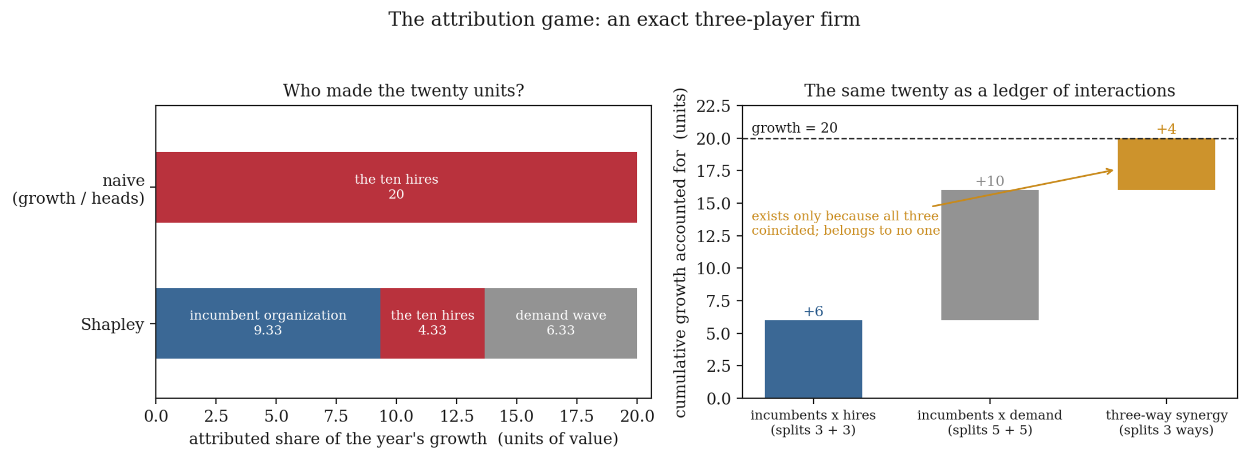

Figure 1: The attribution game. Left: the year's 20 units of growth divided by the naive rule, growth over headcount, against the exact Shapley division: incumbents 9.33, the 10 hires 4.33, the demand wave 6.33, a per-hire contribution of 0.43 against the naive 2. Right: the same 20 as Harsanyi's ledger of interactions: 6 units of incumbent-hire synergy, 10 of incumbent-demand synergy, and a three-way dividend of 4 units that exists only because all three coincided and belongs to no factor alone; each dividend splits equally among the factors it needs, which is how the Shapley division on the left arises.

Two boundary cases calibrate the intuition. Cohen's case for the labor theory's crisis, land yielding 1 alone, labor nothing alone, 10 together, divides under the same rule into 5.5 for land and 4.5 for labor, and the lesson is symmetric: the factor that "does nothing," that merely sits there being scarce and necessary, collects more than the active principle, and the activist intuition that labor's activity earns the whole has no theorem on its side (Cohen, 1995). The opposite limit is the pure complement pair, an engineer and a salesperson worth nothing apart and 100 together, which divides 50 and 50: where value is entirely joint, standalone productivity measures nothing and attribution is symmetry all the way down. Between these poles sits every real firm, which is why both folk theories, pay-by-output and pay-by-ownership, miss systematically. The demand wave's 6.33 deserves a last look before the factors become machines: it is paid, in the actual economy, to whoever holds the residual claim, which is to say that a third of this firm's growth is captured luck, and the well-documented human habit of booking luck as skill (Frank, 2016) has, in the Shapley frame, a name: assigning the dividends of M to one's own line.

5. Intelligence as the Universal Complement

Machine intelligence enters this frame as a factor, and the question the frame forces is the right one: what does adding I to a coalition do to v, across all coalitions, and what does its presence do to the value of everything else. The economics of general purpose technologies says the historical pattern: steam, electricity, and computing created value not as products but as upstream capabilities whose arrival repriced every downstream activity by its ability to use them, with long lags while complementary investment accumulated (Bresnahan and Trajtenberg, 1995). The productivity J-curve names the accounting shadow of those lags, years in which the technology's contribution is real, intangible, and invisible to output statistics (Brynjolfsson, Rock, and Syverson, 2021). On this reading, machine intelligence is the next general purpose technology and the analysis is continuity.

The discontinuity candidate is specific and worth stating with care. Growth theory separates the production of goods from the production of ideas, output from the law of motion of technology itself, and automation matters differently depending on which function it enters; an intelligence that improves the ideas production function participates in producing the thing that produces everything else, and the models in which that loop closes are the models with singularities in them (Aghion, Jones, and Jones, 2019). The ladder is observable in the present tense: systems that execute tasks, then systems that restructure workflows, then systems that act as agents, and, in narrow domains already, systems inside the discovery process itself, of which protein-structure prediction is the clean example, a research bottleneck of fifty years' standing absorbed into a tool (Jumper, Evans, Pritzel, Green, Figurnov, et al., 2021). Task-exposure measurement finds the unusual signature at the labor end: unlike prior automation, which climbed from routine manual work upward, large-model exposure concentrates in cognitive, educated, white-collar tasks (Eloundou, Manning, Mishkin, and Rock, 2024). And the input side has become macroeconomic, with corporate investment in the technology running to hundreds of billions of dollars a year and adoption reported by a large majority of surveyed firms (Maslej, Fattorini, Perrault, Gil, Parli, et al., 2025).

From inside the investment logic, a one-product conjecture suggests itself, and it deserves to be stated before it is corrected, because it organizes a great deal of present behavior. If intelligence can in principle design the better pin factory, the better product, the better sales motion, and eventually the better intelligence, then the expected return on improving intelligence dominates the expected return on improving any particular thing, since one increment of generality pays out across every domain at once. Why invent a better object when you can invest in the thing that invents better objects? On this view the economy compresses toward a single trade: everything converts to intelligence, every other asset becomes fungible against it, and the terminal state is a one-product game. The conjecture has the structure of a limit claim about attribution, and limit claims about attribution can be computed. Acemoglu and Restrepo (2019) supply the variable the computation needs: the boundary between tasks machines can absorb and tasks that remain human is the moving frontier of automation, displacement pushing it forward, the creation of new human tasks pushing it back. Make that boundary a parameter and the one-product conjecture becomes a theorem's worth of exact statements, most of them corrections.

6. The Compression Path

Let the economy be a continuum of tasks of total weight one, and let four factors play: I, intelligence, the models and agents; C, compute, the physical substrate that runs them; E, energy; and H, human labor. A share g of the task space is automatable at the current frontier. Humans perform any task at productivity one. The full machine stack, intelligence with its compute and its energy, performs automatable tasks at productivity $\alpha > 1$ and cannot touch the rest. Each task is performed by the best producer present, so the characteristic function is

$$v_g(S) \;=\; (1-g)\,\mathbf{1}[H\in S] \;+\; g\,\max\Big\{\alpha\,\mathbf{1}[\{I,C,E\}\subseteq S],\; \mathbf{1}[H\in S]\Big\},$$

a one-line model whose entire content is three assumptions worth stating in words: some work only humans can do, shrinking as g grows; machine production needs the whole stack, intelligence and compute and energy jointly or nothing; and humans retain an outside option on every task, the ability to do it the old way at productivity one. Exact enumeration over all orderings, verified against closed forms, gives for every g and α

$$\phi_H(g) \;=\; 1-\frac{g}{4}, \qquad \phi_I(g) \;=\; \phi_C(g) \;=\; \phi_E(g) \;=\; \frac{g\,(4\alpha-3)}{12}, \qquad v_g(N) \;=\; 1+(\alpha-1)\,g.$$

Read at the main calibration $\alpha = 6$, the path is the whole pre-singularity story in one picture (Figure 2). Total value multiplies sixfold between $g = 0$ and $g = 1$. Labor's absolute attribution barely moves, falling from 1 to 0.75, because the outside option keeps humans pivotal in most orderings; labor's share of value collapses from everything to 0.125, an eighth. The two facts together are the political economy of the transition in miniature: an attribution-faithful economy could hold labor's absolute compensation roughly constant while the economy sextuples around it, and the question of whether that is prosperity or immiseration is the question of whether people live on absolutes or on shares, to which the honest answer, for creatures who rent housing from each other and bid against each other for position, is shares (Korinek and Stiglitz, 2019). Intelligence's attributed share rises from nothing and passes labor's at $g^* = 3/\alpha$, which at this calibration is 0.5: when half the task space is automatable, the machine factor and the human factor weigh the same in the counterfactual ledger.

Two structural results sit on top of the curves, and they are the paper's center, though both are consequences of one modeling choice that must be named before they are read. The model makes the stack factors perfect complements: intelligence, compute, and energy are jointly essential and individually useless, a unanimity structure whose machine term is Leontief, the whole stack or nothing. That assumption is not innocent, and everything below holds within it rather than unconditionally. The first result is the symmetry, and it is not a discovery about the world but a direct consequence of that choice: $\phi_I = \phi_C = \phi_E$ identically, at every g and every α, because once the three factors are defined as jointly essential and individually useless they are interchangeable inside every coalition that uses them, and the symmetry axiom then forces their values equal. The equality is built into the value function; the axiom only reads it out. What the model exhibits, then, is conditional and still illuminating: under perfect complementarity attribution is blind to glamour, the value cannot be made to say that the model weights are the magic and the megawatts are the commodity, and in the counterfactual ledger the dumb complement and the brilliant one weigh the same wherever both are strictly necessary. The premium our intuitions award to intelligence is, in this game and only because of how the game is built, a sentiment the arithmetic does not share. The practical reading holds regardless: the closer intelligence comes to needing its complements, the more valuable those non-substitutable complements become, in lockstep, which is why the capability race arrives accompanied by a compute race, an energy race, and a fabrication race rather than instead of them. Value migrates to whatever lets intelligence act.

The second result is the ceiling, and it inherits the same conditionality, because it is the same perfect-complementarity assumption seen from the other side. Intelligence's share of total value at full automatability is $(4\alpha - 3)/12\alpha$: it is 0.25 at $\alpha = 3$, 0.292 at $\alpha = 6$, 0.3125 at $\alpha = 12$, 0.328 at $\alpha = 48$, and it approaches but never reaches one-third, 0.333, as $\alpha \to \infty$. No finite productivity multiple, however astronomical, carries intelligence's attributed share past the ceiling set by two essential complements, and the general statement, computed here for k from one to five, is that with k essential complements the limit game is the unanimity game and intelligence's ceiling is exactly $1/(k{+}1)$: a half, a third, 0.25, 0.2, 0.167. The $1/(k{+}1)$ form is a fact about the unanimity game, not a law of production: it follows from the complements being strictly essential, and the moment a complement becomes substitutable rather than essential the limit game stops being a unanimity game and the ceiling moves up. This is the path the model excludes by assumption rather than by argument, and it is a genuine one. Capability growth that erodes a complement's essentiality, an algorithmic advance that lets the same output be produced with far less energy or with commodity rather than frontier compute, makes that complement partly substitutable, lowers the effective k, and so raises intelligence's ceiling by the capability route the perfect-complement model rules out. Define, with that boundary marked, the object the title names: the *Shapley singularity* is the limit $\Psi_t \to 1$, where $\Psi_t = \phi_I(v_t)/v_t(N)$ is intelligence's share of attributed value across the frontier economy's production games. Within the perfect-complement model the limit cannot be approached along the capability axis at all: raising α raises output without bound while Ψ saturates strictly below the ceiling, and the only motions that raise the ceiling are changes in the player set itself, complements being absorbed, the energy contract, the fabrication capacity, the datacenter, and eventually the embodiment, brought inside the boundary of the bidding entity so that the game's essential triple becomes one player. Relax the complementarity and a second motion opens, capability that dissolves a complement's essentiality, so the clean claim is conditional: holding the stack's complements fixed and essential, the singularity of attribution, if it ever arrives, is an ownership event, a fact about the boundaries of firms rather than about how smart the machines got. The thesis is that within the complement structure the economy actually has, ownership and not capability moves the ceiling; whether that structure holds, or whether capability erodes it from inside, is the empirical question the model frames rather than settles. The empirical singularity literature, testing for growth accelerations and finding the brakes still engaged (Nordhaus, 2021), and the growth-theory observation that the binding constraint migrates to whatever remains essential and hard to improve (Aghion, Jones, and Jones, 2019), are the macro shadows of the same structure.

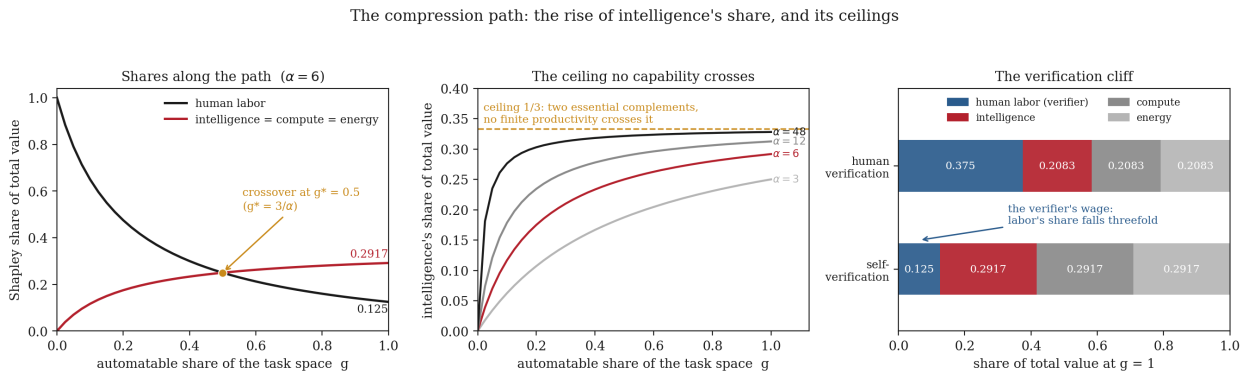

Figure 2: The compression path, computed exactly at $\alpha = 6$ except where marked. Left: Shapley shares of total value along the automatable share g; labor falls from all of it to 0.125 while the three stack factors, forced equal by symmetry at every point, rise together to 0.292 each, crossing labor at $g^* = 3/\alpha = 0.5$. Center: intelligence's share against g for $\alpha = 3, 6, 12, 48$, saturating at $(4\alpha - 3)/12\alpha$, beneath the one-third ceiling that two essential complements impose on any finite capability. Right: the verification regimes of §7 at $g = 1$: while machine output requires human verification, labor is the single best-attributed factor at 0.375 of value against 0.208 for each stack factor; when the stack verifies itself, labor falls threefold to 0.125 and intelligence rises to 0.292.

7. The Verifier's Wage and the Bargaining Zone

The compression game hides one human monopoly inside H's outside option, and the present moment makes it worth isolating: verification. Machine output that no one can certify is, in most deployed settings, worth little; the signature of the current deployment phase is the human in the loop, the engineer reviewing the generated code, the radiologist signing the read, the lawyer who must stand behind the filing. Model the frontier at $g = 1$, all tasks automatable at $\alpha = 6$, under two regimes. Under human verification, machine output counts only when a human certifies it, so the productive coalition needs the entire set, intelligence, compute, energy, and the human verifier, while humans alone retain their manual productivity of 1. Under self-verification, the stack certifies its own output and the game is the compression game's endpoint. The exact division under human verification makes labor the single best-attributed factor on the frontier: $\phi_H = 2.25$ of the 6 units, a share of 0.375, against 1.25, a share of 0.208, for each stack factor. The flip to self-verification is one step in capability and the largest single redistribution on the whole path: labor falls to 0.75, a share of 0.125, threefold down a cliff of 1.5 units, a quarter of total value, while intelligence rises to 1.75 and 0.292. The last wage, the arithmetic says, is the verifier's wage, and the qualitative reason it may persist is the one Polanyi gave for why we know more than we can tell: certification trades on tacit knowledge, the capacity to recognize soundness without being able to specify the procedure, and tacit capacities are the hard residue of automation (Polanyi, 1966). The institutional reason is liability, which law assigns to persons; the verifier is partly a knower and partly an address for blame, and both components carry Shapley weight as long as the coalition needs them.

Attribution, however, is a point, and power is a zone, and the second result of this section is the distance between them. The core of a game is the set of allocations no coalition can profitably walk out on (Gillies, 1959). In the human-verification game, labor's core constraint is only its outside option: allocations are stable for any labor payment from 1, a floor share of 0.167, up to the entire product, because every coalition without the verifier is worth nothing and the verifier alone is worth 1. The Shapley point, 2.25 at a share of 0.375, sits inside that wide interval as the unique fair division; nothing in the stability of the game selects it. In the displacement limit, humans fully substituted, the game is the unanimity game on the three stack factors, and the core degenerates to the whole simplex: every split of the 6 units among intelligence, compute, and energy, including any one of them taking everything, is core-stable, while the Shapley point sits at 2 each, a share of one-third apiece, 0.333. Contribution is determinate and appropriation is indeterminate, and the indeterminacy is total where the production is most joint. This is the sense in which the market cannot answer the distributional question of the machine economy: prices in that zone do not measure contribution, they measure bargaining position, scarcity engineering, and institutional leverage, the things Hayek correctly said prices aggregate, which were never contribution even in the human economy (Hayek, 1945). The political map follows from the geometry of the zone rather than from the old class line: owners of deployable intelligence, owners of the bottleneck complements through which intelligence must act, and suppliers of whatever has become substitutable, with each boundary's rents set by how essential, and how absorbable, the factor behind it is. Writing P_i for what factor i's owner extracts and ϕ_i for its attributed contribution, the auditable quantity is the gap $P_i - \phi_i$, rent stated exactly, and the oldest intuition of political economy, that production is social while appropriation is private, survives its formalization intact: the Shapley value computes the social production, the core exhibits the private appropriation's degrees of freedom, and nothing in the mathematics closes the gap, because closing it is not a mathematical act.

8. An Index, and the Boundaries of the Audit

If pre-singularity economics is the trajectory of Ψ_t, intelligence's share of attributed value, then the honest empirical program is to estimate Ψ_t and its companions rather than to keep reading GDP, which was designed for other questions and answers this one badly; its own architect warned at the start that welfare and national income are different objects (Kuznets, 1934). GDP asks what was sold. The production-game frame asks what made the sale possible, and its dashboard is a vector of estimated shares: intelligence's share of value by sector, weighted across the economy; the same share inside the ideas-producing sectors specifically, since that is where the recursive loop would close; the automatable fraction of the task space, for which the exposure measurements already exist (Eloundou, Manning, Mishkin, and Rock, 2024); the share of frontier model improvement itself attributable to machine rather than human research labor, the recursion term; and the tightness of each bottleneck complement, compute concentration, energy queue depth, fabrication capacity, and the human verification burden per unit of machine output, whose decline is the cliff of §7 announcing itself in advance.

Every term of such an index is an estimated counterfactual, and the limits should be stated with the same flatness as the program. Coalition values must be modeled, from quasi-experiments, adoption discontinuities, and structural assumptions, and the assumptions rather than the arithmetic carry the conclusions. The value function is plural: an economy has a GDP game, a profit game, a welfare game, a capability game, and a survival game, and Ψ_t differs across them, so a single-number index is already an editorial act. And any such index, once published and used, becomes a target, at which point the regularities it summarizes begin to decay, which is Goodhart's law in its original monetary habitat (Goodhart, 1975) and in the audit-culture generalization that when a measure becomes a target it ceases to be a good measure (Strathern, 1997). An attribution index for the machine economy would be gamed by the most resourceful actors ever to game an index. That is an argument for building it with adversarial care and plural definitions, and it is also the strongest reason the construction belongs in public hands rather than only in the private dashboards where approximations of it certainly already run: the actors best placed to compute the true ledger are the actors with the least interest in publishing it.

The frame inherits objections from each tradition it crosses, and the serious ones mark the boundaries of what the audit can do. The deepest is that the whole construction presupposes a well-defined v, and the economy does not come with one. What counts as value is itself the contested object, and unpriced production, the care work and the maintenance and the commons, enters every official v at zero, from which the Shapley value will dutifully compute a zero share. The frame does not solve this. It sharpens it, by turning the choice of v into a visible premise rather than a buried one, so that the audit can be re-run under rival definitions to show whose contribution appears and disappears as the definition moves.

A second objection, Reiff's, attacks the morality of the average: the Shapley value pays for marginal contributions in orderings that never happened, and a person owed something is owed it for what they actually did (Reiff, 2013). The answer concedes the description and rejects the demand. Under synergy there is no fact of the matter about an "actual" individual contribution. The actual history produced a joint product and a set of dividends, and any division of a dividend is a convention; the Shapley convention is the one that treats orderings, hence positions, hence incumbency, as morally arbitrary, and Young's theorem prices the alternatives, since abandoning the averaging means abandoning marginality, symmetry, or exhaustion (Young, 1988). Read the value in the deflationary register throughout, as an audit baseline that makes deviations arguable, a floor for the conversation about desert rather than its conclusion.

Then there is the player problem. "Intelligence" names many things at once: weights, architectures, inference systems, agents, data, and the labor congealed in all of them, and the bloc treatment hides the internal ledger the way treating "labor" as a substance once did. The frame is at least self-similar under the objection, since the same machinery runs inside the bloc, over models and datasets and toolchains, and meets the same fog at the bottom that all attribution meets: inside a network, individual neurons are polysemantic, features smear across units in superposition, and the meaningful unit of contribution is itself an open research question (Elhage, Hume, Olsson, Schiefer, Henighan, et al., 2022). Attribution has a chosen grain at every level, and the choice of grain is part of the politics.

The largest limit is the one the two central results rest on: perfect complementarity. The compression game makes the stack factors jointly essential and individually useless, a unanimity structure, and that single assumption produces both the $\phi_I = \phi_C = \phi_E$ symmetry and the $1/(k+1)$ ceiling. Real production rarely sits at that corner. Energy, compute, and algorithms substitute for one another at the margin, a more efficient model buying back megawatts, a better algorithm buying back chips, and the standard instrument for this, the constant-elasticity production function with finite elasticity of substitution between the factors, is exactly the tool this paper does not deploy. With imperfect substitutability the equalities loosen, the three stack values separate by how essential each remains, and the ceiling moves, rising as any complement becomes more substitutable. The ceiling is therefore not a law of capability but a property of the unanimity game, and a capability path that erodes a complement's essentiality, lowering the effective number of strictly necessary complements, is a genuine route to a higher ceiling that the perfect-complement model excludes by construction. The symmetry and the ceiling should be read as the perfect-complement limit of a more general production game whose interior, where the elasticity of substitution is finite and the conclusions soften, is the untested generalization this frame most needs. The ownership-event thesis holds within the perfect-complement model; outside it, capability that dissolves essentiality competes with ownership as a way to move the ceiling, and which dominates is empirical.

Two further limits are macroeconomic and microeconomic twins. On the macro side, demand saturates, and spending migrates to whatever stays scarce and human-bound, which is Baumol's disease as the demand-side counterpart of the bottleneck theorem (Baumol, 1967). Where the cost disease bites, measured value concentrates in the unautomated sectors and Ψ_t stalls for reasons the capability curve never shows; Nordhaus's singularity tests, which keep returning "not yet" on the growth-acceleration evidence, are consistent with this drag (Nordhaus, 2021). On the micro side, verification may simply stay hard, tacit, and legally human, in which case the cliff of §7 never arrives, the verifier's wage compounds, and the frontier economy stabilizes in the human-verification regime, with labor holding the largest single share on the frontier that was supposed to abolish it. The frame's service is to make these futures the same diagram under different parameter motions rather than rival prophecies.

9. The One-Axis Economy

The compression thesis, stated at the end with the computations behind it, is this. The economy approaching machine intelligence does not become a one-product game in which only intelligence is worth owning. It becomes a one-axis economy, in which every factor is repriced by its relation to deployable intelligence, as amplifier, substrate, bottleneck, verifier, or substitute, and the repricing follows attribution arithmetic that can be computed in advance. Along that axis, a handful of results from the minimal games carry the rest. Joint production generates dividends that belong to no one, so every accounting that zeroes them is a hidden politics. As the automatable share of work grows, labor's attributed share collapses by an order while its absolute attribution barely moves, so the distributional outcome is decided by whether institutions pay shares or absolutes, a choice the arithmetic exposes and cannot make. Where intelligence and its complements are jointly essential, attribution cannot distinguish the brilliant factor from its dumb but essential complements, so the premium on intelligence is a fact about bargaining rather than about contribution, and the capability race drags a complement race behind it; loosen the joint essentiality and the values separate, which is the same statement read from its boundary. Under that perfect-complement structure intelligence's share of attributed value is capped at $1/(k+1)$ by k essential complements for any finite capability, so the Shapley singularity, $\Psi \to 1$, is reachable within the model only by redrawing the boundaries of ownership until the complements are inside the machine's owner, which is to say that the singularity worth watching, so long as the complements stay essential, is a merger wave wearing a research program's clothes; the alternative route, capability that dissolves a complement's essentiality and lifts the ceiling from inside, is the one the model sets aside and the world may not. And beneath all of it, the core's indeterminacy where production is most joint means the question of who gets what was never going to be answered by markets or by theorems, because in the zone where everything is everyone's joint product, distribution is chosen.

The failed cultural loop with which the seed of this paper began, that there are Marxists and no Shapleyans, can now be read as a diagnosis rather than a complaint. Marxism offered an identity and could not compute; the attribution tradition could compute and offered nobody an identity, so it sank into the plumbing, indispensable and unseen, running cost allocations and power indices and feature explanations while the distributional argument upstairs proceeded on vibes and flags. An economy in which the principal factor of production is about to change does not need a new flag. It needs the plumbing brought upstairs: public, plural, causally disciplined attribution of the joint product, an audit nobody is recruited by and everybody can check, under value functions chosen in the open, with the gap between contribution and appropriation published as the first number rather than buried as the last. The mathematics will not say who should own the machines, the megawatts, or the right to verify. It will say, exactly and to 4 decimal places, what each of them added under every story we are willing to specify, and an argument about justice that begins from that ledger is the only kind the coming economy will not be able to dismiss as arithmetic error.

References

Acemoglu, D., and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2), 3–30.

Aghion, P., Jones, B. F., and Jones, C. I. (2019). Artificial intelligence and economic growth. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda* (pp. 237–282). University of Chicago Press.

Alchian, A. A., and Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, 62(5), 777–795.

Baumol, W. J. (1967). Macroeconomics of unbalanced growth: The anatomy of urban crisis. *American Economic Review*, 57(3), 415–426.

Bresnahan, T. F., and Trajtenberg, M. (1995). General purpose technologies: 'Engines of growth'? *Journal of Econometrics*, 65(1), 83–108.

Brynjolfsson, E., Rock, D., and Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333–372.

Clark, J. B. (1899). *The Distribution of Wealth: A Theory of Wages, Interest and Profits*. Macmillan.

Cohen, G. A. (1995). *Self-Ownership, Freedom, and Equality*. Cambridge University Press.

Elhage, N., Hume, T., Olsson, C., Schiefer, N., Henighan, T., et al. (2022). Toy models of superposition. *Transformer Circuits Thread*. Anthropic.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of large language models. *Science*, 384(6702), 1306–1308.

Elster, J. (1985). *Making Sense of Marx*. Cambridge University Press.

Foerster, J., Farquhar, G., Afouras, T., Nardelli, N., and Whiteson, S. (2018). Counterfactual multi-agent policy gradients. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*.

Frank, R. H. (2016). *Success and Luck: Good Fortune and the Myth of Meritocracy*. Princeton University Press.

Gale, D., and Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1), 9–15.

Gillies, D. B. (1959). Solutions to general non-zero-sum games. In A. W. Tucker and R. D. Luce (Eds.), *Contributions to the Theory of Games, Volume IV* (pp. 47–85). Princeton University Press.

Goodhart, C. A. E. (1975). Problems of monetary management: The UK experience. In *Papers in Monetary Economics, Volume I*. Reserve Bank of Australia.

Harsanyi, J. C. (1963). A simplified bargaining model for the n-person cooperative game. *International Economic Review*, 4(2), 194–220.

Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519–530.

Heskes, T., Sijben, E., Bucur, I. G., and Claassen, T. (2020). Causal Shapley values: Exploiting causal knowledge to explain individual predictions of complex models. In *Advances in Neural Information Processing Systems 33*.

Holmström, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, 13(2), 324–340.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589.

Korinek, A., and Stiglitz, J. E. (2019). Artificial intelligence and its implications for income distribution and unemployment. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda* (pp. 349–390). University of Chicago Press.

Kuznets, S. (1934). *National Income, 1929–1932*. National Bureau of Economic Research.

Lundberg, S. M., and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30*.

Marx, K. (1867). *Capital: A Critique of Political Economy, Volume I*. Otto Meissner. (B. Fowkes, Trans., Penguin, 1976.)

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., et al. (2025). *The AI Index 2025 Annual Report*. AI Index Steering Committee, Stanford Institute for Human-Centered Artificial Intelligence.

Milgrom, P., and Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review*, 80(3), 511–528.

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1), 8–30.

Neumann, J. von, and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Nordhaus, W. D. (2021). Are we approaching an economic singularity? Information technology and the future of economic growth. *American Economic Journal: Macroeconomics*, 13(1), 299–332.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press.

Platz, J. von (2021). The principle of merit and the capital-labour split. *Economics and Philosophy*.

Polanyi, M. (1966). *The Tacit Dimension*. Doubleday.

Reiff, M. R. (2013). *Exploitation and Economic Justice in the Liberal Capitalist State*. Oxford University Press.

Ricardo, D. (1817). *On the Principles of Political Economy and Taxation*. John Murray.

Roemer, J. E. (1982). *A General Theory of Exploitation and Class*. Harvard University Press.

Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn and A. W. Tucker (Eds.), *Contributions to the Theory of Games, Volume II* (pp. 307–317). Princeton University Press.

Shapley, L. S., and Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, 48(3), 787–792.

Shubik, M. (1962). Incentives, decentralized control, the assignment of joint costs and internal pricing. *Management Science*, 8(3), 325–343.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell.

Strathern, M. (1997). 'Improving ratings': Audit in the British university system. *European Review*, 5(3), 305–321.

Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.

Wicksteed, P. H. (1894). *An Essay on the Co-ordination of the Laws of Distribution*. Macmillan.

Wolpert, D. H., and Tumer, K. (2001). Optimal payoff functions for members of collectives. *Advances in Complex Systems*, 4(2–3), 265–279.

Wood, E. M. (1989). Rational choice Marxism: Is the game worth the candle? *New Left Review*, I/177, 41–88.

Young, H. P. (1988). Individual contribution and just compensation. In A. E. Roth (Ed.), *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. Cambridge University Press.