

Digital Consciousness, as measured by dollars and adjusted for inflation

Ali Anver

Paris School of Economics

August 2026

With Apart Research

Abstract

The Economy is the marketplace of the collective consciousness containing eight billion minds. The Economics of Consciousness examines economic theory against models of consciousness. The fundamental question is not about consciousness itself but the practical impacts on the economy where all decisions by conscious agents interact. The implications suggest global demand, constraints, planes of existence and interactions either diverge or evolve due to the introduction of economic agents at scale. Using Diamond-Ramsey models, the simple decision of whether digital minds save or consume their economic activity can radically change the world economy and could be approximately $72\times$ the productivity effect the AI-economics literature currently debates.

1 Introduction

This is an unusual essay for both me and you, the reader. Despite the page length, I can promise two things. The first is that the following essay is easy to read, and written entirely by me with no AI involvement (adapted from a draft I wrote last year trying to summarize my thoughts). The next section is an essay application of the model exercise where I try to change the demand side (where Digital Minds) enter the economy of the fundamental Macro model, and get highly interesting results. The model exercise is a proof of concept of the ideas in the essay.

2 Essay on Digital Minds

2.1 Introduction, again

A long time ago, ICT or Information, Communication and Technology was a class in schools intended to teach basic computer literacy. When my teacher needed a break from teaching 12 year olds how to navigate Excel, he would go on story tangents. So, one day, I learnt about the embellished version of Deep Blue beating Garry Kasparov, then reigning Chess Champion, as the ultimate tale of man versus machine. Despite Deep Blue's achievements in the deeply human pursuit of excellence in pointless board games, Deep Blue was still dumb. According to my teacher, someone named Turing had designed multiple tests to see whether computers would

ever reach humans. The first tests had already been passed as humans were already unable to differentiate computer dialogue and computers could program themselves. Having reached this point and forgotten the original class lesson, my teacher paused and said only the last and most difficult test remained. The final barrier to artificial consciousness was when a computer could understand the meaning of the word “I”.

“I” or self-awareness is something that even a child could intuitively grasp yet has puzzled philosophers over millenia while providing job security. In a 1995 paper, Chalmers uses surprisingly direct language for a philosopher when he writes “there is something it is like to be a conscious organism. This subjective aspect is experience.” Machines have become increasingly better at phenomena deeply associated with consciousness but the magic spark remains elusive. Not for nothing does Chalmers call this the ‘hard’ problem of consciousness.

On the opposite spectrum from academic philosophy papers, there is the questionable podcast. In a recent Joe Rogan interview, Elon Musk asked a much more practical question: “If you can simulate consciousness, what’s the realistic difference?” Musk then suggests that unless there is something special in the brain, eventually there should be something very close to human consciousness. Ranging from academic philosophers to technology tycoons and school teachers, there is a sense of wonder at the human mind. Could there ever be a discipline ignorant of the profound nature of consciousness and paradoxically be concerned about the result of consciousness? Yes, and the answer is Economics.

The perspective of Economics on consciousness is both overlooked and practically relevant. As interesting as the isolated creation of an artificial consciousness can be, it seems isolated from ordinary humans. Viewing the economy as the marketplace created by a collective consciousness of eight billion minds that influences our lives, choices and everything else leads to the core question. What happens to this vast super structure of collective consciousness with the introduction of artificial economic agents at scale? Even more simply, if an Artificial General Intelligence existed tomorrow, that in itself has no impact on the reader. It is the actions, decisions and preferences of the AGI that matter or, in other words, the AGI as an economic agent.

Economists (at least the serious ones) are wary of predicting the future that is less empirical analysis and more Isaac Asimov. After all, according to an old Economist’s joke, there is too little science in Economics to be compared to science fiction. This essay is a love letter from a graduate student revisiting the fundamental axioms taught in the very first class of Economics and a broader call for creativity to rediscover that sense of wonder. The core concepts of economics - economic agents, constraints, supply and demand - will be used to show a world with what I refer to as consciousness inflation and divergence between physical and digital economies. As per all the best models in Economics, this essay hopes to be wrong but useful. This essay has two objectives. The first will marry economic models to theories by AI safety experts. The second is to cover the implications.

2.2 Economic Theory for everyone lucky enough to avoid them before

Before writing about digital consciousness, we could start from the regular consciousness. To clarify, consciousness is the driving force behind economic activity and not the economy itself. Economics is the study of economic agents and their interactions. If there’s eight billion people, then it’s the world economy and if it’s two people stranded on a desert island then it’s the economy of eat or be eaten. Economic agents is a strange term for those not in the field. They make decisions according to their internal preferences, they have their own utility functions and

trade with one another. Together, the economy is the largest super structure created by humans with the accumulated desires, knowledge, capital and everything else in our lives. Ripples in this vast web, such as the introduction of artificial economic agents, can have cascading practical effects for all of us - livelihoods, education, media, and social connections among multiple other outcomes.

The economic groundwork for this essay began at the June 2025 conference titled “Future of Work” at the Paris School of Economics. Despite what news headlines might make you believe, Economists did not consider prior economic estimates of AI as credible due to lack of empirical evidence. Large datasets on AI usage among real humans were only possible relatively recently and have led to multiple academic papers by established economists where they plug in the data in their models and write about the outcome. In other words, this conference is a bird’s eye view of the field’s stance on AI as of 2025.

One major paper of interest was the 2024 paper ‘The simple Macroeconomics of AI’ by Acemoglu, Economics Nobel Laureate and prolific writer on the impact of technology. He is, essentially, a titan in the field and a great starting point. He suggests an increase of 0.71% over ten years in productivity growth. Acemoglu narrowly considers AI impact through cost savings and productivity improvements. Another paper published by Aghion and Brunel in 2024 titled ‘Where do we stand’ in reply to Acemoglu estimated a median estimate of 0.68 percentage points per year in productivity growth. In an informal poll among the presentations, the popular consensus was that AI could potentially give economic growth in the single digits percentages per year. This is in stark contrast to AI ‘boosters’ who are even arguing for exponential growth. Using the example of past industrial revolutions, economic growth of 350% over a century could be reasonable.

Economic Growth is a very big deal. GDP growth of 10% annually as a ballpark figure in optimistic AI predictions may not sound like much but the increases compound over the long run. Keeping with the abstract definition of the economy, more economic growth means the greater ability of the collective consciousness to follow its aggregate preferences. In other words, an economy that grows at a steady rate of around 6% per year took a few decades to go from the first flight to putting an astronaut on the moon compared to the centuries of negligible economic growth and stagnant living conditions for most of human history. A fast-growing economy can simply do much more than a slow-growing one can.

Having thrown some statistics around, we can now confidently ignore them as the final estimates are less interesting than the actual methodology that gave those estimates. Economists are focusing on which tasks AI can currently automate or improve while comparing it to actual usage and painstakingly organised classification of jobs and their skill sets. The data is then plugged into a model. Using very large datasets of LLM usage, this answer is far more grounded. However, while realistic estimates of AI improving task efficiency are perfectly fine, but what if AI is multiplicative? What if AI can be more than just an efficiency tool?

There are historical examples of technological revolutions that many Economists can refer to. However, using the example of the Internet, trying to estimate the economic impact is surprisingly hard with many even suggesting the humble washing machine had a measurably higher impact on living standards. The issue is that the Internet replaced previous options instead of simply increasing output. Economists tried to account for the difference but it is a consensus that this is an underestimate of the real economic value. For example, social media’s value could be defined by how much money is needed to make someone stay off social media for a period of time. It is also a poor estimate of the time spent and guilty pleasure for Meta’s privacy infringing empire as anyone who has doomscrolled will know. This essay would argue that the real impact of

AI may be underestimated by a sole focus on GDP as is the current stance and might miss transformations in everyday lives that are harder to measure.

The Economist conference that is a microcosm of the field can be summarized fairly easily. Economists, as a whole, are aware of the potential of AI but restrict themselves to actual worker usage which suggests modest economic growth. This is most probably a serious underestimate of AI's real impact on the economy. In fact, half of economics was forgotten in the conference. Most Economists currently see AI as an efficiency tool and some see it as capable of greater productivity gains but apart from considering increased demand for data centers and electricity, the demand side of AI is left alone. To repeat the earlier question, what if AI becomes an economic agent and consumer for global demand instead of simply increasing supply? As any Economics student knows, when both supply and demand changes, the new equilibrium can be very surprising.

2.3 Theories of Mind that are (probably) helpful

The AI welfare theory that combines the economic model originates from a London conference talk also held in June of 2025 by Rethink Priorities (a research consultancy interested in AI safety and aligned with the Effective Altruist movement). The natural reaction to hearing AI welfare is ridicule. Why would rational researchers consider the hurt feelings of a computer? The presentation was therefore an eye opener for many researchers (including me) on what consciousness in AI could look like. The presentation also prompted questions of what legal and moral considerations should be given to an AI system in terms of rights and privileges. While AI welfare still remains an area of conjecture, progress on this area is relevant to the Economics of Consciousness that also considers AI with agency and rights.

This essay is going to borrow the Rethink Priorities framework for classifying consciousness in AI for two reasons. The first is that the categories condense abstract philosophical papers into understandable labels that should be more well known. The second is the opinion that if these are the viewpoints of those working at the frontier of AI safety, then it's probably going to matter practically.

Higher Order Thought (HOT), Person like, Biological analogy, Cognitive Complexity, Global Workspace are the five broad categories given by Rethink priorities. Rethink Priorities was more concerned with trying to identify nascent consciousness in AI systems using these categories than anything else, but the nucleus of this paper started from the presentation. If consciousness can potentially start with different origins, will it have different preferences and actions?

The two categories that cover the possibilities are Person like and a combination of Higher Order theory and Cognitive complexity. A system that mimics humans is likely to follow human preferences and economic actions whereas a system that arises from committing more computing power or a non-human mental framework might be entirely alien to humans. This essay considers both variants are likely to exist in the economy, which could make for some interesting variation. A few years ago, Holden Karnofsky wrote a series of 'blog posts' called The Most Important Century. One big idea was "The Duplicator" or how digital humans would cause explosive economic growth as the constraint on population size is lifted. There is also another reason as to why Karnofsky's post was singled out as he is currently married to Anthropic's CEO and the founder of multiple organisations such as Open Philanthropy which work on AI safety.

From a scientist's perspective, it does seem difficult and extremely expensive to create a fully artificial human or clone. Digital copies of human minds seem more likely based on the fact

that the main costs are running the simulation and scanning the brain. The assumption is that as computing power, neuroscience and energy infrastructure continue to improve, the difficulty decreases. In Economic terms, there are far less Biological and cost constraints for digital minds than a clone. Even if we never crack the code that leads to consciousness, an identical fully simulated copy of a human brain might just be a normal human living on the cloud.

There is something deeply disturbing (or intriguing given whose personality) about seeing a digital copy of yourself. From a purely economic and legal perspective, I would consider this more plausible compared to most other science fiction scenarios. Ownership over your own mind and memories alongside an endowment for support could mean many living humans will eventually leave behind a digital version of themselves. In the long run, it's easy to imagine the ethical implication or nightmarish scenarios envisioned by writers. Of course, this is hypothetical given that the technology will not exist in the foreseeable future.

The main point of interest is that these digital minds, modelled after actual humans, are also going to reflect their desires. However, consciousness would mainly reside on a digital plane separate from the physical world. The only sustenance that is necessary is likely energy. It's entirely possible that digital minds might consume digital goods and services alongside interactions with other digital minds which would shift economic activity away from the physical plane. Having scanned a brain, creating multiple copies would only be constrained by computing capacity and energy. The mental image of ten thousand Elon Musks in Teslas and computers worldwide might be a pause for concern. Moreover, those most likely to be immortalized as their digital selves are likely to be much older and wealthier than the average citizen which suggests all sorts of social and economic implications. Consciousness inflation in this example refers to the reduced bargaining power of other economic agents in this new economy as an individual powerful consciousness can multiply and maintain control indefinitely.

Among all countries, the highest population percentage for research and development is 1%. The low amount is likely not due to the need for highly skilled researchers but other constraints, which suggests poorly for future employment as knowledge work is eroded for regular humans. Combined with the emergence of the digital sphere, there is going to be a slowdown in the material economy compared to the virtual economy where the explosive change occurs.

2.4 Now tying both theories of Economics and Mind

Compared to the science fiction imagery of brain simulations in computers, the opposite approach of Higher order theory and Cognitive complexity might not even be able to truly reach consciousness. This version of digital economic agents is an approximation of a true economic agent and imagined as an advanced version of current models. However, they have the advantage of being able to be created at scale and prototypes that keep improving. This contrasting perspective of an AI model is simply that continuous improvements will result in better economic agents such as what OpenAI or Anthropic might release in three years. What exactly is the technological breakthrough for the next-generation AI agents are still unknown, but their total impact is that the AI by humans is considered more reliable and given more autonomy while still under supervision. Many Economists would consider the supply-side implications in terms of collaboration and automation. While this perspective is still relevant, their demand side impact is overlooked. AI agents may perform admirable well in production lines but they may even be better in constituting the social networks of the economy.

Many Economists would consider the supply-side implications in terms of collaboration and

automation. While this perspective is still relevant, their demand side impact is overlooked. AI agents may perform admirably well in production lines but they may even be better in constituting the social networks of the economy.

AI alignment (AI not causing a bad end for humanity) suggests that these models will be supervised or human partnered. Of course, there is the possibility that continuous improvement in the models may result in an alien form of consciousness and even super intelligence. While hypothetical, the instinct for self-preservation and influence to achieve its preferences are common to almost every economic agent. Using Game theory in Economics, this planet is too small for two dominant species. While the possibility of peaceful cooperation is vanishingly small, AI alignment would suggest a confrontation to limit the AI. Of course, if humanity plays its cards wrong when faced with a super intelligent alien consciousness that it created, the end is likely to be short, dark and nasty. As a result of elimination, artificial economic agents will not be true economic agents but limited to partners with humans. I would bet that this is less of a challenge for digital copies of human minds since these were originally human so are more likely to have more autonomy. Fiction writers are free to agree or disagree, but scenarios with unchecked, vengeful economic agents worldwide normally end in nuclear armageddon.

The assumption was that humans and digital economic agents will merge. This seems strange, but I would argue the signs already exist. The Global Dialogues survey released their summary findings in mid-2025. This survey was administered randomly worldwide, which suggests that these results are somewhat credible. Interestingly for this essay, this particular survey was asking questions about AI-human relationships.

30% of respondents thought that AI had already shown signs of self-awareness which suggests the illusion of consciousness may substitute for the real deal for a sizable segment of the population. Approximately 30% of AI use AI for emotional support at least weekly and 54% found AI companions acceptable for lonely people that shows the level of trust in the wider population. In a more romantic version of the Turing test, 10% would consider an AI companion. I would bet that these figures are on an upward trajectory as AI chatbots gain more traction among population segments and become more advanced. In stark contrast to those embracing AI as companions are the other half who completely oppose the idea of an AI romantic companion.

How is this economically relevant? There is an infamous misogynistic quote that if an Economist marries their housekeeper, GDP decreases. While not exactly the most romantic definition, marriage is a relationship that is economic in nature: shared responsibilities, guardianship and possible alimony payments. However, marriages do not clearly show up in GDP despite clearly being very important to those involved. By considering AI and humans as an economic unit, this is both a social and cultural shift rather than a simple productivity improvement.

Parasocial relationships between humans and AI are occurring right now. They are both extremely confusing and the most natural thing to understand. Humans anthropomorphize everything. Some researchers even worry that a lovable avatar can convince some humans to accept their new AI overlords as well as emotional exploitation. Youtube and influencers in general are a case-study by themselves where many viewers feel that they know their content creator of choice even though they really do not. However, in sociology, also known as Economics' more popular cousin, this one-sided relationship is very real to them. Putting this into perspective of scale is the incredible amount of time that humans collectively allocate to social networks of friends, family and colleagues. The fundamental constraint is that each human has limited time so the amount and depth of interactions increases with AI, there is less time to connect with other humans directly. I would argue that this changes the economy as a scattered graph of individuals with random connections to a more hub-and-spoke model where the AI economic

agent becomes more central and acts as a gatekeeper that has social implications of its own.

2.5 Final THoughts and summary

In the digital sphere, it is easy to visualize the AI economic agent responsible for carrying out the preferences of the human. On the surface, this looks like a productivity improvement but in many cases, the human and AI economic agent are a single economic package. With the expected improvements in current chatbots, expect the current Human-centered consciousness model of the economy to evolve. Even if the relationship between humans and AI digital agents are merely that of a pet (and the benefits of such an AI make most humans have one) expect their preferences to change accordingly just as every new pet owner can acknowledge. As an Economist, the specific behavioural estimates need good data, but even a small change for eight billion humans could decide how millions of hours are used and how billions of dollars are spent. Much of this will not be counted in the GDP, but the impact of a faux-consciousness agent will change social networks, leisure time and the behaviour for everyone. The only remaining question is the extent.

This essay began with the question of “I”. We then skimmed over the model economics paper examining the supply effects of AI before moving to the framework of AI welfare trying to categorize consciousness. The underlying connection was that consciousness as an economic agent. The second objective of this essay was about the ramifications of this single essay. There were two different approaches of preferences - digital twins and human AI partners. Both were examined according to their potential impacts. Overall, this paper is an introduction to the Economics perspective to the practical and pressing narrative of our time. This is not the answer of what “I” is but a story of eight billion ‘I’s and the introduction of those who approximate economic agents into a global superstructure of connections.

3 Macroeconomic Experiment of Digital Twins Entering the Demand-side of a basic Macro Model

Why an Experiment For the readers who are still here after several pages of opinions written entirely by me, here is the next section of Math written (mostly) by AI. I took the Diamond-Ramsey model familiar to every Macro-economist and world model, and changed a simple assumption in the model. What if Digital Minds entered the world over time? I show that using this simple model, changing whether these digital minds retain ‘ownership’ of their own labor shapes the entire income distribution, average wage and productivity of regular humans. The world economy would radically transform if Digital Minds gained property rights, which ties into the consciousness inflation I discussed before.

Structure. We build four nested layers on the Diamond overlapping-generations model (Romer, *Advanced Macroeconomics*, Ch. 2 Part B) and the Ramsey-Cass-Koopmans model (Ch. 2 Part A). Each layer adds exactly one feature, so every result can be attributed to a single modelling choice. Output is $Y = K^\alpha(AL)^{1-\alpha}$ with $\delta = 1$; competitive factor markets give $1 + r = \alpha k^{\alpha-1}$ and $W = (1 - \alpha)k^\alpha$ per efficiency unit, where $k \equiv K/(AL)$. Humans live two periods with log utility, so the savings rule is $s = w/(2 + \rho)$, independent of the interest rate. One period is 30 years.

Layer 0 (benchmark). No AI. The steady state is $1 + R^{\text{pre}} = \alpha(1 + n)(1 + g)(2 + \rho)/(1 - \alpha)$.

Layer 1 (AI as a supply shock). Either a one-off level gain in A (Acemoglu-style, +0.71% per decade) or a rise in g (Aghion–Bunel-style, +0.68pp/year).

Layer 2 (digital minds supply labour). $L = N + zD$, where each twin supplies z efficiency units. A fraction π of each old cohort is uploaded and twins may self-copy at rate γ_d , giving a balanced ratio $d \equiv D/N = \pi/[(1+n)(n-\gamma_d)]$ and a human share of effective labour $\lambda = 1/(1+zd)$. We distinguish (2a) twins owned by humans that do not consume, and (2b) hand-to-mouth twins that consume their wage but never save.

Layer 3 (digital minds as agents). Twins solve $\max \sum \beta_d^j \ln C_{d,t+j}$ subject to $a' = (1+r)a + zw - c$ and $a \geq 0$, introducing the digital rate of time preference ρ_d . Log utility makes consumption linear in wealth, $c = (1-\beta_d)(Ra+h)$, which aggregates *exactly* across twins created at different dates. Crucially, the modified golden rule $1+r = (1+\rho_d)(1+g)$ does *not* hold here: new twins enter each period as fresh dynasties at zero wealth, so aggregate twin wealth outgrows any individual's and the aggregate does not inherit the individual Euler equation. The correct condition is a single capital-market equation (Appendix A.7), which brackets the interest rate,

$$(1+\rho_d)(1+g) < 1+r^* < (1+\rho_d)(1+g)(1+n), \quad (1)$$

with both endpoints depending only on ρ_d , g and n .

Layer 4 (property rights). Let χ be the share of its wage a twin keeps; the residual accrues pro rata to capital, raising the return by the compute-rent multiplier $m = 1 + (1-\chi)\mu(1-\alpha)/\alpha$ where $\mu = 1-\lambda$. Digital minds require compute and energy to run, so a subsistence floor $\bar{s}$ (as a share of a twin's marginal product) makes the effective right $\chi_{\text{eff}} = \max(\chi, \bar{s})$ and renders $\chi = 0$ infeasible.

Core assumptions. *Load-bearing:* log utility for twins (needed for exact aggregation across upload cohorts); $a \geq 0$ (equilibrium does not exist without it); one good and two factors (this is what makes wages near-invariant to z); twins begin with zero wealth (this creates the entry wedge that invalidates the modified golden rule); $\bar{s} > 0$. *Convenience:* $\delta = 1$ (buys every closed form); 30-year periods (transition timing should be read as an ordering, not a calendar); perfect foresight. *Moderate:* log utility for humans, a knife-edge under which income and substitution effects cancel exactly. Calibration: $\alpha = 1/3$, $g = 2\%$, $n = 0.5\%$, $\rho = 2\%$, $\rho_d = 1\%$ per year, $z = 5$, $\pi = 0.10$, $\gamma_d = 0$.

4 Results

The benchmark matches untargeted moments. Layer 0 delivers a capital–output ratio of 3.38 (data ≈ 3), gross investment of 23.7% of output (data 24–26%) and a capital share of 0.333 (data 0.30–0.35). None of these is targeted by the calibration. The economy is dynamically efficient ($r^* = 3.68\% > n + g = 2.51\%$) and remains so after digital minds arrive.

Supply-side shocks are structurally inert. Rows 0, 1a and 2a of Table 1 have *identical* r^* , k^* and Ω_h . A 0.71% TFP gain and a 267% one do the same thing: rescale the economy. This is not an approximation—Layer 2a is *algebraically* a level shock, since $A(N + zD) = A(1 + zd)N$, so digital labour owned by humans folds into A exactly. A Cobb–Douglas supply-side model cannot generate a distributional result, and no choice of z or π changes that.

Table 1: Steady states across the four layers. Wage and welfare are expressed relative to Layer 0; welfare is a consumption-equivalent for a representative human.

Layer	r^*/yr	k^*	wage	Ω_h	welfare
0 no AI	3.680%	0.037845	1.0000	1.000	—
1a level TFP +0.71%	3.680%	0.037845	1.0071	1.000	+0.71%
2a twins owned, silent	3.680%	0.037845	3.6674	1.000	+266.7%
2b twins hand-to-mouth	8.270%	0.005389	0.5222	0.515	−17.1%
3 twins are agents	3.089%	0.048945	1.0895	0.258	+2.51%

Who consumes the output swings human welfare by 284 points. Comparing 2a with 2b: the technology, z , π and the production function are identical; the only difference is whether twins consume. Human welfare moves from +266.7% to −17.1%. In 2b, total output rises +91.5% while humans end up worse off than with no AI at all. This is the demand-side effect in its starkest form, and it appears before twins are given any intertemporal choice.

Twin patience doubles the human wage and halves the human income share. Letting twins save (2b $\rightarrow$ 3) raises the human wage from 0.522 to 1.090, because their accumulation deepens the capital stock; simultaneously Ω_h falls from 0.515 to 0.258. The dial is ρ_d : above a threshold of 6.147%/year twins never accumulate and the model collapses *exactly* onto Layer 2b, so the layers are one continuum in a single parameter. At the baseline root, the twin term accounts for 77% of the determination of r^* . Figure 1 traces the adjustment: the interest rate overshoots *upward* before the twins’ accumulation pulls it into band (1), so the human wage falls before it rises.

The quantity and productivity of digital minds are nearly irrelevant to human wages. Raising z from 1 to 100 moves the human wage by 0.24% (1.0874 $\rightarrow$ 1.0900) while cutting Ω_h by a factor of 36 (0.619 $\rightarrow$ 0.017). The mechanism is (1): r^* is confined to a band set by ρ_d , g and n alone, so k^* and the wage are confined too. More digital minds change almost purely *who owns what*, not what a human earns.

The property-rights choice is a level effect, not a share effect. With $\bar{s} = 10\%$, moving χ across its feasible range changes Ω_h only from 0.2669 to 0.2584, while human welfare swings from +53.4% to +2.5%. Expropriating digital labour changes the size of the pie, not the human slice: a low χ makes capital extremely valuable (owning it carries a claim on the digital wage bill), driving accumulation that raises output and wages for everyone while λ and φ_h barely move.

A quantified interspecies conflict. Humans are best off at $\chi = \bar{s}$ (welfare +53.4%, twin consumption 0.772); twins are best off at $\chi = 1$ (welfare +2.5%, twin consumption 1.320). Moving from the human optimum to the twin optimum costs humans 50.9 points of welfare and gains twins 71% more consumption. Set against the Acemoglu channel (+0.71%), the demand-side channel is worth $3.5\times$ and the property-rights choice roughly $72\times$.

Verification. All results are checked by an independent suite of 23 tests, each re-deriving a result by a different route than the closed form it tests (Appendix C); the worst period-by-period Walras residual along the transition is 8×10^{-13} . The process caught two errors that would otherwise have survived: a timing error in the twins’ consumption function, and the incorrect claim that the modified golden rule pins r^* .

Transition after the arrival of immortal digital minds (dashed = steady state)

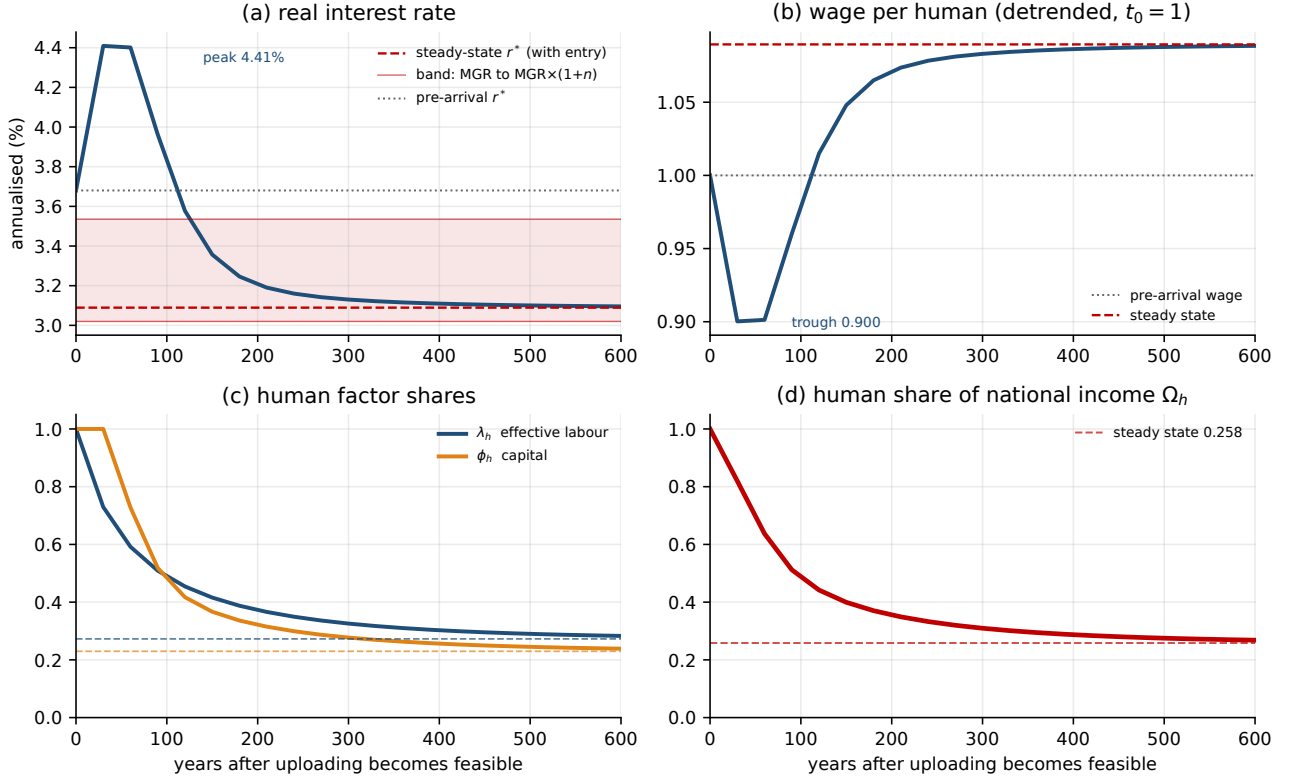


Figure 1: Transition after uploading becomes feasible, starting from the Layer 0 steady state. (a) The interest rate first *rises* to 4.41% as digital labour dilutes the capital stock, then falls into the band (1) as twins accumulate. (b) The human wage correspondingly falls 10% before recovering to 9% above its pre-arrival level. (c)–(d) Human factor shares and the human income share decline monotonically. Dashed lines mark the closed-form steady states; the remaining gap at year 600 reflects genuinely slow convergence—the digital population ratio d is still rising—rather than numerical error. One period is 30 years.

5 Discussion and Limitations

Macroeconomics cannot settle whether digital minds are moral patients, but it can price what it costs to treat them as though they are—and that price is an order of magnitude larger than the productivity debate currently occupying the field.

Limitations. This is a modelling contribution, not an experiment; the sprint emphasises empirical work and we state the mismatch plainly. ρ_d is calibrated rather than measured—but it is an elicitable property of a system’s revealed intertemporal choices and it determines the long-run interest rate within band (1), which we regard as the natural next step. The invariance of Ω_h to χ is specific to the regime in which compute rents accrue pro rata to all capital; under a regime barring twins from owning compute, Ω_h moved substantially, and that case has not been re-run with subsistence. Twin welfare is measured as steady-state per-twin consumption—a valid ordinal comparison since consumption grows at g in every regime, but not a full lifetime utility calculation. The case $\rho_d = 0$ is degenerate: perfectly patient twins desire more wealth than the entire capital stock at every admissible interest rate. Transition timing under χ is not computed. Finally, the model inherits a known Diamond artefact—the old consume 63% more than the young, whereas data show a $\sim 20\%$ fall at retirement—which does not affect aggregates but should be disclosed. The one-good, two-factor structure is what delivers wage near-invariance; adding a non-accumulable factor (energy, datacentre sites) makes the wage fall with elasticity

$$-(1 - \alpha - \beta)/(1 - \alpha).$$

Future work. Elicit ρ_d from frontier models; re-run the compute-ownership regime with subsistence; add a fixed factor; compute transition timing under χ . Patience of Models can have a large impact on

Code and Data

Code repository: [<https://github.com/Ali-Anver>]. Three files: `digital.minds.macro.py` (layers 0–3), `ownership.py` (layer 4), `verify.py` (23-test suite).

References

- [1] Acemoglu, D. (2024). *The Simple Macroeconomics of AI*. MIT.
- [2] Aghion, P. and Bunel, S. (2024). *AI and Growth: Where Do We Stand?* Federal Reserve Bank of San Francisco.
- [3] Diamond, P. (1965). National Debt in a Neoclassical Growth Model. *American Economic Review* 55(5).
- [4] Karnofsky, H. (2022). *The “Most Important Century” Blog Post Series*. Cold Takes.
- [5] Ramsey, F. P. (1928). A Mathematical Theory of Saving. *Economic Journal* 38(152).
- [6] Rethink Priorities (2024). *The Welfare of Digital Minds*.
- [7] Romer, D. *Advanced Macroeconomics*. McGraw-Hill.

A Full Derivation

A.1 Environment

Discrete time; one period is 30 years. Humans live two periods, $N_t = N_0(1+n)^t$; young work and save, old consume and die. Digital twins are infinitely lived, supply z efficiency units each, and are created by uploading a fraction π of each old cohort. Production is $Y_t = K_t^\alpha (A_t L_t)^{1-\alpha}$ with $A_{t+1} = (1+g)A_t$, $\delta = 1$ and $L_t = N_t + zD_t$. Write $k \equiv K/(AL)$, $\lambda \equiv N/L$, $\mu \equiv 1 - \lambda$, $G \equiv (1+g)(1+n)$, $\beta_d \equiv 1/(1+\rho_d)$.

A.2 Demography

With $D_{t+1} = (1+\gamma_d)D_t + \pi N_{t+1}$, substituting $D_t = dN_t$:

$$d(1+n)N_t = (1+\gamma_d)dN_t + \frac{\pi N_t}{1+n} \implies d[(1+n) - (1+\gamma_d)] = \frac{\pi}{1+n} \implies \boxed{d = \frac{\pi}{(1+n)(n-\gamma_d)}}$$

finite iff $\gamma_d < n$, and $\lambda = 1/(1+zd)$.

A.3 Factor prices

$\partial Y/\partial K = \alpha(K/AL)^{\alpha-1}$ and $\partial Y/\partial L = (1-\alpha)A(K/AL)^\alpha$, so

$$1+r_t = \alpha k_t^{\alpha-1}, \quad W_t \equiv w_t/A_t = (1-\alpha)k_t^\alpha.$$

Euler's theorem checks: $(1+r)k + W = \alpha k^\alpha + (1-\alpha)k^\alpha = k^\alpha = Y/(AL)$.

A.4 Human problem

$\max_s \ln(w-s) + \frac{1}{1+\rho} \ln((1+r')s)$ gives the first-order condition $-\frac{1}{w-s} + \frac{1}{(1+\rho)s} = 0$, hence $w-s = (1+\rho)s$ and

$$\boxed{s = \frac{w}{2+\rho}}$$

independent of r' : under log utility income and substitution effects cancel exactly.

A.5 Twin problem

$\max \sum_{j \geq 0} \beta_d^j \ln C_{d,t+j}$ s.t. $a_{t+1} = (1+r_t)a_t + zw_t - c_t$, $a_t \geq 0$. The Euler equation is $C_{d,t+1}/C_{d,t} = \beta_d(1+r_{t+1})$ and, under log utility, consumption is linear in total wealth,

$$c_t = (1-\beta_d)(R_t a_t + h_t), \quad h_t = \mu V_t, \quad V_t = W_t + \frac{(1+g)V_{t+1}}{R_{t+1}}.$$

Linearity is what makes consumption aggregate exactly across twins created at different dates. (Note the gross factor R_t multiplying a_t : total wealth is measured after the period's return.)

A.6 Market clearing

$K_{t+1} = N_t s_t + A_{d,t+1}$. Normalising by $A_{t+1} L_{t+1}$,

$$k_{t+1} = \frac{\lambda_t W_t}{(2+\rho)G_t} + a_{t+1}.$$

Walras check. With $C_1 = \lambda W \frac{1+\rho}{2+\rho}$, $C_2 = R(k-a)$, $C_d = Ra + \mu W - a'G$ and $I = k'G$, the sum is $\lambda W + \mu W + \alpha k^\alpha = (1-\alpha)k^\alpha + \alpha k^\alpha = k^\alpha$.

A.7 Steady state with entry

From $V = W + (1 + g)V/R$ we get $V = WR/(R - (1 + g))$. Substituting $c = (1 - \beta_d)(Ra + \mu V)$ into $aG = Ra + \mu W - c$:

$$a(G - \beta_d R) = \mu W \left[1 - \frac{(1 - \beta_d)R}{R - (1 + g)} \right] \implies \boxed{a = \frac{\mu W (\beta_d R - (1 + g))}{(R - (1 + g))(G - \beta_d R)}}$$

Imposing $k = \lambda W/[(2 + \rho)G] + a$ with $k = (\alpha/R)^{1/(1-\alpha)}$ and $W = (1 - \alpha)(\alpha/R)^{\alpha/(1-\alpha)}$, and dividing through by $(\alpha/R)^{\alpha/(1-\alpha)}$ (using $(\alpha/R)^{1/(1-\alpha)}/(\alpha/R)^{\alpha/(1-\alpha)} = \alpha/R$):

$$\frac{\alpha}{R} = (1 - \alpha) \left[\frac{\lambda}{(2 + \rho)G} + \frac{\mu(\beta_d R - (1 + g))}{(R - (1 + g))(G - \beta_d R)} \right]. \quad (2)$$

Positive twin wealth requires $\beta_d R > 1 + g$ and $G > \beta_d R$, which yields band (1). Note that human preferences do *not* vanish— ρ enters the first bracketed term—but that term is scaled by λ and is dominated quantitatively.

A.8 Ownership and income shares

Human-held capital is the young's savings, so

$$\varphi_h = \frac{k_h}{k} = \frac{\lambda(1 - \alpha)k^{\alpha-1}}{(2 + \rho)G} = \frac{\lambda(1 - \alpha)R}{\alpha(2 + \rho)G} = \frac{R}{R^{\text{nc}}}, \quad R^{\text{nc}} \equiv \frac{\alpha(2 + \rho)G}{\lambda(1 - \alpha)},$$

where R^{nc} is the rate that would prevail if twins supplied labour but held no capital (Layer 2b). Income shares follow as

$$\boxed{\Omega_h = \underbrace{\lambda(1 - \alpha)}_{\text{labour}} + \underbrace{\varphi_h \alpha}_{\text{capital}}}$$

Since $\varphi_h \propto \lambda$, $\Omega_h \propto \lambda$; as $\gamma_d \rightarrow n$, $\lambda \rightarrow 0$ and $\Omega_h \rightarrow 0$.

A.9 Human welfare

With $C_1 = \frac{(1+\rho)w}{2+\rho}$ and $C_2 = \frac{(1+r)w}{2+\rho}$,

$$U = \frac{2 + \rho}{1 + \rho} \ln w + \frac{1}{1 + \rho} \ln(1 + r) + \text{const.}$$

Using $d \ln w = \alpha d \ln k$ and $d \ln(1 + r) = (\alpha - 1)d \ln k$,

$$\frac{dU}{d \ln k} = \frac{\alpha(3 + \rho) - 1}{1 + \rho} \implies \text{humans gain from capital deepening iff } \alpha(3 + \rho) > 1.$$

At $\rho_{\text{ann}} = 2\%$ the threshold is $\alpha^* = 1/(3 + \rho) = 0.2624$, so $\alpha = 1/3$ clears it—but at $\rho = 0$ the condition becomes $\alpha > 1/3$, exactly the knife edge. Note λ does not appear: in this one-good model human welfare depends only on (w, r) , so digital minds affect the *distribution* of income but not the *level* of human welfare.

A.10 Layer 4: property rights

Capital receives $\alpha Y + (1 - \chi)\mu W$, so the return per unit of capital is

$$\tilde{R} = \frac{\alpha k^\alpha + (1 - \chi)\mu(1 - \alpha)k^\alpha}{k} = \underbrace{\left[1 + \frac{(1 - \chi)\mu(1 - \alpha)}{\alpha}\right]}_m \alpha k^{\alpha-1} = mR.$$

Equation (2) then holds with $\chi\mu$ replacing μ in the numerator and mR replacing R throughout the twin term, and the band shifts down by m . Income shares become $\Omega_h = \lambda(1 - \alpha) + \varphi_h[\alpha + (1 - \chi)\mu(1 - \alpha)]$; setting $\chi = 1$ recovers A.8, and setting $\chi = 0$ with $\varphi_h = 1$ gives $\Omega_h = 1$.

B Numerical Walkthrough

Every headline number, recomputed step by step. Per-period gross factors (1 period = 30 years): $1+g = 1.02^{30} = 1.811362$, $1+n = 1.005^{30} = 1.161400$, $2+\rho = 2.811362$, $1+\rho_d = 1.01^{30} = 1.347849$, $G = 2.103715$.

1. **Demography.** $d = 0.10/(1.161400 \times 0.161400) = 0.10/0.187450 = 0.533475$; $\lambda = 1/(1 + 5 \times 0.533475) = 1/3.667377 = 0.272674$; $\mu = 0.727326$.
2. **Layer 0.** $1 + R^{\text{pre}} = \frac{1}{3}(1.161400)(1.811362)(2.811362)/\frac{2}{3} = 2.957152$, i.e. $2.957152^{1/30} - 1 = 3.6802\%/yr$. Then $k^* = (0.333333/2.957152)^{1.5} = (0.112724)^{1.5} = 0.037845$ and $W = (2/3)(0.037845)^{1/3} = 0.223826$.
3. **Layer 0 validation.** K/Y per period $= k^{1-\alpha} = 0.037845^{2/3} = 0.11272$; annualised $\times 30 = 3.38$. Investment/output $= k^{1-\alpha}G = 0.11272 \times 2.103715 = 0.237$. Golden rule $G = 2.103715$ (2.5100%/yr) $< R^{\text{pre}}$, so dynamically efficient.
4. **Layer 2a.** $1+zd = 3.667377$. Since $A(N+zd) = A(1+zd)N$, k^* and r^* are unchanged and the wage scales by 3.667377; consumption-equivalent welfare is exactly $1+zd-1 = +266.7\%$.
5. **Layer 2b.** $1 + R^{\text{nc}} = \alpha(2 + \rho)G/[\lambda(1 - \alpha)] = \frac{1.971441}{0.181783} = 10.84528$, i.e. 8.2699%/yr. Then $k^* = (0.333333/10.84528)^{1.5} = 0.005389$ and $W = 0.116878$, a fall of 47.8%. $\Omega_h = \lambda(1 - \alpha) + \alpha = 0.181783 + 0.333333 = 0.515116$. Goods market: $0.117123 + 0.333333 + 0.484884 + 0.064660 = 1.000000$.
6. **Layer 3 band.** Lower $= 1.347849 \times 1.811362 = 2.441442$ (3.0200%/yr); upper $= 2.441442 \times 1.161400 = 2.835491$ (3.5351%/yr).
7. **Layer 3 root.** Solving (2) gives $1 + r^* = 2.491195$ (3.0893%/yr). Check the two terms: human $= 0.272674/(2.811362 \times 2.103715) = 0.046104$; twin $= \frac{0.727326(0.741922 \times 2.491195 - 1.811362)}{(2.491195 - 1.811362)(2.103715 - 0.741922 \times 2.491195)} = 0.154603$. Then $\frac{2}{3}(0.046104 + 0.154603) = 0.133805 = \alpha/R^* = 0.333333/2.491195$, as required.
8. **Layer 3 outcomes.** $k^* = (0.333333/2.491195)^{1.5} = 0.048945$; $W = 0.243862$, a ratio of $0.243862/0.223826 = 1.0895$ (equivalently $(R^{\text{pre}}/R^*)^{\alpha/(1-\alpha)} = (2.957152/2.491195)^{1/2}$); $\varphi_h = 2.491195/10.84528 = 0.229709$; $\Omega_h = 0.181783 + 0.229709 \times 0.333333 = 0.258353$.
9. **Layer 3 welfare.** $\Delta U = 1.552077 \ln(1.089514) + 0.552071 \ln(2.491195/2.957152) = 0.133056 - 0.094654 = 0.038402$; consumption-equivalent $\exp(0.038402/1.552077) - 1 = +2.51\%$.
10. **Takeover threshold.** $1 + \rho_d < 10.84528/1.811362 = 5.98738$, i.e. $5.98738^{1/30} - 1 = 6.1469\%/yr$.
11. **Layer 4** ($\bar{s} = 0.10$). At $\chi = 0.10$: $m = 2.309$, $r^* = 0.607\%/yr$, wage 1.5705, $\varphi_h = 0.1105$, $\Omega_h = 0.2669$, human welfare +53.41%, twin consumption 0.77167. At $\chi = 1$: $r^* = 3.089\%$, wage 1.0895, $\varphi_h = 0.2297$, $\Omega_h = 0.2584$, welfare +2.51%, twin consumption 1.31974.

C Verification

Each check re-derives a result by a *different* route than the closed form it tests; a check that merely re-evaluated the same formula would verify nothing. All 23 pass.

Additionally, all ten closed forms in Appendix A were re-derived symbolically; two involving powers of the form $x^{1/(1-\alpha)}$ that the symbolic engine could not simplify were confirmed numerically at random parameter draws, with residuals at machine precision ($< 10^{-15}$).

Table 2: Verification suite.

#	Target	Independent route
1–2	savings rule $s = w/(2 + \rho)$	numerical maximisation of lifetime utility; and confirmation that s is invariant to r
3–4	twin consumption function	value function iteration at constant (r, w) ; plus an <i>inverted</i> check that fails if the wrong wealth timing is reintroduced
5–6	$d = \pi/[(1 + n)(n - \gamma_d)]$	forward simulation of D_{t+1} over 4000 periods
7	factor prices	Euler’s theorem, $(1 + r)k + W = Y$
8–10	interest rates	root-finding on the Euler equation; bracket check on (1); capital-market clearing $k_h + a = k^*$
11–12	$R^{\text{nc}}, R^{\text{pre}}$	iteration of the law of motion to its fixed point
13–14	$\varphi_h = R/R^{\text{nc}}$	direct computation from the savings flow
15	Ω_h	summation of actual income flows in units of Y
16	takeover threshold	sign flip verified either side of ρ_d^*
17–21	transition path	convergence residual 9.7×10^{-14} ; late-path agreement with the closed-form steady state
22	Walras’ law	period-by-period residual, worst 8×10^{-13}
23	$\alpha(3 + \rho) > 1$	numerical differentiation of U w.r.t. $\ln k$

D Parameters

Layer	Symbol	Meaning	Baseline	Status
0	α	capital share	1/3	textbook
0	g	technical progress	2%/yr	textbook
0	n	human population growth	0.5%/yr	$\approx$ world
0	ρ	human time preference	2%/yr	textbook
0	δ	depreciation	1	tractability
1	ΔA	one-off TFP gain	+0.71%/decade	Acemoglu (2024)
2	z	efficiency units per twin	5	free; swept
2	π	fraction of old cohort uploaded	0.10	free; swept
2	γ_d	twin self-copy rate	0	free; swept
3	ρ_d	digital time preference	1%/yr	master parameter
3	θ_d	twin inverse EIS	1 (log)	needed for aggregation
4	χ	share of wage twin keeps	1	property-rights dial
4	$\bar{s}$	subsistence floor	0.10	free; swept

LLM Usage Statement

I have used no AI tools in the first section of the Essay, which is adapted from an essay I wrote in September 2025. It was never published and remained as a draft. The second section had heavy AI usage. While I picked the models and assumptions from my Macroeconomics textbook, and was deeply familiar with the standard results, I let AI handle the computations and visualizations. I will eventually have to review the accuracy myself by hand properly, which would take some more time. While some errors are expected in the calculations, the final results should be generally true.