Hedging the AI Singularity

Andrew Y. Chen

Federal Reserve Board

April 2025*

Abstract

This paper explores an alternative explanation for the high valuations of AI stocks: they may serve as hedges against a negative AI singularity scenario rather than merely reflecting future earnings growth. We present a simple disaster risk model where AI stocks provide positive payoffs precisely when other consumption collapses due to revolutionary AI advancements. This hedging value can justify high price-dividend ratios even without dividend growth. Our analysis highlights potential financial market solutions to AI disaster risk beyond universal basic income. Unlike previous work, this short paper is generated by prompting LLMs.

Keywords: Artificial Intelligence, Disaster Risk, Asset Pricing

*email:andrew.y.chen@frb.gov. ChatGPT-o1 and Claude-3.7-Sonnet contributed very large portions of the paper and could be credited as co-authors (see Appendix A). I thank Andrei Goncalves for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the position of the Board of Governors of the Federal Reserve or the Federal Reserve System.

1 Introduction

Artificial intelligence is advancing at a breathtaking pace. Recent developments like DeepSeek's R1 model (DeepSeek-AI et al., 2025), which achieved reasoning capabilities comparable to OpenAI's o1 despite development constraints, and Waymo's autonomous vehicles, which have logged millions of miles without human drivers, hint at a future where AI systems increasingly outperform humans across diverse domains. For workers, this rapid progress raises concerns about wage displacement and economic insecurity (Acemoglu and Restrepo, 2020; Knesl, 2023), as AI threatens to automate an ever-expanding range of tasks previously reserved for human labor.

While technological change has disrupted labor markets before, AI represents a fundamentally different kind of revolution. Unlike previous innovations that enabled humans to produce specific products or services more efficiently, there is no product or service that AI could not, in principle, create. This paper itself exemplifies this distinction—it was written entirely by AI systems, prompted with a few simple instructions.[1] The complete prompts and generation process are available at `https://github.com/chenandrewy/Prompts-to-Paper/`. Such capabilities stand in stark contrast to earlier revolutions like the internet, which primarily transformed information access and connectivity rather than intellectual production itself. Moreover, AI progress may be incredibly sudden—the technological singularity that researchers like Bostrom (2014) and Russell (2019) have warned about for years.

This paper examines how AI stocks might be priced in light of these risks. Specifically, we study the possibility that AI stocks are valued highly in part because they hedge against a negative AI singularity scenario where revolutionary AI advancements devastate consumption for the representative household. While conventional wisdom attributes high AI stock valuations to expectations of future earnings growth, our model demonstrates how such stocks could command substantial premiums even if their dividends never grow, purely because they provide positive payoffs precisely when other consumption collapses.

We should emphasize that we are not predicting a negative singularity will occur. Multiple viewpoints exist regarding the likelihood and nature of transformative AI scenarios, ranging from unalloyed optimism to profound concern (Bengio et al., 2024). Our purpose is merely to explore the asset pricing implications of such risks, which are worth considering even if the probability is small, given the magnitude of the potential impact. Similarly, we are not claiming that the hedging mechanism we describe is already priced into current AI stock valuations. Our model simply illustrates a possible mechanism through which singularity

[1] "We" refers to one human author and multiple LLMs. For a purely human perspective see Appendix A.

risk could affect asset prices, complementing the conventional growth narrative.

Our approach builds upon two interconnected strands of literature. First, recent work on artificial intelligence risk highlights concerns about a potential negative AI singularity scenario. Bengio et al. (2024) warn about extreme AI risks including the potential loss of human control over autonomous systems, while Bostrom (2014) outlines how a sufficiently advanced AI could pose existential threats to humanity. Jones (2024) explicitly frames the tension between AI-driven economic growth and existential risk as a crucial dilemma. These concerns reflect what Russell (2019) calls the "control problem" – ensuring advanced AI systems remain aligned with human values and welfare. If such risks materialize, they could significantly reduce consumption for the representative household through channels such as labor displacement (Acemoglu and Restrepo, 2020), disruption to way of life, or loss of meaning.

This AI singularity risk has important asset pricing implications that connect to the disaster risk literature pioneered by Rietz (1988) and expanded by Barro (2006). Just as rare macroeconomic disasters can explain high equity risk premia, the possibility of an AI singularity event may influence current asset valuations. Wachter (2013) demonstrates how time-varying disaster probability drives stock market volatility and return predictability, a framework that applies to uncertain technological transformations. In the specific context of technology risk, Zhang (2019) shows that firms with routine-task labor maintain a "replacement option" that hedges their value against unfavorable macroeconomic shocks. Similarly, Knesl (2023) finds that firms with high share of displaceable labor have negative exposure to technology shocks. Our model extends these insights to examine how AI-related assets might serve as hedges against negative singularity events, providing an alternative explanation for their high valuations beyond conventional growth narratives. More broadly, our work connects to the theoretical literature on how technology transitions affect economic welfare, as explored by Korinek and Suh (2024).

2 Model

We propose a simple model to explore the asset pricing implications of AI singularity risk. Our goal is to capture the essence of the argument that AI stocks may be valued highly in part because they hedge against a negative AI singularity scenario.

Our model features two types of agents. First, there are AI owners who are fully invested in AI assets and are not marginal investors in the stock market. Second, there is a representative household who is the marginal investor in publicly traded assets. Following the standard approach in asset pricing (**lucas1978asset**; **mehra1985equity**), only the representative household's consumption matters for equilibrium asset prices. We assume the household has constant relative risk aversion (CRRA) preferences:

$$U(C_t) = \frac{C_t^{1-\gamma} - 1}{1-\gamma}$$

where γ is the coefficient of relative risk aversion.

The representative household's gross consumption growth follows a simple process. In normal times, consumption growth is 1 (no growth for simplicity). However, the economy faces the possibility of "disasters" in the form of AI singularity events. When such a disaster occurs, gross consumption growth becomes e^{-b} where $b > 0$. This captures our key assumption that a revolutionary improvement in AI technology could be devastating for the representative household. While the AI owners capture the benefits of these technological improvements, the household experiences a significant decline in welfare through channels such as labor income displacement, disruption to way of life, or loss of meaning. At time $t = 0$, we assume no disasters have yet occurred, meaning the singularity has not yet happened. Importantly, multiple disasters may occur over time, reflecting ongoing uncertainty even after an initial singularity event.

We consider a publicly traded AI asset with the following characteristics. Before any singularity event, the asset's dividend represents a small fraction of aggregate consumption. Each time a disaster occurs, the dividend's fraction of consumption changes by a factor of e^h. Specifically, if we denote the dividend at time t as D_t, then:

$$\frac{D_t}{C_t} = \frac{D_0}{C_0} \cdot (e^h)^{N_t}$$

where N_t is the number of disasters that have occurred by time t. Importantly, h can be negative, capturing a scenario where even AI-related public equities might see their dividends shrink in absolute terms during a disaster. This reflects the possibility that the benefits of AI improvements may be concentrated in privately-held AI assets, with publicly traded AI companies potentially facing disruption themselves.

3 Price-Dividend Ratio Derivation

We now derive the price-dividend ratio for the AI asset in our model. Working in continuous time, we assume disasters (AI-singularity events) arrive according to a Poisson process with constant intensity λ, and the representative household discounts future payoffs at rate ρ.

The no-arbitrage price of the AI asset at time 0 is the expected present value of its future dividends, discounted by the stochastic discount factor. With CRRA preferences, the stochastic discount factor at time t is:

$$M_t = e^{-\rho t} \times (C_t/C_0)^{-\gamma}$$

Since consumption follows $C_t = C_0 \times e^{-bN_t}$ and dividends follow $D_t = D_0 \times e^{hN_t}$, where N_t is the number of disasters by time t, we can write the price-dividend ratio as:

$$P_0/D_0 = E_0 \left[\int_0^\infty e^{-\rho t} \times e^{(b\gamma+h)N_t} dt \right]$$

For a Poisson process with intensity λ, the moment generating function of N_t gives us $E[e^{\alpha N_t}] = \exp(\lambda t(e^\alpha - 1))$ for any constant α. Applying this with $\alpha = b\gamma + h$, we obtain:

$$P_0/D_0 = \int_0^\infty e^{-\rho t} \times \exp(\lambda t(e^{b\gamma+h} - 1))dt$$

This simplifies to:

$$P_0/D_0 = \int_0^\infty \exp(t[-\rho + \lambda(e^{b\gamma+h} - 1)])dt$$

For convergence, we need $\rho > \lambda(e^{b\gamma+h} - 1)$. When this condition is satisfied, the integral evaluates to:

$$P_0/D_0 = \frac{1}{\rho - \lambda(e^{b\gamma+h} - 1)}$$

This expression reveals how the price-dividend ratio depends on disaster risk. The term $e^{b\gamma+h}$ captures the combined effect of (i) the consumption drop during a disaster, which raises marginal utility by a factor of $e^{b\gamma}$, and (ii) the change in the AI asset's dividend, which scales by e^h.

4 Numerical Analysis

To illustrate the quantitative implications of our model, we present price-dividend ratios for various parameter combinations in Table 1. We set $\rho = -\ln(0.96) \approx 0.04082$, $\gamma = 2$, and $h = 0.20$. The table shows how the price-dividend ratio varies with the disaster size b and disaster probability λ.

The table reveals several important patterns. First, when the disaster probability is very low ($\lambda = 0.0001$), the price-dividend ratio remains around 25 regardless of disaster size, close to what we would expect in a model without disasters ($1/\rho \approx 24.5$). Second, as either the disaster probability λ or the disaster size b increases, the price-dividend ratio

Table 1: Price-Dividend Ratios for Different Disaster Sizes and Probabilities

	b				
λ	0.40	0.55	0.70	0.85	0.95
0.0001	25	25	25	25	25
0.005	31	36	48	81	200
0.01	42	71	775	Inf	Inf
0.015	66	1281	Inf	Inf	Inf
0.02	155	Inf	Inf	Inf	Inf

rises dramatically. For sufficiently large values of λ and b, the ratio becomes infinite as the convergence condition fails.

These results illustrate how AI assets that serve as hedges against negative singularity events can command high valuations, even if their expected dividend growth is modest. The hedging value comes from the positive correlation between the asset's dividends and marginal utility during disaster states. This mechanism offers an alternative explanation for high AI stock valuations beyond the conventional growth narrative.

5 Model Discussion

Our model intentionally abstracts from many real-world complexities to highlight a single key mechanism: how AI assets might serve as hedges against negative AI singularity events. While a more elaborate framework could incorporate additional features, the current streamlined approach serves our purpose well.

One important feature of our model is market incompleteness, which is implicit rather than explicitly modeled. This incompleteness is embedded in the disaster magnitude parameter b, which represents the net effect of (1) the AI disaster itself and (2) the AI asset's dividend response. If markets were complete, the representative household could buy shares in all AI assets, including privately-held ones, and not only fully hedge against the singularity but potentially benefit from it. However, in reality, most households cannot invest in many cutting-edge AI labs such as OpenAI, Anthropic, xAI, or DeepSeek. This market incompleteness limits the household's ability to hedge against singularity risk.

A more elaborate model might explicitly represent AI owners, their incentives, and their interactions with the representative household. Such an approach could address questions like: How might AI owners' incentives lead to a negative singularity? Would profit motives accelerate risk-taking in AI development? While these questions are certainly important, addressing them would require additional speculative assumptions that would decorate conjectures with mathematics rather than add substantive insights. The resulting model would be more costly to analyze and to read, while the core economics would remain essentially unchanged.

The current streamlined approach has the added benefit of leaving room for the human-written perspective included in Appendix A. By focusing on the essential mechanism, we avoid distracting readers with model elaborations that do not significantly enhance the central insight: that AI assets could be valued for their hedging properties against singularity risk, regardless of dividend growth expectations.

6 Policy Implications and Conclusion

Our analysis highlights how financial markets might price AI singularity risk and suggests potential market-based approaches to addressing concerns about AI-driven economic disruption. Traditional proposals often focus on universal basic income as a solution to technological unemployment. While such programs may have merits, our model points to complementary financial market solutions that have received less attention in the literature.

If AI stocks indeed serve as hedges against singularity risk, they offer a natural mechanism for households to protect themselves against potential AI-induced consumption declines. Policy could focus on broadening access to such hedging opportunities, perhaps through public investment vehicles that hold diversified portfolios of AI-related assets. However, the effectiveness of this approach is fundamentally limited by the market incompleteness we discussed earlier—as long as many cutting-edge AI developments remain in private hands, the hedging potential of publicly traded assets will be incomplete.

Our model suggests that the high valuations of AI stocks may partly reflect their role as hedges against a negative AI singularity rather than just optimistic growth expectations. This insight offers a more nuanced view of current market dynamics and suggests that investors may be more forward-thinking about technological risks than commonly assumed. Whether such hedging motives actually influence current valuations remains an empirical question, but our framework provides a theoretical foundation for investigating this possibility.

In conclusion, as AI continues its rapid advancement, understanding how financial markets price the associated risks becomes increasingly important. Our simple model demonstrates how AI assets that hedge against singularity risk could command high valuations even without dividend growth. This perspective complements the conventional growth narrative and highlights potential market-based approaches to managing AI risk that deserve greater consideration alongside traditional policy proposals.

References

Acemoglu, Daron and Pascual Restrepo (2020). "Robots and Jobs: Evidence from US Labor Markets". In: *Journal of Political Economy*.

Barro, Robert J. (2006). "Rare Disasters and Asset Markets in the Twentieth Century". In: *Quarterly Journal of Economics*.

Bengio, Yoshua, Geoffrey Hinton, Andrew Yao, Dawn Song, Pieter Abbeel, et al. (2024). "Managing extreme AI risks amid rapid progress". In: *Science* 384.6698. URL: https://arxiv.org/abs/2310.17688.

Bostrom, Nick (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

DeepSeek-AI et al. (Jan. 2025). "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning". In: *arXiv*. URL: https://arxiv.org/abs/2501.12948.

Jones, Charles I. (2024). "The AI Dilemma: Growth versus Existential Risk". In: URL: https://web.stanford.edu/~chadj/existentialrisk.pdf.

Knesl, Jiří (2023). "Automation and the Displacement of Labor by Capital: Asset Pricing Theory and Empirical Evidence". In: *Journal of Financial Economics* 147.2, pp. 271–296.

Korinek, Anton and Donghyun Suh (2024). *Scenarios for the Transition to AGI*. Tech. rep. NBER Working Paper.

Rietz, Thomas (1988). "The Equity Risk Premium: A Solution?" In: *Journal of Monetary Economics*.

Russell, Stuart (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking Press.

Wachter, Jessica A. (2013). "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" In: *Journal of Finance*.

Zhang, Miao Ben (2019). "Labor-Technology Substitution: Implications for Asset Pricing". In: *Journal of Finance* 74.4, pp. 1793–1839.

A A Purely Human Perspective

The following is the README.md file from the GitHub repository:

```
# Prompts-to-Paper

Writes a paper about hedging a negative AI singularity, using AI.

- `make-paper.py` writes a paper

- `plan0403-streamlined.yaml` contains the prompts

- `make-many-papers.py` runs `make-paper.py` many times.

The README is entirely human-written.  Please forgive typos and errors.

# Motivation

On March 8, 2025 I thought I should write a paper about hedging the AI
singularity.

I was worked up.  I had been repeatedly shocked by AI progress.  I was using
AI reasoning, vibe coding, and AI lit reviews in my daily life.  Six months
ago, I had thought each of these things is impossible.

What will happen in the next six years?!  Will my entire job be replaced by
AI? I have no idea.

But I do know that if there are huge disruptions, then tech stocks will
benefit.  So if anything bad happens to my human capital, I could at least
partially hedge.  Strangely, I hadn't heard about this concept before.

I asked a friend if he would be interested in working on this paper.
Unfortunately, he was busy with revision deadlines for the next month.

So, I thought I should use AI to write the paper.  It would be an elegant way
to make my point.  It would also hint at where the research process is going
in this strange age of AI.

This project was inspired by Novy-Marx and Velikov (2025) and Chris Lu et
al.  (2024).  These projects show how AI could generate massive amounts of
academic research.  My goal differs in quality over quantity.  I want to
generate just one paper, but one paper that (I hope) people find is worth
reading.
```

The Paper Generation Process

A natural question is: is this paper **really** written by AI?

I'd say the AI are junior co-authors.

If they were human, I would absolutely have to give o1, Sonnet, and ChatGPT Deep Research credit as co-authors. They did the math, writing, and literature reviews. Sonnet also wrote much of the code (via the Cursor AI IDE).

Of course, I did substantial hand-holding (see `plan0403-streamlined.yaml`). Writing this paper would have been easier if I had done more of the work myself. But that can happen with human co-authors too.

Perhaps in 12-24 months, AI can be equal co-authors. I wouldn't be surprised if I could just ask Gemini 3.5 to "write a model to formalize my intuition," and it would get it right.

Paper Iterations

Like human-written papers, the writing process was iterative. The first formalizations were terrible.

`plan0313-laborshare.yaml` contains prompts for a neoclassical growth model, where the capital share suddenly increases. ChatGPT-o1 patiently explained to me why this is a bad model.

me: > I thought there would be a wage risk effect that leads to higher investment for the more risk averse agent. High capital share means low or even no wage income.

ChatGPT-o1: > Below is an explanation of why one might **expect** a "wage-risk" channel in which **more** risk aversion could lead to **more** *additional* investment (relative to the no-jump benchmark) in the event that the capital share might jump to 1. However, this channel **does not operate** in the usual **representative-agent** version of the model---there, wage and capital income ultimately go to the *same* agent, so there is no meaningful "hedge" of wage risk. Instead, the wage-risk hedge arises naturally if you depart from the pure representative-agent setting (for example, if households receive labor income but must *choose* how much capital they own).

The final `plan0403-streamlined.yaml` uses a simplified Barro-Rietz disaster model, with two agents (though only one is relevant for stock prices). I went through several iterations of this model with Claude 3.7 Sonnet (thinking mode) and ChatGPT-o1. The only derivations I did myself were to check o1's work.

Literature Reviews

A key element was generating lit reviews (`./lit-context/`) to give the AI context. I used ChatGPT's Deep Research (launched Feb 2025) until I ran out of credits. Claude Web Search (launched March 2025, after I began the project) did the remainder.

These new products were a game changer. Both Novy-Marx and Velikov (2025) and Chris Lu et al. (2024) ran into hallucinated citations. OpenAI Deep Research and Claude Web Search had no problems if they were used with care.

More broadly, knowing how to use which AI and when was helpful for generating a good paper.

AI Model Selection

o1 did the theory, and sonnet thinking did the writing. It's well known that these are the strengths of these two models.

Sonnet thinking is OK at economic theory. But I found that it was not assertive enough. It led me down wrong paths because it was too eager to come up with some ideas that for my story (even if they did not make sense).

I briefly tried having Llama 3.1 470b do the writing. It was terrible! It would be extremely difficult to generate a paper worth reading that way.

I did not try many other models, in order to get this paper out quickly. Gemini 2.5's release, at the end of March 2025, was *hype*. I tried it out briefly and was impressed. But I gritted my teeth and ignored it. I'd never get the paper finished if I wanted to really try to explore alternative models.

Picking the best of N papers

The quality writing varies across each run of the code. There is both a good tail and a bad tail. Some drafts, I found quite insightful! Others, had flagrant errors in the economics.

Rather than try to prompt engineer an error free, insightful paper, I decided to just generate N papers and choose the best one.

Lessons about Research

A common response to Novy-Marx and Velikov (2025) is that "people are not ready for this." I heard concerns that peer review process will be inundated with AI-generated slop.

Working on this paper gave me a different perspective. It made me think about the fundamentals. I think the fundamentals are the following:

1. Readers want to learn something interesting and true.

2. Readers don't want to check all the math.

3. A system of author reputations makes 1 and 2 possible.

AI-generated papers don't change any of these fundamentals. Critically, item 3 made me quite cautious about putting my name on AI slop. As a result, I don't think AI-generated papers will change much about peer review, at least not the current generation of AI.

Limitations of the Current AI (April 7, 2025)

This will likely be out of date by the time you read it.

But right now, AI is like a junior co-author with a talent for mathematics and elegant writing, but sub-par economics reasoning. Put another way, the writing can fail to portray the mathematics accurately.

For example, 3.7 Sonnet sometimes fails to recognize that the economic model does not capture an important channel. This is a common scenario in economics writing (no model can capture everything). The standard practice is to dance gingerly around the channel in the writing. A decent PhD student can recognize this. But Sonnet cannot. Instead, 3.7 Sonnet will write beautiful prose about the channel anyway, even though it's not really being studied properly.

AI also cannot generate satisfying mathematics on its own (at least not satisfying to me). I tried asking o1 and Sonnet to generate a model to illustrate the point I'm trying to make. The resulting models were either too simplistic or did not lead to a clean analysis. They often introduced complications that I found unnecessary.

```
There could be models with capabilities that I missed.  But my sense is that
ChatGPT-o1 and Claude 3.7 Sonnet are close to the best for producing economic
research.

But more importantly, how long will these limitations last?

## The Future of AI and Economics Research

At some point, 2024-style economic analysis will be "on tap." You'll be
able to go to a chatbot and ask "write me a paper about hedging AI disaster
risk," and it will return you something like this paper (or perhaps something
better).

"Economics on tap" could be a disaster for the economics labor market.  It
would certainly mean that AI is an extremely cheap substitute for at least
some economists' labor.  It's possible that this would result in a strong
substitution away from labor.

The optimistic argument is that AI also complements economists' labor.
Perhaps, the number of economists will remain the same, but research output
increases in terms of both quantity and quality.

But I think there are reasons why total research output is limited.  Two key
factors in academic publishing are attention and reputation (Klamer and van
Dalen 2001, J of Economic Methodology).  Readers can only pay attention to
so many scholars.  These scholars, in turn, can only pay attention to so may
projects.

I'm not saying that I *expect* a disaster for the economics labor market.  But
it's definitely a scenario that economists should think about.
```

B Prompts Used to Generate This Paper

Each prompt consists of context and instructions. The context consists of the responses to the previous prompts, and may include literature reviews (all AI generated). For writing tasks (using Claude 3.7 Sonnet), a system prompt is also included.

For further details, see `https://github.com/chenandrewy/Prompts-to-Paper/`.

The system prompt and instructions are listed below.

System Prompt (model: claude-3-7-sonnet-20250219)

```
You are an asset pricing theorist who publishes in the top journals
   (Journal of Finance, Journal of Financial Economics, Review of
```

```
   Financial Studies). You think carefully with mathematics and
   check your work, step by step.

Your team is writing a paper with the following main argument: the
   high valuations of AI stocks could be in part because they hedge
   against a negative AI singularity (an explosion of AI development
    that is devastating for the representative investor). This
   contrasts with the common view that AI valuations are high due to
    future earnings growth. Since the AI singularity is inherently
   unpredictable, the paper is more qualitative than quantitative.
   The goal is to just make this point elegantly.

Write in prose. No headings and no bullet points. But do use display
    math to highlight key assumptions. Cite papers using Author (
   Year) format.

Be conversational yet rigorous. Favor plain english. Be direct and
   concise. Remove text that does not add value. Use topic sentences
   . The first sentence of each paragraph should convey the point of
    the paragraph.

Be modest. Do not overclaim.

Format the math nicely. Use we / our / us to refer to the writing
   team.
```

Instruction: 01-model-prose (model: claude-3-7-sonnet-20250219)

```
Draft the model description. The model is purposefully simple and
   captures the essence of the main argument. Only describe the
   assumptions. No results or insights.
  - Two agents
    - AI owners: Fully invested in AI, not marginal investors in
       stocks
    - Representative household: Marginal investor, only their
       consumption matters, CRRA
  - Representative household's gross consumption growth
    - is either 1 or e\\^(-b) (disaster)
```

```
      - A disaster is a revolutionary improvement in AI that is
        devastating for the household
      - Benefits of AI improvement are captured by the AI owners
      - For the household, labor income, way of life, meaning is
        lost
      - At t=0, no disasters have happened (singularity has not
        occurred)
      - Multiple disasters may happen, capturing ongoing uncertainty
         if a singularity occurs
  - A publicly traded AI asset
    - Dividend is a small fraction of consumption before the
      singularity
    - Each time a disaster occurs, the dividend's fraction of
      consumption grows by a factor of e\\^h
    - Meant to capture a worst case scenario, where the dividend may
        actually shrink in each disaster
      - i.e. AI improvements are concentrated in privately-held AI
        assets
```

Instruction: 02-result-notes (model: o1)

```
Find the price/dividend ratio of the AI asset at t = 0. Show the
   derivation, step by step.
```

Instruction: 03-table-notes (model: o3-mini)

```
Make a table of the price/dividend for b from 0.40 to 0.95 and prob
   of disaster from 0.0001 to 0.02. Here, fix h = 0.20, CRRA = 2,
   time preference = 0.96. If the price is infinite, use "Inf".
   Round to the nearest whole number.
```

Instruction: 04-resultandtable-prose (model: claude-3-7-sonnet-20250219)

```
Convert the notes in `02-result-notes` and `03-table-notes` into
   prose. The prose is intended to immediately follow `01-model-
   prose` and should flow naturally. Include the table.
```

Instruction: 05-litreview-prose (model: claude-3-7-sonnet-20250219)

```
Write a short two paragraph lit review based on the "prose-response"
    and "lit-" context.

Be careful to avoid incorrect citations. Make sure the papers cited
   make the claims they are cited for.
```

Instruction: 06-full-paper (model: claude-3-7-sonnet-20250219)

```
Write a paper titled "Hedging the AI Singularity" based on the "
   prose-response" context.

Title page:
- Title: "Hedging the AI Singularity"
- Abstract (less than 100 words)
  - Goal is to make a simple point
  - Secondary goal: bring attention to financial market solutions to
      AI disaster risk
  - At the end, say: unlike previous work, this short paper is
     generated by prompting LLMs.

The start of the Introduction is important. You need to bring the
   reader in, catch their eye, and establish credibility.

Start with background. Describe how AI progress is happening quickly
    (e.g. Deepseek R1, Waymo), and investors may be concerned about
   their wages being displaced (cite papers).

Then describe how technological change has occurred before, but AI
   is distinct because there is no product or service that AI could
   not, in principle create.  An example is the current paper, which
    is entirely written by AI, using a few short prompts. Provide a
   link to the github site, which is https://github.com/chenandrewy/
   Prompts-to-Paper/. This differs from say, the internet revolution
   . AI progress may also be incredibly sudden (the AI singularity).
    Include a footnote: "we" refers to one human author and multiple
    LLMs. For a purely human perspective see \\hyperref[app:readme
   ]\\{\\textcolor\\{blue\\}\\{Appendix \\ref\\{app:readme\\}\\}\\}.
```

```
Then describe what the paper does. It studies how AI stocks are
   priced, given that there is the risk that AI will destroy
   livelihoods and consumption.

Afterwards, the text should discuss:
- We are not saying a negative singularity will happen
  - But it is nevertheless important to consider this scenario
- We are also not saying that this hedging value is priced in
   already
  - Model illustrates a possible mechanism
- Related lit at end of Introduction
  - Cite papers in `05-litreview-prose`
  - Add Jones (2024) "AI Dilemma" and Korinek and Suh (2024) "
     Scenarios" if they're not already cited
- Model is the simplest possible to make the main argument
- Derivation of the key formulas
- High price/dividend ratios, even though dividends never grow
- A "Model Discussion" section that discusses natural model
   extensions and why they are not included
  - Market incompleteness is implicit but important
    - Implicit in the disaster magnitude 'b'
    - 'b' is the *net* effect of (1) AI disaster and (2) AI asset
       dividend
    - If markets were complete, representative household could buy
       shares in all AI assets (including private AI assets), and
       not only fully hedge but benefit from the singularity
    - In reality, most households cannot buy shares in many cutting
       edge labs (e.g. OpenAI, Anthropic, xAI, DeepSeek)
  - A more elaborate model would explicitly model the AI owners,
     their incentives, and interaction with the representative
     household
    - How might AI owners' incentives lead to a negative singularity
       ?
    - But wouldn't this just decorate speculations with math?
    - This would be costly to analyze, as well as to read
    - The core economics will remain the same
  - A short model analysis allows room for the human-written
     Appendix \\ref\\{app:readme\\}
```

```
- A "Policy Implications and Conclusion" section that discusses
   financial market solutions to AI disaster risk
  - These solutions are an alternative to UBI
    - Key economics: this hedge is limited by market incompleteness
  - These solutions to AI disaster risk are not discussed enough in
     the literature (cite papers)
  - Be very centrist (see below)

Text should avoid
- Being overly academic
- Politically-charged topics: sovereign wealth funds, industrial
   policy, redistribution, extolling free markets
- Overselling the model (it's just a simple illustration)
- Incorrect citations
  - Make sure papers cited make the claims they are cited for

Style Notes:
- Be conversational and direct, yet rigorous
- A touch of wit and wry humor are OK
- No bulleted lists
- No subsections (e.g. Section 1.2) though sections are OK (Section
   1)

Output a complete latex document, including preamble. Cite papers
   using \\cite, \\citep, \\citet. Use ‘template.tex‘ and keep the
   appendix that is already in the template.
```