

Transformative AI, existential risk, and asset pricing*

Trevor Chow
Stanford

Basil Halperin
MIT

J. Zachary Mazlish
Oxford and GPI

November 1, 2023
Preliminary

Abstract

We study the implications of transformative artificial intelligence for asset prices, and in particular, how financial market prices can be used to forecast the arrival of such technology. We take into account the double-edged nature of transformative AI: while advanced AI could lead to a rapid acceleration in economic growth, some researchers are concerned that building a superintelligence misaligned with human values could create an existential risk for humanity. We show that under standard asset pricing theory, either possibility – aligned AI accelerating growth or unaligned AI risking extinction – would predict a large increase in *real interest rates*, due to consumption smoothing. The simple logic is that, under expectations of either rapid future growth or future extinction, agents will save less, increasing real interest rates. We contribute a variety of new empirical evidence confirming that, contrary to some recent work, higher growth expectations cause higher long-term real interest rates, as measured using inflation-linked bonds and rich cross-country survey data on inflation expectations. We conclude that monitoring real interest rates is a promising framework for forecasting AI timelines.

*Email: tmychow@stanford.edu, basilh@mit.edu, john.mazlish@economics.ox.ac.uk

Thanks to Daniele Caratelli, Matt Clancy, Tyler Cowen, Tamay Besiroglu, Joel Flynn, Jackson Mejia, Laura Nicolae, Jett Pettus, Karthik Tadepalli, Phil Trammell, and participants at ETGP 2022 for very useful conversations on this topic and feedback. We are grateful to the George and Obie Shultz Fund at MIT for generous financial support; as well as to the Institute for Humane Studies, under grant number 017435.

1 Introduction

Background. Recent rapid progress in generative artificial intelligence has highlighted the possibility that humanity may soon develop “transformative AI”: AI technology that precipitates a transition comparable to the agricultural or industrial revolutions. Leading research labs like OpenAI and Google DeepMind bluntly declare their mission to build “artificial general intelligence” that can perform at or above human level on all tasks (OpenAI 2023; DeepMind 2023). The possibility of relatively short timelines for AGI is taken seriously by leading machine learning researchers, who in two recent surveys gave median estimates of 2061 and 2059 for the year in which AI will outperform humans at all tasks (Grace et al. 2018; Stein-Perlman, Weinstein-Raun, and Grace 2022).

The prospect of such transformative AI is a “double-edged sword”, in the language of Jones (2023). On the one hand, continued AI innovations like those which have occurred in protein folding or text generation could accelerate economic growth and improve well-being. In the same way that growth increased by roughly an order of magnitude with the industrial revolution, some have predicted that transformative AI automating all tasks would increase growth by another order of magnitude, with GDP growth rising to 30% or more per year (Davidson 2021; Davidson 2023). Indeed, standard models of economic growth extended to include human-level AI can predict even economic singularities: infinite output in finite time (Aghion, Jones, and Jones 2018; Trammell and Korinek 2020).

On the other hand, many in the AI research community and in the broader public are concerned that such powerful AI technology could create severe risks, even an “existential risk” for the human species. This concern is driven by the challenge of ensuring that smarter-than-human AI technology pursues goals matching human values, rather than pursuing unintended and undesirable goals: the “AI alignment problem” (Ngo 2022; Yudkowsky 2016). A 2022 survey of machine learning researchers found that – among those who chose to respond – the median believed there to be a 10% chance that human-level AI results in “human extinction or similarly permanent and severe disempowerment of the human species” (Stein-Perlman, Weinstein-Raun, and Grace 2022). This scenario is referred to as *unaligned* AI, in contrast to the growth-enhancing scenario with *aligned* AI.

Most economists, meanwhile, have been notoriously less likely to agree that transformative AI will be developed soon, less optimistic that aligned AI would radically accelerate economic growth, and less pessimistic that unaligned AI could pose an existential risk to human survival, on average (Korinek et al forthcoming).

This paper. We study the implications of transformative AI for asset prices and show how financial market prices can be used to forecast the arrival of such technology. In particular, we show that the prospect of transformative AI would predict a large increase in *real interest rates*, and would do so under expectations of either aligned or unaligned AI. As a result, to the extent that financial markets are efficient information aggregators, the level of long-term real interest rates can be used to forecast the development of transformative AI.

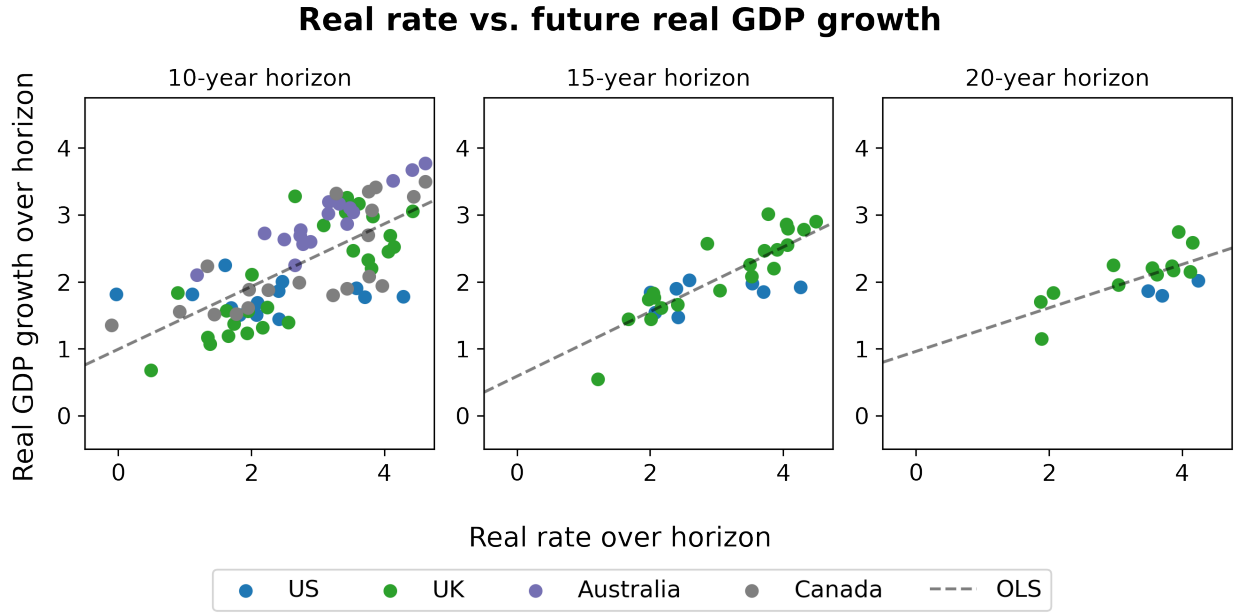


Figure 1: Real interest rates versus future GDP growth. Each subfigure plots a scatterplot of real interest rates of the titular maturity on the x-axis versus *future* annual GDP growth over the same horizon on the y-axis. Real interest rates are measured on the last trading day of each year. The scatter plots show all available data up through 2022, for the US (since 1999), the UK (since 1985), Australia (since 1995), and Canada (since 1991), where the end date of the data depends on the time horizon. More details on data sources are given in section 3.

This predicted rise in real interest rates is a basic implication of all modern asset pricing models, and is simply an application of the logic of consumption smoothing. Consider the case of *aligned* transformative AI: with the prospect of growth-induced high consumption and low marginal utility in the future, agents today would want to borrow more or save less, pushing up real interest rates at the relevant maturity. Similarly, if the market were forecasting future AI to be *unaligned* and to extinguish humanity, then there would also be no desire to save for the future (due to future extinction), again pushing up real interest rates at the appropriate maturity.

Empirical results on real rates. We offer new empirical evidence confirming that indeed, in the data, higher future growth increases real interest rates.

Measuring real interest rates is challenging. Existing work estimates real interest rates by using the nominal yields on nominal bonds and attempting to construct a measure of expected inflation to subtract from the nominal yields. The estimation of expected inflation needed for this, however, is difficult. We tackle this difficulty in two ways.

First, we use real yields from *inflation-linked bonds*, which provide a cleaner direct measurement of real rates compared to estimation based on nominal yields. To our knowledge, prior literature on the topic has not used real rates from inflation-linked bonds only because these bonds are comparatively new, with 20 or 30 years of data available.

Using these real yields directly from inflation-linked bonds, we show that higher real rates today indeed predict higher future GDP growth. Figure 1 shows the correlations for the US, UK, Australia, and Canada at the 10-, 15-, and 20-year horizons, comparing real interest rates over the relevant horizon with future GDP growth at the same horizon. While this data is merely correlational, and the data points are not independent of each other, it is suggestive evidence that growth and real interest rates are significantly linked.

In the body of the paper, we more carefully identify the effect of expected GDP growth on real interest rates using 1) changes in survey measures of long-term expected growth compared to changes in real rates and 2) quasi-exogenous shocks to growth expectations, using the patent shocks of Kogan et al. (2017). Our regressions confirm a clear link between expected growth and real interest rates.

Second, we use rich survey data on long-term inflation expectations from across 89 countries over the last 30 years to construct real interest rates from nominal bonds. The survey data is a unique dataset of forecasts from professional forecasters collected by Consensus Economics. By using forward-looking forecasts of inflation – rather than backwards-looking statistical measures of expected inflation, as in much of the literature – we are able to construct a larger panel of real interest rate data. Our results using this data are similar to the results using inflation-linked bonds, and combined, provide the best evidence we are aware of about the link between ex-ante real rates and expected growth.

Other asset prices. We also briefly discuss the implications of transformative AI for other asset prices. We highlight that the implications of transformative AI for *equity* prices are much more ambiguous than for real interest rates. Among other issues, while the prospect of *aligned* AI leading to rapid growth may increase equity valuations, expectations of *unaligned* AI on the other hand would lower valuations. The net effect is qualitatively ambiguous, making stocks more difficult to use as a barometer for market expectations for AI timelines without taking a quantitative stand on magnitudes. Moreover, even whether higher expected future growth from aligned AI raises or lowers equity valuations is itself unclear, and depends critically on the intertemporal elasticity of substitution: larger future cashflows due to economic acceleration may be more than offset by the higher discount rate previously discussed. We also touch on the implications for land and commodity prices.

Outline. The structure of this paper is as follows. In section 2, we define the “transformative AI” scenario under consideration, and provide a brief overview of existing related work, which may be less familiar to many readers with a background in economics. We also briefly review the relevant and burgeoning literature on the economics of AI. In section 3, we demonstrate the simple result that growth and death risk raise real interest rates in a very broad set of models. Section 4 presents evidence that higher growth expectations raise real rates today, and offers some commentary on existing analysis of this topic. Section 5 reviews relevant literature finding that mortality and savings behavior is

related. Section 6 discusses the implications of transformative AI for equities, land, and commodities. Section 7 concludes.

2 Defining transformative AI and relevant literature

In this section, we define the “transformative AI” scenario under consideration and provide context on existing research on the topic.¹ Much of this work may be unfamiliar to economists; familiar readers may wish to skip to section 3 after reviewing the definition of transformative AI in section 2.1, which is referenced throughout the paper.

2.1 Defining transformative AI

For the purposes of this paper, we consider the prospect of “transformative AI” as defined informally by Karnofsky (2016): artificial intelligence technology that has at least as profound an impact on the human trajectory as did the industrial revolution or agricultural revolution. As Karnofsky (2016) discusses, this term is similar to other concepts such as “artificial general intelligence” and “superintelligence”, but is intended to be more inclusive – capturing technology which is transformative, even if such technology is not able to match all human abilities.

We operationalize this definition of transformative AI by dividing two cases.

Definition (Aligned transformative AI). Aligned transformative AI is technology that causes growth in global GDP in excess of 30% per year.

Definition (Unaligned AI). Unaligned AI is technology that causes the extinction of humanity.

Our definition of aligned transformative AI follows Davidson (2021), who defines “explosive growth” as growth in gross world product of at least 30%, i.e. an increase in growth rates by an order of magnitude.² He discusses the possibility that transformative AI could cause such explosive growth. We take this as our benchmark for the effect of aligned AI, though given the unprecedented magnitude under consideration, these numbers should clearly be taken as rough approximations rather than as precise predictions.³ Our definition of the unaligned AI scenario follows the literature on the topic, which is summarized in section 2.4.

¹Later sections also discuss relevant economics literature in the context of our analysis: section 4 reviews existing work on the relationship between real interest rates and growth; section 5 reviews existing work on the relationship between real interest rates and mortality or catastrophe risk; and section 7 reviews relevant work on other asset prices.

²See also Hanson (2000).

³One alternative would be to use a definition in terms of task models (Zeira 1998) For example, aligned transformative AI could be defined as technology which can perform 100% of “cognitive” tasks that human perform, as in Davidson (2023). Because such technology would plausibly rapidly accelerate GDP growth – see Davidson (2023) for structural estimates – for the purposes of this paper the distinction in definitions is not important.

2.2 Forecasting transformative AI

Analysis of the possibilities for artificial intelligence has a long history. Good (1965) originated the concept of an “intelligence explosion”, a hypothesized phenomenon where AI systems gain the ability to improve their own algorithms and architectures, leading to recursive improvement and rapid increases in intelligence and power. Vinge (1993) originated and Kurzweil (2005) popularized the related concept of a “technological singularity”, referring to an acceleration in technological progress occurring so quickly that it would be difficult to predict *ex ante* how the world would look after. While these earlier analyses were mostly speculative, rapid progress in machine learning over the last decade has resulted in analysis more grounded in the reality of modern AI.

Cotra (2020) provides an influential benchmark forecast for the development of transformative AI. Her framework is based on estimating the number of computations the human brain can perform per second. She then forecasts forward trends in the computational power of computers, using long-run trends like Moore’s Law. She combines these to estimate the date by which computing power could match that of the human brain. Her analysis generates a distribution of estimates, with Cotra (2020) estimating a median arrival date of 2050 for transformative AI, and the updated analysis in Cotra (2022) forecasting a median of 2040. These estimates, however, are highly uncertain: the analysis of Cotra (2020) showed a 10% probability of transformative AI before 2030 and a 20% probability that transformative AI is not developed until after 2100.

Surveys of machine learning researchers are not too far off from the Cotra (2020) estimates. Grace et al. (2018) survey 352 AI researchers on “when unaided machines can accomplish every task better and more cheaply than human workers” and find a median of 2061. Stein-Perlman, Weinstein-Raun, and Grace (2022) run an updated version of this survey with 738 respondents, and find a median of 2058 for the same question. These results again come with significant dispersion.

Davidson (2023) uses a large-scale semi-endogenous growth model, *a la* Jones (1995), to forecast timelines for the development of transformative AI, and has a median forecast of 2043 for the development of transformative AI. This approach to forecasting the path of AI is analogous to the “dynamic integrated-climate economy” (DICE) modeling approach used in the climate literature: it is a computational integrated assessment model with an economics foundation.

Economists generally have been more cautious about forecasting the development of transformative AI. Korinek et al (forthcoming) survey economists and AI researchers about the probability of the development of “human-level machine intelligence”. The median response of AI researchers in this survey was before 2050; for economists, the median response was after 2070.⁴ Another survey by the Centre for Macroeconomics among European economists asked about the implications of progress in AI for global economic growth over the next decade. 64% of respondents answered growth would “increase (to 4-6%)”, though in open-ended comments many of these respondents noted that

⁴These results were sensitive to how the question was asked.

they thought the increase would be smaller; 36% of respondents answered growth would “remain unchanged” (CFM 2023).

2.3 The economics of transformative AI

A small but important economics literature has analyzed the economics of transformative AI. A larger literature studies the economics of prosaic AI more broadly.⁵

The seminal contribution to this literature is Aghion, Jones, and Jones (2018), who consider a range of possible scenarios for the effects of artificial intelligence on economic growth. Of particular interest is their result that if AI automates tasks in the *ideas* production function (rather than the goods production function), then speeding up the rate of automation is equivalent to speeding up the rate of population growth. It is well-known that in semi-endogenous growth models, the rate of growth on the balanced growth path is proportional to population growth. Under some conditions, they show there can be a ‘singularity’ in the sense of reaching infinite output in finite time. They also highlight the critical role of potential bottleneck tasks in preventing (long-run) growth explosions.⁶

Trammell and Korinek (2020) offer a literature review of how the many permutations of different assumptions on the role of AI in growth models have differential implications for growth and for macroeconomic aggregates like the labor share. Their review highlights a wide range of possibilities, depending on what is assumed.⁷

2.4 The alignment problem and the economics of existential risk

Concern over risks from artificial intelligence technology are widespread not just among the public and in fiction, but also among many scientists across many fields. This has recently been captured by the “Statement on AI Risk” signed by a long list of AI scientists and public figures, stating, “Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war” (Center for AI Safety 2023).

The basic concern is that it may be technically challenging to successfully program artificial intelligence technology in such a way that it behaves in line with human values. Just as software bugs can have large negative consequences in more mundane computer

⁵See, for example: Agrawal, McHale, and Oettl (2018), Acemoglu and Restrepo (2018), Beraja et al. (2023), and Brynjolfsson and McAfee (2014).

⁶Clancy (2022) offers a readable summary.

⁷Nordhaus (2021) builds a model of one specific set of assumptions of these many permutations, and considers whether the empirical predictions of those assumptions are born out in the data, as an attempt to forecast “are we approaching an economic singularity?”. Korinek (2019) analyzes the Malthusian implications of AI that can substitute for humans. Besiroglu, Emery-Xu, and Thompson (2022) show that if the capital share in the ideas production function increases, then the long-run growth rate also increases. They show that the capital share of deep learning for computer vision is substantially higher than the capital share for prior research technologies, and estimate that if the capital share of the economy-wide ideas production function rose to the level of that for computer vision, then growth would be three to eight times higher.

systems, software bugs in very powerful artificial intelligence systems could have correspondingly impactful negative consequences.

Ngo (2022) provides an overview of this “AI alignment problem” from the perspective of modern deep learning methods. Bostrom (2014) provides a book-length treatment from just before the deep learning revolution. Yudkowsky (2016) provides a conceptual argument for why the task of ensuring agents which are more intelligent than humans will act in line with human values should be perceived as challenging. Karnofsky (2021) offers a comprehensive and updated summary of these arguments. There is limited analysis of the AI alignment problem from an economics perspective. Hadfield-Menell and Hadfield (2019), Gans (2018), and Ely and Szentes (2023) are three exceptions.

In the economics literature, an important set of papers has analyzed how we should think about the tradeoffs between technology which brings positive benefits but creates existential risks. The important work of Aschenbrenner (2020) builds a model of directed technical change, extending the work of Jones (2016), where society can invest either in technology that increases consumption or technology that reduces the risk of death. Aschenbrenner shows that under reasonable parameters, optimally, existential risk follows a Kuznets-style curve: first rising, as society values consumption, and then falling, as the diminishing marginal utility of additional consumption is outweighed by the benefit of lower existential risk. Trammell (2020) shows a similar result in the context of an exogenous growth model. Jones (2023) summarizes these frameworks. Acemoglu and Lensman (2023) study the optimal regulation of technology adoption when that technology poses a (non-existential) disaster risk, motivated by AI technology. They show that if adoption is irreversible, then the path of adoption should be gradual, taking a ‘wait-and-see’ approach.

3 Real interest rates, growth, and mortality in theory

In this section, we demonstrate that real interest rates are connected to both expected economic growth and mortality across a broad range of modeling frameworks. The connection is driven by the same, simple economic logic across all modeling frameworks: higher expected growth and higher mortality risk both reduce the supply of savings, which pushes up real interest rates. We show this logic holds in the three classes of models, covering the modern asset pricing modeling frameworks:

- (i) Representative agent models
- (ii) Incomplete markets models
- (iii) Overlapping generations models

Since these results are known in the literature, we focus on results and intuition, and refer interested readers to relevant papers for full derivations. We also consider extensions with behavioral frictions (rule-of-thumb behavior and household myopia); nonstandard preferences (recursive preferences and habit formation); and models with new goods, which create nonstationary utility functions.

We conclude the section by emphasizing how the relationship between real rates and growth depends critically on the time horizon. In the short run – at business cycle horizons – *real interest rates that are too high cause low growth*, due to nominal rigidities. In the long run, when nominal rigidities fade, *high growth causes high real rates*. Hence, our empirical analyses in sections 4 and 5 focus on long-term real rates.

3.1 Representative agent models

It is well-known that in the canonical infinitely-lived representative agent model that the real interest rate is closely tied to growth and death.

The Ramsey rule. In the deterministic case with time-separable utility over the level of consumption, the real interest rate has a particularly simple expression – the canonical “Ramsey rule”:

$$r = \rho + \frac{1}{\sigma}g \tag{1}$$

Here, r is the real interest rate over some time horizon, $\rho > 0$ is the rate of pure time preference, $\sigma > 0$ is the elasticity of intertemporal substitution, and g is the growth rate of consumption. With σ usually calibrated somewhere between 0.2 and 2 – an issue to which we return – we see that higher growth implies a higher real rate. Higher existential risk shows up here as a higher rate of time discounting ρ , thus also implying a higher real rate.

Benchmark calibration under transformative AI. Consider briefly a benchmark calibration at the annual frequency with $\sigma = 1$ (log utility) and $\rho = 0.01$ for a back-of-the-envelope calculation. Then, a growth rate of 1% per capita would imply a real interest rate of 2% under the Ramsey rule – not far off the level seen in advanced economies today. Meanwhile, a transformative AI-induced growth explosion causing the growth rate g to rise to 30% (as defined in section 2.1) would raise real interest rates to 31%. This would be an unprecedentedly high level.

The Euler equation. The Ramsey rule analysis above importantly assumed away uncertainty, among other things. Consider an infinitely-lived household with expected, discounted, time-separable utility over the level of consumption. Denote the period utility function $u(C_t)$ where C_t is consumption and u has diminishing marginal utility, $u' > 0$ and $u'' < 0$. Denote the subjective discount rate as β , and the probability of death in period t as δ . In this representative agent framework, the probability of death δ is equivalent to the probability of extinction.

The resulting intertemporal optimality condition is the well-known Euler equation:

$$1 = \beta\delta\mathbb{E}_t \left[\frac{u'(C_{t+1})}{u'(C_t)} \right] (1 + r_t) \quad (2)$$

Suppose the path of consumption does not adjust, as in an endowment economy, for simplicity.

First, observe that higher death risk causes a higher real rate. A higher death risk is a lower probability of surviving until the next period, δ . A lower δ in (2) requires a higher real rate r_t . The intuition is that a higher probability of death shifts in the willingness to supply savings.

Second, observe that higher consumption growth, all else equal, also raises the real rate. Consider a shock that increases next-period consumption in at least one state of the world and shrinks it in none of them. Then, due to diminishing marginal utility, expected marginal utility $\mathbb{E}_t[u'(C_{t+1})]$ decreases, requiring a higher real rate r_t .

A risky shock – one which increases next-period consumption in some states but lowers it in others – does not unambiguously increase the real interest rate. The Euler equation (2) shows that what matters is *expected growth in marginal utility* – i.e. growth expectations taken over the risk-neutral measure. Thanks to diminishing marginal utility, this means that low-growth states of the world are weighted more highly.

For example, if consumption growth is lognormally distributed around the mean g and variance Var , then:

$$r_t = \rho + \frac{1}{\sigma}g - \frac{1}{2\sigma^2}\text{Var} \quad (3)$$

Here, $\rho \equiv -\ln(\beta\delta)$ and σ is once again the intertemporal elasticity of substitution.

The fact that the real rate is decreasing in the variance term shows how a shock which increases expected consumption growth could still push down the real rate if the shock also increases the variance of growth sufficiently. This again is due to the fact that what matters is expected growth in marginal utility. Diminishing marginal utility ensures that low-growth states of the world are weighted more highly. We return to this when discussing inequality.

3.2 Incomplete markets models and heterogeneous agents

The analysis above of the representative agent model demonstrates the importance of savings and borrowing decisions for understanding the effect of expected growth and mortality on real interest rates. This suggests it is worth considering how including realistic borrowing frictions affects the analysis.

Werning (2015) provides a benchmark analysis. In a world where idiosyncratic income risk does not covary with aggregate output, assuming isoelastic utility, and taking the “zero-liquidity limit” so that all agents are hand-to-mouth, then the *slope* of the relationship between growth and the real interest rate is the same, but the *level* is lowered.

The analog to the Ramsey equation (1) is:

$$r = \rho + \frac{1}{\sigma}g - \gamma_1 \quad (4)$$

All the terms are as before, with the addition of $-\gamma_1$. $\gamma_1 > 0$ reflects the idiosyncratic risk facing the “marginal saver”, which is the agent who *most* wants to save. The slope of the relationship between real rates and growth is still governed by the inverse of the intertemporal elasticity of substitution. Thus the real rate still increases with growth g and the existential risk probability embedded in ρ .

Moving away from the Werning (2015) benchmark, if idiosyncratic risk does covary with aggregate output, the relationship is more complicated. Auclert, Rognlie, and Straub (2018) show that an analog of the Ramsey equation can be written, for a particular form of idiosyncratic risk, as:

$$r_t = \rho + \frac{1}{\sigma\gamma_2} [C_{t+1} - \gamma_3 C_t] - \gamma_1 \quad (5)$$

Recall C_t is aggregate consumption at time t . If $\gamma_2 = \gamma_3 = 1$, then (5) is the same as (4), but they diverge due to cyclical income risk. γ_2 is the cyclical income risk of the marginal saver: the elasticity of individual income to aggregate income. γ_3 is the ratio of average cyclical income across all types, relative to the cyclical income of the marginal saver. Holding all else equal, higher C_{t+1} unambiguously increases the real rate.

Summarizing. While the relationship between output or consumption growth and real rates in these models is more complicated, a positive shock to growth still causes higher real rates. An increase in mortality risk has the same effect on the real rate as previously.

3.3 Overlapping generations models

The overlapping generations (OLG) framework is closely related to the incomplete markets framework of the prior section. Consider a simple version of this framework, where each agent lives for two periods and has log utility. There is Cobb-Douglas production technology with capital share α , population growth of n , and exogenous Hicks-neutral productivity growth of g . Then it can be shown that the analog of the Ramsey rule is:

$$r = \rho + g + \gamma_4 \quad (6)$$

Here, the coefficient on growth is 1, since log utility implies that the elasticity of intertemporal substitution is 1. The new term is γ_4 , which is a function of the capital share α and population growth n .⁸ Once again, the slope of the relationship between the real inter-

⁸Population growth does not affect the real rate in the canonical representative agent model, unlike the OLG model. Baker, De Long, and Krugman (2005) discuss how under imperfect altruism, population

est rate and growth is governed by the intertemporal elasticity of substitution; and the relationship between the real rate and mortality risk is direct.⁹

3.4 Recursive preferences and habit formation

Flynn, Schmidt, and Toda (2023) study the relationship between consumption growth and real interest rates under recursive preferences, such as the form studied in Epstein and Zin (1991) and Weil (1989). They show that the relationship is again determined by the elasticity of intertemporal substitution, where this elasticity must be defined appropriately given the recursive nature of preferences. The relationship between real rates and existential risk is unaffected by recursive preferences.

Bhamra and Uppal (2014), Hamilton et al. (2016), and Dennis (2009) study the relationship between consumption growth and real rates under habit formation. They show that with internal habits, the real rate is increasing in consumption growth. On the other hand, consider the extreme case with external habit where utility is determined entirely by the difference between individual consumption and average consumption. In such a world, a rapid acceleration in growth that lifts the consumption of all equally would not lower future marginal utility at all, and would not provide any incentive to save less or borrow more today. The real interest rate would be unaffected by the prospect of *aligned* transformative AI under this assumption, though it would still rise under the prospect of *misaligned*, extinction-causing AI. However, to the extent that preferences are not *purely* based on external habit, then rapid growth caused by transformative AI would still raise the real rate. This discussion emphasizes the importance of whether transformative AI will decrease marginal utility, rather than growth rates per se.

3.5 Myopic consumers

If all agents in the economy are fully myopic and do not recognize an impending acceleration in growth or extinction event, then real interest rates are unaffected by such prospects. However, even if *consumers* are fully myopic, as long as *financial markets* can foresee these events, then these prospects will be priced in to real interest rates. Dupraz, Le Bihan, and Matheron (2022) consider a model where consumers are myopic but financial markets are fully forward-looking.

3.6 New goods

Trammell (2023) shows that the introduction of new goods can keep marginal utility perpetually high, even as consumption grows without bound. In this case, there would not be any incentive to save less or borrow more today in response to higher expected

growth increases the real rate even in the representative agent model.

⁹In general without more assumptions than we have made so far, the OLG framework can lead to multiple or degenerate equilibria (see Acemoglu 2009, ch. 9).

growth. The real interest rate would be unaffected by the prospect of *aligned* transformative AI under this assumption, though it would still rise under the prospect of misaligned, extinction-causing AI.

3.7 Distinguishing the short and long run

The time horizon in question is critical to understanding the relationship between real interest rates and growth. In the short run, nominal rigidities play an important role, while they dissipate at a long enough horizon.

The business cycle literature shows that – under nominal rigidities – when real interest rates are “too” high, short-term growth is lower. Think about central banks raising the nominal policy rate. In a flexible price world, then all nominal prices and wages immediately jump to offset this increase. Inflation rises one-for-one with the nominal rate increase, leaving the real rate unchanged. However, in a world with nominal rigidities, then price or wage inflation cannot immediately jump to fully offset the increase in the nominal rate, so the real interest rate rises. Under standard assumptions about the transversality condition (Cochrane 2017), this jump in the real rate requires the *level* of consumption to jump down immediately, so that the *growth rate* of consumption can be higher, per the Euler equation (2).¹⁰

To be clear, at business cycle horizons, it is not high real rates *per se* which cause low growth, but real rates which are *higher than they would be* in a world without nominal rigidities. (The real rate in such a flexible world – the “natural rate of interest” – is of course unobservable, however.)

In the long run, on the other hand, it is thought that nominal rigidities relax as prices and wages have time to adjust, and it is higher consumption growth that causes higher real rates. Thus, in the short run, too high real rates cause low growth; in the long run, high growth causes high real rates.

Due to this flipping of signs, empirical work must carefully distinguish between short-run analysis and long-run analysis. This is an important issue with existing literature, an issue to which we turn in the next section – on the relationship between real rates and growth in the data.

3.8 Summarizing

Real rates and growth. The common thread across models is: if growth lowers the marginal utility of consumption, then growth increases real interest rates. We showed that this holds broadly across models, and highlighted two ways in which it might not. First, for a shock which increases consumption in some states of the world but lowers it in others, real interest rates could fall depending on how these net out. Second, marginal

¹⁰The textbook references on this topic under the canonical New Keynesian model are Galí (2015) and Woodford and Walsh (2005).

utility could stay high even with rapid consumption growth if utility is a function of relative consumption (i.e. external habit) or if the introduction of new goods keeps marginal utility high.

Real rates and mortality risk. Across all models, higher expected mortality risk raises real interest rates.

4 Empirical evidence on real rates versus growth: r vs. g

In the last section, we presented theoretical intuition for why higher expected growth would result in higher interest rates: expectations for such high growth would lead people to want to save less and borrow more today. In this section, we provide some simple empirical evidence that the predicted relationship holds in the available data.

We first show that the relationship between higher expected growth and higher real interest rates holds in the data: when expected growth is higher, so are real rates. Furthermore, positive changes in growth expectations are associated with positive changes in real rates. We next show evidence that “market expectations” are rational: when 10-year real rates are higher, subsequent 10-year *realized* growth is also higher. Finally, we offer new evidence from a set of quasi-exogenous shocks to growth expectations showing that higher growth expectations *causally* increase real rates.

Before presenting our results, we explain how our measures of real rates differ from those used previously in the literature.

4.1 Measuring real rates

The traditional approach. Most bonds historically have been *nominal*, where the yield is not adjusted for changes in inflation. Therefore, the vast majority of research studying the relationship between real interest rates and growth starts with nominal interest rates and attempts to construct real rates from the nominal rates.

Recalling that real interest rates are nominal interest rates minus expected inflation, such methodology requires estimating inflation expectations to subtract from the nominal rates. However, constructing measures of inflation expectations is extremely difficult.

Papers with this approach typically construct inflation expectations using a backwards-looking statistical model – usually simply a rolling AR(1) forecast based on past inflation. This is the approach used in the careful archival work of Schmelzing (2019); and in the analyses of Lunsford and West (2019) or Borio et al. (2022).

However, such an approach is inherently backwards-looking and fails to capture the forward-looking nature of inflation expectations. For example, consider the environment as of September 2023 where inflation was falling rapidly from its highs of the previous year. An AR(1) forecast using US data would assume that the year-over-year 3.7% CPI inflation rate was an appropriate forecast for the year ahead. However, more direct measurements of inflation expectations show substantially lower inflation expectations. For

example, the Survey of Professional Forecasters shows a consensus inflation forecast of 2.7%. In short, crude autoregressive statistical models often diverge sharply from more direct measures of inflation expectations, for time periods when such ground truth is available.

Finally, another – sometimes severe – problem with measurement of real rates using historical bonds is credit risk. While modern sovereign bonds from countries like the US are closer to risk-free, this is not the case for all sovereign bonds, and especially was not always so historically. This is relevant, for example, in the long-run historical trends estimated by Schmelzing (2019). He estimates a steady long-run decline in real rates using historical sovereign nominal bonds. Besides also finding an explanation in declining time preference (Clark 2007; Stefanski and Trew 2022), this plausibly reflects a long-run decline in credit risk. For example, the estimates of Schmelzing show a sharp rise in real rates during the Napoleonic wars. It seems natural to suspect this reflects heightened credit risk on sovereign bonds during the conflict, rather than a true increase in risk-free real interest rates.

Our primary approach: inflation-linked bonds. We take a more direct approach.

For our primary analysis, we instead use market real interest rates from *inflation-linked bonds*. These are bonds for which the payout is adjusted for realized inflation. As a result, the yields on these bonds directly reflect ex ante *real* interest rates, not nominal interest rates. This allows us to avoid needing to estimate inflation expectations.

To our knowledge, prior literature analyzing the determinants of real interest rates has not used data from inflation-linked bonds only because these bonds are comparatively new. In the United States, inflation-linked bonds (known as Treasury Inflation-Protected Securities, or TIPS) only began to be issued in 1997, for example. Many countries do not issue such bonds at all.

The yields on inflation-linked bonds do not perfectly reflect real interest rates because of various wedges, but these wedges are plausibly not too large. First, yields on such bonds also reflect term premia: risk compensation for uncertainty about the future path of real rates (Christensen, Lopez, and Rudebusch 2010). Additionally, such bonds are commonly thought to offer a (negative) convenience yield due to their relative illiquidity; and they have some embedded optionality due to deflation floors, among other complexities related to the bond structuring. See D’Amico, Kim, and Wei (2018), Christensen, Lopez, and Rudebusch (2010), and Fleckenstein, Longstaff, and Lustig (2014) for more. Beyond the literature finding that such wedges are not too large, the noise introduced is likely smaller than the noise introduced by using nominal bonds and may be stable enough to not disrupt long-horizon analyses.

Inflation-linked bond data. For the US, we use the fitted real yield curve from Gürkaynak, Sack, and Wright (2007), as updated by the Federal Reserve and available online. These fitted rates are available at 10-year, 15-year, and 20-year horizons, since January 1999.¹¹

¹¹Indeed, they are available at any horizon up to 20 years.

For the UK, we use the fitted real yield curve produced by the Bank of England. The 10-year and 15-year horizons that we use are available since January 1985; the 20-year horizon since June 1986.¹² For Australia and Canada, we use fitted 10-year real rates from Augur Labs. These data are available since 1995 and 1991, respectively. We also use Bloomberg “generic 10-year inflation indexed bond” series data from France, Israel, Sweden, Chile, Mexico, South Africa, Brazil, Japan, Germany, and Italy (with varying start dates).

Supplementary data: consensus forecasts. As a supplementary approach, we use consensus forecasts of inflation from a much wider cross-section of countries to construct real rates from nominal rates. By using explicit inflation forecasts, rather than a backwards-looking statistical model, we avoid the criticisms we just levied towards the traditional approach.

We use data from Consensus Economics on long-run inflation – at the 10-year horizon – to subtract from 10-year nominal bonds for 89 countries.

Adjusting for credit risk. For both the inflation-linked bond data and the consensus forecast implied real rates, we subtract measures of credit risk (i.e., the risk of default) in order to best measure the “risk-free” rate.

We have data on 10-year USD credit default swaps for 62 countries in our sample, ranging over different sample horizons. Some of the data comes from Bloomberg and some comes from the replication package of Longstaff et al. (2011). Following the data in Arellano, Bai, and Lizarazo (2017), we assume a 60% recovery rate R on defaulted bonds and approximate the risk-neutral default probability as:

$$\mathbb{P}(\text{default}) = 1 - \exp\left(\frac{-S * T}{1 - R}\right)$$

S represents the CDS spread and T the maturity. Annualized risk-neutral default probabilities are then subtracted from the 10-year interest rate to produce a risk-free rate.

This approximation is imperfect: in particular, assuming a uniform recovery rate across all times and countries is a simplification. Furthermore, since CDS spreads are in USD, they may not precisely reflect default risk on the local currency bond, which is what our 10-year interest rates measure.

However, these adjustments are superior to doing nothing, and particularly when default risk is mild, it remains a good approximation. Our results suggest that making some adjustment for default risk is essential, but exactly how that adjustment is made is not critical, implying the results are fairly robust. For example, if we follow the even simpler approach of Damodaran (2023) and just subtract CDS rates from the 10-year interest rate our results remain significant in the same direction.

¹²Since the 25-year horizon is only available since January 1998, there is not yet enough data to include it in our analysis.

4.2 Real rates and *expected* growth

All theoretical results presented above demonstrate the positive relationship between (risk-free) real interest rates and expected consumption growth. Therefore, we begin our analysis by looking at the relationship between our measures of 10-year real rates and 10-year expected growth. We use measures of both expected *GDP* growth and expected *consumption* growth. If the personal expected consumption growth of marginal investors is better approximated by an economy's GDP growth than consumption growth, the relationship between expected GDP growth and real interest rates could be stronger than the relationship with economy-wide expected consumption growth.

Our data on 10-year expected GDP and consumption growth comes from the same Consensus Economics dataset of 89 countries mentioned above. Surveys of 10-year expectations (for growth and inflation) are conducted twice a year before 2014 and quarterly since then.

Given that our interest rate data does not always align with the exact dates of survey dates and that our default-risk adjustments are imperfect, our measures of risk-free rates around times of extreme volatility are noisy. To avoid outliers, we remove observations where the *change* in expected GDP growth between consecutive surveys is >0.5 pp and where the change in real rates between consecutive observations is >3 pp. The signs of coefficients in the regressions remain the same and mostly remain significant regardless, as shown in appendix X.

Proceeding to the results, we run regressions of the following forms:

$$r_{i,t} = \alpha + \beta \mathbb{E}_t(g_{i,t}) + \epsilon_{i,t} \quad (7)$$

$$\Delta r_{i,t} = \alpha + \beta \mathbb{E}_t(\Delta g_{i,t}) + \epsilon_{i,t} \quad (8)$$

The dependent variable is either the level or the one-period change in country i 's 10-year risk-free interest rate; the independent variable is either the level or the one-period change in their 10-year expected GDP or consumption growth. We denote inflation-indexed measures of real rates r , while nominal interest rates less consensus inflation expectations are $i - \mathbb{E}(\pi)$. Expected consumption growth is $\mathbb{E}(g_c)$ while expected GDP growth is $\mathbb{E}(g)$. We use Δ to denote the change in a variable's value across Consensus survey dates and Δ_t to denote the change in a variable's value from one consensus survey to the survey t years (and zero months) later. Standard errors are Newey-West to account for auto-correlation.

Note that if there is no endogeneity and the country representative agent has standard power utility, the coefficient β measures their inverse elasticity of intertemporal substitution.

Table 1 shows the results from (7) where the regression is in levels.

The first two columns have our sample of inflation-indexed interest rates as the dependent variable; the third and fourth column use nominal interest rates less inflation expectations. Coefficients on both expected consumption growth and expected GDP growth are uniformly positive and significant.

Table 1: Expected Growth vs. Real Rate

	r	r	$i - \mathbb{E}(\pi)$	$i - \mathbb{E}(\pi)$
$\mathbb{E}(g_c)$		0.65*** (0.07)		0.19*** (0.04)
$\mathbb{E}(g)$	0.60*** (0.08)		0.19*** (0.04)	
const	-2.09*** (0.19)	-2.13*** (0.18)	-2.04*** (0.11)	-1.97*** (0.13)
Observations	418	396	2036	1704
R^2	0.21	0.27	0.01	0.01

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

The size of the coefficients when using nominal interest rates less expected inflation is about a third of the size of the inflation-indexed β coefficient. Some of this discrepancy may be due to Consensus surveys of inflation not accurately reflecting market expectations, some due to noise in our measures of default-risk. The nominal interest rate sample is much larger and includes many more emerging market countries. Both inflation risk and default risk are a larger component of emerging market interest rates, making our measures less precise. However, measurement error likely also exists in the level of inflation-indexed real rates, due to the default-risk adjustments.

As a result of this imprecision in the *level* of real rates, we believe the best estimates comes from regressions of changes in real rates on changes in expected growth. Tables 2 through 4 present our regression results where the dependent variable is the *change* in real interest rates and the independent variable is the *change* in growth expectations. Table 2 focuses on changes across consecutive surveys; table 3 has changes across two-years; and table 4 has changes across five-years.

There are pros and cons to focusing on the shorter versus longer horizon changes. The advantage of looking at shorter horizon changes is that it gives us the largest sample. The advantage of looking at longer horizon changes is that they are more likely to purge the noise from short-term variation in risk-premia (such as CDS liquidity).

All β coefficients are positive and all are significant at the $p < 0.05$ level other than the coefficient in the fourth column of table 3, which is still significant at the $p < 0.1$ level. Estimates of the inverse IES range from 0.2 to 1, with an average estimate of 0.53. The average within only the inflation-indexed bonds regressions is 0.56. We do not place too much stock in a specific estimate of the inverse IES, but the clear and consistent pattern of positive and significant β coefficients shows that the basic mechanism underlying the link between expected growth and risk-free real interest rates holds in the data. Given our large cross-country sample and our improved estimates of risk-free real rates we consider this to be some of the best existing evidence on the link between changes in growth and changes in real rates, a fundamental relationship in macroeconomics.

Table 2: Change in Expected Growth vs. Change Expected Real Rate

	r	r	$i - \mathbb{E}(\pi)$	$i - \mathbb{E}(\pi)$
$\Delta \mathbb{E}(c)$		0.50** (0.22)		0.23** (0.09)
$\Delta \mathbb{E}(g)$	0.99*** (0.26)		0.42*** (0.12)	
const	0.07** (0.03)	0.05 (0.03)	0.03** (0.01)	0.02 (0.02)
Observations	389	370	1945	1636
R^2	0.06	0.02	0.01	0.01

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 3: 2-Year Change in Expected Growth vs. 2-Year Change Expected Real Rate

	r	r	$i - \mathbb{E}(\pi)$	$i - \mathbb{E}(\pi)$
$\Delta_2 \mathbb{E}(c)$		0.44*** (0.15)		0.28* (0.15)
$\Delta_2 \mathbb{E}(g)$	0.41*** (0.15)		0.33** (0.16)	
const	0.00 (0.08)	-0.03 (0.08)	0.11** (0.06)	0.09* (0.05)
Observations	293	284	1546	1326
R^2	0.03	0.05	0.01	0.02

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4: 5-Year Change in Expected Growth vs. 5-Year Change Expected Real Rate

	r	r	$i - \mathbb{E}(\pi)$	$i - \mathbb{E}(\pi)$
$\Delta_5 \mathbb{E}(c)$		0.63*** (0.14)		0.82*** (0.21)
$\Delta_5 \mathbb{E}(g)$	0.40** (0.18)		0.89*** (0.22)	
const	-0.37*** (0.14)	-0.31** (0.13)	0.22 (0.19)	0.05 (0.14)
Observations	183	183	1009	890
R^2	0.03	0.10	0.04	0.08

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

The low R^2 values for all regressions do point to the fact that the basic Ramsey equation for real interest rates is unlikely to be the full story: much else seems to be contributing to real-rate variation. However, we would note that the R^2 values are highest for the link between the 5-year change in expected consumption growth and the 5-year change in real rates (under either measure). Since this is the least noisy estimate of the theoretical relationship of interest, we are encouraged that it has the highest R^2 value. The inverse IES estimates on those regressions are slightly higher; 0.63 using inflation-indexed bonds and 0.82 using nominal rates less inflation expectations.

4.3 Correlational evidence

In this subsection, we present some brief correlational evidence showing that real rates and future *realized* growth are tightly linked in the available data. The link between real rates and *realized* growth relies on growth expectations representing rational forecasts. Therefore, given the above evidence that real rates respond to changes in expected growth, the evidence we now provide is evidence that those growth expectations were indeed rational.

Using inflation-linked bonds. Using the real interest rates from the fitted zero-coupon yield curves, we compare with future GDP growth, by country. We do this at the 10-year, 15-year, and 20-year horizon where data is available across the four countries, in figure 1. We take the real rate from the last trading day of the year, and compare it to future GDP growth over the appropriate horizon. We use future GDP growth rather than consumption growth for simplicity, as the two are highly correlated, especially at long horizons. The figures also show lines of best fit.

We do not perform intensive statistical analysis for these correlations, but the figures present a strong relationship between long-horizon real interest rates today and future economic growth over the next 10-20 years. To our knowledge, this very simple stylized fact relating yields on inflation-linked bonds and realized future growth is novel.

Beyond being a simple correlation – i.e., the relationship is not causally identified – several properties of the data should be noted as caveats. Growth and real rates are both persistent variables; the data points contain overlapping periods, since they are measured at annual frequencies; and growth rates in these four countries are cross-sectionally correlated.

Additionally, while we have X, Y, and Z data points at the 10-, 15-, and 20-year horizons respectively, many other factors affect real interest rates and evolved substantially during this period. Population growth has slowed, retirement has lengthened, and populations have aged (Auclert, Malmberg, et al. 2021; Benzell, Kotlikoff, et al. 2015; Benzell and Brynjolfsson 2019). International capital markets may have become more integrated. However, for the purposes of our analysis – studying a potential shock of the magnitude of transformative AI – if markets were forecasting transformative AI, then, we might reasonably expect that this massive of an increase in the growth rate would drown out the

impact of any other factors.¹³

4.4 Causally identified evidence from patents

In this subsection, we use novel patent announcements to construct an instrument for growth expectations, and show that announcements for new patents increase future growth and also real interest rates.

In progress.

5 Empirical evidence on real rates versus mortality risk

In section 3, we presented theoretical intuition for why higher expected mortality or existential risk would result in higher real interest rates: a heightened probability of death would lead agents to want to save less and borrow more today. In this section, we present a review of already-existing work from a disparate set of literatures which provide evidence in support of the theory.

As a preliminary comment, we clarify that we study the relationship between real interest rates and the probability of truly *existential* risks – the probability of human extinction. We contrast this with the large literature on “rare disasters”, which studies events that have a differential impact on risky assets like equities compared to their impact on risk-free bonds. “Disaster risk” thus provides a potential explanation for the equity premium puzzle (Rietz 1988; Barro 2006; Gourio 2008; Gabaix 2012; Pindyck and Wang 2013).¹⁴ While disaster risk is about events that *differentially* affect stocks versus bonds, existential risk is about events that eliminate agents, thus affecting the return on stocks and bonds equally (and therefore cannot contribute to explaining the equity premium puzzle): existential risk sets the return on both to -100%.^{15,16}

¹³Taking the univariate pooled OLS results from the real rate data far too seriously, the fact that the 10-year real rate in the US ended 2022 at 1.6% would predict average annual real GDP growth of 2.6% over the next 10 years in the US; the analogous interest rate of -0.2% in the UK would predict 0.7% annual growth over the next 10 years in the UK. Such growth rates, clearly, are not compatible with the arrival of transformative aligned AI within this horizon.

¹⁴For a slightly different but related analysis, see Caldara and Iacoviello (2022) on measuring geopolitical risk.

¹⁵There is also a literature on the relationship between violent, non-existential conflict and asset prices. Hirshleifer, Mai, and Pukthuanthong (2023) use natural language processing techniques to study war discourse in newspaper articles and the relationship with equity prices. Ferguson (2008) and Bialkowski and Ronn (2017) provide narrative evidence of the effect of the world wars on financial markets. Rexer, Kapstein, Rivera, et al. (2022) study the relationship between violent conflict and sovereign nominal bonds. Leigh, Wolfers, and Zitzewitz (2003) as well as Wolfers and Zitzewitz (2009) study the relationship between the Iraq War and financial markets.

¹⁶Another literature that studies a non-existential disaster risk and financial markets is the climate literature. Giglio, Kelly, and Stroebe (2021) provide a review.

5.1 Individual mortality risk and savings

In the theory reviewed in section 3, the mechanism by which higher expected mortality risk increased the real interest rate is by reducing savings. With higher probability of nonexistence in the future, agents have lower incentive to save for the future, and this reduced supply of savings increases the real interest rate.

In this subsection we provide evidence on the mechanism: we review existing work showing that reduced mortality risk at the individual level causally increases savings (or equivalently, increases investment). While evidence of lower savings at the *individual* level need not raise interest rates at the *economy-wide* level, it does provide evidence for the mechanism through extinction risk increases interest rates.

One example comes from Baranov and Kohler (2018), who study the provision of a new AIDS therapy in Malawi which caused a significant increase in life expectancy. Using spatial and temporal variation in where and when these therapeutics were rolled out, it was found that increased life expectancy results in more savings and more human capital investment in the form of education spending.

A second example comes from the informational experiment, again in Malawi, of Ciancio et al. (2020). The authors provide information to correct pessimistic priors about life expectancy, and found that higher life expectancy directly caused more savings, via investment in agriculture and livestock.

A third example comes from testing for Huntington's disease, a disease which causes a meaningful drop in life expectancy to around 60 years, in Oster, Shoulson, and Dorsey (2013). Using variation in when people are diagnosed with Huntington's, the authors find that those who learn they carry the gene for Huntington's earlier are 30 percentage points less likely to finish college, which is a significant fall in their human capital investment.

A fourth piece of evidence comes from survey data in the US. Heimer, Myrseth, and Schoenle (2019) find, cross-sectionally, that pessimistic survival beliefs are correlated with a lower savings rate. This is true even after controlling for risk preferences, cognitive ability, and socioeconomic factors.

5.2 Aggregate mortality risk and savings

To our knowledge, there is limited existing work on whether increases or decreases in life expectancy due to health improvements at the population level, rather than individual level, causally affect savings (e.g. due to pandemics or wars), *ceteris paribus*. This would be an interesting question for future research. Jayachandran and Lleras-Muney (2009) is one notable exception: they study the sudden drop in maternal mortality in Sri Lanka between 1946 to 1953. They find that for every additional year of life expectancy, years of education increase by 3%.¹⁷

Additionally, there is tentative *correlational* evidence from surveys during the Cold War that a higher perceived risk of nuclear war was associated with a higher savings rate.

¹⁷See also Hansen (2013) and Hansen and Strulik (2017).

Russett and Slemrod (1993) find this in a 1990 survey data based on $n = 431$ American respondents. Slemrod (1982) as well as Russett, Cowden, et al. (1994) look at the time-series correlation over the course of the Cold War between the U.S. private savings rate and the average of public opinion surveys on nuclear war risk (as well as the correlation with the Bulletin of Atomic Scientists “doomsday clock”) and find positive correlations. Finally, Slemrod (1990) finds a suggestive negative correlation between the national savings rate and the survey average of perceived nuclear war risk in a cross-section of 19 OECD countries in the 1980s.¹⁸

6 Other asset prices

In this section, we consider the possibilities for how transformative AI may affect asset prices other than real interest rates. Our main message is that the sign of the impact on real rates is much clearer the sign of the impact on other asset prices.

6.1 Transformative AI and equity prices

It may be tempting to use for forecasting AI timelines the market capitalization of companies like Alphabet (owner of DeepMind, a leading AI research lab) or that of chip-makers like Nvidia and TSMC. However, extracting AI-related expectations from stock prices is a challenging exercise for four reasons.

Aligned versus unaligned AI. First, and most importantly, AI-related companies will only have the possibility of high profits if transformative AI is aligned. Under *unaligned* AI where humanity is extinguished, the value of stocks along with everything else is converted to zero.

¹⁸There is also work on the relationship between nuclear war risk and *equities*, with particular focus on the Cuban Missile Crisis.

We note first that even a full-blown nuclear war, while it would be doubtless the gravest catastrophe ever known, need not be a true *existential* risk in the sense of wiping out the entirety of the human population. Besides two-sided nuclear exchange possibly being limited, it is still a matter of scientific debate just how much damage such a war and the resulting nuclear winter would cause. Reisner et al. (2018) provides a full-scale analysis; Rodriguez (2019) offers an opinionated summary of the literature.

As a result, the literature reviewed here on nuclear war risk is not necessarily comparable to the truly *existential* risk postulated by unaligned AI. It *may* be closer in nature to the “rare disasters” literature mentioned above.

Finer (2021) studies the cross-section of US equities during the Cuban Missile Crisis. He compares companies with headquarters that are more or less exposed to Cuban missiles, as assessed by secret (at the time) intelligence assessments. He finds that the more exposed stocks fell by more during the crisis.

Burdekin and Siklos (2022) study the Cuban Missile Crisis. They hand collect data on daily equity prices in Canada and Mexico, and together with US data, conclude “markets assigned a very small risk to the crisis leading to the use of nuclear arsenals”.

Profiting versus not. Second, it is not obvious that even in the aligned case that these companies will earn high profits. For instance, OpenAI has committed to a capped profit model, and other AI labs may sign on to a similar ‘Windfall Clause’ promising ex ante to donate profits beyond some threshold (OpenAI 2023; O’Keefe et al. 2020). Beyond corporate altruism, it is plausible that if a private company develops truly transformative AI technology, then the local government may nationalize and expropriate it (or at least attempt to do so) to distribute the benefits more broadly, preventing profits.

Public versus private companies. Third, when considering equity valuations, there is the question of which stock or stocks to consider. Critically, even if one takes a basket of tech companies and averages over them, then this only includes existing public companies. If the market expects transformative AI very soon, but only because it will be developed by a company which is not traded publicly (e.g. OpenAI) then this will not necessarily show up in any index of publicly-traded equities, depending on the affect of such technology on the distribution of firm profits.¹⁹

Higher growth may lower stock prices. Fourth, and quite importantly, it is not obvious whether expectations of transformative AI would raise or lower average equity prices. This is because stock prices reflect the present-discounted value of future profits; and transformative AI may raise those future profits, but – as emphasized throughout this paper – transformative AI would also raise the interest rate used to discount those profits. The net effect on average stock prices is ambiguous, without making more assumptions.

In particular, higher growth causes lower average stock prices if the *intertemporal elasticity of substitution* is greater than one, rather than less than one. This parameter – denoted as σ in section 3 – is subject to significant debate. In particular, while macroeconomics papers often calibrate to $\sigma < 1$, typically asset pricing papers calibrate to $\sigma > 1$. For example, Best et al. (2020) use bunching at mortgage notches to estimate $\sigma = 0.1$. This estimate would imply that stock prices fall strongly, on average, with news about higher future growth (and that real interest rates are very sensitive to changes in growth expectations).

6.2 The price of land and commodities

To the extent that advanced AI is able to substitute for labor but not for land or commodities in production, then the value of land and commodities could skyrocket in the case of aligned AI. However, this does require the auxiliary assumption about the shape of the production function – regarding the non-substitutability with land or commodities – which was not needed previously, and which is highly uncertain.

Additionally, again the value of land and commodities are (of course) directly sensitive

¹⁹Consider the case of a previous transformative technology: the automobile. Suppose you invested in automobile stocks at the turn of the 20th century.

themselves to real interest rates.²⁰ This complicates interpretation of their valuation for the same reason as stock valuations.

Finally, the value of land and commodities are hurt by the prospect of *unaligned* AI. As with equities, the net effect of higher valuation from the prospect of aligned AI versus lower valuation from the prospect of human extinction makes the prices of these assets difficult to use for forecasting AI timelines.

7 Conclusion

In this paper, we do not use any detailed inside knowledge of artificial intelligence technology to provide a forecast of the likely timeline for the development of transformative AI. That is, we do not present an ‘inside view’ on AI timelines (Kahneman 2011).

Instead, we argue that market efficiency provides an ‘outside view’ for forecasting AI timelines. The straightforward economic logic of intertemporal optimization, backed up by simple empirical evidence, shows that the prospect of transformative AI would predict high long-term real interest rates. Such rates can be measured on long-term inflation linked bonds, and used to inform forecasts of transformative AI.

²⁰Relatedly, Giglio, Maggiori, and Stroebel (2015) estimate 999-year risky, nominal discount rates using features of housing market contracts. See also Andersen (2022), Bracke, Pinchbeck, and Wyatt (2018), Feselmeyer, Liu, and Salvo (2016), and Giglio, Kelly, and Stroebel (2021).

References

- Acemoglu, Daron (2009). *Introduction to modern economic growth*. Princeton University Press.
- Acemoglu, Daron and Todd Lensman (2023). *Regulating Transformative Technologies*. Tech. rep. National Bureau of Economic Research.
- Acemoglu, Daron and Pascual Restrepo (2018). “The race between man and machine: Implications of technology for growth, factor shares, and employment”. In: *American economic review* 108.6, pp. 1488–1542.
- Aghion, Philippe, Benjamin F Jones, and Charles I Jones (2018). “Artificial intelligence and economic growth”. In: *The economics of artificial intelligence: An agenda*. University of Chicago Press, pp. 237–282.
- Agrawal, Ajay, John McHale, and Alexander Oettl (2018). “Finding needles in haystacks: Artificial intelligence and recombinant growth”. In: *The economics of artificial intelligence: An agenda*. University of Chicago Press, pp. 149–174.
- Andersen, Carsten (2022). “Do distant rent flows matter? Inferring discount rates from leasehold apartments in Denmark”. In: *Economics Letters* 219, p. 110798.
- Arellano, Cristina, Yan Bai, and Sandra Lizarazo (2017). *Sovereign risk contagion*. Tech. rep. National Bureau of Economic Research.
- Aschenbrenner, Leopold (2020). “Existential risk and growth”. In: *Global Priorities Institute*. <https://globalprioritiesinstitute.org/leopold-aschenbrenner-existential-risk-and-growth>.
- Auclert, Adrien, Hannes MalMBERG, et al. (2021). *Demographics, wealth, and global imbalances in the twenty-first century*. Tech. rep. National Bureau of Economic Research.
- Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2018). *The intertemporal keynesian cross*. Tech. rep. National Bureau of Economic Research.
- Baker, Dean, J Bradford De Long, and Paul R Krugman (2005). “Asset returns and economic growth”. In: *Brookings Papers on Economic Activity* 2005.1, pp. 289–330.
- Baranov, Victoria and Hans-Peter Kohler (2018). “The impact of AIDS treatment on savings and human capital investment in Malawi”. In: *American Economic Journal: Applied Economics* 10.1, pp. 266–306.
- Barro, Robert J (2006). “Rare disasters and asset markets in the twentieth century”. In: *The Quarterly Journal of Economics* 121.3, pp. 823–866.
- Benzell, Seth G and Erik Brynjolfsson (2019). *Digital abundance and scarce genius: Implications for wages, interest rates, and growth*. Tech. rep. National Bureau of Economic Research.
- Benzell, Seth G, Laurence J Kotlikoff, et al. (2015). *Robots are us: Some economics of human replacement*. Tech. rep. National Bureau of Economic Research.
- Beraja, Martin et al. (2023). “AI-tocracy”. In: *The Quarterly Journal of Economics* 138.3, pp. 1349–1402.
- Besiroglu, Tamay, Nicholas Emery-Xu, and Neil Thompson (2022). “Economic impacts of AI-augmented R&D”. In: *arXiv preprint arXiv:2212.08198*.
- Best, Michael Carlos et al. (2020). “Estimating the elasticity of intertemporal substitution using mortgage notches”. In: *The Review of Economic Studies* 87.2, pp. 656–690.

- Bhamra, Harjoat S and Raman Uppal (2014). “Asset prices with heterogeneity in preferences and beliefs”. In: *The Review of Financial Studies* 27.2, pp. 519–580.
- Bialkowski, Jędrzej and EHUD I RONN (2017). “Financial markets facing apocalypse in World War II”. In: *The Journal of European Economic History* 46.2, pp. 9–49.
- Borio, Claudio et al. (2022). “Why so low for so long? A long-term view of real interest rates”. In: *International Journal of Central Banking* 18.3, pp. 47–87.
- Bostrom, Nick (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Bracke, Philippe, Edward W Pinchbeck, and James Wyatt (2018). “The time value of housing: Historical evidence on discount rates”. In: *The Economic Journal* 128.613, pp. 1820–1843.
- Brynjolfsson, Erik and Andrew McAfee (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.
- Burdekin, Richard CK and Pierre L Siklos (2022). “Armageddon and the stock market: US, Canadian and Mexican market responses to the 1962 Cuban missile crisis”. In: *The Quarterly Review of Economics and Finance* 84, pp. 112–127.
- Caldara, Dario and Matteo Iacoviello (2022). “Measuring geopolitical risk”. In: *American Economic Review* 112.4, pp. 1194–1225.
- Center for AI Safety (2023). *Statement on AI Risk*.
- CFM (2023). *Artificial Intelligence and the Economy*. URL: <https://cfmsurvey.org/surveys/artificial-intelligence-and-economy>.
- Christensen, Jens HE, Jose A Lopez, and Glenn D Rudebusch (2010). “Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields”. In: *Journal of Money, Credit and Banking* 42, pp. 143–178.
- Ciancio, Alberto et al. (2020). “Mortality risk information, survival expectations and sexual behaviors”. In.
- Clancy, Matt (2022). *What if we could automate invention? Growth theory in the shadow of artificial general intelligence*. URL: <https://www.newthingsunderthesun.com/pub/2ek4d4s3/release/3>.
- Clark, Gregory (2007). *A farewell to alms: a brief economic history of the world*. Princeton University Press.
- Cochrane, John H (2017). “The new-Keynesian liquidity trap”. In: *Journal of Monetary Economics* 92, pp. 47–63.
- Cotra, Ajeya (2020). *Forecasting TAI with biological anchors*. <https://www.lesswrong.com/posts/KrJfoZzpSDpnrV9va/draft-report-on-ai-timelines>.
- (2022). *Two-year update on my personal AI timelines*. <https://www.lesswrong.com/posts/AfH2oPHCApdKicM4m/two-year-update-on-my-personal-ai-timelines>.
- D’Amico, Stefania, Don H Kim, and Min Wei (2018). “Tips from TIPS: The Informational Content of Treasury Inflation-Protected Securities”. In: *Journal of Financial and Quantitative Analysis* 53.1, pp. 395–436.
- Damodaran, Aswath (2023). “Country Risk: Determinants, Measures and Implications—The 2023 Edition”. In: SSRN.

- Davidson, Tom (2021). “Could advanced AI drive explosive economic growth”. In: *Open Philanthropy*.
- (2023). *What a compute-centric framework says about AI takeoff speeds*. <https://www.lesswrong.com/posts/Gc9FGtdXhK9sCSEYu/what-a-compute-centric-framework-says-about-ai-takeoff>. Accessed on July 7th, 2023.
- DeepMind (2023). *About DeepMind*. Accessed: 2023-07-07. URL: <https://www.deepmind.com/about>.
- Dennis, Richard (2009). “Consumption habits in a new Keynesian business cycle model”. In: *Journal of Money, Credit and Banking* 41.5, pp. 1015–1030.
- Dupraz, Stéphane, Hervé Le Bihan, and Julien Matheron (2022). “Make-up strategies with finite planning horizons but forward-looking asset prices”. In: *Journal of Economic Surveys* 36, pp. 1–30.
- Ely, Jeffrey C. and Balazs Szentes (2023). *Natural Selection of Artificial Intelligence*. Working Paper.
- Epstein, Larry G and Stanley E Zin (1991). “Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis”. In: *Journal of political Economy* 99.2, pp. 263–286.
- Ferguson, Niall (2008). “Earning from history?: Financial markets and the approach of world wars”. In: *Brookings Papers on Economic Activity* 2008.1, pp. 431–477.
- Fesselmeyer, Eric, Haoming Liu, and Alberto Salvo (2016). “How do households discount over centuries? Evidence from Singapore’s private housing market”. In: *Journal of Real Estate Finance* 41.2, pp. 171–190.
- Finer, David Andrew (2021). *No Shock Waves through Wall Street? Market Responses to the Risk of Nuclear War*. The University of Chicago.
- Fleckenstein, Matthias, Francis A Longstaff, and Hanno Lustig (2014). “The TIPS-treasury bond puzzle”. In: *the Journal of Finance* 69.5, pp. 2151–2197.
- Flynn, Joel P, Lawrence DW Schmidt, and Alexis Akira Toda (2023). “Robust comparative statics for the elasticity of intertemporal substitution”. In: *Theoretical Economics* 18.1, pp. 231–265.
- Gabaix, Xavier (2012). “Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance”. In: *The Quarterly journal of economics* 127.2, pp. 645–700.
- Galí, Jordi (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- Gans, Joshua S (2018). *Self-regulating artificial general intelligence*. Tech. rep. National Bureau of Economic Research.
- Giglio, Stefano, Bryan Kelly, and Johannes Stroebe (2021). “Climate finance”. In: *Annual Review of Financial Economics* 13, pp. 15–36.
- Giglio, Stefano, Matteo Maggiori, and Johannes Stroebe (2015). “Very long-run discount rates”. In: *The Quarterly Journal of Economics* 130.1, pp. 1–53.
- Good, IJ (1965). “Franz L. Alt and Morris Rubinfeld, ed.,” *Speculations Concerning the First Ultrainelligent Machine*. In: *Advances in computers*, pp. 31–88.
- Gourio, Francois (2008). “Disasters and recoveries”. In: *American Economic Review* 98.2, pp. 68–73.

- Grace, Katja et al. (2018). “When will AI exceed human performance? Evidence from AI experts”. In: *Journal of Artificial Intelligence Research* 62, pp. 729–754.
- Gürkaynak, Refet S, Brian Sack, and Jonathan H Wright (2007). “The US Treasury yield curve: 1961 to the present”. In: *Journal of monetary Economics* 54.8, pp. 2291–2304.
- Hadfield-Menell, Dylan and Gillian K Hadfield (2019). “Incomplete contracting and AI alignment”. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 417–422.
- Hamilton, James D et al. (2016). “The equilibrium real funds rate: Past, present, and future”. In: *IMF Economic Review* 64, pp. 660–707.
- Hansen, Casper Worm (2013). “Life expectancy and human capital: Evidence from the international epidemiological transition”. In: *Journal of health economics* 32.6, pp. 1142–1152.
- Hansen, Casper Worm and Holger Strulik (2017). “Life expectancy and education: evidence from the cardiovascular revolution”. In: *Journal of Economic Growth* 22, pp. 421–450.
- Hanson, Robin (2000). “Long-term growth as a sequence of exponential modes”. In: *Working manuscript*.
- Heimer, Rawley Z, Kristian Ove R Myrseth, and Raphael S Schoenle (2019). “YOLO: Mortality beliefs and household finance puzzles”. In: *The Journal of Finance* 74.6, pp. 2957–2996.
- Hirshleifer, David, Dat Mai, and Kuntara Pukthuanthong (2023). *War Discourse and the Cross Section of Expected Stock Returns*. Tech. rep. National Bureau of Economic Research.
- Jayachandran, Seema and Adriana Lleras-Muney (2009). “Life expectancy and human capital investments: Evidence from maternal mortality declines”. In: *The Quarterly Journal of Economics* 124.1, pp. 349–397.
- Jones, Chad (2023). “The AI Dilemma: Growth versus Existential Risk”. In.
- Jones, Charles I (1995). “R & D-based models of economic growth”. In: *Journal of political Economy* 103.4, pp. 759–784.
- (2016). “Life and growth”. In: *Journal of political Economy* 124.2, pp. 539–578.
- Kahneman, Daniel (2011). *Thinking, fast and slow*. macmillan.
- Karnofsky, H (2021). “The “most important century” blog post series”. In: *Cold Takes*.
- Karnofsky, Holden (2016). *Some Background on Our Views Regarding Advanced Artificial Intelligence*. <https://www.openphilanthropy.org/research/some-background-on-our-views-regarding-advanced-artificial-intelligence/>.
- Kogan, Leonid et al. (2017). “Technological innovation, resource allocation, and growth”. In: *The Quarterly Journal of Economics* 132.2, pp. 665–712.
- Korinek, Anton (2019). *The rise of artificially intelligent agents*. Tech. rep. working paper, University of Virginia.
- Kurzweil, Ray (2005). “The singularity is near”. In: *Ethics and emerging technologies*. Springer, pp. 393–406.

- Leigh, Andrew, Justin Wolfers, and Eric Zitzewitz (2003). *What do financial markets think of war in Iraq?*
- Longstaff, Francis A et al. (2011). "How sovereign is sovereign credit risk?" In: *American Economic Journal: Macroeconomics* 3.2, pp. 75–103.
- Lunsford, Kurt G and Kenneth D West (2019). "Some evidence on secular drivers of US safe real rates". In: *American Economic Journal: Macroeconomics* 11.4, pp. 113–139.
- Ngo, Richard (2022). "The alignment problem from a deep learning perspective". In: *arXiv preprint arXiv:2209.00626*.
- Nordhaus, William D (2021). "Are we approaching an economic singularity? information technology and the future of economic growth". In: *American Economic Journal: Macroeconomics* 13.1, pp. 299–332.
- O’Keefe, Cullen et al. (2020). "The windfall clause: Distributing the benefits of AI for the common good". In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pp. 327–331.
- OpenAI (2023). *About OpenAI*. Accessed: 2023-07-07. URL: <https://openai.com/about/>.
- Oster, Emily, Ira Shoulson, and E Ray Dorsey (2013). "Optimal expectations and limited medical testing: Evidence from Huntington disease". In: *American Economic Review* 103.2, pp. 804–830.
- Pindyck, Robert S and Neng Wang (2013). "The economic and policy consequences of catastrophes". In: *American Economic Journal: Economic Policy* 5.4, pp. 306–339.
- Reisner, Jon et al. (2018). "Climate impact of a regional nuclear weapons exchange: An improved assessment based on detailed source calculations". In: *Journal of Geophysical Research: Atmospheres* 123.5, pp. 2752–2772.
- Rexer, Jonah M, Ethan B Kapstein, Andres F Rivera, et al. (2022). *Pricing Conflict Risk: Evidence from Sovereign Bonds*. Tech. rep. Empirical Studies of Conflict Project.
- Rietz, Thomas A (1988). "The equity risk premium a solution". In: *Journal of monetary Economics* 22.1, pp. 117–131.
- Rodriguez, Luisa (2019). "How many people would be killed as a direct result of a US-Russia nuclear exchange?" In: *Rethink Priorities*, June 29.
- Russett, Bruce, Jonathan Cowden, et al. (1994). "Did Americans’ expectations of nuclear war reduce their savings?" In: *International Studies Quarterly* 38.4, pp. 587–603.
- Russett, Bruce and Joel Slemrod (1993). "Diminished expectations of nuclear war and increased personal savings: Evidence from individual survey data". In: *The American Economic Review* 83.4, pp. 1022–1033.
- Schmelzing, Paul (2019). "Eight centuries of global real interest rates, RG, and the ‘suprasecular’ decline, 1311–2018". In: *RG, and the ‘Suprasecular’ Decline*, pp. 1311–2018.
- Slemrod, Joel (1982). *Post-war capital accumulation and the threat of nuclear war*. Tech. rep. National Bureau of Economic Research.
- (1990). "Fear of nuclear war and intercountry differences in the rate of saving". In: *Economic inquiry* 28.4, pp. 647–657.
- Stefanski, Radoslaw and Alex Trew (2022). *Selection, Patience, and the Interest Rate*. Adam Smith Business School.

- Stein-Perlman, Zach, Benjamin Weinstein-Raun, and Katja Grace (2022). “expert survey on progress in AI”. In: *AI Impacts*. Available online at: <https://aiimpacts.org/2022-expert-survey-on-progress-in-ai> (accessed December 7, 2022).
- Trammell, Philip (2020). “Existential Risk and Exogenous Growth”. In.
- (2023). “New Products and Long-term Welfare”. In.
- Trammell, Philip and Anton Korinek (2020). “Economic growth under transformative AI”. In: *Global Priorities Institute*. <https://globalprioritiesinstitute.org/philip-trammell-and-anton-korinek-economic-growth-under-transformative-ai>.
- Vinge, Vernor (1993). “Technological singularity”. In: *VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute*, pp. 30–31.
- Weil, Philippe (1989). “The equity premium puzzle and the risk-free rate puzzle”. In: *Journal of monetary economics* 24.3, pp. 401–421.
- Werning, Iván (2015). *Incomplete markets and aggregate demand*. Tech. rep. National Bureau of Economic Research.
- Wolfers, Justin and Eric Zitzewitz (2009). “Using markets to inform policy: The case of the Iraq war”. In: *Economica* 76.302, pp. 225–250.
- Woodford, Michael and Carl E Walsh (2005). “Interest and prices: Foundations of a theory of monetary policy”. In: *Macroeconomic Dynamics* 9.3, pp. 462–468.
- Yudkowsky, Eliezer (2016). “The AI alignment problem: why it is hard, and where to start”. In: *Symbolic Systems Distinguished Speaker* 4.
- Zeira, Joseph (1998). “Workers, machines, and economic growth”. In: *The Quarterly Journal of Economics* 113.4, pp. 1091–1117.