Selection, Patience, and the Interest Rate*

Radoslaw (Radek) Stefanski† Alex Trew‡

September 7, 2023

Abstract

The interest rate has been falling for centuries. A process of natural selection that leads to increasing societal patience is key to explaining this decline. Three observations support this mechanism: patience varies across individuals, is inter-generationally persistent, and is positively related to fertility. A calibrated dynamic, heterogenous-agent model of fertility permits us to isolate the quantitative contribution of this mechanism. We find that selection alone is the key to explaining the decline of the interest rate.

JEL codes: E21; E43; J11; N30; O11.
Keywords: Interest rates; selection; fertility; patience; heterogenous agents.

*We would like to thank the Editor, Joseph Vavra, and four anonymous referees whose comments have greatly improved the paper. We would also like to thank Cornilius Chikwama, John Cochrane, Greg Clark, Richard Dawkins, Kevin Donovan, Byeongju Jeong, Rod McCrorie, Nick Papageorge, Rick van der Ploeg, Sevi Rodriguez Mora, Ctirad Slavik, Tony Smith, Harald Uhlig and Ludo Visschers as well as seminar participants at Cardiff, CERGE-EI, Edinburgh, New Hampshire, Oxford, St Andrews, Barcelona, UEA and conference participants at WEHIA for their comments and suggestions. Stefanski would also like to thank Yale University Department of Economics for its hospitality while writing this paper. The usual disclaimer applies.

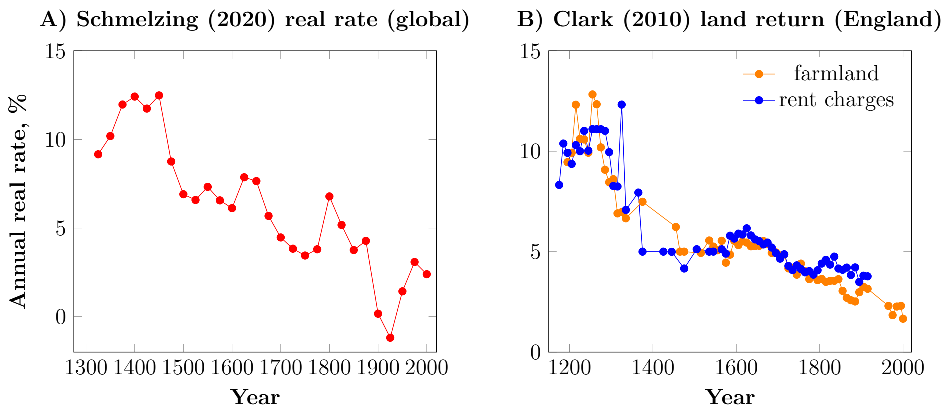
†University of St Andrews, Castlecliffe, The Scores, Fife, KY16 9AR, United Kingdom. Email address: rls7@st-andrews.ac.uk.

‡University of Glasgow and IZA, Glasgow, G12 8QQ, United Kingdom. Email address: alex.trew@glasgow.ac.uk.

1 Introduction

Real interest rates have been falling for at least the last eight centuries (Figure 1). The global real interest rate declined from around 11-12% in the fourteenth century to just 2–3% today (Schmelzing, 2020). The real return on land in England fell from around 10% in the thirteenth century to 1-2% today (Clark, 2010).[1]

Figure 1: Real interest rate, 1175–2000



This large, slow and persistent decline suggests that fundamental economic forces are at play.[2] A standard expression for equilibrium real interest rates comes from the Euler equation in a neoclassical consumption model which, with log utility, is,

$$r_t = g_t - \log \beta. \tag{1}$$

The real interest rate, r_t, is the difference between the growth rate of consumption, g_t, and (the log of) the level of patience, β.[3] Since growth was close to zero up to 1800 and then increased following the onset of the industrial revolution, equation (1) points towards rising levels of patience as the driver of declining real interest rates. We would not normally think of a preference parameter as varying over time at the individual level. We may, however, think of time-varying changes in *societal*

[1] We elaborate on these data in Appendix A. We also report further data across multiple regions and asset classes. Each point to a similar, centuries-long downward trend.

[2] Indeed, Rogoff et al. (2022) finds very limited evidence for structural breaks in the series using the Schmelzing (2020) series.

[3] Of course, a less parsimonious model could incorporate variance in consumption growth, uncertainty of returns, or time-varying risk preferences. As we document in Appendix B, evidence on the long-run changes in each of these additional factors is unable to explain the observed pattern.

levels of patience driven by changing demographics. Blanchard (1985), for example, showed that rising life expectancy can appear as an increase in the effective β where agents with finite-horizons save more. However, as we are able to make explicit, this particular demographic channel does not explain the long decline in rates since life expectancy was flat until the 19th century (Wrigley et al., 1997).

In this paper, we propose a novel demographic channel that can explain the long decline in interest rates. We introduce a model of endogenous fertility, in the spirit of Becker and Barro (1988), where patience levels are heterogenous across agents. Since children are a form of saving, and since patient agents tend to save more than impatient agents, the model implies that patient agents will have more children. Those children in turn inherit part of their parent's higher patience levels, either through genetics, socialization or imitation. Together, these facts imply that the average level of patience in society will increase over time as a result of evolutionary pressures that naturally select the most patient agents. More patient societies are then willing to accept lower rates of interest to induce them to save.

While this mechanism is theoretically plausible, its practical relevance is a quantitative question we are able to address with our model. Specifically, our model's structure allows us to calibrate the historical distribution of patience across individuals using modern, micro-level evidence alone. Using the calibrated model, we demonstrate that the contribution of selection—defined as the difference between our heterogeneous-agent model and a homogenous-agent model—accounts for much of the global interest rate's decline over the past 700 years. While the model can also match the history of population growth, income per capita growth, and life expectancy, it is selection that drives the ability to explain the fall in the interest rate. The fully calibrated model can explains 91.1% of the historical decline in the interest rate. Turning off all channels except the selection mechanism accounts for 86% of the decline in the interest rate in the data. If we retain all channels except selection, the model accounts for only 3.7% of the historical decline. Without a fully-specified and calibrated model, we would be unable to assess the quantitative importance of each of these channels and measure the importance of our selection mechanism.

Understanding the factors driving the real rate of interest over time is crucial for long-term, inter-temporal decisions that are associated with savings and investment choices or future paths of innovation, as well as for the long-run sustainability of public debt. Optimal policies to address very long-term, inter-generational optimization problems, such as those associated with irreversible planetary climate change or social-security funding, often hinge almost entirely on the social discount rate (see Weitzman, 2001; Arrow et al., 2013; and, Millner, 2020). Failing to take account of the declining social discount rate can mean significantly under-valuing a flow of future benefits, relative to adopting a constant social discount rate. Moreover, the further into the future the benefits are realised, the bigger difference accounting for the declining social discount rate makes. By formalizing the process by which societal patience can change, we provide an additional basis for incorporating a declining discount rate in policy.

Related literature First, we contribute to the literature on the role of selection and preferences in economics. Galor and Moav (2002) propose a theory in which there is an evolutionary advantage to traits that are complementary to the escape from the Malthusian trap. Following the demographic transition, higher incomes improve child quality instead of child quantity. Although our model captures a shift in the aggregate time-series correlation between fertility and income, this shift doesn't diminish the central role of selection in explaining the decline in real rates. Closely related is Galor and Özak (2016), which presents a dynamic model in which higher patience leads to better economic outcomes and, consequently, greater reproductive success. Geographical variation in returns on agricultural investment mean that the returns to patience also vary, an implication that Galor and Özak find is consistent with empirical evidence from pre-industrial societies. Our contribution is to understand the relatively more recent dynamics in a way that complements the very long-run comparative analysis in Galor and Özak. Falk et al. (2018) point to some variation in cross-country averages of various preference characteristics, including patience. Sunde et al. (2021) is an example of such variation having significant consequences for comparative development. We focus on the distribution across individuals at a global level, and so explaining differences in cross-country average levels of patience is beyond the scope of this paper. We do show, however, that the interest rate difference implied by the cross-country variation in patience is small compared to the decline in the interest rate observed over time.[4]

We also relate to literature on evolutionarily stable preferences (Becker, 1976, Rogers, 1994, and Robson and Szentes, 2008). Since we calibrate a dynamic model

[4]See Appendix A.3.

that incorporates the shifting distribution of types, we are able to show just how long it can take for such stable preferences to be realized. The closest to our set-up is Hansson and Stuart (1990), in which the population growth of a dynasty is assumed monotonically increasing in per-capita consumption. In our model, population growth is a function of preferences and of the environment. The long-run in our model is the result of a slow process of selection that leads to the most patient dynasty dominating, a result which also echoes the Ramsey (1928) conjecture.[5]

Second, we connect to the economic history literature on the intergenerational transmission of wealth. Clark and Hamilton (2006) shows that families around the beginning of the seventeenth century with more wealth tended to have more surviving children. Records going back to the mid-thirteenth century suggest a similar pattern. For Clark (2007a), variation in reproductive success arises from the Malthusian relationship between wealth and survival. Since innate patience is more deep-rooted than wealth, we view patience as the fundamental driver of differences in both dynastic wealth and household survival.

Third, we relate to the growing literature on family macroeconomics (see Doepke and Tertilt, 2016 for a recent survey). Doepke and Zilibotti (2008) in particular focuses on the intergenerational evolution of patience across and within social classes as it relates to parental decisions to invest in different characteristics of their children. Since this mechanism operates at a shorter time horizon (no more than two to three generations), such a channel is complementary to one driven by selection that operates over much longer periods.

Fourth, our work connects to research on the drivers of the more recent decline in global real rates (see, for example, the chapters in Teulings and Baldwin, eds, 2014). Particularly since Laubach and Williams (2003), a focus in macroeconomics has been to understand the decline in the natural rate since around 1960. Contributions such as Krueger and Ludwig (2007), Del Negro et al. (2018), Carvalho et al. (2016), Mian et al. (2021) and Auclert et al. (2021) seek to understand the role that higher life expectancy, increased risk and declining growth play in driving the decline in the real rate. Over a relatively short horizon of decades, such factors can explain significant fluctuations in the equilibrium real interest rate. Mechanisms driven by evolutionary

[5]Ramsey (op. cit., p. 559) conjectured that, in an economy populated by two groups each with different levels of patience, "...equilibrium would be attained by a division of society into two classes, the thrifty enjoying bliss and the improvident at the subsistence level." See also Becker (1980).

forces, and which naturally take longer for changes to be apparent, are less immediately important. As we show, however, over the centuries-long period that we study, while such other channels do exist there is only one – an increase in societal patience – that is able to explain the slow, continual decline since 1300.

Structure In section 2 we introduce the evidence on the distribution and transmission of preferences. In section 3 we develop a Barro-Becker model of fertility where the key departure is to introduce heterogenous dynasties that differ according to their discount factor. We calibrate the model in Section 4 and compare its quantitative implications to the historical record in Section 5. In that Section we also quantify the contribution of selection to the decline in interest rates under different assumptions about the model. Section 6 discusses the implications of imperfect transmission or mutation of preferences. Finally, section 7 offers some concluding remarks.

2 Heterogeneity, transmission and fertility

Two facts motivate our departure from a standard model of endogenous fertility: first, patience varies across individuals and, second, patience is inter-generationally persistent. In terms of heterogeneity, Andersen et al. (2008) elicit time and risk preferences in a representative sample of Danes, while Alan and Browning (2010) use structural estimation and the PSID. Both studies find similar heterogeneity in discount factors across individuals. More recently, Falk et al. (2018) establish the substantial extent to which preferences varies across the globe at the level of individuals. The intergenerational transmission of preferences, either by genetics, imitation or by socialization, has been identified in studies on Danish and Bangladeshi families (respectively, Brenøe and Epper, 2022 and Chowdhury et al., 2022). Dohmen et al. (2011) has shown that other elements of preferences are also persistent intergenerationally. Of course, such transmission is not perfect in reality and we explore in section 6 the consequences of imperfect transmission for our mechanism. Finally, Cronqvist and Siegel (2015) and Giannelis et al. (2023) support both the variability and the transmissibility of patience. These papers explore determinants of the variation in savings disposition, finding that at least a third of variation can be explained by genetics. Given the persistent nature of the savings disposition, Cronqvist and Siegel argue that the origins of savings behavior can be found in the transmitted variation in time preference or self-control.

A third fact is implicit in standard models of fertility such as Becker and Barro (1988): higher levels of patience will drive higher demand for children since they are a form of saving. This is supported by the evidence in Chowdhury et al. (2022) which finds that the number of children in the household is positively relative to the father's patience and in Bauer and Chytilová (2013), which finds the connection among women in Indian villages. However, to the best of our knowledge the direct connection has not been investigated substantially beyond this. To provide further support, in Appendix A.2 we use the German Socio-Economic Panel (SOEP) data and find a robust, positive relationship between self-reported individual patience levels and the number of children. This holds when we control for a large number of additional variables, including age, net income, gender and household status.

The above evidence together implies that parents that are more patient will have more children than the average, and that the offspring of those highly patient parents will be more patient than the average of their generation. This suggests that over time a greater proportion of the population becomes more patient leading to higher societal levels of patience.

3 A heterogenous-agent Barro Becker model

In section 3.1 we present the basic set-up in which dynasties differ in their levels of patience and where productivity, life expectancy and child costs can potentially vary over time. We discuss some of the assumptions before describing the solution of the model and develop a simple expression that captures the mechanisms driving changes in the interest rate. In section 3.2 we then introduce a generalization of this basic set-up to incorporate the implications of potentially imperfect altruism where life expectancy plays a role as in Blanchard (1985). We present calibration of the model in section 4 before discussing quantitative results in section 5. As we will show in that section, the key mechanism driving the decline in interest rates over our period is the selection that leads to greater societal patience over time. Time-varying productivity, child costs and life expectancy are important in matching the path of population and income per capita, but are not important in explaining the decline in interest rates.[6]

[6] Appendix G presents a version of the model with a fixed factor (such as land in a Malthusian setting), in which we also find the role of selection to be key.

3.1 Model set-up

Consider an economy with aggregate population N_t at time t that consists of a finite number of dynasties, indexed by $i = 1, \ldots, I$. Each dynasty i consists of N_t^i identical households. Each dynasty differs in its discount factor, β^i.[7] Without loss of generality, the sequence $\{\beta^i\}_{i=1}^I$ is strictly increasing in i, so dynasty I has the highest discount factor, β^I. Each period every household is endowed with a unit of labor that it inelastically provides in exchange for a wage, w_t, as well as a stock of reproducible capital, k_t^i, that it either retained (if it is a surviving adult) or inherited from its parent (if it is a child) and that it rents out in exchange for a rental rate, r_t. Capital depreciates at rate $\delta > 0$ and can be accumulated via investment of retained output x_t. A single time period in our model is 25 calendar years (a 'generation').

Each household of type i solves the following utility maximization problem in each period t:

$$U_t^i(k_t^i) = \max_{c_t^i, n_{c,t}^i, x_t^i} \alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i) + \beta^i U_{t+1}^i(k_{t+1}^i) \tag{2}$$

$$\text{s.t.} \qquad c_t^i + q_t n_{c,t}^i + x_t^i \leq w_t + r_t k_t^i, \qquad n_{t+1}^i = \pi_t + n_{c,t}^i, \qquad k_{t+1}^i = \frac{k_t^i(1-\delta) + x_t^i}{n_{t+1}^i}.$$

As in Becker and Barro (1988, 1989), households derive utility from their own consumption, c_t^i, from the size of the household at the beginning of the next period, n_{t+1}^i, and from the next period average continuation utility, $U_{t+1}^i(k_{t+1}^i)$. Parents face a trade-off when it comes to children: They enjoy bigger families, but at the same time they derive welfare from children who are wealthier. Given their income from supplying labor, w_t, and renting out capital, $r_t k_t$, households choose the quantity of their consumption, c_t^i, the number of children to have, $n_{c,t}^i$, and the quantity of capital to accumulate, x_t^i. We allow child costs to vary according to an exogenous price $q_t \equiv D_t^{\frac{1}{1-\nu}} a_t$, where D_t is time-varying firm productivity. In this expression, the productivity term captures the rising costs of raising children to adulthood, and a_t is an exogenous shock to child costs which we calibrate to match the path of population

[7]Since households within a dynasty are identical, and since we obtain solutions to the model in terms of dynasty-aggregates, we omit a household index. As we explain below, household-level quantities are lower-case, so, e.g., c_t^i is the time t consumption of an individual household in dynasty i; dynasty-aggregates are upper case, so C_t^i is the sum of consumption by households in dynasty i at time t. Quantities without the index i are economy-wide aggregates, such as C_t as the sum of dynasty consumptions at time t.

growth.[8] Finally, we allow the probability of survival of existing households, π_t, to vary exogenously over time to match the data on life expectancy. The survival probability of children is set to 1 (this can be readily generalized). Together, these assumptions imply that the expected number of people in a household at the beginning of the next period will be $n^i_{t+1} = \pi_t + n^i_{c,t}$. We assume that parents care about their children equally and endow them each with the same share of accumulated capital (which may be negative[9]). Thus, parents face a quantity-quality tradeoff with respect to the number of children à la Becker and Barro (1988, 1989). Finally, we also assume that the child of an adult in dynasty i perfectly inherits the discount factor β^i (we discuss relaxing this assumption in section 6). This transmission can be thought of as coming from genetics, imitation or socialization and, given the lack of clear identification of mechanisms in the empirical literature described above, is left as a reduced form assumption.

Discussion Three aspects of the above model merit further discussion. First, as in Becker and Barro (1988, 1989), we assume in (2) a particular form of altruism. Part of this altruism is that parents enjoy larger families; another part is that parents care about the average utility of children in future periods. This particular choice of utility function follows Tamura (1996), Lucas (2002) and Bar and Leukhina (2010). In those models, parents live for one period only and do not co-exist with their children. In our model, however, parents survive into the future with some probability, alongside their children. As such, our discounting parameter not only captures the present value of the utility of children realized at different points in the future but also captures the future utility of the parent. The implicit altruistic assumption we make is that, in terms of time discounting, the future utility of children is treated the same as the future utility of parents. That is, in each period a parent applies the same *time discount factor* to their own and their children's future utility. The altruism implicit in the equal treatment of own and descendant utility is itself unrelated to β. Given

[8]Since we have assumed that the cost of raising children is paid in terms of the final good, we need to assume that the goods cost of children grows in proportion to income, much like in Lucas (2002) and Bar and Leukhina (2010), to ensure the existence of a balanced growth path. As in those studies, we take q_t to be a reduced form way to capture the cost of an additional child. That cost can capture the parental opportunity cost from spending time with children, the presence of mandatory education, changes in the relative cost of services, laws that prevent children working, and so on. An equivalent approach would be to incorporate the time cost of raising children.

[9]Schoonbroodt and Tertilt (2014) shows how the possibility to endow negative bequests can matter for efficiency in models of endogenous fertility.

this assumption, and the fact that parents live for multiple periods, the β captures the level of time preference rather than the extent of altruism. We refer to this preference structure as 'perfect altruism' since parents take account of all future periods, i.e., they care about their descendants even after they die. In order to consider the sensitivity of our results to this assumption, we relax it in section 3.2 and allow for a form of 'imperfect' altruism in which parents care about their children, but only while they themselves are alive. This introduces a different effective weight on a parent's own future utility compared to the entire future of their descendant's utility.[10] This also permits us to consider a mechanism, such as in Blanchard (1985), where finite horizons can matter in savings decisions; this may be important in explaining a decline in interest rates that arises out of increasing life expectancy observed after 1850. In the quantitative analysis, the contribution of selection to explaining the decline in the interest rate remains similar in both cases.

Second, an assumption implicit in models of endogenous fertility is that parents can always choose the number of children. Whether deliberate birth control existed in the period before the demographic transition is a topic of recent debate.[11] For Clark and Hamilton (2006) and Clark (2007a), differences in *survival* rates across groups, rather than fertility rates, led to changes in the composition of the population. Either interpretation is consistent with our model. In the current set-up, we make the assumption that all children survive to adulthood and so the endogenous choice of the number of 'children' in our model is really a choice of the number of adults in the dynasty in the next period. We can thus otherwise think of this as a choice to allocate the resources in raising a child to adulthood. The mechanism in our model holds whether the variation arises from endogenous birth rates, or whether it arises via endogenous survival to adulthood through different parental investment decisions.[12]

Third, children are a normal good in our model. As such, while more patience leads directly to greater fertility, more patient households also save more in terms of physical

[10]Again, this is distinct from the time preference parameter, as the time-zero household problem shows.

[11]See Cinnirella et al. (2017), Clark and Cummins (2019) and Cinnirella et al. (2019). Clark et al. (2020) found that parity dependent fertility control did not exist within marriage; de la Croix et al. (2019) incorporate additional margins, such as the propensity to marry, the child mortality rate and the rate of childlessness within marriage, and find that the net reproduction rate can vary considerably across social groups, suggesting some fertility control.

[12]We could separately consider these in a model where child mortality existed, and where the 'fertility' choice is partly the number of births and partly an investment in raising children to adulthood.

capital, accumulating greater wealth. That higher wealth then leads further to greater fertility, an indirect consequence of higher patience. That is, there are three channels at work, that between patience and wealth/income, that between wealth/income and fertility, and that between patience and fertility. Each of those channels operate in a way supported by the data. The evidence discussed in Section 2 points to a positive and direct relationship between patience and fertility. The connection from patience to income/wealth is established in Epper et al. (2020) and Sunde et al. (2021). The causal relationship between income, wealth and fertility has also been found to be positive at an individual level. Black et al. (2013) asks 'Are children normal?'. Their analysis using U.S. data uses the oil price shock of the 1970s to isolate exogenous variation in incomes, pointing to a causal effect on fertility of higher men's income. Lovenheim and Mumford (2013) use exogenous shocks to housing wealth to show using individual-level U.S. data a relationship between wealth and fertility. Kolk (2022) finds a positive relationship between lifetime accumulated income and fertility for men in Sweden. Evidence from household-level data using exogenous wealth shocks supports a positive causal relationship between income and fertility. Kearney and Wilson (2018) use the fracking boom to isolate exogenous variation in income, finding a positive relationship between (male) earnings and household fertility. Bennett et al. (2021) use the discovery of oil and gas in the North Sea as an unexpected shock, find a positive relationship between income and the number of children. Cesarini et al. (2023) uses lottery shocks to wealth, finding again a positive connection with fertility.

In sum, along with the positive direct relationship between patience and fertility, the evidence on a causal connection from patience to income/wealth and from income/wealth to fertility implicitly supports the notion that there is an positive and causal indirect connection between patience and fertility via income/wealth.

Time-zero households and dynastic planners Since households care about the outcomes of their future children, we can simplify the above problem and, by iterative substitution, re-write the individual household problem in the framework of a time zero household of each type:

$$\max_{\{c_t^i, n_{c,t}^i, x_t^i\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} (\beta^i)^t \left(\alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i)\right) \tag{3}$$

s.t.

$$c_t^i + q_t n_{c,t}^i + x_t^i \leq w_t + r_t k_t^i, \quad n_{t+1}^i = \pi_t + n_{c,t}^i, \quad k_{t+1}^i = \frac{(1-\delta)k_t^i + x_t^i}{n_{t+1}^i}.$$

The above reflects the choice of an individual time zero adult household. Since there are N_0^i identical households in each dynasty i at time zero, we can also re-write the time zero household problem from the perspective of a single dynastic planner for each type. At time t, there are N_t^i identical members of the dynasty of type i. Next period, the dynasty will be comprised of the number of children produced by each household, $n_{c,t}^i$ (all of whom are assumed to survive and become household heads in their own turn), and the expected number of surviving adults. The number of households in dynasty i at time $t+1$ will thus be given by $N_{t+1}^i = (\pi_t + n_{c,t}^i)N_t^i = n_{t+1}^i N_t^i$. Dynasty-aggregate values are $C_t^i \equiv c_t^i N_t^i$, $N_{c,t}^i \equiv n_{c,t}^i N_t^i$, $K_t^i \equiv k_t^i N_t^i$, $X_t^i \equiv x_t^i N_t^i$ and so we can re-write the time-zero household problem for the dynastic planner of each type as:

$$\max_{\{C_t^i, N_{c,t}^i, X_t^i\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} (\beta^i)^t \left(\alpha \log(C_t^i) + (1-\alpha-\beta^i)\log(N_{t+1}^i)\right) \tag{4}$$

s.t.

$$C_t^i + q_t N_{c,t}^i + X_t^i \leq w_t N_t^i + r_t K_t^i$$

$$N_{t+1}^i = \pi_t N_t^i + N_{c,t}^i$$

$$K_{t+1}^i = (1-\delta)K_t^i + X_t^i.$$

Notice that the discount factor appears both as the term used for discounting the future, but also as a preference weight for children. This reflects the fact that current children are both an investment and a consumption good in this model. In particular, the more patient agents place less weight on current children as they are partially viewed as current consumption goods rather than entirely investment goods for the future. Following Lucas (2002), to ensure strict concavity of the objective we need to assume that $1 - \alpha - \beta^i > 0$.

Firms The representative firm hires workers (N_t) and capital (K_t) to produce final output (Y_t). The profit maximization problem of the firm is given by:

$$\max_{\{K_t, N_t\}} Y_t - w_t N_t - r_t K_t, \tag{5}$$

where $Y_t = D_t K_t^{\nu} N_t^{1-\nu}$ and where $0 < \nu < 1$ is the output elasticity of capital. D_t is the exogenous and potentially time-varying level of technology.

Market clearing Economy-wide aggregate quantities are denoted C_t, etc. The market clearing conditions are given by:

$$\sum_{i=1}^{I} C_t^i = C_t \ , \ \sum_{i=1}^{I} N_t^i = N_t \ , \ \sum_{i=1}^{I} N_{c,t}^i = N_{c,t} \ , \ \sum_{i=1}^{I} K_t^i = K_t,$$

$$C_t + q_t N_{c,t} + X_t = D_t K_t^{\nu} N_t^{1-\nu}. \tag{6}$$

Notice that capital is now produced from output and that producing a child costs and exogenous q_t units of output.

Competitive equilibrium A competitive equilibrium, given a series of child prices $\{q_t\}_{t=0}^{\infty}$ and technology $\{D_t\}_{t=0}^{\infty}$, along with parameter values and initial conditions $\{N_0^1, \ldots, N_0^I, K_0^1, \ldots K_0^I\}$, consists of allocations $\{C_t^i, N_{c,t}^i, N_{t+1}^i, K_{t+1}^i, X_t^i\}_{t=0}^{\infty}$ for each dynasty $i = 1, \ldots, I$ and prices $\{w_t, r_t\}_{t=0}^{\infty}$ such that firms' and dynasties' maximization problems are solved, and all markets clear.

3.1.1 Model solution

To solve the model, we start by deriving the first order conditions of the dynastic planner and the firms. For given parameter values, initial population and capital distributions, the competitive equilibrium of the problem, for each dynasty $i = 1, \ldots, I$, is characterized by consumer first-order conditions with respect to choice of children and consumption as:

$$\frac{(1 - \alpha - \beta^i)}{N_{t+1}^i} + (\pi_{t+1} q_{t+1} + w_{t+1}) \frac{\alpha \beta^i}{C_{t+1}^i} = q_t \frac{\alpha}{C_t^i}, \tag{7}$$

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i (\pi_{t+1} (1 - \delta + r_{t+1}), \tag{8}$$

with consumer budget constraints for each dynasty i:

$$C_t^i + q_t N_{t+1}^i + K_{t+1}^i = (w_t + \pi_t q_t) N_t^i + (1 - \delta + r_t) K_t^i. \tag{9}$$

The firm first-order conditions are:

$$w_t = (1-\nu)D_t K_t^{\nu} N_t^{-\nu} \text{ and } r_t = \nu D_t K_t^{\nu-1} N_t^{1-\nu}. \tag{10}$$

Finally, there are two transversality conditions per dynasty:

$$\lim_{t\to\infty} (\beta^i)^t u'(C_t^i) K_{t+1}^i = 0, \quad \lim_{t\to\infty} (\beta^i)^t u'(C_t^i) N_{t+1}^i = 0, \tag{11}$$

where, $u(C_t^i) = \log(C_t^i)$ is the period utility of consumption.

From the above, we obtain the following two Euler equations that describe the evolution of dynasty consumption and population:

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i R_{t+1}, t \geq 0, \qquad \frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, t \geq 1, \tag{12}$$

where $R_{t+1} \equiv (1 - \delta + r_{t+1})$ is the gross real interest rate on capital and $\tilde{R}_{t+1} \equiv R_{t+1} \frac{q_{t-1}R_t - w_t - q_t\pi_t}{q_t R_{t+1} - w_{t+1} - q_{t+1}\pi_{t+1}}$ is the shadow gross real interest rate on dynasty population.[13]

Since the interest rates are common across dynasties, we can obtain expressions relating the *relative* evolution of total consumption and population for any two dynasties. Using repeated substitution, together with market clearing conditions, we can obtain the shares of consumption and population of each dynasty relative to economy-wide aggregate consumption and population, respectively, as a function of the initial distribution of dynasty-specific consumption and population:

$$\frac{C_t^i}{C_t} = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^{I} (\beta^j)^t C_0^j}, \text{ and, } \quad \frac{N_{t+1}^i}{N_{t+1}} = \frac{(\beta^i)^t N_1^i}{\sum_{j=1}^{I} (\beta^j)^t N_1^j}, \tag{13}$$

for $t \geq 0$. Note that given the initial distributions, the evolution of a particular dynasty's population and consumption shares depends only on that dynasty's patience relative to the patience of other dynasties. In particular, recalling that dynasty I is most patient, the above expressions imply that as $t \to \infty$, so $\frac{N_{t+1}^I}{N_{t+1}} \to 1$ and $\frac{C_{t+1}^I}{C_{t+1}} \to 1$ whilst, for all $i < I$, $\frac{N_{t+1}^i}{N_{t+1}} \to 0$ and $\frac{C_{t+1}^i}{C_{t+1}} \to 0$. This means that the total consumption and population of the most patient dynasty will dominate the economy over time (consistent with the Ramsey (1928) conjecture). As $t \to \infty$ the model collapses to

[13] These two interest rates differ since children are both a consumption and an investment good, whereas capital is only an investment good.

a standard homogenous agent model with discount factor β^I and a standard Barro-Becker steady state. Consequently, if we derive the steady state, the model can be solved with a variation of the reverse-shooting algorithm.[14]

3.1.2 Distribution of β^i

We assume that the distribution of generational discount factors in the population follows a scaled beta distribution defined on $(0, \bar{\beta})$, with cumulative distribution function, $F(\cdot)$ given by:

$$F(\beta; t) = \frac{B\left(\beta/\bar{\beta}, \gamma_t, \delta_t\right)}{B(\gamma_t, \delta_t)}. \tag{14}$$

In the above, $B(\gamma_t, \delta_t)$ and $B\left(\beta/\bar{\beta}, \gamma_t, \delta_t\right)$ are the complete and incomplete beta functions, respectively, and $\gamma_t, \delta_t > 1$ are two potentially time-varying shape parameters that determine the mean and dispersion of the distribution.

There are a number of reasons for choosing this distribution. First, this distribution can be defined on any positive sub-interval, and thus is useful for considering discount factors which are naturally bounded. Second, it is a flexible distribution often used to mimic other distributions, both skewed and centered, given appropriate bounds. Third, a distributional assumption is required for the purposes of calibration, as will become clear below. Finally, the beta distribution is also intimately linked to the evolution of the population distribution implied by our model, as Theorem 1 shows.

Theorem 1. *If $I \to \infty$ and dynastic discount factors within the population are distributed according to a scaled-beta distribution on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}}$ and $\delta_{\bar{t}}$ for some period $\bar{t}$, then dynastic discount factors will also be distributed according to a scaled beta distribution in period $\bar{t}+1$ on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}+1} = \gamma_{\bar{t}} + 1$ and $\delta_{\bar{t}+1} = \delta_{\bar{t}}$.*

Proof. See Appendix E.2. □

Theorem 1 establishes that if discount factors obey a scaled-beta distribution in any one period then they will follow a scaled-beta distribution in all other periods.

[14] Notice that since the model exhibits trend growth in both output per worker and population to derive the steady state it first needs to be de-trended. Appendix C describes derivations in the full model (incorporating the extension introduced below in section 3.2) and its solution in de-trended variables, and documents the aggregation and solution algorithm.

Since the theorem also pins down the evolution of shape parameters over time, the choice of year in which to calibrate is irrelevant. An immediate implication of the Theorem is that we can derive expressions for the mean and variance of generational discount factors at any time t:

$$\mathrm{E}_t(\beta) = \bar{\beta}\frac{\gamma_0 + t}{\gamma_0 + t + \delta_0} \text{ and } \mathrm{var}_t(\beta) = \bar{\beta}^2\frac{(\gamma_0 + t)\delta_0}{((\gamma_0 + t) + \delta_0)^2(\gamma_0 + t + \delta_0 + 1)} \tag{15}$$

Selection A further advantage of this distributional assumption is that we can illustrate the role of selection in driving the evolution of the interest rate with the following simple approximation:[15]

$$R_{t+1} \approx \frac{g_{Nt+1}g_{Dt+1}^{\frac{1}{1-\nu}}}{E_t(\beta)}. \tag{16}$$

This approximation is an analogue to equation (1). It illustrates the three key forces driving changes in the interest rate: 1) the time-varying growth rate of consumption as captured by population growth ($g_{Nt+1} \equiv N_{t+1}/N_t$) and 2) productivity growth ($g_{Dt+1} \equiv D_{t+1}/D_t$); and, 3) selection-driven changing societal patience as captured by the expected value of the discount factor, $E_t(\beta)$. Without our selection mechanism, in a model with homogenous agents, $E_t(\beta)$ is simply a constant and the interest rate is driven by growth alone. Population and productivity growth *increase* over our period, especially after the industrial revolution, and thus cannot help explain falling interest rates. However, with our selection mechanism, $E_t(\beta)$ increases over time. We show that changes in the mean value of beta driven by selection are enough to explain a large part of the observed decline of interest rates, irrespective of the other mechanisms.

3.2 Imperfect altruism and finite-horizons

Recent studies[16] after Laubach and Williams (2003) have found a key role for rising life expectancy in explaining the decline in the natural rate since the 1960s. In such explanations, higher life expectancy causes agents to save more, thus depressing the real interest rate. In a model with perfect altruism, however, since there is

[15] The derivation of which is in Appendix E.3
[16] See under related literature in section 1.

no distinction between utility derived by a future self and a future child, the life expectancy of an adult does not affect their savings decisions. In order to incorporate such a channel, we thus need to depart from the perfect altruism of section 3.1. To do so, we introduce a form of imperfect altruism in which parents care about their children's future average utility, but only while they themselves are alive. Changes in life expectancy then influence the interest rate since an individual parent's own finite horizon enters into their decision-making (see Blanchard, 1985).

Suppose that each household i's problem at time t is now:

$$U_t^i(k_t^i) = \max_{c_t^i, n_{c,t}^i, x_t^i} \alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i) + \beta^i(\pi_t(1-\omega)+\omega)U_{t+1}^i(k_{t+1}^i). \quad (17)$$

$$\text{s.t.} \qquad c_t^i + q_t n_{c,t}^i + x_t^i \leq w_t + r_t k_t^i, \qquad n_{t+1}^i = \pi_t + n_{c,t}^i, \qquad k_{t+1}^i = \frac{k_t^i(1-\delta)+x_t^i}{n_{t+1}^i}.$$

In the above $\omega \in [0,1]$ is a parameter that captures a particular form of imperfect altruism that parents may have for their children. If $\omega = 1$, we return to the baseline preferences of section 3.1 where parents are perfectly altruistic towards their children. Setting $0 \leq \omega < 1$, introduces a form of selfishness; with $\omega = 0$, parents care about their own future utility and the utility of their descendants but only so long as they themselves are alive. In the extreme case that survival probability goes to zero, agents would care only about present consumption and the number of children that they have. Higher life expectancy (captured by an increase in the survival probability, π_t) extends the expected horizon over which parents consider future utility, meaning that parents care more for the future and save more, potentially depressing interest rates. The solution to the above problem and its aggregation follows in a similar fashion to the baseline model for a time zero household and the dynasty planner (see Appendix C for details).

With imperfect altruism, and a similar assumption on the distribution of discount factors, we can obtain a generalization of (16) (see Appendix E.3 for the derivation):

$$R_{t+1} \approx \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{E_t(\beta)(\omega + (1-\omega)\pi_{t+1})}. \quad (18)$$

Relative to equation (16), equation (18) now incorporates the fourth force that can drive changes in the interest rate in our model: life expectancy as captured by changing survival probabilities (π_t) in the presence of imperfect altruism ($\omega < 1$). Just like population and income per capita, life expectancy begins to increase during the demographic transition. As we will see, changes to life expectancy can thus help explain some of the decline in the interest rate after around 1850 but not earlier.

4 Calibration

The calibration aims to replicate the path of world population, income per capita and life expectancy over the period 1300 to 2000. Additionally, it aims to fit the modern variance of patience types using experimental data from around the year 2000. Note that the decline in the interest rate is not itself targeted in our calibration. Model parameters and their calibrated values are summarized in Table 1. Below we outline the calibration procedure, leaving technical details to Appendix D. Calibration of parameters depends on the value we set for ω, that is, whether we have a model with perfect altruism or whether we introduce a form of imperfect altruism as in section 3.2. In the absence of a clear way to calibrate the extent of imperfect altruism, we present the two extreme versions: with perfect altruism ($\omega = 1$) and with imperfect altruism where parents care about their children only while they are themselves alive ($\omega = 0$).

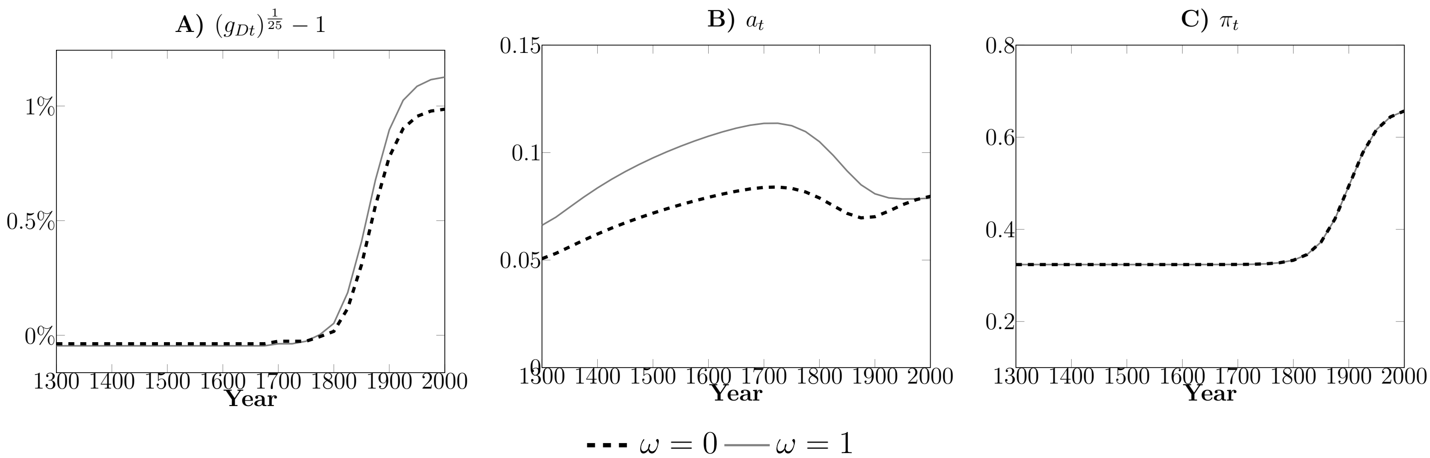
One period in the model is 25 calendar years and period zero in the model corresponds to the year 1300 in the data. The initial level of population is set to be $N_0 = 0.37$ corresponding to a world population of 0.37 billion in 1300 (The Maddison Project, 2013). The capital elasticity of the production function is set to 0.33 to match the capital share in Gollin (2002). We assume that all children survive into adulthood (25 years) and that capital depreciates 10% annually, so that $\delta = 1 - (1 - 0.1)^{25} = 0.928$.

The path for net annualized productivity growth, $(g_{Dt})^{\frac{1}{25}} - 1$, the child cost parameter, a_t, and the survival probability, π_t, are reported in Figure 2. The productivity growth sequence g_{Dt} is set to match the compound annual growth rates of world GDP per capita over three periods that exhibit markedly different growth patterns (1275–1700, 1700–1875, and 1875–2000), using data in The Maddison Project (2013), with a smoothed transition imposed in the model. Note that since in the model the savings rate endogenously grows over time (as average patience increases), the low growth in output per capita in The Maddison Project (2013) up to 1775 is captured

Table 1: Calibrated parameters

Parameter(s)	Value		Target/Description/Source
	Perfect altruism $\omega = 1$	Imperfect altruism $\omega = 0$	
D_0	1	1	Normalization
N_0	0.37	0.37	World population, 1300, The Maddison Project (2013)
δ	0.928	0.928	10% annual depreciation (see text)
g_{Dss}	1.005^{25}	1.005^{25}	Predicted long run productivity growth rates, Crafts and Mills (2017)
$\tilde{k}_0$	0.001	0.001	On saddle path (see text)
a_{ss}	0.345	0.345	Predicted future population growth rates from Herrington (2021) of -1%
ν	0.330	0.330	Capital share, Gollin (2002)
I	10000	10000	Number of Types
$\{\beta^i\}_{i=1}^I$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^I$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^I$	Subdivide domain into grid
π_{ss}	0.667	0.667	Long-run adult life expectancy of 75
α	0.345	0.345	Global consumption share of 0.75 (see App. D)
$\bar{\beta}$	0.512	0.769	Maximum (generational) discount factor, $\frac{1-\alpha}{\pi_{ss}(1-\omega)+\omega}$
$\{\gamma_{28}, \delta_{28}\}$	$\{32.6, 58.8\}$	$\{33.6, 60.4\}$	Standard deviation of discount factors (Andersen et al., 2008; Falk et al., 2018) and long run rate of return
$\{g_{Dt+1}\}_{t=-1}^{28}$	Figure 2	Figure 2	World output per capita growth rates (smoothed), The Maddison Project (2013)
$\{a_t\}_{t=-1}^{28}$	Figure 2	Figure 2	World population growth rates, The Maddison Project (2013)
$\{\pi_t\}_{t=0}^{28}$	Figure 2	Figure 2	English life expectancy, smoothed
$\left\{\frac{N_0^i}{N_0}^i\right\}_{i=1}^I$	See text	See App. D	Andersen et al. (2008) and Falk et al. (2018)
$\left\{\frac{\tilde{k}_0^i}{\tilde{k}_0}\right\}_{i=1}^I$	See text	See App. D	Consistency (see text)

Figure 2: Time-varying parameters



Note: We report the annualized growth rate of the technology level, D_t, calibrated so the model matches the average growth of output per capita in the data. The parameter a_t, along with D_t, determines the cost of children via $q_t \equiv a_t D_t^{\frac{1}{1-\nu}}$. The adult survival probability, π_t, is chosen to match the historical trend in life expectancy. Each of the population, income and life expectancy series are smoothed to capture transitions between different phases, as detailed in the text and elaborated in Appendix D. We report model outputs both with perfect altruism ($\omega = 1$) and with imperfect altruism ($\omega = 0$).

by a very slightly decreasing level of TFP. The cost of children $q_t \equiv D_t^{\frac{1}{1-\nu}} a_t$ is set by choice of a_t to match compound annual population growth rates over three periods that also exhibit markedly different growth patterns (1275–1775, 1775–1875, and 1875–2000) by exogenously varying a_t, again in a smoothed fashion. The path for the survival probability π_t is set to match a smoothed series based on the historical life expectancy. Roser et al. (2013) compiles life expectancy over the period 1543–2020 based on a number of sources, including Wrigley et al. (1997).

Dynasties We assign a discount factor to each dynasty $i \in I$. Recall that we order dynasties such that the sequence $\{\beta^i\}_{i=1}^{I}$ is strictly increasing in i. Given the restriction that $1 - \alpha - \beta^i > 0$, each discount factor is bounded by $0 < \beta^i < \bar{\beta}$, where $\bar{\beta} \equiv 1 - \alpha$. We divide this interval $(0, \bar{\beta})$ into I equally-sized sub-intervals and locate each type's patience level at the central point of every sub-interval, so that, for each i, $\beta^i = \bar{\beta}\frac{(2i-1)}{2I}$. We specify the number of dynasties to be $I = 10,000$.[17] We discuss the

[17]This is largely a computational choice which makes little difference to our results for a large enough number of dynasties. If too few dynasties are chosen, the resulting transitions are non-smooth. Since we view our model as largely approximating a near-continuous distribution of types

calibration of α and the upper bound $\bar{\beta}$ in Appendix D. In practice, the population at this upper bound in our simulations is negligible until the very distant future.

Initial capital and population distribution We present here the calibration of the capital and population distribution only for the model with $\omega = 1$ simply for expositional clarity. We leave the calibration of the general model (which is procedurally similar) to Appendix D. Results in section 5 are presented for both the $\omega = 0$ and $\omega = 1$ cases.

We first need the initial distribution of capital, $\{K_0^i\}_{i=1}^I$, and population, $\{N_0^i\}_{i=1}^I$ across dynasties. This data is not readily available for the year 1300. To obtain the initial distribution of capital we will assume that the model is in equilibrium *prior* to our initial period. We then use the structure of our model to obtain the initial population distribution from modern data.

More specifically, the initial distribution of capital across dynasties determines the population distribution of those dynasties in *subsequent* periods. To obtain a capital distribution in period $t = 0$ we assume that the growth of each dynasty's population is consistent with solutions of the model in the period prior to the initial period.[18] In practice, this means assuming that the second equation in (12) also holds for $t = 0$. This gives I equations that can be used to pin down an initial capital distribution for I dynasties. Notice that this assumption also implies a direct relationship between period zero and period 1 population distributions which will be useful below:

$$\frac{N_1^i}{N_1^j} = \frac{\beta^i}{\beta^j}\frac{N_0^i}{N_0^j}. \tag{19}$$

Since we do not have data on the population distribution of patience in the year 1300 ($t = 0$ in the model), we choose our period-zero distribution of types so that the model replicates evidence on the distribution of types in the year 2000 ($t = 28$ in the model). To do this, note that the second expression in (13) gives the population share of each dynasty over time as a function of the $t = 1$ population share and each dynasty's level of patience. Using this and (19), we obtain the $t = 0$ population share

in the data, we select a large number of types in the calibration.

[18]That is, we assume that outcomes in the period before $t = 0$ are on the equilibrium saddlepath just as much as they are in periods from $t = 0$ on. This simply means that we are ignoring potential shocks, such as wars, famines or pandemics, that may cause population growth from $t = 0$ to deviate from the saddlepath that continues from period $t = 1$.

of each dynasty i relative to dynasty I,

$$\frac{N_0^i}{N_0^I} = \frac{N_t^i}{N_t^I}\left(\frac{\beta^i}{\beta^I}\right)^t. \tag{20}$$

With evidence on the distribution of patience at some later date t, we can thus calibrate the initial distribution of the population across levels of patience.

Patience distribution One problem with using modern data is that it will capture only a censored portion of the full initial distribution of preference types – even the most populous dynasties of the year 1300 could be completely indiscernible in data for the year 2000.[19] To address this issue, we use modern data to parameterize the scaled beta distribution introduced in section 3.1.2. This holds for the $\omega = 1$ version; the $\omega = 0$ version is in the appendix. Note that from (15) the two shape parameters, γ_t and δ_t, of the distribution of generational discount factors may be obtained if we observe the mean and variance of that distribution. Appendix D details the calibration of the variance using the experimental evidence from representative individuals in Denmark (Andersen et al., 2008) and the individual-level data in the Global Preference Survey introduced in Falk et al. (2018). We choose the mean of the patience distribution by matching the model implied rate of return in 2100 to 6.4%. This value is derived from the long run equity returns calculated in Appendix A.5. Appendix D.2 demonstrates robustness to varying γ_t and δ_t around the calibrated values.

5 Quantitative results

Figure 3 reports various key simulation results with $\omega = 0$ (results for $\omega = 1$ are in Appendix Figure A3). Panel A) shows a monotonic increase in the level of societal patience (the mean level of patience in the economy) over time. That increase reflects the changes depicted in panel B) – the shifting distribution of patience across agents at different points in time. In 1300, societal patience is low and virtually no-one belongs to the dynasties with $\beta > 0.2$ (an annual discount factor of around 0.94).

[19]For example, consider two dynasties i and j with discount factors $\beta^i = 0.05$ and $\beta^j = 0.5$. From equation (20), the relative size of the two dynasties in the year 2000 ($t = 28$) and the year 1300 ($t = 0$) will differ by a factor of $\frac{N_0^i/N_0^j}{N_{28}^i/N_{28}^j} = \left(\frac{\beta^i}{\beta^j}\right)^{28} = 10^{-28}$.

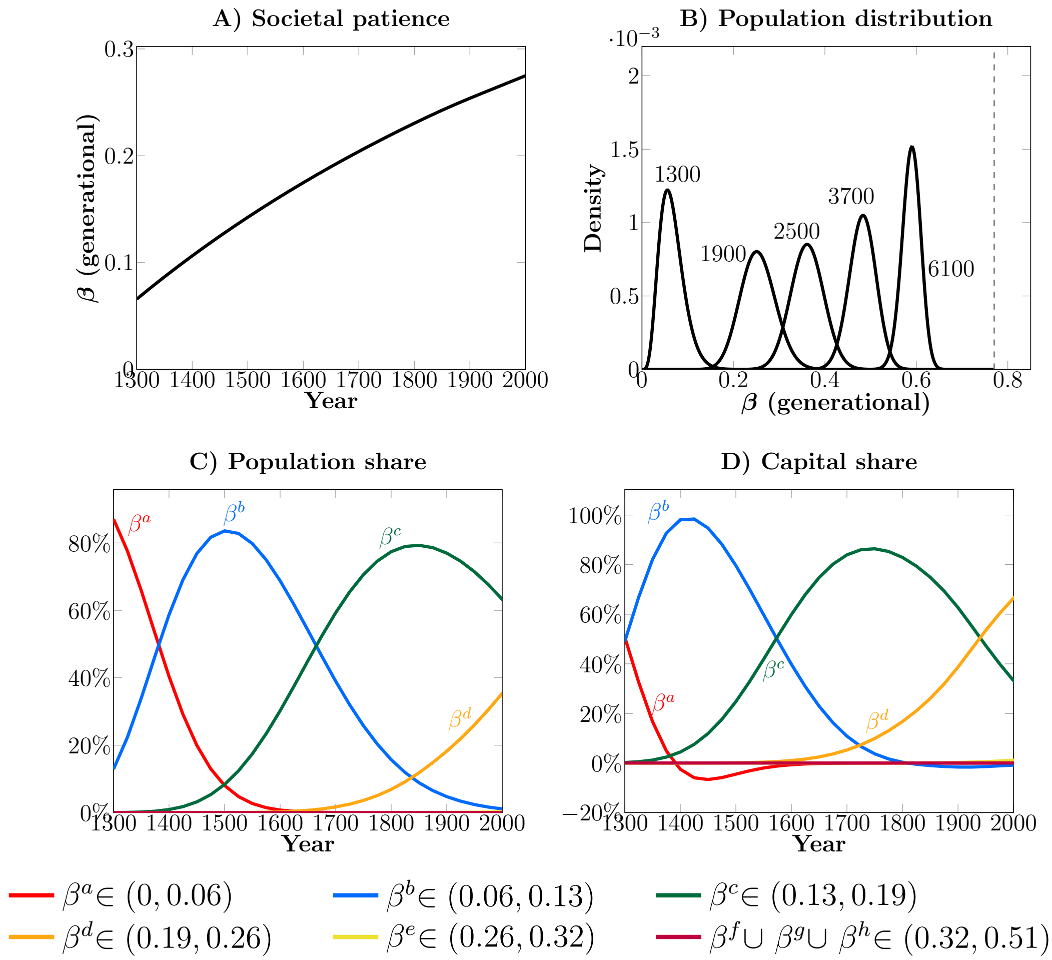
More patient households however, will tend to have more children who in turn will have the same higher levels of patience as their parents. The distribution of the population will thus shift towards higher levels of patience as relatively more patient households are born. By 1900, the median dynasty will have a generational discount factor of around $\beta = 0.25$. Panel B) also makes clear that while there is substantial variation in levels of patience across individuals at a point in time, this cross-sectional variation is substantially less than the change in patience as a whole over the 700 years we study.

To examine the changing composition of the population over time, we aggregate individual dynasties into eight groups by their level of patience.[20] Panel C) of Figure 3 shows the population share of each group over time. Initially, the global population is dominated by the least patient agents in group β^a, who constitute approximately 90% of the total population in 1300. Over time, since the group as a whole has fewer children than more patient groups, the share of these agents falls and the group with the next highest patience level, β^b, takes their place, accounting for more than 80% of the population by around 1500. This process continues until the most patient group of agents eventually dominates the entire population. As can be seen from panel B), the shift from the least patient to the most patient group is gradual. Each dynasty and respective group experiences periods of dominance and decline over several centuries. To make things concrete, we can consider the relative fertility of different groups in the year 2000 in the $\omega = 0$ version, at which point the average number of children per household per period is 0.7 (recall that since adults can live for multiple periods, population growth can be positive even if $n_c < 1$). In 2000, the bulk of the population belongs to the β^c and β^d groups. On average, dynasties in β^d have 0.3 more children than their counterparts in β^c. The most patient agents (β^h) have 2.6 more children than the least patient agents (β^a). Of course, since there are vastly fewer households in the β^h group of dynasties in 2000 it takes many centuries before that fertility differential begins to be manifested as a non-negligible population share.

The key to understanding the cyclical pattern for each group lies in Figure 3 panel D), which reports the evolution of the capital share owned by each group. Given that agents can lend or borrow capital for optimal consumption and fertility

[20]That is, we split the interval $(0, \bar{\beta})$ into eight sub-intervals each containing the same mass of dynasties. The aggregate of dynasties in each sub-interval thus comprises the groups $\{\beta^a, \beta^b, \ldots, \beta^h\}$, where β^a is the group of the least patient dynasties.

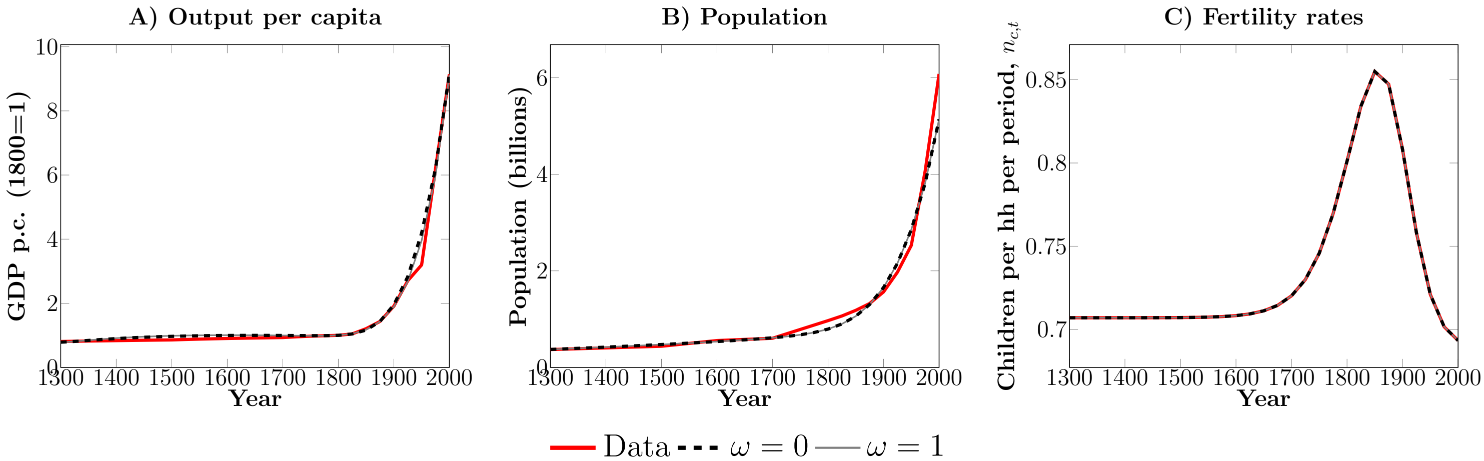
Figure 3: The rise of societal patience ($\omega = 0$)



Note: Panel A) depicts the societal average level of generational patience over time. Panel B) shows the distribution of levels of patience (the generational β) at the labelled years over the period 1300-6100. The dashed vertical line is at $\bar{\beta}$. Each panel C)–D) reports the sum of the model output across all dynasties in the group of dynasties defined in the legend. Results here from the model with imperfect altruism ($\omega = 0$); the results with perfect altruism ($\omega = 1$) are given in Appendix Figure A3.

decisions, the β^a-group initially depletes its capital, later borrowing from more patient dynasties to prioritize current consumption over children. Note that the ability of the more impatient dynasties to increase their consumption hinges on the population size and capital ownership of their relatively more patient counterparts. The growth of the β^b-group thus paves the way for the relative decline of the β^a-group as a more substantial market emerges for the latter's capital. As the β^a-group diminishes, so the β^b-group emerges as the largest population and the dominant owner of capital.

Figure 4: Output per capita, population, fertility and child cost



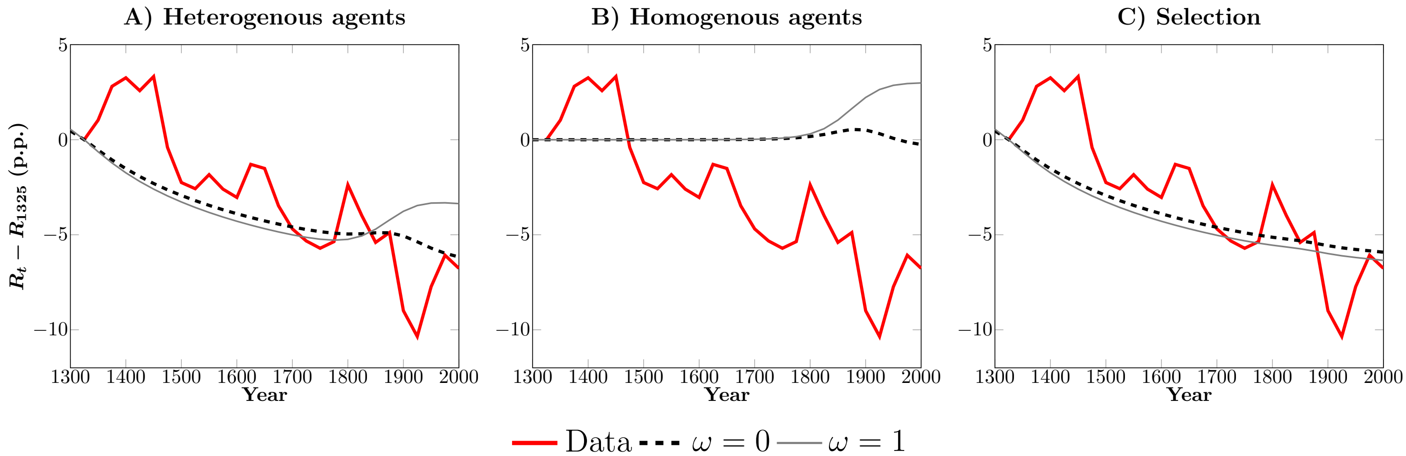
Note: Data for world output per capita and population is from The Maddison Project (2013). The models are calibrated to match exactly the fertility rates in the data, the derivation of which is described in Appendix D. The $n_{c,t}$ is children per household per generation. Since adults can live for multiple generations, population growth can be positive even if the number of children born to each household in each period is less than one. We report model outputs both with perfect altruism ($\omega = 1$) and with imperfect altruism ($\omega = 0$).

The eventual rise of the β^c-group provides the β^b-group with growing opportunities to maintain high consumption by selling their capital holdings. By around 1900, the β^b-group starts borrowing capital, albeit to a lesser degree given their relatively greater patience compared to the β^a-group. As Figure 3 also makes clear, over the period of study, the most patient dynasties represent an insignificant fraction of the population. Appendix Figure A3 presents analogous model outcomes for the scenario where $\omega = 1$. All results remain qualitatively identical.

Figure 4 reports the model outcomes (for both $\omega = 0$ and $\omega = 1$) and the data for aggregate population, output per capita and fertility rates. Panel A) of Figure 4 shows the growth of output per capita as close to zero up to around 1800, thereafter accelerating, as in the data. Panel B) demonstrates the excellent fit in the model to the growth of population in the data. The aggregate fertility rate over time is given in Figure 4 panel C), with an initial increase in fertility at the onset of the industrial revolution followed by a rapid decline that, by the calibration,[21] matches the pattern in the data (see Galor, 2005; Bar and Leukhina, 2010). This figure also makes clear

[21] The aggregate fertility series in the data is calculated using the smoothed data for population and life expectancy. The figure depicts the number of children born to each household in each period.

Figure 5: Selection and the interest rate



Note: Data is the Schmelzing (2020) global real interest rate; we report the 25-yearly median interest rate beginning in 1325 (since the Schmelzing real 'Global R' series begins in 1314). We report the model interest rates in annualized terms. The data and model outputs are normalized to zero in the year 1325. We report model outputs both with perfect altruism ($\omega = 1$) and with imperfect altruism ($\omega = 0$). Panel A) reports results with heterogenous agents whose distribution of patience is calibrated at the year 2000. Results in panel B) are based on a homogeneous-agent set-up where there is only one dynasty calibrated to match the average patience in the year 2000 in the heterogenous-agent model. Panel C) is the contribution of selection, defined as the difference between the heterogenous- and homogenous-agent model results.

that the aggregate path of fertility does not prevent the role that differences in fertility rates across dynasties play in driving higher societal patience over time.

Finally, we examine the model's prediction for the decline in the interest rate. Figure 5 panel A) reports the calibrated model against the Schmelzing (2020) global real interest rate series, each normalized to their values in 1325. Up to 1775, in the presence of a constant growth rate and stable life expectancy, steadily increasing societal patience acts endogenously to depress the equilibrium interest rate over time. After 1775, the growth rate of output per worker increases, putting upward pressure on the interest rate in both versions of the model. In the model with perfect altruism ($\omega = 1$), since there is no life-expectancy channel, the growth-effect overwhelms the ongoing trend in societal patience and interest rates begin to rise. Where life expectancy does play a role in the presence of imperfect altruism ($\omega = 0$), the growth in life expectancy around the same time as the increased growth rate of output per capita means that each mechanism roughly cancels each other out, leaving the trend in societal patience to continue to drive declining rates. While we have no clear way to calibrate ω, a value nearer to $\omega = 0$ thus appears more plausible given its better fit with the interest rate series. Panel B) reports the simulation outcomes in corresponding models with homogenous agents, i.e., where all households have a constant value of β. In this scenario, since there can be no trend in societal patience, the only mechanisms operating to change equilibrium interest rates are growth rates and life expectancy: again, growth pushes up interest rates in the absence of the life-expectancy channel ($\omega = 1$); where life expectancy plays a role, the two channels act in opposite directions and together cause only minimal change in the interest rate. Panel C) of Figure 5 then reports the contribution of selection as the difference between the heterogenous- and homogenous-agent models. As is clear, the selection mechanism alone can account for most of the decline in the interest rate over time. Without preference heterogeneity and the associated selection, we would be unable to explain the historical decline in rates.

5.1 Decomposition

Equation (18) encapsulates the four channels governing the change in the interest rate in the model: population growth, productivity growth, changes to life expectancy, and, changes to average patience. We now turn to quantifying the importance of each of these mechanisms in explaining various elements of the historical data. To do so, we simulate counterfactuals in which we shut down each of these channels separately. In a final counterfactual we shut down all channels except selection.

Table 2 presents the quantitative implications of the model compared with the data under different assumptions. We report the decline in the interest rate over the whole period for which we have data ($\Delta_{2000-1325}R_t$), the average annualized growth rate of income per capita ($\bar{g}_y$), the average annualized growth rate of population ($\bar{g}_N$), the increase in fertility from 1325 to its peak in 1875 ($\Delta_{1875-1325}N_{c,t}$) as well as its decline from 1875 to 2000 ($\Delta_{2000-1875}N_{c,t}$). Figure 6 illustrates the decline of interest rates in each model variant, relative to 1325.

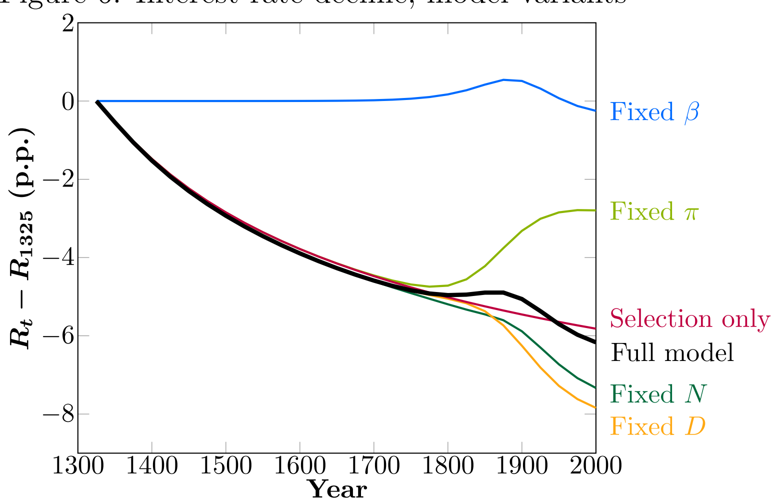
What becomes clear is that time-varying productivity, population growth and life expectancy in the model are key for matching the average growth of per capita income, population and fertility in the data. However, these forces are not able to account for the decline in the interest rate. The model with constant population ('Fixed N') fails – by construction – to match either the average population growth

Table 2: Decomposition of channels

	Data	Full model	Fixed N	Fixed D	Fixed π	Fixed β	Selection only
$\Delta_{2000-1325} R_t$ (p.p.)	-6.77	-6.17	-7.33	-7.84	-2.79	-0.25	-5.82
$\bar{g}_y(\%)$	0.40	0.38	0.38	0.13	0.38	0.38	0.10
$\bar{g}_N(\%)$	0.43	0.39	0.00	0.39	0.39	0.39	0.00
$\Delta_{1875-1300} N_{c,t}$	0.15	0.14	-0.10	0.14	0.24	0.14	0.00
$\Delta_{2000-1875} N_{c,t}$	-0.16	-0.15	-0.23	-0.15	0.08	-0.15	0.00

Note: We report model implications against the data from the Schmelzing (2020) global real interest rate; we report the 25-yearly median interest rate beginning in 1325 (since the Schmelzing real 'Global R' series begins in 1314). The 'Full model' is that with all channels operating (that is, also with imperfect altruism $\omega = 0$). $\bar{g}_X$ denotes the mean annualized growth rates over the period 1300-2000. Each model variant holds fixed the specified channels leaving others to operate. 'Selection only' is a counterfactual where we hold constant all of N, D and π.

Figure 6: Interest rate decline, model variants



Note: The figure depicts the model-implied decline in interest rates since 1325 in each of the model variants specified in Table 2, i.e., the 'Full model' is that with all channels operating (that is, also with imperfect altruism $\omega = 0$). Each model variant holds fixed the specified channels leaving others to operate. 'Selection only' is a counterfactual where we hold constant all of N, D and π.

or the fertility pattern but it can explain the decline in interest rates since selection drives higher average patience. This makes clear that it is not the population size that impacts interest rates but rather its changing composition. The model with constant technology ('Fixed D') struggles to explain most of the average growth in output, but it is able to understand a fall in interest rates as selection continues to operate. Note that in this model variant, some growth in income per capita arises even in the absence of technological progress – as societal patience increases, so the average saving rate increases and capital accumulates leading to per capita output growth. We return to this implication in section 5.2. The model with constant life expectancy ('Fixed π') fails to capture the fall in fertility after 1875, but also means that higher growth in D after 1800 predicts a counterfactually increasing interest rate (as shown in Figure 6).

In terms of explaining the interest rate, the full model explains 91.1% of the historical decline in the interest rate. Turning off all channels except the selection mechanism ('Selection only') yields a model decline in the interest rate of 86% that in the data. If we retain all channels except selection ('Fixed β'), we completely fail to understand the decline in rates – the model then explains only 3.7% of the historical decline. Selection alone is the key to explaining the decline of the interest rate.

5.2 Further implications

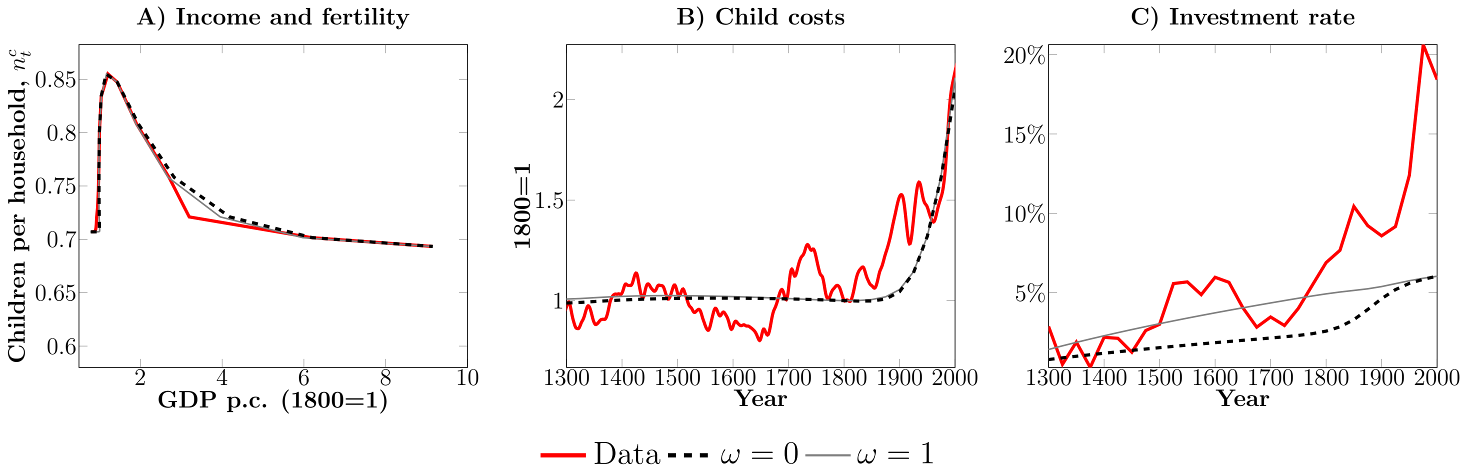
In addition to accurately matching the decline in interest rates and other macroeconomic trends, there are two more aspects of the quantitative findings that merit further discussion.

Fertility and child costs To match the trend in global population over time, we exogenously varied the path for a_t, the parameter that, along with D_t, governs the cost of children for households. Since, by construction, we capture the time path of fertility in a growing economy, we also capture a relationship between income and fertility that switches from positive to negative once a certain threshold level of income is reached. This is shown in Figure 7 panel A), which presents model predicted income-fertility pairs at different points in time. Both time-series and cross-country evidence support this: typically, a negative correlation between income and fertility emerges once countries surpass a certain income level.[22]

Since child costs within the model are not targeted directly but are calculated from other calibrated parameters (since $q_t = a_t D_t^{\frac{1}{1-\nu}}$), we can validate the role that child costs play by comparing the model against the data. One proxy for the cost of children is the relative price of the service sector goods with respect to the Consumer

[22]See, for example, Galor and Weil (2000), Jones and Tertilt (2008) and Manuelli and Seshadri (2009); more recent discussion of potential changes to the relationship is in Doepke et al. (2022).

Figure 7: Further quantitative implications



Note: Panel A) depicts output per capita against fertility from Figure 4. Panel B) reports child costs in the model ($q_t \equiv a_t D_t^{\frac{1}{1-\nu}}$) against the price of services relative to CPI in the data, which we take as a proxy for child costs. Panel C) gives the series for the investment rate which is from Broadberry and de Pleijt (2021) over the period 1300–1825 (which covers England up to 1700 and Great Britain thereafter) and from Thomas and Dimsdale (2017) for 1875–2000. For 1850 we take the mean of these two series. We report model outputs both with perfect altruism ($\omega = 1$) and with imperfect altruism ($\omega = 0$).

Price Index (see Appendix A.6).[23] This captures the idea that the prices of service goods, such as those that affect the costs of raising children, increase with income due to relatively low productivity growth (the Baumol and Bowen (1965) cost disease). Figure 7 panel B) shows a close fit between the relative price of children implied by our model[24] and the proxy for the child cost in the data, suggesting that our use of child costs to fit the population series is plausible. A number of factors outside of our model could underpin this change over time. For example, variation in laws that prevent children working, the introduction of mandatory education, parents spending more time with children, changing social norms, and so on, can all affect the relative cost of having children.

[23]On measurement issues and available data see Deaton and Muellbauer (1986) and Donni (2015).

[24]We calculate the Consumer Price Index in our model as a Paasche Index of consumption goods and children with a base period in the year 2000. We use a Paasche Index instead of a Laspeyres index to take into account the changing consumption weights over the 700 years under consideration which the Laspeyres index holds constant.

Investment ratio An implication of steadily rising average patience is that savings rates should also increase over time. This is consistent with empirical evidence on the relationship between patience and savings (Cronqvist and Siegel, 2015, Sunde et al., 2021). In our model, the endogenous growth in savings rates also sustains significant growth in income per capita even in the absence of technological change (see Table 2). There is also good empirical evidence that savings rates have increased over long periods of time. Sutch, ed (2006) documents an increase in the gross private saving rate in the U.S. from around 10% to over 30% over the period 1834–1909. The savings rate then declines, particularly in the latter half of the twentieth century, largely for reasons extrinsic to our model.[25] In the UK, the household savings rate grew from less than 5% in the nineteenth century to around 10% in 2000 (Thomas and Dimsdale, 2017). Data prior to the nineteenth century on household savings is scarce, but data on investment is available for the UK. Figure 7 panel C) reports the UK investment rate from 1300 to 2000 (using Broadberry and de Pleijt, 2021 and Thomas and Dimsdale, 2017). As can be seen, the trend in the investment rate is increasing over the 700 years, accelerating after 1900. The Figure shows that in our model, the investment rate increases over time as societal patience grows, even if we do not match the level. Given that nothing in the calibration is intended to target these objects, we consider the fit to the data to be further validation of the role of increasing societal patience over time.

6 Imperfect transmission

Thus far, we have assumed that the level of patience within a dynasty is perfectly transmitted across generations. In reality, transmission is likely to be imperfect due to factors such as mutation, changing patterns of socialization or mean-reversion that may arise from imperfect assortative mating. An immediate question is whether such imperfect transmission slows the process of selection and hence diminishes its role in driving the declining interest rate.

To address this issue we consider a version of our model in which some portion of children in dynasty i inherit a mutated level of patience of $\{\beta^i - \varepsilon, \beta^i + \varepsilon\}$ for

[25]Much of the twentieth century variation in savings rates, at least for the U.S., has been shown to be the result of government transfers and of changes to consumption propensities which may be attributed to policy intervention (see, e.g. Gokhale et al., 1996).

some $\varepsilon > 0$. Children with a positive shock have what is called an advantageous mutation (one that increases fitness); whilst those with a negative shock have a deleterious (which decreases fitness) mutation. One consideration is whether such noise is distributed symmetrically or whether it is skewed toward the mean of the population. While Brenøe and Epper (2022) and Chowdhury et al. (2022) find strong transmission in patience across generations, neither study identifies the existence of asymmetric transmission. We thus consider the case of symmetric noisy transmission and its implications.

Specifically, in Appendix F we develop a version of our model which incorporates mutation as an unanticipated shock to an agent's discount factor. In the model setting, such 'mutation' is a reduced form way to consider the implications of imperfect transmission of preferences in general. Those agents that experience a deleterious mutation have very small effects on interest rates. Those that receive an equal-sized, but advantageous mutation can have large and long-lasting consequences since those agents begin to accumulate a greater share of capital and have a larger number of children (who themselves inherit the higher patience level). This highly asymmetric response to a symmetric shock demonstrates that even a small and ongoing process of imperfect transmission would serve to accelerate the pace of selection and the decline in the interest rate.

If some form of asymmetric, mean-reverting transmission did exist, it would need to be very strongly mean-reverting in order to offset the consequence of even a small number of advantageous mutations. Given the complexity that such skewed mutation bring to the model, given that there is no immediate way to calibrate such noisy transmission in general, and given the limited evidence in the literature to guide us in calibrating any potentially asymmetric mutation, we leave a fuller analysis of the implications of skewed imperfect transmission to future research.

7 Concluding remarks

We introduced a simple fertility model with heterogenous preferences, calibrated to the modern-day distribution in patience, and showed that the process of natural selection can explain the trend in the interest rate over the last eight centuries. The role of selection is robust to incorporating a number of extensions and to alternative calibration. There are many further implications to consider. First, in our model the population shift toward more patient types occurs partly via trading in capital. This suggests a potentially important relationship between the constraints on trade or borrowing, the evolution in the population and the interest rate. Second, we have focused on a simple form of the intergenerational transmission of preferences and pointed to some implications of imperfect transmission. We leave to future work a fuller consideration of the role of the strength and bias of transmission across generations. Moreover, we studied heterogenous patience levels as the only time-varying element of societal preferences. The evidence on the heterogeneity of altruism, risk aversion, and other preferences, together with their intergenerational transmission and effect on fertility, suggests that a number of additional further preference heterogeneities could evolve over time alongside time preference. Third, we have focused our model on its implications for the interest rate but our time period encompasses the onset of the industrial revolution and periods of mass migration. While we captured the path of income and population by exogenously varying child cost and productivity, these items could potentially be made endogenous. The role for the evolution of societal preferences in explaining these changes is left for future work.

Finally, we noted in the introduction that understanding social discount rates is critical in formulating optimal policies to address very long-term, inter-generational problems such as those that relate to the funding of social security programmes and that address climate change. Our analysis points to a new way to rationalize the use of a declining discount rate in cost-benefit analysis. Moreover, understanding the short- and long-run relationship between the social discount rate and policy interventions is an important avenue for future research.

8 Data availability

Data and code for replicating the tables and figures in this article can be found in Stefanski and Trew (2023) in the Harvard Dataverse, DATASET URL.

References

Alan, Sule and Martin Browning, "Estimating Intertemporal Allocation Parameters using Synthetic Residual Estimation," *The Review of Economic Studies*, October 2010, *77* (4), 1231–1261.

Allen, Robert, "The Great Divergence in European Wages and Prices from the Middle Ages to the First World War," *Explorations in Economic History*, 2001, *38* (4).

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and Elisabet E. Rutström, "Eliciting Risk and Time Preferences," *Econometrica*, 2008, *76* (3), 583–618.

Arrow, Kenneth J., Maureen L. Cropper, Christian Gollier, Ben Groom, Geoffrey M. Heal, Richard G. Newell, William D. Nordhaus, Robert S. Pindyck, William A. Pizer, Paul R. Portney, Thomas Sterner, Richard S. J. Tol, and Martin L. Weitzman, "Determining Benefits and Costs for Future Generations," *Science*, 2013, *341* (6144), 349—350.

Auclert, Adrien, Hannes Malmberg, Frederic Martenet, and Matthew Rognlie, "Demographics, Wealth, and Global Imbalances in the Twenty-First Century," *Working Paper 29161, NBER, Cambridge, MA.*, 2021.

Bar, Michael and Oksana Leukhina, "Demographic transition and industrial revolution: A macroeconomic investigation," *Review of Economic Dynmics*, 2010, *13*, 424—451.

Barro, Robert J. and Gary S. Becker, "Fertility Choice in a Model of Economic Growth," *Econometrica*, March 1989, *57* (2), 481–501.

Bauer, Michal and Julie Chytilová, "Women, Children and Patience: Experimental Evidence from Indian Villages: Women, Children and Patience," *Review of Development Economics*, November 2013, *17* (4), 662–675.

Baumol, William J. and William G. Bowen, "On the Performing Arts: The Anatomy of Their Economic Problems," *The American Economic Review*, 1965, *55* (1/2).

Becker, Gary S., "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology," *Journal of Economic Literature*, 1976, *14* (3), 817–26.

_ **and Robert J. Barro**, "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics*, February 1988, *103* (1), 1–25.

Becker, Robert A., "On the Long-Run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households," *Quarterly Journal of Economics*, 1980, *95* (2), 375–82.

Bennett, Patrick, Chiara Ravetti, and Po Yin Wong, "Losing in a Boom: Long-term Consequences of a Local Economic Shock for Female Labour Market Outcomes," *Labour Economics*, 2021, *73* (102080).

Bernstein, P. L., *Against the Gods: The Remarkable Story of Risk*, Wiley, 1998.

Black, Dan A., N. Kolesnikova, S. G. Sanders, and Lowell J. Taylor, "Are Children "Normal"?," *Review of Economics and Statistics*, 2013, *95* (1), 21–33.

Blanchard, Olivier J., "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 1985, *93* (2), 223–47.

Brenøe, Anne A. and Thomas Epper, "The Intergenerational Transmission of Time Preferences Across Four Decades," *European Economic Review*, 2022, *148* (104208).

Broadberry, Stephen and John Joseph Wallis, "Growing, Shrinking, and Long Run Economic Performance: Historical Perspectives on Economic Development," Technical Report w23343, National Bureau of Economic Research, Cambridge, MA April 2017.

_ **and Roger Fouquet**, "Seven Centuries of European Economic Growth and Decline," *Journal of Economic Perspectives*, 2015, *29* (4).

Broadberry, Stephen N. and Alexandra de Pleijt, "Capital and Economic Growth in Britain, 1270-1870: Preliminary Findings," *Discussion Paper DP15889, CEPR, London*, 2021.

Buera, Francisco J. and Joseph P. Kaboski, "Scale and the Origins of Structural Change," *Journal of Economic Theory*, March 2012, *147* (2).

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio, "Demographics and Real Interest Rates: Inspecting the Mechanism," *European Economic Review*, 2016, *2016* (September), 208—226.

Cesarini, David, Erik Lindqvist, Robert Östling, and Anastasia Terskaya, "Fortunate Families? The Effects of Wealth on Marriage and Fertility," *Working Paper 31039, NBER, Cambridge, MA*, 2023.

Chowdhury, Shyamal, Matthias Sutter, and Klaus Zimmermann, "Economic Preferences across Generations and Family Clusters: A Large-Scale Experiment in a Developing Country," *Journal of Political Economy*, April 2022, (720395).

Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf, "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England," *Demography*, April 2017, *54* (2), 413–436.

_ , _ , and _ , "Further Evidence of Within-Marriage Fertility Control in Pre-Transitional England," *Demography*, August 2019, *56* (4), 1557–1572.

Clark, Gregory, "The Cost of Capital and Medieval Agricultural Technique," *Explorations in Economic History*, 1988, *25*, 265–294.

_ , *A Farewell to Alms*, Princeton University Press, Princeton and Oxford, 2007.

_ , "Genetically Capitalist? The Malthusian Era, Institutions and the Formation of Modern Preferences," *mimeo*, 2007.

_ , "The Macroeconomic Aggregates for England, 1209-1869," in A. J. Field, ed., *Research in Economic History*, Vol. 27, Emerald Group Publishing, Bingley, 2010, pp. 51—140.

_ **and Gillian Hamilton**, "Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England," *The Journal of Economic History*, 2006, *66* (3), 707–36.

_ **and Neil Cummins**, "Randomness in the Bedroom: There Is No Evidence for Fertility Control in Pre-Industrial England," *Demography*, August 2019, *56* (4), 1541–1555.

_ , _ , and Matthew Curtis, "Twins Support the Absence of Parity-Dependent Fertility Control in Pretransition Populations," *Demography*, August 2020, *57* (4), 1571–1595.

Crafts, Nicholas and Terence Mills, "Trend TFP Growth in the United States: Forecasts versus Outcomes," *Discussion Paper DP12029, CEPR, London*, 2017.

Cronqvist, Henrik and Stephan Siegel, "The Origins of Savings Behavior," *Journal of Political Economy*, February 2015, *123* (1), 123–169.

de la Croix, David, Eric B. Schneider, and Jacob Weisdorf, "Childlessness, Celibacy and Net Fertility in Pre-Industrial England: The Middle-Class Evolutionary Advantage," *Journal of Economic Growth*, September 2019, *24* (3), 223–256.

Deaton, Angus S. and John Muellbauer, "On Measuring Child Costs: With Applications to Poor Countries," *Journal of Political Economy*, August 1986, *94* (4), 720–744.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti, "Global Trends in Interest Rates," 2018. Working Paper 25039, NBER, Cambridge, MA.

Dimson, Elroy, Paul Marsh, and Mike Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.

Doepke, Matthias and Fabrizio Zilibotti, "Occupational Choice and the Spirit of Capitalism," *Quarterly Journal of Economics*, May 2008, *123* (2), 747–793.

_ **and Michele Tertilt**, "Families in Macroeconomics," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, North Holland, Amsterdam, 2016, pp. 1789–1891.

_ **, Anne Hannusch, Michele Tertilt, and Fabian Kindermann**, "The Economics of Fertility: A New Era," 2022. Working Paper 29948, NBER, Cambridge, MA.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "The Intergenerational Transmission of Risk and Trust Attitudes," *The Review of Economic Studies*, 2011, *79* (2), 645–77.

Donni, Olivier, "Measuring the Cost of Children," *World of Labor, IZA, Bonn*, 2015, (132).

Duarte, M. and D. Restuccia, "The Role of the Structural Transformation in Aggregate Productivity," *The Quarterly Journal of Economics*, 2010, *125(1)*, 129–173.

Epper, Thomas, Ernst Fehr, Helga Fehr-Duda, Claus Thustrup Kreiner, David Dreyer Lassen, Søren Leth-Petersen, and Gregers Nytoft Rasmussen, "Time Discounting and Wealth Inequality," *American Economic Review*, April 2020, *110* (4), 1177–1205.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 2018, *133* (4), 1645–92.

Featherstone, Allen M. and Timothy G. Baker, "An Examination of Farm Sector Real Asset Dynamics: 1910–85," *American Journal of Agricultural Economics*, 1987, *69* (3), 532–546.

Galor, Oded, "Chapter 4 From Stagnation to Growth: Unified Growth Theory," in "Handbook of Economic Growth," Vol. 1, Elsevier, Amsterdam, 2005, pp. 171–293.

_ **and David N. Weil**, "Population, technology and growth: From the Malthusian regime to the demographic transition'," *American Economic Review*, 2000, *90* (4), 806–28.

_ **and Omer Moav**, "'Natural Selection and the Origin of Economic Growth'," *Quarterly Journal of Economics*, 2002, *117* (4), 1133–91.

_ **and Özak Özak**, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064–103.

Giannelis, Alexandros, Emily A. Willoughby, Robin Corley, Christian Hopfer, John K. Hewitt, William G. Iacono, Jacob Anderson, Aldo Rustichini, Scott I. Vrieze, Matt McGue, and James J. Lee, "The Association between Saving Disposition and Financial Distress: A Genetically Informed Approach," *Journal of Economic Psychology*, June 2023, *96*, 102610.

Gokhale, Jagadeesh, Laurence J. Kotlikoff, and John Sabelhaus, "Understanding the Postwar Decline in U.S. Saving: A Cohort Analysis," *Working Paper 5571, NBER, Cambridge, MA*, 1996, p. 93.

Gollin, Douglas, "Getting Income Shares Right," *Journal of Political Economy*, April 2002, *110* (2), 458–474.

Griffiths, Anthony J. F., "Mutation," *Encyclopædia Britannica*, 2020.

Hansson, Ingemar and Charles E. Stuart, "'Malthusian Selection of Preferences'," *American Economic Review*, 1990, *80* (3), 529–44.

Herrington, Gaya, "Update to limits to growth: Comparing the World3 model with empirical data," *Journal of Industrial Ecology*, June 2021, *25* (3), 614–626.

Hudson, Michael, "How Interest Rates Were Set, 2500 BC-1000 AD: Máš, tokos and fœnus as Metaphors for Interest Accruals," *Journal of the Economic and Social History of the Orient*, 2000, *43* (2), 132–61.

Jablonka, Eva and Marion J. Lamb, *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life* Life and Mind, Cambridge, Mass: MIT Press, 2005.

Jones, Larry E. and Michèle Tertilt, "An Economic History of Fertility in the United States: 1826–1960," in Peter Rupert, ed., *Frontiers of Family Economics*, Vol. 1, Bingley: Emerald Group Publishing, Bingley, 2008, chapter 5, pp. 165–230.

Kearney, Melissa S. and Riley Wilson, "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence From The Fracking Boom," *Review of Economics and Statistics*, 2018, *100* (4), 678–90.

Kolk, Martin, "The Relationship between Life-Course Accumulated Income and Childbearing of Swedish Men and Women Born 1940–70," *Population Studies*, November 2022, pp. 1–19.

Krueger, Dirk and Alexander Ludwig, "On the Consequences of Demographic Change for Rates of Returns to Capital, and the Distribution of Wealth and Welfare," *Journal of Monetary Economics*, January 2007, *54* (1), 49–87.

Laubach, Thomas and John C. Williams, "Measuring the Natural Rate of Interest," *The Review of Economics and Statistics*, 2003, *85* (4), 1063–1070.

Levy, M., "An evolutionary explanation for risk aversion," *Journal of Economic Psychology*, 2015, *46*, 51–61.

Lovenheim, Michael F and Kevin J. Mumford, "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," *Review of Economics and Statistics*, 2013, *95* (2), 464–75.

Lucas, Robert E., *The industrial revolution: Past and future.*, Harvard University Press, Cambridge, 2002.

Machiavelli, Niccolo, *The Prince*, Dante University Press, 2003.

Manuelli, Rodolfo E. and Ananth Seshadri, "Explaining International Fertility Differences," *Quarterly Journal of Economics*, May 2009, *124* (2), 771–807.

Mian, Atif, Ludwig Straub, Amir Sufi, and Chicago Booth, "What Explains the Decline in R*? Rising Income Inequality versus Demographic Shifts," 2021, p. 47.

Millner, Antony, "Nondogmatic Social Discounting," *American Economic Review*, March 2020, *110* (3), 760–775.

Parker, Geoffrey, *The Grand Strategy of Philip II*, Yale University Press, 2000.

Pinker, S., *The Better Angels of Our Nature: A History of Violence and Humanity*, Penguin, 2012.

Ramsey, Frank P., "A mathematical theory of saving," *Economic Journal*, 1928, *38* (152), 543–59.

Riley, James C., "Estimates of Regional and Global Life Expectancy, 1800–2001," *Population and Development Review*, 2005, *31* (3), 537–543. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1728-4457.2005.00083.x.

Robson, Arthur J., "A Biological Basis for Expected and Non-expected Utility," *Journal of Economic Theory*, 1996, *68*, 397–424.

_ **and Balazs Szentes**, "Evolution of Time Preference by Natural Selection: Comment," *American Economic Review*, 2008, *98* (3), 1178–88.

Rogers, Alan R., "Evolution of Time Preference by Natural Selection," *American Economic Review*, 1994, *84* (3), 460–81.

Rogoff, Kenneth S., Barbara Rossi, and Paul Schmelzing, "Long-Run Trends in Long-Maturity Real Rates 1311-2021," *Working Paper 30475, NBER, Cambridge, MA*, 2022.

Roser, Max, Esteban Ortiz-Ospina, and Hannah Ritchie, "Life Expectancy," *Our World in Data*, May 2013.

Salinger, M. J., "'Climate Variability and Change: Past, Present and Future'," *Climate Change*, 2005.

Schmelzing, Paul, "Eight centuries of global real interest rates, R-G, and the 'suprasecular' decline, 1311–2018," January 2020. No. 845, Bank of England Staff Working Paper, London.

Schoonbroodt, Alice and Michèle Tertilt, "Property Rights and Efficiency in OLG Models with Endogenous Fertility," *Journal of Economic Theory*, March 2014, *150*, 551–582.

Smits, J.P., P.J. Woltjer, and D. Ma, "A Dataset on Comparative Historical National Accounts, ca. 1870-1950: A Time-Series Perspective," *Groningen Growth and Development Centre Research Memorandum GD-107, Groningen: University of Groningen.*, 2009.

Stefanski, Radek and Alex Trew, "Replication Data for: "Selection, Patience, and the Interest Rate"," 2023. Harvard Dataverse, https://doi.org/10.7910/DVN/UGWOHH.

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim, "Patience and Comparative Development," *The Review of Economic Studies*, December 2021, *89* (5), 2806—40.

Sutch, Richard, ed., "Chapter Ce Saving, Capital, and Wealth," in Richard Sutch, ed., *Historical Statistics of the United States*, Millenial Edition On Line, Cambridge University Press, Cambridge, 2006, pp. 3–287–3–297.

Tamura, Robert, "From decay to growth: A demographic transition to economic growth," *Journal of Economic Dynamics and Control*, 1996, *20*, 1237–61.

Teulings, Coen and Richard Baldwin, eds, *Secular Stagnation: Facts, Causes and Cures*, CEPR Press, London, 2014.

T'Hart, Joost Jonker Marjolein and Jan Luiten van Zanden, eds, "A financial history of The Netherlands," 1997.

The Maddison Project, "The Maddison Project, 2013 version," 2013. http://www.ggdc.net/maddison/maddison-project/home.htms.

Thomas, Ryland and Nicholas H. Dimsdale, "A Millennium of UK Data," *OBRA Dataset, Bank of England, London*, 2017.

Timmer, M. P., G. J. de Vries, and K. de Vries, *Patterns of Structural Change in Developing Countries* 2015.

UN, "World Population Prospects 2019, Online Edition. Rev. 1.," 2019.

Vischer, Thomas, Thomas Dohmen, Armin Falk, David Huffman, Jürgen Schupp, Uwe Sunde, and Gert G. Wagner, "Validating an ultra-short survey measure of patience," *Economics Letters*, 2013, *120* (2), 142–145.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp, "The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements," *Schmollers Jahrbuch (Journal of Applied Social Science Studies)*, 2007, *127* (1), 139–69. doi:10.5684/soep.v33.

Weitzman, Martin L., "Gamma Discounting," *American Economic Review*, March 2001, *91* (1), 260–271.

Wrigley, Edward Anthony A., R. S. Davies, Jim E. Oeppen, and Roger S. Schofield, *English Population History from Family Reconstitution 1580–1837*, Cambridge University Press, Cambridge, 1997.

Zijdeman, Richard and Filipa Ribeira da Silva, "Life Expectancy at Birth (Total)," December 2015. Type: dataset.

ONLINE APPENDIX

A Data Appendix

A.1 Detail on interest rate data

Schmelzing (2020) considers a number of measures of real interest rates over time, which vary by the asset class and region. Figure 1 panel A) reports the 25 year medians of the annualized generational returns based on the headline annual series 'Global R' in real terms (based on sheet II column N of the data appendix accompanying Schmelzing, 2020). In this section we describe these measures and supplement them with country-specific real rates of return on land based on the work of Clark (1988). The findings all point to a centuries-long downward trend in real interest rates – regardless of the measure used and regardless of the region under examination.[26] Figure 1 panel B) is from Figure 5 of Clark (2010).

Safe or risk-free rate The main measure introduced by Schmelzing (2020), and the real rate used in our paper in our Figure 1 panel A), is the 'risk-free' measure. Schmelzing describes this as the real interest rate for the historical 'safe asset provider'. The series is constructed by splicing together yields of long-term, marketable, sovereign-bond debt issued by the countries that were considered to be the safest and most reliable in a given period of time. The series runs from 1311 to 2018, using data from Italy, Spain, Holland, UK, Germany and the US. Importantly each of the types of debt was traded on deep secondary markets and the series' "central feature consists of the fact that it remained default-free over its 707 year span" (op. cit., p.18). The nominal rates of return are deflated using country-specific price data from Allen (2001). For details of the assets used, the countries under consideration, the chosen splice points as well as the justification of those countries and dates, see Table A1. Whilst arguably the exact timing of the splice points is somewhat subjective, Schmelzing very carefully lays out the case for the selected countries and their debt being the safest assets available in their given time. He also shows that the return on land consistently coincides with the safest asset.

Country specific Schmelzing extends the data used in the safe-asset calculations to generate a 700 year long series for all countries in that exercise as well as a number

[26]For expositional ease, all results in the section are presented as 50-year averages of generational rates of return.

Table A1: Details of Schmelzing's Global 'Safe Rate'.

Period	Country	Type of Assets	Justification for: Start Date	End Date
1311-1509	Italy	Venetian *Prestisi* and Genoese *Luoghi*. Earliest marketable long-term sovereign bond debt.	Earliest inflation data available from 1311, (Allen, 2001).	Battle of Agnadello (1509). Venice lost "in one day what took them eight hundred years exertion to conquer", (Machiavelli, 2003)
1510-1598	Spain	*Juros* long-term debt (de-facto sovereign debt: sold for cash, established seniority system, traded in secondary market). Cont. serviced unlike short-term debt.	"During the 16th century no other power controlled ... armed forces as powerful or financial resources as vast as Habsburg Spain,"(Parker, 2000).	Philip II's death in 1598 & Spanish decline: "The empire on which the sun never set had become a target on which the sun never set", (Parker, 2000).
1599-1702	Holland	Long term bond debts (*Renten* and obligations) issued by Dutch province	"Financial capital of the world," (Marjolein T'Hart and van Zanden, eds, 1997)	Transition of financial markets from Amsterdam to London
1703-1907	UK	British consol yields	Britain Europe's "most vibrant" economy, (Broadberry and Fouquet, 2015)	Germany overtakes UK in GDP
1908-1913	Germany	German Imperial 3% benchmark	Strongest growth trajectory	World War 1
1914-1918	UK	British consol yields	UK regains GDP primacy	Cost of War, lower GDP
1919-1961	US	10-year treasury bonds	US GDP pc permanently surpasses UK	Great Inflation in US
1962-1980	Germany	10-year government bonds	Revaluation of D-mark, rise of eurodollar market & low inflation rates.	Paul Volcker's successful 'war on inflation'
1981-2018	US	10-year treasury bonds	Largest GDP, low inflation	-

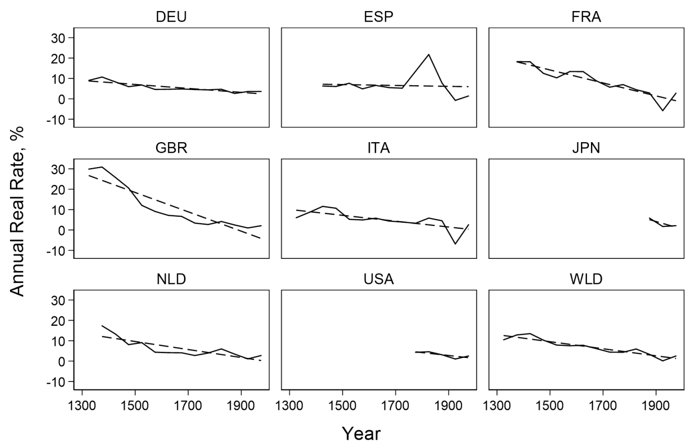
of other economically important countries. In particular he constructs rates for Italy, UK, Holland/NL, Germany, France, United States, Spain and Japan. Data for each country consists of long-term debt yields. For countries and time periods included in the global 'safe' series, the debt instruments remain the same and consist of the sovereign debt discussed above. For countries and/or periods not covered in the 'safe' series, observations are arithmetically weighted on the country-level across data points of long-term consolidated debt (such as debt issued by municipalities or mortgage-like pledge loans) and sovereign personal loans (like loans to the British Crown or French Revolutionary war loans to the United States) until marketable, national bond data becomes available. The nominal rates of return are deflated using country-specific price data from Allen (2001). As can be seen in the first panels graphs of Figure A.1 panel A), the real rates of return are declining in each country under consideration.

Global Schmelzing then constructs a global interest rate series by weighting the country-specific data above using GDP shares derived from The Maddison Project (2013). The GDP share of the eight countries under consideration are on average 80.1%, and for the past 600 years they have never fallen below 52%. As can be seen in the last panel (WLD) of Figure A.1 panel A), the global real rate of return is steadily declining over the entire period.

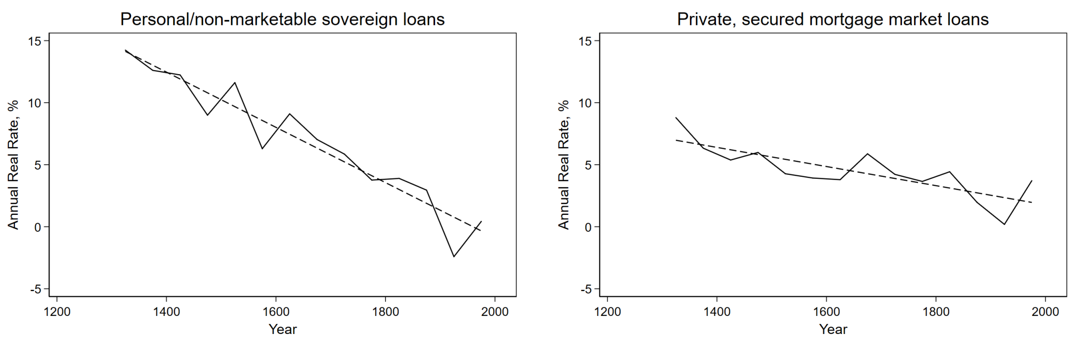
'Personal' or 'Sovereign' non-marketable loans Schmelzing also examines the extent to which the non-marketability of loans can account for the decline of interest rates presented above by examining personal loans to sovereigns (including "pledge loans" and loans from municipalities to the central authorities). These types of loans were very common, outside "of the urban financial centers of Northern and Central Europe in late medieval and early modern times, prior to the consolidation of debt on the national level, (...) especially in war episodes and in the context of weak central bureaucracies, (...) until well into the 17th century (...). Such non-marketable sovereign loans have gone out of fashion over the past two centuries." (op. cit., p.9). As Schmelzing notes, "A 'benchmark' non-marketable instrument today is represented by U.S. savings bonds, which are non-transferable, long-term, and redeemable after 12 months." (p.11) Since there was considerably more scope to distort market prices of capital in these circumstances, it is interesting to see if the rate of decline in these types of loans is any larger than in the safe-series or in the global-series. The analysis

Figure A1: Country specific real rates of return on long-term debt and land. Dashed line show regression trends.

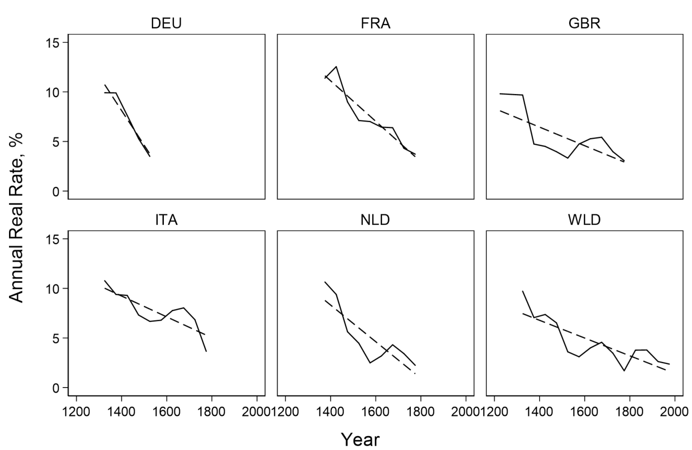
A) Rates of Return on long-term debt



B) Rates of return on personal/non-marketable loans to sovereigns and private debt



C) Rates of return on land (Flanders/Netherlands, Italy) and rent charges (France, Germany)



focuses on 454 non-marketable sovereign loans but excludes 'all intra-governmental loans, loans featuring in-kind payments, forced loans and those which are de facto expropriations'. The prices are adjusted for inflation using arithmetically weighted inflation rates from Allen (2001). The results are shown in the first panel of Figure A.1 panel B); here too we observe falling interest rates. Importantly the rate of decline of interest rates is very similar to other measures of interest rates.

Private, 'non-sovereign' rates Schmelzing also examines non-sovereign (private) real interest rates. In particular, he constructs a consistent series from the private, secured mortgage market over last 700 years within "Carolignian Europe" – mostly Germany, Switzerland, some parts of France and Holland. These debts "all involve the debtor as a private party who pays the recorded interest rate, which is tied to the value of a real estate asset itself, or where the collateral involved consists of a real estate asset. The creditor counterparties involve abbeys, municipalities, or other private individuals." (op. cit., p.25). Contract length is often not specified but is for at least for 'one life'-time, thus this is certainly long-term private debt. The instruments involved historically are *Leibrenten* or *Erbleihen* which changed into Pfandbriefe in the 19th century and still exist today. Inflation data once more comes from Allen (2001). The result is shown in the second panel of Figure A.1 panel B) and also demonstrates a steady decline over time.

Land Using data for nominal returns to farmland and rent-charges reported in Clark (1988) as well as inflation data from Schmelzing, we construct real interest rates on land for various countries. In particular, the first five panels of Figure A.1 panel C) show the real rates of return on land – arguable the 'safest asset' – for 5 countries (Italy, U.K., Flanders, France and Germany).[27] In addition, Schmelzing constructs a real interest rate on land using similar sources, specifically Ward (1960, cited in Schmelzing), Featherstone and Baker (1987), and Clark (1988, 2010), for the 'G-5' countries (Italy, U.K., Flanders, France, U.S.). We report the GDP-weighted average in the last panel of Figure A.1 panel C). The high interest rates in 13th century England that can be seen shown in Figure 1 panel B). are echoed across northern

[27]The GBR series is constructed using the same nominal interest rate data as in Figure 1. Notice also that the real rates data for the Netherlands (i.e. NLD) is constructed using nominal interest rates from Flanders and inflation from Amsterdam - whilst not ideal this is the best we can do due to a lack of other data.

Europe with surprisingly close agreement and the declining pattern of real interest rates on land is a feature in every country in which long-term data is available.

In addition to data for the last eight centuries, there is also evidence of an even longer-run trend from ancient data, as shown in Table A2.

Table A2: Historical interest rates

Period	Place	Rate (%)	Note
3000-1900 BC	Sumer	20–25	Rate of interest on silver[a]
c.2500 BC	Mesopotamia	≥20	Smallest fractional unit[b]
1900–732 BC	Babylonia	10–25	Return on loans of silver[a]
C6th BC	Babylonia	16–20	Interest on loans[a]
C5th-2nd BC	Greece	≥10	Smallest fractional unit[b]
C2nd BC on	Rome	$\geq 8\frac{1}{3}$	Smallest fractional unit[b]
C1st-3rd AD	Egypt	9–12	Land return, interest on loans[a]
C1st-9th AD	India	15-30	Interest on loans[a]
C10th AD	South India	15	Yield on temple endowments[a]
1200 AD	England	10	Return on land, rent charges[a]
1200–1349 AD	Flanders, France, Germany, Italy	10–11	Return on land, rent charges[a]
C15th AD	Various European	9.43	Risk-free rental rate[c]
C16th AD	Ottoman Empire	10–20	Interest on loans[a]
C19th AD	Various European	3.43	Risk-free rental rate[c]
2000 AD	England	4–5	Return on land, rent charges[a]
2000–17 AD	Various European	1.24	Return on land, rent charges[c]

Notes: [a]*Calculated or referenced in Clark (2007b).* [b]*Hudson (2000).* [c]*Schmelzing (2020).*

A.2 The German Socio-Economic Panel

The German Socio-Economic Panel (SOEP) is a longitudinal dataset which has, since 1984, collected information by interview on around 30,000 unique individuals in nearly 11,000 households (see Wagner et al., 2007). Among the data collected is household net income, marital status and age. Of particular use to this paper is a question asking for 'general personal patience' on a scale of 0-10 (where 0 is very impatient and 10 is very patient). This question was asked in 2008 and 2013. We use SOEP-Core version 33.1 which includes data up to 2016. Since there is some variability in self-reported patience of individuals between 2008 and 2013, we use the 2008 measure of patience since it has been validated using experimental methods (Vischer et al., 2013). We then focus on the number of unique children in each household at 2008 plus the number of additional household children up to 2013.

To construct our sample, we merge 2008 and 2013 using the 'never changing person ID'. We calculate the total number of children of each household as the number present at 2008 plus any additional children at 2013. We drop those 41 observations where patience is not observed in 2008 as well as the resident relatives and non-relatives. Our sample of 17,452 individuals thus leaves only the head of the household and their partner. The average number of children in each household is 0.71 (with a standard deviation of 1.00); the average number in a household that has at least one child is 1.71 (s.d. 0.84). The average patience level is 6.1 (s.d. 2.28).

Equation (13) gives the equilibrium relationship between dynasty population dynamics, the dynasty-specific discount rate and the gross real interest rate on children (which is common across dynasties). Since $N^i_{t+1} = N^i_t n^i_t$, we can re-write (13) in terms of the number of children each household has as simple $n^i_t = \beta^i \tilde{R}_{t+1}$. Motivated by this simple relationship, we estimate the following specification,

$$children_{i,2013} = \beta_0 + \beta_1 patience_{i,2008} + \mathbf{X}'_i \boldsymbol{\beta} + \varepsilon_i \tag{A1}$$

where $children_{i,2013}$ is the unique number of children of person i over the period 2008–13, $patience_{i,2008}$ is the self-reported patience in 2008, and $\mathbf{X}$ is a vector of control variables including age, log of net income, as well as dummy variables for gender and marital status.

Table A3 column 1 reports our most parsimonious regression specification, where we restrict the sample to those of child-rearing age (18-40). We can see a statistically strong positive correlation between the patience of an individual and the number of children they have. Columns 2 to 4 include observations of all ages. Column 2 includes a control for age, column 3 adds the log of net income and column 4 adds dummy variables for whether an observation is male, head of the household, married, widowed, divorced or separated. Our preferred specification, in Column 5, reports results with all controls for only those observations aged 18-40. In each of

Table A3: Patience and Children

VARIABLES	(1) totalChildren	(2) totalChildren	(3) totalChildren	(4) totalChildren	(5) totalChildren
HHpatience	0.027**	0.013***	0.017***	0.012***	0.022***
	(0.010)	(0.004)	(0.004)	(0.004)	(0.009)
HHage		-0.024***	-0.021***	-0.030***	0.017***
		(0.001)	(0.001)	(0.001)	(0.005)
lincome			0.414***	0.274***	0.175***
			(0.016)	(0.017)	(0.035)
Observations	4,341	17,224	17,222	17,222	4,340
R^2	0.004	0.176	0.256	0.336	0.312
Controls	no	no	no	yes	yes
Ages	18-40	All	All	All	18-40

Note: Robust standard errors in parentheses. Standard errors are clustered at the household level. Observations are weighted according to SOEP individual person weights. lincome is the log of household post-government income. Controls are dummy variables for whether an observation is male, the household head, married, widowed, divorced or separated.

these specifications, the coefficient on patience is statistically significant and of the expected sign. Table A4 reports the results from an alternative approach to age, where we use dummy variables for age brackets instead of including age as a linear variable.

Table A4: Patience and Children: Age bins

VARIABLES	(1) totalChildren	(2) totalChildren	(3) totalChildren
HHpatience	0.010**	0.016***	0.014***
	(0.004)	(0.004)	(0.004)
mediumyoung	0.573***	0.272***	0.146***
	(0.061)	(0.062)	(0.056)
mediumold	0.884***	0.471***	0.199***
	(0.057)	(0.060)	(0.058)
old	-0.056	-0.362***	-0.729***
	(0.050)	(0.052)	(0.055)
lincome		0.420***	0.312***
Observations	17,224	17,222	17,222
R^2	0.181	0.259	0.317
Controls	yes	no	yes

Note: Robust standard errors in parentheses. Standard errors are clustered at the household level. Observations are weighted according to SOEP individual person weights. lincome is the log of household post-government income. mediumyoung is a dummy equal to 1 if $25 < HHage <= 35$*; mediumold is a dummy equal to 1 if* $35 < HHage <= 45$*; and, old is a dummy equal to 1 if* $45 < HHage$*. Controls are dummy variables for whether an observation is male, the household head, married, widowed, divorced or separated.*

A.3 Cross-country considerations

The denormalized Falk et al. (2018) data points to some variation in average discount factors across countries. These differences are much smaller than the variation implied by the model across time. For example, the average patience level in the top and bottom 10 percent of countries in the denormalized Falk et al. (2018) data is $\beta_{10} = 0.934$ versus $\beta_{90} = 0.949$ (on an annualized basis). Using equation (16) with $g_{Nt} = g_{Dt} = 1$ this implies a difference of approximately 1.7 percentage points in interest rates between the most and the least patient countries $(1/\beta_{10} - 1/\beta_{90})$. Whilst this is not negligible, compared to the average decline of the interest rate of between 7 and 14 percentage points over the 700-year time span under consideration this difference is relatively unsubstantial. As such throughout the paper we abstract from variation in patience levels across countries and we focus only on variation in individual level patience that gives rise to the large downward trend in interest rates over time.

A.4 Steady state consumption share

Data on final consumption expenditures in US dollars (NE.CON.TOTL.CD) and GDP at market prices in US dollars (NY.GDP.MKTP.CD) comes from the World Development Indicators. To match the s^c_{ss} term in the main body of the text, we proceed as follows. We first calculate the ratio of global consumption to global GDP in every year and then calculate the average of world consumption shares for the years 2000-2018 which comes to 75%.

A.5 Calibrating the beta distribution

The annualized variance of generational discount factors We proceed in two steps to calculate a global variance for individual discount rates. A natural source would be the Global Preference Survey described in Falk et al. (2018). This cannot be used directly, however, as its data is normalized (each preference variable has a zero global mean and unit standard deviation). The GPS data is also based on responses to survey questions that are each focused on distinct preference characteristics. This is problematic given the evidence in Andersen et al. and other work that the *joint*-elicitation of time and risk preferences matters for measures of patience. Andersen et al. (2008) report the standard error of their estimate for the discount *rate*, r. Since

Table A5: Annual Rates of Return, un-weighted.

Asset	N	Mean	Median	Std	p90/p10
Equities	2520	0.064	0.056	0.206	0.464
Bonds	2520	0.009	0.006	0.125	0.169
Treasuries	2520	0.016	0.012	0.129	0.248

Table A6: Generational Rates of Return (Annualized), un-weighted.

Asset	N	Mean	Median	Std	p90/p10
Equities	1930	0.049	0.051	0.038	0.094
Bonds	1930	0.001	0.011	0.043	0.092
Treasuries	1930	0.004	0.010	0.054	0.119

$\beta = \frac{1}{1+r}$ in equilibrium, we need to express var $\left(\frac{1}{1+r}\right)$ as a function of the mean $\mathrm{E}(r)$ and variance var(r). We use a first-order Taylor expansion of the second moment of the transformed variable to find var $\left(\frac{1}{1+r}\right) = \frac{1}{(1+\mathrm{E}(r))^4}\mathrm{var}_t(r)$. Thus we use the time preference evidence in Andersen et al. to 'de-normalize' the Falk et al. data by fixing the GPS variation across individuals in Denmark to that found in the experiments. We then obtain a measure of the global variation across individuals, having taken account of region-specific fixed effects. We find the median standard deviation across countries is 0.0053.

The long run interest rate To find data on the long run interest rates we use the Credit Suisse Global Investment Returns Yearbook (Dimson et al., 2002). This publication provides cumulative real returns from 1900 to 2015 for equities, bonds and treasury bills for 23 major economies that cover 98% of the world equity market in 1900 and 92% at the end of 2015. Furthermore, the yearbook provides an "all-country world equity index denominated in a common currency, in which each of the 23 countries is weighted by its starting-year equity market capitalization. (It) also compute(s) a similar world bond (and treasury) index, weighted by GDP."

For each country (c), year (t) and asset class (s), we are given a cumulative real return, $R^s_{c,t}$. We then use this to calculate both the annual rate of return ($r^s_{c,t}$) and the annualized 25-year generational rate of return ($\bar{r}^s_{c,t}$) as:

$$r^s_{c,t+1} = \left(\frac{R^s_{c,t+1}}{R^s_{c,t}}\right) - 1, \tag{A2}$$

and

$$\bar{r}^s_{c,t+25} = \left(\frac{R^s_{c,t+25}}{R^s_{c,t}} \right)^{\frac{1}{25}} - 1. \tag{A3}$$

Tables A5 and A6 show summary statistics for both the annualized and generational rates of return. Notice that as usual returns are highest for equities. For annual data, it is also true that the variation in returns is much higher in equities than in either bonds or treasuries. Generational return on equities however (these are the annualized rates of return from making and holding an investment for 25 years) still offer higher average rates of return than bonds or treasuries, but are no longer as volatile - the variation in generational equity returns is either smaller or indistinguishable from variation in returns on treasuries or bonds. This motivates why we choose to calibrate our model to average, generational returns on equities - dynastic planners have a long time horizon and rates of returns of equities over this horizon are higher than of bonds or treasuries - and their variation is no higher.

The rate of return used in the calibration of the main body of the paper is obtained as follows. We calculate the (weighted) generational rate of returns of the world equity index, $\bar{r}^s_{W,t}$, in every year and then find the average of the implied rates of return between 2000 and 2015 which is equal to an annualized 6.4%.

A.6 The price of service-sector goods

We construct a relative price index of the service sector with respect to the consumer price index (CPI) for the UK. To do this we construct constant and current prices sectoral value added measures for the service (S) sector. Taking a ratio of these gives us a price index for the S sector relative to the CPI. Then, dividing the S price index by the CPI results in the required data series. We proceed as follows.

1. We calculate (constant price) service sector value added shares:

   (a) Obtain 1949-2009 constant (2005) price sector value added for the S sector from Timmer et al. (2015).

   (b) From Smits et al. (2009) obtain the 1855-1965 sector size index normalized to 1965=1.

   (c) Using quantity data from (a) transform the value added indices in (b) into constant price value added of S.

(d) Combine the two series using data from Timmer et al. (2015) for 1949-2009 and data from Smits et al. (2009) for 1855-1948 to obtain constant (2005) price sectoral value added for S for 1855-2009.

(e) Use this to calculate (constant price) sectoral value added shares.

2. We calculate current price sectoral value added shares:

   (a) Use current price value added data from Timmer et al. (2015) to calculate current price value added shares for S for 1960-2009

   (b) Combine this data with current price value added shares from Buera and Kaboski (2012) from 1800 to 1959, interpolating missing values. (They in turn obtain this data from Mitchell). This gives current price sectoral value added shares from 1800 to 2009.

3. From Thomas and Dimsdale (2017) obtain current price and constant price GDP for 1700-2016 (tabs A8, A9). Re-base the constant price GDP measure to the year 2005. Next, multiply constant price sectoral shares from 1. by this constant price GDP to obtain sectoral value added in constant 2005 prices. Next, multiply current price sectoral shares from 2. by this current price GDP measure to obtain current price sectoral value added. We do this, following Duarte and Restuccia (2010), to remain consistent between the current and constant price sectoral value added series.

4. Divide the S sector's current price value added by its corresponding constant price value added series from point 3 above to obtain a sectoral price index for S for 1855-2009.

5. Obtain the sectoral price index for S for 1270-1855 from Thomas and Dimsdale (2017) in tabs A6 and A7. Use the implied growth rates from this series to extend the 1855-2009 sectoral price index data from 4. above backwards to 1270.

6. Obtain a Consumer Price Index from Thomas and Dimsdale (2017) (Tab A47, preferred measure). Then, dividing the service sector price from the previous point by this consumer prices index allows us to calculate the relative price of service to consumption goods from 1270 to 2009. This is the data that we use in the main body of our paper.

B Equation (1): derivation and discussion

In the main text we posited an expression for the real interest rate as a function of growth and the discount rate:

$$r_t = g_t - \ln \beta.$$

In more general terms, the real interest rate on an asset L takes the form,

$$\tilde{r}_t^L = \gamma g_t - \frac{\gamma^2}{2}\sigma_t^2 - \ln \beta + \gamma d_{L,t}. \tag{A4}$$

where γ is the relative risk aversion coefficient, σ^2 is the variance of consumption growth, $d_{L,t}$ is related to the covariance between the consumption growth and the return on asset L. While this is a standard expression, below we present its derivation for completeness. We also discuss the evidence on these other parameters and the role they play in driving declining interest rates.

B.1 Derivation

Consider a household that maximizes the present value of a flow utility by choice of a portfolio of assets comprised of the risky asset, L and risk-free bonds, B,

$$\max_{L_t, B_t} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t U(C_t) \tag{A5}$$

subject to,

$$L_{t+1} + B_{t+1} = R_t^L L_t + R_t^f B_t + W_t - C_t \tag{A6}$$

where R_t^L and R_t^f are gross returns on risky assets and bonds, respectively, and where W_t is an income endowment each period. R_t^f is known at period $t-1$; only the probability distribution of R_t^L is known at period $t-1$.

Optimal portfolio choices satisfy,

$$R_{t+1}^f \mathbb{E}_t \frac{\beta U'(C_{t+1})}{U'(C_t)} = 1, \tag{A7}$$

$$\mathbb{E}_t R_{t+1}^L \frac{\beta U'(C_{t+1})}{U'(C_t)} = 1. \tag{A8}$$

To obtain an expression in certainty-equivalent form, we make two assumptions.

First, we impose CRRA utility of the form,

$$U(C_t) = \frac{1}{1-\gamma} C_t^{1-\gamma}, \tag{A9}$$

and so the optimal portfolio satisfies,

$$R_{t+1}^f \mathbb{E}_t \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} = 1, \tag{A10}$$

$$\mathbb{E}_t R_{t+1}^L \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} = 1. \tag{A11}$$

Second, let $r_{t+1}^L = \ln R_{t+1}^L$ and $g_{t+1} = \ln(C_{t+1}) - \ln(C_t)$ and assume that these are jointly Normally distributed,

$$\begin{bmatrix} g_{t+1} \\ r_{t+1}^L \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{g}_{t+1} \\ \bar{r}_{t+1}^L \end{bmatrix}, \begin{bmatrix} \sigma_{g,t}^2, \sigma_{g,L,t}^2 \\ \sigma_{g,L,t}^2, \sigma_{L,t}^2 \end{bmatrix} \right). \tag{A12}$$

where $\bar{x}_t$ is the mean of x, $\sigma_{x,t}^2$ is the variance of x, and $\sigma_{x,y,t}^2$ is the covariance of x and y at time t.

Given these assumptions, we can re-write the first order conditions as,

$$\beta \exp\left\{ r_{t+1}^f - \gamma \bar{g}_{t+1} + \frac{1}{2} \text{var}_t \left(-\gamma g_{t+1} \right) \right\} = 1 \tag{A13}$$

$$\beta \exp\left\{ \bar{r}_{t+1}^L - \gamma \bar{g}_{t+1} + \frac{1}{2} \text{var}_t \left(r_{t+1}^L - \gamma g_{t+1} \right) \right\} = 1. \tag{A14}$$

Note that from (A13) we have the following expression for the real rate,

$$r_t^f = \gamma g_t - \frac{\gamma^2}{2} \sigma_{g,t}^2 - \ln \beta. \tag{A15}$$

where with log utility ($\gamma \to 1$) and no consumption growth variance ($\sigma_{g,t}^2 = 0$), we have the expression for the real rate given above as equation (1).

The two first order conditions together give a relationship between the risk-free rate and the return on L,

$$\bar{r}_{t+1}^L + \frac{1}{2} \sigma_{L,t+1}^2 = r_{t+1}^f + \gamma \sigma_{g,L,t+1}^2 \tag{A16}$$

Note that $\bar{r}^L_{t+1} = \mathbb{E}_t r^L_{t+1}$ and, since r^L_t is Normally distributed, we can write $\ln \mathbb{E}_t R^L_{t+1} = \bar{r}^L_{t+1} + \frac{1}{2}\sigma^2_L$ and so,

$$\ln \mathbb{E}_{t-1} R^L_t = r^f_t + \gamma \sigma^2_{g,L,t} \tag{A17}$$

which, with $\tilde{r}^L_t = \ln \mathbb{E}_{t-1} R^L_t$ and $d_{L,t} = \sigma^2_{g,L,t}$, is the expression given in equation (A4).

B.2 Discussion

As we discussed in the paper, and as we develop in the extended versions of the model, the historical record for per capita growth and life expectancy are unable to explain the fall in rates over time. Equation (A4) suggests a number of additional potential channels.

Variance of consumption growth If the variance of consumption growth ($\sigma^2_{g,t}$) increased over time, this could explain a fall in real rates. However, shocks to consumption, assets and production have either remained stable or declined over time. Climate variability has been relatively constant over the last millennium, at least up until the 20th century (Salinger, 2005). Levels of violence and warfare have systematically declined (Pinker, 2012). Moreover, the emergence of sophisticated insurance markets have improved the resilience of agents to shocks (Bernstein, 1998). Each of these changes lead to lower, not higher, variance in consumption growth. Broadberry and Wallis (2017) provides direct evidence of the consequence. Using cross-country data for the later 19th century, and long-run historical data for a number of European countries, Broadberry and Wallis shows that sustained increases in growth are the result of fewer episodes of negative growth, rather than more episodes of positive growth.

Risk aversion Note that the relationship between relative risk aversion (γ) and the risk-free rate depends, by (A4), on the sign of $(\bar{g}_t - \gamma\sigma^2)$. Maddison (2013) data suggests that the country-level average annual variance in per capita incomes since 1800 are at least one order of magnitude less than the average level of annual growth. So a fall in risk aversion may explain a portion of the decline in rates. In the same way as the level of patience is not normally time-varying, the deep risk aversion parameters are usually considered fixed over time. There is evidence that risk aversion is intergenerationally transmitted, but the direction of the effect on fertility is not clear and so there is no clear route in the manner of a Barro-Becker fertility model of the sort introduced in the paper. However, we can see the required direction of any potential societal shift: the evidence on risk aversion is that it has, if anything, emerged and grown over time as an evolutionary adaptation (Robson, 1996; Levy, 2015). This would make the decline in the real interest rate harder to explain.

Declining risk We might see a decline in interest rates if our data are historical returns on assets that become steadily closer to being risk-free over time. This would manifest itself through a decline in d_t and hence falling interest rates.[28] There are a number of reasons for thinking this is not the case, however. First, a key contribution of Schmelzing (2020) is in constructing a dataset of the global risk-w rate by careful study of financial history, taking into account the shifts in stable global financial systems. Thus the series is constructed from the rates of returns on sovereign debt in 14th century Genoa, 18th century UK and 20th century US. Clark (2010), in contrast, uses data for one country and calculates returns on the safest assets within a single country. Second, Clark (2010) makes the case for England that the risk of expropriation of land was very stable in the long run and did not change significantly over this period. For Clark (p.44), "The medieval land market offered investors a practically guaranteed ... real rate of return with almost no risk."

C General model derivation

In this section we derive the solution to and calibrate the general model. As we will see below, this general model nests that developed in section 3.1. Setting $\omega = 1$ will give us the model of section 3.1 with perfect altruism; and setting $\omega < 1$ introduces a form of imperfect altruism as introduced in section 3.2. In simulations we set $\omega = 0$,

[28]Importantly a falling d_t is *not* caused by declining idiosyncratic risk. When we speak of the declining risk of an asset we are not referring to returns becoming less volatile over time, but rather returns on the risky asset become less (positively) correlated with consumption growth. Risk that is uncorrelated with consumption growth rates will generate no premium on returns - and changes in this type of risk will not result in changes in the interest rate. So, for example, if the probability of expropriation of an asset declines over time - this would *not* be reflected in declining interest rates. Instead, we would need to observe a decline in expropriation probability in 'bad' times i.e. when a negative shock hits consumption growth.

capturing the possibility that agents care about children only while they themselves are alive.

Time zero household problem By iterative substitution, we re-write the individual household problem (17) in the framework of a time zero household of each type. The time-zero household solves:

$$\max_{\{c_t^i, n_{c,t}^i, x_t^i\}_{t=0}^\infty} \sum_{t=0}^{\infty} (\beta^i)^t \left(\prod_{j=0}^{t} (\pi_j(1-\omega)+\omega) \right) \left(\alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i)\right) \tag{A18}$$

s.t.

$$c_t^i + q_t n_{c,t}^i + x_t^i \leq w_t + r_t k_t^i, \quad n_{t+1}^i = \pi + n_{c,t}^i, \quad k_{t+1}^i = \frac{(1-\delta)k_t^i + x_t^i}{n_{t+1}^i}.$$

As noted in the text, the altruism component, common across all agents and captured in ω and π_j, is distinct from the pure time preference β^i, which varies by dynasty.

Dynastic planner problem We can re-write the time zero household problem from the perspective of a single dynastic planner for each type of dynasty i. Dynasty-aggregate values are $C_t^i \equiv c_t^i N_t^i$, $N_{c,t}^i \equiv n_{c,t}^i N_t^i$, $K_t^i \equiv k_t^i N_t^i$, $X_t^i \equiv X_t^i N_t^i$. The number of households in dynasty i at time $t+1$ will be given by $N_{t+1}^i = (\pi_t + n_{c,t}^i) N_t^i = n_{t+1}^i N_t^i$. The problem for dynasty i is then given by:

$$\max_{\{C_t^i, N_{c,t}^i, X_t^i\}_{t=0}^\infty} \sum_{t=0}^{\infty} (\beta^i)^t \theta(t,\omega) \left(\alpha \log(C_t^i) + (1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))\log(N_{t+1}^i)\right) \tag{A19}$$

s.t.

$$C_t^i + q_t N_{c,t}^i + X_t^i \leq w_t N_t^i + r_t K_t^i$$

$$N_{t+1}^i = \pi_t N_t^i + N_{c,t}^i$$

$$K_{t+1}^i = (1-\delta)K_t^i + X_t^i,$$

where in the above $\theta(t,\omega) \equiv \prod_{j=0}^{t}(\pi_j(1-\omega)+\omega)$.

Firms The representative firm hires workers (N_t) and capital (K_t) to produce final output (Y_t). The profit maximization problem of the firm is given by:

$$\max_{\{K_t,N_t\}} Y_t - w_t N_t - r_t K_t, \tag{A20}$$

where $Y_t = D_t K_t^{\nu} N_t^{1-\nu}$ is a standard Cobb-Douglas production function where $0 < \nu < 1$ is the output elasticity of capital. D_t is the exogenous and time-varying level of technology.

Market clearing The market clearing conditions are given by:

$$\sum_{i=1}^{I} C_t^i = C_t \ , \ \sum_{i=1}^{I} N_t^i = N_t \ , \ \sum_{i=1}^{I} N_{c,t}^i = N_{c,t} \ , \ \sum_{i=1}^{I} K_t^i = K_t,$$

$$C_t + q_t N_{c,t} + X_t = D_t K_t^{\nu} N_t^{1-\nu}. \tag{A21}$$

Notice that capital is now produced from output and that producing a child costs an exogenous q_t units of output.

Competitive equilibrium A competitive equilibrium, given a series of child prices $\{q_t\}_{t=0}^{\infty}$ and technology $\{D_t\}_{t=0}^{\infty}$, parameter values and initial conditions $\{N_0^1, \dots, N_0^I, K_0^1, \dots K_0^I\}$, consists of allocations $\{C_t^i, N_{c,t}^i, N_{t+1}^i, K_{t+1}^i, X_t^i\}_{t=0}^{\infty}$ for each dynasty $i = 1, \dots, I$ and prices $\{w_t, r_t\}_{t=0}^{\infty}$ such that firms' and dynasties' maximization problems are solved, and all markets clear.

C.1 Solution

To solve the model, we start by deriving the first order conditions of the dynastic planner and the firms. For given parameter values, initial population and capital distributions, the competitive equilibrium of the problem, for each dynasty $i = 1, \dots, I$, is characterized by consumer first-order conditions with respect to choice of children and consumption as:

$$\frac{(1 - \alpha - \beta^i(\pi_{t+1}(1-\omega) + \omega))}{N_{t+1}^i} + (\pi_{t+1} q_{t+1} + w_{t+1}) \frac{\alpha \beta^i (\pi_{t+1}(1-\omega) + \omega)}{C_{t+1}^i} = q_t \frac{\alpha}{C_t^i}, \tag{A22}$$

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i(\pi_{t+1}(1-\omega)+\omega)(1-\delta+r_{t+1}), \tag{A23}$$

with consumer budget constraints for each dynasty i:

$$C_t^i + q_t N_{t+1}^i + K_{t+1}^i = (w_t + \pi_t q_t) N_t^i + (1-\delta+r_t) K_t^i. \tag{A24}$$

The firm first-order conditions are:

$$w_t = (1-\nu) D_t K_t^{\nu} N_t^{-\nu} \text{ and } r_t = \nu D_t K_t^{\nu-1} N_t^{1-\nu}. \tag{A25}$$

The market clearing conditions are:

$$\sum_{i=1}^{I} C_t^i = C_t \ , \ \sum_{i=1}^{I} N_t^i = N_t \ , \ \sum_{i=1}^{I} K_t^i = K_t,$$

$$C_t + q_t(N_{t+1}^i - \pi_t N_t^i) + (K_{t+1}^i - (1-\delta)K_t^i) = D_t K_t^{\nu} N_t^{1-\nu}. \tag{A26}$$

Finally, there are two transversality conditions per dynasty:

$$\lim_{t\to\infty} (\beta^i)^t u'(C_t^i) K_{t+1}^i = 0, \quad \lim_{t\to\infty} (\beta^i)^t u'(C_t^i) N_{t+1}^i = 0, \tag{A27}$$

where, $u(C_t^i) = \log(C_t^i)$ is the period utility of consumption.

Dynamics Since π_t can be time-varying, we define a mortality-adjusted population measure as $\tilde{N}_t^i \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i$, where π_{ss} is the long-run survival probability, i.e. $\pi_{ss} = \lim_{t\to\infty} \pi_t$. Making this adjustment allows us to derive an Euler equation for the evolution of the adjusted population that depends only on interest rates which, as we show below, allows for a simple solution to the problem via aggregation. Specifically, from equations (A22) and (A23), we can obtain two Euler equations that describe the evolution of dynasty consumption and adjusted population:

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i \bar{R}_{t+1}, t \geq 0, \tag{A28}$$

$$\frac{\tilde{N}_{t+1}^i}{\tilde{N}_t^i} = \beta^i \hat{R}_{t+1}, t \geq 1. \tag{A29}$$

where $R_{t+1} \equiv (1-\delta+r_{t+1})$, $\bar{R}_{t+1} \equiv R_{t+1}(\pi_{t+1}(1-\omega)+\omega)$ and $\hat{R}_{t+1} \equiv R_{t+1}(\pi_t(1-\omega)+\omega)\frac{q_{t-1}R_t - q_t\pi_t - w_t}{q_t R_{t+1} - q_{t+1}\pi_{t+1} - w_{t+1}}$. Note that when $\omega = 1$ we are back to the baseline model and $\tilde{N}_t^i = N_t^i$.

Since the interest rates are common across dynasties, we can obtain expressions relating the *relative* evolution of total consumption and adjusted population for any two dynasties, $\{i, j\}$ which is true for all $t \geq 0$ for the first expression and for $t \geq 1$ for the second expression:

$$\frac{C_{t+1}^i}{C_t^i} = \frac{\beta^i}{\beta^j}\frac{C_{t+1}^j}{C_t^j}, \quad \text{and,} \quad \frac{\tilde{N}_{t+1}^i}{\tilde{N}_t^i} = \frac{\beta^i}{\beta^j}\frac{\tilde{N}_{t+1}^j}{\tilde{N}_t^j}. \tag{A30}$$

Using repeated substitution, together with market clearing conditions, we can obtain the shares of consumption and adjusted population of each dynasty relative to economy-wide aggregate consumption and adjusted population, respectively, as a function of the initial distribution of dynasty-specific consumption and adjusted population:

$$\frac{C_t^i}{C_t} = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I (\beta^j)^t C_0^j}, \text{ and,} \quad \frac{\tilde{N}_{t+1}^i}{\tilde{N}_{t+1}} = \frac{(\beta^i)^t \tilde{N}_1^i}{\sum_{j=1}^I (\beta^j)^t \tilde{N}_1^j}, \tag{A31}$$

for $t \geq 0$ where $\tilde{N}_t \equiv \sum_{i=1}^I \tilde{N}_t^i$. Note that given the initial distributions, the evolution of a particular dynasty's adjusted population and consumption shares depends only on that dynasty's patience relative to the patience of other dynasties. In particular, recalling that dynasty I is that with the highest patience, the above expressions imply that as $t \to \infty$, so $\frac{\tilde{N}_{t+1}^I}{\tilde{N}_{t+1}} \to 1$ and $\frac{C_{t+1}^I}{C_{t+1}} \to 1$ whilst, for all $i < I$, $\frac{\tilde{N}_{t+1}^i}{\tilde{N}_{t+1}} \to 0$ and $\frac{C_{t+1}^i}{C_{t+1}} \to 0$. Given the above, it is easy to show that the same relationship also holds for the un-adjusted population. From the definition of $\tilde{N}_t^i$ and $\tilde{N}_t$ we can show that:

$$\frac{N_t^i}{N_t} = \frac{\frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}\frac{\tilde{N}_t^i}{\tilde{N}_t}}{\sum_{i=1}^I \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}\frac{\tilde{N}_t^i}{\tilde{N}_t}}. \tag{A32}$$

Together with the previous results this implies that as $t \to \infty$, so $\frac{N_{t+1}^I}{N_{t+1}} \to 1$ whilst, for all $i < I$, $\frac{N_{t+1}^i}{N_{t+1}} \to 0$. This means that the consumption and population of the most patient type will dominate the economy over time. As $t \to \infty$ the model collapses to standard homogenous-agent model with discount factor β^I and a standard Barro-Becker steady state (with the appropriate de-trending, discussed below).

Detrending Since our model exhibits equilibrium growth both in output per worker and in population, in order to solve it we need to first de-trend all variables (that is, to write them in units per effective worker). We define de-trended variables as follows: $\tilde{k}_t^i \equiv \frac{K_t^i}{D_t^{\frac{1}{1-\nu}} N_t}$, $\tilde{k}_t \equiv \frac{\sum_{i=1}^I K_t^i}{D_t^{\frac{1}{1-\nu}} N_t}$, $\tilde{c}_t^i \equiv \frac{C_t^i}{D_t^{\frac{1}{1-\nu}} N_t}$, $\tilde{c}_t \equiv \frac{\sum_{i=1}^I C_t^i}{D_t^{\frac{1}{1-\nu}} N_t}$, $\tilde{w}_t \equiv \frac{w_t}{D_t^{\frac{1}{1-\nu}}}$. We also let $\eta_t^i \equiv \frac{N_t^i}{N_t}$ and denote gross growth rates of aggregate population and as TFP $g_{Nt+1} \equiv \frac{N_{t+1}}{N_t}$ and $g_{Dt+1} \equiv \frac{D_{t+1}}{D_t}$. We then proceed to re-write the first order conditions of the model in terms of the above variables. The de-trended first order conditions and budget constraint of the dynastic planner are:

$$\alpha \frac{\tilde{c}_{t+1}^i}{\tilde{c}_t^i} g_{Nt+1} a_t = \alpha \beta^i (\pi_{t+1}(1-\omega)+\omega) \left(\tilde{w}_{t+1} + \pi_{t+1} a_{t+1}\right) + (1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega)) \frac{\tilde{c}_{t+1}^i}{\eta_{t+1}^i}. \tag{A33}$$

$$\frac{\tilde{c}_{t+1}^i}{\tilde{c}_t^i} g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}} = \beta^i (\pi_{t+1}(1-\omega)+\omega) \left(1-\delta+r_{t+1}\right). \tag{A34}$$

$$\tilde{c}_t^i = \left(\tilde{w}_t + \pi_t a_t\right) \eta_t^i + \left(1-\delta+r_t\right) \tilde{k}_t^i - g_{Nt+1} \left(a_t \eta_{t+1}^i + \tilde{k}_{t+1}^i g_{Dt+1}^{\frac{1}{1-\nu}}\right). \tag{A35}$$

The de-trended first order conditions of the firm are:

$$r_t = \nu \tilde{k}_t^{\nu-1} \text{ and } \tilde{w}_t = (1-\nu) \tilde{k}_t^{1-\nu}. \tag{A36}$$

Finally, the market clearing conditions in terms of de-trended variables are:

$$\sum_{i=1}^I \tilde{c}_t^i = \tilde{c}_t \ , \ \sum_{i=1}^I \eta_t^i = 1 \ , \ \sum_{i=1}^I \tilde{k}_t^i = \tilde{k}_t,$$

$$\tilde{c}_t + a_t(g_{Nt+1} - \pi_t) + (\tilde{k}_{t+1} g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}} - (1-\delta)\tilde{k}_t) = \tilde{k}_t^{\nu}. \tag{A37}$$

Steady state If we assume that the child-cost parameter, the survival probability and TFP growth rates converge to constants (i.e. $a_t \to a_{ss}$, $\pi_t \to \pi_{ss}$ and $g_{Dt} \to g_{Dss}$) we can solve for the steady-state levels of the de-trended model. Denoting steady state values as $\tilde{k}_{ss}$, etc. we have:

$$g_{Nss}^I = g_{Nss} \text{ and } g_{Nss}^i = 0 \ \ \forall i < I \tag{A38}$$

$$\tilde{k}_{ss}^I = \tilde{k}_{ss} \text{ and } \tilde{k}_{ss}^i = 0 \ \ \forall i < I \tag{A39}$$

$$\tilde{c}^I_{ss} = \tilde{c}_{ss} \text{ and } \tilde{c}^i_{ss} = 0 \quad \forall i < I. \tag{A40}$$

Using the above along with the de-trended dynasty first order conditions and budget constraints (A33)-(A35), and the firm's de-trended first order conditions (A36) we obtain a set of 6 equations that characterize the steady state and can be solved for six unknowns, $g_{Nss}, \tilde{k}_{ss}, \tilde{w}_{ss}, r_{ss}, \tilde{y}_{ss}$ and $\tilde{c}_{ss}$:

$$\alpha g_{Nss} a_{ss} = \alpha \beta^I (\pi_{ss}(1-\omega)+\omega)\left(\tilde{w}_{ss} + \pi_{ss} a_{ss}\right) + (1-\alpha-\beta^I(\pi_{ss}(1-\omega)+\omega))\tilde{c}_{ss}. \tag{A41}$$

$$g_{Nss} g_{Dss}^{\frac{1}{1-\nu}} = \beta^I (\pi_{ss}(1-\omega)+\omega)\left(1-\delta+r_{ss}\right). \tag{A42}$$

$$\tilde{c}_{ss} + a_{ss}(g_{Nss} - \pi_{ss}) + (\tilde{k}_{ss} g_{Nss} g_{Dss}^{\frac{1}{1-\nu}} - (1-\delta)\tilde{k}_{ss}) = \tilde{k}^{\nu}_{ss} \tag{A43}$$

$$\tilde{y}_{ss} = \tilde{k}^{\nu}_{ss} \tag{A44}$$

$$\tilde{w}_{ss} = (1-\nu)\tilde{k}^{1-\nu}_{ss} \tag{A45}$$

$$r_{ss} = \nu \tilde{k}^{\nu-1}_{ss}. \tag{A46}$$

Note that the above steady state is identical to the steady state which would arise in an economy populated by only one dynasty with discount factor β^I.

Initial population and consumption A useful relationship for solving the model is that between the relative dynastic consumption in period zero and relative dynastic population size in period 1. We can take the equations in (A31) and (A32) and use them to replace $\tilde{c}^i_t$, $\tilde{c}^i_{t+1}$ and η^i_{t+1} in equation (A33). Then, taking the limit of the resulting expression as $t \to \infty$ and using the steady-state first order conditions (A41)-(A46) we obtain the following relationship for each i:

$$\frac{C^i_0}{C^I_0} = \frac{1-\alpha-\beta^I(\pi_1(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_1(1-\omega)+\omega)} \frac{N^i_1}{N^I_1}. \tag{A47}$$

Aggregation It is convenient to solve the model in two stages: first, by deriving aggregate variables and, second, by calculating dynasty-specific variables. We start by re-writing the first order condition (A33) and (A34) for dynasty I in terms of aggregate population growth rates, g_{Nt}, and de-trended capital, $\tilde{k}_t$, only. To do this, we use equations (A31) and (A47) as well as the definitions of the de-trended variables above to relate dynasty- and aggregate-level de-trended variables via weighted averages of time zero dynasty-level de-trended consumption:

$$\tilde{c}_t^i = \frac{(\beta^i)^t \tilde{c}_0^i}{\sum_{j=1}^I (\beta^j)^t \tilde{c}_0^j} \tilde{c}_t, \text{ and,} \quad \eta_{t+1}^i = \frac{(\beta^i)^t (1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))\tilde{c}_0^i}{\sum_{j=1}^I (\beta^j)^t (1-\alpha-\beta^j(\pi_{t+1}(1-\omega)+\omega))\tilde{c}_0^j}. \tag{A48}$$

Substituting (A36), (A37) and (A48) into (A33) and (A34), all evaluated with $i = I$, gives us two first order condition in terms of aggregate population growth rates, $\{g_{Nt}\}_{t=1}^{\infty}$, de-trended capital $\{\tilde{k}_t\}_{t=0}^{\infty}$ and the initial de-trended consumption distributions, $\{\tilde{c}_0^i\}_{i=1}^I$, only. Assuming that the model converges to its steady state after T periods, we use a reverse-shooting algorithm to solve for $\{g_{Nt}\}_{t=1}^T$ and $\{\tilde{k}_t\}_{t=1}^{\infty}$ as a function of $\{\tilde{c}_0^i\}_{i=1}^I$. Given this, we can then use (A48), the firm first order condition (A36) and market clearing condition (A37) to solve for $\{\tilde{c}_t^i, \eta_{t+1}^i, \tilde{c}_t, \tilde{w}_t, r_t\}_{t=0}^T$ as functions of $\{\tilde{c}_0^i\}_{i=1}^I$.

Given the above and the assumption that the model converges to steady-state after T periods,[29] we can use the dynasty specific budget constraints to derive sequences of each dynasty's capital stock, $\{\tilde{k}_t^i\}_{t=1}^T$, as functions of $\{\tilde{c}_0^i\}_{i=1}^I$:

$$\tilde{k}_t^i = \frac{\tilde{c}_t^i + g_{Nt+1}\left(a_t \eta_{t+1}^i + \tilde{k}_{t+1}^i g_{Dt+1}^{\frac{1}{1-\nu}}\right) - (\tilde{w}_t + \pi_t a_t)\,\eta_t^i}{(1-\delta+r_t)} \tag{A49}$$

Finally, given a distribution of period zero capital across dynasties, evaluating (A35) at $t = 0$, enables us to infer the dynasty distribution of initial consumption:

$$\tilde{c}_0^i = (1-\delta+r_0)\,\tilde{k}_0^i - g_{N1}\left(a_0 \eta_1^i + \tilde{k}_1^i g_{D1}^{\frac{1}{1-\nu}}\right) + (\tilde{w}_0 + \pi_0 a_0)\,\eta_0^i. \tag{A50}$$

We can thus solve the problem for any initial distribution of capital and population.

D Calibration detail

The following section provides further details on the calibration of the model. It is written-up in terms of the general model in which ω can take any value. We also present the detail on the data sources used in the calibration.

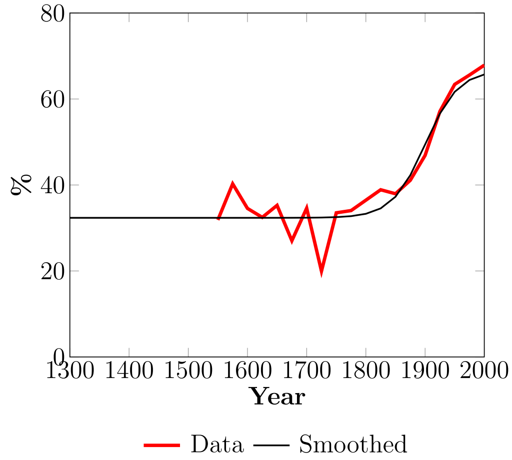
[29] So that $\tilde{k}_{T+1}^I = \tilde{k}_{ss}$ and $\tilde{k}_{T+1}^i = 0$ for all $i \neq I$.

Productivity Productivity in the model, D_t, is chosen to match the average compound generational growth rates of world GDP per capita during three time periods that exhibited markedly different growth patterns – 1275 to 1775, 1775 to 1875, and 1875 to 2000. Specifically, we calculate the average generational GDP per capita growth rates during each of those periods (1.06%, 9.75% and 44.85% respectively) using data from The Maddison Project (2013). We then estimate a generalised logistic function using three corresponding productivity growth rates in the model (-1.57%, 8.71% and 32.59% in the $\omega = 1$ model and -1.29%, 6.92% and 27.97% in the $\omega = 0$ model) such that when the model is fed in this implied productivity growth path it generates the observed GDP per capita growth over the three time periods.[30] Finally, the model also requires choosing a long-run productivity growth rate, g_{Dss}. As Crafts and Mills (2017) argue, predicting future TFP growth rates from past data can be difficult. Nonetheless, both they and an extensive literature have shown consistently declining productivity growth rates that may stay low into the future. They estimate that productivity growth between 2005 and 2016 was approximately 0.5% per year. As such we set our $g_{Dss} = 1.005^{25}$ and assume that the productivity growth rate in the model drops to this level after 2100.

Survival probabilities We calculate survival probabilities in the model by using data on life expectancy for England and the UK for the period 1543-2020 from Roser et al. (2013) who in turn compile data from Riley (2005), Zijdeman and Ribeira da Silva (2015) and the UN (UN, 2019). We use English data as this offers the longest time span available. Wrigley et al. (1997) Table 6.21 finds life expectancy from birth in England fluctuates between 37 and 40 with no clear trend up to 1800. We linearly interpolate this data and smooth it using the Hodrick Prescott filter with smoothing parameter of 100. Then, assuming that one generation is 25 years, we calculate the generational expected probability of death in the model as $\Pi_t = 1 - 25/l_t$ - where l_t is life expectancy for generations from 1550 to 2000 at 25 year intervals. Finally, we fit a generalized logistic function to this data in order to generate a smooth transition in

[30] The productivity logistic function is given by: $g(t) \equiv A + \frac{K-A}{1+e^{-B(\frac{t-1275}{25}-M)}}$, where in the model with $\omega = 0$, $A \equiv 0.98$ is the minimum asymptote, $K \equiv 1.33$ is the maximum asymptote whilst $B = 0.91$ and $M = 22.63$ are the fitted values. In the model with $\omega = 1$ the corresponding values are $A \equiv 0.99$, $K \equiv 1.28$, $B = 1.01$ and $M = 22.76$.

Figure A2: Generational Survival Probability, England.



life expectancies.[31] We assume a long run survival probability of $\pi_{ss} = 0.667$ which implies a life expectancy of 75 years. Also notice that $\pi_{ss} > \pi_t$ – a fact that we use in our calibration of the patience grid below. We extrapolate this logistic function back to 1300 giving us survival probabilities from 1300 to 2000. The survival probability implied by the data is presented in Figure A2, along with the smoothed series we use to calibrate the model (Figure 2 panel C)).

Child cost In both the model with $\omega = 0$ and with $\omega = 1$, we generate a demographic transition by choosing child costs, a_t, in such a way as to exactly replicate the evolution of population growth. Specifically, generational population growth rates are chosen to match world population growth rates during three periods that exhibited markedly different growth patterns – 1275-1775, 1775-1875, and 1875 to 2000. We calculate the average generational population growth rate during each of those periods (3.06%, 11.87% and 35.64% respectively) using data from The Maddison Project (2013). We then smooth the transition between these growth rates using a generalized logistic function. Finally, we choose period-by-period a_t in order to replicate the observed (smoothed) growth rates of population. The model also requires choosing a long-run population growth rate which enables us to pin down long-run child cost, a_{ss}, in the next paragraph below. Herrington (2021) updates the Club of Rome's

[31] This logistic function is given by: $g(t) \equiv A + \frac{K-A}{1+e^{-B(\frac{t-1550}{25}-M)}}$, where $A \equiv 0.32$ is the minimum asymptote chosen to match the average probabilities in the first 250 years, $K \equiv 2/3$ is the maximum asymptote chosen to match a long-run life expectancy of 75 years whilst $B = 0.89$ and $M = 14.01$ are the fitted values.

1972 'Limits To Growth' report and estimates that global population between 2020 and 2100 may fall by 1% per year. As such we set the long run growth rate to be $g_{Nss} = 0.99^{25}$ and assume that population growth rates in the model drop to this level after 2100.

Patience grid We assign a discount factor to each dynasty $i \in I$. Recall that we order dynasties such that the sequence $\{\beta^i\}_{i=1}^I$ is strictly increasing in i. In order to ensure the strict concavity of the objective function we need to assume that $1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega) > 0$. Given this restriction as well as the assumptions that $\lim_{t\to\infty}\pi_t \to \pi_{ss}$ and that $\pi_{ss} \geq \pi_t$, each discount factor can bounded by $0 < \beta^i < \bar{\beta}$, where $\bar{\beta} \equiv \frac{1-\alpha}{\pi_{ss}(1-\omega)+\omega}$. We sub-divide the interval $(0, \bar{\beta})$ into I equally-sized sub-intervals and locate each type's patience level at the central point of every sub-interval, so that, for each i, $\beta^i = \bar{\beta}\frac{(2i-1)}{2I}$. We set the number of types to be $I = 10,000$ in order to obtain a good approximation of continuous distribution. To pin down the sequence of β^is, we need to find values for α and $\bar{\beta}$, β^I, $\tilde{k}_{ss}$ and a_{ss}. We can solve for these five unknowns by noting first that the share of expenditure on consumption relative to aggregate income in the steady-state, $s^c_{ss} \equiv \lim_{t\to\infty} C_t/Y_t$, is an implicit function of the above values and given by:

$$s^c_{ss} \equiv \frac{\tilde{k}^{\nu}_{ss} - (g_{Nss} g_{Dss}^{\frac{1}{1-\nu}} \tilde{k}_{ss} - (1-\delta)\tilde{k}_{ss} - a_{ss}(g_{Nss} - \pi_{ss})}{\tilde{k}^{\nu}_{ss}}. \tag{A51}$$

We assume that capital depreciates 10% annually, so that $\delta = 1-(1-0.1)^{25} = 0.928$. Then, combining this equation with the first two equations of the steady state first order conditions (A41) and (A42) as well as the expression $\beta^I = \bar{\beta}\frac{(2I-1)}{2I}$ and $\bar{\beta} \equiv \frac{1-\alpha}{\pi_{ss}(1-\omega)+\omega}$ and setting $s^c_{ss} = 0.75$ to match the average global steady-state income share post-2000,[32] allows us to solve for these unknowns: $\alpha = 0.488$, $\bar{\beta} = 0.512$, $\beta^I = 0.512$, $a_{ss} = 0.345$ and $\tilde{k}_{ss} = 0.082$ in the baseline model and $\alpha = 0.488$, $\bar{\beta} = 0.769$, $\beta^I = 0.769$, $a_{ss} = 0.345$ and $\tilde{k}_{ss} = 0.082$ in the model with $\omega = 0$.

Capital distribution We choose total (de-trended) capital so that initial (de-trended) capital stocks pre-1300 lie on a saddle path. To do this we proceed by starting the model in 1275 (instead of 1300) with a guess of the initial capital stock. Since capital stocks adjusts very quickly (one period in the model is 25 years) capital

[32]See Appendix A for details.

stock in 1300 (i.e. $t = 0$) will be on the saddle path.[33] The initial distribution of capital across dynasties determines the population distribution of those dynasties in *subsequent* periods. To obtain the initial capital distribution, we assume that the growth of each dynasty's population is *consistent* with solutions of the model in the period prior to the initial period. Specifically, we assume that outcomes in the period before $t = 0$ are on the equilibrium saddlepath just as much as they are in periods from $t = 0$ on. The initial distribution of capital is thus chosen such that population growth rates are solutions of the model from period $t = 0$. In practice, this means assuming that equation (A29) also holds for $t = 0$. This assumption then allows us to re-write equation (A47) as:

$$\frac{C_0^i}{C_0^I} = \frac{\beta^i}{\beta^I}\left(\frac{1-\alpha-\beta^I(\pi_0(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_0(1-\omega)+\omega)}\right)\frac{N_0^i}{N_0^I}. \tag{A52}$$

or in de-trended terms as:

$$\frac{\tilde{c}_0^i}{\tilde{c}_0^I} = \frac{\beta^i}{\beta^I}\left(\frac{1-\alpha-\beta^I(\pi_0(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_0(1-\omega)+\omega)}\right)\frac{\eta_0^i}{\eta_0^I}. \tag{A53}$$

Given this consistency assumption, the initial population distribution tells us what initial consumption distribution should be. Finally, we can use dynastic budget constraints (A50) to determine what this initial distribution of consumption implies about the initial distribution of (de-trended) capital, $\{\tilde{k}_0^i\}_{i=1}^I$.

Patience distribution As in the the model with $\omega = 1$ we do not have data on the population distribution of patience in the year 1300 ($t = 0$ in the model), we choose our period-zero distribution of types so that the model replicates evidence on the distribution of types in the year 2000 ($t = 28$ in the model). Under the assumption of consistency we made in the previous paragraph we obtain the following expression relating the relative distribution of dynasties in time t with respect to their distribution in time zero:

$$\frac{\eta_t^i}{\eta_t^I} = \frac{(\beta^i)^t(1-\alpha-\beta^i(\pi_t(1-\omega)+\omega))(1-\alpha-\beta^I(\pi_0(1-\omega)+\omega))}{(\beta^I)^t(1-\alpha-\beta^I(\pi_t(1-\omega)+\omega))(1-\alpha-\beta^i(\pi_0(1-\omega)+\omega))}\frac{\eta_0^i}{\eta_0^I}. \tag{A54}$$

[33]In practice this means setting $\tilde{k}_{-1} = \tilde{k}_0$ - but then only considering results from 1300.

Thus, given evidence on the distribution of patience at some later date t, we can infer the initial distribution of the population across levels of patience. However, modern data will capture only a censored portion of the full initial distribution of preference types. To address this issue, we once more impose a distribution on the data. Unlike the baseline model with $\omega = 1$ however, we will impose a distribution of generational discount factors on the *mortality-adjusted* population, $\tilde{N}_t^i \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i$. Similarly to the baseline we prove the following theorem in the extended model which shows that if in any one period the distribution of patience levels in the mortality-adjusted population takes the form of a scaled beta distribution, $\tilde{f}_t(\beta; \gamma_t, \delta_t)$, then for a fine enough grid, it will follow a scaled-beta distribution in the mortality-adjusted population in all other periods with shape parameters given by $\gamma_{t+1} = \gamma_t + 1$ and $\delta_{t+1} = \delta_t$.

Theorem 2. *If $I \to \infty$ and dynastic discount factors within the mortality-adjusted population are distributed according to a scaled beta distribution on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}}$ and $\delta_{\bar{t}}$ for some period $\bar{t}$, then dynastic discount factors will also be distributed according to a scaled beta distribution in the mortality-adjusted population in period $\bar{t}+1$ on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}+1} = \gamma_{\bar{t}} + 1$ and $\delta_{\bar{t}+1} = \delta_{\bar{t}}$.*

Proof. See Appendix E. □

Notice that when $\omega = 1$ the mortality-adjusted and the unadjusted populations are identical. We can relate the distribution of patience within the adjusted population to the distribution of patience in the non-adjusted population for any $0 \leq \omega \leq 1$ using the following theorem.

Theorem 3. *If $I \to \infty$ and dynastic discount factors are distributed according to $\tilde{f}(\beta)$ within the mortality-adjusted population, then dynastic discount factors will be distributed according to the following distribution in the un-adjusted population:* $f_t(\beta) = \frac{\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}}{E_{\tilde{f}_t}\left(\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}\right)} \tilde{f}_t(\beta)$.

Proof. See Appendix E. □

Furthermore, recalling that in our calibration we have $1 - \alpha = \bar{\beta}(\omega + (1-\omega)\pi_{ss})$, using Theorem 3 we can derive an explicit expression for the distribution of patience levels in the un-adjusted population:

$$f_t(\beta) \equiv f(\beta; \gamma_t, \delta) = \frac{(1-\delta)(\bar{\beta}(\pi_{ss}(1-\omega)+\omega) - \beta(\pi_t(1-\omega)+\omega))}{(\bar{\beta}-\beta)((\pi_{ss}(1-\omega)+\omega)(1-\gamma_t-\delta_t)+\gamma_t(\pi_t(1-\omega)+\omega))}\tilde{f}(\beta; \gamma_t, \delta_t). \tag{A55}$$

As in the baseline model with $\omega = 1$ we can then calculate analytical expressions for the expected value and variance of both generational and annualized β. When $t \to \infty$, the mean generational beta converges to $\bar{\beta}$ and the variance goes to zero. We set $\text{var}_{28}(\beta^{\frac{1}{25}}) = 0.0053^2$ to match experimental evidence from representative individuals in Denmark (Andersen et al., 2008) and the individual-level data in the Global Preference Survey (GPS) described in Falk et al. (2018) and explained in greater detail in Appendix A.5. We can also derive an approximate expression (see Appendix E.3) for the annualized gross interest rate:

$$R_t^{\frac{1}{25}} \approx \left(\frac{\gamma_t - 1 + \delta_t}{\bar{\beta}\gamma_t} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{\omega + (1-\omega)\pi_t} \right)^{\frac{1}{25}}. \tag{A56}$$

Given this, we set $R_{32}^{\frac{1}{25}} - 1 = 0.064$ to match the average (annualized) generational rates of return from global equity holdings calculated in Appendix A.5. Notice that this pins down a level of the interest rate so is akin to a normalization. It in no way influences the observed decline of the interest rate. We can use this expression and the expression for model-implied variance to determine the parameters of mortality-adjusted distribution of patience: $\gamma_{28} = 32.6$ and $\delta_{28} = 58.8$ in the model with $\omega = 1$ and $\gamma_{28} = 33.6$ and $\delta_{28} = 60.4$ in the model with $\omega = 0$. Given these parameters, we can use the CDF of the distribution of the discount factors in the adjusted-population, $\tilde{F}$, to approximate, for some I, the proportion of the population assigned to each dynasty i relative to the most patient dynasty I, in the year 2000 (i.e. period $t = 28$) by:

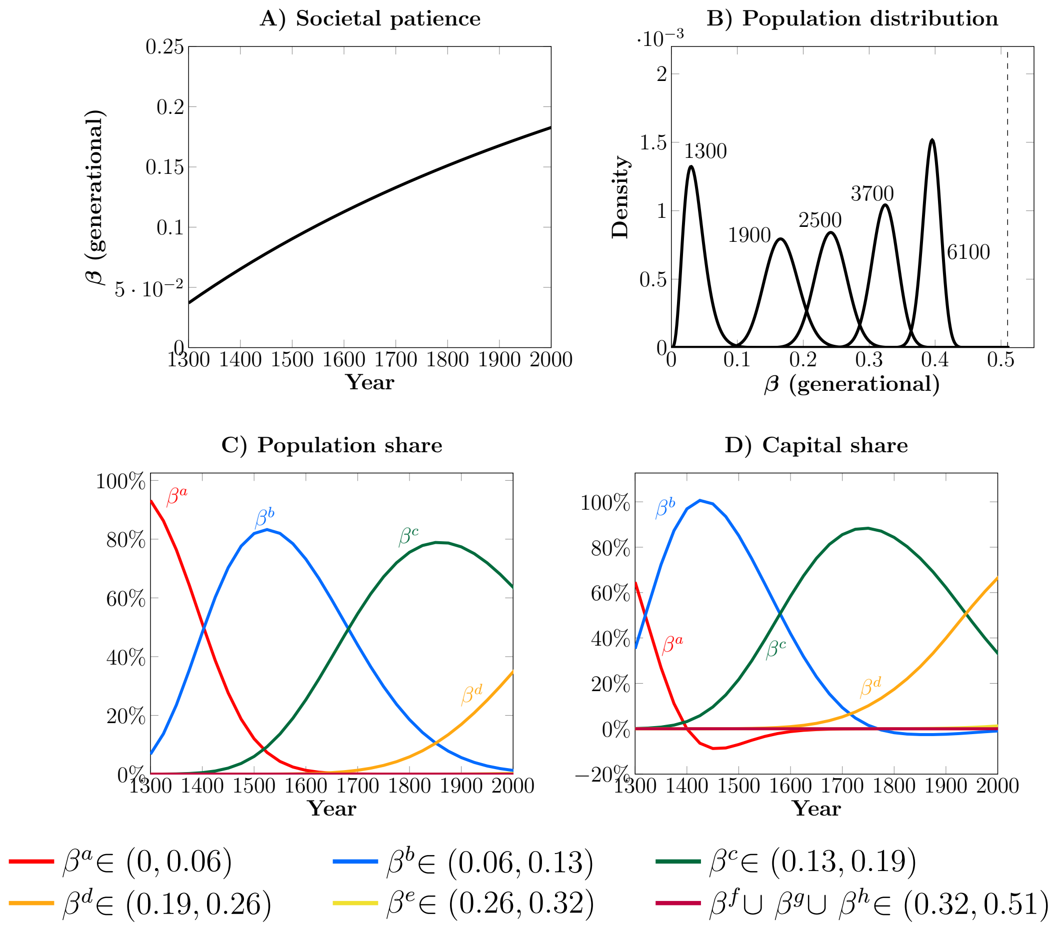
$$\frac{\eta_{28}^i}{\eta_{28}^I} = \frac{N_{28}^i}{N_{28}^I} = \frac{\frac{1-\alpha-\beta^i(\pi_{28}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}}{\frac{1-\alpha-\beta^I(\pi_{28}(1-\omega)+\omega)}{1-\alpha-\beta^I(\pi_{ss}(1-\omega)+\omega)}} \frac{\tilde{F}\left(\beta^i + \frac{\bar{\beta}}{2I}; 28\right) - \tilde{F}\left(\beta^i - \frac{\bar{\beta}}{2I}; 28\right)}{\tilde{F}\left(\beta^I + \frac{\bar{\beta}}{2I}; 28\right) - \tilde{F}\left(\beta^I - \frac{\bar{\beta}}{2I}; 28\right)}. \tag{A57}$$

With the above proportions in hand, we can then calculate the $t = 0$ distribution of population using equation (A54) with $t = 28$, and proceed to solve the model.

D.1 Additional figures

We include additional Figure A3 associated with the $\omega = 1$ model, the equivalent to Figure 3 in the main text. As is clear, nothing qualitatively changes in terms of the path of societal patience, the changing population distributions, nor the population and capital shares of each group of dynasties.

Figure A3: The rise of societal patience ($\omega = 1$)



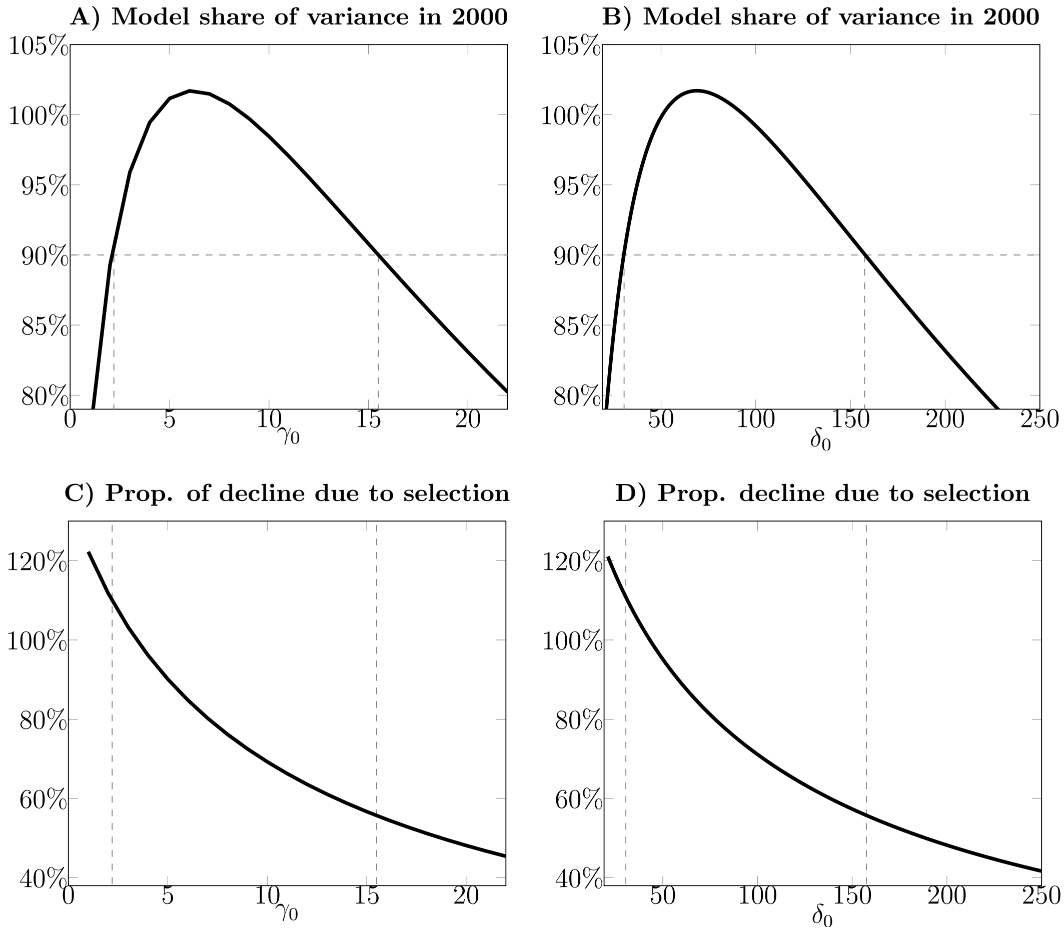
Note: These results are in the model with $\omega = 1$. Panel A) depicts the societal average level of generational patience over time. Panel B) shows the distribution of levels of patience (the generational β) at the labelled years over the period 1300-6100. The dashed vertical line is at $\bar{\beta}$. Each panel C)–D) reports the sum of the model output across all dynasties in the group of dynasties defined in the legend. The results with $\omega = 0$ are given in Figure 3.

D.2 Robustness of calibration

We examine what share of the decline in interest rates is explained by selection when we allow the variance of patience to deviate by no more than 10 percentage points from the variance in the GPS data.[34] Our robustness exercise consists of independently varying one of either δ_0 or γ_0 in the scaled-beta distribution, $\tilde{f}_t(\beta)$, whilst re-calibrating the other parameter so that the model matches the interest rate predicted by the $\omega = 0$ model in 1325 (though quantitatively and qualitatively similar results hold in the corresponding exercises with $\omega = 1$ version of the model). All other parameters in the calibration are independent of these two parameters and thus remain unchanged. For a given change in each parameter we plot two graphs: 1) the proportion of the decline in the interest rate between 1325 and 2000 that can be explained by selection; and 2) the proportion of the variance in patience observed in the year 2000 (in the GPS data) that can be explained given our parameter choice. Figure A4 shows the results of the above exercise. Panels A) and B) show the proportion of year 2000 variance in patience explained by the model when changing γ_0 and δ_0 respectively while panels C) and D) show the proportion of the decline explained by selection when changing γ_0 and δ_0 respectively. Panels A) and B) point to a hump-shaped relationship between these parameters and the success of the model distribution in capturing the variance in the data. Figures C) and D), however, show that the higher the δ_0 and the higher the γ_0 the smaller the role of selection. Recall that in the calibration of the model with $\omega = 0$ we take $\gamma_{28} = 33.6$ (so that $\gamma_0 = 5.6$) and $\delta_{28} = 60.4$ (so that $\delta_0 = 60.4$). Figure A4 points to the substantial robustness in the model to alternative calibration. The model distribution captures more than 90% of the variance in the data so long as γ_0 is approximately within $(2, 16)$ and δ_0 is approximately within $(30, 158)$. Despite such variation, the contribution of selection to explaining the decline in the interest rate remains strong. The predicted role of selection within both sets of bounds varies from approximately 56%-112%. Thus, for a large range of parameters determining the key object of the model – the distribution of patience across individuals in the year 2000 – the model points to an important role for selection.

[34] This is a larger than likely range for the variance; the bootstrapped 95% confidence interval for the variance is $[98.6, 101.0\%]$; the 90% confidence interval is $[98.3, 101.4\%]$.

Figure A4: Varying γ_0 and δ_0



Note: Panels A) and B) give the variance in the distribution of patience in the model at the year 2000, as a share of the corresponding variance in the data. Panels C) and D) report the proportion of the decline in the interest rate over 1325–2000 that can be explained by the selection mechanism (difference between heterogenous- and homogenous-agent model). Panels A) and C) vary γ_0 holding δ_0 constant; panels B) and D) vary δ_0 holding γ_0 constant. The model is that in section 3.2 with imperfect altruism ($\omega = 0$). The dashed lines give us the range of the selection-explained decline in the interest rate that results from a deviation from the variance in the data of 10 percentage points.

E Asymptotic results

E.1 General model

The following Lemma is used in the proof of Theorem 2.

Lemma 1. *Let $\tilde{f}_t(\beta)$ be the continuous probability density function depicting the* *distribution of dynastic discount factors in the mortality-adjusted population at time t as the number of agents $I \to \infty$. The probability density function of dynastic discount factors in the mortality-adjusted population at time $t+1$ will then satisfy the following relationship:*

$$\tilde{f}_{t+1}(\beta) = \frac{\beta \tilde{f}_t(\beta)}{E_t(\beta)}. \tag{A58}$$

Proof. Suppose that there are I dynasties with discount factors, β^i, distributed evenly along a grid so that $\beta(i;I) = \bar{\beta}\frac{2i-1}{2I}$ for $i = 1, \ldots, I$. Notice that the distance between any two points is simply: $\Delta(I) \equiv \beta(i;I) - \beta(i-1;I) = \frac{\bar{\beta}}{I}$. We define the following function: $\tilde{\nu}_t(\beta(i;I)) \equiv \tilde{\nu}_t^I(\beta^i) \equiv \frac{\tilde{N}_t^i}{\tilde{N}^t}$, where $\tilde{N}_t^i \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i$ and $\tilde{N}_t \equiv \sum_{i=1}^I \tilde{N}_t^i$, which maps the discount factor of a particular dynasty to the mortality-adjusted population share of that dynasty i at time t. Notice, that we can think of this function as a probability mass function of a discrete random variable with realization, $\beta(i;I)$, on the domain $\{\bar{\beta}\frac{2i-1}{2I} | i = 1, \ldots, I\}$. We wish to characterize the evolution of the asymptotic function, $\frac{\tilde{\nu}_t(\beta(i;I))}{\Delta(I)}$, over time as $I \to \infty$ – that is as the number of dynasties or types becomes infinite.[35] For each agent i, we can re-write equation (A29) as:

$$\tilde{N}_{t+1}^i = \beta^i \tilde{R}_{t+1} \tilde{N}_t^i. \tag{A59}$$

Summing these expressions over all agents gives, $\tilde{N}_{t+1} = \tilde{R}_{t+1} \sum_{j=1}^I \beta^j \tilde{N}_t^j$, which can also be written as:

$$\tilde{N}_{t+1} = \tilde{R}_{t+1} \tilde{N}_t \sum_{j=1}^I \beta^j \tilde{\nu}_t^I(\beta^j). \tag{A60}$$

Dividing equation (A59) by equation (A60) we obtain:

$$\tilde{\nu}_{t+1}^I(\beta^i) = \frac{\beta^i \tilde{\nu}_t^I(\beta^i)}{\sum_{j=1}^I \beta^j \tilde{\nu}_t^I(\beta^j)}. \tag{A61}$$

This recursive formulation defines the evolution of the probability mass function over time. We are interested in the properties of this function as $I \to \infty$. To aid us in this investigation, notice that the cumulative distribution function of β^i in the adjusted

[35]The idea here is that although our model will be solved numerically, and thus, we will always need to construct a grid and hence choose a finite number of types, we wish to emphasize that the choice of the size of the grid will be less and less relevant as long as it is relatively large. Furthermore, later we will wish to calibrate the model at a particular point in time, and hence it will be useful to show that a form of stability for the distribution function of types exists over time. This is easier to do in a continuous setting than a discrete case.

population at time t for a grid of size I is:

$$\tilde{F}_t^I(\beta^i) \equiv \sum_{j=1}^{i} \tilde{\nu}_t^I(\beta^i). \tag{A62}$$

This also means that $\tilde{\nu}_t^I(\beta^i) = \tilde{F}_t^I(\beta^i) - \tilde{F}_t^I(\beta^{i-1}) = \tilde{P}_t^I(\beta^{i-1} < \beta \leq \beta^i)$, where the right-hand side represents the probability that the random variable β in takes on a value on the semi-closed interval $(\beta^{i-1}, \beta^i]$. Given the above, notice that (A61) can be re-written as:

$$\frac{\tilde{\nu}_{t+1}^I(\beta^i)}{\Delta^i(I)} = \frac{\beta^i \frac{\tilde{\nu}_t^I(\beta^i)}{\Delta^i(I)}}{\sum_{j=1}^{I} \beta^j \tilde{P}_t^I(\beta^{i-1} < \beta \leq \beta^i)}. \tag{A63}$$

Taking the limit of both sides of the above as $I \rightarrow \infty$ we obtain the following expression:

$$\tilde{f}_{t+1}(\beta) = \frac{\beta \tilde{f}_t(\beta)}{E_{\tilde{f},t}(\beta)}, \tag{A64}$$

where $\tilde{f}_t$ is the continuous probability density function corresponding to the discrete mass function $\tilde{\nu}_t^I$[36] and $E_{\tilde{f},t}(\beta) \equiv \int_0^1 u\tilde{f}_t(u)du = \lim_{I\rightarrow\infty} \sum_{j=1}^{I} \beta^j \tilde{P}_t^I(\beta^j < \beta \leq \beta^{j+1})$, is simply the mean of the corresponding continuous random variable. □

E.1.1 Theorem 2.

The functional equation (A58) in the above Lemma describes the evolution of the distribution of the limit function of discount factors in the mortality-adjusted population over time. It is easy to show that a time invariant solution $\tilde{f}(\beta)$ of the above does not exist. Instead, we are interested in a solution that takes the following form $\tilde{f}_t(\beta) \equiv \tilde{f}(\beta; \boldsymbol{\theta_t})$, where $\boldsymbol{\theta_t}$ is a vector of potentially time varying parameters of the distribution $\tilde{f}$. In other words, we are looking for a solution to the above that remains of a fixed type, with only its parameters changing. The following Theorem proposes one such candidate solution for the extended model.

Theorem 2. *If $I \rightarrow \infty$ and dynastic discount factors within the mortality-adjusted population are distributed according to a scaled beta distribution on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}}$ and $\delta_{\bar{t}}$ for some period $\bar{t}$, then dynastic discount factors will*

[36] To see this, notice that $\lim_{I\rightarrow\infty} \frac{\tilde{\nu}_t(\beta(i+1;I))}{\Delta(I)} = \lim_{I\rightarrow\infty} \frac{\tilde{F}_t(\beta(i+1;I)) - \tilde{F}_t(\beta(i;I))}{\beta(i+1;I) - \beta(i;I)} = \lim_{I\rightarrow\infty} \frac{\tilde{F}_t(\beta(i;I)+\Delta(I)) - \tilde{F}_t(\beta(i;I))}{\Delta(I)} = F_t'(\beta(i+1;I))$

also be distributed according to a scaled beta distribution in the mortality-adjusted population in period $\bar{t}+1$ *on* $(0, \bar{\beta})$ *with shape parameters* $\gamma_{\bar{t}+1} = \gamma_{\bar{t}}+1$ *and* $\delta_{\bar{t}+1} = \delta_{\bar{t}}$*.*

Proof. Consider the scaled beta distribution defined on $(0, \bar{\beta})$ with probability density function $\tilde{f}$ is given by:

$$\tilde{f}_t(\beta; \boldsymbol{\theta_t}) \equiv \tilde{f}(\beta; \gamma_t, \delta_t) = \frac{(\bar{\beta} - \beta)^{\delta_t - 1}\beta^{\gamma_t - 1}}{\bar{\beta}^{\delta_t + \gamma_t - 1}B(\gamma_t, \delta_t)}, \tag{A65}$$

where $B(\gamma_t, \delta_t)$ is the beta function. The mean of this distribution is given by:

$$E_{\tilde{f},t}(\beta; \gamma_t, \delta_t) = \bar{\beta}\frac{\gamma_t}{\gamma_t + \delta_t}. \tag{A66}$$

Using the result of Lemma 1 and equations (A65)-(A66), we can write the PDF of discount factors at time $t+1$ as:

$$\begin{aligned}
\tilde{f}_{t+1}(\beta; \gamma_t, \delta_t) &= \frac{\beta \tilde{f}_t(\beta)}{E_{\tilde{f},t}(\beta)} \\
&= \frac{\beta(\bar{\beta} - \beta)^{\delta_t - 1}\beta^{\gamma_t - 1}}{\bar{\beta}\frac{\gamma_t}{\gamma_t + \delta_t}\bar{\beta}^{\delta_t + \gamma_t - 1}B(\gamma_t, \delta_t)} \\
&= \tilde{f}(\beta; \gamma_{t+1}, \delta_{t+1})
\end{aligned} \tag{A67}$$

where, $\gamma_{t+1} = 1 + \gamma_t$ and $\delta_{t+1} = \delta_t$. The third equality follows from a beta function identity that $B(1+x, y) = \frac{x}{x+y}B(x, y)$. Thus, one solution to the functional equation (A58) is the beta distribution with parameters given by $\gamma_{t+1} = 1+\gamma_t$ and $\delta_{t+1} = \delta_t$. □

E.1.2 Theorem 3

The following theorem establishes a relationship between the distribution of discount factors in the mortality-adjusted and the unadjusted populations.

Theorem 3. *If* $I \to \infty$ *and dynastic discount factors are distributed according to* $\tilde{f}(\beta)$*, within the mortality-adjusted population, then dynastic discount factors will be distributed according to the following distribution in the un-adjusted population:* $f_t(\beta) = \frac{\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}}{E_{\tilde{f}_t}\left(\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}\right)}\tilde{f}_t(\beta)$*.*

Proof. Suppose that there are I dynasties with discount factors, β^i, distributed evenly along a grid so that $\beta(i; I) = \bar{\beta}\frac{2i-1}{2I}$ for $i = 1, \cdots, I$. Notice that the distance between any two points is simply: $\Delta(I) \equiv \beta(i+1;I) - \beta(i;I) = \frac{\bar{\beta}}{I}$. We define the following two functions. First, $\tilde{\nu}_t(\beta(i;I)) \equiv \tilde{\nu}_t^I(\beta^i) \equiv \frac{\tilde{N}_t^i}{\tilde{N}^t}$ where $\tilde{N}_t^i \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i$ and $\tilde{N}_t \equiv \sum_{i=1}^I \tilde{N}_t^i$ which maps the discount factor of a particular dynasty to the adjusted population of that dynasty i at time t. Second, $\nu_t(\beta(i;I)) \equiv \nu_t^I(\beta^i) \equiv \frac{N_t^i}{N^t}$, which maps the discount factor of a particular dynasty to the fraction of the total population of that dynasty i at time t. Notice, that we can think of these functions as probability mass functions of discrete random variables with realizations, $\beta(i;I)$, on the domain $\{\frac{2i-1}{2I}|i = 1, \cdots, I\}$. We wish to characterize the evolution of the asymptotic functions, $\frac{\tilde{\nu}_t(\beta(i;I))}{\Delta(I)}$ and $\frac{\nu_t(\beta(i;I))}{\Delta(I)}$, over time as $I \to \infty$ - that is as the number of dynasties or types becomes infinite. We first derive a relationship between these two distributions. Since $\tilde{\nu}_t^I(\beta^i) \equiv \frac{\tilde{N}_t^i}{\tilde{N}_t} = \frac{\frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i}{\sum_{j=1}^I \frac{1-\alpha-\beta^j(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^j(\pi_t(1-\omega)+\omega)} N_t^j}$, we can re-write this expression to obtain the total population in dynasty i as:

$$N_t^i = \left(\sum_{j=1}^I \frac{1-\alpha-\beta^j(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^j(\pi_t(1-\omega)+\omega)} N_t^j \right) \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i). \tag{A68}$$

Summing the above over all i and simplifying we obtain the following expression for total population at time t:

$$N_t = \left(\sum_{j=1}^I \frac{1-\alpha-\beta^j(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^j(\pi_t(1-\omega)+\omega)} N_t^j \right) \left(\sum_{i=1}^I \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i) \right). \tag{A69}$$

Taking the ratio of these two expressions we obtain an expression for the proportion of workers in each dynasty:

$$\nu_t^I(\beta^i) \equiv \frac{N_t^i}{N_t} = \frac{\frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i)}{\sum_{i=1}^I \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i)}. \tag{A70}$$

Dividing both sides by $\Delta^i(I)$ and taking the limit of the above as $I \to \infty$ the above becomes:

$$f_t(\beta) = \frac{\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}}{E_{\tilde{f}_t}\left(\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}\right)} \tilde{f}_t(\beta), \tag{A71}$$

where f_t and $\tilde{f}_t$ are the continuous probability density functions corresponding to the discrete mass functions ν_t^I and $\tilde{\nu}_t^I$ respectively and $E_{\tilde{f}_t}(\beta)$ is the mean of the latter corresponding continuous variable. □

E.2 Model with $\omega = 1$

The following Theorem proposes a candidate solution for the distribution of discount factors in the model of section 3.1. Note that since this version of the model is nested in the general model of section 3.2 (with $\omega = 1$), the proof follows from the proofs for the general model presented above.

Theorem 1. *Within the baseline model with $\omega = 1$, if $I \to \infty$ and dynastic discount factors within the population are distributed according to a scaled beta distribution on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}}$ and $\delta_{\bar{t}}$ for some period $\bar{t}$, then dynastic discount factors will also be distributed according to a scaled beta distribution in period $\bar{t}+1$ on $(0, \bar{\beta})$ with shape parameters $\gamma_{\bar{t}+1} = \gamma_{\bar{t}} + 1$ and $\delta_{\bar{t}+1} = \delta_{\bar{t}}$.*

Proof. The baseline model is simply a special case of the extended model with $\omega = 1$. Theorem 3 above thus implies that the distribution of discount factors in the mortality-adjusted and the un-adjusted populations are the same. That is, if $I \to \infty$, $\omega = 1$, dynastic discount factors are distributed according to the probability density function $\tilde{f}(\beta)$ within the mortality-adjusted population and the probability density function $f(\beta)$ in the un-adjusted population then according to Theorem 3 $\tilde{f}(\beta) = f(\beta)$. Given this, the proof of the current theorem follows directly from Theorem 2 which is identical to this theorem but refers to the mortality-adjusted population distribution which by the above argument is identical to the un-adjusted population in the baseline case. □

E.3 Asymptotic expression for the rate of interest

This section derives the approximate expressions for the interest rate (16) and (18) used in the main body of the paper. We show derivations for the extended model, but note that when $\omega = 1$, the extended model and the its approximations collapse to those of the baseline model. Suppose that there are I dynasties with discount factors, β^i, distributed evenly along a grid so that $\beta(i; I) = \bar{\beta}\frac{2i-1}{2I}$ for $i = 1, \cdots, I$. Notice that the distance between any two points is simply: $\Delta(I) \equiv \beta(i+1; I) - \beta(i; I) = \frac{\bar{\beta}}{I}$.

From the de-trended first order condition (A34) of the extended model, we can write:

$$R_{t+1} = \frac{\tilde{c}^i_{t+1}/\tilde{c}^i_t}{\beta^i} \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{\omega + (1-\omega)\pi_{t+1}} = \frac{\left(\frac{\tilde{\kappa}^I_{t+1}(\beta^i)/\Delta(I)}{\tilde{\kappa}^I_t(\beta^i)/\Delta(I)}\right) \frac{\tilde{c}_{t+1}}{\tilde{c}_t}}{\beta^i} \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{\omega + (1-\omega)\pi_{t+1}} \tag{A72}$$

where $\tilde{\kappa}^I_t(\beta^i) \equiv \tilde{c}^i_t/\tilde{c}_t$ and $R_{t+1} \equiv (1 - \delta + r_{t+1})$. In addition we can use equations (A48), (A53) and (A54) to write:

$$\frac{\tilde{c}^i_t}{\tilde{c}_t} = \frac{\frac{\beta^i}{1-\alpha-\beta^i(\omega+(1-\omega)\pi_{ss})} \tilde{N}^i_t}{\sum_{j=1}^{I} \frac{\beta^j}{1-\alpha-\beta^j(\omega+(1-\omega)\pi_{ss})} \tilde{N}^j_t}, \tag{A73}$$

where, as before, $\tilde{N}^i_t \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N^i_t$ is the mortality-adjusted population. Finally, equation (A73) can be written as:

$$\frac{\tilde{\kappa}^I_t(\beta^i)}{\Delta(I)} = \frac{\frac{\beta^i}{1-\alpha-\beta^i(\omega+(1-\omega)\pi_{ss})} \frac{\tilde{\nu}^I_t(\beta^i)}{\Delta(I)}}{\sum_{j=1}^{I} \frac{\beta^j}{1-\alpha-\beta^j(\omega+(1-\omega)\pi_{ss})} \tilde{\nu}^I_t(\beta^j)}, \tag{A74}$$

where, as before, we define the following function: $\tilde{\nu}_t(\beta(i;I)) \equiv \tilde{\nu}^I_t(\beta^i) \equiv \frac{\tilde{N}^i_t}{\tilde{N}^t}$ and $\tilde{N}_t \equiv \sum_{i=1}^{I} \tilde{N}^i_t$, which maps the discount factor of a particular dynasty to the mortality-adjusted population share of that dynasty i at time t. Taking the limit of both sides of the above as $I \to \infty$ we obtain the following expression:

$$\tilde{f}_{ct}(\beta) = \frac{\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})} \tilde{f}_t(\beta)}{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}, \tag{A75}$$

where $\tilde{f}_t$ and $\tilde{f}_{ct}$ are the continuous probability density function corresponding to the discrete mass functions $\tilde{\nu}^I_t$ and $\tilde{\kappa}^I_t$. Note also that using the relationship between $\tilde{f}_{t+1}(\beta)$ and $\tilde{f}_t(\beta)$ in equation (A58) as well as equation (A75) we can write the following expression:

$$\frac{\tilde{f}_{ct+1}(\beta)}{\tilde{f}_{ct}(\beta)} = \beta \frac{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}{E_{\tilde{f}_t}\left(\frac{\beta^2}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)} \tag{A76}$$

Finally, taking the limit of both sides of (A72) as $I \to \infty$ and substituting the expression from equation (A76) we obtain:

$$R_{t+1} = \frac{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}{E_{\tilde{f}_t}\left(\frac{\beta^2}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)} \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{(\omega+(1-\omega)\pi_t)} \frac{\tilde{c}_{t+1}}{\tilde{c}_t} \tag{A77}$$

Note that over time the growth rate of aggregate consumption converges to 1. In particular for high enough t the approximation $\frac{\tilde{c}_{t+1}}{\tilde{c}_t} \approx 1$ holds. Consequently, for high enough t this expression becomes:

$$R_{t+1} \approx \frac{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}{E_{\tilde{f}_t}\left(\frac{\beta^2}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)} \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{(\omega+(1-\omega)\pi_{t+1})}. \tag{A78}$$

Given our assumption that $\tilde{f}_t$ is the PDF of the scaled-beta distribution with shape parameters γ_t and δ_t we can solve for the expectation expressions in the above. Then for high enough t we can write the gross interest rate as:

$$R_{t+1} \approx \left(\frac{\gamma_t+\delta_t}{\bar{\beta}(1+\gamma_t)} \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{(\omega+(1-\omega)\pi_{t+1})}\right). \tag{A79}$$

Next, notice that the mean of beta is given by:

$$E_{f,t}(\beta) = \frac{\bar{\beta}\gamma_t}{\gamma_t+\delta_t} \frac{(1-\delta_t)\omega + (\delta_t+\gamma_t)\left(\pi_t \frac{1+\gamma_t}{\gamma_t+\delta_t} - \pi_{ss}\right)(1-\omega)}{(1-\delta_t)\omega + (\delta_t+\gamma_{t-1})\left(\pi_t \frac{1+\gamma_{t-1}}{\gamma_{t-1}+\delta_t} - \pi_{ss}\right)(1-\omega)} \approx \frac{\bar{\beta}\gamma_t}{\gamma_t+\delta_t}, \tag{A80}$$

where the final relationship holds exactly if $\omega = 1$ or approximately if either $\pi_t \approx \pi_{ss}$ or if δ_t or γ_t are large enough. Furthermore, since $\frac{1+\gamma_t}{\gamma_t+\delta_t} \approx \frac{\gamma_t}{\gamma_t+\delta_t}$ for high enough δ_t or γ_t, we can approximate the gross interest in (A79) rate as:

$$R_{t+1} \approx \frac{g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}}}{E_t(\beta)(\omega+(1-\omega)\pi_{t+1})}.$$

This expression is equation (18) in the paper and, with $\omega = 1$, it is equation (16).

F Mutation

The model introduced in section 3 showed how natural selection favored more patient dynasties and drove the observed fall in the interest rate. For simplicity, we abstracted from an important part of the evolutionary process – mutation. In biology, a mutation is "an alteration in the genetic material of a cell of a living organism that is more or less permanent and that can be transmitted to the cell's (...) descendants" (Griffiths, 2020). In our model setting, such 'mutation' is a reduced form way to consider the implications of imperfect transmission of preferences in general. Mutation is one of the fundamental forces of evolution since it helps contribute to the variability of traits within populations. As mutations occur, the process of natural selection determines which of these will thrive and which will die out by selecting the most advantageous mutations for the given environment. In this section we introduce mutation into our model and examine the role it has on the process of natural selection and the economy. Specifically, we allow for the possibility that a proportion of some dynasty exogenously, unexpectedly and permanently experiences a mutation in its discount factor from one period to the next.

Our experiment can also be interpreted without reference to genetics. What we refer to as a 'mutation' may be thought of as forms of imperfect transmission, i.e., changes in the discount factor brought about by parental decisions or peer influence through education or parental investment (i.e., different forms of imitation and socialization) that can also be horizontal or oblique channels of transference (Jablonka and Lamb, 2005). They could also be interpreted as immigration, invasion or colonization, where a small number of outsiders arrive with different discount factors that differ from those of the existing population.[37] As such, this section can also be interpreted as examining the effects of a new variant of dynasty no-matter its source.

Setup We consider the version of the model with imperfect altruism ($\omega = 0$). For simplicity we assume that productivity growth and survival probability are at their constant, steady-state values throughout. We model a mutation as an unanticipated shock to an agent's discount factor. Instead of attempting to match the rate at

[37] In this case, the comparison is not exact, as migration would additionally increase the size of the population while in our mutations the population remains fixed. Since only a very small number of agents are assumed to mutate, the results are quantitatively and qualitatively almost indistinguishable from a migration story.

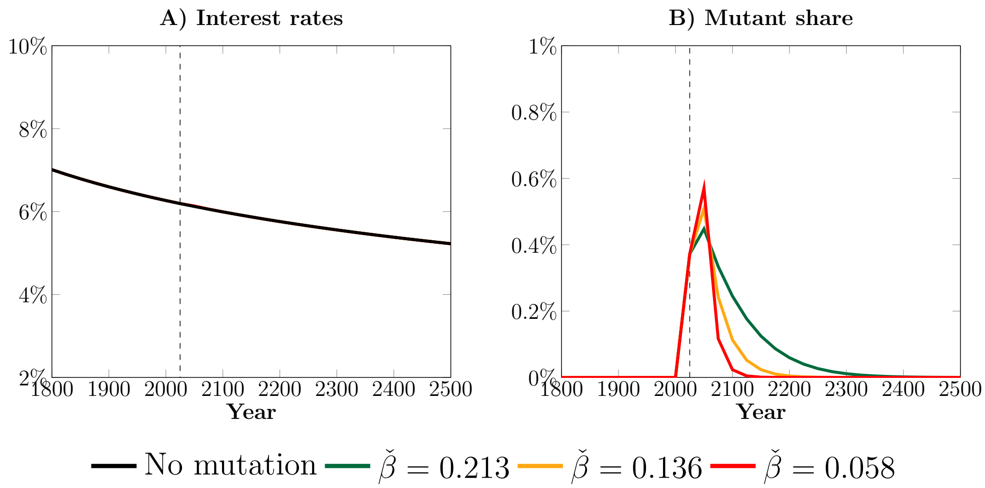
which mutations occur in nature (something which would be difficult to calibrate) we instead consider the consequences of different types of one-off mutations. We assume mutations occur at only one point in time. Each mutation counterfactual involves an unexpected but permanent change in discount factor for 1% of agents belonging to the dynasty with that period's median discount factor. These mutants then form a new dynasty, retaining their net capital per capita from the previous period.[38] Mutations can be divided into two categories based on the impact they have on an agent's 'fitness' or reproductive success: deleterious and advantageous mutations. We introduce different mutations in the year 2025; the median dynasty in that period has $\beta = 0.291$ and 1% of that dynasty comprises around 0.4% of the aggregate population.

Deleterious mutations First, agents from the median dynasty can mutate to lower levels of patience. In the biological literature these types of mutations are known as 'deleterious' since the mutants have lower fitness than before: agents mutating to a lower level of patience will have fewer children over their lifetime than agents from that same dynasty who did not mutate. The aggregate effects of these deleterious mutations are short-lived and quantitatively small. Figure A5 reports the effect on interest rates of three separate mutations of the 2025 median dynasty to three different levels of lower patience. It also shows the proportion of mutants in the population after the shock. Notice that the mutations – even that to very low levels of patience – have practically no effect on interest rates. Furthermore, selective pressure works against the low-patience mutants. Agents with lower patience will choose to have fewer children and their share will quickly diminish in the population: the lower the mutant's discount factor, the faster they will disappear.

Advantageous mutations Second, agents from the 2025 median dynasty can mutate to higher levels of patience. These mutations are knows as 'advantageous' in the biological literature as they increase the fitness of the dynasty: agents mutating to

[38]For tractability, we allow mutations only on our grid of discount factors. Thus, after mutation there will be two dynasties with the same discount factor, but potentially different capital stocks. The assumption that mutants take their capital with them is quantitatively unimportant – we could otherwise assume that mutated agents are 'shunned' by their dynasties and start life with no capital or that mutants are favoured children gifted with above average capital stocks. In both extremes the quantitative results are almost indistinguishable as agents quickly adjust their capital holdings according to their time preference.

Figure A5: Deleterious mutations in 2025

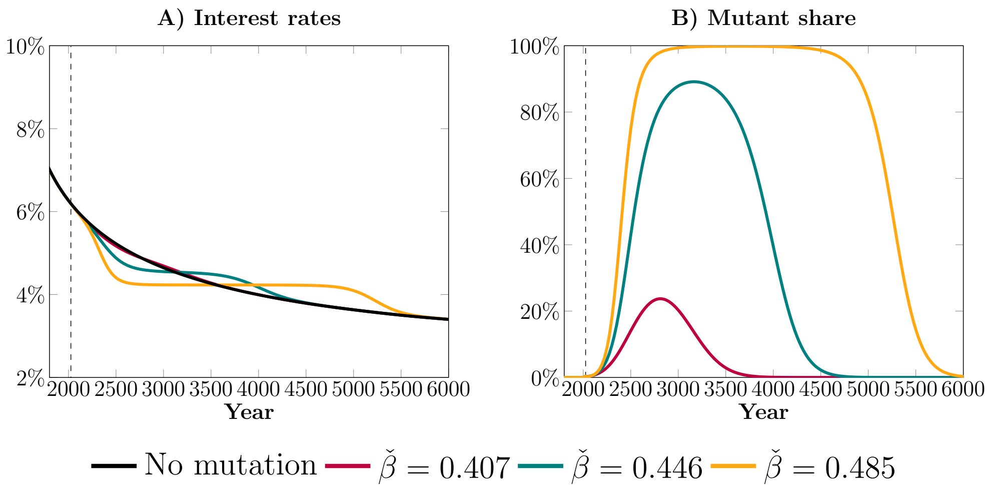


Note: Figures report the simulation output with an unexpected mutation in the year 2025 (dashed line). Each line is a different mutation counterfactual. A mutation causes 1% of the dynasty with the median level ($\beta = 0.291$) of patience in 2025 to wake up in 2025 with the level of patience $\check{\beta}$ denoted in the Figure legend. Panel A) reports the aggregate interest rate; panel B) reports the population share of the mutant as a percentage of the aggregate population.

this higher level of patience will have more children over their lifetime than agents from the same dynasty who remain un-mutated. Advantageous mutations can have large and very long-lasting effects. Figure A6 shows the effects on interest rates of a mutation to successively higher discount factors as well as the share of mutants in the population. Notice that a mutation to a discount factor that is 2/3 higher than the median – 0.485 – pushes interest rates forward in the evolutionary process by thousands of years. Since at the time of the mutation, so few agents belong to this higher patience type that a 1% mutation of the median dynasty to this higher level of patience is an enormous shock. The economy is suddenly inhabited by a relatively large proportion of the agents who are very patient. These agents quickly amass all the capital in the economy and begin to have large numbers of children which dominate the population. This process would have happened without the mutation, but it would have lasted hundreds if not thousands of years more. With mutation the process lasts less than 500 years.

Mutations to particularly high (but not the maximum) levels of patience give rise to some especially interesting dynamics. Agents mutated in this manner can come to dominate the population for significant periods of time (see for example Figure

Figure A6: Advantageous mutations in 2025



Note: Figures report the simulation output with an unexpected mutation in the year 2025 (dashed line). Each line is a different mutation counterfactual. A mutation causes 1% of the dynasty with the median level ($\beta = 0.291$) of patience in 2025 to wake up in 2025 with the level of patience $\check{\beta}$ denoted in the Figure legend. Panel A) reports the aggregate interest rate; panel B) reports the population share of the mutant as a percentage of the aggregate population.

A6), where mutants with discount factor 0.485 practically dominate the population for three thousand years or so before beginning to lose their dominance to dynasties with higher betas still. The effects of these types of mutations look initially like a shift to a new steady state where mutated agents seem to dominate the population forever and interest rates reflect that mutant dynasty's domination for many generations. However, since these are not the most patient agents in the population, their domination is not permanent and a transition eventually takes place to agents with even higher patience levels. This results in multiple oscillations of the interest rate with results first 'converging' to an intermediate steady-state-like phase and then only slowly shifting to the true steady state where the most patient agent dominates.

G Model with fixed aggregate capital

In the model of section 3 we allowed capital to be produced using retained output. Below, we present a minimal version of the model where we fix aggregate capital to $\bar{K}$, which we might consider to be something more akin to land. We set $\omega = 1$ for simplicity. Each household of type i now solves the following utility maximization problem in each period t:

$$U_t^i(k_t^i) = \max_{c_t^i, n_{c,t}^i, x_t^i} \alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i) + \beta^i U_{t+1}^i(k_{t+1}^i) \tag{A81}$$

$$\text{s.t.} \qquad c_t^i + q_t n_{c,t}^i + p_t x_t^i \leq w_t + r_t k_t^i, \qquad n_{t+1}^i = \pi_t + n_{c,t}^i, \qquad k_{t+1}^i = \frac{k_t^i + x_t^i}{n_{t+1}^i}.$$

In this version of the model, there is no capital depreciation, x_t^i is the quantity of capital to accumulate or decumulate via trading with other households, and the price of purchasing capital stock is given by p_t.

Aggregation to time-zero dynastic planners is as in the section 3 model. The firm problem and market clearing is also the same, except that now we impose,

$$\sum_{i=1}^{I} K_t^i = K_t = \bar{K},$$

As in the model with reproducible capital, the time-varying parameters D_t, π_t and q_t are not important for the extent to which selection matters for the decline of the interest rate. We thus keep D_t and π_t constant. Given this, the model does not exhibit trend growth and so we can also normalize the cost of children, q_t to one. The solution method follows as before and, as in the general model, as the number of dynasties becomes large and assuming specific distributional forms on patience we can obtain the following simple approximation for the gross interest rate:

$$R_{t+1} \approx \frac{1}{E_t(\beta)}. \tag{A82}$$

The calibration procedure is the same as before, though we have different values for some parameters (and note that ν is now a land elasticity). Calibrated parameters are reported in Table A7.

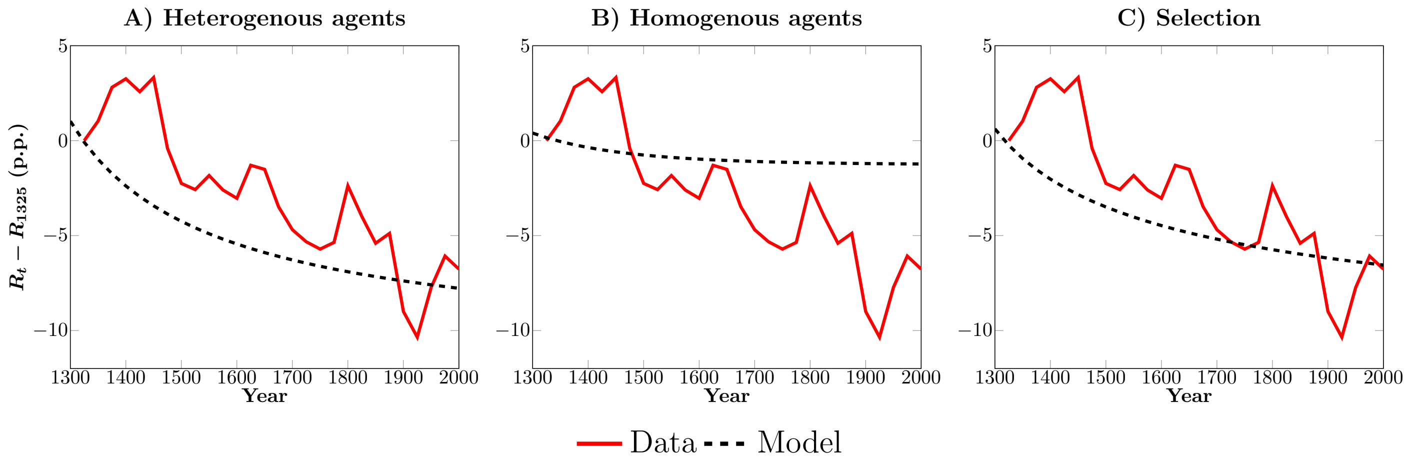
We report results for the model in Figure A7. The model with heterogenous agents can account for a nine percentage point drop in interest rates, while the decline in the homogenous agent model is less than two percentage points. The decline in the homogenous agent model occurs due to a (counterfactually) slowing growth rate that arises because of convergence dynamics. Specifically, as population increases, consumption grows but it does so at a decreasing rate as the economy approaches its steady state capital-labour ratio. By equation (1), this slowing growth rate of

Table A7: Model parameters

Parameter(s)	Value	Target/Description/Source
D_t	1	Normalization
q_t	1	Normalization
N_0	0.370	Aggregate population, 1300, The Maddison Project (2013)
$\bar{K}$	11.722	Aggregate population, 2000, The Maddison Project (2013)
ν	0.190	Land share, Caselli (2005)
π_t	0.667	Adult life expectancy of 75
I	10,000	Number of types
$\{\beta^i\}_{i=1}^{I}$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^{I}$	Subdivide domain into grid
α	0.427	Consumption share (see Appendix)
$\bar{\beta}$	0.573	Max. (generational) discount factor
$\{\gamma_{28}, \delta_{28}\}$	$\{32.089, 53.531\}$	Standard deviation of discount factors (Andersen et al., 2008; Falk et al., 2018) and long run rate of return (see Appendix)
$\left\{\frac{N_0^i}{N_0}\right\}_{i=1}^{I}$	See paper	Andersen et al. (2008) & Falk et al. (2018)
$\left\{\frac{K_0^i}{K_0}\right\}_{i=1}^{I}$	See paper	Consistency (see paper)

consumption depresses the real rate, even with a fixed β. As is clear, selection plays the same role as in the model with endogenous capital. Compared to Figure 5, the results on the contribution of selection in panel C) are similar. In the 'Selection Only' version of the model with endogenous capital (Table 2), the results capture 86% of the decline in interest rates between 1325 and 2000. In the model with fixed aggregate capital, the decline in the results is 7.69p.p., which is 114% of that observed in the data.

Figure A7: Selection and the interest rate (fixed capital)



Note: Data is the Schmelzing (2020) global real interest rate; we report the 25-yearly median interest rate beginning in 1325 (since the Schmelzing real 'Global R' series begins in 1314). We report the model interest rates in annualized terms. The data and model outputs are normalized to zero in the year 1325. We report outputs with fixed aggregate capital and with $\omega = 1$. *Panel A) reports results with heterogenous agents whose distribution of patience is calibrated at the year 2000. Results in panel B) are based on a homogeneous-agent set-up where there is only one dynasty calibrated to match the average patience in the year 2000 in the heterogenous-agent model. Panel C) is the contribution of selection, defined as the difference between the heterogenous- and homogenous-agent model results.*

References

Alan, Sule and Martin Browning, "Estimating Intertemporal Allocation Parameters using Synthetic Residual Estimation," *The Review of Economic Studies*, October 2010, *77* (4), 1231–1261.

Allen, Robert, "The Great Divergence in European Wages and Prices from the Middle Ages to the First World War," *Explorations in Economic History*, 2001, *38* (4).

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and Elisabet E. Rutström, "Eliciting Risk and Time Preferences," *Econometrica*, 2008, *76* (3), 583–618.

Arrow, Kenneth J., Maureen L. Cropper, Christian Gollier, Ben Groom, Geoffrey M. Heal, Richard G. Newell, William D. Nordhaus, Robert S. Pindyck, William A. Pizer, Paul R. Portney, Thomas Sterner, Richard S. J. Tol, and Martin L. Weitzman, "Determining Benefits and Costs for Future Generations," *Science*, 2013, *341* (6144), 349—350.

Auclert, Adrien, Hannes Malmberg, Frederic Martenet, and Matthew Rognlie, "Demographics, Wealth, and Global Imbalances in the Twenty-First Century," *Working Paper 29161, NBER, Cambridge, MA.*, 2021.

Bar, Michael and Oksana Leukhina, "Demographic transition and industrial revolution: A macroeconomic investigation," *Review of Economic Dynmics*, 2010, *13*, 424—451.

Barro, Robert J. and Gary S. Becker, "Fertility Choice in a Model of Economic Growth," *Econometrica*, March 1989, *57* (2), 481–501.

Bauer, Michal and Julie Chytilová, "Women, Children and Patience: Experimental Evidence from Indian Villages: Women, Children and Patience," *Review of Development Economics*, November 2013, *17* (4), 662–675.

Baumol, William J. and William G. Bowen, "On the Performing Arts: The Anatomy of Their Economic Problems," *The American Economic Review*, 1965, *55* (1/2).

Becker, Gary S., "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology," *Journal of Economic Literature*, 1976, *14* (3), 817–26.

_ **and Robert J. Barro**, "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics*, February 1988, *103* (1), 1–25.

Becker, Robert A., "On the Long-Run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households," *Quarterly Journal of Economics*, 1980, *95* (2), 375–82.

Bennett, Patrick, Chiara Ravetti, and Po Yin Wong, "Losing in a Boom: Long-term Consequences of a Local Economic Shock for Female Labour Market Outcomes," *Labour Economics*, 2021, *73* (102080).

Bernstein, P. L., *Against the Gods: The Remarkable Story of Risk*, Wiley, 1998.

Black, Dan A., N. Kolesnikova, S. G. Sanders, and Lowell J. Taylor, "Are Children "Normal"?," *Review of Economics and Statistics*, 2013, *95* (1), 21–33.

Blanchard, Olivier J., "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 1985, *93* (2), 223–47.

Brenøe, Anne A. and Thomas Epper, "The Intergenerational Transmission of Time Preferences Across Four Decades," *European Economic Review*, 2022, *148* (104208).

Broadberry, Stephen and John Joseph Wallis, "Growing, Shrinking, and Long Run Economic Performance: Historical Perspectives on Economic Development," Technical Report w23343, National Bureau of Economic Research, Cambridge, MA April 2017.

_ **and Roger Fouquet**, "Seven Centuries of European Economic Growth and Decline," *Journal of Economic Perspectives*, 2015, *29* (4).

Broadberry, Stephen N. and Alexandra de Pleijt, "Capital and Economic Growth in Britain, 1270-1870: Preliminary Findings," *Discussion Paper DP15889, CEPR, London*, 2021.

Buera, Francisco J. and Joseph P. Kaboski, "Scale and the Origins of Structural Change," *Journal of Economic Theory*, March 2012, *147* (2).

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio, "Demographics and Real Interest Rates: Inspecting the Mechanism," *European Economic Review*, 2016, *2016* (September), 208—226.

Cesarini, David, Erik Lindqvist, Robert Östling, and Anastasia Terskaya, "Fortunate Families? The Effects of Wealth on Marriage and Fertility," *Working Paper 31039, NBER, Cambridge, MA*, 2023.

Chowdhury, Shyamal, Matthias Sutter, and Klaus Zimmermann, "Economic Preferences across Generations and Family Clusters: A Large-Scale Experiment in a Developing Country," *Journal of Political Economy*, April 2022, (720395).

Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf, "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England," *Demography*, April 2017, *54* (2), 413–436.

_ , _ , and _ , "Further Evidence of Within-Marriage Fertility Control in Pre-Transitional England," *Demography*, August 2019, *56* (4), 1557–1572.

Clark, Gregory, "The Cost of Capital and Medieval Agricultural Technique," *Explorations in Economic History*, 1988, *25*, 265–294.

_ , *A Farewell to Alms*, Princeton University Press, Princeton and Oxford, 2007.

_ , "Genetically Capitalist? The Malthusian Era, Institutions and the Formation of Modern Preferences," *mimeo*, 2007.

_ , "The Macroeconomic Aggregates for England, 1209-1869," in A. J. Field, ed., *Research in Economic History*, Vol. 27, Emerald Group Publishing, Bingley, 2010, pp. 51—140.

_ **and Gillian Hamilton**, "Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England," *The Journal of Economic History*, 2006, *66* (3), 707–36.

_ **and Neil Cummins**, "Randomness in the Bedroom: There Is No Evidence for Fertility Control in Pre-Industrial England," *Demography*, August 2019, *56* (4), 1541–1555.

_ , _ , **and Matthew Curtis**, “Twins Support the Absence of Parity-Dependent Fertility Control in Pretransition Populations,” *Demography*, August 2020, *57* (4), 1571–1595.

Crafts, Nicholas and Terence Mills, “Trend TFP Growth in the United States: Forecasts versus Outcomes,” *Discussion Paper DP12029, CEPR, London*, 2017.

Cronqvist, Henrik and Stephan Siegel, “The Origins of Savings Behavior,” *Journal of Political Economy*, February 2015, *123* (1), 123–169.

de la Croix, David, Eric B. Schneider, and Jacob Weisdorf, “Childlessness, Celibacy and Net Fertility in Pre-Industrial England: The Middle-Class Evolutionary Advantage,” *Journal of Economic Growth*, September 2019, *24* (3), 223–256.

Deaton, Angus S. and John Muellbauer, “On Measuring Child Costs: With Applications to Poor Countries,” *Journal of Political Economy*, August 1986, *94* (4), 720–744.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti, “Global Trends in Interest Rates,” 2018. Working Paper 25039, NBER, Cambridge, MA.

Dimson, Elroy, Paul Marsh, and Mike Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.

Doepke, Matthias and Fabrizio Zilibotti, “Occupational Choice and the Spirit of Capitalism,” *Quarterly Journal of Economics*, May 2008, *123* (2), 747–793.

_ **and Michele Tertilt**, “Families in Macroeconomics,” in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, North Holland, Amsterdam, 2016, pp. 1789–1891.

_ **, Anne Hannusch, Michele Tertilt, and Fabian Kindermann**, “The Economics of Fertility: A New Era,” 2022. Working Paper 29948, NBER, Cambridge, MA.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, “The Intergenerational Transmission of Risk and Trust Attitudes,” *The Review of Economic Studies*, 2011, *79* (2), 645–77.

Donni, Olivier, "Measuring the Cost of Children," *World of Labor, IZA, Bonn*, 2015, (132).

Duarte, M. and D. Restuccia, "The Role of the Structural Transformation in Aggregate Productivity," *The Quarterly Journal of Economics*, 2010, *125(1)*, 129–173.

Epper, Thomas, Ernst Fehr, Helga Fehr-Duda, Claus Thustrup Kreiner, David Dreyer Lassen, Søren Leth-Petersen, and Gregers Nytoft Rasmussen, "Time Discounting and Wealth Inequality," *American Economic Review*, April 2020, *110* (4), 1177–1205.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 2018, *133* (4), 1645–92.

Featherstone, Allen M. and Timothy G. Baker, "An Examination of Farm Sector Real Asset Dynamics: 1910–85," *American Journal of Agricultural Economics*, 1987, *69* (3), 532–546.

Galor, Oded, "Chapter 4 From Stagnation to Growth: Unified Growth Theory," in "Handbook of Economic Growth," Vol. 1, Elsevier, Amsterdam, 2005, pp. 171–293.

_ **and David N. Weil**, "Population, technology and growth: From the Malthusian regime to the demographic transition'," *American Economic Review*, 2000, *90* (4), 806–28.

_ **and Omer Moav**, "'Natural Selection and the Origin of Economic Growth'," *Quarterly Journal of Economics*, 2002, *117* (4), 1133–91.

_ **and Özak Özak**, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064–103.

Giannelis, Alexandros, Emily A. Willoughby, Robin Corley, Christian Hopfer, John K. Hewitt, William G. Iacono, Jacob Anderson, Aldo Rustichini, Scott I. Vrieze, Matt McGue, and James J. Lee, "The Association between Saving Disposition and Financial Distress: A Genetically Informed Approach," *Journal of Economic Psychology*, June 2023, *96*, 102610.

Gokhale, Jagadeesh, Laurence J. Kotlikoff, and John Sabelhaus, "Understanding the Postwar Decline in U.S. Saving: A Cohort Analysis," *Working Paper 5571, NBER, Cambridge, MA*, 1996, p. 93.

Gollin, Douglas, "Getting Income Shares Right," *Journal of Political Economy*, April 2002, *110* (2), 458–474.

Griffiths, Anthony J. F., "Mutation," *Encyclopædia Britannica*, 2020.

Hansson, Ingemar and Charles E. Stuart, "'Malthusian Selection of Preferences'," *American Economic Review*, 1990, *80* (3), 529–44.

Herrington, Gaya, "Update to limits to growth: Comparing the World3 model with empirical data," *Journal of Industrial Ecology*, June 2021, *25* (3), 614–626.

Hudson, Michael, "How Interest Rates Were Set, 2500 BC-1000 AD: Máš, tokos and fœnus as Metaphors for Interest Accruals," *Journal of the Economic and Social History of the Orient*, 2000, *43* (2), 132–61.

Jablonka, Eva and Marion J. Lamb, *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life* Life and Mind, Cambridge, Mass: MIT Press, 2005.

Jones, Larry E. and Michèle Tertilt, "An Economic History of Fertility in the United States: 1826–1960," in Peter Rupert, ed., *Frontiers of Family Economics*, Vol. 1, Bingley: Emerald Group Publishing, Bingley, 2008, chapter 5, pp. 165–230.

Kearney, Melissa S. and Riley Wilson, "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence From The Fracking Boom," *Review of Economics and Statistics*, 2018, *100* (4), 678–90.

Kolk, Martin, "The Relationship between Life-Course Accumulated Income and Childbearing of Swedish Men and Women Born 1940–70," *Population Studies*, November 2022, pp. 1–19.

Krueger, Dirk and Alexander Ludwig, "On the Consequences of Demographic Change for Rates of Returns to Capital, and the Distribution of Wealth and Welfare," *Journal of Monetary Economics*, January 2007, *54* (1), 49–87.

Laubach, Thomas and John C. Williams, "Measuring the Natural Rate of Interest," *The Review of Economics and Statistics*, 2003, *85* (4), 1063–1070.

Levy, M., "An evolutionary explanation for risk aversion," *Journal of Economic Psychology*, 2015, *46*, 51–61.

Lovenheim, Michael F and Kevin J. Mumford, "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," *Review of Economics and Statistics*, 2013, *95* (2), 464–75.

Lucas, Robert E., *The industrial revolution: Past and future.*, Harvard University Press, Cambridge, 2002.

Machiavelli, Niccolo, *The Prince*, Dante University Press, 2003.

Manuelli, Rodolfo E. and Ananth Seshadri, "Explaining International Fertility Differences," *Quarterly Journal of Economics*, May 2009, *124* (2), 771–807.

Mian, Atif, Ludwig Straub, Amir Sufi, and Chicago Booth, "What Explains the Decline in R*? Rising Income Inequality versus Demographic Shifts," 2021, p. 47.

Millner, Antony, "Nondogmatic Social Discounting," *American Economic Review*, March 2020, *110* (3), 760–775.

Parker, Geoffrey, *The Grand Strategy of Philip II*, Yale University Press, 2000.

Pinker, S., *The Better Angels of Our Nature: A History of Violence and Humanity*, Penguin, 2012.

Ramsey, Frank P., "A mathematical theory of saving," *Economic Journal*, 1928, *38* (152), 543–59.

Riley, James C., "Estimates of Regional and Global Life Expectancy, 1800–2001," *Population and Development Review*, 2005, *31* (3), 537–543. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1728-4457.2005.00083.x.

Robson, Arthur J., "A Biological Basis for Expected and Non-expected Utility," *Journal of Economic Theory*, 1996, *68*, 397–424.

— **and Balazs Szentes**, "Evolution of Time Preference by Natural Selection: Comment," *American Economic Review*, 2008, *98* (3), 1178–88.

Rogers, Alan R., "Evolution of Time Preference by Natural Selection," *American Economic Review*, 1994, *84* (3), 460–81.

Rogoff, Kenneth S., Barbara Rossi, and Paul Schmelzing, "Long-Run Trends in Long-Maturity Real Rates 1311-2021," *Working Paper 30475, NBER, Cambridge, MA*, 2022.

Roser, Max, Esteban Ortiz-Ospina, and Hannah Ritchie, "Life Expectancy," *Our World in Data*, May 2013.

Salinger, M. J., "'Climate Variability and Change: Past, Present and Future'," *Climate Change*, 2005.

Schmelzing, Paul, "Eight centuries of global real interest rates, R-G, and the 'suprasecular' decline, 1311–2018," January 2020. No. 845, Bank of England Staff Working Paper, London.

Schoonbroodt, Alice and Michèle Tertilt, "Property Rights and Efficiency in OLG Models with Endogenous Fertility," *Journal of Economic Theory*, March 2014, *150*, 551–582.

Smits, J.P., P.J. Woltjer, and D. Ma, "A Dataset on Comparative Historical National Accounts, ca. 1870-1950: A Time-Series Perspective," *Groningen Growth and Development Centre Research Memorandum GD-107, Groningen: University of Groningen.*, 2009.

Stefanski, Radek and Alex Trew, "Replication Data for: "Selection, Patience, and the Interest Rate"," 2023. Harvard Dataverse, https://doi.org/10.7910/DVN/UGWOHH.

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim, "Patience and Comparative Development," *The Review of Economic Studies*, December 2021, *89* (5), 2806—40.

Sutch, Richard, ed., "Chapter Ce Saving, Capital, and Wealth," in Richard Sutch, ed., *Historical Statistics of the United States*, Millenial Edition On Line, Cambridge University Press, Cambridge, 2006, pp. 3–287–3–297.

Tamura, Robert, "From decay to growth: A demographic transition to economic growth," *Journal of Economic Dynamics and Control*, 1996, *20*, 1237–61.

Teulings, Coen and Richard Baldwin, eds, *Secular Stagnation: Facts, Causes and Cures*, CEPR Press, London, 2014.

T'Hart, Joost Jonker Marjolein and Jan Luiten van Zanden, eds, "A financial history of The Netherlands," 1997.

The Maddison Project, "The Maddison Project, 2013 version," 2013. http://www.ggdc.net/maddison/maddison-project/home.htms.

Thomas, Ryland and Nicholas H. Dimsdale, "A Millennium of UK Data," *OBRA Dataset, Bank of England, London*, 2017.

Timmer, M. P., G. J. de Vries, and K. de Vries, *Patterns of Structural Change in Developing Countries* 2015.

UN, "World Population Prospects 2019, Online Edition. Rev. 1.," 2019.

Vischer, Thomas, Thomas Dohmen, Armin Falk, David Huffman, Jürgen Schupp, Uwe Sunde, and Gert G. Wagner, "Validating an ultra-short survey measure of patience," *Economics Letters*, 2013, *120* (2), 142–145.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp, "The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements," *Schmollers Jahrbuch (Journal of Applied Social Science Studies)*, 2007, *127* (1), 139–69. doi:10.5684/soep.v33.

Weitzman, Martin L., "Gamma Discounting," *American Economic Review*, March 2001, *91* (1), 260–271.

Wrigley, Edward Anthony A., R. S. Davies, Jim E. Oeppen, and Roger S. Schofield, *English Population History from Family Reconstitution 1580–1837*, Cambridge University Press, Cambridge, 1997.

Zijdeman, Richard and Filipa Ribeira da Silva, "Life Expectancy at Birth (Total)," December 2015. Type: dataset.