

Selection, Patience, and the Interest Rate*

Radoslaw (Radek) Stefanski[†] Alex Trew[‡]

Original: February 8, 2020

Revised: June 14, 2021

Abstract

The interest rate has been falling for centuries. The key to explaining this decline is increasing societal patience, driven by a process of natural selection. Three observations support this mechanism: patience varies across individuals, is inter-generationally persistent, and is positively related to fertility. To establish the importance of this channel, we introduce a dynamic, heterogeneous-agent model of fertility. The structure of our model enables us to use modern, micro-level data to calibrate the historical distribution of patience. Our quantitative results match the centuries-long fall in the interest rate, highlighting the crucial role of selection in this historical, and ongoing, trend.

JEL codes: E21; E43; J11; N30; O11.

Keywords: Interest rates; selection; fertility; patience; heterogeneous agents.

*This paper was previously circulated under the title ‘Patience Breeds Interest’. We would like to thank 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 and Ludo Visschers as well as seminar participants at Cardiff, CERGE-EI, Edinburgh, New Hampshire, Oxford, St Andrews 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 finalizing this paper. The usual disclaimer applies.

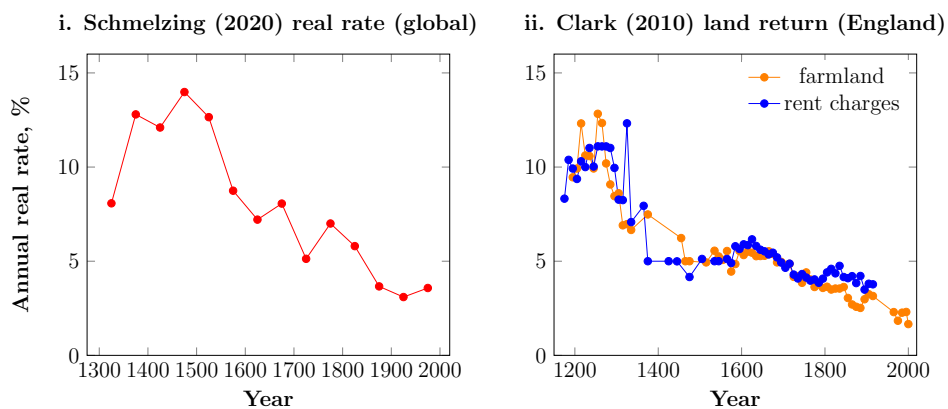
[†]School of Economics and Finance, Castlecliffe, The Scores, University of St Andrews, Fife, KY16 9AR, United Kingdom. Email address: rls7@st-andrews.ac.uk.

[‡]Adam Smith Business School, University of Glasgow, 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 12% in the fourteenth century to just over 1% 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 returns, 1175–2000



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

$$r_t = g_t - \log \beta. \quad (1)$$

The real interest rate, r_t , is the difference between the growth rate of per capita consumption, g_t and (the log of) the level of patience, β .² Since income per capita growth was close to zero up to 1800 and then increased following the onset of the industrial revolution (The Maddison Project, 2013), 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 level of an individual. However, this parameter could also reflect time-varying

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

²Of 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.

changes in *societal* levels of patience driven by changing demographics. Blanchard (1985), for example, shows that a declining probability of death can appear as an increase in effective β , thus contributing to a falling interest rate. Although the historical evidence does not support this particular channel as being a driver of the centuries-long decline in interest rates,³ it does highlight that a changing population composition could have an important impact on societal levels of patience.

In this paper, we propose a novel demographic channel that, we find, can explain the decline in interest rates. In particular, we introduce a model of endogenous fertility, in the spirit of Barro and Becker (1988), where patience levels are heterogeneous across agents. In this model, societal patience increases over time as a result of evolutionary pressures that naturally select the most patient agents. Our calibrated model explains most of the decline in the real rate.

We develop our findings in three steps. First, we motivate the model by drawing together the following empirical facts: patience varies across individuals; more patient individuals have more children; and, patience is at least partially inter-generationally persistent. Together, these facts imply that patience levels should increase over time as more patient households come to form an increasing proportion of the population. Our mechanism is thus *theoretically* plausible. Second, we construct a fertility model à la Barro and Becker (1988) that captures these empirical facts. Levels of patience are assumed to be heterogeneous across households, but are transmitted inter-generationally. Since children are a form of saving in this model, and since patient agents tend to save more than impatient agents, the model implies that patient agents will have more children. We then calibrate this model to experimental micro-level evidence on the modern distribution of patience, and use the structure of the model to find the initial (historical) distribution. Finally, we show that our calibrated model captures the timing and magnitude of the decline in real rates over eight centuries, despite this not being a target of the calibration. Our mechanism is thus also *quantitatively* plausible. The calibrated model is also consistent with the magnitude of the historical relationship between wealth at death (in bequests) and number of surviving children, lending further credence to our fertility mechanism. Crucially, without a fully-specified and calibrated model, we would be unable to arrive at the conclusion

³Life expectancy was flat until the 19th century (and even falling in England and Wales between 1550 and 1750). Only after 1750 did a ‘mortality revolution’ take place that resulted in a rapid increase in life expectancy (Wrigley et al., 1997) and potentially contributed to falling rates.

that our mechanism is quantitatively important and to measure its role in driving the declining interest rate. Indeed, a key contribution of our model is a quantitative theory of the evolution of the population distribution of patience.

Understanding the level and path of the real rate of interest over time is crucial for long-term inter-temporal decisions that are associated with long term savings and investment choices or future paths of innovation, as well as for the long-run sustainability of public debt. Furthermore, 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 rate at which the future is discounted (see Weitzman, 2001; Arrow et al., 2013; and, Millner, 2019). Small changes in assumptions can give rise to vastly different policy recommendations and potentially result in devastatingly different outcomes in the distant future. Having a clear understanding of the factors driving the path of societal patience can help us construct optimal policies today to cope with these types of problems and to provide better solutions for tomorrow.

Related literature Our paper is related to a number of different strands of the literature. First, we contribute to the role of evolution and selection in economics. Galor and Moav (2002) propose a theory in which natural selection entails an evolutionary advantage to human traits that are complementary to the escape from the Malthusian trap. In the Malthusian era, higher income leads endogenously to higher fertility. Following an endogenous demographic transition, higher incomes improve child quality (greater human capital) instead of child quantity. In our paper, fertility decisions depend on inherited traits that determine time preference. We find that human capital is not critical to explaining the decline in real rates over a period before and after sustained high growth emerges.⁴

In a related paper, Galor and Özak (2016) present a model in which higher patience leads to better economic outcomes and, consequently, greater reproductive success. Geographical variation in returns to agricultural investment mean that the returns to patience also varies. Since patience is partly inherited, and since better economic outcomes lead to more children, locations that offer greater returns to patience observe an increasing share of long-term orientated individuals over time.

⁴Further on the connection between fertility, human capital and development see also Becker et al. (1990) and Tamura (2006).

Galor and Özak present empirical evidence which shows that cross-sectional variation in measures of long-term orientation can be explained by historical differences in crop yields. Outcomes that benefit from patience, such as technological adoption, are also connected with agricultural productivity. While Galor and Özak can thus explain a portion of the *level* differences in patience around the world, our contribution is to understand the dynamics of the evolution of patience in a quantitative model that can match the data over time.

We also relate to the further literature on the biological (evolutionary) basis of stable preferences (for example, see Becker, 1976, Rogers, 1994, Robson and Szentes, 2008, and Robson et al., 2012). In this literature, choices are determined wholly by inherited traits, whilst preferences that survive in the long-run are defined as evolutionarily stable. Given that we build a calibrated, dynamic model that incorporates the shifting distribution of types, we are able to show just how long – *millenia* – it can take for such a steady state to arrive. The closest to our set-up is Hansson and Stuart (1990), in which the fitness of a dynasty (its population growth) is assumed to be monotonically increasing in the per-capita consumption of the dynasty. The steady-state is thus one dominated by dynasties that maximise per capita consumption growth. In our model, the population growth rate of a dynasty is a function of the fundamental preferences and of the environment. Agents behave optimally for given discount rates and societal preferences change over time because dynasties have different levels of fitness. 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 at the turn of the seventeenth century with more wealth tended to have more children. Moreover, records on royal tenants in England (whose wealth would have been greater than average), suggest that a positive relationship between wealth and fertility goes back at least to the mid-thirteenth century. For Clark (2007a), variation in the number of children per household comes purely from the Malthusian relationship between income and fertility. Since innate patience is more deep-rooted than wealth, and since the

⁵Ramsey (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) and Mitra and Sorger (2013).

accumulation of wealth is a direct consequence of higher patience, we view patience as the fundamental driver of differences in both dynastic wealth and household fertility. The same evolutionary pressures yield a society that is wealthier, more literate and more patient, but the mechanism is ‘survival of the patient’ rather than ‘survival of the wealthiest’.

Third, our model draws a connection between family-level decisions and their macroeconomic consequences. As such, we relate to the growing literature on family macroeconomics (see Doepke and Tertilt, 2016 or Doepke et al., 2019 for recent surveys). While our treatment of the complexities of family decision-making is simplified, our study suggests another way in which changes over time in the nature of fertility decisions can manifest themselves in significant changes to macroeconomic variables.

Fourth, our work connects to recent research on whether the decline in global real interest rates in the past few decades is a result of long-run trends or cyclical shifts (see, for example, the chapters in Teulings and Baldwin, eds, 2014). Del Negro et al. (2018) study the determinants of the interest rate using a VAR analysis of data since 1870 for advanced countries. Del Negro et al. isolate the role of growing risk and declining growth rates in explaining the decline of the last ten years, but limited role for a stochastic discount factor. As we will show, the evolution of society toward the more patient generates a decline in the interest rate that slows over time and is thus hard to discern even in data since 1870. Carvalho et al. (2021) consider the demographic channel over 1990 to 2014, finding a quantitatively significant role for growing life expectancy in explaining declining rates.

Structure In section 2 we briefly introduce empirical evidence on the distribution and transmission of levels of patience. In section 3 we develop a Barro-Becker model of fertility with heterogeneous dynasties that differ according to their discount factor, while section 4 presents the solution to the model. Section 5 calibrates the model to existing modern data on the distribution of patience and section 6 presents quantitative results, comparing various characteristics of the model to the historical record. We also present a validation exercise, showing that relationship between wealth and fertility is similar to that in the data. Section 7 considers two extensions. First, we consider a number of shocks to the population and capital (such as might arise in a pandemic, war or a revolution). Second, we explore the impact of the mutation of a

small portion of a dynasty to a higher or lower level of patience. Finally, section 8 offers some concluding remarks.

2 Evidence on patience

We suggested above that a dynamic model of societal preferences may explain the decline in the interest rate if three conditions are met: patience is heterogeneous; preferences are inter-generationally persistent; and, patience is related to fertility. We offer a brief discussion of the empirical evidence that each of these conditions are met.

First, modern empirical studies find that there is indeed significant heterogeneity of patience. Andersen et al. (2008) use experimental evidence from a representative sample of Danes to elicit time and risk preferences. Alan and Browning (2010) use structural estimation on data in the longitudinal PSID survey. Both studies find similar levels of heterogeneity in discount factors across individuals, whether or not estimating discount factors jointly with risk attitude. More recently, Falk et al. (2018) establish the substantial extent to which preferences vary both across the globe and within countries.

Second, the strong intergenerational transmission of preferences, either by genetics, imitation or by socialization, has been identified in a number of studies. Brenøe and Epper (2018) find substantial transmission of patience across generations of Danish families. Chowdhury et al. (2018) find the same based on experimental evidence in Bangladesh. Other elements of preferences are also persistent intergenerationally: Dohmen et al. (2011) show a strong connection between generations of a family of attitudes to risk and trust.⁶

Third, since altruism toward children make them appear as a ‘normal good’, the equilibrium of standard models of fertility such as Barro and Becker (1988) implies that higher levels of patience will drive higher demand for future consumption including via the consumption of future children. To our knowledge, this connection has not been investigated empirically. In Appendix A, we use German Socio-Economic Panel (SOEP) data to show that there is a robust, positive relationship between individual patience levels and the quantity of offspring. The SOEP is a longitudinal

⁶See also Barth et al. (2020) and Fagereng et al. (2021) on the mechanisms that can drive the intergenerational transmission of wealth.

dataset which collects information by interview from 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. In 2008 and 2013, the interviews included a question asking for ‘general personal patience’ on a scale of 0-10 (where 0 is very impatient and 10 is very patient). The 2008 measure has been validated using experimental methods (Vischer et al., 2013). We find a statistically strong positive correlation between the self-reported patience of an individual and the number of children they have. This holds when we control for a large number of additional variables, including age, net income, gender and household status.

The above facts imply that parents that are highly 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 and to falling interest rates. In the next section, we build a quantitative model that captures these three facts and enables us to measure the importance of this selection mechanism in accounting for the observed decline of the real interest rate.

3 Baseline model

Consider an economy with aggregate population N_t at time t . The population consists of a finite number of dynasties, indexed by $i = 1, \dots, I$. A dynasty i consists of N_t^i equally-sized households. Households within a dynasty are identical, but dynasties differ in their discount factors, 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 non-reproducible capital (or land), k_t^i , that it inherited from its parent and that it rents out in exchange for a rental rate, r_t . Each household of type i solves the following utility maximization problem

⁷Since 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 .

in each period t :

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

s.t.

$$\begin{aligned} c_t^i + n_{c,t}^i + p_t x_t^i &\leq w_t + r_t k_t^i \\ n_{t+1}^i &= \pi + n_{c,t}^i \\ k_{t+1}^i &= \frac{k_t^i + x_t^i}{n_{t+1}^i}. \end{aligned}$$

As in Barro and Becker (1988, 1989), households derive utility from their own consumption, c_t^i , from the size of the household at beginning of the next period, n_{t+1}^i (since households are altruistic), and from the next generation's average continuation utility, $U_{t+1}^i(k_{t+1}^i)$. This particular choice of utility function follows Tamura (1996), Lucas (2002) and Bar and Leukhina (2010). 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 . For simplicity, we assume that the cost of a child is the same as the cost of a unit of consumption.⁸ The price of purchasing capital stock is given by p_t . We also assume that the exogenous survival probability for existing households, π , is age independent and constant across dynasties. The survival probability of children is 1 (this can readily be generalized). Together, these imply that the expected number of people in a household at the end of the period (and the beginning of the subsequent period) will be $n_{t+1}^i = \pi + n_{c,t}^i$. We assume that parents care about their children equally and endow them each with the same share of accumulated capital. Thus, parents face a quantity-quality tradeoff with respect to the number of children à la Barro and Becker (1988, 1989). Finally, we also assume that the child of an adult in dynasty i perfectly inherits the discount factor β^i (we relax this assumption in section 7.2). 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

⁸This has no impact on our key findings regarding interest rates, but naturally does affect steady state populations in an obvious way; counterfactuals are available from the authors.

reduced form assumption.

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 as follows:

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

s.t.

$$\begin{aligned} c_t^i + n_{c,t}^i + p_t x_t^i &\leq w_t + r_t k_t^i \\ n_{t+1}^i &= \pi + n_{c,t}^i \\ k_{t+1}^i &= \frac{k_t^i + x_t^i}{n_{t+1}^i}. \end{aligned}$$

The above reflects the choice of an individual time zero adult household. Since households within a dynasty are identical, and since there are N_0^i identical members of each dynasty i at time zero, we can re-write the time zero household problem as the problem facing 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 which are assumed to survive), and the expected number of surviving adults. The number of people in dynasty i at time $t + 1$ will thus be given by $N_{t+1}^i = (\pi + n_{c,t}^i) N_t^i = n_{t+1}^i N_t^i$. 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 re-write the time-zero household problem for the dynastic planner of each type as:

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

s.t.

$$\begin{aligned} C_t^i + N_{c,t}^i + p_t X_t^i &\leq w_t N_t^i + r_t K_t^i \\ N_{t+1}^i &= \pi N_t^i + N_{c,t}^i \\ K_{t+1}^i &= K_t^i + X_t^i. \end{aligned}$$

Just as in Lucas (2002), to ensure strict concavity of the objective we need to assume that $1 - \alpha - \beta^i > 0$. 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 effectively 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.

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:

$$\max_{\{K_t, N_t\}} Y_t - w_t N_t - r_t K_t, \quad (5)$$

where $Y_t = DK_t^\nu N_t^{1-\nu}$ is a standard Cobb-Douglas production function, D is the exogenous level of technology and $0 < \nu < 1$ is the output elasticity of capital. In our setup, we think of capital as a fixed, non-reproducible and scarce quantity akin to land.

Market Clearing Finally, the market clearing conditions are standard and given by:

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

$$C_t + N_{c,t} = DK_t^\nu N_t^{1-\nu}. \quad (6)$$

One point to emphasize once more is that we assume there exists a fixed quantity of capital, \bar{K} . This is an important way of introducing scarcity into the model. Since natural selection works through adjustments in how agents respond to scarcity, this will be a crucial part of our mechanism.

Competitive Equilibrium A competitive equilibrium, for given 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 $\{w_t, r_t, p_t\}_{t=0}^\infty$ such that firms' and dynasties' maximization problems are solved, and all markets clear.

4 Solution

To solve the model, we derive the first order conditions of firms and the dynastic planner (see Appendix C). For given parameter values, initial dynasty populations $\{N_0^1, \dots, N_0^I\}$ and stocks of capital $\{K_0^1, \dots, K_0^I\}$, 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:

$$\frac{(1 - \alpha - \beta^i)}{N_{t+1}^i} + (\pi + w_{t+1}) \frac{\alpha \beta^i}{C_{t+1}^i} = \frac{\alpha}{C_t^i}, \quad (7)$$

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i \frac{p_{t+1} + r_{t+1}}{p_t}, \quad (8)$$

with consumer budget constraints for each dynasty i :

$$C_t^i + N_{t+1}^i + p_t K_{t+1}^i \leq (w_t + \pi) N_t^i + (r_t + p_t) K_t^i. \quad (9)$$

The firm first-order conditions are:

$$w_t = (1 - \nu) D K_t^\nu N_t^{-\nu}, \quad (10)$$

$$r_t = \nu D K_t^{\nu-1} N_t^{1-\nu}. \quad (11)$$

The market clearing conditions are:

$$\sum_{i=0}^I N_t^i = N_t, \quad (12)$$

$$\sum_{i=0}^I K_t^i = K_t = \bar{K}. \quad (13)$$

Finally, there are two transversality conditions per dynasty:

$$\lim_{t \rightarrow \infty} (\beta^i)^t u'(C_t^i) K_{t+1}^i = 0, \quad (14)$$

$$\lim_{t \rightarrow \infty} (\beta^i)^t u'(C_t^i) N_{t+1}^i = 0, \quad (15)$$

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 dynasty population:

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i R_{t+1}, t \geq 0, \quad (16)$$

$$\frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, t \geq 1, \quad (17)$$

where in the above $R_{t+1} \equiv \left(\frac{p_{t+1} + r_{t+1}}{p_t} \right)$ is the gross real interest rate on capital whilst $\tilde{R}_{t+1} \equiv R_{t+1} \frac{R_t - (w_t + \pi)}{R_{t+1} - (w_{t+1} + \pi)}$ is the shadow gross real interest rate on dynasty population. These two interest rates differ since children are both a consumption good and an investment good, whereas capital is only an investment good.

Given the above Euler equations, and since the interest rates are common across dynasties, we can write the following expressions relating the *relative* evolution of consumption and 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:

$$\frac{C_{t+1}^i}{C_t^i} = \frac{\beta^i C_{t+1}^j}{\beta^j C_t^j}, \quad \text{and,} \quad \frac{N_{t+1}^i}{N_t^i} = \frac{\beta^i N_{t+1}^j}{\beta^j N_t^j}. \quad (18)$$

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:

$$\frac{C_t^i}{C_t} = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I (\beta^j)^t C_0^j}, \quad \text{and,} \quad \frac{N_{t+1}^i}{N_{t+1}} = \frac{(\beta^i)^t N_1^i}{\sum_{j=1}^I (\beta^j)^t N_1^j}, \quad (19)$$

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 that with the highest patience, the above expressions imply that as $t \rightarrow \infty$, so $\frac{N_{t+1}^I}{N_{t+1}} \rightarrow 1$ and $\frac{C_{t+1}^I}{C_{t+1}} \rightarrow 1$ whilst, for all $i < I$, $\frac{N_{t+1}^i}{N_{t+1}} \rightarrow 0$ and $\frac{C_{t+1}^i}{C_{t+1}} \rightarrow 0$. This means that the consumption and population of the most patient type will dominate the economy over time (consistent with the Ramsey (1928) conjecture). As $t \rightarrow \infty$ the model collapses to standard homogeneous agent model with discount factor β^I and a

standard Barro-Becker steady state. Consequently, if we derive the steady state values (see Appendix C.1), the model can be solved with a reverse-shooting algorithm.

Aggregation It is convenient to solve the model in two stages: first, by deriving aggregate variables and, second, by calculating prices and dynasty-specific variables.

We start by re-writing the first order condition (7) for dynasty I in terms of aggregate population only. To do this, we use equations (19) and the derivations shown in Appendix C to relate dynasty- and aggregate-level variables via weighted averages of time zero dynasty-level consumption:

$$C_t^i = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I (\beta^j)^t C_0^j} C_t, \text{ and, } N_{t+1}^i = \frac{(\beta^i)^t (1 - \alpha - \beta^i) C_0^i}{\sum_{j=1}^I (\beta^j)^t (1 - \alpha - \beta^j) C_0^j} N_{t+1}. \quad (20)$$

Substituting (6), (10) and (20) into (7), all evaluated with $i = I$, gives us a first order condition in terms of aggregate population, $\{N_t\}_{t=0}^\infty$, and initial consumption distributions, $\{C_0^i\}_{i=1}^I$, only. Assuming that the model converges to its steady state after T periods, we use a reverse-shooting algorithm to solve for $\{N_t\}_{t=0}^T$, as a function of $\{C_0^i\}_{i=1}^I$. Given this, we can then use (20), market clearing condition (6) and the firm first order condition (10) to solve for $\{C_t^i, N_{t+1}^i, C_t, w_t\}_{t=0}^T$ as functions of $\{C_0^i\}_{i=1}^I$.

Next, given the above solutions, we use household I first order condition (8) and the firm first order condition (11) to derive the solutions for $\{p_t\}_{t=0}^\infty$ and $\{r_t\}_{t=0}^\infty$ as functions of $\{C_0^i\}_{i=1}^I$. Given the above and the assumption that the model converges to steady-state after T periods,⁹ we can use the dynasty specific budget constraints to derive sequences of each dynasty's capital stock, $\{K_t^i\}_{t=1}^T$, as functions of $\{C_0^i\}_{i=1}^I$:

$$K_t^i = \frac{C_t^i + N_{t+1}^i + p_t K_{t+1}^i - (w_t + \pi) N_t^i}{(r_t + p_t)}. \quad (21)$$

Finally, since we know the distribution of period zero capital across dynasties, then (21) evaluated at $t = 0$, can be used to infer the dynasty distribution of initial consumption:

$$C_0^i = (r_0 + p_0) K_0^i - N_1^i - p_0 K_1^i + (w_0 + \pi) N_0^i. \quad (22)$$

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

⁹So that $K_{T+1}^I = \bar{K}$ and $K_{T+1}^i = 0$ for all $i \neq I$.

5 Calibration

The key aims of the calibration are to replicate the increase in world population between the years 1300 and 2000, to fit the distribution of patience types using contemporaneous experimental data, and, to match the remaining technological and preference parameters to reproduce various key moments in the data.

Model parameters and their calibrated values are summarized in Table 1. We take one period in the model to be 25 years (a generation) and we assume that period zero in the model corresponds to the year 1300 in the data. We normalize the level of technology so that $D = 1$. The initial level of population is set to be $N_0 = 0.370$ corresponding to a world population of 0.37 billion in 1300 and the aggregate land supply, $\bar{K} = 11.780$, is chosen so that the model reproduces a global population of 6.08 billion at period 28 (the year 2000) in the model (The Maddison Project, 2013). The land elasticity of the production function is set to $\nu = 0.190$ to match the share of land in value added found by Caselli (2005). We assume that all children survive into adulthood (25 years) and set $\pi = 0.67$ to yield an expected lifetime of 75 years.¹⁰

We specify the number of dynasties to be $I = 2000$. This is largely a computational choice which makes little difference to our results for a large enough number of dynasties.¹¹ 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 our requirement 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 , $\beta^i = \bar{\beta} \frac{(2i-1)}{2I}$. To pin down the sequence of β^i 's, we need to find values for α and $\bar{\beta}$. We can solve for these two unknowns by noting first that the share of expenditure on consumption relative to aggregate income in the steady-state, $s_{ss}^c \equiv C_{ss}/Y_{ss}$, is a function of α , β^I , and other calibrated parameters:

$$s_{ss}^c \equiv \frac{\alpha (1 - \beta^I (1 - \nu(1 - \pi)))}{(1 - \pi(1 - \alpha)) (1 - \beta^I)}. \quad (23)$$

¹⁰This is higher than historical evidence would suggest, but since survival probability is exogenous, and since it has a consequence principally for the steady state of the model, targeting modern rather than historical life expectancy makes more sense.

¹¹If 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 in the data, we select a large number of types in the calibration.

Table 1: Model parameters

Parameter(s)	Value	Target/Description/Source
D	1	Normalization
N_0	0.370	Aggregate population, 1300, The Madison Project (2013)
\bar{K}	11.780	Aggregate population, 2000, The Madison Project (2013)
ν	0.190	Land share, Caselli (2005)
π	0.667	Adult life expectancy of 75
I	2000	Number of types
$\{\beta^i\}_{i=1}^I$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^I$	Subdivide domain into grid
α	0.428	Consumption share (see text)
$\bar{\beta}$	0.572	Maximum (generational) discount factor
$\{\gamma_{28}, \delta_{28}\}$	$\{36, 60\}$	Standard deviation of discount factors (Andersen et al., 2008; Falk et al., 2018) and long run rate of return (see text)
$\left\{\frac{N_0^i}{N_0}\right\}_{i=1}^I$	See text	Andersen et al. (2008) and Falk et al. (2018)
$\left\{\frac{K_0^i}{K}\right\}_{i=1}^I$	See text	Consistency (see text)

Note also that the highest discount factor in our grid, β^I , is related to the upper bound of the discount factors, $\bar{\beta}$, by the expression $\beta^I = \bar{\beta} \left(\frac{2I-1}{2I}\right)$ where $\bar{\beta} \equiv 1 - \alpha$. With $s_{ss}^c = 0.75$ chosen to match to the average global steady-state income share post-2000,¹² we can thus solve the above equations simultaneously to obtain: $\alpha = 0.428$ and $\bar{\beta} = 0.572$.

Finally, we need the initial distribution across dynasties of capital, $\{K_0^i\}_{i=1}^I$, and population, $\{N_0^i\}_{i=1}^I$. This data is not readily available for the year 1300. Instead, our calibration strategy will rely, first, on an assumption that the model was in equilibrium *prior* to our initial period, and, second, on using the model to obtain the relative initial population of each dynasty from contemporaneous data.

Capital distribution 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 popula-

¹²See Appendix A for details.

tion is consistent with solutions of the model in the period prior to the initial period. 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 $t = 1$. 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 (17) also holds for $t = 0$ which in turn implies that the second expression in (18) also holds at $t = 0$:

$$\frac{N_1^i}{N_1^j} = \frac{\beta^i N_0^i}{\beta^j N_0^j}. \quad (24)$$

There are three steps to see why the above consistency assumption is necessary to pin down the distribution of initial capital stock across dynasties. First, in Appendix C we establish a relationship between the population of each dynasty (relative to the most patient dynasty) in the first period and the consumption of each dynasty (relative to the most patient dynasty) in period zero:

$$\frac{C_0^i}{C_0^I} = \frac{N_1^i}{N_1^I} \left(\frac{1 - \alpha - \beta^I}{1 - \alpha - \beta^i} \right). \quad (25)$$

Second, the consistency assumption (24) along with (25) amounts to fixing the initial distribution of consumption (relative to the most patient dynasty) according to the following:

$$\frac{C_0^i}{C_0^I} = \frac{\beta^i}{\beta^I} \left(\frac{1 - \alpha - \beta^I}{1 - \alpha - \beta^i} \right) \frac{N_0^i}{N_0^I}. \quad (26)$$

That is, given the consistency assumption, the initial population distribution tells us what initial consumption distribution should be. Finally, we can use dynastic budget constraints (22) to determine what this initial distribution of consumption implies about the initial distribution of capital, $\{K_0^i\}_{i=1}^I$.

Population distribution 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 (which we describe below) on the distribution of types in the year 2000 ($t = 28$ in the model). Equation (19) gives the population share of each dynasty over time as a function of the $t = 1$ pop-

ulation share and each dynasty's level of patience. Using this and (24), we have the $t = 0$ population share of each dynasty i relative to dynasty I :

$$\frac{N_0^i}{N_0^I} = \frac{N_t^i}{N_t^I} \left(\frac{\beta^i}{\beta^I} \right)^t, \quad (27)$$

With evidence on the distribution of patience at some later date t , we could thus calibrate the initial distribution of the population across levels of patience. One problem with this approach is that modern data 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 2000.¹³ To address this issue, 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:

$$F(\beta; t) = \frac{B(\beta/\bar{\beta}, \gamma_t, \delta_t)}{B(\gamma_t, \delta_t)}. \quad (28)$$

In the above, $B(\gamma_t, \delta_t)$ and $B(\beta/\bar{\beta}, \gamma_t, \delta_t)$ 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, it is a distribution that 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 that is often used to mimic other distributions, both skewed and centered, given appropriate bounds. Finally, the the beta distribution is also intimately linked to the evolution of the population distribution implied by our model, as the following Theorem shows:

Theorem 1. *If $I \rightarrow \infty$ and dynastic discount factors 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 $\delta_{\bar{t}+1} = \delta_{\bar{t}}$.*

¹³For example, consider two dynasties i and j with discount factors $\beta^i = 0.05$ and $\beta^j = 0.5$. From equation (27), 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 $\frac{N_0^i/N_0^j}{N_{28}^i/N_{28}^j} = \left(\frac{\beta^i}{\beta^j} \right)^{28} = 10^{-28}$.

Proof. See Appendix E. □

Theorem 1 establishes that, for a fine enough grid, if discount factors obey a scaled beta distribution in any one period then they will follow a scaled-beta distribution in all other periods. If the beta distribution fits the data well in a given year, the model predicts it will fit the data well in any other year. Furthermore, because the model pins down the evolution of parameters of the scaled beta distribution, our choice of year to calibrate the scaled beta distribution (here the year 2000) will be irrelevant – in principle, the same parameters (adjusted for time) would emerge if we were to recalibrate the model using data at another point in time. 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 :

$$E_t(\beta) = \bar{\beta} \frac{\gamma_0 + t}{\gamma_0 + t + \delta} \text{ and } \text{var}_t(\beta) = \bar{\beta}^2 \frac{(\gamma_0 + t)\delta}{((\gamma_0 + t) + \delta)^2(\gamma_0 + t + \delta + 1)} \quad (29)$$

As $t \rightarrow \infty$, the mean beta converges to $\bar{\beta}$ and the variance goes to zero: thus the agent with the highest discount factor comes to entirely dominate the economy.

Note that the two shape parameters of the distribution of generational discount factors may be obtained if we observe the mean and variance of that distribution. Our measure of the variance is derived from data on *annual* discount rates. Our target for the mean is a function of the prevailing long-run interest rate in the economy. We thus need expressions for the variance of the *annualized* generational discount factor and for the long-run interest rate in terms of the parameters of the distribution of *generational* discount factors. The variance of the annualized generational discount factor, $\beta^{\frac{1}{25}}$, is given by:

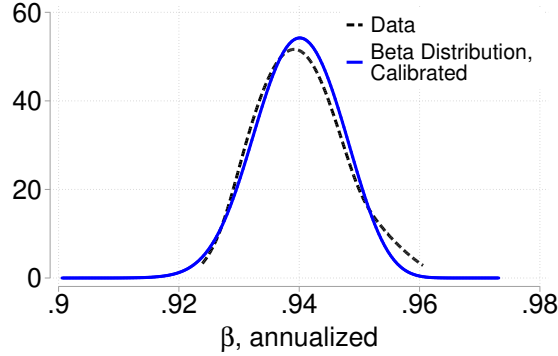
$$\text{var}_t(\beta^{\frac{1}{25}}) = \bar{\beta}^{\frac{2}{25}} \frac{\Gamma(\gamma_t + \delta_t)}{\Gamma(\gamma_t)^2} \left(\frac{\Gamma(\gamma_t)\Gamma(\frac{2}{25} + \gamma)}{\Gamma(\frac{2}{25} + \gamma_t + \delta_t)} - \frac{\Gamma(\gamma_t + \delta_t)\Gamma(\frac{1}{25} + \gamma_t)^2}{\Gamma(\frac{1}{25} + \gamma_t + \delta_t)^2} \right), \quad (30)$$

and an approximate expression (see Appendix E) for the annualized gross interest rate:

$$R_t^{\frac{1}{25}} \approx 1 + \left(\frac{\gamma_t - 1 + \delta_t}{\bar{\beta}\gamma_t} \right)^{\frac{1}{25}}. \quad (31)$$

As described in Appendix A, we set $\text{var}_{28}(\beta^{\frac{1}{25}}) = 0.005^2$ to match experimental evidence from representative individuals in Denmark (Andersen et al., 2008) and the

Figure 2: Distribution of annualized discount factors in model and data



Notes: The figure depicts the calibrated scaled beta distribution used in the model period $t = 28$ (solid line) against the bandwidth filtered data described in the text with $BW = 0.005$ (dashed line).

individual-level data in the Global Preference Survey (GPS) described in Falk et al. (2018). We set $R_{28}^{\frac{1}{25}} - 1 = 0.063$ to match the average (annualized) generational rates of return on global equities.¹⁴ Together, these two equations imply the following shape parameters of the beta distribution: $\gamma_{28} = 36$ and $\delta_{28} = 60$. As can be seen in Figure 2, there is a good fit between the annualized distribution of generational discount factors in the year 2000.

Once parameters γ_t and δ_t have been calibrated, we can use the CDF to approximate, for some I , the proportion of the population assigned to each dynasty i in the year 2000 (i.e. period $t = 28$) by:

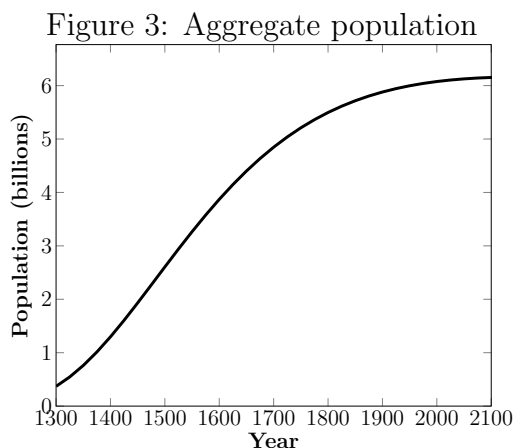
$$\frac{N_{28}^i}{N_{28}} = F\left(\beta^i + \frac{\bar{\beta}}{2I}; 28\right) - F\left(\beta^i - \frac{\bar{\beta}}{2I}; 28\right). \quad (32)$$

With the above proportions in hand, we can then calculate the $t = 0$ distribution of population using equation (27) with $t = 28$, and proceed to solve the model.

¹⁴In Appendix A we show that over the time spans under consideration by dynastic planners – a basket of global equities was just as safe as bonds or treasuries but offered higher rates of return. Specifically, the variation in the global rates of return on equities over 25 year periods are either smaller or statistically indistinguishable from rates of return on government bonds or treasuries. Since we are focusing on dynasty planners that have a horizon of 25 years or more, we calibrate to the higher rates of equity return.

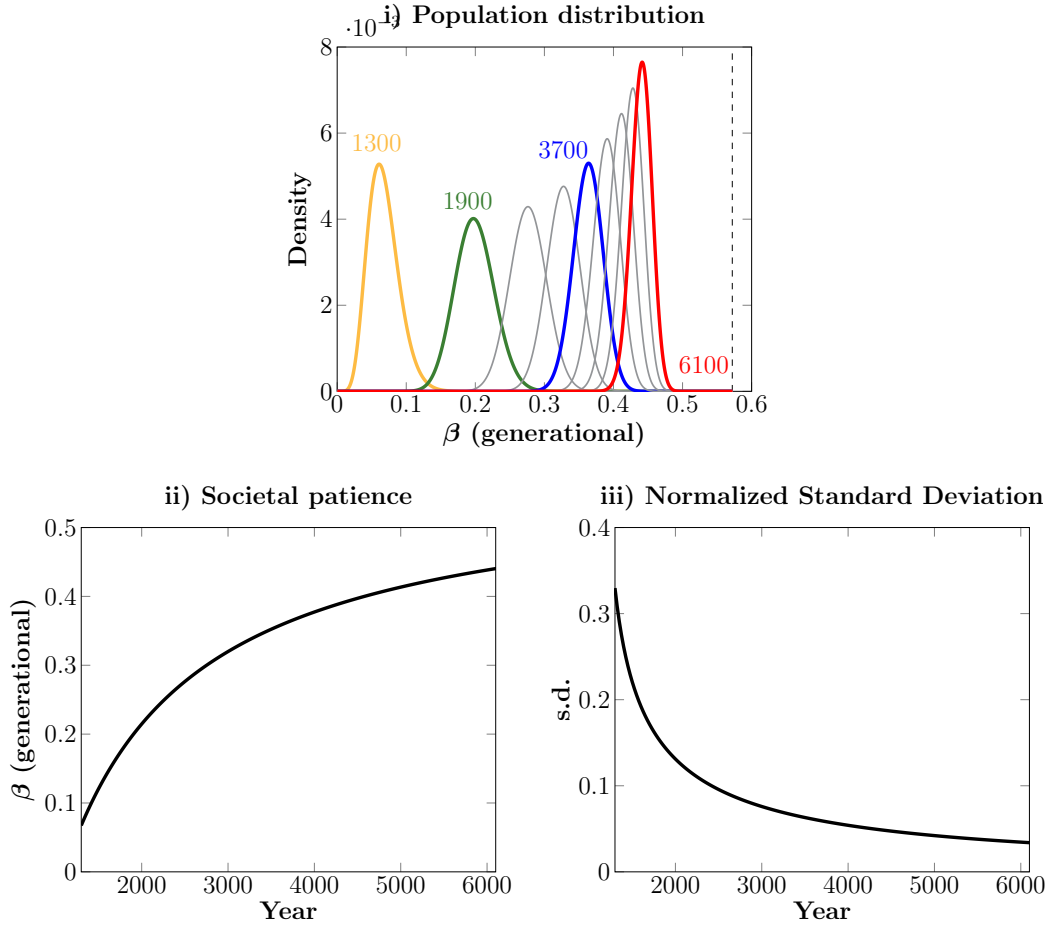
6 Quantitative results

Figure 3 shows the increase in aggregate population over time generated by the model. Since the model was calibrated to match the levels of global population in the years 1300 and 2000, we match the increase in world population over the period. Since the level of technology in the model is constant, the increase predicted by the model is more smooth than observed in the data.



Next, we examine the predictions of the model for the distribution of patience levels in the population. Figure 2 showed the distribution of discount factors across the population in both the model and the data in the year 2000. A key implication of our model, is that this distribution changes over time: the mean patience of the population increases, whilst the variance (normalized by the mean) decreases as is shown in equation (29). Figure 4 panel i) depicts this evolution over time. In our initial period, 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). 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 now has a discount factor of around $\beta = 0.2$. The consequence of this, as shown in panel ii) of Figure 4, is that the level of societal patience monotonically increases over time. The (un-normalized) standard deviation of patience in the population will first gradually grow over time (as more patient agents have more children) and then later it will become more concentrated over time (as the mass of the

Figure 4: The rise of societal patience



Notes: Panel i) shows the population distribution of levels of patience (the generational β) at 600-year intervals starting in the year 1300 and ending in 6100. The dashed vertical line is at $\beta = 0.572$. Panel ii) depicts the societal average level of generational patience over time. Panel iii) shows the mean-normalized standard deviation of patience in the population over time.

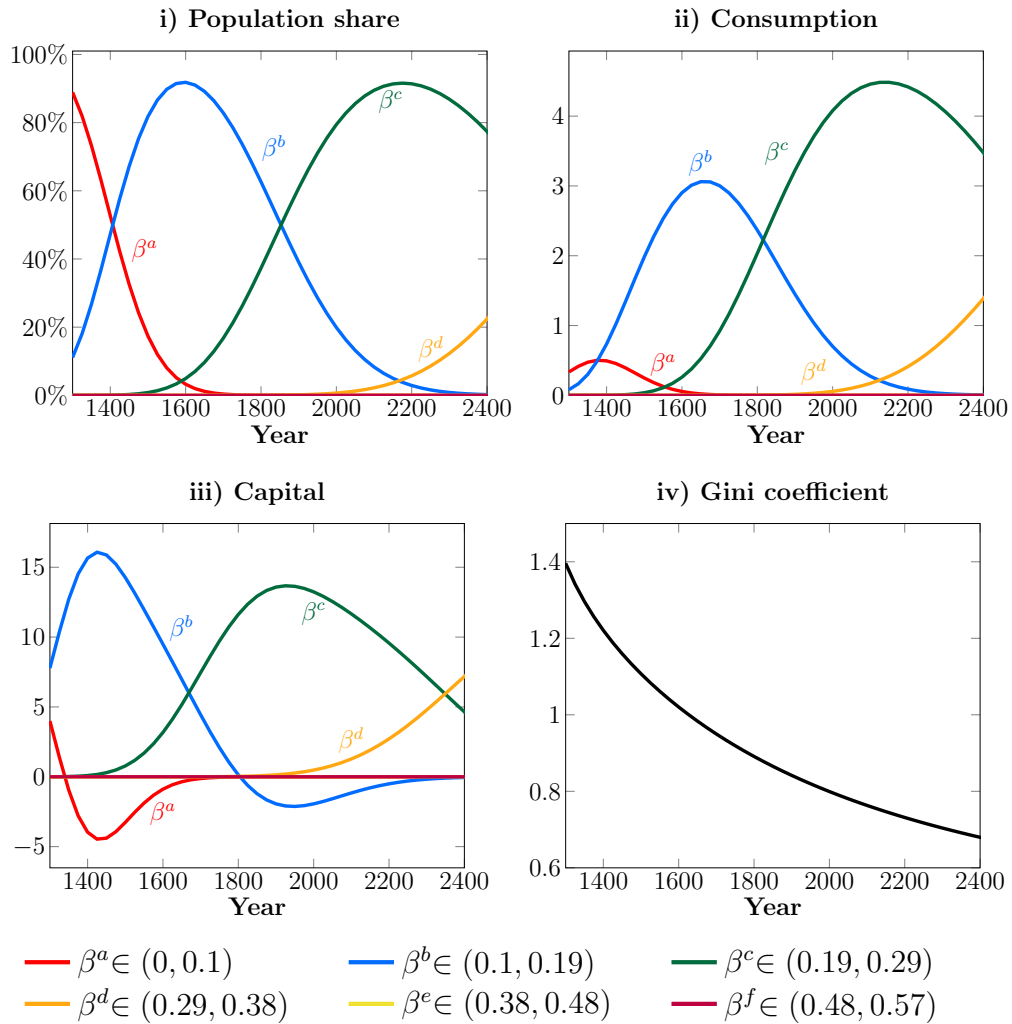
population reaches the upper limit of patience, $\bar{\beta}$). As is shown in panel iii) of Figure 4, the mean-normalized standard deviation of patience decreases monotonically as the population becomes concentrated in the most patient dynasties, eventually reaching a mass point at the most patient dynasty. We can also show that the first and second derivatives of the mean with respect to time are positive and negative respectively. In other words, mean patience is increasing, but at a decreasing rate.

The key parameter governing this evolution is the shape parameter of the scaled beta distribution, γ_t , which, as we argued in Theorem 1, evolves approximately ac-

According to the first-order difference equation, $\gamma_{t+1} = \gamma_t + 1$. In the population genetics literature this type of evolution of an organism’s phenotype (or observable trait) over time, first identified and discussed by Darwin (1859), is known as ‘directional selection’. This is a form of natural selection in which extreme characteristics of agents are favored over less-extreme characteristics (in a given environment) and which in turn causes the relative frequency of the extreme variant of an agent to shift over time in the direction of that particular agent type. Under this sort of selection the numbers of the advantageous type of agent increase as a consequence of differences in survival and reproduction abilities among different types. In our simplified case, survival probabilities are the same across agents and only reproduction abilities vary. Another feature of this type of directional-selection, which also holds in our model, is that the increase in the share of the dominant type is independent of the dominance of the particular type at any given moment (Molles, 2010). This fact follows directly from the above first-order difference equation which is independent on the population share of the dominant type of agent.

To aid our discussion of the changes within the population over time, we split dynasties into six groups according to their level of patience. This permits us to examine the characteristics of low, medium and high-patience types over time as represented by groups of dynasties. Figure 5 panel i) gives the share of each group as a percentage of the total population over time. There is a clear cyclical pattern over time in the types of patience that dominate. The world starts out being dominated by the least patient agents, group β^a , who initially account for approximately 90% of the total population in the year 1300. Over time, since they have 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 90% of all agents in the years 1600. Later still, the dominance of this group is broken by the rise of the β^c -group which in turn comes to overtake the population over the subsequent 400 years. This wave-like pattern continues into the future until, eventually, the entire population is dominated by the most patient group of agents. This figure emphasizes the findings shown in Figure 4, which demonstrates that the mean level of population patience shifts steadily upwards. The transition from least to most patient is not instantaneous – each dynasty and group of dynasties has their rise to and their fall from dominance of the overall population. Panel ii) of Figure 4 shows the consumption levels of each group over time, and makes clear that the waves depicted in panel i)

Figure 5: Characteristics of groups and inequality over 1300–2100

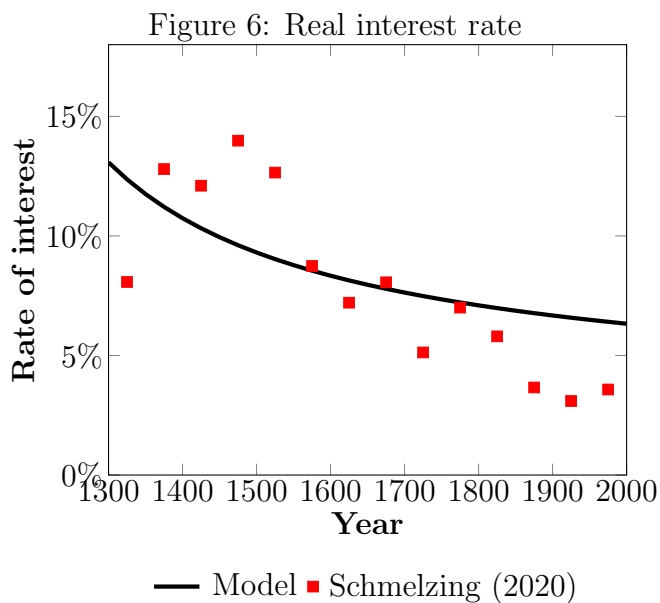


Note: Each panel i)–iii) reports the sum of the model output across all dynasties in the group of dynasties defined in the legend. Panel iv) reports the Gini coefficient calculated based on individual wealth.

occur along with substantial growth in the aggregate population.

The key to understanding the cyclical outcome for groups lies in Figure 5 panel iii), which reports the capital owned by each group over time. Since agents are able to lend and borrow capital in making optimal choices of consumption and children, the β^a -group of dynasties at first begins to borrow from the more patient dynasties in order to substitute away from children toward the current consumption good. The extent to which the most impatient dynasties can increase their consumption depends then on the population size of, and the capital owned by, the relatively more

patient types. The growth of the β^b -group thus facilitates the (relative) decline of the β^a -group since there emerges a larger and larger market for their capital. As the β^a -group diminishes, so the β^b -group emerges as the largest population and the dominant owner of capital. The eventual emergence of the β^c -group then yields to the β^b -group the increasing opportunity to sustain high consumption through sale of their capital holdings. In panel iv), we summarize the evolution of capital via a measure of inequality. We can see that inequality, as measured by the Gini coefficient, declines over time. Since within a dynasty wealth is evenly distributed, as the population distribution across types becomes more concentrated, so the overall distribution of wealth becomes more even.



Finally, Figure 6 reports the model fit against the interest rate data in Schmelzing (2020). We observe a significant fall in the implied rate as the level of societal patience grows. This decline is approximated by equation (31). Note that in addition to matching the decline, the model also captures the slowing rate of decline in the rate. We do not capture fluctuations around the long-run trend, as our model does not include factors such as time-varying growth rates, risk levels or cyclical shocks such as wars and pandemics. We explore the impact of such shocks to population and capital in Section 7.1. As we show, an unexpected decline in the population of 30% can result in a decline in the interest rate of around 1.5 percentage points. From this perspective, the Black Death (in the mid-fourteenth century) could explain a

portion of difference between what the model suggests for interest rates in that period and what is observed in the Schmelzing (2020) data. Most importantly our model, calibrated to evidence on the distribution of types of patience in the year 2000 (and other macroeconomic data) successfully captures the main trend in the real rate over the course of eight centuries.

6.1 Wealth and fertility

An important advantage of using a quantitative model is that its implications can be compared with further data not targeted by the calibration. This provides an additional external validation of the mechanisms at work. One such test is of the implied relationship between wealth and the number of children.

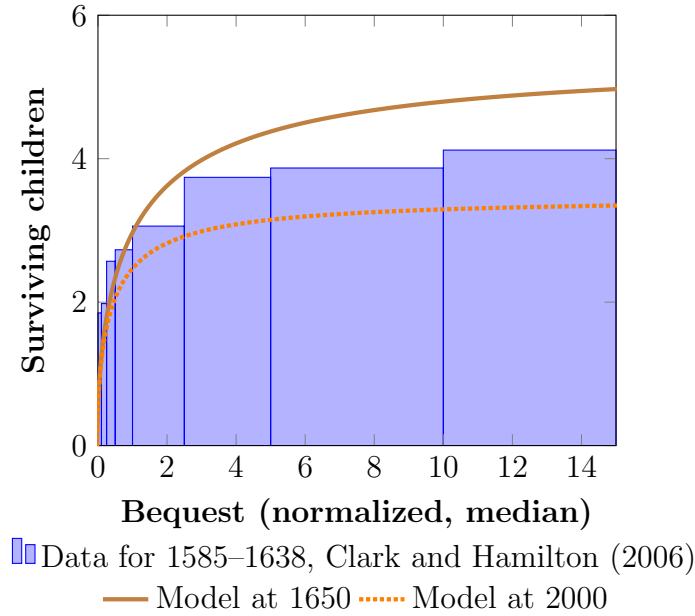
As described above, Clark and Hamilton (2006) uses English probate records over the period 1585–1638 to show that richer households tended to have more children than poorer households. The typical will in the Clark and Hamilton data includes the names of nearly all *surviving* children of the testator, along with the value of bequests both of the main property in the estate as well as small items for each child. Using this data, they show that the number of children surviving at a father’s point of death is positively related to the total value of the assets bequeathed to those children.¹⁵

To compare the results of our model with the bequest data, in Appendix D we obtain an expression for the expected number of surviving children, and the size of bequests, at the period of death of the parent. Figure 7 reports our model’s prediction for the year 1650, normalizing bequest levels by their median size at that point in time. Qualitatively, both the model and the data predict a non-linear, positive relationship between the value of bequests and the number of surviving children. Moreover, up to bequests of five times the median level, the quantitative predictions of the model are very close to the data. For higher levels of bequests, the model over-predicts by around one surviving child per testator. While there is much real-world complexity that is absent in the model, one candidate for the difference between the model and the data in this part of the wealth distribution is that wealthier men tended to marry younger, potentially more fertile women, something absent from our model.¹⁶

¹⁵While the likelihood of writing a will is naturally contingent on having positive assets to bequeath, Clark and Hamilton documents that wills were made by a cross section of society of all social classes, including labourers with limited possessions.

¹⁶Another potential explanation would be the existence of a positive relationship between child survival or life expectancy and wealth. Clark (2007a) reports evidence that these do not vary

Figure 7: Wealth and children



The mechanism generating the positive relationship between wealth and children in our model stems from the positive connection between patience and fertility as well as between patience and wealth. More patient households place a greater value on future consumption and hence save more, both in terms of physical capital and in terms of children. While ‘survival of the richest’ does indeed hold empirically, the driving mechanism for our relationship is the ‘survival of the patient’.

Figure 7 also reports the implication of the model for the same relationship in the year 2000 – children are still a ‘normal’ good, as in Becker (1960). Richer households continue to have more children, but just not as many more as in the past. This is supported by household-level evidence that identifies the relationship using exogenous variation in income or wealth, such as Black et al. (2013), Lovenheim and Mumford (2013), Kearney and Wilson (2018) and Bennett et al. (2020).¹⁷

significantly by the level of assets at death at this time.

¹⁷Cross-sectional and time-series evidence can find a negative correlation; see Galor and Weil (2000), Manuelli and Seshadri (2009) and Doepke and Tertilt (2016).

7 Extensions

In this section we consider two types of extension: First, we look at the impact of unexpected shocks to aggregate population and the land stock. Second, we extend our baseline set-up to consider a form of mutation, which we model as an expected change in the discount rate in a portion of the median-type dynasty.

7.1 The consequences of pandemic, war and revolution

Jordà et al. (2020) estimates the medium-run consequences of pandemics and wars on real rates of return, using the data in Schmelzing (2020). Our first set of extensions thus considers whether our model is consistent with these results. We can go further with the model, however, and also compare our results to a recent empirical study on wealth inequality of the Black Death in Medieval Europe (Alfani and Murphy, 2017).

Pandemic We model a pandemic as an event where a constant fraction of each dynasty (and hence the population) unexpectedly dies at the start of a period.¹⁸ The net capital holdings of the deceased households are re-distributed equally among remaining members of the dynasty. We suppose that the disease hits in 2025 and has a death rate of 30%. This size of shock is chosen in order to generate a large, unexpected pandemic similar in magnitude to the medieval Black Death.

We report results in Figure 8. Panel i) shows aggregate population. Immediately after the negative shock to population, households choose to have more children as the returns on children increase relative to those of land. No underlying parameters of the model change and so the long run steady state level of population is as it was prior to the pandemic. Only around the year 2800 does the population reach close to the same level it would have been. The effects of pandemics can, along certain dimensions, be very long lasting. Panel ii) gives the interest rate. Since capital is immediately more abundant relative to labour the interest rate drops in the period of the shock. Subsequently, the interest rate is marginally higher than in the baseline, driven by a higher population growth rate as the economy returns to its pre-shock growth path.¹⁹ This result for interest rates is consistent with the findings in Jordà

¹⁸For computational convenience in this and the subsequent section we reduce the grid of discount factors from $I = 2000$ to $I = 20$ dynasties. For the time period under consideration, the sparser grid significantly reduces computational time and leaves results quantitatively nearly identical.

¹⁹Higher population growth rates result in higher interest rates since they make an investment in

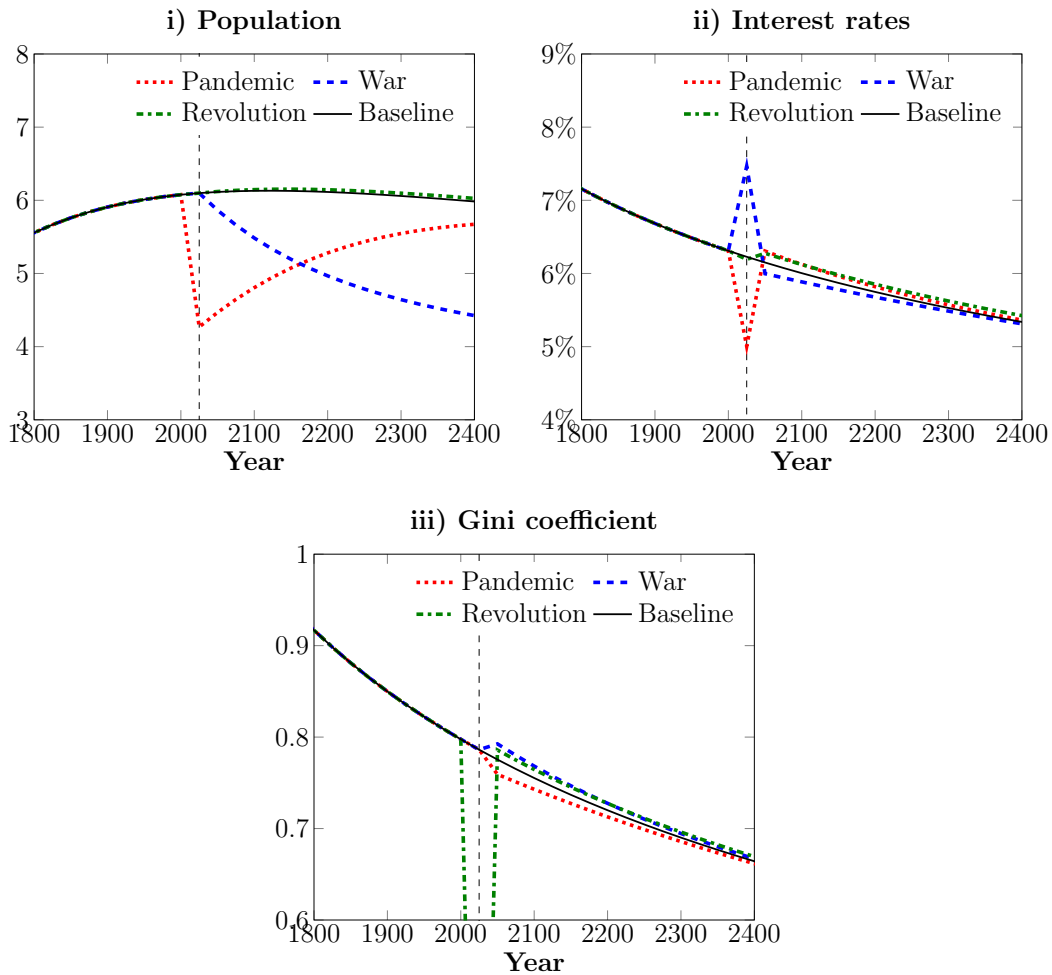
et al. (2020) who show that the immediate response of the interest rate to a typical pandemic is a fall of the real rate, with effects lasting up to 40 years after the end of the pandemic on average. Finally, panel iii), shows that a pandemic acts to reduce the level of wealth inequality. The scarcity of workers after a pandemic drives up wages for those who survive. Households that rely more on wages than rental income to accumulate greater quantities of capital thus reducing wealth inequality. Again, this is consistent with the work of Alfani and Murphy (2017) who finds a large decline in economic inequality driven by a similar mechanism in much of Europe during and after the Black Death.

War Next, we consider a counterfactual that is more akin to a large war, a shock that results in the permanent destruction of capital (land). We model a war as the destruction of 30% of each dynasty’s net capital holdings at the start of a period. This experiment thus sheds light on how a heterogeneous agent economy adjusts to a sudden decrease in the capital-labor ratio caused by a decrease in the capital stock. The results in Figure 8 panel i) show that in this case the long-run level of population does not recover; the long run capital to labour ratio is unchanged and, since there is no mechanism for capital accumulation in our model, this implies that the long-run total population must be lower. The results in panels ii)-iii), mirror the results (with reversed intuition) from a pandemic; we see an immediate spike in interest rates followed by lower than baseline rates driven by negative population growth, and an increase in inequality that decays over time due to lower wages. Again, the estimates in Jordà et al. (2020) for the impact of war on the interest rate are qualitatively consistent with the predictions of our our model.

Revolution Finally, we consider the effects of a revolution that we model as an unexpected uniform (re)distribution of capital across agents in the year 2025, à la Piketty (2014). Figure 8 depicts the consequences for population, interest rates and the Gini coefficient. By design, in the period of the shock the Gini coefficient falls to zero. Thereafter, inequality, as measured by the Gini coefficient, is marginally higher. Since over time capital is transferred from less patient to more patient households, the return of capital to dynasties that had previously disposed of it shifts the economy

capital today be worth more tomorrow since capital will be relatively more scarce, given the higher population.

Figure 8: Pandemic, war and revolution



Note: Figures depict the long run consequence of a one-off and unanticipated pandemic (a loss of 30% of the total population), a war (a permanent loss of 30% of the capital stock), and a revolution (an equal redistribution of capital across agents) in the year 2025 (dashed vertical line). In panel iii) the Revolution treatment causes the Gini coefficient to drop to zero in 2025, before returning to above the baseline in 2050, the next (generational) period.

back to an equilibrium similar to that in prior years. The decline in the real interest rate is also shifted back to that in previous generations, for similar reasons: the interest rate in the economy is, over the extended transition to the same long run interest rate (as defined by the upper level of patience), higher than that predicted in the baseline.

One-off redistributive policies mechanically reduce inequality during the gener-

ation of their implementation. Thereafter, our model predicts that they result in higher levels of inequality. Furthermore, even though these policies are only ‘one-off’ they nonetheless result in long lasting changes to the economy resulting in higher interest rates and lower capital to labor ratio for hundreds of years.

7.2 Mutation

Our baseline model 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). 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 allow for the possibility that a group of agents 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. Mutations can be thought of as changes in the discount factor brought about by parental or peer influence through education or parental investment (i.e., different forms of imitation and socialization). 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.²⁰ Thus, whilst primarily motivated by genetic mutation, this section can also be interpreted as examining the effects of a new variant of dynasty no-matter its source.

Setup We model a mutation as an unanticipated, one-off and permanent shock to an agent’s discount factor. Instead of attempting to match the rate at which mutations occur in nature (something which would be difficult to calibrate) we instead consider the consequences of different types of one-off mutations. Each mutation

²⁰In 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.

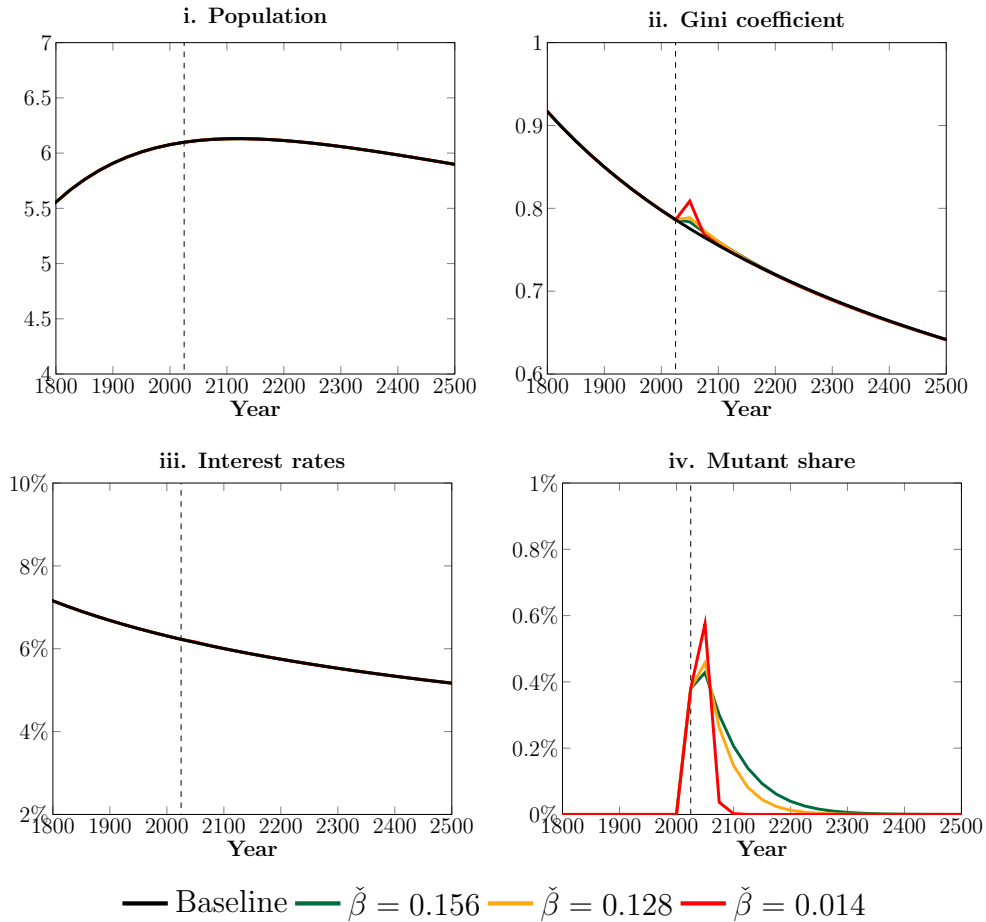
counterfactual involves an unexpected but permanent change in discount factor for 1% of agents belonging to the dynasty with that period's median discount factor. These mutants then form a new dynasty, retaining their net capital per capita from the previous period.²¹ These types of mutations can be divided into two categories based on the impact they have on an agent's 'fitness' or reproductive success: deleterious and advantageous mutations.

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 9 reports the effect on population, inequality and 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 the lowest patience – have very small effects on population, interest rates and inequality. 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 known as 'advantageous' in the biological literature as they increase the fitness of the dynasty: agents mutating to 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 10 shows the effects on population, interest rates and inequality of a mutation to successively higher discount factors as well as

²¹For 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 9: 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 ($\beta = 0.213$) of patience in 2025 to wake up in 2025 with the level of patience $\check{\beta}$ denoted in the Figure legend.

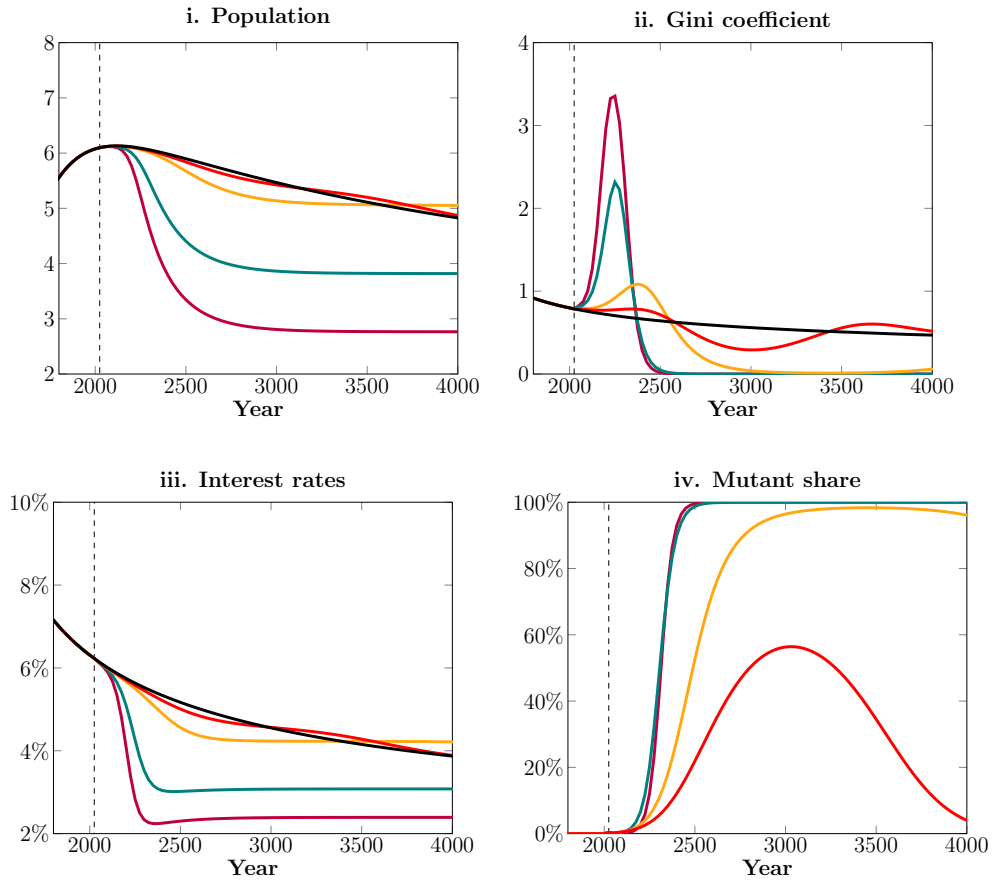
the share of mutants in the population. Notice that a mutation to the highest level of patience pushes the economy forward in the evolutionary process by thousands of years. Since at the time of the mutation, so few agents are of the most patient type, a 1% mutation of the median dynasty to the highest-patience dynasty is an enormous shock. The economy is suddenly inhabited by a relatively large proportion of the agents of the most patient type. These agents quickly amass all the capital in the economy and begin to have large numbers of children which thereafter dominate the population. This process would have happened without the mutation, but

it would have lasted thousands of years more. With mutation the process lasts less than a thousand years. Thus we shift from today's economy to one in which the most patient agents dominate. Population and interest rates approach the long run steady state. In response to this shock, wealth inequality (as measured by the Gini coefficient) first spikes to levels of nearly 3.5 then falls to practically zero. This occurs because the mutated agents very quickly start purchasing capital from all the agents in the economy. This results in all existing agents getting into debt and substituting children for consumption. Since the remaining lower-patience dynasties (who are now in debt) continue to make up a relatively large part of the population, wealth inequality rises. After about 500 years however, all but the most patient dynasty have been out-populated. Since by then there is only one type of dynasty, wealth inequality falls to zero.

Mutations to levels of patience that are higher-than-median but not the highest, give rise to some especially interesting dynamics. Agents mutated in this manner can come to dominate the population for some time (see for example Figure 10), where mutants with discount factor 0.355 practically dominate the population for a thousand years or so before being overtaken by 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 dominate the population forever and interest rates and Gini coefficients 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. In the case of the outcomes this results in multiple oscillation 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.

Timing and environment The above discussion points to the importance of the pre-existing environment when it comes to the impact of mutation. The exact same mutation can have vastly different effects on outcomes depending on when it takes place. What may be a highly advantageous mutation in an environment where the median dynasty is especially impatient might not be nearly as advantageous, or might even be deleterious, in an environment where the median dynasty is very patient. To emphasize this point, Figure 11 shows the effects of the same mutation occurring in one of three different years. In particular, we consider the same mutation of 1% of

Figure 10: Advantageous mutations in 2025

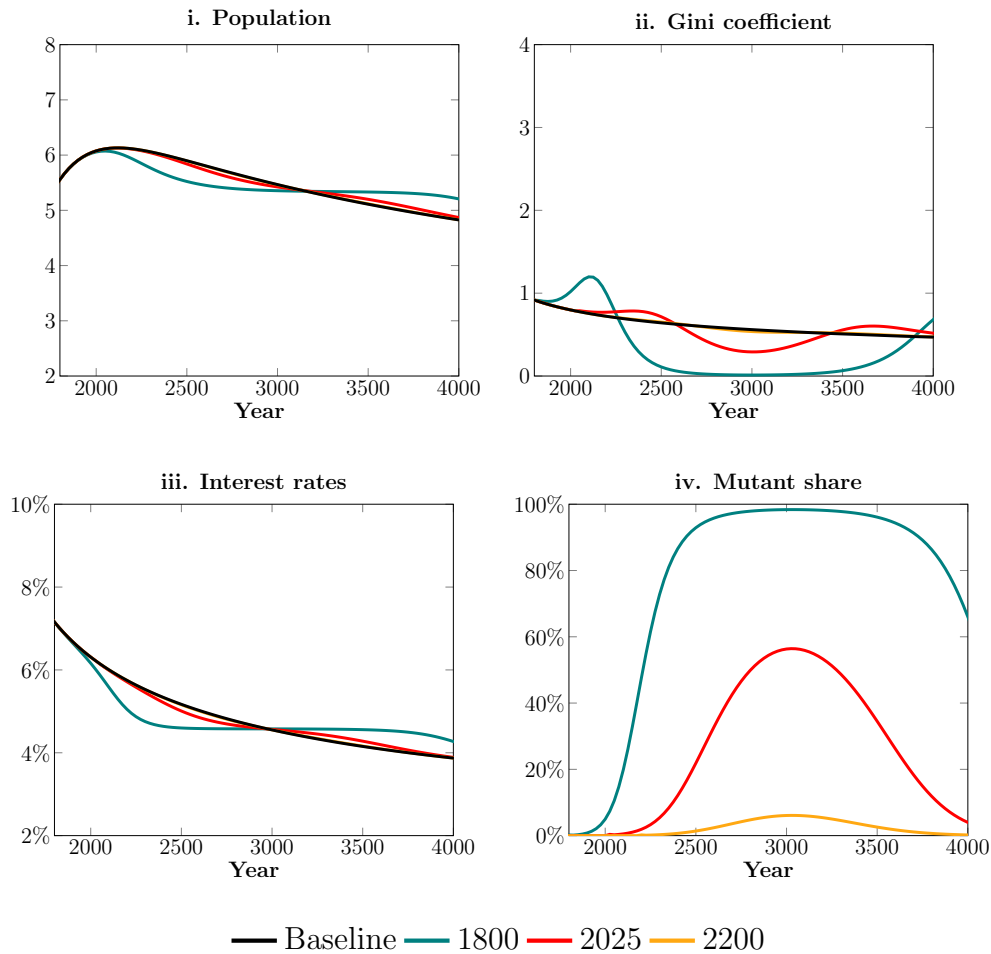


— Baseline — $\check{\beta} = 0.554$ — $\check{\beta} = 0.466$ — $\check{\beta} = 0.355$ — $\check{\beta} = 0.327$

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 ($\beta = 0.213$) of patience in 2025 to wake up in 2025 with the level of patience $\check{\beta}$ denoted in the Figure legend.

the dynasty with a discount factor of 0.212 (the 2025 median in the baseline) to a discount factor of 0.327, but change the year in which it takes place to 1800, 2025 and 2200. If the mutation takes place in 1800 then mutants dominate the population to a far larger extent and for a far longer period of time than if the mutation takes place in 2025 or 2200. A mutation in 1800 also has a very sizeable economic impact affecting population, inequality and interest rates for more than a thousand years. The same mutation by 2200 however has almost no discernible effect. Thus whether mutations are deleterious or advantageous is not predetermined but depends on the

Figure 11: Same mutation, different periods



Note: Figures report the simulation output with the same unexpected mutation in the year 1800, 2025 and 2200 respectively. Each line is the same mutation counterfactual that takes place in different years. A mutation causes 1% of the dynasty with discount factor of $\beta = 0.213$ to wake up in either 1800, 2025 or 2200 with the level of patience $\hat{\beta} = 0.327$.

structure of the rest of the population and hence on the environment at the time of mutation.²²

Implications for evolution One final point that emerges from this last exercise is that mutation in the distant past can give rise to long periods of stability, where evolution seems to stop only for it to seemingly start up once more many hundreds of

²²The only exception being a mutation to the very highest level of patience: this will always (eventually) dominate the economy irrespective of the environment.

years later. For example, as Figure 11 panel iv shows, the mutation in 1800 generates a period where the population is almost entirely composed of one type of agent (the mutant) between the years 2400 and 3600. This period of time is associated with practically constant population and interest rates as well as zero inequality. Looking at this sort of data might lead one to mistakenly infer that the economy is in a steady state, and that the process of natural selection had concluded. The process however is only paused. The mutation in 1800 results in the economy ‘leap-frogging’ the evolutionary process and the selection process once more begins to apply to mutants as the share of the more patient non-mutants comes to dominate thousands of years after the initial mutation. Thus, over the course of human history, it may be quite reasonable to expect very long periods of stability in terms of interest rates and patience, only to be followed by a ‘gradual then sudden’ change. All one needs for this to happen is a mutation to a particularly advantageous discount factor in the past. These mutants then dominate the economy for long enough periods to give rise to the illusion that evolution has halted.

8 Concluding remarks

We introduced a simple fertility model with heterogeneous 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. There are many further implications to consider. First, in our model the population shift toward more patient types occurs partly via trading in the fixed asset, land. This suggests a potentially important relationship between the constraints on trade or borrowing, the evolution in the population and the interest rate. Second, the historical trends in growth and risk work to increase, not decrease, the real interest rate. With a more general model and with data for the evolution in risk and growth, we may conduct an exercise to attribute portions of the trend to different causes. Third, we have focused on a simple form of the intergenerational transmission of preferences. More likely than perfect transmission is some form of partial transmission, either by genetic mutation or environmental adaptation or imitation. Moreover, we studied heterogeneous patience levels as the only time-varying element of societal preferences. The evidence on the heterogeneity of risk aversion, together with its intergenerational transmission and affect on fertility, suggests that this could be an additional further

preference heterogeneity that evolves over time alongside patience. A more general model could account for the evolution of the distribution across patience and risk aversion. Fourth, we have focused our model on its implications for the interest rate but our time period encompasses the onset of the industrial revolution. The role for the evolution of societal preferences in explaining potentially endogenous technological progress is left for future work.

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. What is clear from our analysis is that such policy should take into account not only that the social discount rate evolves over time in a predictable fashion, but that that path is not independent from some policy interventions. Understanding the short- and long-run relationship between the social discount rate and policy interventions is an important avenue for future research.

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.
- Alfani, Guido and Tommy E. Murphy**, “Plague and lethal epidemics in the pre-industrial world,” *Journal of Economic History*, 2017, *77* (1), 314–43.
- 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, S., G. W. Harrison, M. I. Lau, and E. E. Rutström**, “Eliciting Risk and Time Preferences,” *Econometrica*, 2008, *76* (3), 583–618.
- Arrow, K. J., M. Cropper, C. Gollier, B. Groom, G. Heal, R. Newell, W. Nordhaus, R. Pindyck, W. Pizer, P. Portney, T. Sterner, R. S. J. Tol, and M. Weitzman**, “Determining Benefits and Costs for Future Generations,” *IPCC Policy Forum*, 2013.

- Bar, Michael and Oksana Leukhina**, “Demographic transition and industrial revolution: A macroeconomic investigation,” *Review of Economic Dynamics*, 2010.
- Barro, Robert J. and Gary S. Becker**, “A Reformulation of the Economic Theory of Fertility,” *Quarterly Journal of Economics*, February 1988, *103* (1), 1–25.
- and –, “Fertility Choice in a Model of Economic Growth,” *Econometrica*, March 1989, *57* (2), 481–501.
- Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom**, “Genetic Endowments and Wealth Inequality,” *Journal of Political Economy*, 2020, *128* (4), 1474–1522.
- Becker, Gary S.**, “An Economic Analysis of Fertility,” in Gary S. Becker, ed., *Demographic and Economic Change in Developed Countries*, Princeton University Press, 1960, pp. 1789 – 1891.
- , “Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology,” *Journal of Economic Literature*, 1976, *14* (3), 817–26.
- , **Kevin M. Murphy, and Robert Tamura**, “Human Capital, Fertility and Economic Growth,” *Journal of Political Economy*, 1990, *98* (5), S12–37.
- 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,” *NHH Discussion Paper 03/2020*, 2020.
- 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, A. A. and T. Epper**, “The Intergenerational Transmission of Time Preferences Across Four Decades,” 2018. Mimeo.
- 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).
- Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio**, “Demographics and Real Interest Rates: Inspecting the Mechanism,” *Federal Reserve Bank of San Francisco Working Paper*, 2021, p. 32.
- Caselli, Francesco**, “Accounting for Cross-country income Differences,” in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth, Volume 1A.*, Elsevier B.V., 2005, pp. 680–738.
- Chowdhury, S., M. Sutter, and K. F. Zimmerman**, “Evaluating intergenerational persistence of economic preferences: A large scale experiment with families in Bangladesh,” 2018. IZA Discussion Paper 11337.
- 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, 2007.
- , “Genetically Capitalist? The Malthusian Era, Institutions and the Formation of Modern Preferences,” *mimeo*, 2007.
- , “The Macroeconomic Aggregates for England, 1209-1869,” *Research in Economic History*, 2010.
- **and Gillian Hamilton**, “Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England,” *The Journal of Economic History*, 2006, 66 (3), 707–36.
- Darwin, Charles**, *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.*, London: John Murray, 1859.

- Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti**, “Global Trends in Interest Rates,” September 2018. Federal Reserve Bank of New York Staff Reports, No. 866.
- Dimson, Paul Marsh Elroy and Mike Staunton**, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.
- Doepke, M. and M. Tertilt**, “Chapter 23 - Families in Macroeconomics,” in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2 of *Handbook of Macroeconomics*, Elsevier, 2016, pp. 1789 – 1891.
- Doepke, Matthias, Giuseppe Sorrenti, and Fabrizio Zilibotti**, “The Economics of Parenting,” *Annual Review of Economics*, 2019.
- Dohmen, T., A. Falk, D. Huffman, and U. Sunde**, “The Intergenerational Transmission of Risk and Trust Attitudes,” *The Review of Economic Studies*, 2011, 79 (2), 645–77.
- Fagereng, Andreas, Magne Mogstad, and Marte Rønning**, “Why Do Wealthy Parents Have Wealthy Children?,” *Journal of Political Economy*, 2021, 129 (3), 703–56.
- Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. 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, O. and D. N. Weil**, “Population, technology and growth: From the Malthusian regime to the demographic transition’,” *American Economic Review*, 2000, 90 (4), 806–28.
- **and O. Moav**, “Natural Selection and the Origin of Economic Growth’,” *Quarterly Journal of Economics*, 2002, 117 (4), 1133–91.
- **and O. Özak**, “The Agricultural Origins of Time Preference,” *American Economic Review*, 2016, 106 (10), 3064–103.

- Griffiths, Anthony J. F.**, “Mutation,” *Encyclopædia Britannica*, 2020.
- Hansson, I. and C. Stuart**, “‘Malthusian Selection of Preferences’,” *American Economic Review*, 1990, *80* (3), 529–44.
- Homer, Sidney and Richard Sylla**, *A History of Interest Rates*, 4 ed., Wiley, 2005.
- 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.
- Jordà, Oscar, Sanjay R. Singh, and Alan M. Taylor**, “Longer-Run Economics Consequences of Pandemics,” *Federal Reserve Bank of San Francisco Working Paper*, 2020, No. 2020-09.
- 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.
- 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.
- Millner, Antony**, “Non-dogmatic social discounting,” *American Economic Review*, 2019.
- Mitra, Tapan and Gerhard Sorger**, “On Ramsey’s conjecture,” *Journal of Economic Theory*, 2013, *148*, 1953–76.

- Mokyr, Joel**, *A Culture of Growth*, Princeton University Press, 2016.
- Molles, MC**, *Ecology Concepts and Applications*, McGraw-Hill Higher Learning, 2010.
- Nunn, Nathan**, “History as Evolution,” in Alberto Bisin and Giovanni Federico, eds., *The Handbook of Historical Economics*, Academic Press, January 2021, pp. 41–91.
- Parker, Geoffrey**, *The Grand Strategy of Philip II*, Yale University Press, 2000.
- Piketty, Thomas**, *Capital in the Twenty-first Century*, Harvard University Press, 2014.
- 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.
- 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.
- , – , and **Emil Iantchev**, “The Evolutionary Basis of Time Preference: Intergenerational Transfers and Sex,” *American Economic Journal: Microeconomics*, 2012, 4 (4), 172–201.
- Rogers, Alan R.**, “Evolution of Time Preference by Natural Selection,” *American Economic Review*, 1994, 84 (3), 460–81.
- 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. Bank of England Staff Working Paper No. 845.

- Tamura, Robert**, “From decay to growth: A demographic transition to economic growth,” *Journal of Economic Dynamics and Control*, 1996, *20*, 1237–61.
- , “Human capital and economic development,” *Journal of Development Economics*, 2006, *79*, 26–72.
- Teulings, Coen and Richard Baldwin**, eds, *Secular Stagnation: Facts, Causes and Cures*, CEPR Press, 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.htm>s.
- 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, E. A., R. S. Davies, J. E. Oeppen, and R. S. Schofield**, *English Population History from Family Reconstitution 1580–1837*, Cambridge University Press, 1997.

ONLINE APPENDIX

A Data Appendix	45
A.1 Detail on interest rate data	45
A.2 The German Socio-Economic Panel	50
A.3 Steady state consumption share	54
A.4 Calibrating the beta distribution	54
B Equation (1): derivation and discussion	57
B.1 Derivation	57
B.2 Discussion	59
C Model derivations	60
C.1 Steady State	60
C.2 Model solution	61
D Surviving children and bequests	64
E Asymptotic results	65
E.1 Proof of Theorem 1	65
E.2 Asymptotic expression for the rate of interest	68

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. 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.²³

Safe or risk-free rate The main measure introduced by Schmelzing (2020), and the real rate used in our paper in our Figure 1(i), 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 2. 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 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

²³For expositional ease, all results in the section are presented as 50-year averages of generational rates of return.

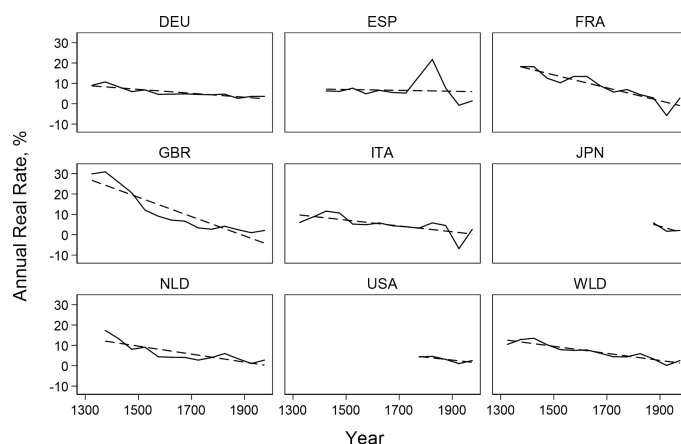
Table 2: Details of Schmelzing’s Global ‘Safe Rate’.

Period	Country	Type of Assets	Justification for:	
			Start Date	End Date
1311-1509	Italy	Venetian <i>Prestisi</i> and Genoese <i>Luoghi</i> . 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	<i>Juros</i> 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 (<i>Renten</i> 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	-

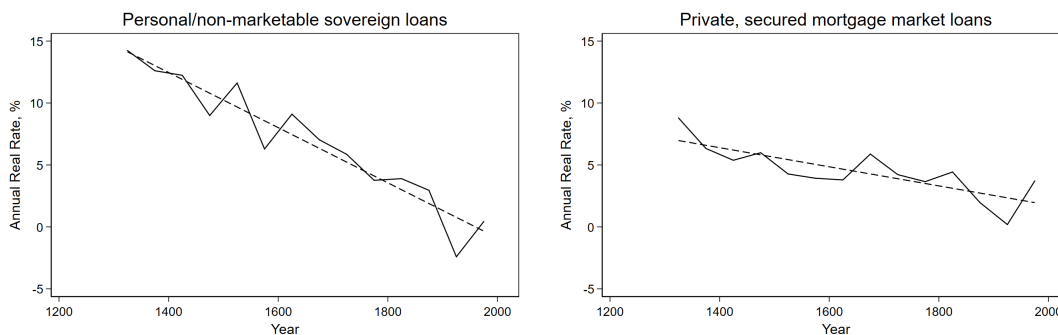
in the global 'safe' series, the debt instruments remain the same and consist of the sovereign debt discussed above. For countries and/or periods not covered in the 'safe' series, observations are arithmetically weighted on the country-level across data points of long-term consolidated debt (such as 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 12i., 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 12i., the global real rate of return is steadily declining over the entire period.

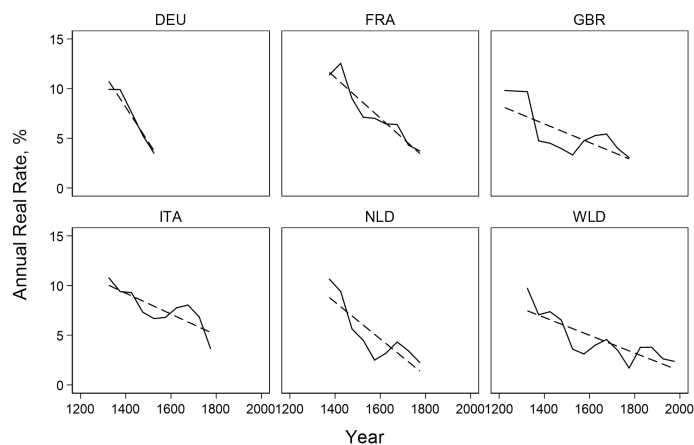
'Personal' or 'Sovereign' non-marketable loans Schmelzing also examines the extent to which the non-marketability of loans can account for the decline of interest rates presented above by examining personal loans to sovereigns (including "pledge loans" and loans from municipalities to the central authorities). 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 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



i. Rates of Return on long-term debt.



ii. Rates of return on personal/non-marketable loans to sovereigns and private debt.



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

Figure 12: Country specific real rates of return on long-term debt and land. Dashed line show regression trends.

inflation rates from Allen (2001). The results are shown in the first panel of Figure 12ii.; 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 12ii. 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 12iii. show the real rates of return on land – arguable the ‘safest asset’ – for 5 countries (Italy, U.K., Flanders, France and Germany).²⁴ 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 12iii.. The high interest rates in 13th century England that can be seen shown in Figure 1ii. are echoed across northern 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

²⁴The 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.

longer-run trend from ancient data, as shown in Table A.1.

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 Euro- pean	9.43	Risk-free rental rate ^c
C16th AD	Ottoman Empire	10–20	Interest on loans ^a
C19th AD	Various Euro- pean	3.43	Risk-free rental rate ^c
2000 AD	England	4–5	Return on land, rent charges ^a
2000–17 AD	Various Euro- pean	1.24	Return on land, rent charges ^c

Notes: ^aCalculated or referenced in Clark (2007b). ^bHudson (2000).
^cSchmelzing (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 (17) 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 $N_{t+1}^i = N_t^i n_t^i$, we can re-write (17) in terms of the number of children each household has as simple $n_t^i = \beta^i \tilde{R}_{t+1}$. Motivated by this simple relationship, we estimate the following specification,

$$children_{i,2013} = \beta_0 + \beta_1 patience_{i,2008} + \mathbf{X}'_i \boldsymbol{\beta} + \varepsilon_i \quad (33)$$

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 3 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 these specifications, the coefficient on patience is statistically significant and of the expected sign. Table 4 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 3: Patience and Children

VARIABLES	(1)	(2)	(3)	(4)	(5)
	totalChildren	totalChildren	totalChildren	totalChildren	totalChildren
HHpatience	0.027** (0.010)	0.013*** (0.004)	0.017*** (0.004)	0.012*** (0.004)	0.022*** (0.009)
HHage		-0.024*** (0.001)	-0.021*** (0.001)	-0.030*** (0.001)	0.017*** (0.005)
lincome			0.414*** (0.016)	0.274*** (0.017)	0.175*** (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

*** p<0.01, ** p<0.05, * p<0.1

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.

Table 4: Patience and Children: Age bins

VARIABLES	(1) totalChildren	(2) totalChildren	(3) totalChildren
HHpatience	0.010** (0.004)	0.016*** (0.004)	0.014*** (0.004)
mediumyoung	0.573*** (0.061)	0.272*** (0.062)	0.146*** (0.056)
mediumold	0.884*** (0.057)	0.471*** (0.060)	0.199*** (0.058)
old	-0.056 (0.050)	-0.362*** (0.052)	-0.729*** (0.055)
lincome		0.420*** (0.017)	0.312*** (0.017)
Observations	17,224	17,222	17,222
R^2	0.181	0.259	0.317
Controls	yes	yes	yes

*** p<0.01, ** p<0.05, * p<0.1

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 $25 < HHage \leq 35$; mediumold is a dummy equal to 1 if $35 < HHage \leq 45$; and, mediumyoung 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 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_{ss}^c 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.4 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 $\beta = \frac{1}{1+r}$ in equilibrium, we need to express $\text{var}\left(\frac{1}{1+r}\right)$ as a function of the mean $E(r)$ and variance $\text{var}(r)$. We use a first-order Taylor expansion of the second moment of the transformed variable to find $\text{var}\left(\frac{1}{1+r}\right) = \frac{1}{(1+E(r))^4} \text{var}_t(r)$. 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 mean standard deviation across countries is 0.005.

The long run interest rate To find data on the long run interest rates we use the Credit Suisse Global Investment Returns Yearbook (Elroy Dimson and Staunton, 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

Table 5: 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

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, $R_{c,t}^s$. We then use this to calculate both the annual rate of return ($r_{c,t}^s$) and the annualized 25-year generational rate of return ($\bar{r}_{c,t}^s$) as:

$$r_{c,t+1}^s = \left(\frac{R_{c,t+1}^s}{R_{c,t}^s} \right) - 1, \quad (34)$$

and

$$\bar{r}_{c,t+25}^s = \left(\frac{R_{c,t+25}^s}{R_{c,t}^s} \right)^{\frac{1}{25}} - 1. \quad (35)$$

Tables 5 and 6 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, $\bar{r}_{W,t}^s$, in every year and then find the average of the implied rates of return between 1975 and 2015 which is equal to annualized 6.3%.

Table 6: 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

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,

$$\tilde{r}_t^L = \gamma g_t - \frac{\gamma^2}{2} \sigma_t^2 - \ln \beta + \gamma d_{L,t}. \quad (36)$$

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 ,

$$\max_{L_t, B_t} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t U(C_t) \quad (37)$$

subject to,

$$L_{t+1} + B_{t+1} = R_t^L L_t + R_t^f B_t + W_t - C_t \quad (38)$$

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,

$$R_{t+1}^f \mathbb{E}_t \frac{\beta U'(C_{t+1})}{U'(C_t)} = 1, \quad (39)$$

$$\mathbb{E}_t R_{t+1}^L \frac{\beta U'(C_{t+1})}{U'(C_t)} = 1. \quad (40)$$

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

First, we impose CRRA utility of the form,

$$U(C_t) = \frac{1}{1-\gamma} C_t^{1-\gamma}, \quad (41)$$

and so the optimal portfolio satisfies,

$$R_{t+1}^f \mathbb{E}_t \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} = 1, \quad (42)$$

$$\mathbb{E}_t R_{t+1}^L \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} = 1. \quad (43)$$

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,

$$\begin{bmatrix} g_{t+1} \\ r_{t+1}^L \end{bmatrix} \sim N \left(\begin{bmatrix} \bar{g}_{t+1} \\ \bar{r}_{t+1}^L \end{bmatrix}, \begin{bmatrix} \sigma_{g,t}^2, \sigma_{g,L,t}^2 \\ \sigma_{g,L,t}^2, \sigma_{L,t}^2 \end{bmatrix} \right). \quad (44)$$

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,

$$\beta \exp \left\{ r_{t+1}^f - \gamma \bar{g}_{t+1} + \frac{1}{2} \text{var}_t (-\gamma g_{t+1}) \right\} = 1 \quad (45)$$

$$\beta \exp \left\{ \bar{r}_{t+1}^L - \gamma \bar{g}_{t+1} + \frac{1}{2} \text{var}_t (r_{t+1}^L - \gamma g_{t+1}) \right\} = 1. \quad (46)$$

Note that from (45) we have the following expression for the real rate,

$$r_t^f = \gamma g_t - \frac{\gamma^2}{2} \sigma_{g,t}^2 - \ln \beta. \quad (47)$$

where with log utility ($\gamma \rightarrow 1$) 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 ,

$$\bar{r}_{t+1}^L + \frac{1}{2} \sigma_{L,t+1}^2 = r_{t+1}^f + \gamma \sigma_{g,L,t+1}^2 \quad (48)$$

Note that $\bar{r}_{t+1}^L = \mathbb{E}_t r_{t+1}^L$ and, since r_t^L is Normally distributed, we can write $\ln \mathbb{E}_t R_{t+1}^L = \bar{r}_{t+1}^L + \frac{1}{2}\sigma_L^2$ and so,

$$\ln \mathbb{E}_{t-1} R_t^L = r_t^f + \gamma \sigma_{g,L,t}^2 \quad (49)$$

which, with $\tilde{r}_t^L = \ln \mathbb{E}_{t-1} R_t^L$ and $d_{L,t} = \sigma_{g,L,t}^2$, is the expression given in equation (36).

B.2 Discussion

As we discussed in the paper, the historical record for per capita growth and life expectancy are unable to explain the fall in rates over time. Equation (36) suggests a number of additional potential channels.

Variance of consumption growth If the variance of consumption growth ($\sigma_{g,t}^2$) 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 (36), 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.²⁵ 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 Model derivations

C.1 Steady State

Denoting steady state values as N_{ss} , etc. we have:

$$N_{ss}^I = N_{ss} \text{ and } N_{ss}^i = 0 \quad \forall i < I \quad (50)$$

$$K_{ss}^I = K_{ss} = \bar{K} \text{ and } K_{ss}^i = 0 \quad \forall i < I \quad (51)$$

²⁵Importantly 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.

$$C_{ss}^I = C_{ss} \text{ and } C_{ss}^i = 0 \quad \forall i < I. \quad (52)$$

Using the above with the first order conditions and budget constraints (7)-(9), along with the firm's first order conditions (10) and (11), it follows that the steady state is characterized by:

$$N_{ss} = \left(\frac{D(1 - \alpha - \beta^I + \alpha\beta^I(1 - \nu))}{(1 - \pi(1 - \alpha))(1 - \beta^I)} \right)^{\frac{1}{\nu}} \bar{K} \quad (53)$$

$$C_{ss} = (D\bar{K}^\nu N_{ss}^{-\nu} + \pi - 1)N_{ss} \quad (54)$$

$$Y_{ss} = D\bar{K}^\nu N_{ss}^{1-\nu} \quad (55)$$

$$p_{ss} = \nu \frac{\beta^I}{1 - \beta^I} D\bar{K}^{\nu-1} N_{ss}^{1-\nu} \quad (56)$$

$$w_{ss} = (1 - \nu)D\bar{K}^\nu N_{ss}^{-\nu} \quad (57)$$

$$r_{ss} = \nu D\bar{K}^{\nu-1} N_{ss}^{1-\nu}. \quad (58)$$

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 .

C.2 Model solution

The following expands on elements of the model solution, as described in Sections 3-4.

Household Problem We can re-write the household consumer maximization problem (4) by substituting out for $N_{c,t}^i$ and X_t^i so that the problem for each dynasty i becomes:

$$\max_{C_t^i, K_{t+1}^i, N_{t+1}^i} \sum_{t=0}^{\infty} (\beta^i)^t (\alpha \log(C_t^i) + (1 - \alpha - \beta^i) \log(N_{t+1}^i)) \quad (59)$$

$$C_t^i + N_{t+1}^i + p_t K_{t+1}^i \leq (w_t + \pi)N_t^i + (r_t + p_t)K_t^i. \quad (60)$$

The first order conditions for this problem are given by:

$$\lambda_t^i = \frac{\alpha(\beta^i)^t}{C_t^i}, \quad (61)$$

$$\frac{(1 - \alpha - \beta^i)(\beta^i)^t}{N_{t+1}^i} + (\pi + w_{t+1})\lambda_{t+1}^i = \lambda_t^i \quad (62)$$

$$p_t \lambda_t^i = (p_{t+1} + r_{t+1})\lambda_{t+1}^i, \quad (63)$$

where, λ_t^i is the Lagrange multiplier on the constraint (60). Now, substituting out for λ_t^i in the last two FOCs using the first FOC, we obtain:

$$\frac{(1 - \alpha - \beta^i)}{N_{t+1}^i} + (\pi + w_{t+1})\frac{\alpha\beta^i}{C_{t+1}^i} = \frac{\alpha}{C_t^i} \quad (64)$$

and

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i \frac{p_{t+1} + r_{t+1}}{p_t}. \quad (65)$$

The above hold for all $t \geq 0$ and for all i . Defining $R_{t+1} \equiv \frac{p_{t+1} + r_{t+1}}{p_t}$ we obtain equation (16) in the main text.

Firm Problem From the firm's problem in (5) we obtain the following first order conditions for all $t \geq 0$:

$$w_t = (1 - \alpha)DK_t^\alpha N_t^{-\alpha} \quad (66)$$

and

$$r_t = \alpha DK_t^{\alpha-1} N_t^{1-\alpha}. \quad (67)$$

Population Euler Equation To derive equation (17) we proceed as follows. We re-write FOC (7) as

$$N_{t+1}^i = \frac{(1 - \alpha - \beta^i)}{\alpha \left(\frac{C_{t+1}^i}{C_t^i} - \pi\beta^i - \beta^i w_{t+1} \right)} C_{t+1}^i,$$

and use the Euler Equation, (65), to substitute out for $\frac{C_{t+1}^i}{C_t^i}$ to obtain an expression for N_{t+1}^i :

$$N_{t+1}^i = \frac{(1 - \alpha - \beta^i)}{\alpha\beta^i (R_{t+1} - \pi - w_{t+1})} C_{t+1}^i.$$

Bringing the above equation forward one period in time we obtain:

$$N_{t+2}^i = \frac{(1 - \alpha - \beta^i)}{\alpha \beta^i (R_{t+2} - \pi - w_{t+2})} C_{t+2}^i.$$

Taking the ratio of these two equations and substituting for $\frac{C_{t+2}^i}{C_{t+1}^i}$ from the Euler equation, (65), we obtain:

$$\frac{N_{t+2}^i}{N_{t+1}^i} = \beta^i \tilde{R}_{t+2}, \quad (68)$$

where in the above $\tilde{R}_{t+2} \equiv R_{t+2} \frac{R_{t+1} - (w_{t+1} + \pi)}{R_{t+2} - (w_{t+2} + \pi)}$. The above equation holds for all $t \geq 0$. We can also re-write it as:

$$\frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, \quad (69)$$

where in the above $\tilde{R}_{t+1} \equiv R_{t+1} \frac{R_t - (w_t + \pi)}{R_{t+1} - (w_{t+1} + \pi)}$, as long as $t \geq 1$. This is equation (17) in the main text.

Initial Population and Consumption To obtain equation (25) in the main text, we plug in equation (19) into (7).

$$\frac{(1 - \alpha - \beta^i)}{\frac{(\beta^i)^t N_1^i}{\sum_{j=1}^I (\beta^j)^t N_1^j}} N_{t+1} + (\pi + w_{t+1}) \frac{\alpha \beta^i}{\frac{(\beta^i)^{t+1} C_0^i}{\sum_{j=1}^I (\beta^j)^{t+1} C_0^j} C_{t+1}} = \frac{\alpha}{\frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I (\beta^j)^t C_0^j} C_t}, \quad (70)$$

Simplifying and re-writing this expression relative to the highest discount factor among agents results in:

$$\frac{(1 - \alpha - \beta^i)}{\frac{N_1^i}{\sum_{j=1}^I (\frac{\beta^j}{\beta^I})^t N_1^j}} N_{t+1} + (\pi + w_{t+1}) \frac{\alpha}{\frac{C_0^i}{\beta^I \sum_{j=1}^I (\frac{\beta^j}{\beta^I})^{t+1} C_0^j} C_{t+1}} = \frac{\alpha}{\frac{C_0^i}{\sum_{j=1}^I (\frac{\beta^j}{\beta^I})^t C_0^j} C_t}, \quad (71)$$

Now as $t \rightarrow \infty$ the above equation becomes:

$$\frac{(1 - \alpha - \beta^i)}{\frac{N_1^i}{N_1^I} N_{ss}} + (\pi + w_{ss}) \frac{\alpha}{\frac{C_0^i}{\beta^I C_0^I} C_{ss}} = \frac{\alpha}{\frac{C_0^i}{C_0^I} C_{ss}}. \quad (72)$$

Then, substituting from the solutions of the steady state shown in equations (53)-(58) into the above, for each $i < I$ we can then show that:

$$\frac{C_0^i}{C_0^I} = \frac{N_1^i}{N_1^I} \frac{1 - \alpha - \beta^I}{1 - \alpha - \beta^i}. \quad (73)$$

Relating Dynasty to Aggregate Measures To obtain equation (20) in the text, simply substitute (73) into (19).

D Surviving children and bequests

In Section 6.1 we compare the model to the testamentary data in Clark (2007a). In order for us to do this, this appendix calculates the expected number of surviving children and the expected value of bequest to those children at parents' expected time of death in the model. For each adult, there is a probability $(1 - \pi)$ of death each period. Each child born to that adult more than one period ago has the same periodic probability of death. We call an agent who becomes a parent in period t (i.e., that was born in period $t - 1$), a t -parent. For each t -parent in each dynasty i , the expected number of children surviving at the point of his death is,

$$\mathbb{E} [s^i(t)] \equiv \sum_{j=1}^{\infty} p(t, j) s^i(t, j), \quad (74)$$

where $p(t, j)$ is the probability of a t -parent dying j periods later and $s^i(t, j)$ is the expected number of children of a t -parent that survive to that point of death. For $j \geq 1$, we have,

$$p(t, j) \equiv \pi^{j-1} (1 - \pi). \quad (75)$$

The t -parent had a sequence of children in each period from t to $t + j$, each of which children, after one period, has a probability of survival to $t + j$. I.e., if a t -parent survives for two periods ($j = 2$), he has had $n_{c,t}^i$ children in period t and $n_{c,t+1}^i$ in period $t + 1$. The children born in t , survive to $t + 2$ with probability π ; the children born in period $t + 1$ survive to $t + 2$ with probability 1. For each $j \geq 1$, we have the expected number of surviving children of a t -parent,

$$s^i(t, j) \equiv \sum_{k=0}^{j-1} \pi^k n_{c,t+j-(k+1)}^i, \quad (76)$$

where $n_{c,t+j-(k+1)}^i$ is the number of children born $k + 1$ periods before $t + j$.

Given the above we can calculate the expected number of surviving children of a

parent born in time t , $s^i(t)$, at their expected time of death:

$$\mathbb{E} [s^i(t)] = \frac{1}{1 + \pi} \sum_{m=0}^{\infty} \pi^m n_{c,t+m}^i. \quad (77)$$

Note that if $n_{c,t+m}^i$ were constant over time, $n_{c,t+m}^i = \bar{n}_c^i$ for all $m > 0$, then the this equation reduces to

$$\mathbb{E} [s^i(t)] = \frac{1}{1 - \pi^2} \bar{n}_c^i. \quad (78)$$

We can also calculate the total bequests of a t -parent, $b^i(t)$, at their expected time of death (and only at that time):

$$\mathbb{E} [b^i(t)] = (1 - \pi) \sum_{m=0}^{\infty} \pi^m k_{t+m+1}^i n_{t+m+1}^i. \quad (79)$$

Total bequests of t -parent at time $t + j$ (and only at that time), for $j \geq 1$:

$$b^i(t, j) \equiv n_{t+j}^i k_{t+j}^i \quad (80)$$

E Asymptotic results

E.1 Proof of Theorem 1

In the baseline calibration of the model we assumed a discrete number of types of agents. In this section, we consider what happens when the number of types of agents approaches infinity, in order to prove Theorem 1.

Theorem 1. *If $I \rightarrow \infty$ and dynastic discount factors 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. Suppose that there are n dynasties with discount factors, β^i , distributed evenly along a grid so that $\beta(i; n) = \frac{2i-1}{2n}$ for $i = 1, \dots, n$. Notice that the distance between any two points is simply: $\Delta(n) \equiv \beta(i+1; n) - \beta(i; n) = \frac{1}{n}$. We define the following function: $\nu_t(\beta(i; n)) \equiv \frac{N_t^i}{N^i}$, 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 this function as a probability mass function of a discrete random variable

with realization, $\beta(i; n)$, on the domain $\{\frac{2i-1}{2n} | i = 1, \dots, n\}$. We wish to characterize the evolution of the asymptotic function, $\frac{\nu_t(\beta(i; n))}{\Delta(n)}$, over time as $n \rightarrow \infty$ - that is as the number of dynasties or types becomes infinite. 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.

For each agent i , we can re-write equation (17) as:

$$N_{t+1}^i = \beta^i \tilde{R}_{t+1} N_t^i. \quad (81)$$

Summing these expressions over all agents, we obtain the following, $N_{t+1} = \beta^i \tilde{R}_{t+1} \sum_{j=1}^n \beta^j N_t^j$, which can also be written as:

$$N_{t+1} = \beta^i \tilde{R}_{t+1} N_{t+1} \sum_{j=1}^n \beta^j \nu_t^n(\beta^j). \quad (82)$$

Dividing equation (81) by equation (82) we obtain:

$$\nu_{t+1}^n(\beta^i) = \frac{\beta^i \nu_t^n(\beta^i)}{\sum_{j=1}^n \beta^j \nu_t^n(\beta^j)}. \quad (83)$$

This recursive formulation defines the evolution of the probability mass function over time. We are interested in the properties of this function as $n \rightarrow \infty$. To aid us in this investigation, notice that the cumulative distribution function of β^i at time t for a grid of size n is:

$$F_t^n(\beta^i) \equiv \frac{\sum_{j=1}^i \beta^j \nu_t^n(\beta^j)}{\sum_{j=1}^n \beta^j \nu_t^n(\beta^j)}. \quad (84)$$

This also means that:

$$\nu_t^n(\beta^i) = F_t^n(\beta^{i+1}) - F_t^n(\beta^i) = P_t^n(\beta^i \leq \beta \leq \beta^{i+1}). \quad (85)$$

Given the above, notice that (83) can be re-written as:

$$\frac{\nu_{t+1}^n(\beta^i)}{\Delta^i(n)} = \frac{\beta^i \frac{\nu_t^n(\beta^i)}{\Delta^i(n)}}{\sum_{j=1}^n \beta^j P_t^n(\beta^j \leq \beta \leq \beta^{j+1})}. \quad (86)$$

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

$$f_{t+1}(\beta) = \frac{\beta f_t(\beta)}{E_t(\beta)}, \quad (87)$$

where f_t is the continuous probability density function corresponding to the discrete mass function ν_t^n ²⁶ and $E_t(\beta) \equiv \int_0^1 u f_t(u) du = \lim_{n \rightarrow \infty} \sum_{j=1}^n \beta^j P_t^n(\beta^j \leq \beta \leq \beta^{j+1})$, is simply the mean of the corresponding continuous random variable. Notice that the above functional equation describes the evolution of the distribution of the limit function over time. It is easy to show that a time invariant solution $f(\beta)$ of the above does not exist (see appendix). Instead, we are interested in a solution that takes the following form $f_t(\beta) \equiv f(\beta; \boldsymbol{\theta}_t)$, where $\boldsymbol{\theta}_t$ is a vector of potentially time varying parameters of the distribution f . In other words, we are looking for a solution to the above that remains of a fixed type, with only its parameters changing.

Below, we show that one solution to the above functional equation is the scaled beta distribution defined on $(0, \bar{\beta})$ with cumulative distribution function, $F(\cdot)$ given in the main body of the text in equation (28). The corresponding probability density function of this distribution f is given by:

$$f_t(\beta; \boldsymbol{\theta}_t) \equiv f(\beta; \gamma_t, \delta_t) = \frac{(\bar{\beta} - \beta)^{\delta_t - 1} \beta^{\gamma_t - 1}}{\bar{\beta}^{\delta_t + \gamma_t - 1} B(\gamma_t, \delta_t)}, \quad (88)$$

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

$$E(\beta; \gamma_t, \delta_t) = \bar{\beta} \frac{\gamma_t}{\gamma_t + \delta_t}. \quad (89)$$

²⁶To see this, notice that $\lim_{n \rightarrow \infty} \frac{\nu_t(\beta(i;n))}{\Delta(n)} = \lim_{n \rightarrow \infty} \frac{F_t(\beta(i+1;n)) - F_t(\beta(i;n))}{\beta(i+1;n) - \beta(i;n)} = \lim_{n \rightarrow \infty} \frac{F_t(\beta(i;n) + \Delta(n)) - F_t(\beta(i;n))}{\Delta(n)} = F_t'(\beta(i;n))$

Using equations (87)-(89), we can write the pdf of discount factors at time $t + 1$ as:

$$\begin{aligned}
f_{t+1}(\beta; \gamma_t, \delta_t) &= \frac{\beta(\bar{\beta} - \beta)^{\delta_t-1} \beta^{\gamma_t-1}}{\bar{\beta}^{\frac{\gamma_t}{\gamma_t+\delta_t}} \bar{\beta}^{\delta_t+\gamma_t-1} B(\gamma_t, \delta_t)} \\
&= \frac{(\bar{\beta} - \beta)^{\delta_t-1} \beta^{\gamma_t}}{\frac{\gamma_t}{\gamma_t+\delta_t} \bar{\beta}^{\delta_t+\gamma_t} B(\gamma_t, \delta_t)} \\
&= \frac{(\bar{\beta} - \beta)^{\delta_t-1} \beta^{\gamma_t}}{\bar{\beta}^{\delta_t+\gamma_t} B(\gamma_t + 1, \delta_t)} \\
&= f(\beta; \gamma_{t+1}, \delta_{t+1})
\end{aligned} \tag{90}$$

where, $\gamma_{t+1} = 1 + \gamma_t$ and $\delta_{t+1} = \delta_t \equiv \delta$. The second 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 (87) is the beta distribution with parameters given by $\gamma_{t+1} = 1 + \gamma_t$ and $\delta_t \equiv \delta$.

□

E.2 Asymptotic expression for the rate of interest

In the model, the mean discount factor influences the interest rate. Recall that

$$R_{t+1} = \frac{C_{t+1}^i / C_t^i}{\beta^i} = \frac{\left(\frac{\kappa_{t+1}^I(\beta^i) / \Delta(I)}{\kappa_t^I(\beta^i) / \Delta(I)} \right) \frac{C_{t+1}}{C_t}}{\beta^i} \tag{91}$$

where $\kappa_t^I(\beta^i) \equiv C_t^i / C_t$. Note also that we can write:

$$\frac{\kappa_t^I(\beta^i)}{\Delta(I)} = \frac{\frac{\beta^i}{1-\alpha-\beta^i} \frac{\nu_t^I(\beta^i)}{\Delta(I)}}{\sum_{j=1}^I \frac{\beta^j}{1-\alpha-\beta^j} \nu_t^I(\beta^j)}. \tag{92}$$

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

$$f_{ct}(\beta) = \frac{\frac{\beta}{1-\alpha-\beta} f_t(\beta)}{E_t\left(\frac{\beta}{1-\alpha-\beta}\right)}, \tag{93}$$

where f_t and f_{ct} are the continuous probability density function corresponding to the discrete mass functions ν_t^I and κ_t^I . Note also that using the relationship derived

between $f_{t+1}(\beta)$ and $f_t(\beta)$ in the Appendix we have the following expression:

$$\frac{f_{ct+1}(\beta)}{f_{ct}(\beta)} = \beta \frac{E_t(\beta/(\bar{\beta} - \beta))}{E_t(\beta^2/(\bar{\beta} - \beta))} \quad (94)$$

Taking the limit of both sides of (91) as $I \rightarrow \infty$ we obtain:

$$R_{t+1} = \frac{E_t(\beta/(\bar{\beta} - \beta))}{E_t(\beta^2/(\bar{\beta} - \beta))} \frac{C_{t+1}}{C_t}. \quad (95)$$

Note that over time the growth rate of aggregate consumption converges to 1. In particular for high enough t the approximation $\frac{C_{t+1}}{C_t} \approx 1$ holds. Consequently, we can write the following expression for mean generational gross interest rates for high enough t :

$$R_{t+1} \approx \frac{E_t(\beta/(\bar{\beta} - \beta))}{E_t(\beta^2/(\bar{\beta} - \beta))}. \quad (96)$$

If we assume that the discount factors follow a beta distribution, then for high enough t we can write the annualized gross interest rate as:

$$R_{t+1}^{\frac{1}{25}} \approx \left(\frac{\gamma_t + \delta_t}{\bar{\beta}(1 + \gamma_t)} \right)^{\frac{1}{25}}. \quad (97)$$