School of Economics and Finance Discussion Papers

Patience Breeds Interest: The Rise of Societal Patience and the Fall of the Risk-free Interest Rate

Radoslaw Stefanski and Alex Trew

School of Economics and Finance Discussion Paper No. 2002
8 Feb 2020
JEL Classification: E21,E43,J11,N30,O11
Keywords: Heterogeneous agents; interest rates; patience; selection

University of St Andrews
School of Economics and Finance
Online Discussion Paper Series
issn 2055-303X
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

Patience Breeds Interest: The Rise of Societal Patience and the Fall of the Risk-free Interest Rate[1]

Radoslaw (Radek) Stefanski[2]
University of St. Andrews and University of Oxford (OxCarre)

Alex Trew[3]
University of Glasgow

February 8, 2020

Abstract

The risk-free rate of return has been declining in real terms over millennia. We isolate the role of time preference – or patience – in explaining this decline. Three facts support our approach: experimental evidence finds significant heterogeneity in patience; individual preference characteristics are highly intergenerationally persistent; and, longitudinal data shows that patience is positively related with fertility decisions. Together these suggest we should expect average *societal* levels of patience to increase over time as the composition of the population shifts towards ever more patient dynasties. We test this mechanism in a Barro-Becker model of fertility with heterogeneous dynasties. We use the present day distribution of patience to calibrate the model. We are able match – both quantitatively and qualitatively – the decline in the risk-free return over the last eight centuries.

JEL codes: E21; E43; J11; N30; O11.
Keywords: Heterogeneous agents; interest rates; patience; selection.

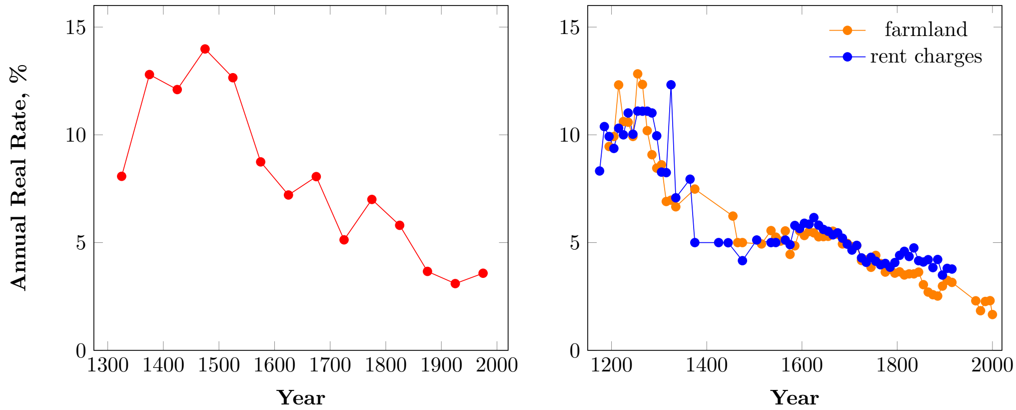
[1]We would like to thank Rod McCrorie, Byeongju Jeong, Rick van der Ploeg, Ctirad Slavik, Tony Smith as well as the participants of seminars at the University of Oxford, the University of St. Andrews, CERGE-EI and conference participants at WEHIA 2019 for their comments and suggestions.

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

[3]Adam Smith Business School, University of Glasgow, Glasgow, G12 8QQ, United Kingdom. **Email address**: alex.trew@glasgow.ac.uk.

1 Introduction

Risk-free interest rates have been falling for millennia. Successive ancient civilizations have been characterized by ever lower rates of interest. From 20–25% per annum in Sumer to 10-20% in Babylon and 9–12% in 1st century AD Egypt (see Appendix Table 2). Evidence for the last eight centuries follows a similar pattern. Figure 1 shows that the global risk-free real interest rate has declined from around 12% in the fourteenth century to just over 1% today (Schmelzing, 2017). Within England, the return on land, one of the safest assets available, fell from around 10% in the thirteenth century to 1-2% today (Clark, 2010).



(a) Global real risk-free rate (Schmelzing (2017)) (b) Real return on land (Clark (2010))

Figure 1: Real risk-free returns, 1175–2000

The list of potential explanations for this decline is short: the risk-free interest rate is a function of the expected rate or growth, the level of risk in an economy and individual time- and risk-preferences. In a standard representative-agent endowment economy where individuals have CRRA preferences and face log-normal shocks to growth, we obtain the following expression for the risk-free rate (see Appendix C):

$$r_t^f = \gamma \bar{g}_t - \frac{\gamma^2}{2}\sigma_t^2 - \ln \beta, \tag{1}$$

where r_t^f is the log of the risk-free rate, $\bar{g}_t$ is the average growth in consumption, σ_t^2 is the variance of consumption growth, β is the subjective discount factor and γ is the coefficient of relative risk aversion.[4]

We naturally think of risk and growth as time-varying. That is, the decline in the rate of interest may be explained by at least one of either a steady decline in expected growth or by a continual increase in risk over time. Neither of these explanations seem plausible. First, the

[4]Given our use of historical data, we may be concerned that our observed interest rates are not rates on assets that are completely risk-free, where the amount of risk in that asset is itself time-varying. See Appendix C.

evidence on per capita incomes up until around 1750 is that they were flat, and potentially growing slowly in some countries (The Maddison Project, 2013). After 1750 growth rates rose sharply with the start of the industrial revolution. Second, shocks to consumption, assets and production have either remained stable or declined over time. Climate variability has been relatively constant over the last millennium up until the 20th century, with air temperature exhibiting an average variation of 0.5 degrees centigrade (Salinger, 2005). Levels of violence and warfare have also systematically declined (Pinker, 2012). Furthermore, the development of technologies to understand probabilities and the emergence of sophisticated insurance markets have improved the resilience of agents to shocks (Bernstein, 1998).

So the directions of change in expected growth and volatility have, if anything, made the declining risk-free rate *harder* to explain. The remaining components of the equilibrium risk-free rate are preference parameters: the discount factor and the coefficient of relative risk aversion. We would not normally think of these parameters as changing over time at the individual level. However, in an economy populated by heterogeneous agents that have different preference parameters, we may think about the evolution of average *societal* patience and risk aversion if those parameters are partly inherited and if they also affect fertility. If, for example, more patient types have more children that are themselves also more patient, then the share of more patient types in an economy will grow over time, causing the average level of patience to increase.

The relationship between relative risk aversion and the risk-free rate depends, by (1), on the sign of $\left(\bar{g}_t - \gamma\sigma^2\right)$. The available 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 the declining risk-free rate may be explained by risk aversion falling over time. The evidence is, however, that risk aversion has, if anything, emerged and grown over time as an evolutionary adaptation (Robson, 1996; Levy, 2015).

We are left with the discount factor as a potential explanation for the decline in the risk-free rate. If patience is indeed heterogeneous, if it is inter-generationally persistent, and if it is related to fertility, then we may indeed think of a time-varying societal level of patience, β_t, which increases over time and in turn drives falling risk-free interest rates.

Evidence on patience First, modern empirical studies find that there is 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.

Third, while there are theoretical reasons (as we discuss below) to think households with higher patience will optimally choose to have more children, the empirical evidence on the connection is more limited. In Appendix B, 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 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.

Related literature A number of other papers have identified importance of the long-run evolution of patience. Using testamentary data for Suffolk (England), Clark (2007) shows that families at the turn of the seventeenth century with more wealth tended to have more children. Moreover, records on Royal tenants (whose wealth would have been greater than average), suggest that a relationship between wealth and fertility goes back at least to the mid-thirteenth century. Thus, for Clark, the economy evolved in a Darwinian 'survival of the richest'. In an economy that increasingly valued literacy, numeracy and patience, this evolution laid the foundations for the later industrial revolution. For Clark, however, 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 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'.

Galor and Özak (2016) present a model in which, similar to Clark (2007), higher patience leads to better economic outcomes and, by consequence, 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 over time a larger share of long-term orientated individuals. 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.

Our model also 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.

Finally, our work also relates to research on whether the decline in the risk-free rate 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 risk-free rate that slows over time and is thus hard to discern even in data since 1870.

Structure The rest of the paper is structured as follows. In section 2 we develop a Barro-Becker model of fertility with heterogenous dynasties that differ according to their discount factor, while section 3 presents the solution of the model. Section 4 calibrates the model to existing modern data on the distribution of patience and section 5 presents quantitative results. Section 6 offers some concluding remarks.

2 Baseline Model

Consider an economy of I dynasties, indexed by $i = 1, \ldots, I$, each populated by N_t^i households at time t. Households within a dynasty are identical, but dynasties differ in their discount factors, β^i. Without loss of generality, the sequence $\{\beta^i\}_{i=1}^I$ is strictly increasing in i, so dynasty I has the greatest discount factor β^I. Each household[5] is endowed each period with a unit of labor that it inelastically provides on the market in exchange for a wage, w_t, as well as some capital stock (or land), k_t^i, that it inherited from its parent and that it rents out on the market in exchange for a rental rate, r_t. Each household of type i then solves the following problem in every period

[5]Since households within a dynasty are identical, and since we obtain solutions to the model in terms of dynasty-aggregates, we omit a household index here. 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.

t:

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

$$c_t^i + n_{c,t}^i + p_t x_t^i \leq w_t + r_t k_t^i$$

$$n_{t+1}^i = \pi + n_{c,t}^i$$

$$k_{t+1}^i = \frac{k_t^i + x_t^i}{n_{t+1}^i}.$$

As in Barro and Becker (1988, 1989), households derive utility from their own consumption, c_t^i, the expected size of the household at the end of the current and the beginning of the next period, n_{t+1}^i, and children's average continuation utility, $U_{t+1}^i(k_{t+1}^i)$. This particular choice of utility function follows Lucas (1) 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 (this can be readily modified). The price of purchasing capital stock is given by p_t. We also assume that the survival probability for existing households is age independent, constant across dynasties and exogenous and given by π and that the survival probability of children is 1 (this can easily 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 equally about their children and endow them with an equal share of accumulated capital. Thus, parents face a quantity-quality tradeoff with respect to the number of children a la Barro and Becker (1988, 1989). Finally, we also assume that the child of an adult in dynasty i perfectly inherit the discount factor β^i. 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, is left as a reduced form assumption.

At time zero, there are N_0^i identical members of the dynasty of type i. Each individual in the dynasty chooses the number of children, $n_{c,t}$, all of which survive. In the subsequent period the number of households in a dynasty will depend on the number of children that each household in the dynasty chose to have as well as the expected number of surviving adults. The number of people in dynasty i at time $t+1$ will be given by $N_{t+1}^i = n_{t+1}^i N_t^i$.

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:

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

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

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. We know that the total number of people of each type evolves according to: $N_{t+1}^i = N_t^i n_{t+1}^i$. 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:

$$\max \sum_{t=0}^{\infty} (\beta^i)^t \left(\alpha \log(C_t^i) + (1 - \alpha - \beta^i) \log(N_{t+1}^i)\right) \tag{4}$$
$$C_t^i + N_{c,t}^i + p_t X_t^i \leq w_t N_t^i + r_t K_t^i$$
$$N_{t+1}^i = \pi N_t^i + N_{c,t}^i$$
$$K_{t+1}^i = K_t^i + X_t^i.$$

Just as in Lucas (1), 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.

Economy-wide aggregates of consumption, population, children and capital are the sum across dynasty-aggregates and denoted, respectively, $C_t \equiv \sum_i C_t^i$, $N_t \equiv \sum_i N_t^i$, $N_{c,t} \equiv \sum_i N_{c,t}^i$ and $K_t \equiv \sum_i K_t^i$.

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

$$\max_{\{K_t, N_t\}} D K_t^{\nu} N_t^{1-\nu} - w_t N_t - r_t K_t. \tag{5}$$

In our setup, we will be thinking of capital as a fixed, non-reproducible and scarce quantity more akin to land rather than what we normally think of as reproducible capital.

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

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

$$C_t + N_{c,t} = DK_t^{\nu} N_t^{1-\nu}$$

The above are entirely standard. 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 $\{N_0^1, \ldots, N_0^I, K_0^1, \ldots K_0^I\}$, consists of allocations $\{C_t^i, N_{c,t}^i, N_{t+1}^i, K_{t+1}^i, X_t^i\}_{t=0}^{\infty}$ for each dynasty $i = 1, \ldots, I$ and prices $\{w_t, r_t, p_t\}_{t=0}^{\infty}$ such that firms' and dynasties' maximization problems are solved, and all markets clear.

3 Solution

Characterization To solve the model, we derive the first order conditions of firms and the dynastic planner (see Appendix C). For given parameter values, and given initial distributions of the size of dynasties $\{N_0^1, \ldots N_0^I\}$ and stock of capital $\{K_0^1, \ldots K_0^I\}$, the competitive equilibrium of the problem is, for each dynasty $i = 1, \ldots, I$, characterized by consumer first-order conditions with respect to choice of children and consumption as:

$$\frac{(1-\alpha-\beta^i)}{N_{t+1}^i} + (\pi + w_{t+1})\frac{\alpha\beta^i}{C_{t+1}^i} = \frac{\alpha}{C_t^i}, \tag{6}$$

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i \frac{p_{t+1} + r_{t+1}}{p_t}, \tag{7}$$

with consumer budget constraints for each dynasty i:

$$C_t^i + N_{t+1}^i + p_t K_{t+1}^i \leq (w_t + \pi)N_t^i + (r_t + p_t)K_t^i. \tag{8}$$

The firm first-order conditions are:

$$w_t = (1-\nu)DK_t^{\nu}N_t^{-\nu} \tag{9}$$

$$r_t = \nu DK_t^{\nu-1}N_t^{1-\nu}. \tag{10}$$

The market clearing conditions are:

$$\sum_{i=0}^{I} N_t^i = N_t, \tag{11}$$

$$\sum_{i=0}^{I} K_t^i = K_t = \bar{K}. \tag{12}$$

Finally, there two transversality conditions per dynasty:

$$\lim_{t\to\infty} (\beta^i)^t u'(C_t^i) K_{t+1}^i = 0, \tag{13}$$

$$\lim_{t\to\infty} (\beta^i)^t u'(C_t^i) N_{t+1}^i = 0, \tag{14}$$

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

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

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i R_{t+1}, t \geq 0, \tag{15}$$

$$\frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, t \geq 1 \tag{16}$$

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 gross real interest rate on children. Note that 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:

$$\frac{C_{t+1}^i}{C_t^i} = \frac{\beta^i}{\beta^j}\frac{C_{t+1}^j}{C_t^j} \quad \text{and} \quad \frac{N_{t+1}^i}{N_t^i} = \frac{\beta^i}{\beta^j}\frac{N_{t+1}^j}{N_t^j}. \tag{17}$$

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

$$\frac{C_t^i}{C_t} = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^{I}(\beta^j)^t C_0^j}, \text{ and } \quad \frac{N_{t+1}^i}{N_{t+1}} = \frac{(\beta^i)^t N_1^i}{\sum_{j=1}^{I}(\beta^j)^t N_1^j} \tag{18}$$

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, recall that dynasty I is that with the highest patience, the above expressions imply that as $t \to \infty$, so $\frac{N_{t+1}^I}{N_{t+1}} \to 1$ and $\frac{C_{t+1}^I}{C_{t+1}} \to 1$ whilst, for all $i < I$, $\frac{N_{t+1}^i}{N_{t+1}} \to 0$ and $\frac{C_{t+1}^i}{C_{t+1}} \to 0$. In a result which echoes the Ramsey (1928) conjecture, this means that the consumption and population of the most patient type will dominate the economy over time.[6] As $t \to \infty$ the model collapses to standard homogenous agent model with discount factor β^I and a standard Barro-Becker steady state. Consequently, the model can be solved with a reverse-shooting algorithm.

Steady State We can derive the steady-state equilibrium as $t \to \infty$ since the economy becomes entirely dominated by that dynasty with the highest patience. In particular, denoting steady state values as N_{ss}, etc., we have:

$$N^I_{ss} = N_{ss} \text{ and } N^i_{ss} = 0 \quad \forall i < I \tag{19}$$

$$K^I_{ss} = K_{ss} = \bar{K} \text{ and } K^i_{ss} = 0 \quad \forall i < I \tag{20}$$

$$C^I_{ss} = C_{ss} \text{ and } C^i_{ss} = 0 \quad \forall i < I. \tag{21}$$

Using the above with the first order conditions and budget constraints (6)-(8), along with the firm's first order conditions (9) and (10), it follows that in the steady state:

$$N^I_{ss} = N_{ss} = \left(\frac{D(1 - \alpha - \beta^I + \alpha\beta^I(1-\nu))}{(1 - \pi(1-\alpha))(1-\beta^I)} \right)^{\frac{1}{\nu}} \bar{K} \tag{22}$$

$$C^I_{ss} = C_{ss} = (D\bar{K}^\nu N_{ss}^{-\nu} + \pi - 1) N_{ss} \tag{23}$$

$$p_{ss} = \nu \frac{\beta^I}{1-\beta^I} D\bar{K}^{\nu-1} N_{ss}^{1-\nu} \tag{24}$$

$$w_{ss} = (1-\nu) D\bar{K}^\nu N_{ss}^{-\nu} \tag{25}$$

$$r_{ss} = \nu D\bar{K}^{\nu-1} N_{ss}^{1-\nu}. \tag{26}$$

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. The equilibrium of the heterogenous agent model converges to an equilibrium of the homogenous agent model with discount factor.

Aggregation It is numerically convenient to solve the model in two stages: first, we solve the model for economy-wide aggregate variables and prices; and, second, given prices and economy-wide aggregates, we calculate the evolution of dynasty-specific aggregate variables.

In Appendix C we show how using the first order conditions and the steady-properties of the problem we can establish a relationship between the population of each dynasty (relative to that of the most patient dynasty) in the first period to the consumption of each dynasty (relative to that of the most patient dynasty) in the first period:

$$\frac{C^i_0}{C^I_0} = \frac{N^i_1}{N^I_1} \left(\frac{1 - \alpha - \beta^I}{1 - \alpha - \beta^i} \right). \tag{27}$$

[6] Ramsey (op. cit., p. 559) conjectured that, in an economy populated by two groups each with different levels of patience, "...equilibrium would be attained by a division of society into two classes, the thrifty enjoying bliss and the improvident at the subsistence level." See also Becker (1980) and Mitra and Sorger (2013).

Given this expression we can re-write the equations in (18) only in terms of aggregate population and the distribution of populations in periods 1:

$$C_t^i = \frac{(\beta^i)^t(1-\alpha-\beta^i)N_1^i}{\sum_{j=1}^{I}(\beta^j)^t(1-\alpha-\beta^j)N_1^j}C_t, \text{ and } \quad N_{t+1}^i = \frac{(\beta^i)^t N_1^i}{\sum_{j=1}^{I}(\beta^j)^t N_1^j}N_{t+1}, \tag{28}$$

where, from the market clearing condition, aggregate consumption is given by:

$$C_t = DK_t^{\nu}N_t^{1-\nu} - (N_{t+1} - \pi N_t). \tag{29}$$

Recall that $K_t = \bar{K}$ is fixed and note that equation (29) pins down aggregate consumption as a function of aggregate population. Then, substituting equations (28), along with equilibrium wages from equation (9), into equation (6) for $i = I$, gives us a single, second order difference equation in aggregate population. Then, given aggregate population at $t = 1$, the distribution of population across dynasties at $t = 1$ and the assumption that the model converges to its steady state value after some finite number of periods T, we can thus solve for $\{N_t\}_{t=0}^T$ using a standard reverse-shooting algorithm. With solutions for aggregate population, we can solve for aggregate consumption over time from (29). Then, substituting equations (28) into equation (7) for $i = I$ and using the solutions to $\{N_t\}_{t=0}^T$ gives us a first order condition in prices of capital, that can be solved for $\{p_t\}_{t=0}^T$, again under the assumption that the model converges to its steady state value after some finite number of periods T.

Given the preceding solutions for economy-wide aggregates and prices, the equations (18) can be used to find dynasty aggregates of consumption and population for $t \geq 1$. Given that, the budget constraint of each dynasty can be used to back out dynastic capital for $t \geq 0$. Note that the only subtlety in the above is that we need to first choose period 1 population distributions, and only given those will we then be able to back out period zero population distributions. Thus, if we are trying to match the distribution of population in period zero, we need to do this indirectly by first 'guessing' a distribution of time 1 population and then seeing if that guess gives rise to the 'correct' distribution in period zero.

4 Calibration

The key aims of the calibration are to reproduce the increase in world population between the years 1300 and 2000, to capture the distribution of patience types using contemporaneous experimental data, and, to match the remaining technological and preference parameters to reproduce various key moments in global 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 total amount of land is chosen to be $\bar{K} = 11.780$ so that the model reproduces a global population of 6.08 billion at period 28 in the model, which is the year 2000 (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.[7]

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.[8] We assign a discount factor to each dynasty and, without loss of generality, order them 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 find these by noticing that the share of expenditure on consumption relative to aggregate income in the steady-state, s^c_{ss}, is given by:

$$s^c_{ss} \equiv \frac{\alpha\left(1-\beta^I(1-\nu(1-\pi))\right)}{(1-\pi(1-\alpha))\left(1-\beta^I\right)}, \tag{30}$$

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 $\beta^I = \bar{\beta}\left(\frac{2I-1}{2I}\right)$ and that $\bar{\beta} \equiv 1-\alpha$. With $s^c_{ss} = 0.75$ chosen to match to the average global steady-state income share post-2000,[9] we can 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, $\left\{K_0^i\right\}_{i=1}^I$, and population, $\left\{N_0^i\right\}_{i=1}^I$. Note that this data is not directly 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 of those dynasties in subsequent periods. To obtain this initial capital distribution, we assume that the growth of each dynasty's population is *always* consistent with solutions of the model. That is, we assume that outcomes in periods before $t = 0$ are on the equilibrium saddlepath just as much as it is in periods from $t = 0$ on. This simply means that we are ignoring potential shocks, such as wars or famines, that may cause population growth from $t = 0$ to be determined by a different process than that from period $t = 1$ (which makes sense in our fully

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

[8]If too few dynasties are chosen, the resulting transitions are non-smooth. Since we view our model as largely approximating a continuous-like distribution of types in the data, we select a large number of types in the model.

[9]See Appendix B for details.

Table 1: Model parameters

Parameter	Value	Target/Description/Source
D	1	Normalization
N_0	0.370	Aggregate population, 1300
$\bar{K}$	11.780	Aggregate population, 2000
ν	0.190	Land share, Caselli (2005)
π	0.667	Adult life expectancy of 75
I	2000	Number of types
$\left\{\beta^i\right\}_{i=1}^{I}$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^{I}$	Subdivide domain into grid
α	0.428	Consumption share (see text)
$\bar{\beta}$	0.572	Maximum (generational) discount factor
$\{\gamma, \delta\}$	{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)
$\left\{\frac{K_0^i}{\bar{K}}\right\}_{i=1}^{I}$	See text	Consistency assumption (see text)

deterministic framework). The initial distribution of capital is thus chosen such that population growth rates are solutions of the model from period $t = 0$. Practically, this means assuming that equation (16) also holds for $t = 0$ which in turn implies that the second expression in (17) also holds at $t = 0$, viz.:

$$\frac{N_1^i}{N_0^i} = \frac{\beta^i}{\beta^j} \frac{N_1^j}{N_0^j} \tag{31}$$

and hence the share of each type of population in period zero satisfies the following:

$$\frac{N_0^i}{N_0} = \frac{(\beta^i)^{-1} N_1^i}{\sum_{j=1}^{I} (\beta^j)^{-1} N_1^j}. \tag{32}$$

Our assumption on consistency of solutions means that we choose $\left\{K_0^i\right\}_{i=1}^{I}$ so that equation (32) holds for each 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 (18) gives the population share of each dynasty over time as a function of the $t = 1$ population share and each dynasty's level of patience. Using this and

(32), we have the $t = 0$ population share of each dynasty i relative to dynasty I:

$$\frac{N_0^i}{N_0^I} = \frac{N_t^i}{N_t^I}\left(\frac{\beta^i}{\beta^I}\right)^t, \tag{33}$$

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.[10] To address this issue, we let the distribution of generational discount factors in the population be a scaled beta distribution defined on $(0, \bar{\beta})$ with cumulative distribution function, $F(\cdot)$ given by:

$$F(\beta; t) = \frac{B\left(\beta/\bar{\beta}, \gamma_t, \delta_t\right)}{B(\gamma_t, \delta_t)}. \tag{34}$$

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

There are a number of reasons for choosing this distribution. First, 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 Lemma shows:

Lemma 1. *If $I \to \infty$ 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. See Appendix D. □

Lemma 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

[10] For example, consider two dynasties i and I with discount factors $\beta^i = 0.05$ and $\beta^I = 0.5$. From equation (33), 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 $\frac{N_0^i/N_0^I}{N_{28}^i/N_{28}^I} = \left(\frac{\beta^i}{\beta^I}\right)^{28} = 10^{-28}$.

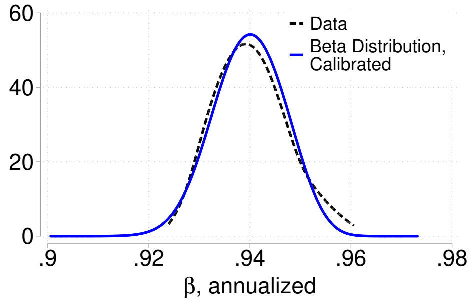


Figure 2: Distribution of annualized discount factors in model and data, BW=0.005. (Source: see text)

the Lemma is that we can derive expressions for the mean and variance of generational discount factors at for any t:

$$\mathrm{E}_t(\beta) = \bar{\beta}\frac{\gamma_0 + t}{\gamma_0 + t + \delta} \text{ and } \mathrm{var}_t(\beta) = \bar{\beta}^2 \frac{(\gamma_0 + t)\delta}{((\gamma_0 + t) + \delta)^2(\gamma_0 + t + \delta + 1)} \tag{35}$$

As $t \to \infty$, 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, just as Ramsey (1928) conjectured.

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:

$$\mathrm{var}_t(\beta^{\frac{1}{25}}) = \bar{\beta}^{\frac{2}{25}} \frac{\Gamma(\gamma_t + \delta_t)}{\Gamma(\gamma_t)^2}\left(\frac{\Gamma(\gamma_t)\Gamma(\frac{2}{25} + \gamma)}{\Gamma(\frac{2}{25} + \gamma_t + \delta_t)} - \frac{\Gamma(\gamma_t + \delta_t)\Gamma(\frac{1}{25} + \gamma_t)^2}{\Gamma(\frac{1}{25} + \gamma_t + \delta_t)^2}\right). \tag{36}$$

and an approximate expression (see Appendix D) for the annualized gross risk free interest rate:

$$R_t^{\frac{1}{25}} \approx 1 + \left(\frac{\gamma_t - 1 + \delta_t}{\bar{\beta}\gamma_t}\right)^{\frac{1}{25}}. \tag{37}$$

As described in Appendix B, we set $\mathrm{var}_{28}(\beta^{\frac{1}{25}}) = 0.005^2$ to match experimental evidence from representative individuals in Denmark (Andersen et al., 2008) and the individual-level data in the Global Preference Survey (GPS) described in Falk et al. (2018). We set $R_{28}^{\frac{1}{25}} - 1 = 0.063$ to match the average (annualized) generational rates of return on global equities.[11] Together,

[11]In Appendix B 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 practically 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.

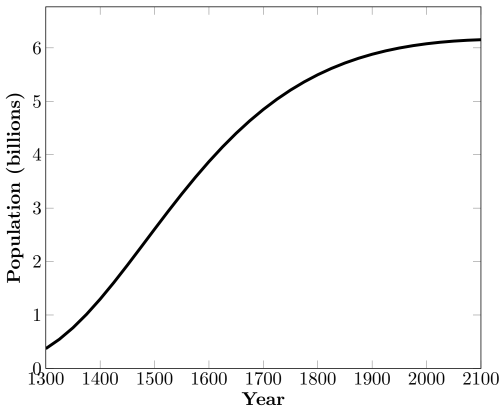


Figure 3: Aggregate population

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:

$$\frac{N_{28}^i}{N_{28}} = F\left(\beta^i + \frac{\bar{\beta}}{2I}; 28\right) - F\left(\beta^i - \frac{\bar{\beta}}{2I}; 28\right). \tag{38}$$

With the above proportions in hand, we can then calculate the $t = 0$ distribution of population using equation (33) with $t = 28$, and proceed to solve the model.

5 Results

Figure 3 shows the increase in aggregate population over time generated by the model. Since this was calibrated to match the levels of global population in the years 1300 and 2000, the model matches the increase in world population over the period, although the increase predicted by the model is unsurprisingly more smooth that what we observe 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 (35). Figure 4 depicts this evolution over time. In our initial period, 1300, societal patience is low – almost 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

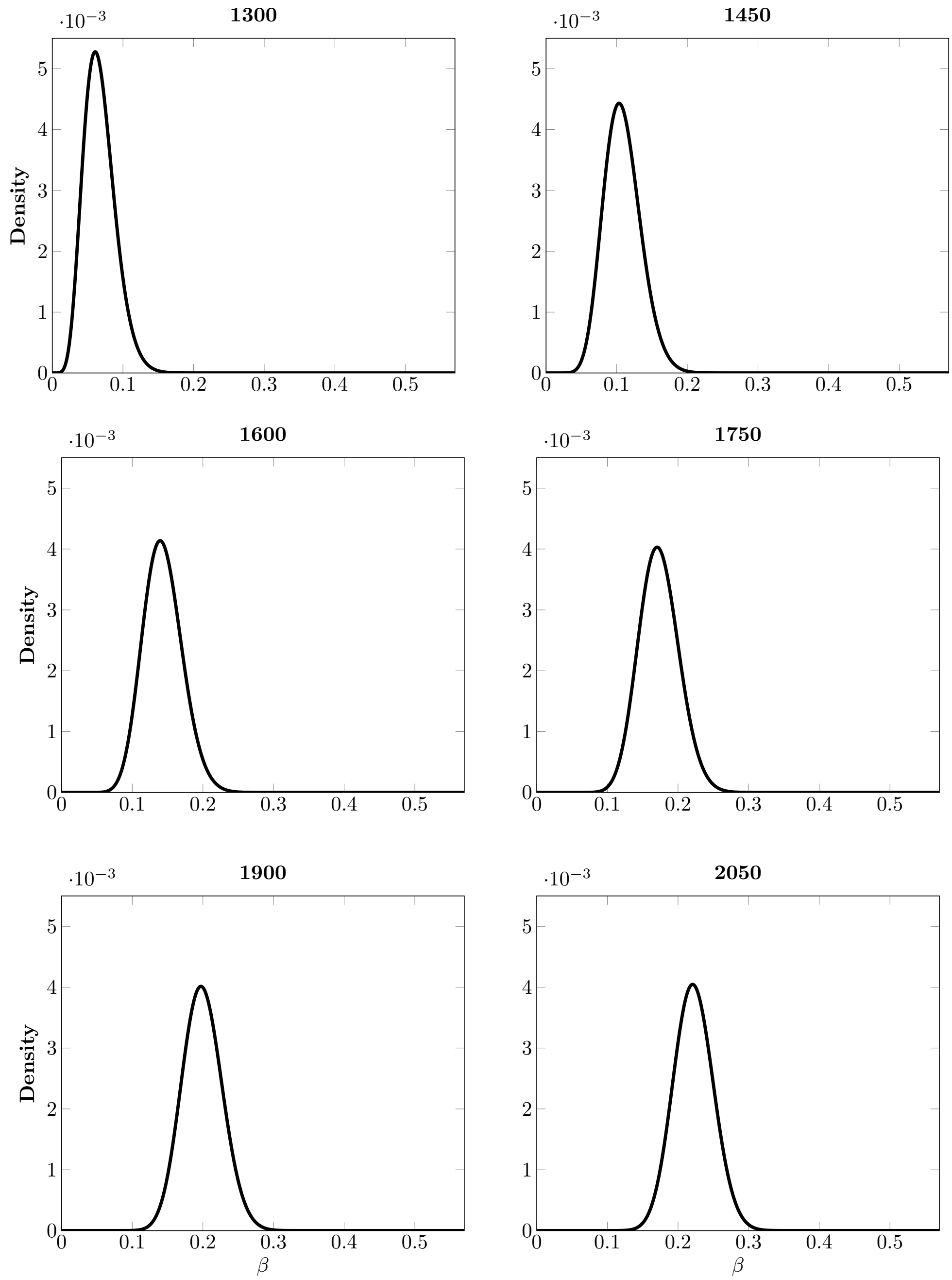


Figure 4: Distribution of patience at different years

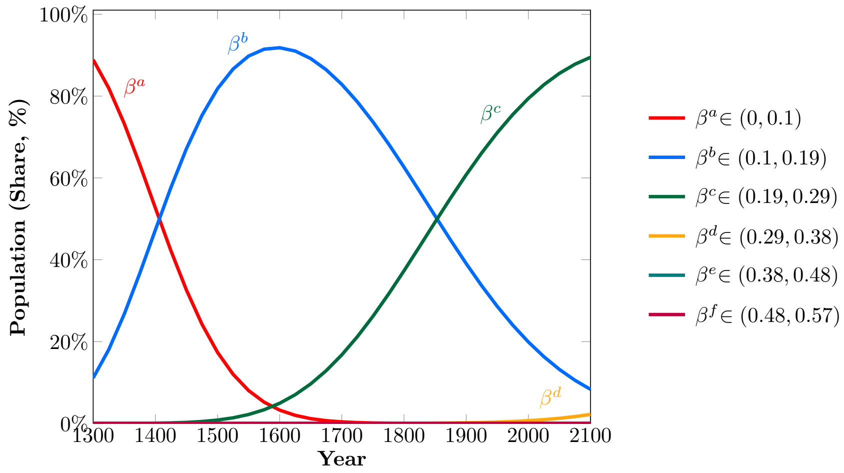


Figure 5: Population share of groups

more patient households are born. By 1900 the median dynasty has a discount factor of $\beta = 0.2$. The (un-normalized) distribution of patience in the population will first gradually becomes less concentrated over time (as more patient agents have more children) and then later it will become more concentrated over time (as the mass of the population reaches the upper limit of patience, $\bar{\beta}$). The mean-normalized patience distribution will decrease monotonically.

The key parameter governing this evolution is the shape parameter of the Scaled-Beta distibution, γ_t, which – as we argued in Lemma 1 – evolves approximately according to the first-order difference equation, $\gamma_{t+1} = \gamma_t + 1$. In the population genetics literature this type of evolution of a characteristic 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 that 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 help present and examine changes in the population over time, we assign agents to one of six groups according to their levels of patience. This allows us to examine the characteristics of low, medium and high-patience types over time. Figure 5 gives the share of each group as a percentage of the total population over time. Notice, that there is a distinctive, cyclical evolution

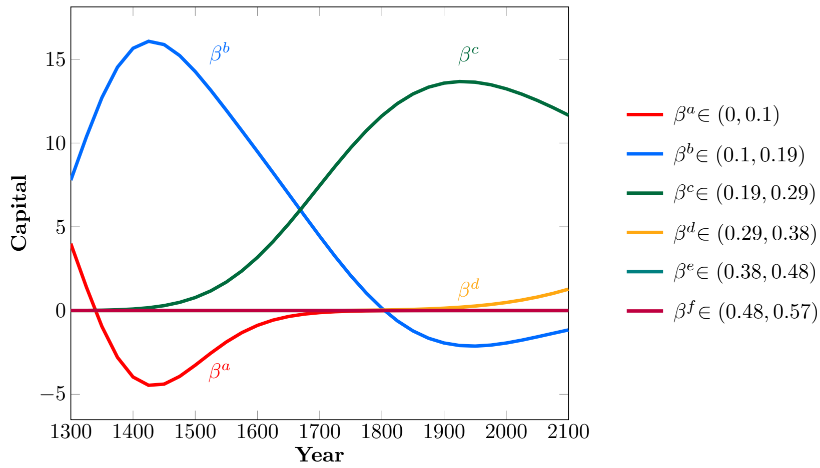


Figure 6: Capital of groups of population

of dominant patience types. The world starts out being dominated by the least patient agents, β^a, who initially account for approximately 90% of the total population in 1300. Over time however, 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. The dominance of this group, however, 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 by changing the importance of individual patience groups. Importantly, the transition from least to most patient is not instantaneous – instead each dynasty and group of dynasties has their rise to and their fall from dominance of the overall population.

The key to understanding this lies in Figure 6, 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 begins to borrow from the more patient dynasties in order to substitute away from children toward the current consumption good. The extent to which impatient dynasties can increase their consumption depends then on the population size of – and the capital owned by – those 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. Figure 7 gives the aggregate of each group of dynasty's consumption levels. While at first the β^a-group is dominant in population share, there is initially a low global population and so its aggregate consumption is low. Successive groups of dynasties rise and fall

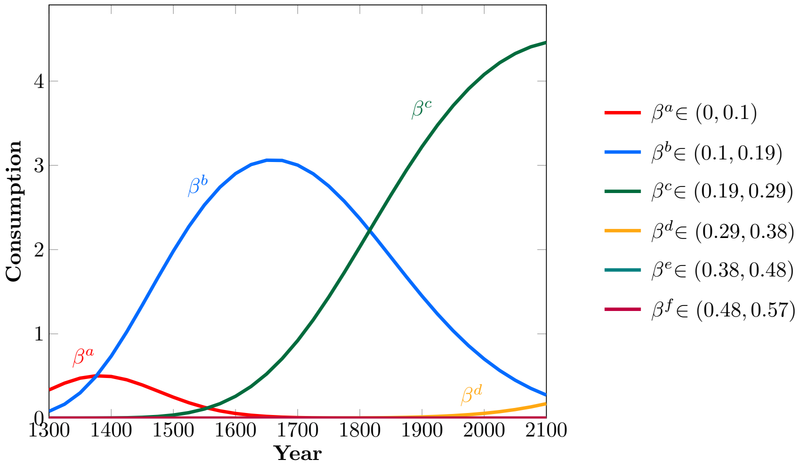


Figure 7: Consumption by groups of population

reaching higher levels of consumption as the aggregate population grows.

Finally, Figure 8 reports the model fit against the interest rate data in Schmelzing (2017). We observe a significant fall in the implied risk free rate as the level of societal patience grows. This decline is approximated by equation (37). Note that in addition to matching the decline, the model also captures the slowing rate of decline in the risk free rate. While there are fluctuations around the long-run trend that we do not capture, our model does not include factors such as time-varying growth rates, risk levels or cyclical shocks such as wars and plagues. Thus the model, calibrated to extant macroeconomic data but, more restrictively, to modern evidence on the distribution of types of patience in the year 2000, successfully captures the main trend in the real risk free rate over the course of eight centuries.

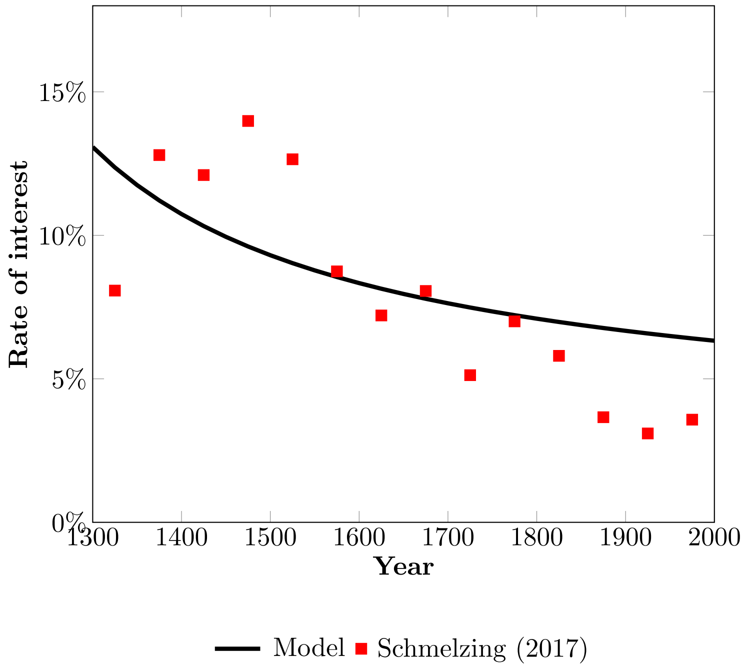


Figure 8: Interest rate

6 Concluding remarks

We have found, using a simple fertility model with heterogeneous preferences calibrated to the modern-day distribution in patience, that we can explain the trend in the risk-free 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. That suggests a potentially important relationship between the constraints on trade or borrowing, the evolution in the population and the risk-free rate. Second, we argued that the time-varying pattern of growth and risk go against the decline in the real 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 studies 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 a clear 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.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström, "'Eliciting Risk and Time Preferences'," *Econometrica*, 2008, *76* (3), 583–618.

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

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.

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

Brenøe, A. A. and T. Epper, "'The Intergenerational Transmission of Time Preferences Across Four Decades'," 2018. Mimeo.

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, "'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.

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.

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.

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

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.

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

Lucas, Robert E., *The industrial revolution: Past and future.*, Harvard University Press, Cambridge, 1.

Mitra, Tapan and Gerhard Sorger, "On Ramsey's conjecture," *Journal of Economic Theory*, 2013, *148*, 1953–76.

Molles, MC, *Ecology Concepts and Applications*, McGraw-Hill Higher Learning, 2010.

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

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

Schmelzing, Paul, "Eight centuries of the risk-free rate: bond market reversals from the Venetians to the 'VaR shock'," October 2017. Bank of England Staff Working Paper No. 686.

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

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

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.

APPENDIX

A Additional Tables

Table 2: Rates of Return, 3000 BC to 2000 AD

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

Notes: [a]*Calculated or referenced in Clark (2007).* [b]*Hudson (2000).* [c]*Schmelzing (2017).*

B Data Appendix

B.1 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 (16) gives the equilibrium relationship between dynasty population dynamics, the dynasty-specific discount rate and the gross real interest rate on children (which is common across dynasties). Since $N^i_{t+1} = N^i_t n^i_t$, we can re-write (16) in terms of the number of children each household has as simple $n^i_t = \beta^i \tilde{R}_{t+1}$. Motivated by this simple relationship, we estimate the following specification,

$$children_{i,2013} = \beta_0 + \beta_1 patience_{i,2008} + \mathbf{X}_i' \boldsymbol{\beta} + \varepsilon_i \tag{39}$$

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. Based on the coefficient in the preferred specification, Column 5, a one standard deviation change in patience is associated with 0.05 standard deviations increase in the number of children.[12] 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.

B.2 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

[12]For those aged 18-40, the standard deviation of patience is 2.37; the standard deviation of children number is 1.09.

Table 3: Patience and Children

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

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

Notes: 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

	(1)	(2)	(3)
VARIABLES	totalChildren	totalChildren	totalChildren
HHpatience	0.010**	0.016***	0.014***
	(0.004)	(0.004)	(0.004)
mediumyoung	0.573***	0.272***	0.146***
	(0.061)	(0.062)	(0.056)
mediumold	0.884***	0.471***	0.199***
	(0.057)	(0.060)	(0.058)
old	-0.056	-0.362***	-0.729***
	(0.050)	(0.052)	(0.055)
lincome		0.420***	0.312***
		(0.017)	(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

Notes: Robust standard errors in parentheses. Standard errors are clustered at the household level. Observations are weighted according to SOEP individual person weights. lincome is the log of household post-government income. mediumyoung is a dummy equal to 1 if $25 < HHage <= 35$*; mediumold is a dummy equal to 1 if* $35 < HHage <= 45$*; and, 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.*

match the s^c_{ss} term in the main body of the text, we proceed as follows. We first calculate the ratio of global consumption to global GDP in every year and then calculate the average of world consumption shares for the years 2000-2018 which comes to 75%.

B.3 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 var $\left(\frac{1}{1+r}\right)$ as a function of the mean $\mathrm{E}(r)$ and variance var(r). We use a first-order Taylor expansion of the second moment of the transformed variable to find var $\left(\frac{1}{1+r}\right) = \frac{1}{(1+\mathrm{E}(r))^4}\mathrm{var}_t(r)$. Thus we use the time preference evidence in Andersen et al. to 'de-normalize' the Falk et al. data by fixing the GPS variation across individuals in Denmark to that found in the experiments. We then obtain a measure of the global variation across individuals, having taken account of region-specific fixed effects. We find the 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 each of the 23 countries is weighted by its starting-year equity market capitalization. (It) also compute(s) a similar world bond (and treasury) index, weighted by GDP."

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

$$r^s_{c,t+1} = \left(\frac{R^s_{c,t+1}}{R^s_{c,t}}\right) - 1, \tag{40}$$

and

$$\bar{r}^s_{c,t+25} = \left(\frac{R^s_{c,t+25}}{R^s_{c,t}}\right)^{\frac{1}{25}} - 1. \tag{41}$$

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

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

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, $\bar{r}^s_{W,t}$, 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%.

C Model derivations

C.1 Equation (1)

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

$$r_t^f \gamma \bar{g}_t - \frac{\gamma^2}{2}\sigma_t^2 - \ln\beta.$$

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

$$\tilde{r}_t^L = \gamma \bar{g}_t - \frac{\gamma^2}{2}\sigma_t^2 - \ln\beta + \gamma d_t. \tag{42}$$

where d_t is related to the covariance between the consumption growth and the return on asset L. While this is a standard expression, we present here its derivation for completeness.

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

$$\max_{L_t, B_t} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t U(C_t) \tag{43}$$

subject to,

$$L_{t+1} + B_{t+1} = R_t^L L_t + R_t^f B_t + W_t - C_t \tag{44}$$

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

Optimal portfolio choices satisfy,

$$R_{t+1}^f \mathbb{E}_t \frac{\beta U'(C_{t+1})}{U'(C_t)} = 1, \tag{45}$$

$$\mathbb{E}_t R_{t+1}^L \frac{\beta U'(C_{t+1})}{U'(C_t)} = 1. \tag{46}$$

To obtain an expression in certainty-equivalent form, we make two assumptions. First, we impose CRRA utility of the form,

$$U(C_t) = \frac{1}{1-\gamma} C_t^{1-\gamma}, \tag{47}$$

and so the optimal portfolio satisfies,

$$R_{t+1}^f \mathbb{E}_t \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} = 1, \tag{48}$$

$$\mathbb{E}_t R_{t+1}^L \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} = 1. \tag{49}$$

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

$$\begin{bmatrix} g_{t+1} \\ r_{t+1}^L \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{g}_{t+1} \\ \bar{r}_{t+1}^L \end{bmatrix}, \begin{bmatrix} \sigma_{g,t}^2, \sigma_{g,L,t}^2 \\ \sigma_{g,L,t}^2, \sigma_{L,t}^2 \end{bmatrix}\right). \tag{50}$$

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

$$\beta \exp\left\{r_{t+1}^f - \gamma \bar{g}_{t+1} + \frac{1}{2}\text{var}_t\left(-\gamma g_{t+1}\right)\right\} = 1 \tag{51}$$

$$\beta \exp\left\{\bar{r}_{t+1}^L - \gamma \bar{g}_{t+1} + \frac{1}{2}\text{var}_t\left(r_{t+1}^L - \gamma g_{t+1}\right)\right\} = 1. \tag{52}$$

Note that from (51) we have the expression for the risk free rate given above as equation (1),

$$r_t^f = \gamma \bar{g}_t - \frac{\gamma^2}{2}\sigma_{g,t}^2 - \ln \beta. \tag{53}$$

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

$$\bar{r}_{t+1}^L + \frac{1}{2}\sigma_{L,t+1}^2 = r_{t+1}^f + \gamma \sigma_{g,L,t+1}^2 \tag{54}$$

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,

$$\ln \mathbb{E}_{t-1} R_t^L = r_t^f + \gamma \sigma_{g,L,t}^2 \tag{55}$$

which, with $\tilde{r}_t^L = \ln \mathbb{E}_{t-1} R_t^L$ and $d_t = \gamma \sigma_{g,L,t}^2$, is the expression given in equation (42).

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. There are a number of reasons for thinking this is not the case, however. First, a key contribution of Schmelzing (2017) is in constructing a dataset of the global risk free 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.2 Model solution

The following expands on elements of the model solution, as described in Sections 2-3.

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:

$$\max_{C_t^i, K_{t+1}^i, N_{t+1}^i} \sum_{t=0}^{\infty} (\beta^i)^t \left(\alpha \log(C_t^i) + (1-\alpha-\beta^i)\log(N_{t+1}^i) \right) \tag{56}$$

$$C_t^i + N_{t+1}^i + p_t K_{t+1}^i \leq (w_t + \pi) N_t^i + (r_t + p_t) K_t^i. \tag{57}$$

The first order conditions for this problem are given by:

$$\lambda_t^i = \frac{\alpha(\beta^i)^t}{C_t^i}, \tag{58}$$

$$\frac{(1-\alpha-\beta^i)(\beta^i)^t}{N_{t+1}^i} + (\pi + w_{t+1})\lambda_{t+1}^i = \lambda_t^i \tag{59}$$

$$p_t \lambda_t^i = (p_{t+1} + r_{t+1})\lambda_{t+1}^i, \tag{60}$$

where, λ_t^i is the Lagrange multiplier on the constraint (57). Now, substituting out for λ_t^i in the last two FOCs using the first FOC, we obtain:

$$\frac{(1-\alpha-\beta^i)}{N_{t+1}^i} + (\pi + w_{t+1})\frac{\alpha\beta^i}{C_{t+1}^i} = \frac{\alpha}{C_t^i} \tag{61}$$

and

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i \frac{p_{t+1} + r_{t+1}}{p_t}. \tag{62}$$

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 (15) in the main text.

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

$$w_t = (1-\alpha) D K_t^{\alpha} N_t^{-\alpha} \tag{63}$$

and

$$r_t = \alpha D K_t^{\alpha-1} N_t^{1-\alpha}. \tag{64}$$

Population Euler Equation To derive equation (16) we proceed as follows. We re-write FOC (6) 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, (62), to substitute out for $\frac{C_{t+1}^i}{C_t^i}$ to obtain and expression for N_{t+1}^i:

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

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

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

Taking the ratio of these two equations and substituting for $\frac{C_{t+2}^i}{C_{t+1}^i}$ from the Euler equation, (62), we obtain:

$$\frac{N_{t+2}^i}{N_{t+1}^i} = \beta^i \tilde{R}_{t+2}, \tag{65}$$

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:

$$\frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, \tag{66}$$

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 (16) in the main text.

Initial Population and Consumption To obtain equation (27) in the main text, we plug in equation (18) into (6).

$$\frac{(1-\alpha-\beta^i)}{\frac{(\beta^i)^t N_1^i}{\sum_{j=1}^I(\beta^j)^t N_1^j}N_{t+1}} + (\pi+w_{t+1})\frac{\alpha\beta^i}{\frac{(\beta^i)^{t+1}C_0^i}{\sum_{j=1}^I(\beta^j)^{t+1}C_0^j}C_{t+1}} = \frac{\alpha}{\frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I(\beta^j)^t C_0^j}C_t}, \tag{67}$$

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

$$\frac{(1-\alpha-\beta^i)}{\frac{N_1^i}{\sum_{j=1}^I(\frac{\beta^j}{\beta^I})^t N_1^j}N_{t+1}} + (\pi+w_{t+1})\frac{\alpha}{\frac{C_0^i}{\beta^I\sum_{j=1}^I(\frac{\beta^j}{\beta^I})^{t+1}C_0^j}C_{t+1}} = \frac{\alpha}{\frac{C_0^i}{\sum_{j=1}^I(\frac{\beta^j}{\beta^I})^t C_0^j}C_t}, \tag{68}$$

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

$$\frac{(1-\alpha-\beta^i)}{\frac{N_1^i}{N_1^I}N_{ss}} + (\pi+w_{ss})\frac{\alpha}{\frac{C_0^i}{\beta^I C_0^I}C_{ss}} = \frac{\alpha}{\frac{C_0^i}{C_0^I}C_{ss}}. \tag{69}$$

Then, substituting from the solutions of the steady state shown in equations (22)-(26) into the above, for each $i < I$ we can then show that:

$$\frac{C_0^i}{C_0^I} = \frac{N_1^i}{N_1^I}\frac{1-\alpha-\beta^I}{1-\alpha-\beta^i}. \tag{70}$$

D Asymptotic results

D.1 Proof of Lemma 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 Lemma 1.

Lemma 1. *If $I \to \infty$ 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 $i = 1, \cdots, n$. 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: $\nu_t(\beta(i;n)) \equiv \frac{N_t^i}{N^t}$, which maps the discount factor of a particular dynasty to the fraction of the total population of that dynasty i at time t. Notice, that we can think of this function as a probability mass function of a discrete random variable with realization, $\beta(i;n)$, on the domain $\{\frac{2i-1}{2n} | i = 1, \cdots, n\}$. We wish to characterize the evolution of the asymptotic function, $\frac{\nu_t(\beta(i;n))}{\Delta(n)}$, over time as $n \to \infty$ - 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 (16) as:

$$N_{t+1}^i = \beta^i \tilde{R}_{t+1} N_t^i. \tag{71}$$

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:

$$N_{t+1} = \beta^i \tilde{R}_{t+1} N_{t+1} \sum_{j=1}^{n} \beta^j \nu_t^n(\beta^j). \tag{72}$$

Dividing equation (71) by equation (72) we obtain:

$$\nu_{t+1}^n(\beta^i) = \frac{\beta^i \nu_t^n(\beta^i)}{\sum_{j=1}^n \beta^j \nu_t^n(\beta^j)}. \tag{73}$$

This recursive formulation defines the evolution of the probability mass function over time. We are interested in the properties of this function as $n \to \infty$. 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:

$$F_t^n(\beta^i) \equiv \frac{\sum_{j=1}^i \beta^j \nu_t^n(\beta^j)}{\sum_{j=1}^n \beta^j \nu_t^n(\beta^j)}. \tag{74}$$

This also means that:

$$\nu_t^n(\beta^i) = F_t^n(\beta^{i+1}) - F_t^n(\beta^i) = P_t^n(\beta^i \leq \beta \leq \beta^{i+1}). \tag{75}$$

Given the above, notice that (73) can be re-written as:

$$\frac{\nu_{t+1}^n(\beta^i)}{\Delta^i(n)} = \frac{\beta^i \frac{\nu_t^n(\beta^i)}{\Delta^i(n)}}{\sum_{j=1}^n \beta^j P_t^n(\beta^j \leq \beta \leq \beta^{j+1})}. \tag{76}$$

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

$$f_{t+1}(\beta) = \frac{\beta f_t(\beta)}{E_t(\beta)}, \tag{77}$$

where f_t is the continuous probability density function corresponding to the discrete mass function ν_t^n[13] and $E_t(\beta) \equiv \int_0^1 u f_t(u) du = \lim_{n\to\infty} \sum_{j=1}^n \beta^j P_t^n(\beta^j \leq \beta \leq \beta^{j+1})$, 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 (34). The corresponding probability density function of this distribution f is given by:

$$f_t(\beta; \boldsymbol{\theta_t}) \equiv f(\beta; \gamma_t, \delta_t) = \frac{(\bar{\beta} - \beta)^{\delta_t - 1} \beta^{\gamma_t - 1}}{\bar{\beta}^{\delta_t + \gamma_t - 1} B(\gamma_t, \delta_t)}, \tag{78}$$

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

$$E(\beta; \gamma_t, \delta_t) = \bar{\beta} \frac{\gamma_t}{\gamma_t + \delta_t}. \tag{79}$$

Using equations (77)-(79), we can write the pdf of discount factors at time $t+1$ as:

$$\begin{aligned} f_{t+1}(\beta; \gamma_t, \delta_t) &= \frac{\beta(\bar{\beta} - \beta)^{\delta_t - 1} \beta^{\gamma_t - 1}}{\bar{\beta} \frac{\gamma_t}{\gamma_t + \delta_t} \bar{\beta}^{\delta_t + \gamma_t - 1} B(\gamma_t, \delta_t)} \\ &= \frac{(\bar{\beta} - \beta)^{\delta_t - 1} \beta^{\gamma_t}}{\frac{\gamma_t}{\gamma_t + \delta_t} \bar{\beta}^{\delta_t + \gamma_t} B(\gamma_t, \delta_t)} \\ &= \frac{(\bar{\beta} - \beta)^{\delta_t - 1} \beta^{\gamma_t}}{\bar{\beta}^{\delta_t + \gamma_t} B(\gamma_t + 1, \delta_t)} \\ &= f(\beta; \gamma_{t+1}, \delta_{t+1}) \end{aligned} \tag{80}$$

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 (77) is the beta distribution with parameters given by $\gamma_{t+1} = 1 + \gamma_t$ and $\delta_t \equiv \delta$.

□

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

D.2 Asymptotic expression for the rate of interest

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

$$R_{t+1} = \frac{C^i_{t+1}/C^i_t}{\beta^i} = \frac{\left(\frac{\kappa^I_{t+1}(\beta^i)/\Delta(I)}{\kappa^I_t(\beta^i)/\Delta(I)}\right)\frac{C_{t+1}}{C_t}}{\beta^i} \tag{81}$$

where $\kappa^I_t(\beta^i) \equiv C^i_t/C_t$. Note also that we can write:

$$\frac{\kappa^I_t(\beta^i)}{\Delta(I)} = \frac{\frac{\beta^i}{1-\alpha-\beta^i}\frac{\nu^I_t(\beta^i)}{\Delta(I)}}{\sum_{j=1}^I \frac{\beta^j}{1-\alpha-\beta^j}\nu^I_t(\beta^j)}. \tag{82}$$

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

$$f_{ct}(\beta) = \frac{\frac{\beta}{1-\alpha-\beta} f_t(\beta)}{E_t(\frac{\beta}{1-\alpha-\beta})}, \tag{83}$$

where f_t and f_{ct} are the continuous probability density function corresponding to the discrete mass functions ν^I_t and κ^I_t. Note also that using the relationship derived between $f_{t+1}(\beta)$ and $f_t(\beta)$ in the Appendix we have the following expression:

$$\frac{f_{ct+1}(\beta)}{f_{ct}(\beta)} = \beta\frac{E_t\left(\beta/(\bar{\beta}-\beta)\right)}{E_t\left(\beta^2/(\bar{\beta}-\beta)\right)} \tag{84}$$

Taking the limit of both sides of (81) as $I \to \infty$ we obtain:

$$R_{t+1} = \frac{E_t\left(\beta/(\bar{\beta}-\beta)\right)}{E_t\left(\beta^2/(\bar{\beta}-\beta)\right)}\frac{C_{t+1}}{C_t}. \tag{85}$$

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:

$$R_{t+1} \approx \frac{E_t\left(\beta/(\bar{\beta}-\beta)\right)}{E_t\left(\beta^2/(\bar{\beta}-\beta)\right)}. \tag{86}$$

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:

$$R^{\frac{1}{25}}_{t+1} \approx \left(\frac{\gamma_t + \delta_t}{\bar{\beta}(1+\gamma_t)}\right)^{\frac{1}{25}}. \tag{87}$$