University of Glasgow

Adam Smith Business School

WORKING PAPER SERIES

Selection, Patience, and the Interest Rate

Radoslaw (Radek) Stefanski and Alex Trew

Paper No. 2023-01
January 2023

Selection, Patience, and the Interest Rate*

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

January 19, 2023

Abstract

The interest rate has been steadily falling for centuries. A process of selection that leads to increasing societal patience is key to explaining this decline. Three observations point to the potential role of this mechanism: patience varies across individuals, patience is inter-generationally persistent, and patience is positively related to fertility. A calibrated dynamic, heterogenous-agent model of fertility permits us to isolate the quantitative contribution of this mechanism. Selection can explain most of the decline in the interest rate over seven centuries, a fact that is robust to a number of model extensions. Quantitative implications are consistent with other facts, such as the steady increase in the investment rate since 1300.

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

*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, Barcelona, UEA and conference participants at WEHIA for their comments and suggestions. Stefanski would also like to thank Yale University Department of Economics for its hospitality while writing this paper. The usual disclaimer applies.

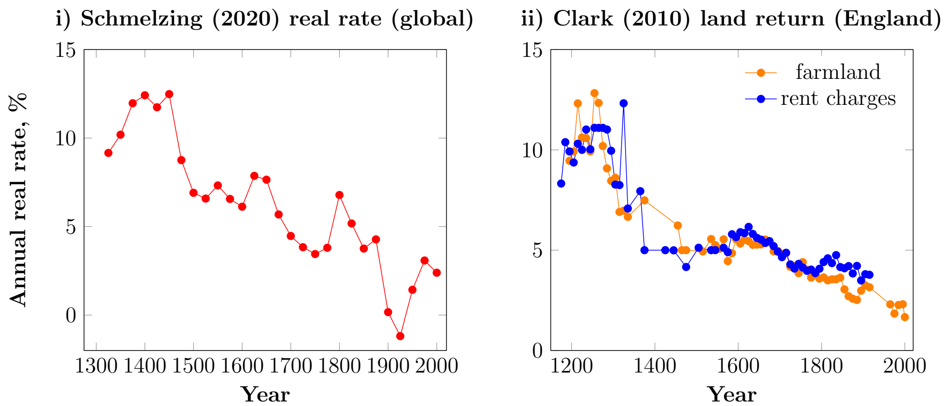
[†]School of Economics and Finance, Castlecliffe, The Scores, University of St Andrews, Fife, KY16 9AR, United Kingdom. Email address: rls7@st-andrews.ac.uk.

[‡]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 11-12% in the fourteenth century to just 2–3% today (Schmelzing, 2020). The real return on land in England fell from around 10% in the thirteenth century to 1-2% today (Clark, 2010).[1]

Figure 1: Real interest rate, 1175–2000



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

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

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

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

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

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

patience driven by changing demographics. As Blanchard (1985) showed, rising life expectancy can appear as an increase in effective β where agents with finite-horizons save more. However, as we are able to make explicit, this channel does not explain the long decline in rates since life expectancy was flat until the 19th century (Wrigley et al., 1997).[4]

We propose a novel demographic channel that can explain the long decline in interest rates. We introduce a model of endogenous fertility, in the spirit of Becker and Barro (1988), where patience levels can be heterogenous across agents. Since children are a form of saving, and since patient agents tend to save more than impatient agents, the model implies that patient agents will have more children. If those children in turn inherit part of their parent's higher patience levels, then the average level of patience in society will increase over time as a result of evolutionary pressures that naturally select the most patient agents. While this mechanism is theoretically plausible, its relevance in practice is a quantitative question. We use the structure of our model to calibrate the distribution of patience across individuals to experimental micro-level evidence.[5] We find that the contribution of selection – the difference between the implications of our heterogeneous-agent model and a model of one dynasty – can explain much of the fall in the global interest rate. This conclusion is robust to a number of extensions to the baseline set-up, including calibration to the historical path of income and population as well as to accounting for the demographic transition via changing fertility and life expectancy. The role of selection is also robust to incorporating a Blanchard (1985)-type mechanism where imperfect altruism means that growing life expectancy can depress interest rates. Without a fully-specified and calibrated model, we would be unable to assess the quantitative importance of each of these channels.

Understanding the factors driving the real rate of interest over time is crucial for long-term inter-temporal decisions that are associated with savings and investment choices or future paths of innovation, as well as for the long-run sustainability of public debt. For example, optimal policies to address very long-term, inter-generational

[4]An additional mechanism could be a relationship between consumption and time preference (Epstein and Hynes, 1983); again, the timing of the increase in consumption would not help explain falling rates before around 1850.

[5]Specifically, the calibration of the distribution of patience in our model is based on the individual-level data in the Global Preference Survey (GPS) described in Falk et al. (2018). We leave the detail to section 3.2.

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).

Related literature First, we contribute to the literature on the the role of selection and preferences in economics. Galor and Moav (2002) propose a theory in which there is an evolutionary advantage to traits that are complementary to the escape from the Malthusian trap. Following the demographic transition, higher incomes improve child quality (greater human capital) instead of child quantity. While we can capture a switch in the aggregate time-series correlation between fertility and income, doing so does not affect the key role that selection plays in explaining the decline in real rates. Closely related is Galor and Özak (2016), which presents a dynamic model in which higher patience leads to better economic outcomes and, consequently, greater reproductive success. Geographical variation in returns to agricultural investment mean that the returns to patience also varies, an implication that Galor and Özak find is consistent with empirical evidence from pre-industrial societies. Our contribution is to understand the relatively more recent dynamics in a way that complements the very long-run comparative analysis in Galor and Özak. Falk et al. (2018) point to some variation in cross-country averages of various preference characteristics, including patience, and Sunde et al. (2021) is an example of such variation having significant consequences for comparative development. Explaining cross-country differences in average levels of patience is beyond the scope of this paper, but we show that the interest rate difference implied by the cross-country variation in patience is small compared to the decline in the interest rate observed over time.

We also relate to the literature on the formation and evolutionary stability of preferences. Becker and Mulligan (1997) study time preference as determined in part by agent choice of effort to increase the 'appreciation' of the future. In that paper, wealth causes patience; as we will see, the timing of the increase in growth that occurs following the industrial revolution does not help us to explain the path of interest rates. Our paper focuses on what Becker and Mulligan term the 'endowed discount factor' alone. Doepke and Zilibotti (2008) focus on the intergenerational evolution of patience across and within social classes as it relates to parental decisions to invest in different characteristics of their children. Since this mechanism operates at a shorter time horizon (no more than two to three generations), such a channel is complementary to one driven by selection that operates over much longer periods. Other work focuses on preferences as an evolutionarily stable outcome (Becker, 1976, Rogers, 1994, and Robson and Szentes, 2008). Since we calibrate a dynamic model that incorporates the shifting distribution of types, we are able to show just how long it can take for such stable preferences to be realized. The closest to our set-up is Hansson and Stuart (1990), in which the population growth of a dynasty is assumed monotonically increasing in per-capita consumption. In our model, population growth is a function of preferences and of the environment. The long-run in our model is the result of a slow process of selection that leads to the most patient dynasty dominating, a result which also echoes the Ramsey (1928) conjecture.[6]

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

Third, we relate to the growing literature on family macroeconomics (see Doepke and Tertilt, 2016 for a recent survey), of which Doepke and Zilibotti (2008) noted above is an early example. 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 over the very long run.

Fourth, our work connects to research on the drivers of the more recent decline in global real rates (see, for example, the chapters in Teulings and Baldwin, eds, 2014). Del Negro et al. (2018) isolate the role of growing risk and declining growth rates in explaining the decline of the last ten years. Carvalho et al. (2021) consider the demographic channel over 1990 to 2014, finding a significant role for life expectancy. Our contribution incorporates the mechanisms thought to matter for the last few decades but also points to an additional mechanism that has operated over centuries.

[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).

Structure In section 2 we introduce the evidence on the distribution and transmission of preferences. In section 3 we develop a Barro-Becker model of fertility where the key departure is to introduce heterogenous dynasties that differ according to their discount factor. This simple model serves to lay out the role of the selection mechanism in driving changes in the interest rate. We calibrate the model and compare its quantitative implications to the historical record. Section 4 extends the baseline model to incorporate endogenous capital accumulation and calibrates child costs, life expectancy and productivity to match the historical path of population and income per capita. We also introduce Blanchard (1985)-type mechanism with a form of imperfect altruism in which life expectancy impacts effective patience. We then quantify the contribution of selection to the decline in interest rates in each of these models. Section 5 explores external validity by comparing other implications of the model to historical data such as the investment rate. Section 6 discusses the implications of imperfect transmission or mutation of preferences. Finally, section 7 offers some concluding remarks.

2 Heterogeneity, transmission and fertility

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 heterogenous; preferences are inter-generationally persistent; and, patience is related to fertility. We offer a brief discussion of the empirical basis of each of these conditions.

Our departure from a standard model of endogenous fertility is based on the following: first, patience varies across individuals and, second, patience is intergenerationally persistent. Andersen et al. (2008) elicit time and risk preferences in a representative sample of Danes, while Alan and Browning (2010) use structural estimation and the PSID. Both find similar heterogeneity in discount factors across individuals. More recently, Falk et al. (2018) establish the substantial extent to which preferences vary across individuals.[7] Intergenerational transmission of preferences, either by genetics, imitation or by socialization, has been identified in studies on Danish and Bangladeshi families (respectively, Brenøe and Epper, 2018 and Chowdhury et

[7]There is evidence in Falk et al. (2018) of country-level variation in average levels of patience, but we find that this cross-sectional variation is relatively small in terms of the implied interest rate differences when compared to the observed fall in the interest rate over time (see Appendix E.3).

al., 2022). Dohmen et al. (2011) has shown that other elements of preferences are also persistent intergenerationally.

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

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

3 A heterogenous-agent Barro-Becker model

In this section we present a baseline model which introduces heterogenous agents into a Barro-Becker fertility framework and allows us to explore the role of selection alone in explaining the path of the interest rate. In Section 4 we extend this simple model to match the path of population, income per capita, life expectancy, capital accumulation and a form of imperfect altruism. Our baseline results continue to hold: the selection mechanism can explain a large portion of the overall decline in the interest rate.

Baseline set-up 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.[8] 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 in each period t:

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

s.t.

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

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

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

As in Becker and Barro (1988) and Barro and Becker (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.[9] 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

[8] Since households within a dynasty are identical, and since we obtain solutions to the model in terms of dynasty-aggregates, we omit a household index. As we explain below, household-level quantities are lower-case, so, e.g., c_t^i is the time t consumption of an individual household in dynasty i; dynasty-aggregates are upper case, so C_t^i is the sum of consumption by households in dynasty i at time t.

[9] This has no impact on our key findings regarding interest rates, but we generalize it in the extended model of section 4.

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 $n^i_{t+1} = \pi + n^i_{c,t}$. 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 Becker and Barro (1988) and Barro and Becker (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 6). This transmission can be thought of as coming from genetics, imitation or socialization and, given the lack of clear identification of mechanisms in the empirical literature described above, is left as a reduced form assumption.

Discussion Two aspects of the above model merit further discussion. First, as in Becker and Barro (1988) and Barro and Becker (1989), we assume in (2) a form of altruism. Part of altruism is that parents enjoy larger families; another part is that parents care about the average utility of children in future periods. In our model, however, parents survive into the future with some probability, alongside their children. As such, our discounting parameter not only captures the utility of children realized at different points in the future but also captures the future utility of the parent. The implicit altruistic assumption we make is that, in terms of time discounting, the future utility of children is treated the same as the future utility of parents. That is, in each period a parent applies the same *time discount factor* to their own and their children's future utility, but the altruism implicit in equal-treatment of own and descendant utility is itself unrelated to β.[10] Given this assumption, and the fact that parents live for multiple periods, the β captures time preference rather than altruism. We refer to this form of altruism as 'perfect' since parents take account of all future periods, i.e., they care about their descendants even after the parent dies. In order to consider the sensitivity of our results to this assumption, we relax it in section 4 and allow for a form of 'imperfect' altruism in which parents care about their children, but only while they themselves remain alive. This introduces a different effective weight on a parent's own future utility compared to the entire future of their descendant's utility.[11] This also permits us to consider a Blanchard

[10]The time-zero household problem in the appendix, equation (76), makes this clear.

[11]Again, this is distinct from the time preference parameter, as the time-zero household problem in equation (76) shows.

(1985) mechanism which may be important in explaining a decline in interest rates that arises out of increasing life expectancy observed after 1850. In the quantitative analysis, the contribution of selection to explaining the decline in the interest rate remains similar in both cases.

Second, an assumption implicit in models of endogenous fertility is that parents can always choose the number of children. Whether deliberate birth control existed in the period before the demographic transition is a topic of recent debate.[12] For Clark and Hamilton (2006) and Clark (2007a), differences in *survival* rates across groups, rather than fertility rates, led to changes in the composition of the population. Either interpretation is consistent with our model. In the current set-up, we make the assumption that all children survive to adulthood and so the endogenous choice of the number of 'children' in our model is really a choice of the number of adults in the dynasty in the next period. We can thus otherwise think of this as a choice to allocate the resources in raising a child to adulthood. The mechanism in our model holds whether the variation arises from endogenous birth rates, or whether it arises via endogenous survival to adulthood through different parental investment decisions.[13]

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:

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

s.t.

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

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

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

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

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

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. 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:

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

s.t.

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

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

$$K_{t+1}^i = K_t^i + X_t^i.$$

Just as in Lucas (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:

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

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 given by:

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

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

We introduce endogenous capital formation in section 4.

3.1 Solution

We define a standard competitive equilibrium of the above in Appendix C.1. 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 is characterized by the following household first-order conditions with respect to choice of children and consumption for each dynasty $i = 1, \dots, I$:

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

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

as well as household budget constraints, firm first order conditions, market clearing conditions and standard transversality conditions all derived in Appendix C.2. From the household first-order conditions, we obtain the following two Euler equations that describe the evolution of dynasty consumption and population:

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i R_{t+1}, t \geq 0, \qquad \frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, t \geq 1, \tag{9}$$

where $R_{t+1} \equiv \left(\frac{p_{t+1}+r_{t+1}}{p_t}\right)$ is the gross real interest rate on capital and $\tilde{R}_{t+1} \equiv R_{t+1}\frac{R_t-(w_t+\pi)}{R_{t+1}-(w_{t+1}+\pi)}$ is the shadow gross real interest rate on dynasty population.[14]

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

$$\frac{C_t^i}{C_t} = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I (\beta^j)^t C_0^j}, \text{ and, } \quad \frac{N_{t+1}^i}{N_{t+1}} = \frac{(\beta^i)^t N_1^i}{\sum_{j=1}^I (\beta^j)^t N_1^j}, \tag{10}$$

for $t \geq 0$. Note that given the initial distributions, the evolution of a particular dynasty's population and consumption shares depends only on that dynasty's patience relative to the patience of other dynasties. In particular, recalling that dynasty I is most patient, the above expressions imply that as $t \to \infty$, so $\frac{N_{t+1}^I}{N_{t+1}} \to 1$ and $\frac{C_{t+1}^I}{C_{t+1}} \to 1$ whilst, for all $i < I$, $\frac{N_{t+1}^i}{N_{t+1}} \to 0$ and $\frac{C_{t+1}^i}{C_{t+1}} \to 0$. This means that the total consumption and population of the most patient dynasty will dominate the economy over time (consistent with the Ramsey (1928) conjecture). As $t \to \infty$ the model collapses to standard homogenous agent model with discount factor β^I and a standard Barro-Becker steady state. Consequently, if we derive the steady state, the model can be solved with a variation of the reverse-shooting algorithm.[15]

3.2 Calibration of baseline model

The key aim of the calibration is to replicate the increase in world population from 1300 to 2000 and to fit the variance of patience types using modern experimental data. Section 4 extends the calibration to match the exact *paths* of global population and income per capita. Model parameters and their calibrated values are summarized in Table 1 and discussed in more detail in Appendix C.4.

One period in the model is 25 calendar years and period zero in the model corresponds to the year 1300 in the data. We normalize the level of technology so that

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

[15]For details of this derivation and specifics of the algorithm see Appendices C.2 and C.3.

Table 1: Model parameters

Parameter(s)	Value	Target/Description/Source
D	1	Normalization
N_0	0.370	Aggregate population, 1300, The Maddison Project (2013)
$\bar{K}$	11.722	Aggregate population, 2000, The Maddison Project (2013)
ν	0.190	Land share, Caselli (2005)
π	0.667	Adult life expectancy of 75
I	10000	Number of types
$\{\beta^i\}_{i=1}^I$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^I$	Subdivide domain into grid
α	0.427	Consumption share (see Appendix)
$\bar{\beta}$	0.573	Max. (generational) discount factor
$\{\gamma_{28}, \delta_{28}\}$	{32.089,53.531}	Standard deviation of discount factors (Andersen et al., 2008; Falk et al., 2018) and long run rate of return (see Appendix)
$\left\{\frac{N_0^i}{N_0}\right\}_{i=1}^I$	See text	Andersen et al. (2008) & Falk et al. (2018)
$\left\{\frac{K_0^i}{K_0}\right\}_{i=1}^I$	See text	Consistency (see text)

$D = 1$. 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, $\bar{K} = 11.722$, 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.

We specify the number of dynasties to be $I = 10,000$.[16] We discuss the grid of βs and the bound on this grid in appendix C.4. 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

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

dynasty from modern data.

Initial capital and population distribution As described in Appendix C.4, the initial distribution of capital is chosen such that population growth rates are solutions of the model from period $t = 0$. In practice, this means assuming that the second equation in (9) holds for $t = 0$ which in turn implies that the expression for relative population growth also holds at $t = 0$:

$$\frac{N_1^i}{N_1^j} = \frac{\beta^i}{\beta^j}\frac{N_0^i}{N_0^j}. \tag{11}$$

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 variance of types in the year 2000 ($t = 28$ in the model). The second expression in (10) 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 (11), we have the $t = 0$ population share of each dynasty i relative to dynasty I:

$$\frac{N_0^i}{N_0^I} = \frac{N_t^i}{N_t^I}\left(\frac{\beta^i}{\beta^I}\right)^t, \tag{12}$$

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.[17] 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:

$$F(\beta; t) = \frac{B\left(\beta/\bar{\beta}, \gamma_t, \delta_t\right)}{B(\gamma_t, \delta_t)}. \tag{13}$$

[17]For example, consider two dynasties i and j with discount factors $\beta^i = 0.05$ and $\beta^j = 0.5$. From equation (12), 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 $\frac{N_0^i/N_0^j}{N_{28}^i/N_{28}^j} = \left(\frac{\beta^i}{\beta^j}\right)^{28} = 10^{-28}$.

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

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

Theorem 1. *If $I \to \infty$ and dynastic discount factors 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 F. □

Theorem 1 establishes that if discount factors obey a scaled-beta distribution in any one period then they will follow a scaled-beta distribution in all other periods. Since the theorem also pins down the evolution of shape parameters over time, the choice of year in which to calibrate is irrelevant. An immediate implication of the Theorem is that we can derive expressions for the mean and variance of generational discount factors at any time t:

$$\mathrm{E}_t(\beta) = \bar{\beta}\frac{\gamma_0 + t}{\gamma_0 + t + \delta} \text{ and } \mathrm{var}_t(\beta) = \bar{\beta}^2 \frac{(\gamma_0 + t)\delta}{((\gamma_0 + t) + \delta)^2(\gamma_0 + t + \delta + 1)} \tag{14}$$

Interest rate approximation It also follows (see Appendix F.3) that we can approximate the gross interest rate by:

$$R_{t+1} \approx \frac{1}{E_t(\beta)}. \tag{15}$$

Over time, societal patience increases and leads to falling interest rates. As $t \to \infty$, the mean beta converges from below to $\bar{\beta}$ and the variance goes to zero: thus the agent with the highest discount factor comes to entirely dominate the economy. We set γ_{28} and δ_{28} to match global variance of discount factors using experimental data from representative individuals in Denmark (Andersen et al., 2008) and individual-level data from the Global Preference Survey (GPS) (Falk et al., 2018) as well as average generational rates of return on global equities. Appendix C.4 provides further details. Appendix D.4 demonstrates robustness to alternative calibrations of the extended models that we introduce in section 4.

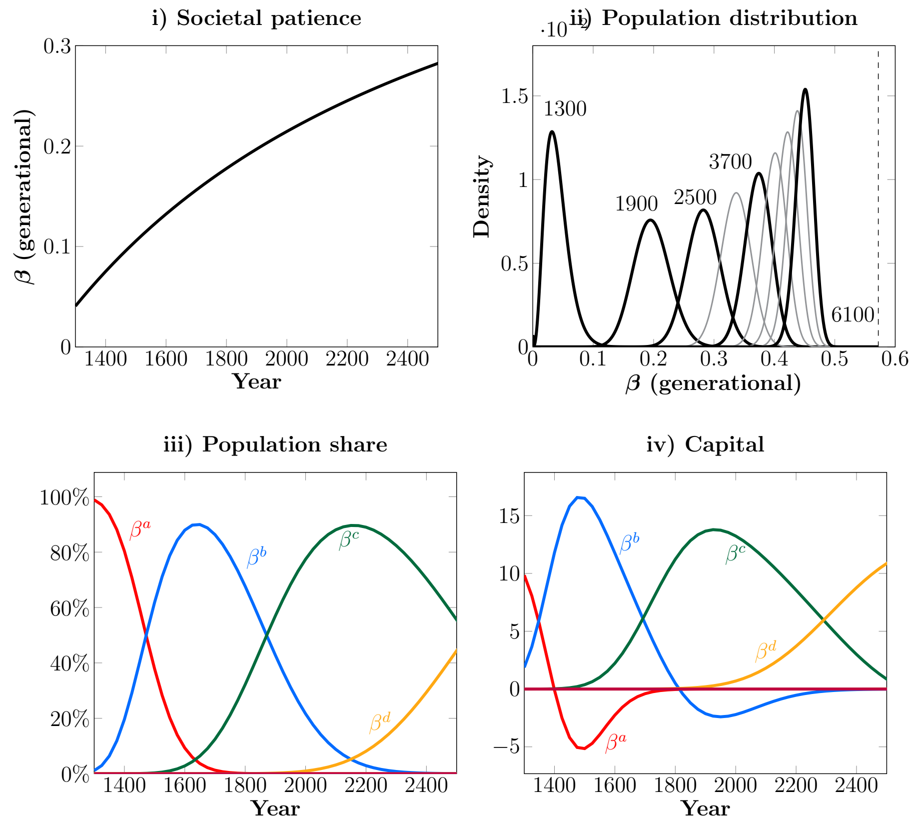
3.3 Quantitative results

Figure 2 panel i) shows a monotonic increase in the level of societal patience (the mean level of patience in the economy) over time and panel ii) shows the distribution across agents at different points in time. In 1300, societal patience is low and virtually no-one belongs to the dynasties with $\beta > 0.2$ (an annual discount factor of around 0.94). More patient households however, will tend to have more children who in turn will have the same higher levels of patience as their parents. The distribution of the population will thus shift towards higher levels of patience as relatively more patient households are born. By 1900, the median dynasty will have a generational discount factor of around $\beta = 0.2$. Panel ii) also makes clear that while there is substantial variation in levels of patience across individuals at a point in time, this cross-sectional variation is substantially less than the change in patience as a whole *over time*.

To examine the changing composition of the population over time, we split dynasties into six groups by their level of patience. Panel iii) of Figure 2 shows the relative size of each group over time. The world starts out being dominated by the least patient agents, group β^a, who in the year 1300 account for around 90% of the total population. 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 the population by 1600. This continues until, eventually, the entire population is dominated by the most patient group of agents. 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.

The key to understanding the cyclical outcome for groups lies in Figure 2 panel iv), which reports the evolution of capital owned by each group. 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.

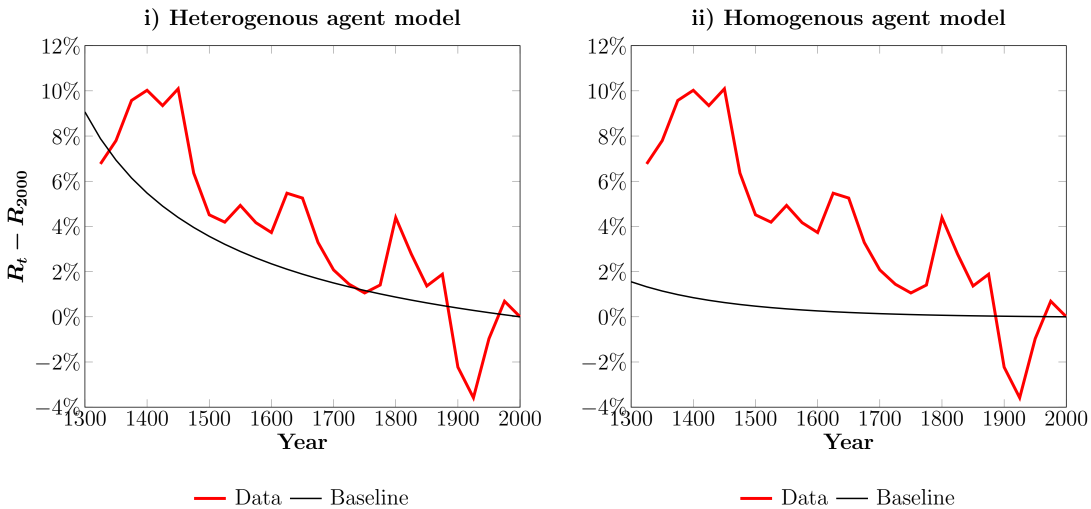
Figure 2: The rise of societal patience



Note: Panel i) depicts the societal average level of generational patience over time. Panel ii) shows the 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 $\bar{\beta}$. Each panel iii)–iv) reports the sum of the model output across all dynasties in the group of dynasties defined in the legend.

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.

Figure 3: Interest rate decline in the baseline model



Note: Data is the Schmelzing (2020) global real interest rate. The model in panel i) is our baseline model with heterogenous agents whose distribution of patience is calibrated at the year 2000. The model in panel ii) is a homogeneous-agent baseline, in which there is one dynasty calibrated to match the average patience in the year 2000 in the heterogenous agent model. Data and model are normalized to zero in the year 2000. We report annualized interest rates.

The interest rate and the role of selection Finally, we turn to the performance of the model in matching the decline in the historical interest rate and examine the role played by selection. Figure 3 panel i) depicts the change in the interest rate predicted by the heterogenous agent model against the Schmelzing (2020) global real interest rate.[18] The model can account for a nine percentage point drop in interest rates, compared to an averaged decline in the data of nearly eleven percentage points. Despite its simplicity, our heterogenous-agent fertility model succeeds in explaining most of the decline in the interest rate.

To isolate the role of selection, we compare our heterogenous agent model to a corresponding homogenous agent model comprised of only one dynasty. In the homogenous agent model, we set the discount factor of the single dynasty to match the mean level of patience in the heterogenous agent model in 2000. Figure 3 panel ii) shows the interest rate in the homogenous agent model. This points to a decline of less than two percentage points, which occurs due to a (counterfactually) slowing

[18]We normalize the interest rate to the year 2000 since we do not target the year 2000 interest rate in Schmelzing explicitly.

growth rate that arises because of convergence dynamics. Specifically, as population increases, consumption grows but it does so at a decreasing rate as the economy approaches its steady state capital-labour ratio. By equation (1), this slowing growth rate of consumption depresses the real rate, even with a fixed β.

The difference between the heterogenous and homogenous models arises solely from the selection mechanism at work. Note that in addition to matching the decline of the interest rate, our model also captures the slowing rate of the decline. In section 4 we introduce extensions to the baseline model in order to incorporate a number of potentially salient features of this period. We then turn to a number of ways to quantify of the role of selection in the baseline and extended models to isolate its contribution to understanding the decline of the interest rate.

4 Extended model

In section 3 we presented a minimal deviation from a standard Barro-Becker fertility model in order to demonstrate the role that selection can play in driving falling interest rates. In doing so, we neglected a number of other key changes over the period. In particular, we now introduce endogenous capital and calibrate the model to fit population growth, income per capita growth and changing life expectancy. We also allow for a Blanchard (1985)-type mechanism in which life expectancy impacts savings. We then evaluate the quantitative performance of these models in accounting for the data.

Extended household problem First, we make capital endogenous by assuming that it is reproducible, depreciates at a rate $\delta > 0$ and can be accumulated via investment of retained output, x_t. Second, we assume that firm productivity, D_t, can vary exogenously over time. Third, we allow child costs to vary according to an exogenous price $q_t \equiv D_t^{\frac{1}{1-\nu}} a_t$. In this expression, the productivity term captures the rising costs of raising children to adulthood, and a_t is an exogenous shock to child costs which we calibrate to match the path of population growth.[19] Fourth, we

[19]Since we have assumed that the cost of raising children is paid in terms of the final good, we need to assume that the goods cost of children grows in proportion to income, much like in Lucas (2002) and Bar and Leukhina (2010), to ensure the existence of a balanced growth path. As in those studies, we take changes in q_t to be a reduced form means to capture laws that prevent children working, the introduction of mandatory education, parents spending more time with children, etc.

allow the probability of death, π_t, (and hence life expectancy) to vary exogenously over time to match the historical evidence. Finally, we allow for a form of imperfect altruism in which parents care about their children's future average utility, but only while they themselves are alive. Changes in life expectancy then influence the interest rate if an individual parent's own finite horizon enters into their decision-making (see Blanchard, 1985). Given the above, each household i's problem at time t is now:

$$U_t^i(k_t^i) = \max_{c_t^i, n_{c,t}^i, x_t^i} \alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i) + \beta^i(\pi_t(1-\omega)+\omega)U_{t+1}^i(k_{t+1}^i). \quad (16)$$

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

In the above $\omega \in [0,1]$ is a parameter that captures a particular form of imperfect altruism that parents may have for their children. If $\omega = 1$, we return to the baseline preferences of section 3 where parents are perfectly altruistic towards their children. Setting $\omega = 0$, introduces a form of selfishness in the sense that parents care about their own future utility and the utility of their descendants, but only so long as they themselves are alive. In the extreme case that survival probability goes to zero, agents would care only about present consumption and the number of children that they have. As such an increase in life expectancy (captured by an increase in the survival probability, π_t) extends the horizon over which parents consider future utility, meaning that parents care more for the future and save more, potentially depressing interest rates. The above problem then aggregates up in a standard fashion for a time zero household and the dynasty planner (see Appendix D for details) with corresponding capital variables representing dynasty aggregates.

The firm's problem remains largely unchanged, with the exception of the exogenously varying D_t and the endogenously varying K_t. Market clearing conditions are also the same, with the exception of the following:

$$C_t + q_t N_{c,t} + X_t = D_t K_t^{\nu} N_t^{1-\nu}, \qquad \sum_{i=1}^{I} K_t^i = K_t. \quad (17)$$

An equivalent approach would be to incorporate the time cost of raising children.

4.1 Calibration of extended model

The calibration, which we summarize here, targets global data and is similar to the baseline (full details are in Appendix D.2). First, productivity D_t is chosen to match observed output growth rates. Second, the depreciation rate is chosen to match a standard annual rate of depreciation of 10% so that $\delta = 1 - (1 - 0.1)^{25} = 0.93$ and the capital share is set to $\nu = 0.33$ (Gollin, 2002). Third, the initial capital stock in the year 1300 is assumed to lie on the saddle path; that is we are assuming no initial jumps in capital stock in 1300.[20] Fourth, we set π_t exogenously to reproduce life expectancy in the data. Fifth, while we do not model the demographic transition explicitly, we capture the path of population by exogenously changing the price of children, q_t, by choosing a_t to match the observed global population over time. The above allows the model to reproduce observed output per worker, fertility rates and population growth in the data.

For all of our quantitative results, we report two forms of this extended model, one with and one without the Blanchard (1985) mechanism. What we refer to as the 'Growth' model, where we set $\omega = 1$, is the baseline model with endogenous capital accumulation and the calibration to match population, income per capita, and life expectancy; what we call the 'Blanchard' model, where $\omega = 0$, is the Growth model with the addition of the full Blanchard mechanism.[21] As can be see in Figure 4, both versions of the extended model match the path of global population and income per capita. The calibrated fertility rate over time is given in Figure 4 panel iii), with an initial increase in fertility at the onset of the industrial revolution followed by a rapid decline that, by construction, matches the pattern in the data (see Galor, 2005; Bar and Leukhina, 2010).

4.2 The interest rate in the extended model

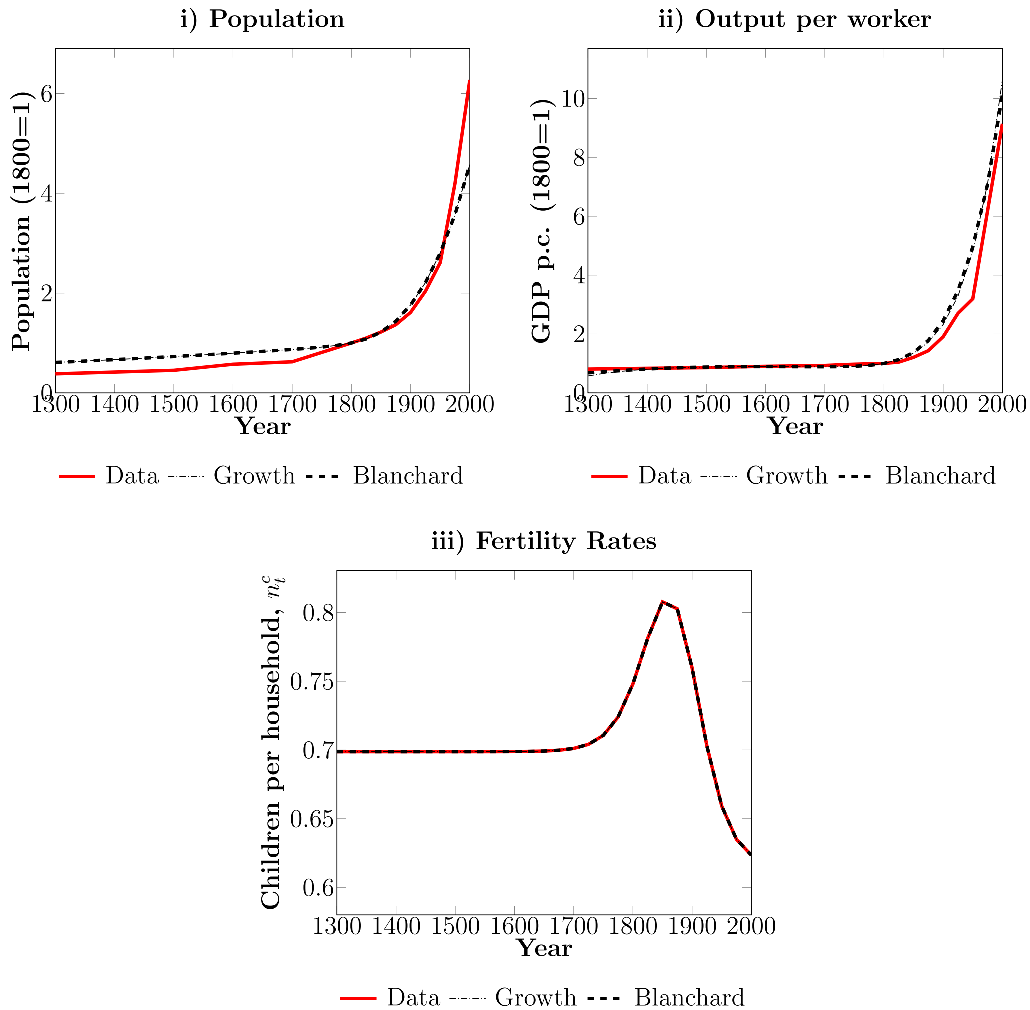
The interest rate in the extended model can be approximated, as we show in Appendix F.3, by the following:

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

[20] This is a simplification that quantitatively has very little impact on the result, since capital adjusts quickly.

[21] Since there is no clear way to calibrate ω, we report the two extreme cases.

Figure 4: Population, output per worker and fertility



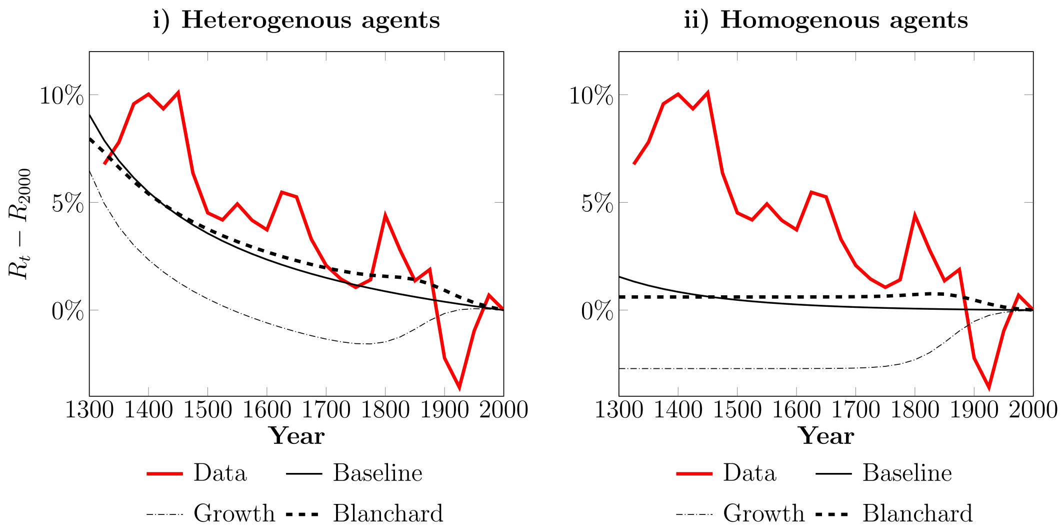
Note: Data for population and GPD is from The Maddison Project (2013). The models are calibrated to match exactly the fertility rates in the data, the derivation of which is described in Appendix D.2. Growth and Blanchard are the models in section 4 with $\omega = 1$ *and* $\omega = 0$ *respectively.*

This approximation is analogous to equation (1) and is a generalization of that derived in the baseline model, equation (15). It illustrates the three key forces driving changes in the interest rate: 1) the time-varying growth rate of consumption as captured by population (g_{Nt}) and productivity growth (g_{Dt}); 2) changes in life expectancy as captured by changing survival probabilities (π_t) in the presence of imperfect altruism ($\omega < 1$); and, 3) selection-driven changing societal patience as captured by the expected value of the discount factor, $E_t(\beta)$. Without our selection mechanism, in a model with homogenous agents, $E_t(\beta)$ is simply a constant and the interest rate is driven by factors 1) and 2) alone. Population and productivity growth increase, especially after the industrial revolution, and thus cannot help explain falling interest rates. Life expectancy also increases during the demographic transition (which occurs after the industrial revolution) and thus can help explain the decline post 1850 but not earlier and not to a quantitatively significant degree. However, with our selection mechanism, $E_t(\beta)$ increases over time and we show that changes in the mean value of beta driven by selection are large enough to explain a large part of the observed decline of interest rates, irrespective of the other mechanisms.[22]

The quantitative results for the path of the interest rate in all models (both the heterogeneous- and homogenous-agent versions) are shown in Figure 5. In the 'Growth' model with heterogenous agents we see a decline in the interest rate of over six percentage points. This is less than in the baseline because, as can be seen in Figure 5 panel i), the increase in productivity growth after 1750 generates a counter-factual increase in interest rates. The homogenous-agent version of this extended model does even worse in explaining the fall in rates, as Figures 5 panel ii) makes clear. In the absence of selection, the rise in productivity growth points to an increase in interest rates of more than two percentage points over this period, which is counter to the data. Finally, the 'Blanchard' model goes some way to mitigate the effect that rising productivity has in causing interest rates to increase. In Figure 5 panel i), the heterogenous-agent Blanchard model points to an eight percentage point decline. In the homogenous-agent model, rising life expectancy completely counteracts growing productivity and population, resulting in the two channels largely cancelling themselves out. Without a selection mechanism, the homogenous-agent model results in a fall from 1300 to 2000 of only 61 basis points in the interest rate as can be seen in Figure 5 panel ii). Thus, here too, selection generates nearly all the change in the interest rate. As is clear, although less than half of one percentage point of the remaining portion of the decline may be explained by increasing life expectancy in the Blanchard (1985) extension (mostly in the latter half of the series), the majority of the decline in the model arises entirely from the effect of selection.

[22]The derivation of which is in Appendix F.3

Figure 5: Interest rate decline in the baseline and extended models

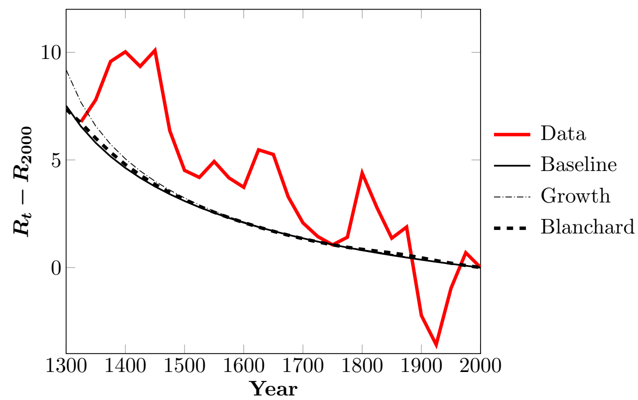


Note: Data is the Schmelzing (2020) global real interest rate. The models in panel i) are the extended models with heterogenous agents whose distribution of patience is calibrated at the year 2000. The models in panel ii) are homogeneous-agent versions, in which there is one dynasty calibrated to match the average patience in the year 2000 in the heterogenous agent model. Data and models are normalized to zero in the year 2000. We report annualized interest rates. Baseline is the model in section 3. Growth and Blanchard are the models in section 4 with $\omega = 1$ *and* $\omega = 0$ *respectively.*

4.3 The role of selection

We measure the contribution of selection as the difference between interest rates in the heterogenous and homogenous versions of each model. The top panel of Figure 6 depicts the contribution of selection against the trend in the data. Selection alone clearly explains a large portion of the total decline in the interest rate in all three models. The bottom panel in Figure 6 quantifies the role of selection. We report three potential measures of success: 1) the absolute decline between the years over which the Schmelzing (2020) data is observed; 2) the average decline (based on a the regression on a time trend); and, 3) the absolute difference between the maximum and the minimum interest rates over the entire period. Selection in the models explains between 97 and 116% of the decline using the first measure, between 59 and 69% using the second measure and between 53 and 69% using the third measure.

Figure 6: Selection and the interest rate



	Data	Baseline	Growth	Blanchard
$R_{1325} - R_{2000}$	6.77	6.56	7.65	6.72
Avg. decline	10.88	6.42	7.32	6.49
max − min	13.67	7.52	9.16	7.36

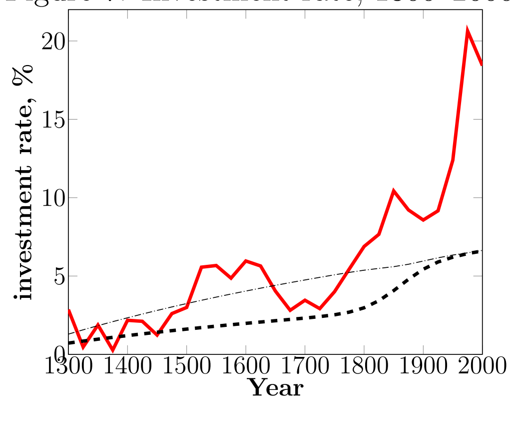
Note: Data is the Schmelzing (2020) global real interest rate. Baseline is the model in section 3. Growth and Blanchard are the models in section 4 with $\omega = 1$ and $\omega = 0$ respectively. The data and model outputs are normalized to zero in the year 2000. We report annualized interest rates. The contribution of selection is the difference between the heterogenous- and homogenous-agent model results. $R_{1325} - R_{2000}$ is the difference between the initial and final observation; Avg. decline is calculated from a regression of the model or data on a time trend, which is then used to obtain an average decline over the whole period; max − min *is the maximum less the minimum over the whole period. Appendix Table 8 reports the metrics for heterogenous-agent and homogenous-agent models separately.*

5 Further quantitative implications

In the extensions to the baseline model, we found that selection remained key to matching the decline in the interest rate. An advantage to using a quantitative model is that its implications can be compared to additional data not targeted by the calibration which provides further external validity to the model's mechanisms. Below we briefly discuss four such exercises by examining the path of the savings rate, the relationship between fertility and income and between wealth and surviving children, and, the response of the interest rate to shocks.

Saving and investment An implication of steadily rising average patience is that savings rates should also increase over time. Sutch, ed (2006) documents an increase in the gross private saving rate in the U.S. from around 10% to over 30% over the

Figure 7: Investment rate, 1300–2000



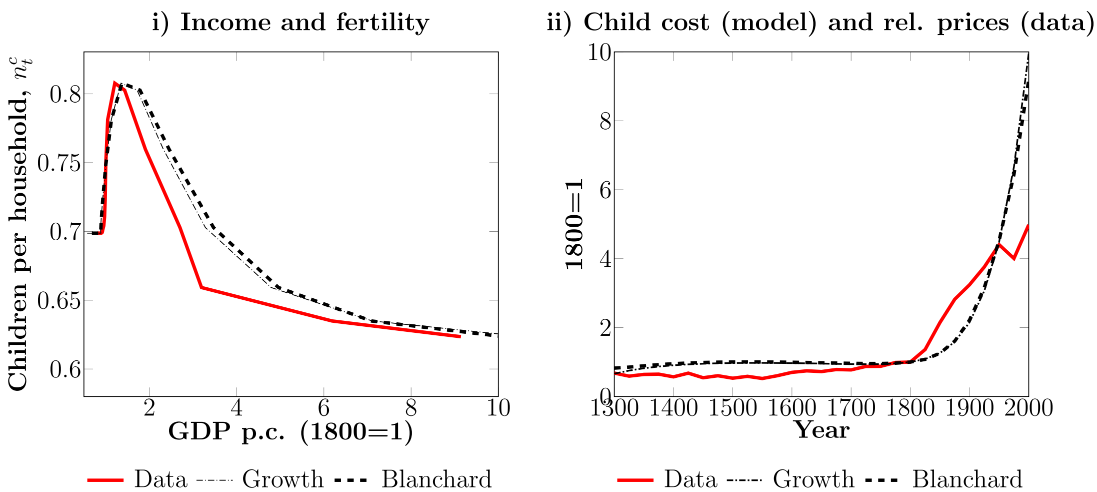
Note: The data for the investment rate is for Broadberry and de Pleijt (2021) over the period 1300–1825 (which covers England up to 1700 and Great Britain thereafter) and from Thomas and Dimsdale (2017a) for 1875–2000. For 1850 we take the mean of these two series. Growth and Blanchard are the models in section 4 with $\omega = 1$ and $\omega = 0$ respectively. Recall that aggregate capital is fixed in the baseline model thus aggregate investment rate is zero and hence not shown.

period 1834–1909. The savings rate then declines, particularly in the latter half of the twentieth century.[23] In the UK, the household savings rate grew from less than 5% in the nineteenth century to around 10% in 2000 (Thomas and Dimsdale, 2017a). Data prior to the nineteenth century on household savings is scarce, but data on investment is available for the UK. Figure 7 reports the UK investment rate from 1300 to 2000 (using Broadberry and de Pleijt, 2021 and Thomas and Dimsdale, 2017a). As can be seen, the trend in the investment rate is increasing over the 700 years, accelerating after 1900. The investment rate in the models with capital accumulation are also reported. As would be expected, the investment rate increases over time as societal patience grows. The 'Growth' model captures the level and the trend in the investment rate up to around 1800. Given that nothing in the calibration is intended to target these objects, we consider the fit to the data to be validation of the role of increasing societal patience.

Fertility, child costs and income The evolution of fertility rates over time is shown above in Figure 4 panel iii). In the data, both time-series and cross-country

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

Figure 8: Fertility, child costs and income



Note: Output per worker and fertility are as in Figure 4. The child cost is the q_t which via a_t calibrated to match the population in the data. The relative price series is the cost of services relative to industry as calculated in Appendix D.3. Both series are normalized to 1800=1. In all panels, Growth and Blanchard are the models in section 4 with $\omega = 1$ and $\omega = 0$ respectively. See text and Appendix D.2-D.3 for details of the calibration.

evidence point to a negative correlation between income and fertility that arises once countries have passed a certain level of income.[24] To match the trend in global population over time in Section 4, we exogenously varied the path for the parameter a_t that governs the cost of children. Since, by construction, we capture the time path for fertility in a growing economy, we also capture this negative relationship, as shown in Figure 8 panel i). A proxy for the cost of children[25] is the price level of services relative to industry (see Appendix D.3). Figure 8 panel ii) shows a close fit between the relative price of children implied by our model and the proxy for the child cost in the data, suggesting that our use of child costs to fit population series is plausible.

Our model is thus able to capture the negative correlation between income and fertility at the aggregate level. However, since higher patience leads to greater accumulation of wealth, as well as higher fertility, our model points to a positive individual-level correlation between wealth and fertility. Recent evidence in Kolk (2022) finds a positive relationship between lifetime accumulated income and fertility for men in

[24] See, for example, Galor and Weil (2000), Jones and Tertilt (2008) and Manuelli and Seshadri (2009).

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

Sweden. Evidence from household-level data using exogenous wealth shocks supports a positive causal relationship between income and fertility.[26] Kearney and Wilson (2018) use the fracking boom to isolate exogenous variation in income, finding a positive relationship between (male) earnings and household fertility. Bennett et al. (2021) use the discovery of oil and gas in the North Sea as an unexpected shock, find a positive relationship between income and the number of children. Variation in the cost of children over time could proxy for an increase in investment per child akin to the growing importance of human capital. It could also be a reduced form means to account for laws that prevent children working, the introduction of mandatory education, parents spending more time with children, and so on.

In short, the model captures the negative relationship between income and fertility at the aggregate level over time, while simultaneously remaining consistent with the micro evidence on a positive individual-level relationship between wealth and the number of children.

Wealth and survival 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.[27]

To compare the results of our model with the bequest data, in Appendix E.1 we obtain an expression for the expected number of surviving children, and the size of bequests, at the period of death of the parent. Appendix Figure 13 reports our model prediction (in the baseline set-up) 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.

The mechanism generating the positive relationship between wealth and children

[26] See Black et al. (2013) and Lovenheim and Mumford (2013).

[27] While 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.

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'.

Shocks: pandemic and war While our models are entirely deterministic, we can examine the effects of unexpected shocks, like those arising from a pandemic or war, on the distribution of population or capital and compare them to data. We conduct these counterfactuals in the baseline model and summarize the implications below. Appendix E.2 provides all the details.

We model a pandemic as an event where a constant fraction of each dynasty unexpectedly dies at the start of a period. The net capital holdings of the deceased households are re-distributed equally among remaining members of the dynasty. We suppose a 30% mortality rate, a shock of similar magnitude to the medieval Black Death. A pandemic means that capital is relatively abundant and so the rate of return on capital, the interest rate, drops in the period of the shock by nearly 1.5 percentage points. 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.[28] This pattern for interest rates is consistent with the findings in Jordà 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, a pandemic acts to reduce the level of wealth inequality in the economy. The relative scarcity of workers after a pandemic drives up wages for those who survive and results in households that rely more on wages than rental income to accumulate greater quantities of capital. 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.

Next, we consider a counterfactual akin to a large war. We model this as a permanent destruction of 30% of each dynasty's net capital holdings. We see an immediate increase in interest rates followed by lower than baseline rates driven by lower population growth, and an increase in inequality that decays over time due

[28] Higher population growth rates lead to higher interest rates since they make an investment today be worth more tomorrow since capital will be relatively more scarce, given the higher population.

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 model. Furthermore, Vandenbroucke (2014), finds that fertility falls as a result of war due to the chance of lost future household income. In our counterfactual, war causes a permanent decline in steady state output, and lower population in the long run. While not exactly the same mechanism, our simple counterfactual and the careful empirical analysis in Vandenbroucke generate similar implications.

6 Imperfect transmission

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

To address this issue we consider a simple extension to the baseline model where some portion of children in dynasty i inherit their level of patience from $\{\beta^i - \varepsilon, \beta^i + \varepsilon\}$ for some $\varepsilon > 0$. Children with a positive shock have what is called an advantageous mutation (one that increases fitness); whilst those with a negative shock have a deleterious (which decreases fitness) mutation. One consideration is whether such noise is distributed symmetrically or whether it is skewed toward the mean of the population. While Brenøe and Epper (2018) and Chowdhury et al. (2022) find strong transmission in patience across generations, neither study identifies the existence of asymmetric transmission. We thus consider the case of symmetric noisy transmission and its implications.

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

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

7 Concluding remarks

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

Finally, we noted in the introduction that understanding social discount rates is critical in formulating optimal policies to address very long-term, inter-generational problems such as those that relate to the funding of social security programmes and that address climate change. 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 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 Dynmics*, 2010.

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

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

Becker, G. S. and C. B. Mulligan, "The Endogenous Determination of Time Preference," *The Quarterly Journal of Economics*, August 1997, *112* (3), 729–758.

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

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

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

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

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).

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

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

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio, "Demographics and Real Interest Rates: Inspecting the Mechanism," *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, Shyamal, Matthias Sutter, and Klaus Zimmermann, "Economic Preferences across Generations and Family Clusters: A Large-Scale Experiment in a Developing Country," *Journal of Political Economy*, April 2022, p. 720395.

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

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

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

_ , *A Farewell to Alms*, Princeton University Press, 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.

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

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

Crafts, Nicholas and Terence Mills, "Economic models vs 'techno-optimism': Predicting medium-term total factor productivity rates in the US," July 2017.

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

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

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti, "Global Trends in Interest Rates," September 2018. Federal Reserve Bank of New York Staff Reports, No. 866.

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

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.

Donni, Olivier, "Measuring the Cost of Children," *izawol*, 2015.

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

Epstein, Larry G. and J. Allan Hynes, "The Rate of Time Preference and Dynamic Economic Analysis," *Journal of Political Economy*, August 1983, *91* (4), 611–635.

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.

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

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

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

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

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

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

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

Jones, Larry E. and Michèle Tertilt, "Chapter 5 An Economic History of Fertility in the United States: 1826–1960," in "Frontiers of Family Economics," Vol. 1, Bingley: Emerald (MCB UP), 2008, pp. 165–230.

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.

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

Kuralbayeva, Karlygash and Radoslaw Stefanski, "Windfalls, Structural Transformation and Specialization," *Journal of International Economics*, 2013, *90*, 273–301.

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.

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

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

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

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

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

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

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

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

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

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

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

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

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim, "Patience and Comparative Development," *The Review of Economic Studies*, December 2021, p. rdab084.

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

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

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

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

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

Thomas, R. and N. Dimsdale, "A Millenium of UK Data," *Bank of England OBRA Dataset*, 2017.

Thomas, R and N Dimsdale, "A Millennium of UK Data," 2017.

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

Vandenbroucke, Guillaume, "Fertility and Wars: The Case of World War I in France," *American Economic Journal: Macroeconomics*, April 2014, *6* (2), 108–136.

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

World Population Prospects 2019, Online Edition. Rev. 1.*, 2019.*

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.*

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

ONLINE APPENDIX

A Data Appendix

A.1 Detail on interest rate data

Schmelzing (2020) considers a number of measures of real interest rates over time, which vary by the asset class and region. Figure 1i) reports the 25 year medians of the headline annual series 'Global R' in real terms (see sheet II column N of the data appendix accompanying Schmelzing, 2020). In this section we describe these measures and supplement them with country-specific real rates of return on land based on the work of Clark (1988). The findings all point to a centuries-long downward trend in real interest rates – regardless of the measure used and regardless of the region under examination.[29]

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

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

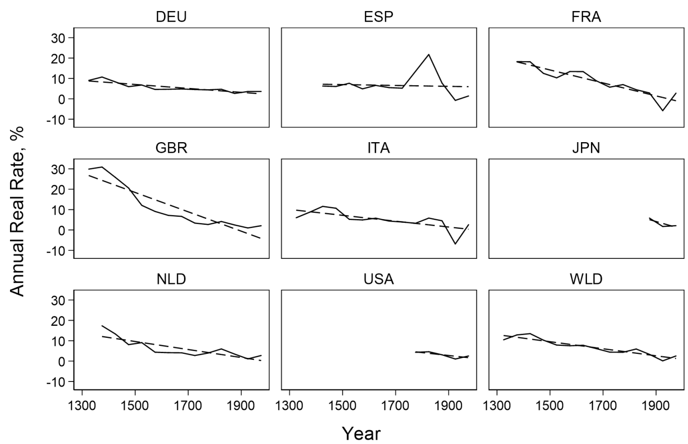
UK, Holland/NL, Germany, France, United States, Spain and Japan. Data for each country consists of long-term debt yields. For countries and time periods included in the global 'safe' series, the debt instruments remain the same and consist of the sovereign debt discussed above. For countries and/or periods not covered in the 'safe' series, observations are arithmetically weighted on the country-level across data points of long-term consolidated debt (such as debt issued by municipalities or mortgage-like pledge loans) and sovereign personal loans (like loans to the British Crown or French Revolutionary war loans to the United States) until marketable, national bond data becomes available. The nominal rates of return are deflated using country-specific price data from Allen (2001). As can be seen in the first panels graphs of Figure 9**i)**, 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 9**i)**, the global real rate of return is steadily declining over the entire period.

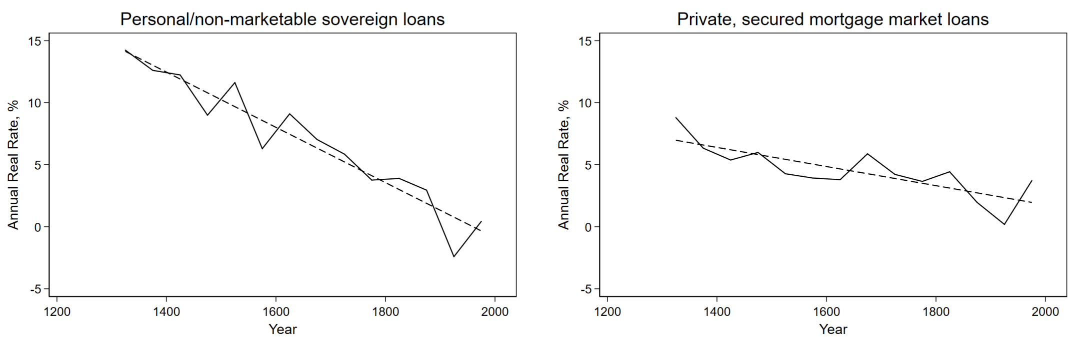
'Personal' or 'Sovereign' non-marketable loans Schmelzing also examines the extent to which the non-marketability of loans can account for the decline of interest rates presented above by examining personal loans to sovereigns (including "pledge loans" and loans from municipalities to the central authorities).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

Figure 9: Country specific real rates of return on long-term debt and land. Dashed line show regression trends.

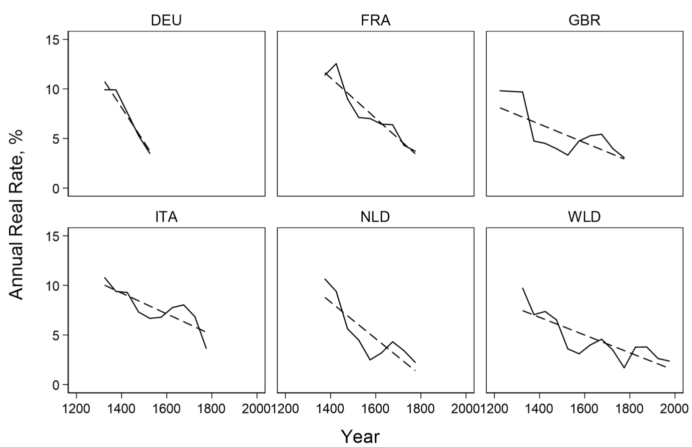
i) Rates of Return on long-term debt



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)



loans, loans featuring in-kind payments, forced loans and those which are de facto expropriations'. The prices are adjusted for inflation using arithmetically weighted inflation rates from Allen (2001). The results are shown in the first panel of Figure 9**ii)**; 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 9**ii)** 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 9**iii)** show the real rates of return on land – arguable the 'safest asset' – for 5 countries (Italy, U.K., Flanders, France and Germany).[30] 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 9**iii)**. 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

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

in every country in which long-term data is available.

In addition to data for the last eight centuries, there is also evidence of an even longer-run trend from ancient data, as shown in Table A.1.

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

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

A.2 The German Socio-Economic Panel

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

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

Equation (9) 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 (9) 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,

$$children_{i,2013} = \beta_0 + \beta_1 patience_{i,2008} + \mathbf{X}_i' \boldsymbol{\beta} + \varepsilon_i \tag{19}$$

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) 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

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**	0.016***	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	no	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 $25 < HHage <= 35$*; mediumold is a dummy equal to 1 if* $35 < HHage <= 45$*; and, old is a dummy equal to 1 if* $45 < HHage$*. Controls are dummy variables for whether an observation is male, the household head, married, widowed, divorced or separated.*

A.3 Steady state consumption share

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

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

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

23 countries is weighted by its starting-year equity market capitalization. (It) also compute(s) a similar world bond (and treasury) index, weighted by GDP."

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

$$r^s_{c,t+1} = \left(\frac{R^s_{c,t+1}}{R^s_{c,t}}\right) - 1, \tag{20}$$

and

$$\bar{r}^s_{c,t+25} = \left(\frac{R^s_{c,t+25}}{R^s_{c,t}}\right)^{\frac{1}{25}} - 1. \tag{21}$$

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%.

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

B Equation (1): derivation and discussion

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

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

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

$$\tilde{r}_t^L = \gamma g_t - \frac{\gamma^2}{2}\sigma_t^2 - \ln \beta + \gamma d_{L,t}. \tag{22}$$

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

B.1 Derivation

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

$$\max_{L_t, B_t} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t U(C_t) \tag{23}$$

subject to,

$$L_{t+1} + B_{t+1} = R_t^L L_t + R_t^f B_t + W_t - C_t \tag{24}$$

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

Optimal portfolio choices satisfy,

$$R_{t+1}^{f}\mathbb{E}_{t}\frac{\beta U'(C_{t+1})}{U'(C_{t})}=1, \tag{25}$$

$$\mathbb{E}_{t}R_{t+1}^{L}\frac{\beta U'(C_{t+1})}{U'(C_{t})}=1. \tag{26}$$

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

$$U(C_t)=\frac{1}{1-\gamma}C_t^{1-\gamma}, \tag{27}$$

and so the optimal portfolio satisfies,

$$R_{t+1}^{f}\mathbb{E}_{t}\beta\left(\frac{C_{t+1}}{C_{t}}\right)^{-\gamma}=1, \tag{28}$$

$$\mathbb{E}_{t}R_{t+1}^{L}\beta\left(\frac{C_{t+1}}{C_{t}}\right)^{-\gamma}=1. \tag{29}$$

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,

$$\left[\begin{array}{c} g_{t+1} \\ r_{t+1}^{L} \end{array}\right]\sim N\left(\left[\begin{array}{c} \bar{g}_{t+1} \\ \bar{r}_{t+1}^{L} \end{array}\right],\left[\begin{array}{c} \sigma_{g,t}^{2},\sigma_{g,L,t}^{2} \\ \sigma_{g,L,t}^{2},\sigma_{L,t}^{2} \end{array}\right]\right). \tag{30}$$

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,

$$\beta\exp\left\{r_{t+1}^{f}-\gamma\bar{g}_{t+1}+\frac{1}{2}\mathrm{var}_{t}\left(-\gamma g_{t+1}\right)\right\}=1 \tag{31}$$

$$\beta\exp\left\{\bar{r}_{t+1}^{L}-\gamma\bar{g}_{t+1}+\frac{1}{2}\mathrm{var}_{t}\left(r_{t+1}^{L}-\gamma g_{t+1}\right)\right\}=1. \tag{32}$$

Note that from (31) we have the following expression for the real rate,

$$r_{t}^{f}=\gamma g_{t}-\frac{\gamma^{2}}{2}\sigma_{g,t}^{2}-\ln\beta. \tag{33}$$

where with log utility ($\gamma \to 1$) and no consumption growth variance ($\sigma^2_{g,t} = 0$), 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,

$$\bar{r}^L_{t+1} + \frac{1}{2}\sigma^2_{L,t+1} = r^f_{t+1} + \gamma\sigma^2_{g,L,t+1} \tag{34}$$

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

$$\ln \mathbb{E}_{t-1} R^L_t = r^f_t + \gamma\sigma^2_{g,L,t} \tag{35}$$

which, with $\tilde{r}^L_t = \ln \mathbb{E}_{t-1} R^L_t$ and $d_{L,t} = \sigma^2_{g,L,t}$, is the expression given in equation (22).

B.2 Discussion

As we discussed in the paper, and as we develop in the extended versions of the model, the historical record for per capita growth and life expectancy are unable to explain the fall in rates over time. Equation (22) suggests a number of additional potential channels.

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

Risk aversion Note that the relationship between relative risk aversion (γ) and the risk-free rate depends, by (22), 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.[31] 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."

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

C Baseline model

C.1 Competitive equilibrium

In the baseline model, 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.

C.2 Model derivations

The following expands on elements of the model solution, as described in Sections 3-3.1.

Time zero household problem Since households care about the outcomes of their future children, we can simplify the problem in (2) by iterative substitution, and re-write the individual household problem in the framework of a time zero household of each type as follows:

$$\max_{\{c_t^i, n_{c,t}^i, x_t^i\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} (\beta^i)^t \left(\alpha \log(c_t^i) + (1-\alpha) \log(n_{t+1}^i)\right) \tag{36}$$

s.t.

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

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

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

The above reflects the choice of an individual time zero adult household.

Household problem We can re-write the dynastic 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:

$$\max_{C_t^i, K_{t+1}^i, N_{t+1}^i} \sum_{t=0}^{\infty} (\beta^i)^t \left(\alpha \log(C_t^i) + (1-\alpha-\beta^i) \log(N_{t+1}^i)\right) \tag{37}$$

$$C_t^i + N_{t+1}^i + p_t K_{t+1}^i \leq (w_t + \pi) N_t^i + (r_t + p_t) K_t^i. \tag{38}$$

The first order conditions for this problem are given by:

$$\lambda_t^i = \frac{\alpha (\beta^i)^t}{C_t^i}, \tag{39}$$

$$\frac{(1 - \alpha - \beta^i)(\beta^i)^t}{N_{t+1}^i} + (\pi + w_{t+1}) \lambda_{t+1}^i = \lambda_t^i \tag{40}$$

$$p_t \lambda_t^i = (p_{t+1} + r_{t+1}) \lambda_{t+1}^i, \tag{41}$$

where, λ_t^i is the Lagrange multiplier on the constraint (38). Now, substituting out for λ_t^i in the last two FOCs using the first FOC, we obtain:

$$\frac{(1 - \alpha - \beta^i)}{N_{t+1}^i} + (\pi + w_{t+1}) \frac{\alpha \beta^i}{C_{t+1}^i} = \frac{\alpha}{C_t^i} \tag{42}$$

and

$$\frac{C_{t+1}^i}{C_t^i} = \beta^i \frac{p_{t+1} + r_{t+1}}{p_t}. \tag{43}$$

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 (9) in the main text.

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

$$w_t = (1 - \alpha) D K_t^{\alpha} N_t^{-\alpha} \tag{44}$$

and

$$r_t = \alpha D K_t^{\alpha - 1} N_t^{1 - \alpha}. \tag{45}$$

Market clearing The market clearing conditions are:

$$\sum_{i=0}^{I} N_t^i = N_t, \tag{46}$$

$$\sum_{i=0}^{I} K_t^i = K_t = \bar{K}. \tag{47}$$

Transversality conditions Finally, there are two transversality conditions per dynasty:

$$\lim_{t\to\infty} (\beta^i)^t u'(C_t^i) K_{t+1}^i = 0, \tag{48}$$

$$\lim_{t\to\infty} (\beta^i)^t u'(C_t^i) N_{t+1}^i = 0, \tag{49}$$

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

Population Euler equation To derive equation (9) 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, (43), 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, (43), we obtain:

$$\frac{N_{t+2}^i}{N_{t+1}^i} = \beta^i \tilde{R}_{t+2}, \tag{50}$$

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:

$$\frac{N_{t+1}^i}{N_t^i} = \beta^i \tilde{R}_{t+1}, \tag{51}$$

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 (9) in the main text.

Steady state In the long run the dynasty with the highest discount factor will come to dominate the entire population, whilst the population of the other dynasties will tend to zero. In particular, continuing to denote by I the dynasty with the highest discount factor, 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:

$$N_{ss}^I = N_{ss} \text{ and } N_{ss}^i = 0 \quad \forall i < I \tag{52}$$

$$K_{ss}^I = K_{ss} = \bar{K} \text{ and } K_{ss}^i = 0 \quad \forall i < I \tag{53}$$

$$C_{ss}^I = C_{ss} \text{ and } C_{ss}^i = 0 \quad \forall i < I. \tag{54}$$

Using the above with the first order conditions (42) and (43) as well as the budget constraint (38), along with the firm's first order conditions, (44)-(45), it follows that the steady state is characterized by:

$$N_{ss} = \left(\frac{D(1 - \alpha - \beta^I + \alpha\beta^I(1-\nu))}{(1 - \pi(1-\alpha))(1-\beta^I)} \right)^{\frac{1}{\nu}} \bar{K} \tag{55}$$

$$C_{ss} = (D\bar{K}^{\nu} N_{ss}^{-\nu} + \pi - 1) N_{ss} \tag{56}$$

$$Y_{ss} = D\bar{K}^{\nu} N_{ss}^{1-\nu} \tag{57}$$

$$p_{ss} = \nu \frac{\beta^I}{1-\beta^I} D\bar{K}^{\nu-1} N_{ss}^{1-\nu} \tag{58}$$

$$w_{ss} = (1-\nu) D\bar{K}^{\nu} N_{ss}^{-\nu} \tag{59}$$

$$r_{ss} = \nu D\bar{K}^{\nu-1} N_{ss}^{1-\nu}. \tag{60}$$

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

Initial population and consumption A useful relationship that we use in solving the model is that between the relative dynastic consumption and relative dynastic population size. To derive this relationship we start by plugging in equation (10) into (7).

$$\frac{(1-\alpha-\beta^i)}{\frac{(\beta^i)^t N_1^i}{\sum_{j=1}^{I}(\beta^j)^t N_1^j}N_{t+1}} + (\pi + w_{t+1})\frac{\alpha\beta^i}{\frac{(\beta^i)^{t+1}C_0^i}{\sum_{j=1}^{I}(\beta^j)^{t+1}C_0^j}C_{t+1}} = \frac{\alpha}{\frac{(\beta^i)^t C_0^i}{\sum_{j=1}^{I}(\beta^j)^t C_0^j}C_t}, \tag{61}$$

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

$$\frac{(1-\alpha-\beta^i)}{\frac{N_1^i}{\sum_{j=1}^{I}(\frac{\beta^j}{\beta^I})^t N_1^j}N_{t+1}} + (\pi + w_{t+1})\frac{\alpha}{\frac{C_0^i}{\beta^I\sum_{j=1}^{I}(\frac{\beta^j}{\beta^I})^{t+1}C_0^j}C_{t+1}} = \frac{\alpha}{\frac{C_0^i}{\sum_{j=1}^{I}(\frac{\beta^j}{\beta^I})^t C_0^j}C_t}, \tag{62}$$

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

$$\frac{(1-\alpha-\beta^i)}{\frac{N_1^i}{N_1^I}N_{ss}} + (\pi + w_{ss})\frac{\alpha}{\frac{C_0^i}{\beta^I C_0^I}C_{ss}} = \frac{\alpha}{\frac{C_0^i}{C_0^I}C_{ss}}. \tag{63}$$

Then, substituting from the solutions of the steady state shown in equations (55)-(60) into the above, for each $i < I$ we can then show that:

$$\frac{C_0^i}{C_0^I} = \frac{N_1^i}{N_1^I}\frac{1-\alpha-\beta^I}{1-\alpha-\beta^i}. \tag{64}$$

C.3 Aggregation and solution algorithm

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 (10) and (64) to relate dynasty- and aggregate-level variables via weighted averages of time zero dynasty-level consumption:

$$C_t^i = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^{I}(\beta^j)^t C_0^j}C_t, \text{ and,} \quad N_{t+1}^i = \frac{(\beta^i)^t(1-\alpha-\beta^i)C_0^i}{\sum_{j=1}^{I}(\beta^j)^t(1-\alpha-\beta^j)C_0^j}N_{t+1}. \tag{65}$$

Substituting (6), (65) and firm first-order conditions 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 (65), market clearing condition (6) and the firm first order condition (44) 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 (45) 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,[32] 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 $\{C_0^i\}_{i=1}^I$:

$$K_t^i = \frac{C_t^i + N_{t+1}^i + p_t K_{t+1}^i - (w_t + \pi) N_t^i}{(r_t + p_t)}. \tag{66}$$

Finally, since we know the distribution of period zero capital across dynasties, then (66) evaluated at $t = 0$, can be used to infer the dynasty distribution of initial consumption:

$$C_0^i = (r_0 + p_0) K_0^i - N_1^i - p_0 K_1^i + (w_0 + \pi) N_0^i. \tag{67}$$

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

C.4 Calibration

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

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:

$$s_{ss}^c \equiv \frac{\alpha \left(1 - \beta^I (1 - \nu(1-\pi))\right)}{(1 - \pi(1-\alpha))\left(1 - \beta^I\right)}. \tag{68}$$

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-

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

2000,[33] we can thus solve the above equations simultaneously to obtain: $\alpha = 0.427$ and $\bar{\beta} = 0.573$.

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 population 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 (9) also holds for $t = 0$ which in turn implies that the second expression for relative population growth also holds at $t = 0$:

$$\frac{N_1^i}{N_1^j} = \frac{\beta^i}{\beta^j}\frac{N_0^i}{N_0^j}. \tag{69}$$

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:

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

Second, the consistency assumption (11) along with (70) amounts to fixing the initial distribution of consumption (relative to the most patient dynasty) according to the following:

$$\frac{C_0^i}{C_0^I} = \frac{\beta^i}{\beta^I}\left(\frac{1-\alpha-\beta^I}{1-\alpha-\beta^i}\right)\frac{N_0^i}{N_0^I}. \tag{71}$$

That is, given the consistency assumption, the initial population distribution tells us what initial consumption distribution should be. Finally, we can use dynastic budget

[33] See Appendix A for details.

constraints (67) to determine what this initial distribution of consumption implies about the initial distribution of capital, $\{K_0^i\}_{i=1}^I$.

Population distribution An immediate implication of Theorem 1 is that we can derive expressions for the mean and variance of generational discount factors at any time t:

$$\mathrm{E}_t(\beta) = \bar{\beta}\frac{\gamma_0 + t}{\gamma_0 + t + \delta} \text{ and } \mathrm{var}_t(\beta) = \bar{\beta}^2 \frac{(\gamma_0 + t)\delta}{((\gamma_0 + t) + \delta)^2(\gamma_0 + t + \delta + 1)} \tag{72}$$

Notice that as $t \to \infty$, the mean beta converges to $\bar{\beta}$ and the variance goes to zero. 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. Given that generational discount factors are distributed according to a scaled beta distribution, we can show that the variance of the annualized generational discount factor, $\beta^{\frac{1}{25}}$, is given by:

$$\mathrm{var}_t(\beta^{\frac{1}{25}}) = \bar{\beta}^{\frac{2}{25}} \frac{\Gamma(\gamma_t + \delta_t)}{\Gamma(\gamma_t)^2} \left(\frac{\Gamma(\gamma_t)\Gamma(\frac{2}{25} + \gamma)}{\Gamma(\frac{2}{25} + \gamma_t + \delta_t)} - \frac{\Gamma(\gamma_t + \delta_t)\Gamma(\frac{1}{25} + \gamma_t)^2}{\Gamma(\frac{1}{25} + \gamma_t + \delta_t)^2} \right), \tag{73}$$

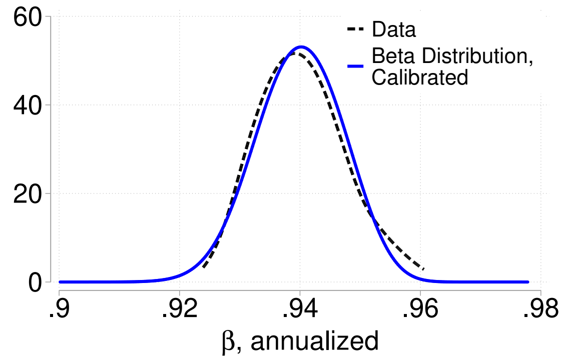
and an approximate expression (see Appendix F) for the annualized gross interest rate:

$$R_t^{\frac{1}{25}} \approx \left(\frac{\gamma_t - 1 + \delta_t}{\bar{\beta}\gamma_t} \right)^{\frac{1}{25}}. \tag{74}$$

As described in Appendix A, we set $\mathrm{var}_{28}(\beta^{\frac{1}{25}}) = 0.0053^2$ to match experimental evidence from representative individuals in Denmark (Andersen et al., 2008) and the individual-level data in the Global Preference Survey (GPS) described in Falk et al. (2018). We set $R_{28}^{\frac{1}{25}} - 1 = 0.063$ to match the average (annualized) generational rates of return on global equities.[34] Together, these two equations imply the following shape parameters of the beta distribution: $\gamma_{28} = 32.089$ and $\delta_{28} = 53.531$. As can be seen

[34]In 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.

Figure 10: 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).

in Figure 10, there is a good fit between the annualized distribution of generational discount factors in the year 2000.

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:

$$\frac{N_{28}^i}{N_{28}} = F\left(\beta^i + \frac{\bar{\beta}}{2I}; 28\right) - F\left(\beta^i - \frac{\bar{\beta}}{2I}; 28\right). \tag{75}$$

With the above proportions in hand, we can then calculate the $t = 0$ distribution of population using equation (12) with $t = 28$, and proceed to solve the model.

D Extended models

In this section we derive the solution and demonstrate the calibration to the extended model in section 4. As we will see below, setting $\omega = 1$ will give us the baseline model where we make capital endogenous and calibrate TFP and child cost; and setting $\omega < 1$ introduces a form of imperfect altruism.

Time zero household problem The time-zero household solves:

$$\max_{\{c_t^i, n_{c,t}^i, x_t^i\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} (\beta^i)^t \left(\prod_{j=0}^{t} (\pi_j(1-\omega)+\omega) \right) \left(\alpha \log(c_t^i) + (1-\alpha)\log(n_{t+1}^i) \right) \tag{76}$$

s.t.

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

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

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

As noted in the text, where the discount factor is common the altruism component, captured in ω and π_j, is distinct from the time preference β.

Dynastic planner problem Dynasty-aggregate values are $C_t^i \equiv c_t^i N_t^i$, $N_{c,t}^i \equiv n_{c,t}^i N_t^i$, $K_t^i \equiv k_t^i N_t^i$, $X_t^i \equiv X_t^i N_t^i$. We can re-write the time-zero household problem for the dynastic planner of each type in the same way as the baseline for each dynasty i:

$$\max_{\{C_t^i, N_{c,t}^i, X_t^i\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} (\beta^i)^t \theta(t,\omega) \left(\alpha \log(C_t^i) + (1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))\log(N_{t+1}^i) \right) \tag{77}$$

s.t.

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

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

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

where in the above $\theta(t,\omega) \equiv \prod_{j=0}^{t}(\pi_j(1-\omega)+\omega)$. Setting $\omega = 1$ reverts to perfect altruism (i.e., shuts down the Blanchard (1985) mechanism). Setting $\omega < 1$ introduces a form of imperfect altruism in which parents care about the outcomes of their children, but only when the parents themselves are alive.

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

$$\max_{\{K_t,N_t\}} Y_t - w_t N_t - r_t K_t, \tag{78}$$

where $Y_t = D_t K_t^{\nu} N_t^{1-\nu}$ is a standard Cobb-Douglas production function where $0 < \nu < 1$ is the output elasticity of capital. D_t is the exogenous and – in contrast to the baseline – potentially time-varying level of technology. Furthermore, in contrast to the baseline, capital is entirely reproducible.

Market clearing The market clearing conditions are given by:

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

$$C_t + q_t N_{c,t} + X_t = D_t K_t^{\nu} N_t^{1-\nu}. \tag{79}$$

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

Competitive equilibrium A competitive equilibrium, given a series of child prices $\{q_t\}_{t=0}^{\infty}$ and technology $\{D_t\}_{t=0}^{\infty}$, 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.

D.1 Solution

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

$$\frac{(1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))}{N^i_{t+1}}+(\pi_{t+1}q_{t+1}+w_{t+1})\frac{\alpha\beta^i(\pi_{t+1}(1-\omega)+\omega)}{C^i_{t+1}}=q_t\frac{\alpha}{C^i_t}, \quad (80)$$

$$\frac{C^i_{t+1}}{C^i_t}=\beta^i(\pi_{t+1}(1-\omega)+\omega)(1-\delta+r_{t+1}), \quad (81)$$

with consumer budget constraints for each dynasty i:

$$C^i_t+q_tN^i_{t+1}+K^i_{t+1}=(w_t+\pi_tq_t)N^i_t+(1-\delta+r_t)K^i_t. \quad (82)$$

The firm first-order conditions are:

$$w_t=(1-\nu)D_tK^\nu_tN^{-\nu}_t \text{ and } r_t=\nu D_tK^{\nu-1}_tN^{1-\nu}_t. \quad (83)$$

The market clearing conditions are:

$$\sum_{i=1}^{I}C^i_t=C_t\ ,\ \sum_{i=1}^{I}N^i_t=N_t\ ,\ \sum_{i=1}^{I}K^i_t=K_t,$$

$$C_t+q_t(N^i_{t+1}-\pi_tN^i_t)+(K^i_{t+1}-(1-\delta)K^i_t)=D_tK^\nu_tN^{1-\nu}_t. \quad (84)$$

Finally, there are two transversality conditions per dynasty:

$$\lim_{t\to\infty}(\beta^i)^tu'(C^i_t)K^i_{t+1}=0, \quad (85)$$

$$\lim_{t\to\infty}(\beta^i)^tu'(C^i_t)N^i_{t+1}=0, \quad (86)$$

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

The long-run From the above, as in the baseline model we can obtain two Euler equations that describe the evolution of dynasty consumption and dynasty population:

$$\frac{C^i_{t+1}}{C^i_t}=\beta^i\bar{R}_{t+1}, t\geq 0, \quad (87)$$

$$\frac{N^i_{t+1}}{N^i_t}=\beta^i\frac{(1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))}{(1-\alpha-\beta^i(\pi_t(1-\omega)+\omega))}\tilde{R}_{t+1}, t\geq 1. \quad (88)$$

where $\bar{R}_{t+1} \equiv (1-\delta+r_{t+1})(\pi_{t+1}(1-\omega)+\omega)$ and $\tilde{R}_{t+1} \equiv R_{t+1}(\pi_t(1-\omega)+\omega)\frac{q_{t-1}R_t-q_t\pi_t-w_t}{q_tR_{t+1}-q_{t+1}\pi_{t+1}-w_{t+1}}$.

Given the above Euler equations, and since the interest rates are common across dynasties, we can derive 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:

$$\frac{C_t^i}{C_t} = \frac{(\beta^i)^t C_0^i}{\sum_{j=1}^I (\beta^j)^t C_0^j}, \text{ and, } \quad \frac{N_{t+1}^i}{N_{t+1}} = \frac{(\beta^i)^t \frac{(1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))}{(1-\alpha-\beta^i(\pi_1(1-\omega)+\omega))} N_1^i}{\sum_{j=1}^I (\beta^j)^t \frac{(1-\alpha-\beta^j(\pi_{t+1}(1-\omega)+\omega))}{(1-\alpha-\beta^j(\pi_1(1-\omega)+\omega))} N_1^j}, \tag{89}$$

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 $t \to \infty$, so $\frac{N_{t+1}^I}{N_{t+1}} \to 1$ and $\frac{C_{t+1}^I}{C_{t+1}} \to 1$ whilst, for all $i < I$, $\frac{N_{t+1}^i}{N_{t+1}} \to 0$ and $\frac{C_{t+1}^i}{C_{t+1}} \to 0$. This means that as with the baseline model the consumption and population of the most patient type will dominate the economy over time. As $t \to \infty$ the extended model collapses to standard homogenous agent model with discount factor β^I and a standard Barro-Becker steady state.

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

$$\alpha \frac{\tilde{c}_{t+1}^i}{\tilde{c}_t^i} g_{Nt+1} a_t = \alpha\beta^i(\pi_{t+1}(1-\omega)+\omega)\left(\tilde{w}_{t+1} + \pi_{t+1}a_{t+1}\right) + (1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))\frac{\tilde{c}_{t+1}^i}{\eta_{t+1}^i}. \tag{90}$$

$$\frac{\tilde{c}_{t+1}^i}{\tilde{c}_t^i} g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}} = \beta^i(\pi_{t+1}(1-\omega)+\omega)\left(1-\delta+r_{t+1}\right). \tag{91}$$

$$\tilde{c}_t^i = \left(\tilde{w}_t + \pi_t a_t\right)\eta_t^i + \left(1-\delta+r_t\right)\tilde{k}_t^i - g_{Nt+1}\left(a_t\eta_{t+1}^i + \tilde{k}_{t+1}^i g_{Dt+1}^{\frac{1}{1-\nu}}\right). \tag{92}$$

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

$$r_t = \nu \tilde{k}_t^{\nu-1} \text{ and } \tilde{w}_t = (1-\nu)\tilde{k}_t^{1-\nu}. \tag{93}$$

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

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

$$\tilde{c}_t + a_t(g_{Nt+1} - \pi_t) + (\tilde{k}_{t+1} g_{Nt+1} g_{Dt+1}^{\frac{1}{1-\nu}} - (1-\delta)\tilde{k}_t) = \tilde{k}_t^{\nu}. \tag{94}$$

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

$$g_{Nss}^I = g_{Nss} \text{ and } g_{Nss}^i = 0 \ \ \forall i < I \tag{95}$$

$$\tilde{k}_{ss}^I = \tilde{k}_{ss} \text{ and } \tilde{k}_{ss}^i = 0 \ \ \forall i < I \tag{96}$$

$$\tilde{c}_{ss}^I = \tilde{c}_{ss} \text{ and } \tilde{c}_{ss}^i = 0 \ \ \forall i < I. \tag{97}$$

Using the above along with the de-trended dynasty first order conditions and budget constraints (90)-(92), and the firm's de-trended first order conditions (93) we obtain a set of 6 equations that characterize the steady state and can be solved for six unknowns, $g_{Nss}, \tilde{k}_{ss}, \tilde{w}_{ss}, r_{ss}, \tilde{y}_{ss}$ and $\tilde{c}_{ss}$:

$$\alpha g_{Nss} a_{ss} = \alpha \beta^I (\pi_{ss}(1-\omega)+\omega)\left(\tilde{w}_{ss} + \pi_{ss} a_{ss}\right) + (1-\alpha-\beta^I(\pi_{ss}(1-\omega)+\omega))\tilde{c}_{ss}. \tag{98}$$

$$g_{Nss} g_{Dss}^{\frac{1}{1-\nu}} = \beta^I (\pi_{ss}(1-\omega)+\omega)\left(1-\delta+r_{ss}\right). \tag{99}$$

$$\tilde{c}_{ss} + a_{ss}(g_{Nss} - \pi_{ss}) + (\tilde{k}_{ss} g_{Nss} g_{Dss}^{\frac{1}{1-\nu}} - (1-\delta)\tilde{k}_{ss}) = \tilde{k}_{ss}^{\nu} \tag{100}$$

$$\tilde{y}_{ss} = \tilde{k}_{ss}^{\nu} \tag{101}$$

$$\tilde{w}_{ss} = (1-\nu)\tilde{k}_{ss}^{1-\nu} \tag{102}$$

$$r_{ss} = \nu \tilde{k}_{ss}^{\nu-1}. \tag{103}$$

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

Initial population and consumption Much like we did for the baseline model, we can take the equations in (89) and use them to replace C_t^i and N_{t+1}^i in equation (80). Then taking the limit of the resulting expression as $t \to \infty$ and using the steady-state first order conditions (98)-(103) we obtain the following relationship for each i:

$$\frac{C_0^i}{C_0^I} = \frac{1-\alpha-\beta^I(\pi_1(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_1(1-\omega)+\omega)}\frac{N_1^i}{N_1^I}. \tag{104}$$

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

We start by re-writing the first order condition (90) and (91) for dynasty I in terms of aggregate population growth rates, g_{Nt}, and de-trended capital, $\tilde{k}_t$, only. To do this, we use equations (89) and (104) as well as the definitions of the de-trended variables above to relate dynasty- and aggregate-level de-trended variables via weighted averages of time zero dynasty-level de-trended consumption:

$$\tilde{c}_t^i = \frac{(\beta^i)^t\tilde{c}_0^i}{\sum_{j=1}^I(\beta^j)^t\tilde{c}_0^j}\tilde{c}_t, \text{ and,} \quad \eta_{t+1}^i = \frac{(\beta^i)^t(1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega))\tilde{c}_0^i}{\sum_{j=1}^I(\beta^j)^t(1-\alpha-\beta^j(\pi_{t+1}(1-\omega)+\omega))\tilde{c}_0^j}. \tag{105}$$

Substituting (93), (94) and (105) into (90) and (91), all evaluated with $i = I$, gives us two first order condition in terms of aggregate population growth rates, $\{g_{Nt}\}_{t=0}^\infty$, de-trended capital $\{\tilde{k}_t\}_{t=0}^\infty$ and initial de-trended consumption distributions, $\{\tilde{c}_0^i\}_{i=1}^I$, only. Assuming that the model converges to its steady state after T periods, we use a reverse-shooting algorithm to solve for $\{g_{Nt}\}_{t=0}^T$ and $\{\tilde{k}_t\}_{t=1}^\infty$ as a function of $\{\tilde{c}_0^i\}_{i=1}^I$. Given this, we can then use (105), the firm first order condition (93) and market clearing condition (94) to solve for $\{\tilde{c}_t^i, \eta_{t+1}^i, \tilde{c}_t, \tilde{w}_t, r_t\}_{t=0}^T$ as functions of $\{\tilde{c}_0^i\}_{i=1}^I$.

Given the above and the assumption that the model converges to steady-state after T periods,[35] we can use the dynasty specific budget constraints to derive sequences

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

of each dynasty's capital stock, $\{\tilde{k}_t^i\}_{t=1}^T$, as functions of $\{\tilde{c}_0^i\}_{i=1}^I$:

$$\tilde{k}_t^i = \frac{\tilde{c}_t^i + g_{Nt+1}\left(a_t\eta_{t+1}^i + \tilde{k}_{t+1}^i g_{Dt+1}^{\frac{1}{1-\nu}}\right) - (\tilde{w}_t + \pi_t a_t)\,\eta_t^i}{(1-\delta+r_t)} \tag{106}$$

Finally, since we know the distribution of period zero capital across dynasties, then (92) evaluated at $t = 0$, can be used to infer the dynasty distribution of initial consumption:

$$\tilde{c}_0^i = (1-\delta+r_0)\,\tilde{k}_0^i - g_{N1}\left(a_0\eta_1^i + \tilde{k}_1^i g_{D1}^{\frac{1}{1-\nu}}\right) + (\tilde{w}_0 + \pi_0 a_0)\,\eta_0^i. \tag{107}$$

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

D.2 Calibration

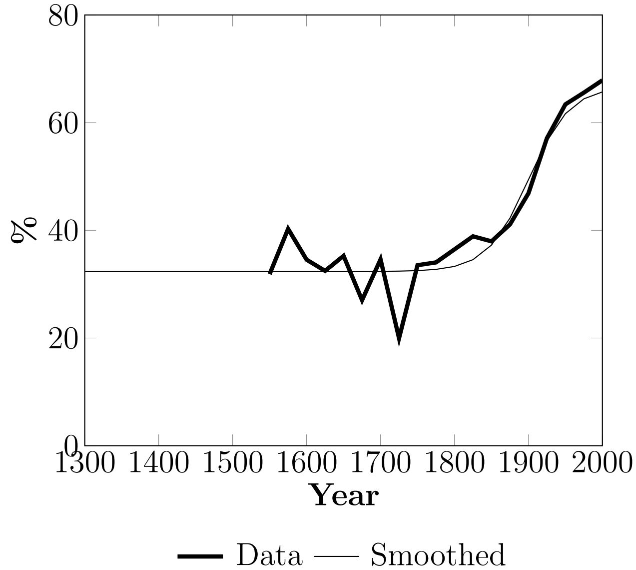
The calibration of the extended models is very similar to the baseline and the calibration of each extended model is also similar to each other. To obtain the calibration of what we call the 'Growth model' we set $\omega = 1$ and to obtain the calibration of the 'Blanchard model' we set $\omega = 0$. We then proceed to choose exogenous parameters which we will feed into the model. All parameter values are summarized in Table 7.

Productivity Productivity in the model, D_t, is chosen to match the median growth rates of world GDP per capita during three time periods that exhibited markedly different growth patterns – 1300 to 1775, 1776 to 1875, and 1876 to 2000. Specifically, we calculate the median generational GDP per capita growth rates during each of those periods (0.84%, 10.06% and 41.48% respectively) using data from The Maddison Project (2013). We then estimate a generalised logistic function using three corresponding productivity growth rates in the model (0.98%, 13.07% and 30.86% in the model with TFP and endogenous capital and 0.98%, 10.46% and 25.89% in the model with the Blanchard mechanism) such that when the model is fed in this implied productivity growth path it generates the observed GDP per capita growth over the three time periods.[36] Finally, the model also requires choosing a long-run

[36] The productivity logistic function is given by: $g(t) \equiv A + \frac{K-A}{1+e^{-B(\frac{t-1275}{25}-M)}}$, where in the model with TFP and endogenous capital $A \equiv 0.98$ is the minimum asymptote, $K \equiv 1.28$ is the maximum asymptote whilst $B = 0.79$ and $M = 22.31$ are the fitted values. In the Blanchard model the

productivity growth rate, g_{Dss}. As Crafts and Mills (2017) argue, predicting future TFP growth rates from past data can be difficult. Nonetheless, both they and an extensive literature have shown consistently declining productivity growth rates that may stay low into the future. They estimate that productivity growth between 2005 and 2016 was approximately 0.5% per year. As such we set our $g_{Dss} = 1.005^{25}$ and assume that the productivity growth rate in the model drops to this level after 2100.

Figure 11: Generational Survival Probability, England.



Survival probabilities We calculate survival probabilities in the model by using data on life expectancy for England and the UK for the period 1543-2020 from Roser et al. (2013) who in turn compile data from Riley (2005), Zijdeman and Ribeira da Silva (2015) and the UN (*World Population Prospects 2019, Online Edition. Rev. 1.*, 2019). We use English data as this offers the longest time span available. We linearly interpolate this data and smooth it using the Hodrick Prescott filter with smoothing parameter of 100. Then, assuming that one generation is 25 years, we calculate the generational expected probability of death in the model as $\Pi_t = 1 - 25/l_t$ - where l_t is life expectancy for generations from 1550 to 2000 at 25 year intervals. Finally, we fit a generalized logistic function to this data in order to generate a smooth transition in

corresponding values are $A \equiv 0.98$, $K \equiv 1.26$, $B = 0.71$ and $M = 21.83$.

life expectancies.[37] We assume a long run survival probability of $\pi_{ss} = 0.667$ which implies a life expectancy of 75 years — this is the same as in the baseline model. Also notice that $\pi_{ss} > \pi_t$ – a fact that we use in our calibration of the patience grid below. We extrapolate this logistic function back to 1300 giving us survival probabilities from 1300 to 2000. The survival probability implied by the data, and the smoothed series we use in the model, is presented in Figure 11.

Child cost Both in the model with endogenous capital and TFP growth as well as the Blanchard model, we generate a demographic transition by choosing child costs, a_t, in such a way as to exactly replicate the evolution of population growth. Specifically, generational population growth rates are chosen to match world population growth rates during three periods that exhibited markedly different growth patterns - 1300-1775, 1775-1875, and 1875 to 2000. We calculate the median population growth rate during each of those periods (2.21%, 10.7% and 28.17% respectively) using data from The Maddison Project (2013). We then smooth the transition between these growth rates using a generalized logistic function. Finally, we choose period-by-period a_t in order to replicate the observed growth rates of population. The model also requires choosing a long-run population growth rate. Herrington (2021) updates the Club of Rome's 1972 'Limits To Growth' report and estimates that global population between 2020 and 2100 may fall by 1% per year. As such we set the long run growth rate to be $g_{Nss} = 0.99^{25}$ and assume that population growth rates in the model drop to this level after 2100.

Patience grid We assign a discount factor to each dynasty $i \in I$. Recall that we order dynasties such that the sequence $\{\beta^i\}_{i=1}^{I}$ is strictly increasing in i. In order to ensure the strict concavity of the objective function we need to assume that $1-\alpha-\beta^i(\pi_{t+1}(1-\omega)+\omega) > 0$. Given this restriction as well as the assumptions that $\lim_{t\to\infty} \pi_t \to \pi_{ss}$ and that $\pi_{ss} \geq \pi_t$, each discount factor can bounded by $0 < \beta^i < \bar{\beta}$, where $\bar{\beta} \equiv \frac{1-\alpha}{\pi_{ss}(1-\omega)+\omega}$. We sub-divide the interval $(0, \bar{\beta})$ into I equally-sized sub-intervals and locate each type's patience level at the central point of every sub-interval, so that, for each i, $\beta^i = \bar{\beta}\frac{(2i-1)}{2I}$. We set the number of types to be $I = 10,000$ in

[37]This logistic function is given by: $g(t) \equiv A + \frac{K-A}{1+e^{-B(\frac{t-1550}{25}-M)}}$, where $A \equiv 0.323$ is the minimum asymptote chosen to match the average probabilities in the first 250 years, $K \equiv 2/3$ is the maximum asymptote chosen to match a long-run life expectancy of 75 years whilst $B = 0.89$ and $M = 14.01$ are the fitted values.

order to obtain a good approximation of continuous distribution. To pin down the sequence of β^is, we need to find values for α and $\bar{\beta}$, β^I, $\tilde{k}_{ss}$ and a_{ss}. We can solve for these five unknowns by noting first that the share of expenditure on consumption relative to aggregate income in the steady-state, $s^c_{ss} \equiv \lim_{t\to\infty} C_t/Y_t$, is an implicit function of the above values and given by:

$$s^c_{ss} \equiv \frac{\tilde{k}^{\nu}_{ss} - (g_{Nss} g_{Dss}^{\frac{1}{1-\nu}} \tilde{k}_{ss} - (1-\delta)\tilde{k}_{ss} - a_{ss}(g_{Nss} - \pi_{ss})}{\tilde{k}^{\nu}_{ss}}. \tag{108}$$

Combining this equation with the first two equations of the steady state first order conditions (98) and (99) as well as the expression $\beta^I = \bar{\beta}\frac{(2I-1)}{2I}$ and $\bar{\beta} \equiv \frac{1-\alpha}{\pi_{ss}(1-\omega)+\omega}$ and, as in the baseline, setting $s^c_{ss} = 0.75$ to match the average global steady-state income share post-2000,[38] allows us to solve for these unknowns: $\alpha = 0.488$, $\bar{\beta} = 0.512$, $\beta^I = 0.512$, $a_{ss} = 0.345$ and $\tilde{k}_{ss} = 0.082$ in the growth model and $\alpha = 0.488$, $\bar{\beta} = 0.7686$, $\beta^I = 0.7685$, $a_{ss} = 0.345$ and $\tilde{k}_{ss} = 0.082$ in the Blanchard model.

Capital distribution We choose total (de-trended) capital so that initial (de-trended) capital stocks pre-1300 lie on a saddle path. To do this we proceed by starting the model in 1275 (instead of 1300) with a guess of the initial capital stock. Since capital stocks adjusts very quickly (one period in the model is 25 years) capital stock in 1300 (i.e. $t = 0$) will be on the saddle path.[39] The initial distribution of this capital stock across dynasties is chosen exactly as in the baseline model. That is, the initial distribution of capital across dynasties determines the population distribution of those dynasties in *subsequent* periods. To obtain the initial capital distribution, we assume that the growth of each dynasty's population is *consistent* with solutions of the model in the period prior to the initial period. Specifically, we assume that outcomes in the period before $t = 0$ are on the equilibrium saddlepath just as much as they are in periods from $t = 0$ on. The initial distribution of capital is thus chosen such that population growth rates are solutions of the model from period $t = 0$. In practice, this means assuming that equation (88) also holds for $t = 0$. This assumption then allows us to re-write equation (104) as:

$$\frac{C^i_0}{C^I_0} = \frac{\beta^i}{\beta^I}\left(\frac{1-\alpha-\beta^I(\pi_0(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_0(1-\omega)+\omega)}\right)\frac{N^i_0}{N^I_0}. \tag{109}$$

[38]See Appendix A for details.

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

or in de-trended terms as:

$$\frac{\tilde{c}_0^i}{\tilde{c}_0^I} = \frac{\beta^i}{\beta^I}\left(\frac{1-\alpha-\beta^I(\pi_0(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_0(1-\omega)+\omega)}\right)\frac{\eta_0^i}{\eta_0^I}. \tag{110}$$

Given this consistency assumption, the initial population distribution tells us what initial consumption distribution should be. Finally, we can use dynastic budget constraints (107) to determine what this initial distribution of consumption implies about the initial distribution of (de-trended) capital, $\{\tilde{k}_0^i\}_{i=1}^I$.

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

$$\frac{\eta_t^i}{\eta_t^I} = \frac{(\beta^i)^t(1-\alpha-\beta^i(\pi_t(1-\omega)+\omega))(1-\alpha-\beta^I(\pi_0(1-\omega)+\omega))}{(\beta^I)^t(1-\alpha-\beta^I(\pi_t(1-\omega)+\omega))(1-\alpha-\beta^i(\pi_0(1-\omega)+\omega))}\frac{\eta_0^i}{\eta_0^I}. \tag{111}$$

Thus, given evidence on the distribution of patience at some later date t, we can infer the initial distribution of the population across levels of patience. However, we face the same problem as before: modern data will capture only a censored portion of the full initial distribution of preference types. To address this issue, we once more impose a distribution on the data. Unlike the baseline however, we will impose a distribution of generational discount factors on the *mortality-adjusted* population, $\tilde{N}_t^i \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}N_t^i$. Notice that when $\omega = 1$ the adjusted and the unadjusted population are identical. When $\omega < 1$ we can relate the distribution of patience within the adjusted population to the distribution of patience in the non-adjusted population using the following theorem.

Theorem 2. *If $I \to \infty$ and dynastic discount factors are distributed according to $\tilde{f}(\beta)$ within the mortality-adjusted population, then dynastic discount factors will be distributed according to the following distribution in the un-adjusted population:* $f_t(\beta) = \frac{\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}}{E_{\tilde{f}_t}\left(\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}\right)}\tilde{f}_t(\beta).$

Proof. See Appendix F. □

Notice that if in any one period the distribution of patience levels in the adjusted-population takes the form of a scaled beta distribution, $\tilde{f}_t(\beta; \gamma_t, \delta_t)$, then from equation (88) the distribution of patience levels in the mortality-adjusted population will satisfy Theorem 1 and, for a fine enough grid, it will follow a scaled-beta distribution in the mortality-adjusted population in all other periods with shape parameters given by $\gamma_{t+1} = \gamma_t + 1$ and $\delta_{t+1} = \delta_t$. Furthermore, recalling that in our calibration we have $1 - \alpha = \bar{\beta}(\omega + (1 - \omega)\pi_{ss})$, using Theorem 2 we can derive an explicit expression for the distribution of patience levels in the un-adjusted population:

$$f_t(\beta) \equiv f(\beta; \gamma_t, \delta) = \frac{(1 - \delta)(\bar{\beta}(\pi_{ss}(1 - \omega) + \omega) - \beta(\pi_t(1 - \omega) + \omega))}{(\bar{\beta} - \beta)((\pi_{ss}(1 - \omega) + \omega)(1 - \gamma - \delta) + \gamma(\pi_t(1 - \omega) + \omega))} \tilde{f}(\beta; \gamma_t, \delta_t). \tag{112}$$

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

$$R_t^{\frac{1}{25}} \approx \left(\frac{\gamma_t - 1 + \delta_t}{\bar{\beta}\gamma_t} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{\omega + (1 - \omega)\pi_t} \right)^{\frac{1}{25}}. \tag{113}$$

Given this, we set $R_{32}^{\frac{1}{25}} - 1 = 0.06$ to match the average (annualized) generational rates of return generated by the baseline model. This is so that both the extended and baseline models predict the same rates of interest in the long run, i.e. when productivity growth rates and child costs have reached their long-run values. We can use this expression and the expression for the variance to determine the parameters of mortality-adjusted distribution: $\gamma_{28} = 31$ and $\delta_{28} = 48$ in the growth model and $\gamma_{28} = 32$ and $\delta_{28} = 50$ in the Blanchard model. Given these parameters, we can use the CDF of the distribution of the discount factors in the adjusted-population, $\tilde{F}$, to approximate, for some I, the proportion of the population assigned to each dynasty

Table 7: Parameters for extended models

Parameter(s)	Value		Target
	Growth	Blanchard	
ω	1	0	Switch between Growth & Blanchard models
D_0	1	1	Normalization
N_0	0.37	0.37	World population, 1300, The Maddison Project (2013)
$\{g_{Dt+1}\}_{t=-1}^{28}$	See Text	See Text	World output per worker growth rates (smoothed), The Maddison Project (2013)
g_{Dss}	1.005^{25}	1.005^{25}	Predicted long run productivity growth rates, Crafts and Mills (2017)
$\tilde{k}_0$	0.003	0.0005	On saddle path (see text)
$\{a_t\}_{t=-1}^{28}$	See Text	See Text	World population growth rates, The Maddison Project (2013)
a_{ss}	0.345	0.345	Predicted future population growth rates from Herrington (2021)
$\{\pi_t\}_{t=0}^{28}$	See Text	See Text	English life expectancy, smoothed
π_{ss}	0.667	0.667	Adult life expectancy of 75
ν	0.33	0.33	Gollin (2002)
I	10,000	10,000	Number of Types
$\{\beta^i\}_{i=1}^{I}$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^{I}$	$\left\{\frac{\bar{\beta}(2i-1)}{2I}\right\}_{i=1}^{I}$	Subdivide domain into grid
α	0.488	0.488	Consumption share (see text)
$\bar{\beta}$	0.512	0.769	Maximum (generational) discount factor
$\{\gamma_{28}, \delta_{28}\}$	$\{31, 48\}$	$\{32, 50\}$	Standard deviation of discount factors (Andersen et al., 2008; Falk et al., 2018) and long run rate of return (see text)
$\{\eta_0^i\}_{i=1}^{I}$	See Text	See Text	Andersen et al. (2008) and Falk et al. (2018)
$\left\{\frac{\tilde{k}_0^i}{\tilde{k}_0}\right\}_{i=1}^{I}$	See Text	See Text	Consistency (see text)

i relative to the most patient dynasty I, in the year 2000 (i.e. period $t = 28$) by:

$$\frac{\eta_{28}^i}{\eta_{28}^I} = \frac{N_{28}^i}{N_{28}^I} = \frac{\frac{1-\alpha-\beta^i(\pi_{28}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{F}\left(\beta^i + \frac{\bar{\beta}}{2I}; 28\right) - \tilde{F}\left(\beta^i - \frac{\bar{\beta}}{2I}; 28\right)}{\frac{1-\alpha-\beta^I(\pi_{28}(1-\omega)+\omega)}{1-\alpha-\beta^I(\pi_{ss}(1-\omega)+\omega)} \tilde{F}\left(\beta^I + \frac{\bar{\beta}}{2I}; 28\right) - \tilde{F}\left(\beta^I - \frac{\bar{\beta}}{2I}; 28\right)}. \tag{114}$$

With the above proportions in hand, we can then calculate the $t = 0$ distribution of population using equation (111) with $t = 28$, and proceed to solve the model.

D.3 Details on relative sectoral prices

We wish to construct a relative price index of the service sector with respect to the industrial sector for the UK. To do this we construct both constant and current prices sectoral value added measures for agriculture (A), industry (I) and service (S) sectors. We then take the ratio of current to constant price sectoral value added which gives us price indices for each sector. Then, dividing the price index of services by the price index of industry results in the required data series.

We proceed as follows.

1. First, we calculate (constant price) sectoral value added shares
    (a) We obtain 1949-2009 constant (2005) price sector value added for A, I, S from Timmer et al. (2015).
    (b) Next using Smits et al. (2009) we obtain 1855-1965 sector size indices normalized to 1965=1.
    (c) Using quantity data from (a) allows us to transform the value added indices in (b) into constant price value added of A, I and S.
    (d) Then we combine the two series using data from Timmer et al. (2015) for 1949-2009 and data from Smits et al. (2009) for 1855-1948. This gives us constant (2005) price sectoral value added for the A, I and S sectors for the years 1855-2009.
    (e) We then use this data to calculate (constant price) sectoral value added shares.
2. Second, we calculate current price sectoral value added shares:

(a) We use current price value added data from Timmer et al. (2015) to calculate current price value added shares for A, I and S for 1960-2009

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

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

4. Fourth, we divide each sector's current price value added by its corresponding constant price value added series to obtain a price index for each sector. This gives sectoral price indices for A, I and S in the UK from 1855-2009.

5. Finally, we obtain sectoral price indices for agriculture, industry and services for the period 1270-1855 from Thomas and Dimsdale (2017b) on tabs A6 and A7. We use the implied growth rates from these series to extend our 1855-2009 sectoral price index data from 4. above backwards to 1270. This can then be used to calculate an index of the relative price of industry to services from 1270 to 2009. This is the data that we use in the main body of our paper.

D.4 Robustness of calibration

Our calibration exercise depends on capturing the modern variation of patience in the population. As we explained in the paper in section 3.2 and in the appendix in section D.2, the calibration of the distribution of patience is based on individual-level data in the Global Preference Survey (GPS) described in Falk et al. (2018). In this section we present a robustness exercise in which we vary the parameters that govern this distribution in order to consider the sensitivity of our results. We do this for the 'Blanchard' version of the model alone, though quantitatively and qualitatively similar results hold in the corresponding exercises with the 'Baseline' and 'Growth' versions of the model. We examine what share of the decline in interest rates is explained by selection when we allow the variance of patience to deviate by no more than 10 percentage points from the variance in the GPS data?[40]

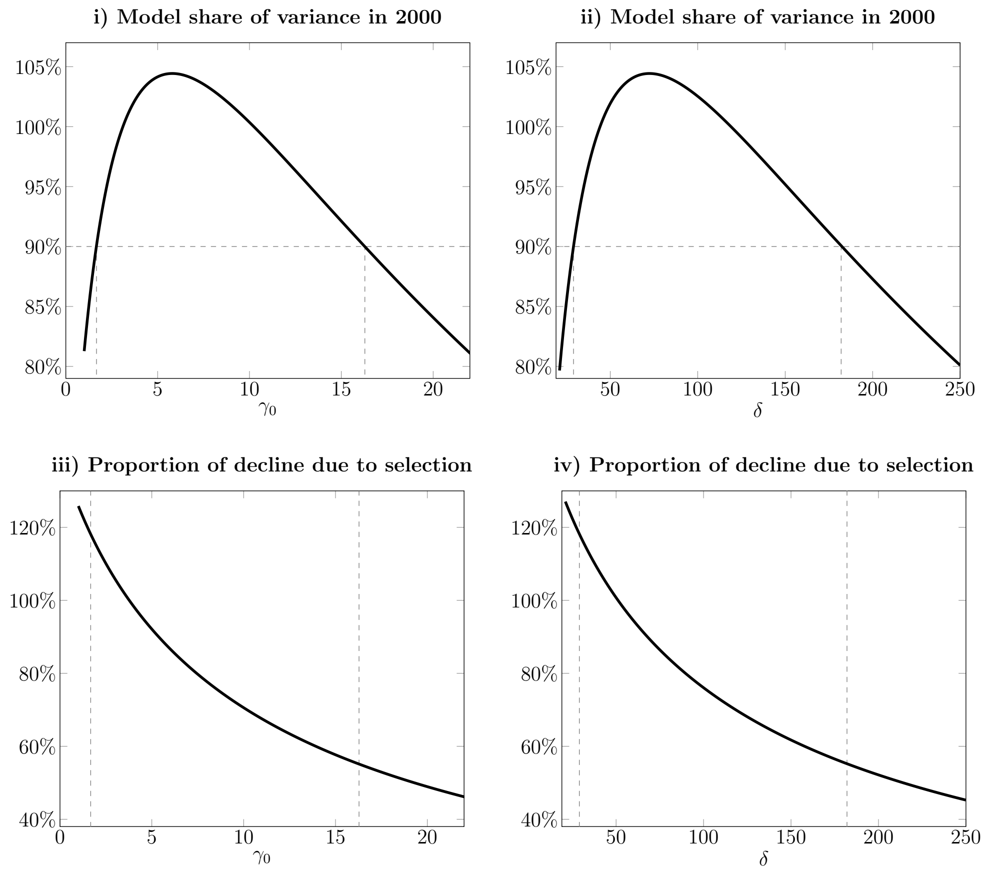
Our robustness exercise consists of independently varying one of either δ or γ_0 in the scaled-beta distribution, $\tilde{f}_t(\beta)$, whilst re-calibrating the other parameter so that the model matches the interest rate predicted by the original Blanchard model in 1325. All other parameters in the calibration are independent of these two parameters and thus remain unchanged. For a given change in each parameter we plot two graphs: 1) the proportion of the decline in the interest rate between 1325 and 2000 that can be explained by selection; and 2) the proportion of the variance in patience observed in the year 2000 (in the GPS data) that can be explained given our parameter choice.

Figure 12 shows the results of the above exercise. Panel i) and ii) the proportion of year 2000 variance in patience explained by the model when changing γ_0 and δ respectively while panel iii) and iv) show the proportion of the decline explained by selection when changing γ_0 and δ respectively. Panels i) and ii) point to a hump-shaped relationship between these parameters and the success of the model distribution in capturing the variance in the data. Figures ii) and iv), however, show that the higher the δ and the higher the γ the smaller the role of selection.

Recall that in the calibration of the Blanchard model we take $\gamma_{28} = 32$ (so that $\gamma_0 = 4$) and $\delta = 50$. Figure 12 points to the substantial robustness in the model to alternative calibration. The model distribution captures more than 90% of the variance in the data so long as γ_0 is within $(1.7, 16.3)$ and δ is within $(29, 182)$. Despite such variation, the contribution of selection to explaining the decline in the interest rate remains strong. The predicted role of selection within both sets of bounds varies from 55%-118%. Thus, for a large range of parameters determining the key object of the model – the distribution of patience across individuals in the year 2000 – the model points to an important role for selection.

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

Figure 12: Blanchard model, varying γ_0 and δ



Note: Panels iii) and iv) report the share of the decline in the interest rate over 1325–2000 that can be explained by the selection mechanism (difference between heterogenous- and homogenous-agent Blanchard model). Panels i) and ii) give the variance in the distribution of patience in the model at the year 2000, as a share of the corresponding variance in the data. Panels i) and iii) vary γ_0 holding δ constant; panels ii) and iv) vary δ holding γ_0 constant. The model is the 'Blanchard' model in section 4 where we set $\omega = 0$. The dashed lines give us the range of the selection-explained decline in the interest rate that results from a deviation from the variance in the data of 10 percentage points.

D.5 Additional figures and tables

In this section we include additional figures and Tables associated with the extended models.

Table 8: Individual model performance against the interest rate

	Data	Heterogenous			Homogenous			Selection		
		A	B	C	A'	B'	C'	ΔA	ΔB	ΔC
$R_{1325} - R_{2000}$	6.77	7.89	5.08	7.33	1.33	-2.57	0.61	6.56	7.65	6.72
Avg. decline	10.88	7.68	4.94	6.80	1.26	-2.37	0.31	6.42	7.32	6.49
max − min	13.67	9.07	8.15	7.96	1.55	-2.57	0.76	7.52	9.16	7.36

Note: Data is the Schmelzing (2020) global real interest rate. Model A is the baseline model with heterogenous agents. Model B is the baseline model where we introduce endogenous capital and calibrate child cost, TFP and life expectancy but set $\omega = 1$ (called "Growth" in Table 6). Model C is model B with the additional mechanism to incorporate imperfect altruism, setting $\omega = 0$ (called "Blanchard" in Table 6). The prime models – A', and so on – are the same models but with homogenous agents (only one dynasty). The ΔA are the selection effects, i.e., the difference between the heterogenous- and homogenous-agent model results. Avg. decline is the result of a regression of the model or data on a time trend, then uses the coefficient on time to obtain an average decline over the whole period; $R_{1325} - R_{2000}$ is the difference between the initial and final observation; max − min *is the maximum less the minimum over the whole period (except for B', the homogenous growth model, since the interest rate increases, where we report the minimum less the maximum).*

E Additional quantitative implications

E.1 Surviving children and bequests

In Section 5 we compare the model to the testamentary data in Clark and Hamilton (2006) and Clark (2007a). In order for us to generate this object in the model, we need to calculate the expected number of surviving children and the expected value of bequest to those children at parents' expected time of death in the model.

To calculate the expected number of children at the point of death, recall that 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,

$$\mathbb{E}\left[s^i(t)\right] \equiv \sum_{j=1}^{\infty} p(t, j) s^i(t, j), \tag{115}$$

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,

$$p(t,j) \equiv \pi^{j-1}(1-\pi). \tag{116}$$

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 $n^i_{c,t}$ children in period t and $n^i_{c,t+1}$ 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,

$$s^i(t,j) \equiv \sum_{k=0}^{j-1} \pi^k n^i_{c,t+j-(k+1)}, \tag{117}$$

where $n^i_{c,t+j-(k+1)}$ 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:

$$\mathbb{E}\left[s^i(t)\right] = \frac{1}{1+\pi} \sum_{m=0}^{\infty} \pi^m n^i_{c,t+m}. \tag{118}$$

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

$$\mathbb{E}\left[s^i(t)\right] = \frac{1}{1-\pi^2} \bar{n}^i_c. \tag{119}$$

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):

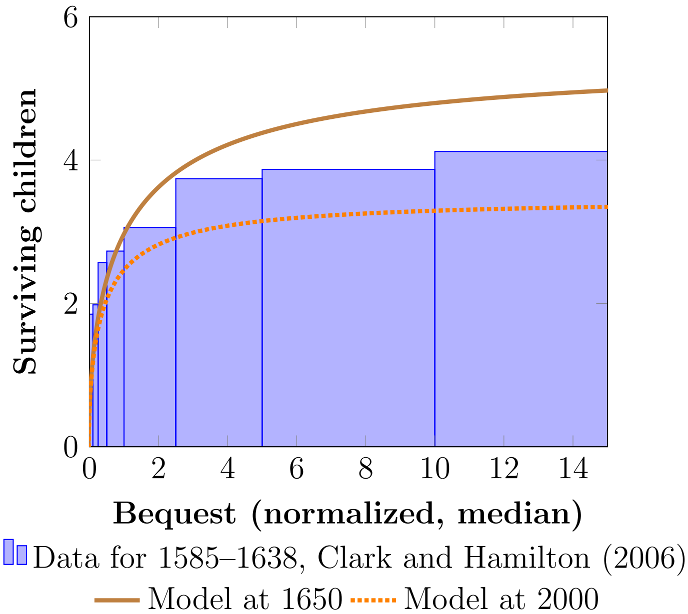
$$\mathbb{E}\left[b^i(t)\right] = (1-\pi) \sum_{m=0}^{\infty} \pi^m k^i_{t+m+1} n^i_{t+m+1}. \tag{120}$$

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

$$b^i(t,j) \equiv n^i_{t+j} k^i_{t+j} \tag{121}$$

Figure 13 reports the model implication in the baseline against the data in Clark and Hamilton (2006) and Clark (2007a).

Figure 13: Wealth and surviving children



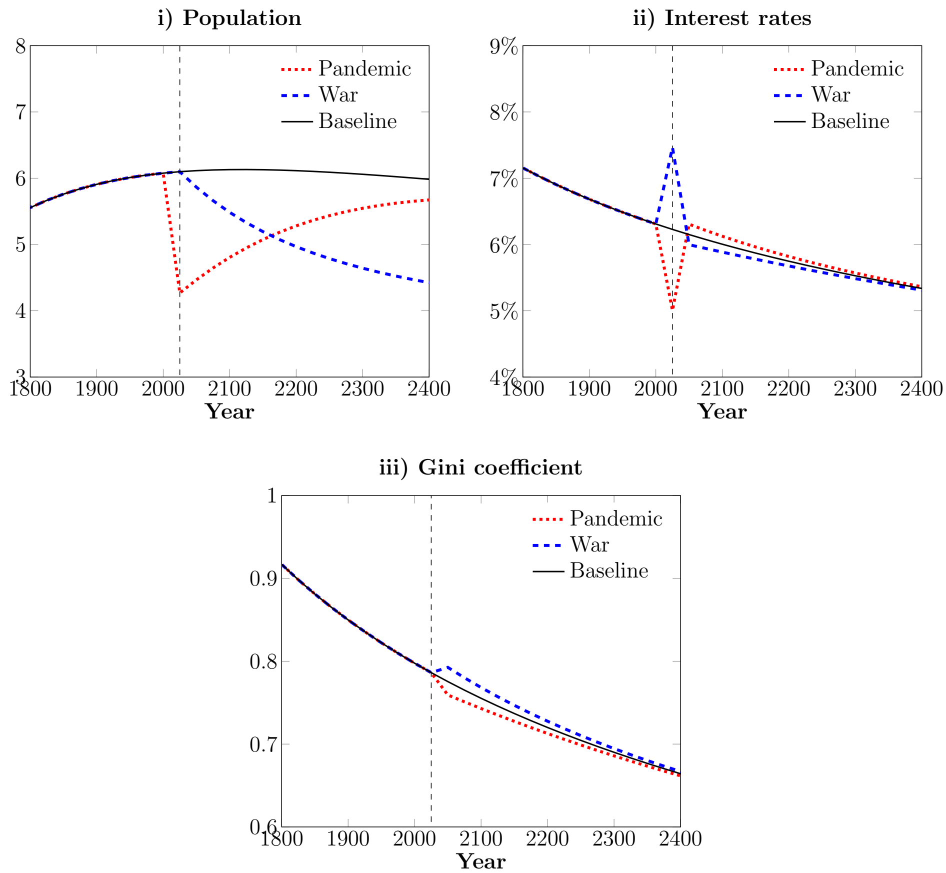
E.2 Shocks

In section 5 we discussed how an unexpected shock to the distribution of population, such as that which might result from a pandemic, and an unexpected shock to the distribution of capital, such as that which might arise from the result of warfare, will influence our economy. Here we describe more fully these counterfactual exercises and their results.

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.[41] 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. First, in Figure 14 panel i), we examine the impact on aggregate population. Immediately after the negative shock to population, households choose to have more children as

[41] For 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.

Figure 14: Pandemic or war



Note: Figures depict the long run consequence of a one-off and unanticipated pandemic (a loss of 30% of the total population) or war (a permanent loss of 30% of the capital stock) in the year 2025 (dashed vertical line).

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. Note, however, that the recovery of the level of population is not instantaneous and remains lower than it would have been without the pandemic for several hundred years. Only around the year 2800 does the population reach close to the same level it would have been. Thus, the effects of pandemics can, along certain dimensions, be very long lasting.

In Figure 14 panel ii), we report the interest rate. A pandemic results in an increase of the capital-labor ratio. Since capital is now abundant, the rate of return on capital, 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.[42] The result for interest rates is consistent with the findings in Jordà 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 inequality. The relative scarcity of workers after a pandemic drives up wages for those who survive and results in households that rely more on wages than rental income to accumulate greater quantities of capital thus reducing wealth inequality. 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. As with a pandemic, 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 heterogenous agent economy adjusts to a sudden decrease in the capital-labor ratio caused by a decrease in the capital stock. The results shown in Figure 14 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 lower 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 qualitatievely consistent with the predictions of our model. This finding is also related to Vandenbroucke (2014), which finds that fertility falls as a result of a war due to the chance of lost future household income. In our model, the war causes a permanent decline in steady state output, and lower population in the long run. While not exactly the same mechanism, our simple counterfactual and the careful empirical analysis in Vandenbroucke generate similar implications.

[42]Higher population growth rates result in higher interest rates since they make an investment in capital today be worth more tomorrow since capital will be relatively more scarce, given the higher population.

E.3 Cross-country considerations

The denormalized Falk et al. (2018) data points to some variation in average discount factors across countries. These differences are much smaller than the variation implied by the model across time. For example, the average patience level in the top and bottom 10 percent of countries in the denormalized Falk et al. (2018) data is $\beta_{10} = 0.934$ versus $\beta_{90} = 0.949$ (on an annualized basis). Using equation (15) this implies a difference of approximately 1.7 percentage points in interest rates between the most and the least patient countries $(1/\beta_{10} - 1/\beta_{90})$. Whilst this is not negligible, compared to the average decline of the interest rate of between 7 and 14 percentage points over the 700-year time span under consideration this difference is relatively unsubstantial. As such throughout the paper we abstract from variation in patience levels across countries and we focus only on variation in individual level patience that gives rise to the large downward trend in interest rates over time.

F Asymptotic results

F.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 $I \to \infty$ 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 I dynasties with discount factors, β^i, distributed evenly along a grid so that $\beta(i; I) = \frac{2i-1}{2I}$ for $i = 1, \cdots, I$. Notice that the distance between any two points is simply: $\Delta(I) \equiv \beta(i+1; I) - \beta(i; I) = \frac{1}{I}$. We define the following function: $\nu_t(\beta(i; I)) \equiv \nu_t^I(\beta^i) \equiv \frac{N_t^i}{N^t}$, which maps the discount factor of a particular dynasty to the fraction of the total population of that dynasty i at time t. Notice, that we can think of this function as a probability mass function of a discrete random variable with realization, $\beta(i; I)$, on the domain $\{\frac{2i-1}{2I} | i = 1, \cdots, I\}$. We wish to characterize the evolution of the asymptotic function, $\frac{\nu_t(\beta(i;I))}{\Delta(I)}$, over time as $I \to \infty$

- that is as the number of dynasties or types becomes infinite. 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 (9) as:

$$N_{t+1}^i = \beta^i \tilde{R}_{t+1} N_t^i. \tag{122}$$

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

$$N_{t+1} = \tilde{R}_{t+1} N_t \sum_{j=1}^{I} \beta^j \nu_t^I(\beta^j). \tag{123}$$

Dividing equation (122) by equation (123) we obtain:

$$\nu_{t+1}^I(\beta^i) = \frac{\beta^i \nu_t^I(\beta^i)}{\sum_{j=1}^{I} \beta^j \nu_t^I(\beta^j)}. \tag{124}$$

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

$$F_t^I(\beta^i) \equiv \frac{\sum_{j=1}^{i} \beta^j \nu_t^I(\beta^j)}{\sum_{j=1}^{I} \beta^j \nu_t^I(\beta^j)}. \tag{125}$$

This also means that:

$$\nu_t^I(\beta^i) = F_t^I(\beta^{i+1}) - F_t^I(\beta^i) = P_t^I(\beta^i \leq \beta \leq \beta^{i+1}). \tag{126}$$

Given the above, notice that (124) can be re-written as:

$$\frac{\nu_{t+1}^I(\beta^i)}{\Delta^i(I)} = \frac{\beta^i \frac{\nu_t^I(\beta^i)}{\Delta^i(I)}}{\sum_{j=1}^I \beta^j P_t^I(\beta^j \leq \beta \leq \beta^{j+1})}. \tag{127}$$

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

$$f_{t+1}(\beta) = \frac{\beta f_t(\beta)}{E_t(\beta)}, \tag{128}$$

where f_t is the continuous probability density function corresponding to the discrete mass function ν_t^I[43] and $E_t(\beta) \equiv \int_0^1 u f_t(u) du = \lim_{I\to\infty} \sum_{j=1}^I \beta^j P_t^I(\beta^j \leq \beta \leq \beta^{j+1})$, 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 (13). The corresponding probability density function of this distribution f is given by:

$$f_t(\beta; \boldsymbol{\theta_t}) \equiv f(\beta; \gamma_t, \delta_t) = \frac{(\bar{\beta} - \beta)^{\delta_t - 1} \beta^{\gamma_t - 1}}{\bar{\beta}^{\delta_t + \gamma_t - 1} B(\gamma_t, \delta_t)}, \tag{129}$$

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

$$E(\beta; \gamma_t, \delta_t) = \bar{\beta} \frac{\gamma_t}{\gamma_t + \delta_t}. \tag{130}$$

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

Using equations (128)-(130), we can write the pdf of discount factors at time $t+1$ as:

$$\begin{aligned} f_{t+1}(\beta;\gamma_t,\delta_t) &= \frac{\beta(\bar{\beta}-\beta)^{\delta_t-1}\beta^{\gamma_t-1}}{\bar{\beta}\frac{\gamma_t}{\gamma_t+\delta_t}\bar{\beta}^{\delta_t+\gamma_t-1}B(\gamma_t,\delta_t)} \\ &= \frac{(\bar{\beta}-\beta)^{\delta_t-1}\beta^{\gamma_t}}{\frac{\gamma_t}{\gamma_t+\delta_t}\bar{\beta}^{\delta_t+\gamma_t}B(\gamma_t,\delta_t)} \\ &= \frac{(\bar{\beta}-\beta)^{\delta_t-1}\beta^{\gamma_t}}{\bar{\beta}^{\delta_t+\gamma_t}B(\gamma_t+1,\delta_t)} \\ &= f(\beta;\gamma_{t+1},\delta_{t+1}) \end{aligned} \tag{131}$$

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

□

F.2 Proof of Theorem 2

The following theorem applies to the extended models and establishes a relationship between the distribution of discount factors in the mortality-adjusted and unadjusted population.

Theorem 2. *If $I \to \infty$ and dynastic discount factors are distributed according to $\tilde{f}(\beta)$, within the mortality-adjusted population, then dynastic discount factors will be distributed according to the following distribution in the un-adjusted population:* $f_t(\beta) = \frac{\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}}{E_{\tilde{f}_t}\left(\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}\right)}\tilde{f}_t(\beta)$.

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

We first derive a relationship between these two distributions. Since $\tilde{\nu}_t^I(\beta^i) \equiv \frac{\tilde{N}_t^i}{\tilde{N}_t} = \frac{\frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i}{\sum_{j=1}^I \frac{1-\alpha-\beta^j(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^j(\pi_t(1-\omega)+\omega)} N_t^j}$, we can re-write this expression to obtain the total population in dynasty i as:

$$N_t^i = \left(\sum_{j=1}^I \frac{1-\alpha-\beta^j(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^j(\pi_t(1-\omega)+\omega)} N_t^j\right) \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i). \tag{132}$$

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

$$N_t = \left(\sum_{j=1}^I \frac{1-\alpha-\beta^j(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^j(\pi_t(1-\omega)+\omega)} N_t^j\right) \left(\sum_{i=1}^I \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i)\right). \tag{133}$$

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

$$\nu_t^I(\beta^i) \equiv \frac{N_t^i}{N_t} = \frac{\frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i)}{\sum_{i=1}^I \frac{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)} \tilde{\nu}_t^I(\beta^i)}. \tag{134}$$

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

$$f_t(\beta) = \frac{\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}}{E_{\tilde{f}_t}\left(\frac{1-\alpha-\beta(\pi_t(1-\omega)+\omega)}{1-\alpha-\beta(\pi_{ss}(1-\omega)+\omega)}\right)} \tilde{f}_t(\beta), \tag{135}$$

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

□

F.3 Asymptotic expression for the rate of interest

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

$$R_{t+1} = \frac{C^i_{t+1}/C^i_t}{\beta^i} = \frac{\left(\frac{\kappa^I_{t+1}(\beta^i)/\Delta(I)}{\kappa^I_t(\beta^i)/\Delta(I)}\right)\frac{C_{t+1}}{C_t}}{\beta^i} \tag{136}$$

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

$$\frac{\kappa^I_t(\beta^i)}{\Delta(I)} = \frac{\frac{\beta^i}{1-\alpha-\beta^i}\frac{\nu^I_t(\beta^i)}{\Delta(I)}}{\sum_{j=1}^I \frac{\beta^j}{1-\alpha-\beta^j}\nu^I_t(\beta^j)}. \tag{137}$$

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

$$f_{ct}(\beta) = \frac{\frac{\beta}{1-\alpha-\beta}f_t(\beta)}{E_t(\frac{\beta}{1-\alpha-\beta})}, \tag{138}$$

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:

$$\frac{f_{ct+1}(\beta)}{f_{ct}(\beta)} = \beta\frac{E_t\left(\beta/(\bar{\beta}-\beta)\right)}{E_t\left(\beta^2/(\bar{\beta}-\beta)\right)} \tag{139}$$

Taking the limit of both sides of (136) as $I \to \infty$ we obtain:

$$R_{t+1} = \frac{E_t\left(\beta/(\bar{\beta}-\beta)\right)}{E_t\left(\beta^2/(\bar{\beta}-\beta)\right)}\frac{C_{t+1}}{C_t}. \tag{140}$$

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:

$$R_{t+1} \approx \frac{E_t\left(\beta/(\bar{\beta}-\beta)\right)}{E_t\left(\beta^2/(\bar{\beta}-\beta)\right)}. \tag{141}$$

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:

$$R_{t+1}^{\frac{1}{25}} \approx \left(\frac{\gamma_t + \delta_t}{\bar{\beta}(1+\gamma_t)} \right)^{\frac{1}{25}}. \tag{142}$$

Finally, a further useful approximation can be made. For high enough δ_t or γ_t it holds that $\frac{1+\gamma_t}{\gamma_t+\delta_t} \approx \frac{\gamma_t}{\gamma_t+\delta_t}$. Given this, and using equation (130) we can write the annualized gross interest rate as:

$$R_{t+1}^{\frac{1}{25}} \approx \left(\frac{1}{E_t(\beta)} \right)^{\frac{1}{25}}. \tag{143}$$

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

$$R_{t+1} = \frac{\tilde{c}_{t+1}^i/\tilde{c}_t^i}{\beta^i} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{\omega + (1-\omega)\pi_t} = \frac{\left(\frac{\tilde{\kappa}_{t+1}^I(\beta^i)/\Delta(I)}{\tilde{\kappa}_t^I(\beta^i)/\Delta(I)} \right) \frac{\tilde{c}_{t+1}}{\tilde{c}_t}}{\beta^i} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{\omega + (1-\omega)\pi_t} \tag{144}$$

where $\tilde{\kappa}_t^I(\beta^i) \equiv \tilde{c}_t^i/\tilde{c}_t$. Note also that using equations (105), (110) and (111) we can write:

$$\frac{\tilde{c}_t^i}{\tilde{c}_t} = \frac{\frac{\beta^i}{1-\alpha-\beta^i(\omega+(1-\omega)\pi_{ss})} \tilde{N}_t^i}{\sum_{j=1}^I \frac{\beta^j}{1-\alpha-\beta^j(\omega+(1-\omega)\pi_{ss})} \tilde{N}_t^j}, \tag{145}$$

where, as before, $\tilde{N}_t^i \equiv \frac{1-\alpha-\beta^i(\pi_{ss}(1-\omega)+\omega)}{1-\alpha-\beta^i(\pi_t(1-\omega)+\omega)} N_t^i$. Finally, this equation can in turn be written as:

$$\frac{\tilde{\kappa}_t^I(\beta^i)}{\Delta(I)} = \frac{\frac{\beta^i}{1-\alpha-\beta^i(\omega+(1-\omega)\pi_{ss})} \frac{\tilde{\nu}_t^I(\beta^i)}{\Delta(I)}}{\sum_{j=1}^I \frac{\beta^j}{1-\alpha-\beta^j(\omega+(1-\omega)\pi_{ss})} \tilde{\nu}_t^I(\beta^j)}. \tag{146}$$

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

$$\tilde{f}_{ct}(\beta) = \frac{\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})} \tilde{f}_t(\beta)}{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}, \tag{147}$$

where $\tilde{f}_t$ and $\tilde{f}_{ct}$ are the continuous probability density function corresponding to the discrete mass functions $\tilde{\nu}_t^I$ and $\tilde{\kappa}_t^I$. Note also that using the corresponding relationship between $\tilde{f}_{t+1}(\beta)$ and $\tilde{f}_t(\beta)$ that is derived in exactly the same fashion as in equation

(128) we can write the following expression:

$$\frac{\tilde{f}_{ct+1}(\beta)}{\tilde{f}_{ct}(\beta)} = \beta \frac{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}{E_{\tilde{f}_t}\left(\frac{\beta^2}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)} \tag{148}$$

Taking the limit of both sides of (144) as $I \to \infty$ and substituting the expression from equation (148) we obtain:

$$R_{t+1} = \frac{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}{E_{\tilde{f}_t}\left(\frac{\beta^2}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{(\omega+(1-\omega)\pi_t)} \frac{\tilde{c}_{t+1}}{\tilde{c}_t} \tag{149}$$

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

$$R_{t+1} = \frac{E_{\tilde{f}_t}\left(\frac{\beta}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)}{E_{\tilde{f}_t}\left(\frac{\beta^2}{1-\alpha-\beta(\omega+(1-\omega)\pi_{ss})}\right)} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{(\omega+(1-\omega)\pi_t)}. \tag{150}$$

If we assume that the discount factors in the adjusted population follow a beta distribution (that is, if we assume that $\tilde{f}_t$ is the PDF of the scaled-beta distribution with shape parameters γ_t and δ_t) then for high enough t we can write the gross interest rate as:

$$R_{t+1} \approx \left(\frac{\gamma_t + \delta_t}{\bar{\beta}(1+\gamma_t)} \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{(\omega+(1-\omega)\pi_t)} \right). \tag{151}$$

A further useful approximation of the above can be made under some additional assumptions. The mean of beta is given by:

$$E_{f,t}(\beta) = \frac{\bar{\beta}\gamma_t}{\gamma_t+\delta_t} \frac{(1-\delta_t)\omega + (\delta_t+\gamma_t)\left(\pi_t \frac{1+\gamma_t}{\gamma_t+\delta_t} - \pi_{ss}\right)(1-\omega)}{(1-\delta_t)\omega + (\delta_t+\gamma_{t-1})\left(\pi_t \frac{1+\gamma_{t-1}}{\gamma_{t-1}+\delta_t} - \pi_{ss}\right)(1-\omega)} \approx \frac{\bar{\beta}\gamma_t}{\gamma_t+\delta_t}, \tag{152}$$

where the final relationship holds exactly if $\omega = 1$ or approximately if either $\pi_t \approx \pi_{ss}$ or if δ_t or γ_t are large enough. Furthermore, since $\frac{1+\gamma_t}{\gamma_t+\delta_t} \approx \frac{\gamma_t}{\gamma_t+\delta_t}$ for high enough δ_t or

γ_t, we can approximate the gross interest rate as:

$$R_{t+1} \approx \frac{g_{Nt} g_{Dt}^{\frac{1}{1-\nu}}}{E_{f,t}(\beta)(\omega + (1-\omega)\pi_t)}. \tag{153}$$

G 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). In our model setting, such 'mutation' is a reduced form way to consider the implications of imperfect transmission of preferences in general. Mutation is one of the fundamental forces of evolution since it helps contribute to the variability of traits within populations. As mutations occur, the process of natural selection determines which of these will thrive and which will die out by selecting the most advantageous mutations for the given environment. In this section we introduce mutation into our model and examine the role it has on the process of natural selection and the economy. Specifically, we allow for the possibility that a proportion of some dynasty exogenously, unexpectedly and permanently experiences a mutation in its discount factor from one period to the next.

We find that some forms of mutation can have dramatic effects whilst others are minor and short-lasting. Furthermore, the effect and impact of a mutation is highly dependent on its type and on its environment, in terms of the contemporaneous composition of agents. Finally we show that mutations to particularly high levels of patience can give rise to long periods of stability, where mutants can dominate the population for many thousands of years. In such a case, the economy can look to have reached a 'steady-state' only for the longer-run evolutionary process to emerge once more.

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.[44] 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 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. We assume mutations occur at only one point in time. Each mutation counterfactual involves an unexpected but permanent change in discount factor for 1% of agents belonging to the dynasty with that period's median discount factor. These mutants then form a new dynasty, retaining their net capital per capita from the previous period.[45] 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.

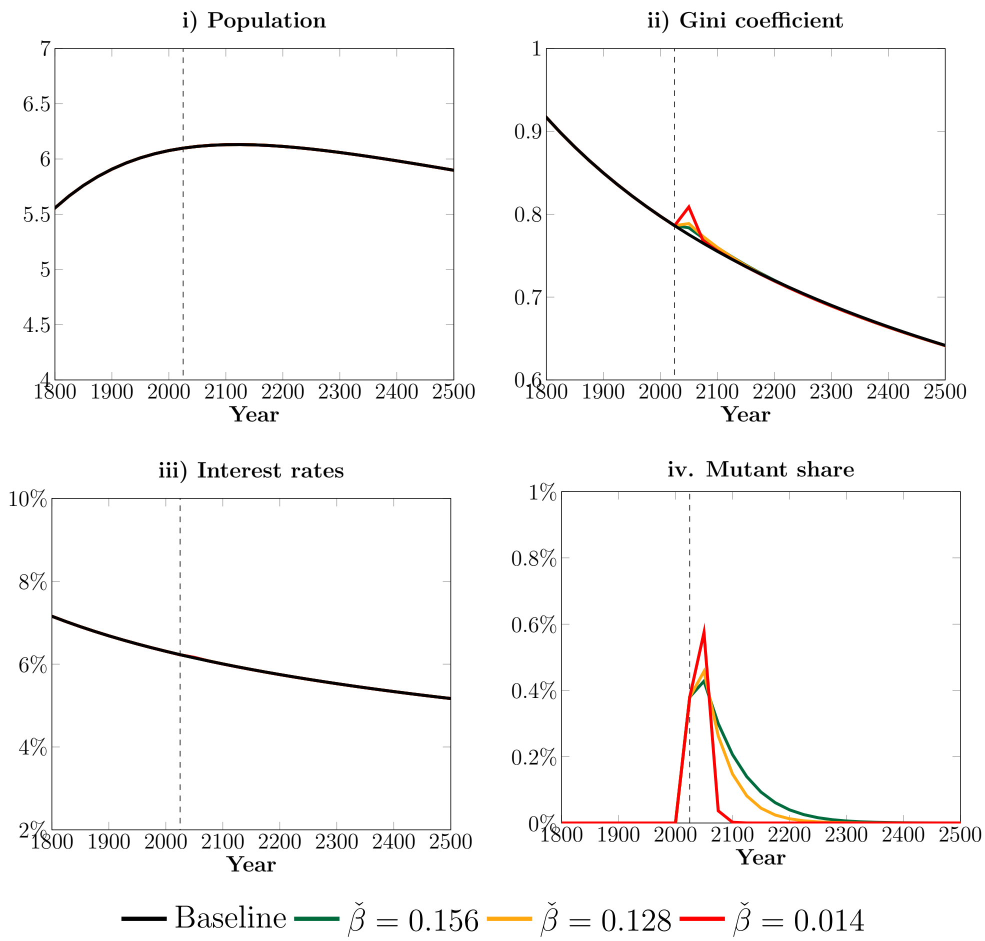
Deleterious mutations First, agents from the median dynasty can mutate to lower levels of patience. In the biological literature these types of mutations are known as 'deleterious' since the mutants have lower fitness than before: agents mutating to a lower level of patience will have fewer children 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 15 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.

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

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

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.

Figure 15: Deleterious mutations in 2025

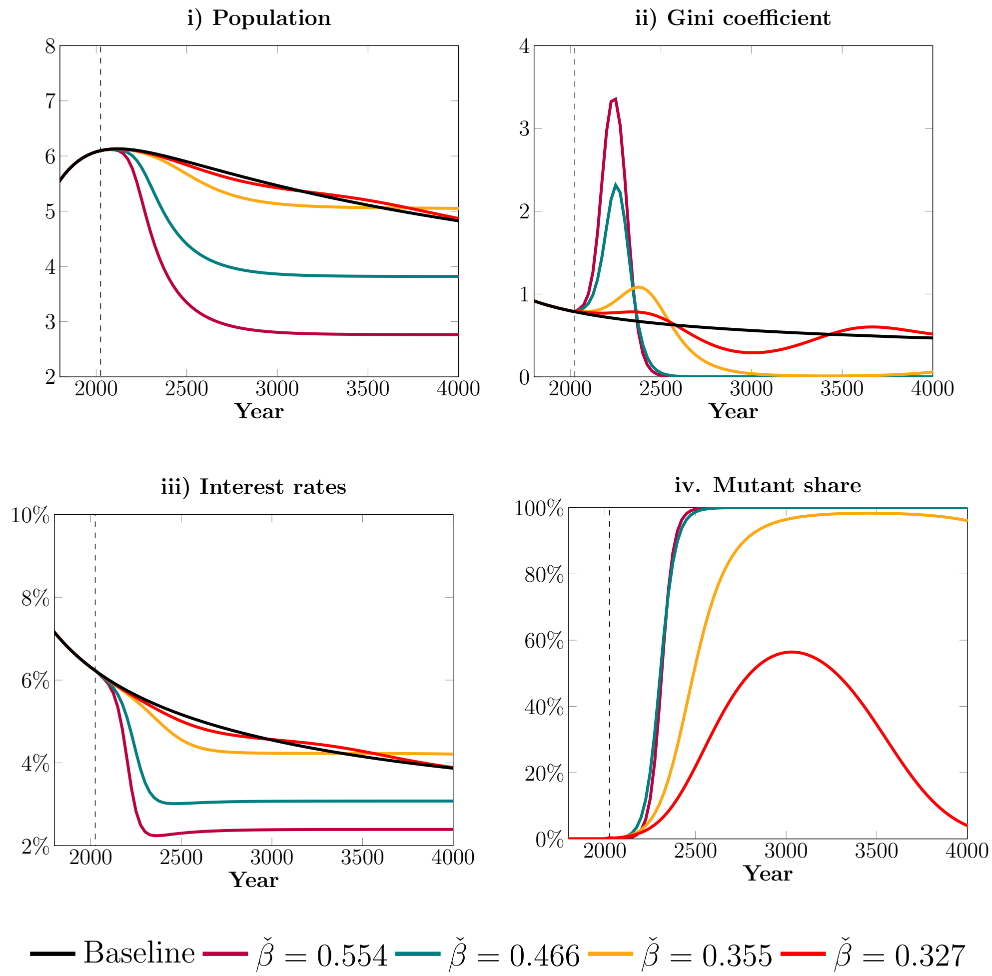


Note: Figures report the simulation output with an unexpected mutation in the year 2025 (dashed line). Each line is a different mutation counterfactual. A mutation causes 1% of the dynasty with the median level ($\beta = 0.213$) of patience in 2025 to wake up in 2025 with the level of patience $\check{\beta}$ denoted in the Figure legend.

Advantageous mutations Second, agents from the 2025 median dynasty can mutate to higher levels of patience. These mutations are knows as 'advantageous' in the biological literature as they increase the fitness of the dynasty: agents mutating to 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 16 shows the effects on population, interest rates and inequality of a mutation to successively higher discount factors as well as 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 16), 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'

Figure 16: Advantageous mutations in 2025



Note: Figures report the simulation output with an unexpected mutation in the year 2025 (dashed line). Each line is a different mutation counterfactual. A mutation causes 1% of the dynasty with the median level ($\beta = 0.213$) of patience in 2025 to wake up in 2025 with the level of patience $\check{\beta}$ denoted in the Figure legend.

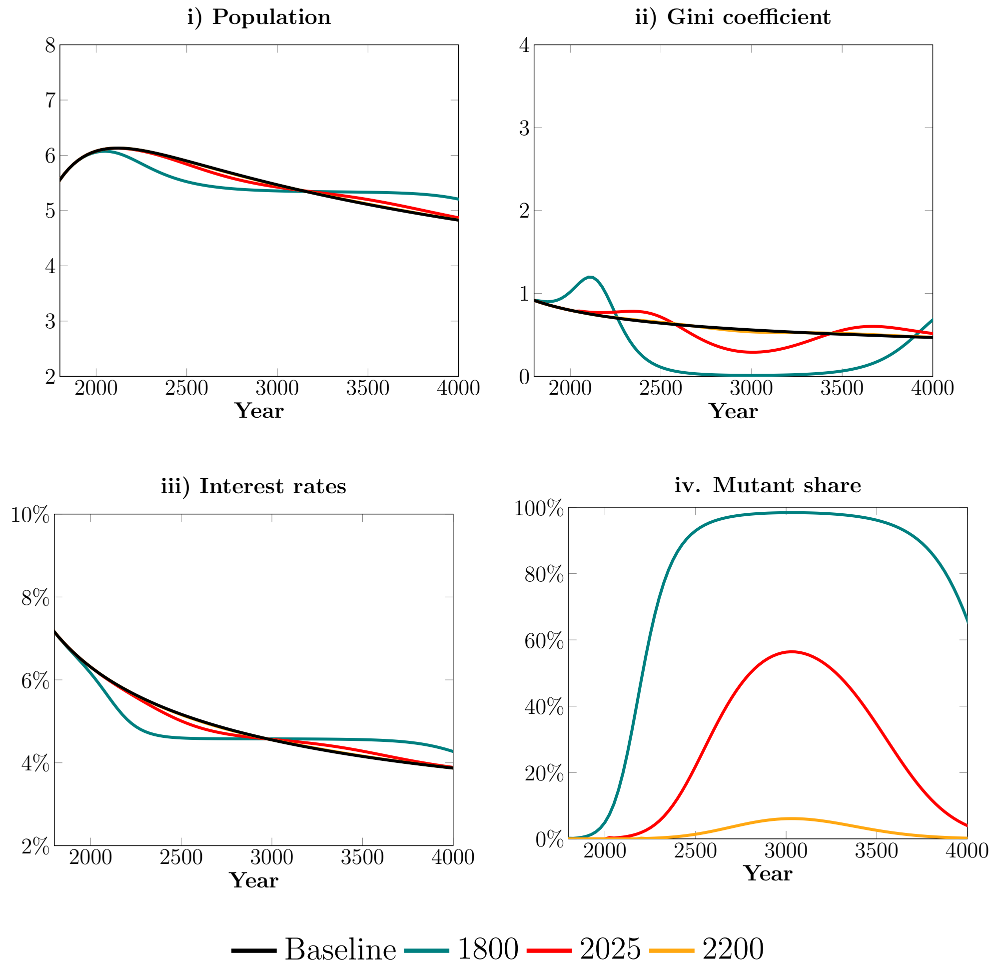
to an intermediate steady-state-like phase and then only slowly shifting to the true steady state where the most patient agent dominates.

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 17 shows the effects of the same mutation occurring in one of three different years. In particular, we consider the same mutation of 1% of 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 structure of the rest of the population and hence on the environment at the time of mutation.[46]

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 years later. For example, as Figure 17 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.

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

Figure 17: 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* $\check{\beta} = 0.327$.