DISCUSSION PAPER SERIES

No. 5106

PATIENCE CAPITAL AND THE DEMISE OF THE ARISTOCRACY

Matthias Doepke and Fabrizio Zilibotti

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP5106.asp**

ISSN 0265-8003

PATIENCE CAPITAL AND THE DEMISE OF THE ARISTOCRACY

Matthias Doepke, UCLA, FRB Minneapolis and CEPR
Fabrizio Zilibotti, IIES Stockholm and CEPR

Discussion Paper No. 5106
June 2005

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 5106

June 2005

ABSTRACT

Patience Capital and the Demise of the Aristocracy*

We model the decision problem of a parent who chooses an occupation and invests in the patience of her children. The two choices complement each other: patient individuals choose occupations with a steep income profile; a steep income profile, in turn, leads to a strong incentive to invest in patience. In equilibrium, society becomes stratified along occupational lines. The most patient people are those in occupations requiring the most education and experience. The theory can account for the demise of the British land-owning aristocracy in the nineteenth century, when rich landowners proved unable to profit from new opportunities arising with industrialization, and were thus surpassed by industrialists rising from the middle classes.

JEL Classification: N23, O14, O15 and Z10
Keywords: British aristocracy, capital accumulation, discount factor, income distribution, Industrial Revolution and patience

Matthias Doepke
Department of Economics
University of California, UCLA
Los Angeles, CA 90095-1477
USA
Tel: (1 310) 794 7278
Fax: (1 801) 469 6100
Email: doepke@econ.ucla.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157169

Fabrizio Zilibotti
Institute International Economic Studies
IIIE, Stockholm University
S-106 91 Stockholm
SWEDEN
Tel: (46 8) 162 225
Email: fabrizio.zilibotti@iies.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=118864

*The authors would like to thank Francesco Caselli, Juan-Carlos Cordoba, Nicola Gennaioli, Maria Saez Marti, Alan Taylor, and seminar participants at the SED Annual Meeting in Florence, the EEA Annual Congress in Madrid, the University of Chicago, the Federal Reserve Bank of Minneapolis, UCLA, USC, Penn State, the Texas Monetary Conference, and Stanford for helpful comments. David Lagakos provided excellent research assistance and Christina Lonnblad provided valuable editorial comments. Financial support by the National Science Foundation (grant SES-0217051), the UCLA Academic Senate, Jan Wallander's and Tom Hedelius' Research Foundation, and the Bank of Sweden Tercentenary Foundation is gratefully acknowledged. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Submitted 02 June 2005

1 Introduction

Humans are born impatient. As parents know well, small children live in an eternal present and are incapable of prefiguring the pleasure that future events can bring. Learning to be future-oriented and to persevere are essential part of our upbringing, and parents spend a substantial amount of time instilling patience into their children. This happens in various forms: deliberate delay of gratification, inducing kids to practice musical instruments and appreciate classical music, religious instruction, and encouragement to work hard in schools are some common examples. Parents' concern for such human assets comes as no surprise, as they turn out to be valuable: empirical evidence shows that individuals who exhibit more patience and perseverance at an early age do better in life.[1]

In this paper, we examine the macroeconomic implications of parental investments in their children's patience. The notion of patience as an asset in which agents can invest, what we term "patience capital," was first introduced in the economic literature by Becker and Mulligan (1997), who consider the problem of a consumer who lives for a finite number of periods and makes a one-time choice of a discount factor. Here, we construct a dynamic dynastic model where the discount factor is treated as a human-capital-like state variable: parents take their own discount factor as given, but can invest in the patience of their children. The focus of the theory is on the interaction of this accumulation process with the choice of occupation and savings. The theory is applied to explain the rise and decline of a class-based society and, more specifically, the decline of the British aristocracy during and after the Industrial Revolution.

The first insight of our analysis is that endogenous accumulation of patience capital can lead to the stratification of societies into "social classes," characterized by different preferences and occupational choices. This occurs even if all individuals are initially identical. The second insight is that such differences in preferences drive the attitudes (or "ethics") displayed by social classes towards investments in physical and human capital. In response to episodes of techno-

[1] See, e.g., Heckman and Rubinstein (2001) and the experimental evidence in Mischel, Shoda, and Rodriguez (1989), discussed below.

logical change (such as the Industrial Revolution), endogenous patience can trigger drastic changes in the income distribution, including the "leapfrogging" of a lower class over the existing wealthy elite.

The key feature of our theory is the association between occupations and consumption profiles. In some professions, lifetime earnings are relatively flat, while in others, especially those requiring the acquisition of skills, high returns are achieved only late in life. These differences affect the incentive of altruistic parents for investing in their children's patience capital: the steeper the consumption profile faced by their children, the stronger the incentive for parents to teach them to be patient. The converse is also true: patient agents have a higher propensity to choose professions entailing steeper earnings and consumption profiles. The dynamic complementarity linking the investment in patience of one generation and the occupational choice of the next leads to the endogenous formation of "social classes." More precisely, dynasties sort into different professions and develop different preferences over time. Financial market development plays a key role: if agents can borrow and perfectly smooth consumption, the link between occupational choice, consumption profile and investment in patience is severed. Thus, class-based societies only emerge when financial markets are shallow, while well-functioning financial markets lead to homogeneous societies.

From a theoretical standpoint, we view investment in patience as a form of human capital investment. In standard human capital theory, agents (dynasties) forego time or utility to increase future enjoyment via higher productivity and consumption. In the case of patience capital, happiness comes through the ability to savor a given stream of future consumption. However, the accumulation of patience capital features important differences. Most notably, if a standard time-separable utility function is assumed, the agents' value functions are convex in patience. In spite of this, we can characterize the problem through a standard recursive formulation with well-defined value functions. The convexity of the value functions turns out to be a surprisingly useful feature to characterize the equilibrium.

We apply our theory to the economic decline of the landed British aristocracy during and after the Industrial Revolution. In the pre-industrial world, wealth, prestige, and political power were associated with the possession of land. Over the nineteenth century the picture changed. A new class of entrepreneurs and businessmen emerged as the new elite. Wealth and political power became progressively detached from land ownership. The new capitalist elite mostly rose from the middle classes: former artisans, merchants, bankers, pre-industrial masters, but also tenant farmers and yeomen. Few aristocrats served as financiers for the new entrepreneurs, and even this became less common as the century progressed. The aristocracy lost ground, first in relative and eventually in absolute terms. At first sight, this decline is quite puzzling, since the aristocracy was the wealthy class and could have been expected to be the main beneficiary of new technological opportunities requiring investments.

We argue that differences in time preference can explain this transformation. The rural aristocracy was too impatient to invest in the new industrial technologies. The pre-industrial middle class, in contrast, had accumulated more patience capital and was culturally better prepared to exploit the new opportunities. These differences, in turn, had their roots in the nature of pre-industrial professions. For centuries, artisans, craftsmen and merchants (the most common activities of the pre-industrial middle class) were used to sacrifice consumption in their youth to acquire skills. In contrast, unskilled laborers, and also landowners, had flat income profiles. Consequently, middle-class parents had the strongest incentive for instilling patience in their children, and the middle class became the patient class. While patience capital was a latent attribute in the pre-industrial world, it became a key asset when new opportunities of enrichment through capital investment arose at the outset of the Industrial Revolution. The rise of the patient bourgeoisie and the demise of the prodigal aristocracy were the consequent outcomes.

In the following section, we relate our work to the existing literature. In Section 3 we analyze an adult's decision problem of choosing an occupation and investing in a child's patience in partial equilibrium. In Section 4 we introduce general equilibrium, and embed the decision problem in a medieval economy populated by landowners, agricultural workers, and artisans. Section 5 introduces capital accumulation. We show that if a new, "capitalist" technology is introduced in the medieval economy, the patient artisans turn into capitalists, while the landowners and workers are left behind. Historical evidence is discussed in Section 6, and Section 7 concludes.

2 Related Literature

A key part of our theory is that patience is important for economic success and can be transmitted from parents to children. Patience can be regarded as a component of a broader set of non-cognitive skills determining how well people can focus on long-term tasks, behave in social interactions, and exert self-restraint. Recent empirical studies emphasize the importance of such human assets for economic success. Heckman and Rubinstein (2001) and Heckman, Hsee, and Rubinstein (2003) use data from the General Educational Development (GED) testing program in the US, and find that non-cognitive skills are responsible for significant differences in wages and education achievements across individuals of equal measured ability (IQ).[2] Similar findings emerge from Segal (2004) using individual measures of non-cognitive abilities at early school age which include proxies for patience. Experimental evidence points at similar conclusions. In a longitudinal study which began in the 1960s at Stanford University, a group of four-year old children were offered a marshmallow, but were told that if they could wait for the experimenter to return after some time, they could have two marshmallows (see Mischel, Shoda, and Rodriguez 1989). Researchers followed the subjects for several years, and found that patient children did significant better in school, marriage and labor market performance.

There is also evidence that non-cognitive skills are affected significantly by nur-

[2]GED is a test that is offered to US high-school dropouts on a voluntary basis. It is devised to test knowledge and academic skills against those of high school graduates. GED recipients can use their test scores to continue education or get better jobs. GED recipients perform on average better than other high-school dropouts: they earn higher wage and attain more education. But this is because they have on average better cognitive skills. If one controls for cognitive abilities (as measured by test scores other than GED), they perform on average worse in both education and professional life than non-recipients. The explanation is that, as documented in the study, that the self-selected population of GED recipients is on average more undisciplined and less future-oriented than that of dropouts who do not take the test.

ture and family upbringing. Heckman (2000) and Heckman and Krueger (2003) review the evidence from a large number of programs targeting disadvantaged children through family development support. They show that most programs were successful in permanently raising the treated children's non-cognitive skills, turning them more motivated to learn, less likely to engage in crime, and altogether more future-oriented than children of non-treated families.[3] These studies show how important the family transmission of this particular form of human capital accumulation – including patience – is. Similar conclusions are reached by a number of studies in child development psychology (see e.g., Goleman 1995, Shonkoff and Philips 2000 and Taylor, McGue, and Iacono 2000). Coleman and Hoffer (1983) argue that the emphasis on patience and self-discipline is the key to the effectiveness of Catholic schools in the US.

If patience is accumulated and transmitted within dynasties, we should expect a positive correlation between parents' and the children's propensity to save and invest. This is consistent with the evidence provided by Knowles and Postlewaite (2004), who report that in the PSID parental savings behavior is an important determinant of education and savings choices of their children's households, after controlling for standard individual characteristics. They interpret their findings as suggestive of large differences in discount factors and of an important role of intergenerational transmission of preferences.

Some recent studies cast light on the socio-economic characteristics of patient agents, suggesting that agents with steeper income profiles are more patient. For instance, a field experiment conducted on Danish households by Harrison, Lau, and Williams (2002) using real monetary rewards shows that highly educated adults have time discount rates (which are inversely related to the discount factor) as low as two thirds as those of less educated agents. This is in line with the key mechanism of our theory that agents who have steeper income profiles have stronger incentives for to invest in patience.[4]

[3] On the other hand, the programs were less successful in raising cognitive skills as measured by IQ test scores. The extent to which parental effort can affect (beyond genetic factors) cognitive skills and social attitudes is more controversial and is the subject of a long-standing debate (see, e.g., Richerson and Boyd (2005) and Bowles and Gintis (2002)). The evidence discussed in the text suggests however that this is less of an issue when one comes to non-cognitive skills.

[4] Other evidence which is consistent with a positive correlation between steep income profiles

A growing literature, both theoretical and empirical, has shown the importance of heterogeneity in preferences, in particular in discount factors, for understanding macroeconomic puzzles in modern economies. Preference heterogeneity has been shown to be necessary to reconcile the quantitative prediction of calibrated macroeconomic models with incomplete markets with the empirical extent of wealth heterogeneity (see Krusell and Smith, Jr. (1998) and De Nardi (2004)).[5] In these models, the heterogeneity of preferences is an exogenous feature. Our theory is complementary to these papers, as it can provide a mechanism through which differences in patience accumulate and persist across agents.

Our paper is related to the growing literature on cultural transmission (e.g., Bisin and Verdier 2000 and 2001, Fernández, Fogli, and Olivetti 2005, Hauk and Saez-Marti 2002, Saez-Marti and Zenou 2004). In this literature, parents evaluate their children's life prospects from the standpoint of their own preferences, and actively try to manipulate children's preference to induce choices that parents regard as desirable. As these papers, we argue that economic incentives are crucial in determining the effort parents exert in affecting their children's preferences. However, in our model, parents exhibit a standard type of altruism as in mainstream dynastic models: parents make no external value judgment on their children's choices. The intertemporal transmission of patience is, like other forms of human capital, a gift that parents pass through to their children. Closer to the tradition of Becker and Mulligan (1997) is the recent paper by Haaparanta and Puhakka (2003), where agents invest in their own patience. In their model, multiple equilibria can arise due to the complementarity between investments in patience and investments in health that prolong the lifetime of individuals.

The importance of cultural and religious aspects in determining which groups

and patience include Carroll and Summers (1991) and Becker and Mulligan (1997). The former document that in both Japan and the United States consumption-age profiles are steeper when economic growth is high. The latter show that consumption grows faster for richer families and adult consumption grows faster for children of the rich (see Table 1). And the Danish study discussed in the text shows that patience increaes with income.

[5] See also Gourinchas and Parker (2002) and Samwick (1998). The relationship of the empirical literature to calibrated macro models is discussed in Browning, Hansen, and Heckman (1999). A different viewpoint is expressed by Ameriks, Caplin, and Leahy (2002) who question whether patience is the key determinant of saving behavior, and argue the key factor to be a psychological attitude which they call "ability to plan". However, they admit that the difference is somewhat hard to identify empirically.

thrived during the Industrial Revolution is at the heart of the celebrated work of Max Weber (1930), who emphasizes how Protestantism, and especially Calvinism, promoted values that were conducive to high savings and wealth accumulation. While we do not focus on religion, our approach echoes the traditional Weberian thesis.[6]

Our theory provides a new perspective of the effects of inequality on development in the face of financial market imperfections. A number of existing theories point out that if financial markets are absent, poor individuals may be unable to finance otherwise profitable investment projects, and are therefore forced to enter less productive professions (see Banerjee and Newman 1993 and Galor and Zeira 1993). Matsuyama (2003) applies similar ideas to the rise and fall of class societies. A common feature of the existing literature is that the rich (who are least constrained by credit market imperfections) generally do best, and should be the first beneficiaries from new investment opportunities. Therefore, these theories cannot account for the fact that the British aristocracy, at a time when wealth inequality was quite extreme and financial markets shallow by modern standards, was rapidly surpassed by middle-class entrepreneurs. In contrast, our theory predicts that under absent financial markets the middle class becomes the patient class, which ultimately results in economic dominance. The two views are complementary in the sense that lack of funds for investment, while not relevant for the middle class, may help explain why the working class was largely excluded from entrepreneurship.

Finally, our paper also relates to a series of recent papers proposing unified theories of the transition from stagnation to growth concentrated on developing joint explanations for the evolution of output and population (see Galor and Weil 2000, Hansen and Prescott 2002, and Doepke 2004). It also relates to a recent literature that emphasizes the role of preference formation for long-run development, but relies on selection instead of conscious investment as the mechanism(see Ga-

[6]In line with the Weberian notion that religious values affect economic behavior, Guiso, Sapienza, and Zingales (2003) use the World Values Surveys to identify the relationship between intensity of religious beliefs and economic attitudes, and find that on average, religion is conducive to higher productivity and growth. Cavalcanti, Parente, and Zhao (2003) question however that differences in preferences arising from religious affiliation can explain large differences in the timing and extent of the Industrial Revolution across countries.

lor and Moav 2002 and Clark and Hamilton 2004). We view the selection and investment approaches to endogenous preference formation as complementary, because they operate on different time scales and lead to distinct implications.

3 Occupational Choice and Time Preference

In this section, we discuss the joint determination of income profiles (through the choice of an occupation) and patience. We first describe the model, and then characterize the solution of a dynamic individual choice problem for a dynasty. Then, in Section 4, we extend the analysis to general equilibrium.

3.1 Preferences, Timing, and Occupations

The model economy is populated by overlapping generations of altruistic agents who live for four periods, two as children and two as adults. Every adult has one child at the beginning of her adulthood. All agents in the economy have the same "basic" preferences. However, a particular aspect of the preferences, namely the time discount factor, is endogenous. In particular, an agent's discount factor is formed during her early childhood, and depends on the time parents decide to spend on increasing the patience of their children.

For simplicity, we assume that agents consume and make economic decisions only when they are adult. Adults work and consume in both adult periods. The amount of time they spend at work is fixed and identical across occupations. The remaining time, which is normalized to unity, can be allocated to either child-rearing l or leisure $1-l$. The motive for child rearing is to increase the patience of the child. Agents' preferences are represented by a time-separable utility function. The period utility (felicity) of an adult agent depends on her consumption and leisure, which are assumed to be multiplicatively separable. More formally, the felicity is given by:

$$w(c,l) = u(c) \cdot h(1-l),$$

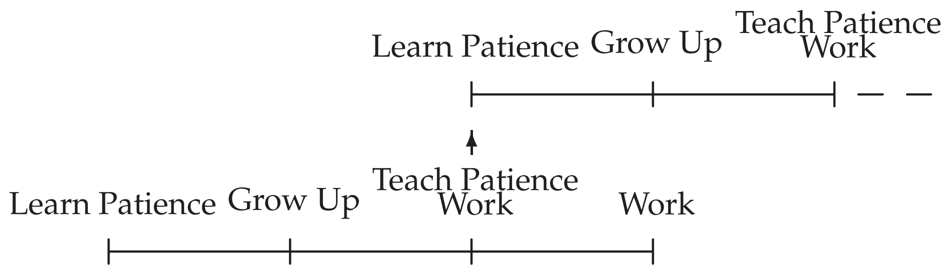


Figure 1: The Timing of Investment in Patience

where $h(1) = 1$, implying that $u(c)$ is the felicity of an agent who does not invest in her child's patience. In addition to their own felicity, adults also care about the utility of their child.

Let $\{c_1, c_2\}$ and $\{l_1, l_2\}$ denote, respectively, the consumption and time invested in patience by an adult in the first and second period of her life. To simplify the analysis, we assume the investment in patience to take place in the first period only, i.e., $l_2 = 0$, as depicted in the time line in Figure 1. This assumption is motivated by the observation that children are most "formative" in their early years, as recently emphasized by Heckman (2000).[7]

The lifetime utility of a young adult endowed with a discount factor given by B can then be represented as follows:

$$u(c_1)h(1-l) + Bu(c_2) + zU(B'(l, B)).$$

Here, z is an altruism parameter which captures the weight of the child in parental utility, $B'(l, B)$ is the "production function" for patience (i.e., the discount factor of the child as a function of the patience and time investment of the parent), and $U(B')$ represents the utility of the child as a function of its discount factor. Notice that discounting within the adult's lifetime is governed by parameter B, while discounting across generations depends on the (exogenous) parameter z. Since parents are altruistic towards their children, the choice problem can be given a

[7]This assumption is not essential, though: our results generalize to a framework where parents invest in their children's patience over two periods, and the formation of patience occurs in both early and late childhood.

"dynastic" interpretation, where the head of the dynasty makes decisions for all subsequent generations. [8]

We assume that $B'(l, B)$ is of the form:

$$B'(l, B) = (1 - \nu)B + f(l), \tag{1}$$

where $\nu \in (0, 1]$ is a constant "depreciation rate" for the time discount factor, and f is a non-negative increasing function. [9] This functional form implies that there exists an upper bound $B_{\max}$ for the discount factor, given by: $B_{\max} = \nu^{-1} f(1)$. We also place the following restrictions on functional forms:[10]

Assumption 1 *The function $u : \mathbf{R}^+ \to \mathbf{R}$ is continuous, differentiable, non-negative, strictly increasing, and weakly concave. The function $h : [0, 1] \to \mathbf{R}$ is continuous, differentiable, non-negative, strictly increasing, strictly concave, and satisfies $h(1) = 1$. The function $f : [0, 1] \to \mathbf{R}^+$ is continuous, differentiable, non-negative, strictly increasing, and weakly concave. The parameters z and ν satisfy $0 < z < 1$ and $0 < \nu < 1$.*

Apart from investment in patience, the second main element of the young adult's decision problem is the choice of an occupation. An occupation i is characterized by an income profile $\{y_{1,i}, y_{2,i}\}$, where we assume $y_{1,i}$ and $y_{2,i}$ to be strictly positive. There is a finite number I of occupations from which to choose. We ignore occupations featuring a dominated income profile, i.e., a profile such that there exists an alternative occupation yielding higher income in one period and at least

[8]It could be argued that investments in patience also affect altruism (hence, we could have B^2 where we have z). Numerical analysis suggests that this formulation would lead to qualitatively similar results, but such change would come at a loss of analytical tractability.

[9]The intergenerational persistence in the discount factor captures the notion that, to some extent, children learn by imitating parental attitudes. Thus, part of the parents' patience is transmitted effortlessly to the child.

[10]The only assumption that may appear to be non-standard is that all felicities are constrained to be positive. Our analysis relies on a cardinal notion of utility. If felicities were negative, it would not be desirable for an altruistic agent to increase the ability of his offspring to savor the future. We believe that this assumption could be relaxed by modeling patience in terms of a relative preference for future vis-a-vis present utility. For instance, lifetime utility could be written as $(1 - \tilde{B})u(c2) + \tilde{B}u(c_2)$, where $\tilde{B}$ is the alternative notion of discounting. In this case, $u(c)$ could be negative.

as high an income in the other period. This is without loss of generality, as no agent would ever choose such an occupation.

Occupations are indexed by consecutive non-negative integers, i.e., $i \in \{1, 2, ..., I\}$, and ordered according to the steepness of the income profile. More formally, we assume:

Assumption 2 *The income profiles satisfy $y_{1,i} > 0$, $y_{2,i} > 0$ for all i. Moreover, a higher index denotes a steeper income profile, i.e., $j > i$ implies:*

$$y_{1,j} < y_{1,i} \quad \text{and} \quad y_{2,j} > y_{2,i}.$$

Adults jointly choose their occupation and their children's patience, so as to maximize utility. We will start our analysis of the adult's choice problem in partial equilibrium, meaning that the income profiles $\{y_{1,i}, y_{2,i}\}$ are taken as given and do not change over time. Later, we will extend the analysis to a general equilibrium economy where the income profiles are endogenously determined.

3.2 Outcomes with Missing Financial Markets

As will become clear below, the development of financial markets plays a key role in our analysis. We start under the assumption that financial markets are absent. In other words, households cannot borrow or lend to smooth out consumption, nor can they leave physical assets to their children. Later, we will contrast the results to outcomes with richer financial markets.

In this environment, consumption is equal to income in each period, $c_1 = y_{1,i}$ and $c_2 = y_{2,i}$, and patience B is the only state variable for a dynasty. The choice problem of a young adult can be represented by the following Bellman equation:

$$v(B) = \max_{i \in I, 0 \leq l \leq 1} \{u(y_{1,i})h(1-l) + Bu(y_{2,i}) + zv(B')\} \tag{2}$$

subject to:

$$B' = (1-\nu)B + f(l). \tag{3}$$

Our decision problem is therefore a dynamic programming problem with a single state variable in the interval $[0, B_{\max}]$, and it can be analyzed using standard techniques. Alternatively, the choice problem can be represented in sequential form by repeatedly substituting for v in (2). While we will mostly work with the recursive formulation, the sequential version is sometimes useful for deriving first-order conditions. The sequential version is written out and shown to be equivalent to the recursive version in the mathematical appendix.

Later on, we will examine the implications of more restricted functional forms for utility (in particular, constant relative risk aversion). Most of our results, however, hold for general functional forms.

Proposition 1 *The value function v is strictly increasing and convex.*

The proof for the proposition is contained in the mathematical appendix.

Intuitively, the convexity of the value function follows from two features of our decision problem: the discount factor enters utility in a linear fashion, and there is a complementarity between the choice of patience and the choice of income profiles. To gain some intuition for the results, consider the decision problem without an occupational choice, that is, with a fixed income profile $\{y_1, y_2\}$. If we vary the discount factor B of the initial generation, while holding constant the investment choices l of all generations, the utility of the initial generation is a linear function of patience B. The reason is that initial utility is a linear function of present and future discount factors, while the initial discount factor, in turn, has a linear effect on future discount factors through the depreciation factor $1-\nu$. The situation is therefore as in the dotted line in Figure 2. If the income profile is constant, it in fact turns out to be optimal to choose a constant l. This is due once again to linearity: the marginal return to investing in patience in a given period is given by $zu(y_2)$, which does not depend on the current level of patience. Thus, if it is optimal in our occupational choice model to hold current and future occupational choices constant over some range of B, the value function is linear over this range.

In general, the optimal income profile is not constant. What turns out to be optimal is to choose a steep income profile (large i) when B is high, and a flat profile

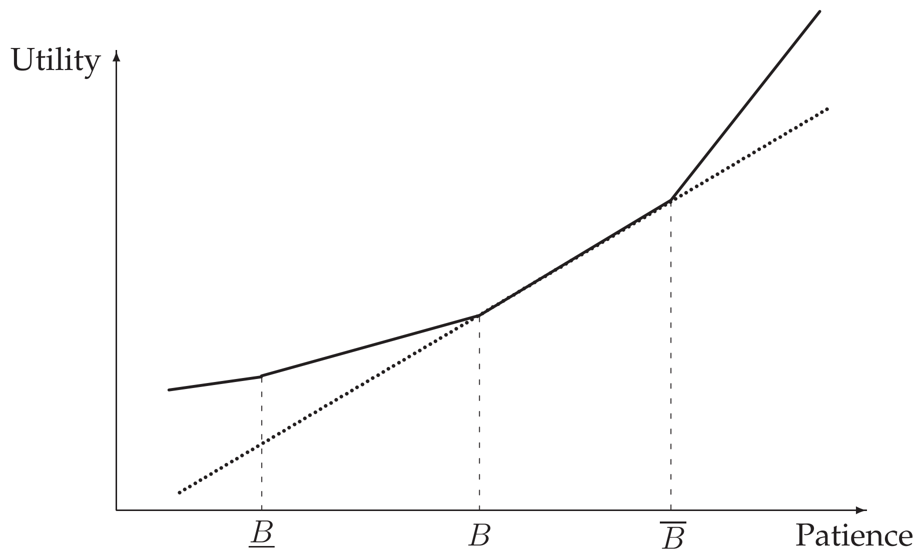


Figure 2: Convexity of the Value Function

when B is low. This is not unexpected, given that a high B implies that more weight is placed on utility late in life. As we increase B, each time a steeper profile is chosen (either in the present or in the future), the value function also becomes steeper in B. The optimal l increases at each step, because the cost of providing patience declines with the steepness of the income profile, while the marginal benefit increases. Since there is only a finite set of profiles, the value function is piecewise linear, where the linear segments correspond to ranges of B for which the optimally chosen present and future income profiles are constant. In Figure 2, the true value function is therefore represented by the solid line, where the points $\underline{B}$, B, and $\overline{B}$ correspond to points where either the current or a future income profile changes. At each of the kinks, some member of the dynasty is indifferent between (at least) two different profiles. Since the choice of l depends on the chosen income profiles, there may be multiple optimal choices l at a B where the value function has a kink, whereas in between kinks the optimal choice of l is unique.

The next propositions summarize our results regarding the optimal choice of income profiles and investment in patience.

Proposition 2 *The solution to the program* (2) *has the following properties: (i) The steepness of the optimal income profile, $y_{2,i}/y_{1,i}$, is non-decreasing in B; (ii) The optimal investment in patience $l = l\,(B)$ is non-decreasing in B.*

Proposition 3 *The state space $[0, B_{\max}]$ can be subdivided into countably many closed intervals $[\underline{B}, \overline{B}]$, such that over the interior of any range $[\underline{B}, \overline{B}]$ the occupational choice of each member of the dynasty (i.e., parent, child, grandchild and so on) is constant and unique (though possibly different across generations), and $l\,(B)$ is constant and single-valued. The value function $v\,(B)$ is piece-wise linear, where each interval $[\underline{B}, \overline{B}]$ corresponds to a linear segment. Each kink in the value function corresponds to a switch to an occupation with a steeper income profile by a present or future member of the dynasty. At a kink, the optimal choices of occupation and l corresponding to both adjoining intervals are optimal (thus, the optimal policy functions are not single-valued at a kink).*

The proposition implies that the optimal policy correspondence $l\,(B)$ is a non-decreasing step function, which takes multiple values only at a step. Propositions 2 and 3 allow us to characterize the equilibrium law of motion for patience. Recall that we assumed that $B' = (1-\nu)\,B + f\,(l)$. Since the policy correspondence $l\,(B)$ is monotone, the dynamics of B are also monotone and converge to a steady state from any initial condition.

Proposition 4 *The law of motion of patience-capital is described by the following difference equation:*

$$B' = (1-\nu)\,B + f\,(l\,(B))\,,$$

where $l\,(B)$ is a non-decreasing step function (as described in Proposition 3). Given an initial condition B_0, the economy converges to a steady state with constant B where parents and children choose the same profession. Multiple steady states are possible.

Notice that while the discount factor of a dynasty always converges, the steady state does not have to be unique, even for a given B_0. For example, if the initial generation is indifferent between two different income profiles, the steady state can depend on which income profile is chosen.

Up to this point, we have not made any use of differentiability assumptions. Given the optimal occupational choices of parents and children, the optimal choice of l must satisfy first-order conditions, which allows us to characterize more sharply the decisions on patience. In particular, we obtain the following first-order condition for l_0:

$$u\left(y_{1,0}\right) h'\left(1-l_0\right) = f'(l_0) \sum_{t=1}^{\infty} z^t (1-\nu)^{t-1} u(y_{2,t}). \tag{4}$$

Here, the left-hand side is the marginal cost of providing patience, and the right-hand side the marginal benefit. Notice that, reflecting our earlier results, the marginal cost is declining in the steepness of the first-generations income profile ($y_{1,0}$ declines when the profile becomes steeper), whereas the marginal benefit increases in the steepness of all subsequent generations' income profiles ($y_{2,t}$ increases in the steepness of the profiles).

Since B_t always converges to a steady state, there must be a time T such that the occupational choice of all members of a dynasty is constant from T onwards. Denoting the constant income profile from this time onwards as $\{y_1, y_2\}$, the steady-state investment in patience $\bar{l}$ must satisfy:

$$u\left(y_1\right) h'\left(1-\bar{l}\right) = f'(\bar{l}) \frac{z}{1-z(1-\nu)} u(y_2) \tag{5}$$

or:

$$\frac{h'\left(1-\bar{l}\right)}{f'(\bar{l})} = \frac{z}{1-z(1-\nu)} \frac{u(y_2)}{u\left(y_1\right)}. \tag{6}$$

Here, the left-hand side is strictly increasing in $\bar{l}$, and the right-hand side is strictly increasing in $u(y_2)/u(y_1)$. The equation therefore pins down $\bar{l}$ as an increasing function of the steepness of the steady-state income profile. The dynamics of B are particularly simple once the occupational choice is constant. Since the law of motion is given by:

$$B_{t+1} = (1-\nu)B_t + f(\bar{l}),$$

then, patience converges to a steady-state $\bar{B}$ given by $\bar{B} = f(\bar{l})/\nu$. Substituting back for $f(\bar{l})$, we can see that patience converges to this steady state at a constant rate:

$$B_{t+1} = (1 - \nu)B_t + \nu\bar{B}.$$

3.3 The Role of Missing Financial Markets

In the preceding analysis, we found that members of different professions face different incentives for investing in patience, provided that the steepness of income profiles differs across professions. A key assumption underlying this result is that access to financial markets to smooth consumption is limited. What determines the the incentive to invest in patience is not the income profile *per se*, but the lifetime profile of period-by-period utilities (felicity). If, however, financial markets are absent, a steep income profile directly translates into a steep utility profile, and thus leads to high incentives to invest.

We now want to make this point more precise by returning to the analysis of Section 3.2, while moving to the opposite extreme in terms of assumptions on financial markets; namely, we allow unrestricted borrowing and lending *within each cohort* at the fixed return R.[11] We will see that in this financial market setup, the choices of patience and occupation no longer interact.

In the environment with borrowing and lending, the Bellman equation describing the young adult's decision problem is given by:

$$v(B) = \max_{i \in I, 0 \leq l \leq 1, s} \left\{ u(y_{1,i} - s)h(1 - l) + Bu(y_{2,i} + Rs) + zv(B') \right\}, \tag{7}$$

subject to:

$$B' = (1 - \nu)B + f(l). \tag{8}$$

The next proposition establishes that the introduction of a perfect market for borrowing and lending removes any link between patience and occupational choice.

Proposition 5 *The value function* v *defined in* (7) *is increasing and convex. The only income profiles that are chosen in equilibrium are those that maximize the present value*

[11] The possibility of wealth transmission *across generations* is discussed in Section 5.1

of income, $y_{1,i} + y_{2,i}/R$. The set of optimal income profiles is independent of patience B. The choice of occupation does not affect the investment in patience.

The intuition for this result is simple: with perfect borrowing and lending, every adult will choose the income profile that yields the highest present value of income, regardless of patience. The proposition shows that at least some degree of financial market imperfection is necessary for occupational choice and investments in patience to be interlinked. It is not necessary, however, to assume the entire absence of financial markets, as we did in the preceding section for analytical convenience. As long as the steepness of an income profile is at least partially transmitted to consumption profiles, the basic mechanism is at work.

A positive implication of this finding is that the degree of discount-factor heterogeneity in a population depends on the development of financial markets. In an economy where financial markets are mostly absent, incentives to invest in patience vary widely across members of different professions, and consequently we would expect to observe a large corresponding variation in actual acquired preferences. In modern times with richer financial markets, these differences should be smaller. For example, while engaging in a lengthy program of study (such as medical school) which leads to high future incomes may still require a certain degree of patience and perseverence, today's students have access to educational loans and credit cards. Hence, the modern-day artisans are able to consume some of their future rewards already in the present, and consequently they and their parents face a smaller incentive to invest in specialized preferences.

4 General Equilibrium with Two Technologies

The results up to this point demonstrate that there exists a basic complementarity between the acts of investing in patience and choosing a profession. Dynasties starting out patient choose professions that are characterized by a steep income profile which, in turn, increases further the incentive to invest in patience. The self-reinforcing nature of the two aspects of our decision problem suggests the possibility that different dynasties may diverge and end up in different steady states. However, multiple steady states are not a necessary feature of our model. What the preceding section does show is that if different dynasties choose separate professions with different income profiles, they will also end up with different levels of patience. Whether different dynasties choose different professions depends both on the level and the steepness of possible income profiles. Therefore, if we wish to determine whether dynasties diverge or converge, we must move beyond partial equilibrium and endogenize the incomes derived in different professions.

The point of this section is to show that general equilibrium forces can adjust the returns to working in different professions such that at least some agents find it optimal to work in each profession. Given the different choices of profession in the population, divergence in patience then necessarily follows. Outcomes of this type naturally occur if the reward to being in a profession is a decreasing function of the number of members of the profession, i.e., if there are decreasing returns. While this result could be derived in many different economic environments, we establish this point within a specific environment geared towards our application to the demise of the aristocracy.

4.1 Analytical Results

We assume that there are three occupations, with occupational mobility across only two of them. We parameterize preferences over consumption by a utility function featuring constant relative risk aversion (CRRA), i.e., $u(c) = c^{\sigma}$, where $0 < \sigma \leq 1$.[12] Finally, we concentrate on equilibria starting from an initial condition where patience is identical across agents. Apart from simplifying the analysis, this focus is coherent with our aim of showing that preference stratification necessarily arises through the process of sorting the population into different occupations, even if everybody is initially identical.

We name occupations and technologies in a way hinting at the application that will be discussed later in the paper. The two modes of production are called

[12] Allowing case $\sigma \leq 0$ would violate our assumption of the period utility function being nonnegative. While the results can in principle extended to richer utility functions, we focus on the case covered by Assumption 1.

agriculture and artisanry. For simplicity, we assume agricultural output, Y_A, and the production of artisans, Y_M, to be perfect substitutes, $Y = Y_A + Y_M$. The two technologies differ in terms of the inputs used. The agricultural technology uses unskilled labor, L , and land, Z, and is described by the following production function:

$$Y_A = Z^{1-\alpha} L^{\alpha}, \tag{9}$$

where $\alpha \in (0, 1)$. The artisan technology is linear in skilled labor:

$$Y_M = qH. \tag{10}$$

The total amount of land is fixed at $Z = 1$. Land is not traded and is owned by a fixed number of dynasties, where each landowner bequeaths the land he owns to his child when he passes away. Land is only productive if the owner monitors production; therefore, landowners do not supply skilled or unskilled labor alongside using their land. Thus, landowning is just another profession characterized by a lifetime profile of rental income. There is, however, no occupational mobility between landowners and the other classes. Since the supply of land is fixed, the decisions of landowners (on investing in patience) have no general-equilibrium implications. We will therefore concentrate on the "lower classes" for now.

The main difference between skilled and unskilled labor is the lifetime labor supply profile. An unskilled worker supplies one unit of agricultural labor in each adult period. For skilled workers, in contrast, the first adult period is partially used for acquiring skills and experience. Effective labor supply is therefore one unit in the first adult period, and $\gamma > 1$ units in the second adult period. In every period the mass of labor-market participants is equal to one (the total mass of landless agents is two, but only half of them are adults). Labor markets are assumed to be competitive.

Define, next, an equilibrium with constant wages across dynasties. We focus on constant-wage equilibria because in this case the analysis of the preceding section (which was for a decision problem with a fixed set of occupational income profiles) directly applies to the decision problem of agents in our general-equilibrium economy. Since the marginal product of each type of labor is a function of labor supply, a constant-wage equilibrium is characterized by a constant number of each type of worker over time.

Definition 1 *An Equilibrium with Constant Wages (ECW) is a time invariant distribution of wages per effective unit of labor and a time-invariant distribution of landless adults between the two occupational choices, such that (a) all working members of the landless dynasties optimally choose their occupation, (b) all parents optimally choose the investment in patience, and (c) all markets clear.*

In an ECW, the income profile of agricultural workers is flat. Landowners receive the same amount of rent every period, and therefore have a flat income profile, just like the workers. In contrast, artisans have an increasing earning profile and, hence, they have a stronger incentive to invest in patience. Given the CRRA preference specification, only the steepness, but not the level of income matters for the investments in patience. The following proposition follows from the definition of ECW and from the analysis of Section 3.

Proposition 6 *An ECW is characterized by occupational segregation, i.e., parents and their children choose the same profession. Under CRRA preferences, the distribution of discount factors converges to a steady state where all worker and landowner dynasties have a discount factor $\bar{B}_A$, whereas artisans have a discount factor $\bar{B}_M > \bar{B}_A$.*

The proposition establishes that if an ECW exists and the number of workers and artisans is strictly positive, we indeed observe diverging patience in the population, with each group converging to a profession-specific discount factor. We still need to establish whether such an ECW actually exists. Since an ECW is a particular type of equilibrium, its existence depends on the initial conditions. We will now show that a unique ECW exists if all dynasties start out with the same initial patience $\tilde{B}$, provided that $\bar{B}_A \leq \tilde{B} \leq \bar{B}_M$. This encompasses the case of an economy where, before time zero, only the agricultural activity was pursued (e.g., because the productivity of artisanry q was very low), and investment in patience had settled down to the level $\tilde{B} = \bar{B}_A$. Then, at time zero, q unexpectedly increases, and occupational sorting starts.

In an ECW, employment and productivity per efficiency unit of labor is constant in each sector. In particular, if we denote by $\mu \in [0, 1]$ the proportion of landless adults employed in agriculture, the competitive wages per efficiency units of labor in artisanry and agriculture are $w_M = q$ and $w_A = \alpha\mu^{\alpha-1}$. Thus, an artisan earns, respectively, q and γq in the first and second period of his life, whereas an unskilled worker earns a flat wage of $\alpha\mu^{\alpha-1}$. Notice that the definition of ECW does not require the age distribution of the adults employed in each profession to be time invariant. For instance, an ECW is consistent with a larger number of young adults choosing artisanry in even than in odd periods (or vice versa), as long as the *total* number of workers engaged in each occupation is time invariant.[13] Establishing the existence of an ECW now amounts to showing that there exists a $\mu \in [0, 1]$, such that all conditions of Definition 1 are satisfied. The following proposition summarizes the result.

Proposition 7 *Suppose that the economy starts out with everyone having the same discount factor $\tilde{B}$, where $\bar{B}_A \leq \tilde{B} \leq \bar{B}_M$. Then, there exists a unique ECW such that:*

- *either $\mu = 1$, $w_A = \alpha$, and all landless adults in all periods weakly prefer to work in agriculture,*
- *or $\mu < 1$, $w_A = \alpha\mu^{\alpha-1}$, $w_M = q$, and μ is such that the initial generation of adults is indifferent between agricultural labor and artisanry, and all children weakly prefer their parents' profession.*

Which of the two possible outcomes is obtained is a function of the productivity of artisanry q. If this productivity is sufficiently high, there will be a positive number of artisans in equilibrium, and preferences will diverge across professions.

4.2 A Medieval Economy

In this section, we illustrate the general-equilibrium results with outcomes in a parameterized version of our economy. The economy is populated by measure

[13] This implies fluctuations in the aggregate manufacturing output, whereas agricultural production remains constant, since young and old adults are equally productive.

one of landless adults, who are either agricultural workers or artisans, and measure a of landowners, each of whom owns an equal share of one unit of land. The functional form for the accumulation of patience is given by:

$$f(l) = \phi\left(1 - (1-l)^{\xi}\right),$$

where we require $\phi > 0$ and $\xi > 1$ to meet the restrictions in Assumption 1. The functional form was chosen because it implies that the marginal productivity of investing in patience converges to zero as the time investment l approaches one. While this property is not required for any of our results, it is useful to ensure that the solution for l is interior. The production technologies are given by (9) and (10). The period utility functions are $u(c) = c^{\sigma}$ for consumption (as mentioned before) and:

$$h(1-l) = (1-l)^{\eta}$$

for leisure, where we require $0 < \eta < 1$ to satisfy Assumption 1. Table 1 summarizes our choices for all parameter values. The number of landowners a is left unspecified, because it only sets the income level that each landowner receives, without any effect on other outcomes.

σ	η	z	γ	q	α	ν	ϕ	ξ
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{4}{3}$

Table 1: Parameter Values for Medieval Economy

In a constant-wage equilibrium, agricultural workers and landowners face a flat income profile, while given our choice of $\gamma = 2$ artisans have twice the income when old compared to when young. We can use the first-order condition (6) to compute the steady-state patience for each profession. Given our functional form assumptions, the condition is given by:

$$\frac{\eta}{\phi\xi(1-\bar{l})^{\xi-\eta}} = \frac{z}{1 - z(1-\nu)}\left(\frac{y_2}{y_1}\right)^{\sigma}.$$

Solving this equation for $\bar{l}$ and plugging in all parameter values, we obtain solutions of $\bar{l}_W = 0.18$ for agricultural workers and landowners and $\bar{l}_A = 0.46$ for artisans. In steady state, patience is given by $\bar{B} = f(\bar{l})/\nu$, so that these investments translate into long-run discount factors of $\bar{B}_W = 0.32$ and $\bar{B}_A = 0.75$. If we interpret the length of a period to be ten years, these numbers correspond to annual discount rates of 0.89 and 0.97, respectively.

We now proceed to compute a constant-wage equilibrium from an initial condition where everybody is equally patient. In particular, we assume all dynasties to start out with patience $B = 0.5$, right in the middle between the two steady states for workers and artisans. Such an initial condition could be justified if initially both agricultural and artisan tasks were carried out by each dynasty, resulting in an income profile of intermediate steepness. The initial condition captures the transition of such an economy from a point where a strict division of labor is introduced. Proposition 7 guarantees that a unique constant-wage equilibrium exists. In the equilibrium, about 55 percent of the landless adults are agricultural workers. The income of an artisan is $q = 0.5$ in the first period and $\gamma q = 1$ in the second period, while an agricultural worker receives a wage of $w_W = 0.67$ in each period. Notice that workers have a lower average income than artisans; they still prefer to be workers because they value the flat income profile.

Figure 3 shows the value function (top panel) and the law of motion for patience (bottom panel) for members of the landless class in our economy. As shown in Section 3.2, the value function is piecewise-linear and convex, with the kink at $B = 0.5$ corresponding to the threshold above which adults choose to be artisans, and below which they become agricultural workers. That the kink is at the initial patience of $B = 0.5$ is, of course, no accident: in the initial period, the number of workers and artisans and, therefore, wages adjust such that each member of the initial generation is just indifferent between being a worker and an artisan. The law of motion for patience jumps at the threshold of $B = 0.5$, which is in line with Proposition 3 . In equilibrium, there is persistence in the occupational choice: the children of first-generation artisans become artisans, while the children of workers become workers. Notice that from the second generation forward, patience diverges from the threshold $B = 0.5$. Hence, only the initial generation is indifferent between being a worker or an artisan; the first generation's children strictly prefer their parent's occupation to the alternative. As shown in Proposition 4,

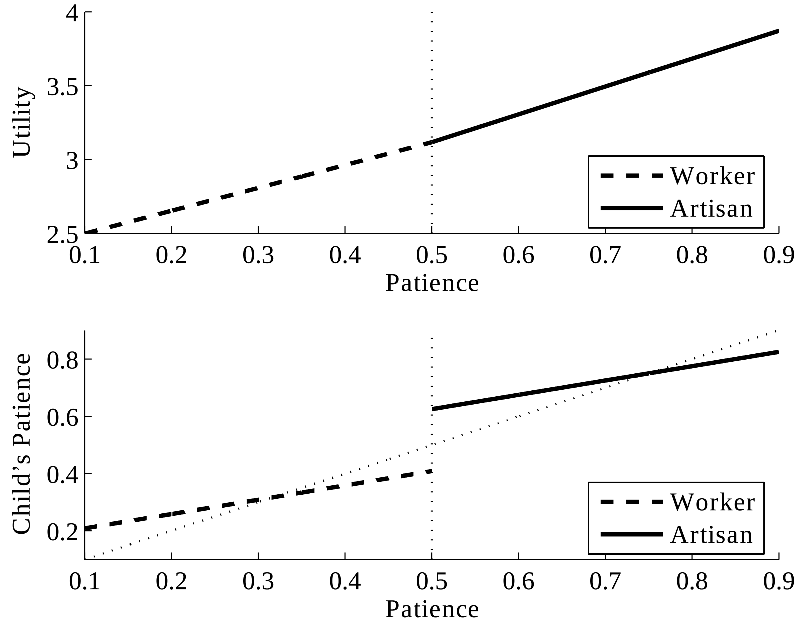


Figure 3: Value Function and Law of Motion for Patience for Lower Classes

the law of motion is linear subject to being in a dynasty with a given occupation. Consequently, patience approaches its steady state of B_W or B_A at a constant rate, as displayed in Figure 4.

The landowners face the same incentives for investing in patience as the agricultural workers. Since they do not face an occupational choice, their law of motion for patience (not shown) is linear. In particular, it is identical to the workers' law of motion up to the threshold $B = 0.5$, and is given by the linear extension of the worker's law of motion above the threshold. Over time, the landowners' patience evolves just like the workers' patience in Figure 4.

Thus, in our two-technology medieval economy preferences diverge across professions from the second generation onward, despite the fact that initially everyone has the same preferences. We end up with a society stratified along occupational lines. As our preceding analysis shows, this stratification is a quite general outcome, provided that income profiles differ across occupations, and financial markets do not allow perfect consumption smoothing.

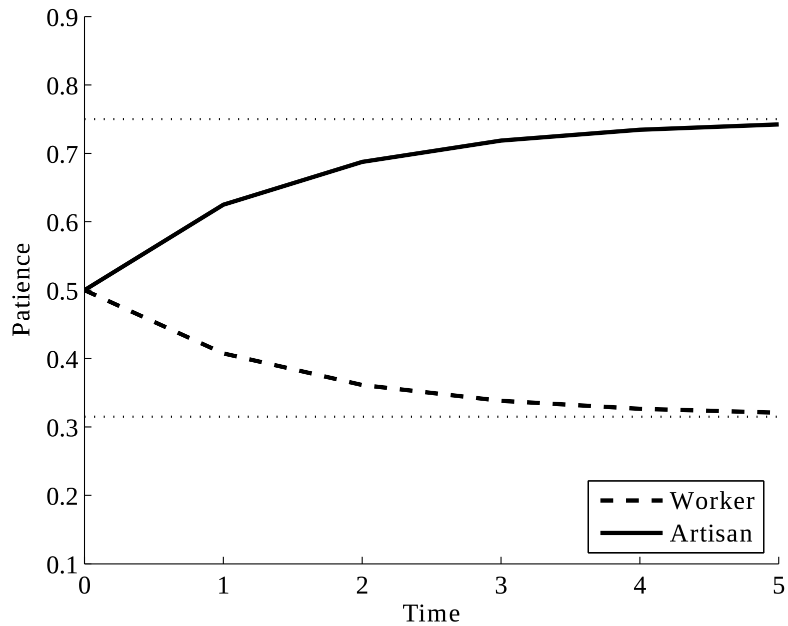


Figure 4: Patience Over Time

In the medieval economy, class differences are only important to the extent that they determine the professional choice of individuals. Patience becomes of central importance, however, when technological change gives rise to new investment opportunities (the "Industrial Revolution"). In the following section, we examine the fate of the different classes in our economy after the arrival of a new "capitalist" technology.

5 Capitalism and the Decline of the Aristocracy

Much of modern macroeconomics is built around a basic workhorse model: the neoclassical growth model with optimizing consumers. In this model, agents face a single intertemporal decision, namely, an investment choice. It is well-known that in this type of model small differences in time preferences across agents have a large impact on wealth accumulation. In particular, if preferences are CRRA and there are different groups which are only distinguished by their time preference, the most patient group ultimately owns all the wealth.

In the theory built in the preceding sections, societies become stratified along occupational lines, where the different classes are distinguished by their time preference. Up to this point, however, there was no possibility of capital accumulation. In this section, we explore what happens if an investment opportunity exists. In particular, we focus on a version of the "Industrial Revolution", in which the investment opportunity is unexpectedly introduced in a medieval society where the classes have already acquired different levels of patience. The result will be unsurprising in the light of standard macroeconomic theory: the patient classes, i.e., the artisan middle classes, are in the best position to take advantage of the new opportunity since they possess the "spirit of capitalism." The artisans therefore leapfrog over the declining landed elite, and achieve economic dominance.

5.1 The General Model with Wealth Accumulation

Consider the choice problem of a dynasty that can accumulate two assets: physical capital and patience. The rate of return to capital is constant and equal to A. Capital is assumed to depreciate at the rate δ. Young adults inherit capital from their parents, and decide how much of their first-period income they consume and how much they invest. Investments in physical capital are assumed to be *irreversible*: agents can consume the return to physical capital (as well as their labor income), but the capital stock cannot be liquidated and turned into consumption. Thus, we interpret our model of capital accumulation as investment in a family-run entrepreneurial activity. The capital owned by an old agent is bequeathed—up to depreciation—to his child.[14] We continue to assume that, because of capital market imperfections, agents cannot borrow.

[14] Dynastic enterprises were very common in the early days of the Industrial Revolution. Caselli and Gennaioli (2003) discuss this phenomenon, and argue that it is due to the underdevelopment of financial market that makes it unprofitable for parents to liquidate their business instead of leaving it to the children. In our model, the irreversibility constraint implies that differences in investments across families cause differences in initial assets for the next generation. Similar results could be obtained under reversible investment if the altruism parameter z (the intergenerational discount factor) were an increasing function of patience B (the intragenerational discount factor).

More formally, let $K \geq 0$ denote the bequest of capital received by a young adult. The sequence of budget constraints facing this agent is given by:

$$c_1 + K' = (1 - \delta + A)K + y_{1i}, \tag{11}$$

$$c_2 = AK' + y_{2i}, \tag{12}$$

subject to

$$K' \geq (1 - \delta)\, K. \tag{13}$$

Consider the budget constraint in the first adult period, (11). The total income consists of labor income, y_{1i}, plus capital income, $(1 - \delta + A)K$. The latter, in turn, consists of the stock of capital remaining after depreciation ($(1 - \delta)K$), and the return to the productive activity (AK). However, because of the irreversibility constraint, (13), he cannot liquidate and consume the stock (namely, $c_1 \leq AK + y_{1i}$). In the second period, (12), the agent earns the labor income y_{1i} and the return to capital, AK'. Since the capital stock cannot be liquidated, the agent bequeaths the capital stock, $K'' = (1 - \delta)K'$, to his child. [15]

We maintain our earlier assumption that utility is CRRA in consumption. The recursive representation of the problem of a young adult endowed with the discount factor B and the capital stock K is therefore given by the following Bellman equation:

$$v(B, K) = \max_{i \in I, l, K'} \left\{ c_1^{\sigma}\ h(1 - l) + B\ c_2^{\sigma} + z\ v(B', (1 - \delta)K') \right\} \tag{14}$$

subject to (3), (11), (12), (13), $c_1 \geq 0$ and $0 \leq l \leq 1$.

Unfortunately, it is impossible to provide an analytical characterization of this program. We will then proceed in two steps. First, we study the case in which there is no occupational choice and no labor income. In this case, some insightful

[15] Note that, in principle, parents could bequeath additional resources to their offspring. However, we focus on economies where the irreversibility of the capital stock is a binding constraint for the old adults. Namely, in the last period of their lives agents would like to liquidate part of the capital stock and consume it, but they are instead forced to leave it to their children as an involuntary bequest. Clearly, in such economies, agents do not leave any additional bequests. Formally, this outcome can be guaranteed to be optimal by choosing the altruism factor z appropriately.

characterization can be attained. Then, we move to the medieval economy of section 4.2 where agents choose between two occupations. This case, which can only be analyzed numerically, provides an interpretation to the decline of the aristocracy and the triumph of the bourgeouisie during the industrial revolution.

5.2 Analytical Results

Suppose that $y_{1i} = y_{2i} = 0$. The adult's decision problem of investing in patience and physical capital, (14), simplifies to:

$$v(B,K) = \max_{l,K'} \left\{ ((1-\delta+A)K - K')^{\sigma} \; h(1-l) + B \; (AK')^{\sigma} + z \, v(B', (1-\delta)K') \right\} \quad (15)$$

subject to (3), (13) and $K' \leq (1-\delta+A)K$, and $0 \leq l \leq 1$. The following proposition sums up the basic characterization of this decision problem.

Proposition 8 *Suppose that returns are sufficiently low to guarantee bounded utility, i.e.,* $z[(1-\delta)(1-\delta+A)]^{\sigma} < 1$. *Then the functional equation defined by* (15) *has a unique fixed point* v, *which is a solution of the corresponding sequence problem. Let* $k \equiv K'/K$. *The value function can be expressed as* $v(B,K) = K^{\sigma} V(B)$, *where*

$$V(B) = \max_{\substack{l \in [0,1], \\ k \in [1-\delta, 1-\delta+A]}} \left\{ ((1-\delta+A) - k)^{\sigma} \; h(1-l) + B \; (Ak)^{\sigma} + z \, k^{\sigma} V((1-\nu)B + f(l)) \right\} \quad (16)$$

is increasing and convex in B. *The optimal policy functions* $l(B)$ *and* $k(B)$ *are continuous and non-decreasing.*

This proposition has the important consequence that from any initial condition, patience B and the growth rate of capital K'/K necessarily converge to a steady state. Moreover, both variables depend on patience B only. The existing capital stock K only enters as a scale parameter, due to the fact that utility is CRRA in consumption and the investment technology is linear. We also find a version of the complementarity result of the occupational choice model: a high level of patience leads to higher investment, and therefore a steeper lifetime utility profile.

Both unique and multiple steady states are possible, depending on parameters. The force favoring multiple steady states is the self-reinforcing mechanism which links savings and investment in patience. A patient person (high B) tends to choose a high investment rate. High investment, in turn, induces a steep consumption profile, and this in turn provides an incentive for altruistic parents to invest in patience.[16] The forces opposed to multiple steady states are the desire to smooth consumption and decreasing returns to investing in patience. When an agent has a strong preference for a smooth consumption profile, the complementarity between savings and patience can be overcome, and the standard forces leading to a unique steady state prevail.

Note, finally, that there are important similarities between the choice of investing in a family firm and our original choice problem of choosing an occupation. In our framework, choosing an occupation amounts to choosing a lifetime income profile. Choosing a level of investment achieves the same purpose, by trading off a reduction in present consumption with a future gain. As we have seen, the same complementarities that characterized the joint decision problem on occupation and patience reappear in the analysis of this section. Nevertheless, there is an important difference. The choice of an occupation has no direct effect on the children (other than the indirect one through patience), whereas investing increases the beginning-of-life assets of the offspring. Differences previously only manifesting themselves in occupational choices and preferences have now an effect on the *growth rate* of assets within dynasties.

[16] With capital accumulation, not only the lifetime income profile of a given cohort, but also the income profile *across generations* is endogenous. If investment is high, the next generation starts out with more capital then the current generation. This shifts future utilities upward relative to current utility. Since the cost of investing in patience is in terms of current utility, while the benefit is proportional to future utilities, this effect increases the complementarity between savings and patience.

5.3 The Decline of the Aristocracy

In this section, we introduce the capitalist technology into the parameterized medieval economy that was discussed in Section 4.2, thereby combining occupational choice and investment into a common framework. While general analytical results cannot be found for this combined model, we provide numerical results to show that its implications are just what the preceding analysis would lead us to expect. In particular, when we start the economy in the medieval steady state where the society is stratified in terms of preferences, we find that only the artisans make use of the new technological opportunity. The result is the economic triumph of the most patient social class, i.e., the urban bourgeoisie, over the original landed elite.

The recursive representation of the problem of a young adult endowed with the discount factor B and capital stock K, and choosing between an agricultural and an artisan technology, is given by:

$$v(B,K) = \max_{i\in\{A,M\},l,K'} \big\{ (y_{1i} + (1-\delta+A)K - K')^{\sigma} \ (1-l)^{\eta} \\ + B \ (y_{2i} + AK')^{\sigma} + z\, v(B',(1-\delta)K')\big\}, \quad (17)$$

subject to (3), $(1-\delta)K \leq K' \leq y_{1i} + (1-\delta+A)K$, and $0 \leq l \leq 1$. The agricultural and artisan technology continue to exist alongside the capitalist technology, and are specified as in Section 4.2. Notice that adults can, and will, continue to work in one of the existing professions even while they are investing in the new technology. This assumption is for realism; in particular, we want to allow aristocrats to earn rents from their land and invest the proceeds in a capital market, so as to not exclude them from investment from the outset. The landowners face the same decision problem as members of the lower classes, with the exception of the occupational choice component: landowners derive income y_1 and y_2 from renting out their land, and do not choose any of the other two professions.

The economy is parameterized as in Section 4.2 (see Table 1, and each class starts out with its steady-state discount factor (B_W for workers and landowners and B_A for artisans). Even though the additional income source implies that the choice

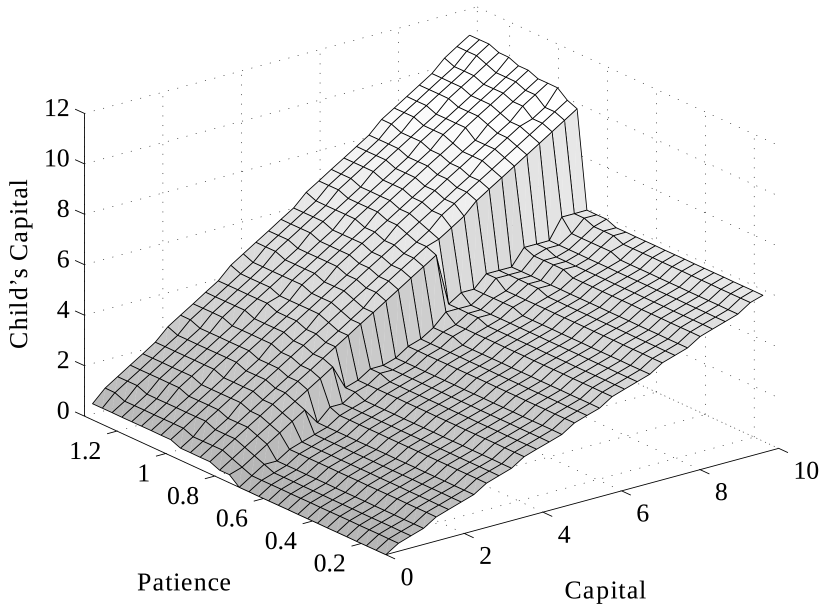


Figure 5: The Policy Function for Capital Investment

problem described by equation (17) is not equivalent to the capital-only model analyzed in Section 5.1, the implications for the interaction between capital accumulation and patience are quite similar. Figure 5 shows (for the landless classes) the policy function for the decision on investment (the child's capital on the vertical axis) as a function of the adult's patience and inherited capital stock. The assumed depreciation rate is $\delta = 0.2$, and the return on the investment technology is $A = 0.6$. The conspicuous feature of Figure 5 is a "cliff." If we hold the inherited capital stock K constant and increase patience B, the irreversibility constraint initially binds for the young adult, so that they do not invest and only carry forward the depreciated capital to the next period. Once a critical level of patience has been reached, however, the invested capital stock increases quite rapidly, until a plateau is reached where savings increase slowly with patience. The reason for this behavior of the policy function is the complementarity between savings and patience. The low plateau for low patience corresponds to savings decisions implying that wealth declines from the current generation to the next. Consequently, current utility is high relative to future utility. The high

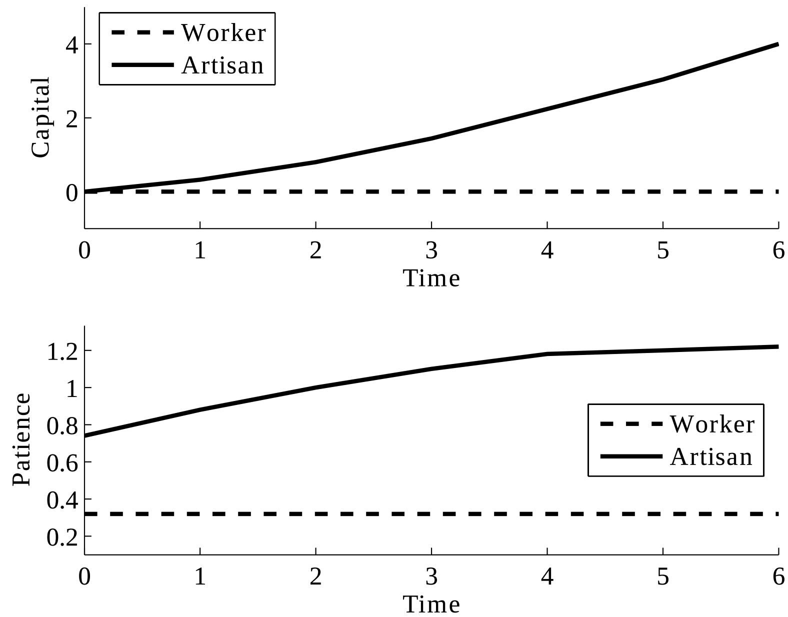


Figure 6: Patience and Capital Over Time

plateau depicts savings choices that imply an increasing profile of wealth from the current to the next generation. Parents therefore enjoy a lower level of utility than their children, which makes investing in the children's patience highly attractive. The face of the cliff is precisely the region where the children's wealth rises above the parents' wealth, which amplifies the effect of rising patience in investment in patience.

Apart from investment in physical capital, our original decisions of choosing an occupation and investing in patience are still present in the model. As previously, for any level of capital there is a critical level of patience B above which an agent chooses to be an artisan. This critical level is hardly affected by the possibility of investment and is still around $B = 0.5$. The dynamics for patience are also close to unchanged for low levels of patience up to about $B = 0.6$. For higher patience, however, the complementarity with physical investment comes into play, leading to high investment in patience for investors. The overall dynamics are divergent: dynasties starting out with low patience stay workers forever, and their patience approaches the medieval worker steady-state B_W. Dynasties that start out sufficiently patient first become artisans, and ultimately also invest. The wealth of the dynasty is increasing over time, and patience approaches a new, higher steady state.

Figure 6 displays the dynamics of patience and capital for a worker and artisan dynasty starting out with zero capital and the medieval steady-state levels of patience, B_W and B_A. For the worker, nothing changes: the dynasty does not invest, and patience remains at the steady state. The artisan dynasty, however, is sufficiently patient to find investment in capital attractive right away. Investment in capital increases the incentive for investing in patience, so that both patience and the growth rate of capital increase during a few periods, approaching a new balanced growth path.[17]

The landowners behave just like the workers. They have the same flat income profile (although possibly a higher income level) and the same low patience. At the return offered by the new technology, the landowners would actually want to borrow, but they cannot due to the borrowing constraint. The landowners therefore continue to live off their land rents, and are soon overtaken by the rising class of capitalists as the economically dominant group in society.

As the discussion in the preceding section should have made clear, some of the qualitative features of the case presented are not fully general. For example, the return to capital is, for obvious reasons, a key parameter: for very low returns, not even the artisans would want to use the new technology, and for extremely high returns even the workers and landowners would turn into capitalists. Likewise, we would not expect to observe a continuing divergence between the capitalists and the other classes if agents were highly risk averse, and the production function for patience highly inelastic. However, despite these caveats, it is a general conclusion that the most patient groups will be the first to make use of a new investment opportunity. Thus, even if the environment were such that ultimately even landowners invest, we would still expect the patient middle class to get a head start, and possibly overtake the landowning class in the process.

[17]Notice that the level of patience exceeds one after a few periods; this does not cause any problems in our overlapping generations economy, since the intergenerational discount factor is still fixed at z.

Thus, we believe that our theory can offer a plausible explanation for the failure of the landed aristocracy in Britain to benefit from the Industrial Revolution. To examine this hypothesis in more detail, we now turn to the historical circumstances that accompanied the decline of the British aristocracy.

6 The Historical Context

The thesis of this paper is that the failure of the aristocracy to take a more active role in the Industrial Revolution stemmed from a form of "cultural inappropriateness" which, in turn, had its roots in centuries of reliance on relatively flat and safe rents, which gave little incentives to parents to educate their children in terms of patience and frugality. While our focus on class-specific preferences may appear unusual from the standpoint of modern economic theory, this argument is very much in line with the perception of many contemporary observers. There are countless examples, both in scientific and fictional writing, of portrayals of members of the landowning class as inherently different from other people, and ill-disposed for commercial activity. To mention just one writer, Adam Smith (1776) describes the cultural attitudes of a landed aristocrat in the following words:

> *The situation of such a person naturally disposes him to attend rather to ornament which pleases his fancy, than to profit for which he has so little occasion. The elegance of his dress, of his equipage, of his house, and household furniture, are objects which from his infancy he has been accustomed to have some anxiety about.* (p. 410)

Landowners were not only preoccupied with conspicuous consumption, leaving them little time to consider profitable investment opportunities, but (according to Smith) did not consider money as something to be profitably invested to begin with:

> *A merchant is accustomed to employ his money chiefly in profitable projects; whereas a mere country gentleman is accustomed to employ it chiefly in expense. The one often sees his money go from him and return to him again*

> *with a profit: the other, when once he parts with it, very seldom expects to see any more of it.* (p. 432)

Smith, as well as many other classical economists and contemporary observers, found it entirely natural to think of members of different classes as essentially distinct beings whose behavior was governed by class-specific rules. In modern economic analysis, preferences are usually assumed to be fixed, and no systematic variation in tastes across different groups of people is allowed for. This view would lead one to believe that Smith and his contemporaries were simply wrong about the peculiar tastes and preferences of the aristocracy. In contrast, we hypothesize that the contemporary view was correct. We do not take the differences in class-specific preferences as exogenously given, however, but provide an economic rationale that explains how the preferences of the aristocracy and the other classes diverged in the first place. To assess the plausibility of our hypothesis, we now turn to historical evidence regarding the main assumptions and implications of our theory. We start by discussing evidence on the driving force behind our mechanism, namely, difference in class-specific income profiles.

6.1 Income Profiles Before the Industrial Revolution

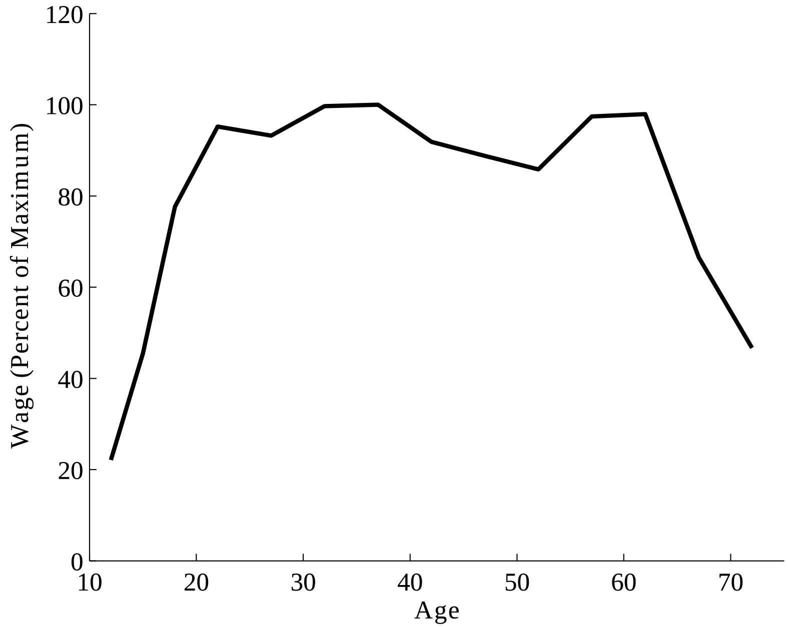
We start by taking a closer look at the life of pre-industrial artisans and craftsmen, i.e., the classes that, according to our theory, developed a high degree of patience as a consequence of the steep income profiles they faced over their lifetime. Pre-industrial artisans were, on average, a middle class, poorer than the aristocrats, merchants and bankers, but richer than unskilled workers. Their lives were very uneven: a hard youth was compensated by the hope for prosperity later in life. In most of Europe, an artisan's career was regulated by the statutes of local guilds and consisted of three periods: apprenticeship, journeymanship, and mastership. The life of an apprentice was not glamorous. "Upon payment of a placement fee, apprentices took their place in their master's household, agreeing to obey and respect him as a father [...] Not all apprentices reached mastership, but this does not gainsay the fact that the purpose of apprenticeship was selection and the goal a direct route to mastership [...]" (Farr 2000, p. 33). The length of the apprenticeship period varied over time, locations and professions. It would typically be 5–6 years, but in some professions one would remain an apprentice for up to 12 years (Epstein 1991).

After apprenticeship, artisans would become journeymen, and travel around European cities, serving as employees at some master's shop. This wandering period was usually compulsory, and would last for a minimum of 3–4 years (Friedrichs 1995). Frugality was essential for journeymen who hoped to become a master one day. "Unless he was able to count on substantial inheritance or fortunate marriage, a journeyman's primary interest was to amass capital for opening their shop or business" (Epstein 1991, p. 115). Having completed these wander-years, the journeyman could apply for admission to mastership, which was in itself a very expensive process.[18] Only at that point, if successful, could the journeyman become a master and a new guild member, and open a shop at his own expense. Not all journeymen managed their life to this happy conclusion, and many remained journeymen forever.

These accounts already show that the life of an artisan was highly investment-intensive. More direct evidence of the steepness of their consumption profile can be inferred from data on wage differences between journeymen and masters. Unfortunately, the availability of these data is limited to the complexity of contracts and the fact that journeymen were often paid piece-rate (nor do tax rolls help much, since they fail to report whether the taxpaying artisan was a journeyman).[19] Perhaps more informative are the data on wealth inequality. These suggest that both earnings and consumption inequality over the life-cycle were, on average, quite large. Farr (2000) shows that in the seventeenth century, the average dowry that daughters of master artisans of Dijon brought to their marriage

[18]The applicants owed the payment of a series of application fees, the completion of a masterpiece according to the regulation of each guild and the outlay (if the masterpiece was accepted) of a luxurious banquet for the masters he hoped to join. In addition, he had to submit the name of a proposed bride, which the guild was supposed to examine and approve.

[19]The only exceptions of which we are aware are the studies of Phelps Brown and Hopkins (1957) and Munro (2004). Based on the same dataset on builders wages in Southern England and the Southern Low Countries between 1346–1500, they report that in Bruges, the craftsmens' journeymen earned half of the master's wage, while the English journeymen earned two-thirds of their masters' wage. It is difficult to generalize these differences to other professions. In most of them, the relationship between shopowning masters and their employees was more hierarchical, so it is reasonable to guess that the wage differences were larger than in the construction sector.



Source: Burnette (2002)

Figure 7: Estimated Earnings Profile for Farm Laborers in England, 1830–1840

was three times as large as that of daughters of journeymen. Approximately the same ratio is documented to exist over a century later in a study on Lyon. Twice as large differences are found in the value of houses owned by masters and journeymen: a ratio of about 6:1 is evident in Delft between 1620 and 1644. In summary, the most conservative estimates point at a master-journeyman wealth ratio of 3:1.

In contrast, the wage profile of agricultural workers was relatively flat. Figure 7 describes the age-earning profile of a farm laborer in England in the early nineteenth century, based on the estimates of Burnette (2002). As shown by the figure, the profile is essentially flat between age 20 and 60. Consistent with our theory, this suggests that appreciation of the future was less valuable for rural workers than for urban ones.

As far as the aristocracy and landed gentry is concerned, the available evidence suggests that their income and consumption profiles were fairly flat as well. Members of this class derived their income mostly from owning land and, to a smaller extent, from mining projects (see Beckett 1986). Annual variation in income can therefore be linked to two dominant sources: fluctuation in land rental rates, and changes in the size of the estate through land sales or purchases. While there were always some families who managed their estates particularly successfully and were able to increase the size of their holdings, most aristocrats contented themselves with preserving the estate, ultimately passing to the next generation just as much as they once inherited. In periods of rising land rental rates, the income of the aristocracy as a class would be increasing as well; but given that rents tended to change only slowly over time, these movements would not generate the steep lifetime income profiles that were typical for artisans and craftsmen.

Taken by itself, the fact that the overall income of a given family was fairly constant over time does not imply that consumption profiles were flat as well. In particular, steep lifetime consumption profiles would arise if aristocracts started to consume heavily only after inheriting their estates, while living frugally during their younger years. However, the available evidence suggests that, if anything, the opposite was true. Young aristocrats typically did not work during their childhood and young adulthood. During this period of their lives, sons and daughters were supported by their parents. These family support payments tended to be quite large, and played a large role in aristocratic indebtedness: "family payments were not the only cause of aristocratic indebtedness, but contemporaries usually regard them as playing a crucial role" (Beckett 1986, p. 298). Thus, aristocrats usually lived in some comfort during their entire lives, and did not experience the stark contrast of a sober adolescence with relative prosperity during adulthood that was so typical for urban artisans and craftsmen.

Our theory also relies on the possibility for agents to choose between different occupations, especially urban versus rural activities. Clearly, in the pre-industrial world there were some barriers to geographical mobility. Nevertheless, mobility from countryside to cities and between cities was quite substantial. Large migratory flows are implied by the observation that, after the Black Death, the size of the urban population expanded dramatically, in spite of high mortality and low fertility that would have placed cities, absent migration, below the natural reproduction rate. A large inflow of people was necessary to maintain the demographic balance as "in spite of their vitality in the economic, politic, artistic, and cultural spheres, from a purely biological point of view the cities of pre-industrial Europe were large graveyards" (Cipolla 1994, p.134).

6.2 Who Were the New Industrialists?

It has long been part of the conventional wisdom on the Industrial Revolution that the majority of the entrepreneurs who came in to riches at the beginning of the nineteenth century were frugal individuals from the middle and lower classes: "The early industrialists were for the most part men who had their origin in the same social strata from which their workers came. They lived very modestly, spent only a fraction of their earnings for their households and put the rest back into the business" (von Mises 1963, p.622).

More recent studies suggest that von Mises may have exaggerated the role of self-made men coming from the lower classes. In a study of founders of large industrial undertakings in Britain between 1750 and 1850, Crouzet (1985) concludes that "neither the upper class nor the lower orders made a a large contribution to the recruitment of industrialists [...] Indeed, the retreat by landowners from involvement in industry, which had started in the seventeenth century, became more and more pronounced as the eighteenth century progressed" (p.68). The only class that was significantly over-represented among the industrialists was the middle class. Quantitatively, only 2.3 percent of the industrialists in the sample came from peerage and gentry.[20] In contrast, as many as 85 percent of the new industrialists had a middle-class background.[21] The contribution of the

[20] If one adds officers in the Army or Navy, classified by Crouzet as "upper class" the figure goes up to 3.0 percent (see Crouzet's Table 5).

[21] The middle class entrepreneurs in Crouzet's classification are a heterogeneous group including bankers and rich merchants at the upper end, and small artisans and tenant farmers at the lower end. A large proportion of them started with limited financial resources, and made their fortune entirely through self-financed investment. As many as 27 percent of the men who entered large-scale industry and 39 percent of the fathers of industrialists came from the lower ranks of the middle class: shopkeepers, self-employed craftsmen, cultivators of various kind (Crouzet (1985) p.127).

working class was very moderate; no more than 12 percent of the industrialists belonged to this class, which accounted for about 70 percent of the population.

Part of the explanation for the small number of aristocratic entrepreneurs is, of course, that there were few aristocrats to begin with. But the differences in numbers do not explain the extent of the underrepresentation of the upper classes. At the beginning of the nineteenth century, peerage and gentry accounted for about 1.4 percent of the population, while the middle class made up slightly less than 30 percent. Thus, a much larger share of the middle class than of the peerage and gentry ended up as entrepreneurs. If we relate the participation of the upper class to their share of wealth owned instead of their share of the population, their representation is surprisingly thin.

Even the estimate of 2.3 percent peers and gentry among the industrialists may overstate the true involvement of the upper classes in entrepreneurial activities. In many cases where landowners are listed among the founders of industrial undertakings, they became involved only by virtue of owning the land on which an industrial activity was to take place. In the majority of these cases, the aristocrats did not actively participate in the entrepreneurial process: "If they owned blast-furnaces, forges and other establishments, they tended to lease them to tenants rather than to operate them through salaried managers . . . [They] were rather passive lessors and investors than active business leaders" (Crouzet 1985 p. 68). Similarly, in the textile industries "they were content to build (or to help to build) mills and to lease them out" (Crouzet 1985 p. 74). Even the aristocracy's involvement in the financing of mining and railways throughout the nineteenth century rarely took the form of a direct involvement in entrepreneurial and managerial aspects. Instead, landowners insisted on receiving regular periodical payments over the sums invested, without any participation in the risk or any commitment to financing the growth of the enterprises.

The marginal involvement of the aristocracy is surprising, given the extreme concentration of wealth in the hands of the landowning elite at the time. As late as in 1880, less than 5000 landowners still owned more than 50 percent of all land (Cannadine 1990). Given their enormous advantage in wealth, the aristocrats should have been well placed to profit from new technologies that are ultimately based on capital investment.

6.3 Lack of Patience Versus Alternative Explanations

Clearly, lack of patience capital on behalf of the landowners is not the only possible explanation for the failure of the aristocracy to profit from industrialization. A complementary explanation for the success of the middle class is that urban workers possessed skills that were essential for industrial activities, while the landowners did not. While this factor must have also played a role, a number of observations put into question whether it could, alone, explain such a major social transformation. First of all, a large share of the new industrialists had not previously been involved in manufacturing activity. For instance, as many as 22 percent of the industrialists' fathers were yeomen and farmers, groups with no history of previous involvement in industrial activity (Crouzet 1985).[22] It seems unlikely that these entrepreneurs possessed specific work skills suitable to industrial activity. Moreover, there is evidence of substantial mobility within industrial sectors. For instance, the boom of the textile industry attracted many outsiders into this thriving business. In 1787, 28 percent of the entrepreneurs in the textile industry came from non-textile trades (Crouzet 1985, p.120). It therefore appears that in terms of their skills, landowners were not at a particular disadvantage relative to many of the middle-class entrepreneurs. In fact, during the industrialization, a number of important sectors (such as mining and railways) required land as a major input. In these sectors, if anything, the landowners should have had an advantage over many of the middle-class city dwellers.

A related argument is that the landowners, busy managing their rural estates, may have lacked the time and opportunity to enter industrial activities, which mostly took place in or near cities. However, many landowners did not actively manage their estates, so in many cases this was not a binding constraint. Even

[22]On the other hand, tenant farmers were heavily engaged in investments in agriculture during the eighteenth century, resulting in large productivity improvement. Since their access to capital markets was limited, their consumption profiles must have been steeper than those of landowners and laborers. Thus, it is conceivable that they had also accumulated a great deal of patience capital.

	1752–1799	1800–1849	1850–1899
Church	60	62	38
Land-Owning	14	14	7
Teaching	9	9	12
Law	6	9	14
Administration	3	1	6
Medicine	1	2	7
Banking	0	0	2
Business	0	0	5
Other	7	3	9

Source: Jenkins and Jones (1950), Table 1

Table 2: Professional Choices of Cambridge Graduates, in Percent

more telling, it was not only the heirs who owned estates who shunned business activity, but also second and third sons of landowners. These younger sons had no choice but to enter some activity other than landowning, and were therefore clearly not held back by their obligations to an existing estate; nevertheless, they did not enter business in any larger numbers than their landowning fathers. Table 2 reports the professional choice of Cambridge graduates during the period 1750–1899. The vast majority of students at Cambridge during this period were sons of members of the landowning class, so their professional choices (other than landowning) give us a good idea of which professions younger sons entered. Strikingly, until 1850, not a single graduate got involved in banking or business (widely defined as any "profit-oriented activity"), and even after 1850 the percentage remains surprisingly low.

Our hypothesis of a lack of patience among the aristocracy leads to a number of additional implications that can be compared to historical evidence. In particular, a low time discount factor should make an aristocrat reluctant to invest not only in new industrial enterprises, but also in financial assets in general. Instead, we would expect the aristocracy to borrow from other groups in the population, in order to finance current consumption expenditures. These implications are supported by historical evidence. Well before the Industrial Revolution, the British government was a major borrower in the economy, with multiple issues of government bonds during the seventeenth and eighteenth centuries. This government borrowing was mostly financed by the middle classes. "The bourgeoisie were the most important individual investors in government loans. They included members of the church, civil service and professions, but the majority were merchants and financiers, or, as smaller owners, tradesmen, craftsmen and artisans." (Dickson 1967, page 302). With regards to long-term bond holdings by the landed aristocracy, Dickson concludes: "It seems fair to generalize that the landed classes as a whole were not significant contributors of new capital for public loans" (Dickson 1967, page 302).

The same pattern can be found when we consider investments in public companies. Bowen (1989) describes the composition of stockholders of the East India Company between 1756 and 1791 as being well-represented by "clergymen, bankers, military and naval personnel, officials, brokers, merchants large and small, and retailers." Regarding the upper classes, Bowen concludes: "One issue of central importance is beyond doubt: there was no large-scale investment in the [East India] Company by the landed interest or aristocracy." (p. 195).

In summary, we find that even well before the Industrial Revolution the landed elite plays a surprisingly minor role in financing government borrowing and private enterprise, despite the fact that this group was far wealthier than the middle-class investors. This behavior stands in marked contrast to the wealth elites in modern industrial countries, which generally own a disproportionate share of most classes of assets, including government debt and public stock.[23]

While the lack of financial investment by the aristocracy is indicative of low patience, the evidence is not yet conclusive. If investments in the aristocracy's agricultural estates carried a higher return than financial assets, it may be the case that the upper classes merely held a different (and possibly more profitable) portfolio than the middle classes, even with the same preferences on the consumption-savings margin. However, this interpretation is contradicted by two observa-

[23]For the case of the United States see Carroll (2001) who documents that "rich households are more likely to own virtually every kind of asset."

tions. The first observation is the style in which they managed their own estates. During the course of the 1700, we observe a progressive withdrawal from day-to-day involvement with farming in both France and Britain. Investments and technical innovation in agriculture were carried out almost exlusively by tenant farmers, not by landowners, who remained content with relying on safe flat rents (see Thompson (1994)).

The second observation is the evidence that the landed aristocracy, rather than investing the rents from their estates in interest-bearing activities, borrowed increasing amounts of resources to finance luxury consumption. Borrowing by the aristocracy had been common long before the Industrial Revolution. Beckett (1986) reports that by the mid-eighteenth century "many families already had an accumulation [of debt] several generations old" (p. 300). Typically, the majority of this debt was not taken on to improve existing estates or to buy more land, but resulted from a failure to match expenditure to income: "Rents and royalties were apparently being sucked into conspicuous consumption and frittered away in spiraling marriage contracts; and the gap between getting and spending was filled not by offloading assets such as land, but by borrowing from—in effect—the commercial, industrial and shopkeeping members of the populace." (Beckett 1986, page 316). Indebtness grew severely during the nineteenth Century. One 1847 writer claimed that "between half and two-thirds of English land was encumbered (i.e. mortgaged)" (Beckett 1986, page 315). Cannadine (1994) summarizes the situation as follows: "Whatever might have been the financial state of financial families, it seems clear that the landed aristocracy *as a class* was in debt through the first three-quarters of the nineteenth century" (p. 49). While some of this debt was raised for investment in non-agricultural ventures, "the first [category] was spending which had its objective the enhancement of the social prestige and the fulfillment of the traditional responsibilities of the landowner [...] To the extent that such self-indulgent activities were financed from middle- and working-class savings, [...] this definitely amounted to a 'haemorrhage of capital,' a 'misallocation of resources,' as funds from urban and industrial Britain were diverted to underpin the indulgence of the landed order" (p. 48–49).

These views are corroborated by a number of sources describing the extravagant consumption habits of the aristocracy. In an analysis of landowners in the Glasgow area between 1770-1815, Devine (1971) argues that extensive expenditures can be attributed to what has been dubbed a "Revolution in Manners" on the part of the landed gentry. [24] The same author documents that despite the fact that land rents were rising during the period, a number of old families increased consumption even faster than income, and were ultimately forced to sell their estates. These views are echoed by Porter (1982), who describes the eighteenth century as the golden age for the landed aristocracy, but also notes that the tide of wealth flowed out as fast as it came to finance an increasingly expensive lifestyle.[25]

In contrast, the new industrialists were regarded by their contemporaries as parsimonious beings, eager to accumulate wealth. According to accounts of the time reported by Crouzet (1985), most Mancunian manufacturers of the late eighteenth century "commenced their careers in business with but slenders capitals [...] Patience, industry and perseverance was their principal stock" (p. 37). Similarly, the cotton spinner John Kennedy declared in 1828 that "the only men who have made their fortunes [in Manchester] have been those who started with nothing. They lived only for their businesses, and brought up in the habits of strict economy."

[24] He writes: "The desire for an increased standard of living among the aristocracy and landed gentry led several of them to live above their means. More varied leisure activities, more elaborate clothing, 'improvement' of estates, more exotic diets all required an increased income. Such were the pressures of social competition and convention that often desires were fulfilled on very narrow financial margins" (page 218).

[25] Here are some examples: "Sir Robert Walpole's private expenditure between 1714 an 1718 totaled £90,000. Walpole and his guests at Houghton Hall in Norfolk downed about £1,500 a year in wine [...] Bedford House had forty-two servants who cost £859 a year. Political and electoral expenses ran into thousands [...] Marriage settlements [...] became heavier drains on fortunes [...] Vast capital fortunes were sunk in building [...] Prodigal peers and their heirs ran up astronomical debts [...] Magnates ransacked Europe for paintings, sculpture, furniture, jewelry. They patronized artists and poets, and collected antiques, scientific instruments and books by the roomful" (p. 59–60).

The Demise of the British Aristocracy

During the early phases of the Industrial Revolution, the landowning class retained most of its political influence, and continued to account for a large fraction of overall wealth. While industrial fortunes were growing fast, the new industrialists usually started with little initial wealth, and therefore took some time to catch up with the major landowners. In addition, the rapid population growth increased the demand for food, which led to a substantial, if temporary, increase in the value of land. In the first half of the nineteenth century, large fortunes were by and large still associated with land ownership. Rubinstein (1981) reports that among the 189 individuals who died between 1809 and 1858 with a fortune exceeding one million pounds, 95 percent were wealthy landowners.

After 1880, however, the rental income from land started to decline, and was no longer sufficient to maintain the lavish lifestyle to which the aristocracy had been accustomed. The aristocrats reacted by increasing even further their already chronic indebtness. Many families became overburdened by debt, until they were forced to sell off parts or all of their estates. Cannadine (1990) writes: "Above all, there was a massive dispersal of land, throughout the British Isles, as estates tumbed into the market in the years immediately before and after the First World War. The scale of this territorial transfer was rivaled only by two other landed revolutions in Britain this Millennium: The Norman Conquest and the Dissolution of the Monasteries" (p. 89). As a result, the land distribution in modern times is nowhere near as concentrated as it was early in the nineteenth century (see Cannadine 1990, 1994). A century after the start of the Industrial Revolution, landowners no longer featured prominently among the wealthiest families in the country. Between 1900 and 1939, only 7 percent of the 273 individuals who died as millionaires belonged to the landed elite.[26] Among the non-landed millionaires, about half of the new fortunes were generated in the manufacturing sector, with most of the rest accounted for by commerce and finance.

Economic decline ultimately led to waning political influence. By a series of franchise extensions throughout the nineteenth century, voting rights were extended

[26] See Rubinstein (1981), Tables 3.2, 3.3 and 3.4.

to the entire male population, and as a consequence the landowning class, which made up only a small fraction of total population, could no longer exert much political power. Even during the first half of the nineteenth century, when voting rights were still limited, the major political debates (concerning, for example, the repeal of the Corn Laws) were increasingly resolved in favor of the industrialists, and against the landowning interest. By the dawn of the twentieth century, the aristocracy's economic and political dominance, at its peak a mere hundred years earlier, had been irrevocably lost.

7 Conclusions

In this paper, we develop a theory that relies on endogenous investments in patience to explain the economic decline of the aristocracy after the start of the Industrial Revolution, as well as the the emergence of a class of industrialists from middle-class origins. Our theory provides a link between a literature emphasizing the role of technological change in long-run growth, and a literature that focuses on the role of political reforms and institutions for economic development. Technological change is important in our model, because the latent class differences in terms of patience only become paramount after the arrival of a new investment-based technology. Following this technological impulse, the model provides an account of the relative economic fortune of different classes, which in itself was a driving force behind many of the political and institutional changes that followed the Industrial Revolution.

The theory provides a new perspective of the impact of financial market frictions on economic development. In our model, the absence of financial markets imply that incentives for investing in patience differ across classes and occupations. In an environment where upper-class wealth is based on land ownership (which provides relatively flat returns over the life cycle), it is members of investment-intensive middle-class occupations that become the most patient. As a result, middle-class entrepreneurs ultimately leapfrog over the stagnant upper class and become economically dominant. This view stands in contrast to an existing literature on inequality and development, which emphasizes that financial market imperfections may exclude poorer individuals from profitable investment. The existing literature may provide another reason for the exclusion of the working class from entrepreneurial activity, but it cannot explain why middle-class entrepreneurs surpassed the existing wealth elite at the beginning of the Industrial Revolution.

The theory also highlights a new implication of financial development: by increasing the possibilities for agents to smooth consumption, well-developed financial markets reduce the extent to which incentives to invest in patience vary across families engaged in different occupations. Thus, financial development leads to more homogeneous preferences within the population. To this end, our analysis suggests that in pre-industrial times members of different classes were really distinct from each other on a fundamental level (in addition to the obvious differences in wealth and power), just as the contemporary observers generally seemed to believe. In modern societies with rich financial markets, the assumption that members of all classes have the same underlying preferences (as is usually assumed in economic modeling) may be more appropriate.

A Mathematical Appendix

A.1 The Sequential Formulation of the Decision Problem

The sequential formulation of the decision problem is given by:

$$v^{\star}(B_0) = \max \left\{ \sum_{t=0}^{\infty} z^t \left[u(y_{1,i_t}) h(1 - l_t) + B_t u(y_{2,i_t}) \right] \right\}, \tag{18}$$

subject to:

$$\begin{aligned} B_{t+1} &= (1-\nu)B_t + f(l_t), \\ i_t &\in I, \\ l_t &\in (0,1). \end{aligned}$$

A.2 Proofs for all Propositions

Proof of Proposition 1: The proof is an application of Corollary 1 to Theorem 3.2 in Stokey and Lucas (1989). The Bellman equation (2) defines a mapping T on the space of bounded continuous functions on the interval $[0, B_{\max}]$, endowed with the sup norm, where the mapping is given by:

$$Tv(B) = \sup_{i \in I, 0 \leq l \leq 1} \left\{ u(y_{1,i}) h(1-l) + B u(y_{2,i}) + z v((1-\nu)B + f(l)) \right\}. \tag{19}$$

Since we assume $0 < z < 1$, this mapping is a contraction by Blackwell's sufficient conditions, and it therefore has a unique fixed point by the Contraction Mapping Theorem. Using Corollary 1, we can now establish that the value function (the fixed point of the mapping T) is increasing and weakly convex by establishing that the operator T preserves these properties.

To establish that the value function is increasing, let v be a non-decreasing bounded continuous function. We need to show that Tv is a strictly increasing function. To do this, choose $\overline{B} > \underline{B}$. We now need to establish that $Tv(\overline{B}) > Tv(\underline{B})$. Since the right-hand side of (19) is the maximization of a continuous function over a compact set, the maximum is attained. Let $\underline{l}$ and $\{\underline{y}_1, \underline{y}_2\}$ be choices attaining the maximum for $\underline{B}$. We then have:

$$\begin{aligned} Tv(\overline{B}) &\geq u(\underline{y}_1)h(1-\underline{l}) + \overline{B}u(\underline{y}_1) + zv((1-\nu)\overline{B} + f(\underline{l})) \\ &> u(\underline{y}_1)h(1-\underline{l}) + \underline{B}u(\underline{y}_1) + zv((1-\nu)\underline{B} + f(\underline{l})) = Tv(\underline{B}), \end{aligned}$$

which is the desired result. Here the weak inequality follows because the choices $\underline{l}, \{\underline{y}_1, \underline{y}_2\}$ may not be maximizing at $\overline{B}$, and the strict inequality follows because v is assumed to be increasing, and we have that $\overline{B} > \underline{B}$ and $u(\underline{y}_2) > 0$.

To establish convexity of the value function, let v be a (weakly) convex bounded continuous function. We need to establish that Tv is also a convex function. To show this, choose a number θ such that $0 < \theta < 1$, let $\overline{B} > \underline{B}$, and let $B = \theta\overline{B} + (1-\theta)\underline{B}$. We now need to show that $\theta Tv(\overline{B}) + (1-\theta)Tv(\underline{B}) \geq Tv(B)$. Let l and $\{y_1, y_2\}$ be choices attaining the maximum for B. Since these are feasible, but not necessarily optimal choices at $\overline{B}$ and $\underline{B}$, we have:

$$\begin{aligned}
Tv(\overline{B}) &\geq u(y_1)h(1-l) + \overline{B}u(y_2) + zv((1-\nu)\overline{B} + f(l)), \\
Tv(\underline{B}) &\geq u(y_1)h(1-l) + \underline{B}u(y_2) + zv((1-\nu)\underline{B} + f(l)).
\end{aligned}$$

Working towards the desired condition, we therefore have:

$$\begin{aligned}
\theta Tv(\overline{B}) + (1-\theta)Tv(\underline{B}) \geq& \theta\left[u(y_1)h(1-l) + \overline{B}u(y_2) + zv((1-\nu)\overline{B} + f(l))\right] \\
&+ (1-\theta)\left[u(y_1)h(1-l) + \underline{B}u(y_2) + zv((1-\nu)\underline{B} + f(l))\right] \\
=& u(y_1)h(1-l) + Bu(y_2) \\
&+ z\left[\theta v((1-\nu)\overline{B} + f(l)) + (1-\theta)v((1-\nu)\underline{B} + f(l))\right] \\
\geq& u(y_1)h(1-l) + Bu(y_2) + zv((1-\nu)B + f(l)) \\
=& Tv(B),
\end{aligned}$$

which is the required condition. Here, the last inequality follows from the assumed convexity of v. The operator T therefore preserves convexity, and thus the fixed point must also be convex. Notice that linearity is key to this result: the discount factor enters linearly utility, and the parental discount factor has a linear effect on the discount factor of the child. □

Proof of Proposition 2: We start by showing that the steepness of the optimal income profile is non-decreasing in B, and that the optimal investment in patience $l(B)$ is non-decreasing in B. Fix two current discount factors $\underline{B} < \overline{B}$, and let the corresponding optimal choices be $\underline{l}, \underline{y}_1, \underline{y}_2$ and $\overline{l}, \overline{y}_1, \overline{y}_2$. We want to show that $\underline{l} \leq \overline{l}$, $\underline{y}_1 \geq \overline{y}_1$, and $\underline{y}_2 \leq \overline{y}_2$.

Since the choices are optimizing for the $\underline{B}$ and $\overline{B}$ agents, the following inequalities must be satisfied at the optimal choices:

$$\begin{aligned}
&u(\overline{y}_1)h(1-\overline{l}) + \overline{B}u(\overline{y}_2) + zv((1-\nu)\overline{B} + f(\overline{l})) \\
&\qquad \geq u(\underline{y}_1)h(1-\underline{l}) + \overline{B}u(\underline{y}_2) + zv((1-\nu)\overline{B} + f(\underline{l})), \quad (20)
\end{aligned}$$

$$u(\overline{y}_1)h(1-\bar{l}) + \underline{B}u(\overline{y}_2) + zv((1-\nu)\underline{B} + f(\bar{l}))$$
$$\leq u(\underline{y}_1)h(1-\underline{l}) + \underline{B}u(\underline{y}_2) + zv((1-\nu)\underline{B} + f(\underline{l})), \quad (21)$$

where the first inequality follows from optimization at $\overline{B}$ and the second from optimization at $\underline{B}$. Subtracting (21) from (20) on both sides, we get the following condition:

$$(\overline{B} - \underline{B})\left[u(\overline{y}_2) - u(\underline{y}_2)\right] \geq z\left[v((1-\nu)\overline{B} + f(\underline{l})) - v((1-\nu)\underline{B} + f(\underline{l}))\right]$$
$$- z\left[v((1-\nu)\overline{B} + f(\bar{l})) - v((1-\nu)\underline{B} + f(\bar{l}))\right]. \quad (22)$$

Here, the sign of the left-hand side is equal to the sign of $\overline{y}_2 - \underline{y}_2$, and, because of the convexity of v and the fact that f is increasing, the right-hand side is non-negative if $\bar{l} \leq \underline{l}$, and non-positive if $\bar{l} \geq \underline{l}$. Taken these implications together, (22) implies that we must have $\overline{y}_2 \geq \underline{y}_2$ or $\bar{l} \geq \underline{l}$, because otherwise the left-hand side is negative and the right-hand side is non-negative. Thus, so far we know that at least one of our two claims, namely that patient agents choose steeper income profiles and invest more in patience, must be true. To show that in fact both are true, we now proceed to establish that each implies the other.

Let us therefore assume that $\bar{l} \geq \underline{l}$. Optimization in the choice of the income profile implies the following inequalities:

$$u(\overline{y}_1)h(1-\bar{l}) + \overline{B}u(\overline{y}_2) \geq u(\underline{y}_1)h(1-\bar{l}) + \overline{B}u(\underline{y}_2), \quad (23)$$
$$u(\overline{y}_1)h(1-\underline{l}) + \underline{B}u(\overline{y}_2) \leq u(\underline{y}_1)h(1-\underline{l}) + \underline{B}u(\underline{y}_2). \quad (24)$$

Subtracting the two equations as before, we get:

$$(\overline{B} - \underline{B})\left[u(\overline{y}_2) - u(\underline{y}_2)\right] \geq \left[h(1-\underline{l}) - h(1-\bar{l})\right]\left[u(\overline{y}_1) - u(\underline{y}_1)\right].$$

Since $\bar{l} \geq \underline{l}$ and h is strictly increasing, the first term on the right-hand side is non-negative. Therefore, we must have $\overline{y}_2 \geq \underline{y}_2$ and, consequently, $\overline{y}_1 \leq \underline{y}_1$ since otherwise the left-hand side is negative and the right-hand side non-negative.

Conversely, suppose we already know that $\overline{y}_2 \geq \underline{y}_2$, which also implies that $\overline{y}_1 \leq \underline{y}_1$. We want to establish that $\bar{l} \geq \underline{l}$. Optimization in the choice of l implies:

$$u(\overline{y}_1)h(1-\bar{l}) + zv((1-\nu)\overline{B} + f(\bar{l})) \geq u(\overline{y}_1)h(1-\underline{l}) + zv((1-\nu)\overline{B} + f(\underline{l})), \quad (25)$$
$$u(\underline{y}_1)h(1-\bar{l}) + zv((1-\nu)\underline{B} + f(\bar{l})) \leq u(\underline{y}_1)h(1-\underline{l}) + zv((1-\nu)\underline{B} + f(\underline{l})). \quad (26)$$

Combining the two conditions one more time, we get:

$$\begin{aligned}\left[u(\overline{y}_1) - u(\underline{y}_1)\right] &\left[h(1-\bar{l}) - h(1-\underline{l})\right] \\ &\geq z\left[v((1-\nu)\overline{B} + f(\underline{l})) - v((1-\nu)\underline{B} + f(\underline{l}))\right] \\ &\qquad - z\left[v((1-\nu)\overline{B} + f(\bar{l})) - v((1-\nu)\underline{B} + f(\bar{l}))\right]. \end{aligned} \tag{27}$$

Here, the first term on the left-hand side is non-positive. If $\overline{y}_1 < \underline{y}_1$, we must have $\bar{l} \geq \underline{l}$, because otherwise the left-hand side is negative and the right hand side is non-negative, because of the convexity of v. If, on the other hand, $\overline{y}_1 = \underline{y}_1$, the left-hand side is zero, and we must therefore ensure that the right-hand side is non-positive. Here two cases need to be distinguished, depending on the curvature of v. First, if v is strictly convex anywhere on the interval $[(1-\nu)\underline{B}+f(\min\{\underline{l},\bar{l}\}),(1-\nu)\overline{B} + f(\max\{\underline{l},\bar{l}\})]$, we must have that $\bar{l} \geq \underline{l}$, because otherwise the right-hand side is strictly positive. Second, it is also possible that v is exactly linear over the relevant range, in which case both the left- and right-hand sides of (27) are zero, regardless of $\underline{l}$ and $\bar{l}$. To satisfy (27), in this case both (25) and (26) must hold as inequalities, implying that both agents are indifferent between $\underline{l}$ and $\bar{l}$. Given that v is linear over the relevant range, our concavity assumptions on h (strict concavity) and f imply that (given the optimal income profile) for each agent there is a unique optimal l. Therefore, we must have $\underline{l} = \bar{l}$, and once again the desired condition is satisfied. □

Proof of Proposition 3: In Proposition 2, we established that the steepness of the optimal income profile is increasing in B, and that the optimal choice of investment in patience $l(B)$ is also increasing in B. It then follows that the patience as well as the steepness of the income profiles of all future members of a dynasty (child, grandchild etc.) are increasing in the patience of the current member of a dynasty.

Since there are only finitely many occupations, we can subdivide the state space $[0, B_{\max}]$ into finitely many closed intervals (they are closed because of our continuity assumptions), where each interval corresponds to the choice of a given occupation i. The agent is just indifferent between two occupations at the boundary of two such intervals, and strictly prefers a given occupation in the interior of such an interval. The intervals can be further subdivided according to the occupational choice of the child. Since $l(B)$ may not be singled valued, there may be multiple optimal B' corresponding to a given B today. Nevertheless, since the B' are strictly increasing in B (because of Proposition 2 and $\nu < 1$) and given that there are only finitely many occupations, we can once again subdivide today's state space in finitely many close intervals, each one corresponding to a specific occupational choice of the child, such that the intervals only overlap only at their boundary points. Continuing this way, the state space $[0, B_{\max}]$ can be divided into a countable number of closed intervals (there is a finite number of possible occupations in each of the countably many future generations), where each interval corresponds to a specific occupational choice of each generation. Let $[\underline{B}, \overline{B}]$ be such an interval. We want to establish that the value function is linear over this interval, and that the optimal choice of patience $l(B)$ is single-valued and constant over the interior of this interval.

It is useful to consider the sequential formulation (18) of the decision problem. Taking the present and future occupational choices i_t as given, we can substitute for B_t and write the remaining decision problem over the l_t on the interval $[\underline{B}, \overline{B}]$ as:

$$v(B) = \max \Bigg\{ u(y_{1,i_0})h(1-l_0) + Bu(y_{2,i_0})$$

$$+ \sum_{t=1}^{\infty} z^t \left[u(y_{1,i_t})h(1-l_t) + \left((1-\nu)^t B + \sum_{s=0}^{t} (1-\nu)^{t-s-1} f(l_s) \right) u(y_{2,i_t}) \right] \Bigg\}. \quad (28)$$

For given current and future income profiles, (28) is strictly concave in l_t for all t, since h is assumed to be strictly concave, and f is weakly concave. Moreover, the discount factor B and all expressions involving l_t appear in separate terms in the sum. Therefore, it follows that, given the optimal income profiles, for all t the optimal l_t is unique, and independent of B. Since on the interior of $[\underline{B}, \overline{B}]$, the current and future optimal income profiles are unique, the optimal policy correspondence $l(B)$ is single-valued. At the boundary between two intervals there are (by construction of the intervals) at least two different optimal income profiles for at least one generation, hence $l(B)$ may take on more than one optimal value, one corresponding to each optimal set of income profiles.

The optimal value function v over the interval $[\underline{B}, \overline{B}]$ is given by (28) with income profiles i_t and investment in patience l_t fixed at their optimal (and constant) values. (28) is linear in B; it therefore follows that the value function is piece-wise linear, with each kink corresponding to the boundary between two of our intervals. □

Proof of Proposition 4: The law of motion $g : [0, B_{\max}] \to [0, B_{\max}]$ for B is given by:

$$g(B) = (1-\nu) B + f(l(B)),$$

where $l(B)$ is a non-decreasing step function (as described in Proposition 3). Since f is an increasing function and we assume that $\nu < 1$, the law of motion $g(B)$ is strictly increasing in B. Notice that $g(B)$ may not be single valued for all

B. Strictly increasing here means that $\overline{B} < \underline{B}$ implies $\overline{B}' < \underline{B}'$ for all $\overline{B}' \in g(\overline{B})$ and $\underline{B}' \in g(\underline{B})$, even if $g(\overline{B})$ or $g(\underline{B})$ is a set. For a given B_0, the law of motion g defines (potentially multiple) optimal sequences of discount factors $\{B_t\}_{t=0}^{\infty}$. Any such sequence is a monotone sequence on the compact set $[0, B_{\max}]$, and must therefore converge. Notice, however, that since $l(B)$ is not single-valued everywhere, different steady states can be reached even from the *same* initial B_0. □

Proof of Proposition 5: That the only income profile chosen in equilibrium are those maximizing $y_{1,I} + y_{2,i}/R$ is an immediate consequence of the fact that, under perfect capital markets, agents can allocate consumption optimally given any present value of income. Therefore, only the present value of income matters. Thus, agents only choose professions yielding the maximum attainable present value of income. That the parents investment in patience does not depend on which among the professions that maximize the present value of income is chosen by the future offsprings is also an immediate consequence of the fact that, under perfect capital markets, consumption profiles are decoupled from income profiles.

To establish that the value function is increasing and convex, we proceed as in the proof of Proposition 2. To establish that it is increasing, let v be a non-decreasing bounded continuous function. We show that Tv is a strictly increasing function. For this purpose, choose $\overline{B} > \underline{B}$. As discussed in the other proof, we need to establish that $Tv(\overline{B}) > Tv(\underline{B})$. Let c_1 and c_2 denote consumption in the two periods of adult life, given a professional choice that maximizes the present value of income. Let $\underline{l}$ be the choice that attain the maximum for $\underline{B}$. We then have:

$$\begin{aligned} Tv(\overline{B}) &\geq u\left(c_1\right) h(1-\underline{l}) + \overline{B}u(c_2) + zv((1-\nu)\overline{B} + f(\underline{l})) \\ &> u\left(c_1\right) h(1-\underline{l}) + \underline{B}u(c_2) + zv((1-\nu)\underline{B} + f(\underline{l})) = Tv(\underline{B}), \end{aligned}$$

which is the desired result. As in the proof of Proposition 2, the weak inequality follows because the choice $\underline{l}$, c_1 and c_2 may not be maximizing at $\overline{B}$, and the strict inequality follows because v is assumed to be increasing, and we have $\overline{B} > \underline{B}$ and $u(c_2) > 0$.

To establish the convexity of the value function, let v be a (weakly) convex bounded continuous function. We need to establish that Tv is also a convex function as well. To show this, let θ be such that $0 < \theta < 1$, let $\overline{B} > \underline{B}$, and let $B = \theta\overline{B} + (1-\theta)\underline{B}$. We now need to show that $\theta Tv(\overline{B}) + (1-\theta)Tv(\underline{B}) \geq Tv(B)$. Let l be the choice that attains the maximum for B. Since this is feasible, but not necessarily an optimal choice at $\overline{B}$ and $\underline{B}$, we have:

$$
\begin{aligned}
Tv(\overline{B}) &\geq u(c_1)h(1-l) + \overline{B}u(c_2) + zv((1-\nu)\overline{B} + f(l)), \\
Tv(\underline{B}) &\geq u(c_1)h(1-l) + \underline{B}u(c_2) + zv((1-\nu)\underline{B} + f(l)).
\end{aligned}
$$

Proceeding as in the proof of Proposition 2 leads to establish that:

$$\theta Tv(\overline{B}) + (1-\theta)Tv(\underline{B}) \geq Tv(B),$$

which is the required condition. □

Proof of Proposition 6: Concerning occupational segregation, first notice that for each cohort an equal number of parents and children end up working in each of the two professions, since the relative supply of labor in the two sectors must be constant. Next, since according to Propositions 3 and 4 the choice of occupation is increasing in patience B (when occupations are ordered by the steepness of the income profile), the patience of any artisan is larger than or equal to the patience of any worker. Finally, the patience of any artisan's child is strictly larger than the patience of any worker's child, since the child's patience is strictly increasing in the parent's patience and occupational choice. The artisans' children as a group are therefore strictly more patient than the workers' children, which implies that they must be those who choose to be artisans in the next generation.

That patience must converge in each dynasty follows from Proposition 4. It remains to be shown that agricultural workers and landowners converge to the same patience B_W, whereas artisans converge to $B_A > B_W$. To this end, consider the first-order condition for investment in patience (6) under CRRA preferences:

$$\frac{h'\left(1-\bar{l}\right)}{f'(\bar{l})} = \frac{z}{1-z(1-\nu)}\left(\frac{y_2}{y_1}\right)^{\sigma}.$$

Notice that only the income ratio y_2/y_1 enters the condition; the level is irrelevant. Since agricultural workers and landowners have the same steepness of the income profile ($y_2/y_1 = 1$), they converge to the same patience B_W. Artisans have a steeper profile $y_2/y_1 = \gamma$, and consequently converge to a higher level of patience $B_A > B_W$. □

Proof of Proposition 7: We start be establishing the following claim. Let $u(c) = c^{\sigma}$ with $\sigma \in (0,1)$. Then,

$$
\begin{aligned}
v_A\left(\tilde{B}\,|\mu\right) &= \left(\alpha\mu^{\alpha-1}\right)^{\sigma} V_A\left(\tilde{B}\right), \\
v_M\left(\tilde{B}\,|\mu\right) &= q^{\sigma} V_M\left(\tilde{B}\right),
\end{aligned}
$$

where

$$V_A\left(\tilde{B}\right) = \max_{B'}\left\{h(1-(B'-(1-\nu)\tilde{B}))+\tilde{B}+zV_A(B')\right\} \text{ and} \tag{29}$$

$$V_M\left(\tilde{B}\right) = \max_{B'}\left\{h(1-(B'-(1-\nu)\tilde{B}))+\tilde{B}\gamma^{\sigma}+zV_M(B')\right\} \tag{30}$$

are increasing, convex functions.

Consider, to this aim, the occupational choice of the first (and second) generation given the expectation of an ECW with μ unskilled workers. [27] Since, initially, all landless agents have the same preferences, the following equilibrium condition must hold:

$$v_A\left(\tilde{B}\,|\mu\right) \geq v_M\left(\tilde{B}\,|\mu\right),$$

where $v_A\left(\tilde{B}\,|\mu\right)$, $v_M\left(\tilde{B}\,|\mu\right)$ denote, respectively, the value of being a worker and an artisan conditional on μ and the expectation that all future generations will stick to the occupational choice of their parents. Clearly, there can be no equilibrium with $\mu = 0,$ since the unskilled wage would in this case become arbitrarily large. This rules out the possibility that in equilibrium $v_A\left(\tilde{B}\,|\mu\right) < v_I\left(\tilde{B}\,|\mu\right)$. However, it is possible that $v_A\left(\tilde{B}\,|1\right) > v_M\left(\tilde{B}\,|1\right)$ (i.e., $\alpha^{\sigma}V_A\left(\tilde{B}\right) > q^{\sigma}V_M\left(\tilde{B}\right)$) In this case, $\mu = 1,$ and all dynasties choose to be unskilled workers. Otherwise there exists a unique value of $\mu \in (0,1)$ such that $\left(\alpha\mu^{\alpha-1}\right)^{\sigma}V_A\left(\tilde{B}\right) = q^{\sigma}V_M\left(\tilde{B}\right)$. Thus, if an ECW exists, it is unique.

So far, we have assumed that children will choose the same profession as their parents. The last step to establish the existence of an ECW is to show that this is indeed optimal. Consider the case in which the solution is interior ($\mu \in (0,1)$) – the other case is straighhtforward –. The analysis of section 3 establishes that artisans will invest more in their children's patience than unskilled workers. In fact, since $\tilde{B} \in [\bar{B}_A, \bar{B}_M]$ and convergence in discount factors is monotonic (see Proposition 4), the net accumulation of patience capital is non-positive for the unskilled dynasties and non-negative for the skilled dynasties (and non-zero for at least one group). In particular, let $B_{2,A}$ and $B_{2,M}$ be the discount factors of the generation that becomes adult in period 2. Then, if the parents' choice was

[27] To avoid uninteresting complications arising from corner solutions, we focus on parameter values such that the ECW value of μ is larger than one half. This implies that even if all old adults are unskilled workers, some young adult will also choose to be unskilled. This assumption is historically plausible since well more than half of the population was employed in the rural sector before the Industrial Revolution. Moreover, we ignore throughout the uninteresting case where the equilibrium features no employment in artisanry.

optimal, $B_{2,A} \leq \tilde{B} \leq B_{2,M}$, with at least one inequality being strict. Now, recall that $v_A\left(\tilde{B}\,|\mu\right) = v_M\left(\tilde{B}\,|\mu\right)$. Moreover, by Proposition 2, the steepness of the optimal income profile is non-decreasing in B. Since the income profile is steeper for artisans than for unskilled workers, this implies that

$$\begin{aligned} v_A\left(B_{2,A}\,|\mu\right) &\geq v_M\left(B_{2,A}\,|\mu\right), \\ \nu_A\left(B_{2,M}\,|\mu\right) &\leq v_M\left(B_{2,M}\,|\mu\right), \end{aligned}$$

and that children will stick to their parents' profession. The same argument applies for the following periods. □

Proof of Proposition 8: We start by guessing that $v(B,K) = K^{\sigma}V(B)$ and that V is strictly increasing and convex in B. Then, the Bellman equation, 15, can be written as

$$\begin{aligned} &v(B,K) = K^{\sigma}V(B) = \\ &\max_{l,k}\left\{K^{\sigma}\left((1-\delta+A)-k\right)^{\sigma}\ h(1-l)+B\ K^{\sigma}\left(Ak\right)^{\sigma}+z\ K^{\sigma}k^{\sigma}V\left((1-\nu)B+f(l)\right)\right\}, \end{aligned}$$

subject to (3), (13) and $K' \leq (1-\delta+A)K$, and $0 \leq l \leq 1$. Dividing both sides of the expression by K^{σ}, and substituting in the constraints, we obtain expression (16).

To establish that V is increasing, we need to show that, given the guess that V is strictly increasing, TV is a strictly increasing function. To this aim, let $\overline{B} > \underline{B}$. We need to establish that $TV(\overline{B}) > TV(\underline{B})$. Let $\underline{l}$ and $\underline{k}$ be choices that attain the maximum for $\underline{B}$. Then:

$$\begin{aligned} TV(\overline{B}) &\geq \left((1-\delta+A)-\underline{k}\right)^{\sigma}h(1-\underline{l})+\overline{B}\left(A\underline{k}\right)^{\sigma}+z\ \underline{k}^{\sigma}V((1-\nu)\overline{B}+f(\underline{l})) \\ &> \left((1-\delta+A)-\underline{k}\right)^{\sigma}h(1-\underline{l})+\underline{B}\left(A\underline{k}\right)^{\sigma}+z\ \underline{k}^{\sigma}V((1-\nu)\underline{B}+f(\underline{l})) \\ &= TV(\underline{B}) \end{aligned}$$

where the strict inequality comes from the guess that V is an increasing function.

To establish the convexity of the value function, we need to establish that, given the guess that V is convex, Tv is also a convex function. Let θ be such that $0 < \theta < 1$, let $\overline{B} > \underline{B}$, and let $B = \theta\overline{B} + (1-\theta)\underline{B}$. We now need to show that $\theta TV(\overline{B})+(1-\theta)TV(\underline{B}) \geq TV(B)$. Let l and k be choices that attain the maximum for B. Since these are feasible, but not necessarily optimal choices at $\overline{B}$ and $\underline{B}$, we have:

$$
\begin{aligned}
TV(\overline{B}) &\geq ((1-\delta+A)-k)^{\sigma}\, h(1-l) + \overline{B}\,(Ak)^{\sigma} + z\, k^{\sigma} V((1-\nu)\overline{B} + f(l)),\\
TV(\underline{B}) &\geq ((1-\delta+A)-k)^{\sigma}\, h(1-l) + \underline{B}\,(Ak)^{\sigma} + z\, k^{\sigma} V((1-\nu)\underline{B} + f(l)).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\theta TV(\overline{B}) + (1-\theta)TV(\underline{B})\\
&\quad\geq ((1-\delta+A)-k)^{\sigma}\, h(1-l) + B\,(Ak)^{\sigma} +\\
&\qquad + z\, k^{\sigma}\left[\theta V((1-\nu)\overline{B} + f(l)) + (1-\theta)V((1-\nu)\underline{B} + f(l))\right]\\
&\quad\geq ((1-\delta+A)-k)^{\sigma}\, h(1-l) + B\,(Ak)^{\sigma} + z\, k^{\sigma} V((1-\nu)B + f(l))\\
&\quad= TV(B),
\end{aligned}
$$

which is the required condition. Note that the last inequality follows from the assumed convexity of V.

Next, we want to show that k is non-decreasing in B, and that l is non-decreasing in B. To this aim, once more fix two current discount factors $\underline{B} < \overline{B}$, and let the corresponding optimal choices be $\underline{l}, \underline{k}$ and $\bar{l}, \bar{k}$. We want to show that $\underline{l} \leq \bar{l}$ and $\underline{k} \leq \bar{k}$.

Since the choices are optimizing for the $\underline{B}$ and $\overline{B}$ agents, the following inequalities must be satisfied at the optimal choices:

$$
\begin{aligned}
&\left((1-\delta+A)-\bar{k}\right)^{\sigma} h(1-\bar{l}) + \overline{B}(A\bar{k})^{\sigma} + z\bar{k}^{\sigma} V((1-\nu)\overline{B} + f(\bar{l}))\\
&\qquad \geq ((1-\delta+A)-\underline{k})^{\sigma}\, h(1-\underline{l}) + \overline{B}(A\underline{k})^{\sigma} + z\underline{k}^{\sigma} V((1-\nu)\overline{B} + f(\underline{l})),
\end{aligned} \tag{31}
$$

$$
\begin{aligned}
&\left((1-\delta+A)-\bar{k}\right)^{\sigma} h(1-\bar{l}) + \underline{B}(A\bar{k})^{\sigma} + z\bar{k}^{\sigma} V((1-\nu)\underline{B} + f(\bar{l}))\\
&\qquad \leq ((1-\delta+A)-\underline{k})^{\sigma}\, h(1-\underline{l}) + \underline{B}(A\underline{k})^{\sigma} + z\underline{k}^{\sigma} V((1-\nu)\underline{B} + f(\underline{l})),
\end{aligned} \tag{32}
$$

where the first inequality follows from optimization at $\overline{B}$ and the second from optimization at $\underline{B}$. Subtracting (32) from (31) on both sides, we get the following condition:

$$
\begin{aligned}
(\overline{B}-\underline{B})A^{\sigma}\left(\bar{k}^{\sigma} - \underline{k}^{\sigma}\right) &\geq z\underline{k}^{\sigma}\left[V((1-\nu)\overline{B} + f(\underline{l})) - V((1-\nu)\underline{B} + f(\underline{l}))\right]\\
&\quad - z\bar{k}^{\sigma}\left[V((1-\nu)\overline{B} + f(\bar{l})) - V((1-\nu)\underline{B} + f(\bar{l}))\right].
\end{aligned} \tag{33}
$$

The sign of the left-hand side is equal to the sign of $\left(\bar{k} - \underline{k}\right)$. Furthermore, because of the convexity of V and the fact that f is increasing, the right-hand side is non-negative if $\bar{k} \leq \underline{k}$ and $\bar{l} \leq \underline{l}$, and non-positive if $\bar{k} \geq \underline{k}$ and $\bar{l} \geq \underline{l}$. Taken together, these implications of (33) imply that we must have $\overline{k} \geq \underline{k}$ or $\overline{l} \geq \underline{l}$, or both, because otherwise the left-hand side is negative and the right-hand side non-negative, which would violate the inequality. To show that in fact both are true, we now proceed to establish that each implies the other.

Let us assume, first, that $\overline{l} \geq \underline{l}$. Optimization in the choice of k implies the following inequalities:

$$\begin{aligned}\left((1-\delta+A)-\bar{k}\right)^{\sigma} h(1-\bar{l}) + \overline{B}(A\overline{k})^{\sigma} + z\bar{k}^{\sigma} V((1-\nu)\overline{B} + f(\bar{l})) \\ \geq ((1-\delta+A)-\underline{k})^{\sigma} h(1-\bar{l}) + \overline{B}(A\underline{k})^{\sigma} + z\underline{k}^{\sigma} V((1-\nu)\overline{B} + f(\bar{l})),\end{aligned}$$

$$\begin{aligned}\left((1-\delta+A)-\bar{k}\right)^{\sigma} h(1-\underline{l}) + \underline{B}(A\overline{k})^{\sigma} + z\bar{k}^{\sigma} V((1-\nu)\underline{B} + f(\underline{l})) \\ \leq ((1-\delta+A)-\underline{k})^{\sigma} h(1-\underline{l}) + \underline{B}(A\underline{k})^{\sigma} + z\underline{k}^{\sigma} V((1-\nu)\underline{B} + f(\underline{l})).\end{aligned}$$

Subtracting the two equations as before, we get:

$$\begin{aligned}(\overline{B}-\underline{B})\left[(A\overline{k})^{\sigma} - (A\underline{k})^{\sigma}\right] + \left(z\bar{k}^{\sigma} - z\underline{k}^{\sigma}\right)\left(V((1-\nu)\overline{B} + f(\bar{l})) - V((1-\nu)\underline{B} + f(\underline{l}))\right) \\ \geq \left[h(1-\underline{l}) - h(1-\bar{l})\right]\left[\left((1-\delta+A)-\bar{k}\right)^{\sigma} - ((1-\delta+A)-\underline{k})^{\sigma}\right]\end{aligned}$$

Since $\overline{l} \geq \underline{l}$ and h is strictly increasing, the first term on the right-hand side is non-negative. We must therefore have that $\overline{k} \geq \underline{k}$, because otherwise the left-hand side is negative and the right-hand side non-negative, which would violate the inequality.

Conversely, suppose that $\overline{k} \geq \underline{k}$. We want to establish that this implies that $\overline{l} \geq \underline{l}$. Optimization in the choice of l implies:

$$\begin{aligned}\left((1-\delta+A)-\bar{k}\right)^{\sigma} h(1-\bar{l}) + z\bar{k}^{\sigma} V((1-\nu)\overline{B} + f(\bar{l})) \\ \geq \left((1-\delta+A)-\bar{k}\right)^{\sigma} h(1-\underline{l}) + z\bar{k}^{\sigma} V((1-\nu)\overline{B} + f(\underline{l})),\end{aligned}$$

$$\begin{aligned}((1-\delta+A)-\underline{k})^{\sigma} h(1-\bar{l}) + z\underline{k}^{\sigma} V((1-\nu)\underline{B} + f(\bar{l})) \\ \leq ((1-\delta+A)-\underline{k})^{\sigma} h(1-\underline{l}) + z\underline{k}^{\sigma} V((1-\nu)\underline{B} + f(\underline{l})).\end{aligned}$$

Combining the two conditions one more time, we get:

$$\begin{aligned}\left[\left((1-\delta+A)-\bar{k}\right)^{\sigma} - ((1-\delta+A)-\underline{k})^{\sigma}\right]\left[h(1-\bar{l}) - h(1-\underline{l})\right] \\ \geq z\bar{k}^{\sigma}\left[V((1-\nu)\overline{B} + f(\underline{l})) - V((1-\nu)\underline{B} + f(\underline{l}))\right] \\ - z\underline{k}^{\sigma}\left[V((1-\nu)\overline{B} + f(\bar{l})) - V((1-\nu)\underline{B} + f(\bar{l}))\right].\end{aligned}$$

The first term on the left-hand side is non-positive. Thus, we must have $\bar{l} \geq \underline{l}$, because otherwise the left-hand side is negative and the right hand side is non-negative, because of the convexity of V.

Therefore, we have established that both policy functions $l\left(B\right)$ and $k\left(B\right)$ are weakly increasing, and the proof is completed. $\square$

References

Ameriks, John, Andrew Caplin, and John Leahy. 2002. "Wealth Accumulation and the Propensity to Plan." NBER Working Paper No. 8920.

Banerjee, Abhijit V. and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy* 101 (2): 274–98.

Becker, Gary S. and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics* 112 (3): 729–58.

Beckett, John V. 1986. *The Aristocracy in England: 1660–1914*. New York: Blackwell.

Bisin, Alberto and Thierry Verdier. 2000. "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits." *Quarterly Journal of Economics* 115 (3): 955–88.

———. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97 (2): 298–319.

Bowen, H.V. 1989. "Investment and Empire in the Later Eighteenth Century: East India Stockholding, 1756–1791." *Economic History Review* 42 (2): 186–206.

Bowles, Samuel and Howard Gintis. 2002. "The Inheritence of Inequality." *Journal of Economic Perspectives* 16 (3): 3–30.

Browning, Martin, Lars P. Hansen, and James J. Heckman. 1999. "Micro Data and General Equilibrium Models." Chapter 8 of *Handbook of Macroeconomics Volume 1A*, edited by John B. Taylor and Michael Woodford. Amsterdam: Elsevier.

Burnette, Joyce. 2002. "How Skilled Were English Agricultural Laborers? Wage Profiles from the 1930s and 1940s." Unpublished Manuscript, Wabash College.

Cannadine, David. 1990. *Decline and Fall of the British Aristocracy*. New Haven: Yale University Press.

———. 1994. *Beyond the Country House: Aspects of Aristocracy in Modern Britain*. New Haven: Yale University Press.

Carroll, Christopher D. 2001. "Portfolios of the Rich." In *Household Portfolios: Theory and Evidence*, edited by Luigi Guiso, Michael Haliassos, and Tullio Jappelli. Cambridge: MIT Press.

Carroll, Christopher D. and Lawrence H. Summers. 1991. "Consumption Growth Parallels Income Growth: Some New Evidence." In *National Sav-*

ing and Economic Performance, edited by B. Douglas Bernheim and John B. Shaven, 305–43. Chicago: University of Chicago Press.

Caselli, Franceso and Nicola Gennaioli. 2003. "Dynastic Management." NBER Working Paper No. 9442.

Cavalcanti, Tiago V., Stephen L. Parente, and Rui Zhao. 2003. "Religion in Macroeconomics: A Quantitative Analysis of Weber's Thesis." Unpublished Manuscript, University of Illinois.

Cipolla, Carlo M. 1994. *Before the Industrial Revolution: European Society and Economy, 1000–1700*. 3rd ed. New York: Norton.

Clark, Gregory and Gillian Hamilton. 2004. "Can Institutions Shape Preferences? The Long Run Implications of Property Rights in the Malthusian Era." Unpublished Manuscript, University of California, Davis.

Coleman, James and Thomas Hoffer. 1983. *Public and Private High Schools*. New York: Basic Books.

Crouzet, François. 1985. *The First Industrialists*. Cambridge University Press.

De Nardi, Mariacristina. 2004. "Wealth Inequality and Intergenerational Links." *Review of Economic Studies* 71 (3): 743–68.

Devine, T. M. 1971. "Glasgow Colonial Merchants and Land, 1770–1815." Chapter 6 of *Land and Industry: The Landed Estate and the Industrial Revolution*, edited by J. T. Ward and R. G. Wilson. Newton Abbot: David & Charles.

Dickson, Peter G.M. 1967. *The Financial Revolution in England: A Study in the Development of Public Credit, 1688–1756*. New York: St. Martin's Press.

Doepke, Matthias. 2004. "Accounting for Fertility Decline During the Transition to Growth." *Journal of Economic Growth* 9 (3): 347–83.

Epstein, Steven A. 1991. *Wage Labor and Guilds in Medieval Europe*. Chapel Hill: University of North Carolina Press.

Farr, James R. 2000. *Artisans in Europe, 1300-1914*. Cambridge University Press.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2005. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." Forthcoming, Quarterly Journal of Economics.

Friedrichs, Christopher R. 1995. *The Early Modern City 1450-1750*. New York: Longman.

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

Galor, Oded and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review* 90 (4): 806–28.

Galor, Oded and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies* 60 (1): 35–52.

Goleman, Daniel. 1995. *Emotional Intelligence: Why It Can Matter More Than IQ*. New York: Bantam Books.

Gourinchas, Pierre-Olivier and Jonathan A. Parker. 2002. "Consumption Over the Life Cycle." *Econometrica* 70 (1): 47–89.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2003. "People's Opium? Religion and Economic Attitudes." *Journal of Monetary Economics* 50 (1): 225–82.

Haaparanta, Pertti and Mikko Puhakka. 2003. "Endogenous Time Preference with Investment and Development Traps." Unpublished Manuscript, Helsinki School of Economics.

Hansen, Gary D. and Edward C. Prescott. 2002. "Malthus to Solow." *American Economic Review* 92 (4): 1205–17.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams. 2002. "Estimating Individual Discount Rates in Denmark: A Field Experiment." *American Economic Review* 92 (5): 1606–17.

Hauk, Esther and Maria Saez-Marti. 2002. "On the Cultural Transmission of Corruption." *Journal of Economic Theory* 107 (2): 311–35.

Heckman, James J. 2000. "Policies to Foster Human Capital." *Research in Economics* 54 (1): 3–56.

Heckman, James J. and Alan B. Krueger. 2003. *Inequality in America: What Role for Human Capital Policies?* Cambridge: MIT Press.

Heckman, James J. and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91 (2): 145–9.

Heckman, James J., Jingjing Hsee, and Yona Rubinstein. 2003. "The GED is a Mixed Signal: The Effect of Cognitive Skills and Personality Skills on Human Capital and Labor Market Outcomes." Unpublished Manuscript, University of Chicago.

Jenkins, Hester and D. Caradog Jones. 1950. "Social Class of Cambridge University Graduates of the 18th and 19th Centuries." *British Journal of Sociology* 1 (2): 93–116.

Knowles, John A. and Andrew Postlewaite. 2004. "Do Children Learn to Save from their Parents?" Unpublished Manuscript, University of Pennsylvania.

Krusell, Per and Anthony A. Smith, Jr. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy* 106 (5): 867–896.

Matsuyama, Kiminori. 2003. "On the Rise and Fall of Class Societies." Unpublished Manuscript, Northwestern University.

Mischel, Walter, Yuichi Shoda, and Monica L. Rodriguez. 1989. "Delay of Gratification in Children." Chapter 6 of *Choice Over Time*, edited by George Loewenstein and Jon Elster. New York: Russell Sage Foundation.

Munro, John H. 2004. "Builders Wages in Southern England and the Southern Low Countries, 1346-1500: A Comparative Study of Trends in and Levels of Human Capital." University of Toronto Department of Economics and Institute for Policy Analysis Working Paper No. 21.

Phelps Brown, E.H. and Sheila V. Hopkins. 1957. "Wage Rates and Prices: Evidence for Population Pressure in the Sixteenth Century." *Economica* 24 (96): 289–306.

Porter, Roy. 1982. *English Society in the Eighteenth Century*. New York: Penguin Books.

Richerson, Peter J. and Robert Boyd. 2005. *Not by Genes Alone*. The University of Chicago Press.

Rubinstein, W.D. 1981. *Men of Property*. London: Croom Helm.

Saez-Marti, Maria and Yves Zenou. 2004. "Cultural Transmission and Discrimination." Unpublished Manuscript, IUI Stockholm.

Samwick, Andrew A. 1998. "Discount Rate Heterogeneity and Social Security Reform." *Journal of Development Economics* 57:117–46.

Segal, Carmit. 2004. "Misbehavior, Education, and Labor Market Outcomes." Unpublished Manuscript, Stanford University.

Shonkoff, Jack and Deborah Philips, eds. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, D.C.: National Academy Press.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Edwin Cannan. The University of Chicago Press, 1976.

Stokey, Nancy L. and Robert E. Lucas, Jr. 1989. *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press.

Taylor, J., M. McGue, and W. G. Iacono. 2000. "Sex Differences, Assortative Mating, and Cultural Transmission Effects on Adolescent Delinquency: A Twin Family Study." *Journal of Child Psychology and Psychiatry* 41 (4): 433–40.

Thompson, F. M. L., ed. 1994. *Landowners, Capitalists, and Entrepreneurs: Essays for Sir John Habakkuk*. Clarendon Press: Oxford.

von Mises, Ludwig. 1963. *Human Action: A Treatise on Economics*. 2nd ed. New Haven: Yale University Press.

Weber, Max. 1930. *The Protestant Ethic and the Spirit of Capitalism*. Translated by Talcott Parsons. London: HarperCollins.