Patience Capital, Occupational Choice, and the Spirit of Capitalism*

Matthias Doepke

UCLA and CEPR

Fabrizio Zilibotti

IIES Stockholm and CEPR

October 2005

Abstract

We model the decision problem of a parent who chooses an occupation and teaches patience to her children. The two choices are linked by a strategic complementarity: 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 socio-economic transformation that characterized the British Industrial Revolution, when a new class of entrepreneurs rising from the middle classes and imbued with an ethics emphasizing patience and savings proved most capable of profiting from new economic opportunities, and eventually surpassed the pre-industrial elite.

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 choose actions whose reward is postponed in time is an essential part of our upbringing, and many parents spend substantial effort on instilling patience into their children. This happens, for example, through deliberate delay of gratification and preaching the virtues of saving, or on a more indirect route by inducing children to practice musical instruments, encouraging them to work hard in school, or religious instruction. Parents' concern for their children's patience comes as no surprise, as this human asset turns out to be valuable: empirical evidence shows that individuals who exhibit more patience 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 an occupation and savings.

The first insight of our analysis is that endogenous accumulation of patience capital can lead to the stratification of a society into "social classes," characterized by different preferences and occupational choices. This occurs even if all individuals are initially identical. In the stratified society, attitudes (or "ethics") displayed towards investments in physical or human capital differ across social classes. Our second main finding is that, as a consequence of these differences, episodes of technological change can trigger drastic changes in the income distribution, including the "leapfrogging" of a lower class over the existing elite.

The key property of our theory that leads to these results is an association between occupations and consumption profiles. In some professions, lifetime earnings are relatively flat, while in others, in particular those requiring the acquisition of skills, high

---

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

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 steep 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 profiles and investment in patience is severed. Thus, class-based societies only emerge when financial markets are shallow, while well-functioning financial markets lead to more homogeneous societies.

In an otherwise stationary society, class differences in patience may be of little consequence other than their role in determining occupational choices. Patience becomes paramount, however, when new investment technologies allow the accumulation of wealth across generations. If such an investment opportunity arises in a stratified society, the members of the most patient class will make the greatest use of it, while individuals from particularly impatient classes may forgo it altogether. As a result, the relative economic fortune of different social classes can change drastically in a few generations.

As an application of our theory, we focus on the major transformation in the distribution of income and wealth that occurred during and after the Industrial Revolution in Britain. Before the onset of industrialization, wealth 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 economic elite, and the landed elite of old was left behind. The members of the new capitalist class mostly rose from the middle classes. They were former artisans, merchants, bankers, or pre-industrial masters, and even tenant farmers and yeomen were well represented.

We argue that differences in time preferences can provide a partial explanation for this transformation. The pre-industrial middle class had accumulated patience capital, and was consequently better prepared to exploit the new economic opportunities than the existing elite. The differences in patience, in turn, had their roots in the nature of pre-industrial professions. For centuries, artisans, craftsmen, and merchants were used to

sacrifice consumption and leisure in their youth to acquire skills. In contrast, unskilled laborers, but also rich landowners, had flat lifetime income profiles. Consequently, middle-class parents had the strongest incentive to instill patience into their children, so that the middle class became the patient class. While patience capital was a latent attribute in the pre-industrial world, it became a key asset—a "spirit of capitalism"—when new opportunities of enrichment through capital investment arose at the outset of the Industrial Revolution. The triumph of the patient bourgeoisie was the consequent outcome.

In the following section, we relate our work to the existing literature. In Section 3 we present and analyze the model. In Section 4 we consider the historical application, and Section 5 concludes. All proofs are contained in the mathematical appendix.

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 that determine how well people can focus on long-term tasks, behave in social interactions, and exert self-restraint.[2] 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 such as being future-oriented are responsible for significant differences in wages and education achievements across individuals of equal ability (as measured by IQ).[3] Similar findings emerge from Segal (2004) using individual measures of non-cognitive abilities at early school age, which include proxies for patience. Experimental evidence also supports the value of patience. In a longitudinal study which began in the 1960s at Stanford University, a group of four-year old children were offered a marshmallow,

---

[2]In everyday language, the term "patience" has a number of different meanings. In the theoretical model below, we model patience as a discount factor. Hence, in the context of this paper "patience" should be interpreted as the weight that future consequences have in today's decisions, as reflected in savings rates, for example. This utility weight can be interpreted as a reduced-form representation of a non-cognitive skill, such as the ability to exert self-control (in environments with temptation), or the ability to imagine the future vividly (which is the interpretation used by Becker and Mulligan 1997).

[3]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.

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 terms of school, marriage, and labor market performance.

There is also evidence that non-cognitive skills are affected by nurture 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.[4] These studies also underline the importance of family transmission for this particular form of human capital accumulation, including the acquisition of patience. Similar conclusions are reached by a number of studies in child development psychology (see, for example, 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 their children's propensity to save and invest. This is consistent with the evidence provided by Knowles and Postlewaite (2004), who show 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. Moreover, the correlation is stronger between mothers and children than between fathers and children. Since, on average, mothers are more involved in children's upbringing than fathers, this observation suggests that the cause of this correlation is not merely genetic, but that some form of cultural transmission or conscious investment is taking place.

Some recent studies cast light on the socio-economic characteristics of patient individuals. 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

---

[4]On the other hand, the programs were less successful in raising cognitive skills as measured by IQ test scores. Compared to the impact on non-cognitive skills, the extent to which parental effort can affect cognitive skills and social attitudes is controversial, and is the subject of a long-standing debate (see, e.g., Richerson and Boyd 2005 and Bowles and Gintis 2002).

our theory, which posits that agents who have steeper income profiles have stronger incentives to invest in patience.[5]

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 preferences 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.[6]

The importance of cultural and religious aspects in determining which groups 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.[7]

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

---

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

[6]Also closely related is Mulligan (1997), where parents make a choice of how altruistic they are to their own children. The macroeconomic consequences of inherited (as opposed to chosen) preferences have been examined by de la Croix and Michel (1999, 2001) and Artige, Camacho, and de la Croix (2004). In the latter paper, inherited consumption habits can lead to the downfall of a temporarily wealthy country or region.

[7]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. However, Cavalcanti, Parente, and Zhao (2003) question whether differences in preferences arising from religious affiliation alone can explain large differences in the timing and extent of the Industrial Revolution across countries.

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 (2006) applies similar ideas to the rise and fall of class societies. A common feature of this literature is that the rich (who are least constrained by credit market imperfections) generally do best, and would be the first beneficiaries from new investment opportunities. These theories cannot account for the fact that a new class of entrepreneurs with a middle-class background leapfrogged over the landed pre-industrial elite, at a time when wealth inequality was quite extreme and financial markets shallow by modern standards. In contrast, our theory predicts that under absent financial markets the middle class becomes the patient class, which ultimately results in economic dominance.[8]

A growing literature, both theoretical and empirical, has demonstrated the importance of heterogeneity in preferences, in particular in discount factors, for understanding macroeconomic puzzles in modern economies. In particular, heterogeneous time preference has proved to be important to reconcile the quantitative predictions of calibrated incomplete-market models with the empirical extent of wealth heterogeneity (see Krusell and Smith, Jr. 1998 and De Nardi 2004).[9] 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 across agents accumulate and persist.

The application of our model to the Industrial Revolution relates to a series of recent papers proposing unified theories of the transition from stagnation to growth (see Galor and Weil 2000, Hansen and Prescott 2002, and Doepke 2004). Within this literature, a few authors emphasize the role of preference formation for long-run development, but rely on selection instead of conscious investment as the mechanism (see Galor 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.

---

[8]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.

[9]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 hard to identify empirically.

3 A Theory of Occupational Choice and Endogenous 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. We show that a complementarity between patience capital and occupational choices leads to some unusual technical features; in particular, the agents' value functions are convex in patience. Despite this complication, we can characterize the problem through a recursive formulation with well-defined value and policy functions. In fact, the convexity of the value function turns out to be a surprisingly useful feature to characterize the equilibrium. In Section 3.4, 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 her parent decides to spend on increasing her patience.

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:

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

where $h(1) = 1$, implying that $u(c)$ is the felicity of an agent who does not invest in her
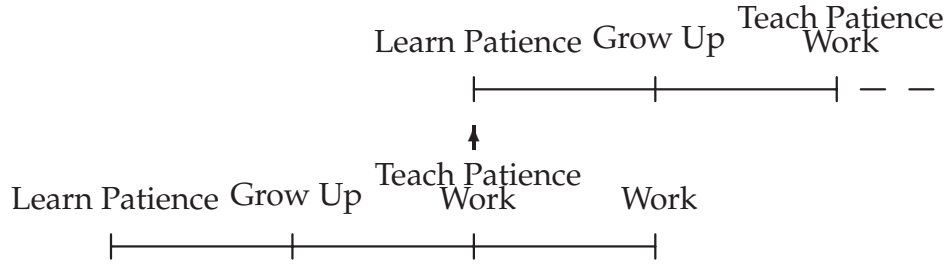
Figure 1: The Timing of Investment in Patience

child's patience.[10] 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).[11]

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

$$u(c_1)h(1-l) + B\, u(c_2) + z\, U(B'(l, B)).$$

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

---

[10]The multiplicative separability between the utilities derived from consumption and leisure is not essential for our theoretical results (the additively separable case is briefly discussed below). However, it is important that the investment in patience requires parents' time. If parents could pay professional educators to transmit patience, and their services were a perfect substitute of parents' time, then rich parents would generally invest more in their children's patience. We regard the assumption that the transmission of patience requires a direct involvement of parents as reasonable, especially since this occurs to a large extent in early childhood, and parents are often more effective role models for children than strangers. An alternative assumption leading to similar results is that the investment in patience requires a costly time investment by the child.

[11]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.

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 "dynastic" interpretation, where the head of the dynasty makes decisions for all subsequent generations.[12]

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. 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. 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:[13]

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

---

[12]It could be argued that investments in patience also affect altruism (e.g., we could have B^2 where we have z). Such a model leads to qualitatively similar results, but the change would come at a loss of analytical tractability.

[13]The only non-standard assumption 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. 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.

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 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 in the mathematical appendix.

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

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 (as depicted by the dotted line in the upper panel of Figure 2). 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 $1 - \nu$. Moreover, given the constant income profile, it is 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. Generalizing from this observation, 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 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 of 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.
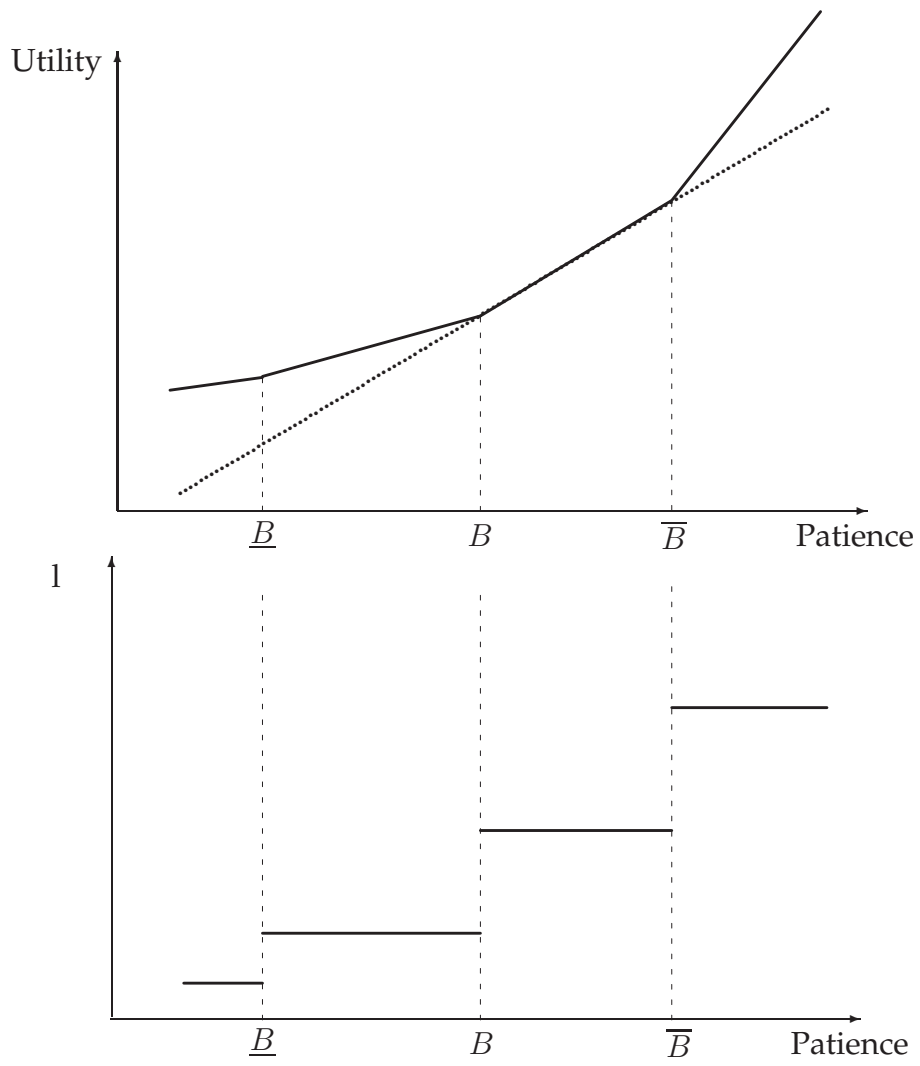
Figure 2: Convexity of the Value Function

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. 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 the decisions on patience more sharply. 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-generation's 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 $B_{t+1} = (1 - \nu)B_t + f(\bar{l})$, patience converges to a steady-state given by $\bar{B} = f(\bar{l})/\nu$.[14] Substituting back for $f(\bar{l})$, we can see that patience converges to this steady state at a

---

[14]If we had assumed $w\left(c, l\right) = u\left(c\right) + h\left(1 - l\right)$, then the first-order condition for l_0 would be:

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

Consequently, (6) would be replaced by

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

i.e., $\bar{l}$ would depend on y_2 but not on y_1. However, the analysis would lead to the same results and interpretation since, by Assumption 2, the ranking of occupations by y_2 is identical to the ranking of occupations by the steepness of earning profile.

14

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 differ-
ent 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 fi-
nancial markets to smooth consumption is limited. What determines the incentive to
invest in patience is not the income profile *per se*, but the lifetime profile of period-by-
period utilities (felicity). A steep income profile directly translates into a steep utility
profile only if financial markets are absent or incomplete.

We now want to make this point more precise by considering the opposite extreme in
terms of assumptions on financial markets. Namely, we allow unrestricted borrowing
and lending *within each cohort* at a fixed rate of return R.[15] We will see that in this finan-
cial 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 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, regard-
less of patience. The proposition shows that at least some degree of financial market

---

[15]The possibility of wealth transmission *across generations* is discussed in Section 4.

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.[16]

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 perseverance, 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.

3.4 General Equilibrium with Two Technologies

The results of the previous section demonstrate that there exists a dynamic complementarity between the choices of investment in patience and occupation. Dynasties that start out patient choose professions that are characterized by a steep income profile, which, in turn, further increases 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, up to this point the level of income derived in each profession has been taken as exogenous. In general equilibrium, price adjustments may strengthen or counteract the tendency for different dynasties to sort into professions with different income profiles, and hence to accumulate patience at different rates. Thus, in order to determine whether dynasties diverge or converge, we need to endogenize the incomes derived in different professions.

---

[16]In addition, given that it is the steepness of lifetime utility profiles that really matters, one can write down simple extensions of the basic framework where different occupations can imply different lifetime utility profiles even with perfect financial markets. This happens, for example, when leisure enters utility in all periods, and different occupations require different lifetime labor supply profiles that cannot be smoothed out.

The main point of this section is to demonstrate that even in a world where agents have initially identical preferences, general equilibrium forces can adjust returns such that agents choose different professions. Given a menu of 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.

We consider a version of the model where people can choose between two occupations. For concreteness, we parameterize preferences over consumption by a utility function featuring constant relative risk aversion (CRRA), i.e., $u(c) = c^\sigma$, where $0 < \sigma \le 1$.[17] 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 label occupations and technologies in a way that hints at the application that will be discussed below. The two modes of production are called 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:

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

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

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

where q is a productivity parameter. The total amount of land is fixed at $Z = 1$. Land is not traded and is owned by a fixed measure of dynasties, each of whom owns an equal share of land, 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. There

---

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

is no occupational mobility between landowners and the other classes.[18]

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.

We now define an equilibrium with constant wages across dynasties. We focus on constant-wage equilibria because in this case the analysis of the preceding section (where the decision problem was over a set of exogenous 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 unskilled workers is flat. In contrast, artisans have an increasing earnings 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. The distribution of discount factors converges to a steady state where all workers (and landowners) have a discount factor $\bar{B}_A$, whereas artisans have a discount factor $\bar{B}_M > \bar{B}_A$.*

---

[18]As landowners make no occupational choice nor do their choices have any general-equilibrium implications, they are inessential for the theoretical analysis of this section. However, we mention them here, since they have a role in the historical application of Section 4.

Given Proposition 6, if an ECW exists and has a positive proportion of both workers and artisans, then it must feature diverging patience in the population, with each group converging to a profession-specific discount factor. However, we still need to establish conditions for the existence of an ECW. As discussed above, we restrict attention to economies where, initially, agents have identical preferences, i.e., they are equally patient. In this case, a unique ECW necessarily exists if $\bar{B}_A \leq \tilde{B} \leq \bar{B}_M$, where $\tilde{B}$ denotes the common initial level of patience in the population. This restriction encompasses the intuitive case of an economy where, before time zero, only the agricultural activity was pursued (say, because the productivity of artisanry q was too low), and patience had reached the steady-state level $\tilde{B} = \bar{B}_A$. The process of occupational sorting can then be triggered by an (unexpected) rise in the productivity level q.

In an ECW, employment and productivity per efficiency unit of labor are constant in each sector. In particular, if we denote the proportion of landless adults employed in agriculture by $\mu \in [0,1]$, the competitive wages per efficiency unit 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 her life, whereas an unskilled worker earns a flat wage of $\alpha\mu^{\alpha-1}$.[19] Establishing the existence of an ECW 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.*

---

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

Which of the two possible outcomes is obtained depends on the productivity of artisanry q. If this productivity is sufficiently high, there will be a positive number of artisans in equilibrium (employment in agriculture is always positive, since the agricultural production function satisfies the appropriate Inada condition). In this case, preferences diverge over time: artisans become the patient class, whereas workers become impatient. The landowners' incentives for investing in patience are the same as those of the agricultural workers, because both groups face a flat income profile (a constant labor income for workers, and a flat stream of rental income for the landowners). Thus, the landowners become impatient as well, regardless of their initial level of patience. A parameterized example illustrating this process will be presented in the next section.

4   An Application: The "Spirit of Capitalism"

In this section we discuss an application of the theory analyzed in the previous sections to the formation of a modern spirit of capitalism. As emphasized by Max Weber (1930), the emergence of a new socio-economic elite endowed with a future-oriented culture and a high propensity to accumulate wealth was a salient feature of the Industrial Revolution. We first show, with the aid of a parameterized version of the model, that our theory is consistent with the emergence of a "capitalist spirit" and subsequent large shifts in the distribution of income and wealth. We also discuss historical evidence about the British Industrial Revolution that supports the main predictions and some auxiliary implications of the theory.

4.1   The Pre-Industrial Economy

We start by presenting a parameterized version of the economy of Section 3.4. 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. This particular functional form 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 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 model has not been calibrated to closely match particular observations; we merely use the parameterized model to illustrate some of the qualitative features of our theory.

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 factors 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 that all dynasties 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 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. Nevertheless, 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
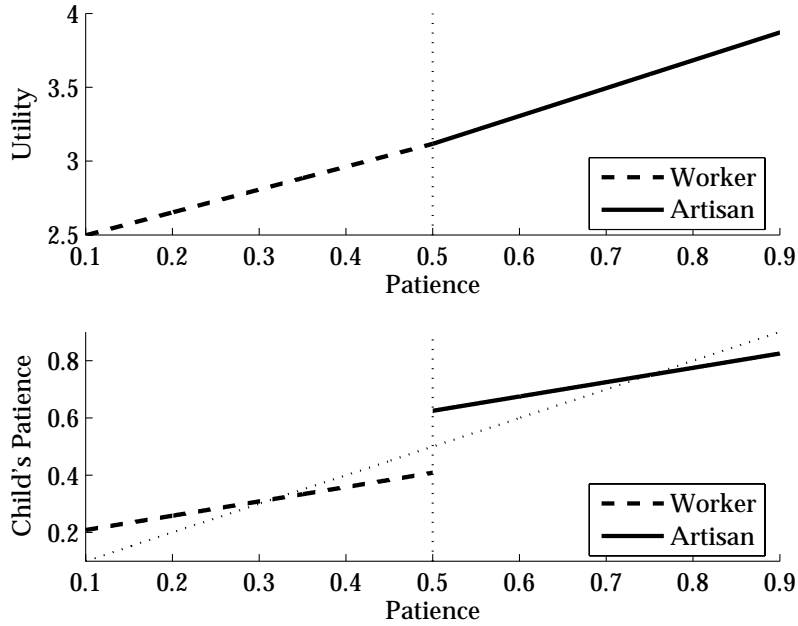
Figure 3: Value Function and Law of Motion for Patience for Lower Classes

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 over the alternative. As shown in Proposition 4, conditional on being in a dynasty with a fixed occupation the law of motion is linear. 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. In particular, they receive the same amount of rent every period, and therefore
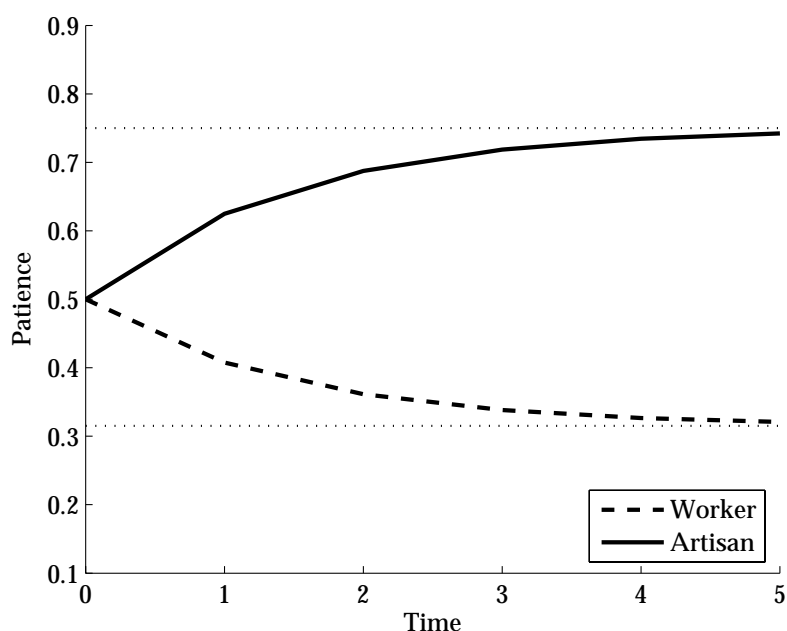
22

Figure 4: Patience Over Time

have a flat income profile, just like the 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. Hence, the landowners end up impatient, regardless of their initial preferences.[20]

Thus, in our two-technology 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. In this "pre-industrial" economy, class differences are only important to the extent that they determine the occupational choice of individuals. Patience becomes of central importance, however, when technological change gives rise to new investment opportunities. In the following section, we examine the fate of the different classes in

---

[20]If we modeled an initial stage of land acquisition, presumably the initial landowners (as the winners in this process) would be unusually patient and hard-working. Over the generations, however, the positive effect of initial patience would wear off, and the flat endowment income generated by land ownership would remove the incentive to continue investing in patience.

our economy after the arrival of a new capital-accumulation technology.

4.2   From Artisan to Capitalist

In this section, we introduce a physical investment technology into the model. The new technology suddenly becomes available in a pre-industrial society where the classes have already acquired different levels of patience. The result is unsurprising in the light of standard economic theory: the patient classes, i.e., the artisans, are the first to take advantage of the new opportunity—they possess the "spirit of capitalism." The artisans leapfrog over the landowning class, and replace them as the economic elite.

After the introduction of the new technology, each dynasty that can accumulate two assets: physical capital and patience. The rate of return to capital is constant and denoted by A (i.e., we us an "AK" technology). Capital depreciates at the rate δ. Young adults inherit capital from their parents, and decide how much of their first-period income to consume and how much to invest. Investments in physical capital are assumed to be irreversible: agents can consume the output of the investment technology (as well as their labor income), but the capital stock itself 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 her child.[21] We continue to assume that, because of capital market imperfections, agents cannot borrow.

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

$$c_1 + K' = (1 - \delta + A)K + y_{1i}, \qquad (11)$$
$$c_2 = AK' + y_{2i}, \qquad (12)$$
$$K' \geq (1 - \delta) K. \qquad (13)$$

---

[21]Dynastic enterprises were common in the early days of the Industrial Revolution. Caselli and Gennaioli (2003) argue that this was due to the underdevelopment of financial markets: it was unprofitable for parents to liquidate their business instead of leaving it to the children. In our model, the irreversibility constraint implies that differences in investment across families lead to different initial assets of the next generation. Under reversible investment, similar results could be obtained if the altruism parameter z (the intergenerational discount factor) was an increasing function of patience B (the intragenerational discount factor).

In the budget constraint (11) for the first adult period, 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 output of the investment technology AK. Because of the irreversibility constraint (13), consumption cannot exceed the sum of current output and labor income: $c_1 \leq AK + y_{1i}$. In the second-period budget constraint (12), the agent earns labor income y_{1i} and produces AK'. Since the capital stock cannot be liquidated, the agent bequeaths the remaining capital stock $K'' = (1 - \delta)K'$ to her child. [22]

We maintain our earlier assumption of constant-elasticity utility in consumption. The recursive representation of the decision problem of a young adult with discount factor B and inherited capital stock K is given by the following Bellman equation:

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

where the maximization is subject to (3), (11), (12), (13), $c_1 \geq 0$ and $0 \leq l \leq 1$. Notice that the agricultural and artisan technology continue to exist alongside the capitalist technology. Thus, adults can, and will, continue to work in one of the existing professions even while they are investing in the new technology. This 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.[23] 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.

Now assume that the economy is parameterized as in Table 1, and that each class starts out with its steady-state discount factor (B_W for workers and landowners and B_A for artisans). Figure 5 shows (for the landless classes) the policy function for the decision on investment (the child's capital $K'' = (1 - \delta)K'$ on the vertical axis) as a function of the

---

[22]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.

[23]The model has to be solved numerically. An analytical characterization can be attained for the special case in which $y_1 = y_2 = 0$. In the working paper version, we prove that in this case the value function is increasing and strictly convex in B and increasing and homogenous of degree σ in K. The physical investment rate as well investment in patience capital is increasing in B, but independent of K.
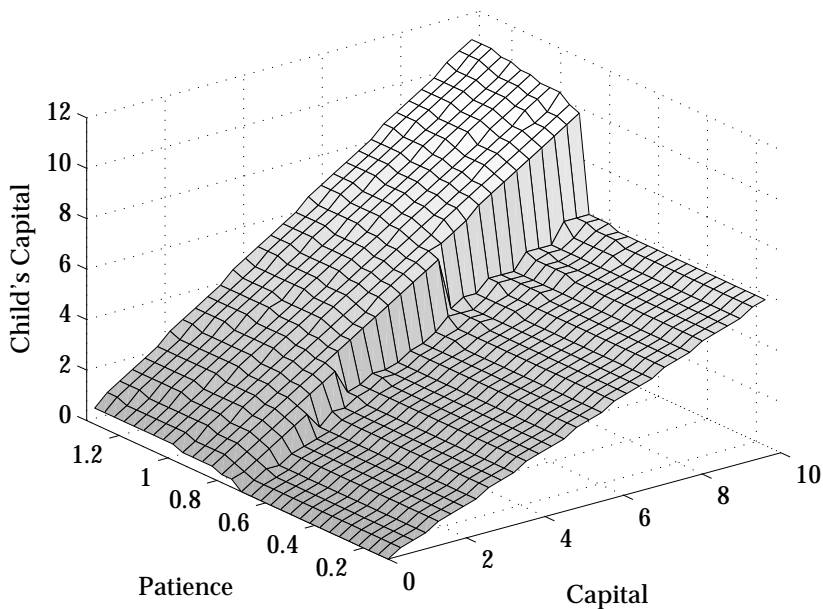
Figure 5: The Policy Function for Capital Investment

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 the policy function in 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 in the first period, so that she makes no additional investment and only carries 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 a complementarity between savings and patience. The low plateau corresponds to savings decisions that imply declining wealth from the current generation to the next. Consequently, current utility is high relative to future utility. The high plateau depicts savings choices that imply an increasing profile of wealth from the current to the next generation. Parents therefore experience 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. The role of wealth profiles across generations (which determine intergenerational utility profiles) is analogous to the role of income profiles in the occupational choice model (where only intragenerational utility profiles were at stake). Phys-
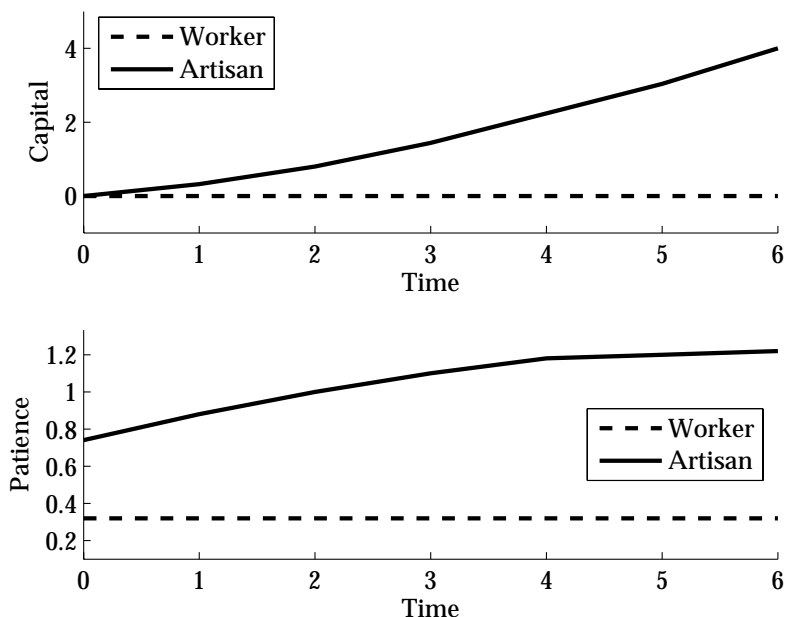
26

Figure 6: Patience and Capital Over Time

ical investment leads to an increasing intergenerational utility profile, and is therefore complementary to investment in patience.

Apart from investment in physical capital, the original decisions of choosing an occupation and investing in patience are still present in the model. As before, 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 such a 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 pre-industrial 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

27

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.[24]

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 do that due to the borrowing constraint. They 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.

An interesting feature of the model is that the same pattern of catch-up and overtaking can also be generated in an environment where the investment technology is available from the outset, instead of being introduced later on. If all dynasties start out sufficiently impatient, the investment technology is initially not used. Some dynasties, however, sort into the artisans professions, and start to accumulate patience. After a few generations, the patience of the artisan middle classes reaches a critical level, at which the artisans start to use the investment technology and turn into capitalists. In this version of the model, it is not the surprise appearance of a new technology, but the endogenous accumulation of patience capital that triggers the Industrial Revolution. Arguably, this sequence of events is closer in spirit to Weber's original hypothesis.

As the preceding discussion 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 the elasticity of intertemporal substitution were extremely low, 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, 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.

To examine our hypothesis in more detail, we now turn to the historical circumstances

---

[24]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.

that accompanied the changing economic fortunes of different social classes during the British Industrial Revolution.

4.3 The Historical Context

It has long been part of the conventional wisdom on the Industrial Revolution that the majority of the new entrepreneurs of the early nineteenth century came from a modest background and had a high propensity to save and accumulate wealth.[25] 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" (p. 68). The only class that was significantly over-represented among the industrialists was the middle class.[26] The minor involvement of the upper class 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 were ultimately based on capital investment.

The new class of industrialists progressively replaced the landed elite as the economically dominant group in society, as reflected, with some lag, in changes in the wealth distribution. In the first half of the nineteenth century, large fortunes were still by and large 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

---

[25]For instance, von Mises (1963) writes: "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" (p. 622).

[26]In the sample analyzed in this study, only 2.3 percent of the industrialists came from peerage and gentry (see Crouzet's Table 5). In contrast, as many as 85 percent of the new industrialists had a middle-class background. This class includes from bankers and rich merchants at the upper end to small artisans and tenant farmers at the lower end. 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 and artisans, cultivators of various kind" (Crouzet 1985, p. 127). The contribution of the working class (about 70 percent of the population) was moderate; no more than 12 percent of the industrialists came from this class. 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 under-representation 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.

percent were wealthy landowners. Only a century later, 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 (see Tables 3.2, 3.3, and 3.4. in Rubinstein 1981). 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.

In the pre-industrial world there were large differences in lifetime income profiles between members of different social classes, which is a key factor in our theory. Artisans and craftsmen, the typical professions of the pre-industrial middle classes, were required to make large human capital investments, and consequently had steep lifetime income profiles. In most of Europe, an artisan's career advanced through three stages: apprenticeship, journeymanship, and mastership.[27] Apprenticeship would on average take 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 would last for a minimum of 3–4 years (Friedrichs 1995). Savings and frugality were 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 an expensive process.[28] 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. These accounts suggest that the life of an artisan was investment-intensive, and the consumption profile very steep (see Phelps Brown and Hopkins 1957, Munro 2004 and Farr 2000 for additional evidence).

In contrast, the wage profile of agricultural workers and landowners was relatively flat. Burnette (2002) documents that the age-earning profile of a farm laborer in England in the early nineteenth century was essentially flat between age 20 and 60. As far as the landed gentry is concerned, the available evidence suggests that their income and con-

---

[27]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).

[28]The applicants owed the payment of a series of application fees, the completion of a masterpiece according to the guild regulation 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.

sumption 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 well 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 landowners 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.[29]

4.4 Patience Capital versus Alternative Explanations

The thesis of this paper is that the differences in lifetime consumption profiles were at the root of the development of cultural differences between social classes. 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. Adam Smith (1776), for instance, dwells on the sharp cultural differences between landowning aristocrats and businessmen. The former are portrayed as obsessed with conspicuous consumption, unwilling to save, and unable to consider money as something to be profitably invested.[30] In con-

---

[29]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 aristocrats 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.

[30]"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." (Adam Smith 1776, p. 432)

trast, new industrialists are described as frugal and thrifty. Crouzet (1985) cites accounts of the time relating that 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).

To these contemporary observers, it seemed entirely natural to think of members of different classes as essentially distinct beings whose behavior was governed by class-specific rules. While these observations are consistent with our hypothesis, they do not necessarily single out patience as the only explanation: differences in other dimensions of preferences (such as work ethics or risk aversion) could also affect an individual's propensity of becoming an entrepreneur. Clearly, the one-dimensional representation of preference heterogeneity in our theory is a simplification. Nevertheless, time preference is a key dimension for intertemporal decisions such as starting or investing in a new enterprise. Moreover, the specific hypothesis of class differences in patience leads to a number of auxiliary implications (other than the selection of the early industrialists) that can be tested. In particular, a low time discount factor should make rich landowners unwilling to invest not only in new industrial enterprises, but in financial assets in general. To the contrary, we would expect them to borrow from the middle class, in order to finance current consumption. 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 urban middle classes, while the contribution of the landed classes as financiers of public loans was insignificant (see Dickson 1967, p. 302). The same pattern can be found when we consider investments in public companies. Bowen (1989) documents that most stockholders of the East India Company between 1756 and 1791 were "clergymen, bankers, military and naval personnel, officials, brokers, merchants large and small, and retailers," whereas "beyond doubt there was no large-scale investment in the [East India] Company by the landed interest or aristocracy." (p. 195). Thus, even before the Industrial Revolution the pre-industrial elite played a surprisingly minor role in financing government borrowing and private enterprise, despite the fact that this group was far wealthier than the middle class. This behavior stands in marked contrast to the wealth elites in modern industrial countries, which generally own a disproportionate share of most types of assets, including government debt and public stock (see Carroll 2001 for

evidence on the United States).[31]

Rather than investing the rents from their estates in interest-bearing assets, rich landowners used the land as collateral to borrow. Such borrowing had been common in Britain long before the Industrial Revolution, and expanded substantially since the Glorious Revolution of 1689. Beckett (1986) reports that by the mid-eighteenth century "many families already had an accumulation [of debt] several generations old" (p. 300). Most 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, p. 316. See also Devine 1971 and Porter 1982.) Aristocratic indebtedness grew severely during the nineteenth century, and a 1847 writer claimed that "between half and two-thirds of English land was encumbered (i.e. mortgaged)" (Beckett 1986, p. 315). Cannadine (1994) summarizes the situation as follows: "Whatever might have been the financial state of individual 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).[32]

One might wonder why, if differences in discount factors were important, the impatient aristocracy did not sell the land to the more patient middle class already prior to the Industrial Revolution.[33] To a large extent, the answer is that this was actually very hard to do. The land market in Britain was subject to pervasive restrictions: most large estates were "entailed," meaning that they could neither be split nor sold by the owner.[34]

---

[31]A possible caveat is that if investments in agricultural estates carried a higher return than financial assets, the upper classes may have merely held a different (and possibly more profitable) portfolio than the middle classes, even with the same preferences on the consumption-savings margin. However, there is little evidence of widespread active involvement of landowners in agricultural investment activity. Thompson (1994) documents that ever since 1700 they progressively withdrew from day-to-day involvement with farming (in both France and Britain) and remained content with relying on safe flat rents. Investments and technical innovation in agriculture which played a key role in the British Industrial Revolution were carried out almost entirely by tenant farmers.

[32]While some of this debt was raised for investment in non-agricultural ventures, according again to Cannadine (1994), "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).

[33]Notice that our theory does not posit that landowners were always impatient; in fact, the first aristocrats in a dynasty, who initially acquired title and estate, may have plausibly been particularly patient.

[34]Through the institution of entail, an aristocratic landowner could effectively prevent his descendents

Mortgaging their land to merchants and banks was therefore the only way in which, *de-facto*, landowners could run down their assets. Eventually, after statutory reforms and changes in the common law made it progressively easier to sell land, many families overburdened by debt did sell off parts or all of their estates. Cannadine (1990) describes the massive sale of land occurred in the first part of the twentieth century as follows: "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). While other factors (taxation, decline of land rents) contributed to this final outcome, a clear thread links the chronic indebtedness of the landed aristocracy over centuries with its eventual decline and inability to hold on to the land.

Clearly, factors other than endogenous preferences also contributed to the success of the middle class during the Industrial Revolution. A particularly important complementary explanation is that urban workers possessed skills that were essential for industrial activities, while the landowners did not. However, the evidence suggest that differences in skills cannot be the only explanation. A significant share of the new industrialists had not previously been involved in any form of manufacturing. For instance, as many as 22 percent of the industrialists' fathers were yeomen and farmers, groups with no experience in industrial activity (Crouzet 1985). Moreover, there is evidence of a substantial mobility across industrial sectors. The boom of the textile industry, for example, 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 industrialization a number of key 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 this was not always a binding constraint. Even 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 not held back by their obligations to an existing estate;

---

from selling part or all of the estate.

34

	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

nevertheless, they did not enter business in any larger numbers than their landowning fathers. For instance, consider Table 2, which reports the occupational choices 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 occupational choices (other than landowning) give us some idea 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. This evidence is corroborated by the study of Crouzet (1985), who documents that few of the new industrialists' fathers were landowners (see footnote 26).

5 Conclusions

In this paper, we have developed a theory of the endogenous accumulation of patience in a dynastic model. We have shown that an interaction between accumulation of patience and occupational choice can lead to the stratification of society in terms of class-specific preferences.

The theory provides a new perspective of the impact of financial market frictions on economic development. By increasing the possibilities for agents to smooth consump-

tion, 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. An application of the theory to the Industrial Revolution period suggests that, in pre-industrial times, members of different classes really were distinct from each other on a fundamental level (in addition to the obvious differences in wealth and power), just as the contemporary observers seemed to believe.

Preference transmission of the type analyzed in our theory may be important for other aspects of individual preferences (such work ethics, risk aversion, and attitudes towards innovation) which might have been equally salient for the emergence of a "spirit of capitalism." Extending the analysis to these additional traits of preferences promises to provide further insights. For instance, financial development would also tend to equalize the attitude towards risk across dynasties engaged in different professions. However, it may induce parents, *ceteris paribus*, to encourage risk-taking behavior in their children, contrary to the case analyzed in this paper, where financial development reduces the return and incentive to invest in patience. The analysis of these and other aspects of endogenous preferences is left to future research.

A Mathematical Appendix

A.1   The Sequential Formulation of the Decision Problem

The sequential formulation of the decision problem is given by:

$$v^\star(B_0) = \max\left\{\sum_{t=0}^{\infty} z^t \left[u(y_{1,i_t})h(1-l_t) + B_t u(y_{2,i_t})\right]\right\}, \tag{15}$$

subject to $i_t \in I$, $l_t \in (0,1)$ and $B_{t+1} = (1-\nu)B_t + f(l_t)$.

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:

$$Tv(B) = \sup_{i \in I, 0 \le l \le 1} \{u(y_{1,i})h(1 - l) + Bu(y_{2,i}) + zv((1 - \nu)B + f(l))\}. \tag{16}$$

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 (16) 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:

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

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 $\theta Tv(\overline{B}) + (1 - \theta)Tv(\underline{B}) \ge Tv(B)$. 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:

$$Tv(\overline{B}) \ge u(y_1)h(1 - l) + \overline{B}u(y_2) + zv((1 - \nu)\overline{B} + f(l)),$$
$$Tv(\underline{B}) \ge u(y_1)h(1 - l) + \underline{B}u(y_2) + zv((1 - \nu)\underline{B} + f(l)).$$

Working towards the desired condition, we therefore have:

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

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

37

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. Q.E.D.

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:

$$u(\overline{y}_1)h(1-\overline{l}) + \overline{B}u(\overline{y}_2) + zv((1-\nu)\overline{B} + f(\overline{l}))$$
$$\geq u(\underline{y}_1)h(1-\underline{l}) + \overline{B}u(\underline{y}_2) + zv((1-\nu)\overline{B} + f(\underline{l})), \quad (17)$$

$$u(\overline{y}_1)h(1-\overline{l}) + \underline{B}u(\overline{y}_2) + zv((1-\nu)\underline{B} + f(\overline{l}))$$
$$\leq u(\underline{y}_1)h(1-\underline{l}) + \underline{B}u(\underline{y}_2) + zv((1-\nu)\underline{B} + f(\underline{l})), \quad (18)$$

where the first inequality follows from optimization at $\overline{B}$ and the second from optimization at $\underline{B}$. Subtracting (18) from (17) on both sides, we get the following condition:

$$(\overline{B} - \underline{B})\left[u(\overline{y}_2) - u(\underline{y}_2)\right] \geq z\left[v((1-\nu)\overline{B} + f(\underline{l})) - v((1-\nu)\underline{B} + f(\underline{l}))\right]$$
$$- z\left[v((1-\nu)\overline{B} + f(\overline{l})) - v((1-\nu)\underline{B} + f(\overline{l}))\right]. \quad (19)$$

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 $\overline{l} \leq \underline{l}$, and non-positive if $\overline{l} \geq \underline{l}$. Taken these implications together, (19) implies that we must have $\overline{y}_2 \geq \underline{y}_2$ or $\overline{l} \geq \underline{l}$, 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 $\overline{l} \geq \underline{l}$. Optimization in the choice of the income profile implies the following inequalities:

$$u(\overline{y}_1)h(1-\overline{l}) + \overline{B}u(\overline{y}_2) \geq u(\underline{y}_1)h(1-\overline{l}) + \overline{B}u(\underline{y}_2), \quad (20)$$
$$u(\overline{y}_1)h(1-\underline{l}) + \underline{B}u(\overline{y}_2) \leq u(\underline{y}_1)h(1-\underline{l}) + \underline{B}u(\underline{y}_2). \quad (21)$$

Subtracting the two equations as before, we get:

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

38

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 $\bar{y}_2 \geq \underline{y}_2$ and, consequently, $\bar{y}_1 \leq \underline{y}_1$ since otherwise the left-hand side is negative and the right-hand side non-negative.

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

$$u(\bar{y}_1)h(1-\bar{l}) + zv((1-\nu)\overline{B} + f(\bar{l})) \geq u(\bar{y}_1)h(1-\underline{l}) + zv((1-\nu)\overline{B} + f(\underline{l})), \qquad (22)$$

$$u(\underline{y}_1)h(1-\bar{l}) + zv((1-\nu)\underline{B} + f(\bar{l})) \leq u(\underline{y}_1)h(1-\underline{l}) + zv((1-\nu)\underline{B} + f(\underline{l})). \qquad (23)$$

Combining the two conditions one more time, we obtain:

$$\left[u(\bar{y}_1) - u(\underline{y}_1) \right] \left[h(1-\bar{l}) - h(1-\underline{l}) \right]$$
$$\geq z \left[v((1-\nu)\overline{B} + f(\underline{l})) - v((1-\nu)\underline{B} + f(\underline{l})) \right]$$
$$- z \left[v((1-\nu)\overline{B} + f(\bar{l})) - v((1-\nu)\underline{B} + f(\bar{l})) \right]. \qquad (24)$$

Here, the first term on the left-hand side is non-positive. If $\bar{y}_1 < \underline{y}_1$, 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, $\bar{y}_1 = \underline{y}_1$, 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 (24) are zero, regardless of $\underline{l}$ and $\bar{l}$. To satisfy (24), in this case both (22) and (23) 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. Q.E.D.

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 (15) 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:

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

$$
\left. + \sum_{t=1}^{\infty} z^t \left[u(y_{1,i_t})h(1 - l_t) + \left((1 - \nu)^t B + \sum_{s=0}^{t} (1 - \nu)^{t-s-1} f(l_s) \right) u(y_{2,i_t}) \right] \right\}. \quad (25)
$$

For given current and future income profiles, (25) 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 (25) with income profiles i_t and investment in patience l_t fixed at their optimal (and constant) values. (25) 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. Q.E.D.

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.   Q.E.D.

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:

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

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:

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

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.   Q.E.D.

Proof of Proposition 6: We start by establishing that in an ECW parents and their children necessarily choose the same profession (occupational segregation). To this aim, first note that an equal number of parents and children in each cohort must be working in each of the two professions, since the relative supply of labor in the two sectors is constant over time in an ECW. 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 $B_A > B_W$.   Q.E.D.

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

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

where

$$V_A(\tilde{B}) = \max_{B'}\{h(1-(B'-(1-\nu)\tilde{B})) + \tilde{B} + zV_A(B')\} \text{ and} \qquad (26)$$

$$V_M(\tilde{B}) = \max_{B'}\{h(1-(B'-(1-\nu)\tilde{B})) + \tilde{B}\gamma^\sigma + zV_M(B')\} \qquad (27)$$

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.[35] Since, initially, all landless agents

---

[35]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

42

have the same preferences, the following equilibrium condition must hold:

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

where $v_A(\tilde{B}|\mu.)$, $v_M(\tilde{B}|\mu.)$ 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 $v_A(\tilde{B}|\mu.) < v_I(\tilde{B}|\mu.)$. However, it is possible that $v_A(\tilde{B}|1.) > v_M(\tilde{B}|1.)$ (i.e., $\alpha^\sigma V_A(\tilde{B}) > q^\sigma V_M(\tilde{B})$) 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 $(\alpha \mu^{\alpha-1})^\sigma V_A(\tilde{B}) = q^\sigma V_M(\tilde{B})$. 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 straightforward. 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 optimal, $B_{2,A} \leq \tilde{B} \leq B_{2,M}$, with at least one inequality being strict. Now, recall that $v_A(\tilde{B}|\mu.) = v_M(\tilde{B}|\mu.)$. 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

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

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

---

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.

References

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

Artige, Lionel, Carmen Camacho, and David de la Croix. 2004. "Wealth Breeds Decline: Reversals of Leadership and Consumption Habits." *Journal of Economic Growth* 9 (4): 423–49.

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 Saving 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.

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 la Croix, David and Philippe Michel. 1999. "Optimal Growth when Tastes are Inherited." *Journal of Economic Dynamics and Control* 23 (4): 519–37.

———. 2001. "Altruism and Self-Restraint." *Annales d'Économie et de Statistique* 63–64:233–59.

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. 2006. "The 2005 Lawrence R. Klein Lecture: Emergent Class Structure." Forthcoming, International Economic Review.

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.

Mulligan, Casey B. 1997. *Parental Priorities and Economic Inequality.* University of Chicago Press.

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.