A Mathematical Appendix (to be published online)

Sequential Formulation of the Decision Problem:

The sequential decision problem corresponding to (4) is given by:

$$V^\star(A_0, B_0) = \max \Bigg\{ \sum_{t=0}^{\infty} z^t \big[\, (1 - B_t) \log(w_{1,i\,t}\; n_{1\,t}) + A_t\,(1 - n_{1\,t}) \\ + B_t \log(w_{2,i\,t}\; n_{2\,t}) + A_t\,(1 - n_{1\,t}) - l_{A\,t} - l_{B\,t} \big] \Bigg\} \qquad (15)$$

subject to $i_t \in I$ $n_{1\,t} \in [0, 1]$ $n_{2\,t} \in [0, 1]$ $l_{A\,t} \in [0, \bar{l}_{At}]$ $l_{B\,t} \in [0, \bar{l}_B]$ $A_{t+1} = \psi \bar{A} + (1 - \psi) A_t + g(l_{A\,t})$ and $B_{t+1} = \psi \bar{B} + (1 - \psi) B_t + f(l_{B\,t})$.

Proofs for all Lemmas and Propositions:

Proof of Lemma 1: Given that optimal labor supply is constant ($n_1 = n_2 = n$), the value function (4) can be written as:

$$V(A, B) = \max_{i \in I, l_A, l_B, n_1, n_2} \big\{ \log(n) + A(1-n) - l_A + (1 - B) \log(w_{1,i}) + B \log(w_{2,i}) - l_B + z\, V(A', B') \big\}.$$

Thus, the return function is additively separable, which implies that the value function is additively separable as well. We therefore have:

$$V(A, B) = v_A\,(A) + v_B\,(B)$$

with:

$$\begin{aligned} v_A(A) &= \max_{l_A, n} \big\{ \log(n) + A(1 - n) - l_A + z\, v_A(A') \big\}, \\ v_B(B) &= \max_{i \in I, l_B} \big\{ (1 - B) \log(w_{1,i}) + B \log(w_{2,i}) - l_B + z\, v_B(B') \big\}, \end{aligned}$$

subject to, respectively, (2) and (3). Q.E.D.

Proof of Proposition 1: We start by establishing that the value function is unique, non-decreasing, and convex. The proof is an application of Corollary 1 to Theorem 3.2 in Stokey and Lucas (1989). The Bellman equation (6) defines a mapping T on the space of bounded continuous functions on the interval $[\bar{B}, B_{\max}]$, endowed with the sup norm, where the mapping is given by:

$$T v_B(B) = \sup_{i \in I, 0 \le l_B \le \bar{l}_B} \big\{ (1 - B) \log(w_{1,i}) + B \log(w_{2,i}) - l_B + z\, v_B(\psi \bar{B} + (1 - \psi) B + f(l_B)) \big\}. \qquad (16)$$

Since we assume $0 < z < 1$, Blackwell's sufficient conditions for a contraction are met, and hence T has a unique fixed point by the Contraction Mapping Theorem. Using Corollary 1, we can now establish that the value function (i.e., the fixed point of the mapping T) is non-decreasing and (weakly) convex by establishing that the operator T preserves these properties.

To establish that the value function is non-decreasing, let v_B be a non-decreasing bounded continuous function. We need to show that $T v_B$ is non-decreasing as well. Choose two points

$B_h > B_l$ from the interval $[\bar{B}, B_{\max}]$,. We want to establish that $Tv_B(B_h) \geq Tv_B(B_l)$. 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 $\{\underline{w}_1, \underline{w}_2\}$ be choices attaining the maximum for B_l. We then have:

$$\begin{aligned} Tv_B(B_h) &\geq (1-B_h)\log(\underline{w}_1) + B_h \log(\underline{w}_2) - \underline{l} + z\, v_B(\psi\bar{B} + (1-\psi)B_h + f(\underline{l})) \\ &\geq (1-B_l)\ \log(\underline{w}_1) + B_l\ \log(\underline{w}_2) - \underline{l} + z\, v_B(\psi\bar{B} + (1-\psi)B_l\ + f(\underline{l})) = Tv_B(B_l), \end{aligned}$$

which is the desired result. Here the first inequality follows because the choices $\underline{l}, \{\underline{w}_1, \underline{w}_2\}$ may not be maximizing at B_h, and the second inequality follows because v_B is assumed to be increasing, and we have that $(1-B_h)\log(\underline{w}_1) + B_h \log(\underline{w}_2) \geq (1-B_l)\log(\underline{w}_1) + B_l \log(\underline{w}_2)$ since $\underline{w}_2 \geq \underline{w}_1$.

To establish that the value function is convex, let v_B be a (weakly) convex bounded continuous function. We need to establish that Tv_B is also a convex function. To show this, choose a number θ such that $0 < \theta < 1$, let $B_h > B_l$, and let $B = \theta B_h + (1-\theta)B_l$. We now need to show that $\theta Tv_B(B_h) + (1-\theta)Tv_B(B_l) \geq Tv_B(B)$. Let l and $\{w_1 w_2\}$ be choices attaining the maximum for B. Since these are feasible, but not necessarily optimal choices at B_h and B_l, we have:

$$\begin{aligned} Tv_B(B_h) &\geq (1-B_h)\log(w_1) + B_h \log(w_2) - l + z\, v_B(\psi\bar{B} + (1-\psi)B_h + f(l)), \\ Tv_B(B_l) &\geq (1-B_l)\log(w_1) + B_l \log(w_2) - l + z\, v_B(\psi\bar{B} + (1-\psi)B_l + f(l)). \end{aligned}$$

Using these results, we have:

$$\begin{aligned} &\theta\, Tv_B(B_h) + (1-\theta)\, Tv_B(B_l) \\ &\qquad \geq \theta \left[(1-B_h)\log(w_1) + B_h \log(w_2) - l + z\, v_B(\psi\bar{B} + (1-\psi)B_h + f(l))\right] \\ &\qquad\quad + (1-\theta) \left[(1-B_l)\log(w_1) + B_l \log(w_2) - l + z\, v_B(\psi\bar{B} + (1-\psi)B_l + f(l))\right] \\ &\qquad = (1-B)\log(w_1) + B\log(w_2) - l \\ &\qquad\quad + z\left[\theta v_B((1-\psi)B_h + f(l)) + (1-\theta)v_B(\psi\bar{B} + (1-\psi)B_l + f(l))\right] \\ &\qquad \geq (1-B)\log(w_1) + B\log(w_2) - l + z\, v_B(\psi\bar{B} + (1-\psi)B + f(l)) = Tv_B(B), \end{aligned}$$

which is the desired condition. Here, the last inequality follows from the assumed convexity of v_B. The operator T therefore preserves convexity, and thus the fixed point must also be convex.

Next, we would like to establish that the steepness of the optimal wage profile as well as the optimal investment in patience are non-decreasing in B. To this end, choose two patience levels $B_h > B_l$, and let the corresponding optimal choices be $\bar{l}, \overline{w}_1, \overline{w}_2$ and $\underline{l}, \underline{w}_1, \underline{w}_2$. We want to prove that $\bar{l} \geq \underline{l}$ and $\overline{w}_2/\overline{w}_1 \geq \underline{w}_2/\underline{w}_1$. Consider first the steepness of the income profile. Optimization in the choice of the income profile implies the following inequalities:

$$\begin{aligned} (1-B_h)\log(\overline{w}_1) + B_h\log(\overline{w}_2) &\geq (1-B_h)\log(\underline{w}_1) + B_h\log(\underline{w}_2), \\ (1-B_l)\log(\overline{w}_1) + B_l\log(\overline{w}_2) &\leq (1-B_l)\log(\underline{w}_1) + B_h\log(\underline{w}_2). \end{aligned}$$

Subtracting the two inequalities yields:

$$(B_h - B_l)\log((\overline{w}_2) - \log(\overline{w}_1)) \geq (B_h - B_l)\left(\log(\underline{w}_2) - \log(\underline{w}_1)\right).$$

or:

$$\frac{\overline{w}_2}{\overline{w}_1} \geq \frac{\underline{w}_2}{\underline{w}_1},$$

which is the desired condition.

Optimization in terms of the investment l implies the following inequalities:

$$\begin{aligned}
-\bar{l} + z\, v_B(\psi\bar{B} + (1-\psi)B_h + f(\bar{l})) &\geq -\underline{l} + z\, v_B(\psi\bar{B} + (1-\psi)B_h + f(\underline{l})),\\
-\bar{l} + z\, v_B(\psi\bar{B} + (1-\psi)B_l + f(\bar{l})) &\leq -\underline{l} + z\, v_B(\psi\bar{B} + (1-\psi)B_l + f(\underline{l})).
\end{aligned}$$

Subtracting the inequalities yields:

$$\begin{aligned}
&v_B(\psi\bar{B} + (1-\psi)B_h + f(\bar{l})) - v_B(\psi\bar{B} + (1-\psi)B_l + f(\bar{l}))\\
&\qquad \geq v_B(\psi\bar{B} + (1-\psi)B_h + f(\underline{l})) - v_B(\psi\bar{B} + (1-\psi)B_l + f(\underline{l})).
\end{aligned}$$

Given that f is increasing and v_B is increasing and convex, it follows that we must have $\bar{l} \geq \underline{l}$.

Given that we have established that the steepness of the optimal income profile w_2/w_1 and the optimal choice of investment in patience $l_B(B)$ is increasing in B, it follows that the patience and 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 as well.

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 in Assumption 1), 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 $l_B(B)$ 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 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 overlap only at their boundary points. Continuing this way, the state space $[\bar{B}, B_{\max}]$ 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 $[B_k, B_{k+1}]$ be such an interval. We now 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_{Bt} on the interval $[B_k, B_{k+1}]$ as:

$$\begin{aligned}
v_B(B) = \max \Bigg\{ &\log(w_{1,i_0}) + B \log\left(\frac{w_{2,i_0}}{w_{1,i_0}}\right) - l_0\\
&+ \sum_{t=1}^{\infty} z^t \left[\log(w_{1,i_t}) + \left(\psi^t \bar{B} + (1-\psi)^t B + \sum_{s=0}^{t} (1-\psi)^{t-s-1} f(l_s)\right) \log\left(\frac{w_{2,i_t}}{w_{1,i_t}}\right) - l_t\right]\Bigg\}. \qquad (17)
\end{aligned}$$

For given current and future income profiles, (17) is concave in l_t for all t, since f is concave. Moreover, patience B and all expressions involving l_{Bt} appear in separate terms in the sum. If f is strictly concave, 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 $[B_k, B_{k+1}]$ the current and future optimal income profiles are unique, the optimal policy correspondence $l_B(B)$ 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 $l_B(B)$ may take on more than one optimal value, one corresponding to each optimal set of income profiles. If f (or a segment of f) is linear, $l_B(B)$ is still generically single-valued on the interior of each interval, as exact indifference only occurs on a zero-measure subset of the parameter space.

The optimal value function v_B over the interval $[B_k, B_{k+1}]$ is given by (17) with income profiles i_t and investment in patience l_t fixed at their optimal (and constant) values. (17) 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 2: The law of motion for B, $f : [\bar{B}, B_{\max}] \to [\bar{B}, B_{\max}]$, is given by:

$$\Theta(B) = \psi \bar{B} + (1-\psi)\, B + f\left(l_B\left(B\right)\right),$$

where $l_B\left(B\right)$ is generically a non-decreasing step function (as described in Proposition 1). Since f is an increasing function and we assume that $\psi < 1$, the law of motion $\Theta(B)$ is strictly increasing in B. Notice that $\Theta(B)$ may fail to be single-valued for some B. Strictly increasing here means that $B_h < B_l$ implies $B'_h < B'_l$ for all $B'_h \in \Theta(B_h)$ and $B'_l \in \Theta(B_l)$, even if $\Theta(B_h)$ or $\Theta(B_l)$ is a set. For a given B_0, the law of motion Θ defines (potentially multiple) optimal sequences of patience $\{B_t\}_{t=0}^{\infty}$. Any such sequence is a monotone sequence on the compact set $[\bar{B}, B_{\max}]$, 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. If f (or a segment of f) is linear, the same results still apply generically, i.e., outside a zero-measure subset of the parameter space. Q.E.D.

Proof of Proposition 3: The strategy of the proof is analogous to the proof of Proposition 1. The Bellman equation (5) defines a mapping T on the space of bounded continuous functions on the interval $[\bar{A}, A_{\max}]$, endowed with the sup norm, where the mapping is given by:

$$Tv_A(A) = \sup_{l_A, n} \left\{ \log(n) + A(1-n) - l_A + z\, v_A(A') \right\}, \tag{18}$$

where $A' = \psi \bar{A} + (1-\psi)A + g(l_A)$. 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.

To establish that the value function is increasing, let v_A be a non-decreasing bounded continuous function. We need to show that Th is a non-decreasing function. Choose $A_h > A_l$. We want to establish that $Tv_B(A_h) > Tv_B(A_l)$. Since the right-hand side of (18) is the maximization of a continuous function over a compact set, the maximum is attained. Let $\underline{l}$ and $\underline{n}$ be the choices attaining the maximum for A_l. We have:

$$\begin{aligned} Tv_A(A_h) &\geq \log(\underline{n}) + A_h(1-\underline{n}) - \underline{l} + z\, v_A(\psi \bar{A} + (1-\psi)A_h + g(\underline{l})) \\ &\geq \log(\underline{n}) + A_l\ (1-\underline{n}) - \underline{l} + z\, v_A(\psi \bar{A} + (1-\psi)A_l\ + g(\underline{l})) = Tv_A(A_l), \end{aligned}$$

which is the desired result. Here the first inequality follows because the choice $\underline{l}$ may not be maximizing at A_h, and the second inequality follows because $A_h > A_l$ and v_A is assumed to be non-decreasing.

To prove that the value function is (weakly) convex, we establish that the operator T preserves convexity. Let v_A be a convex bounded continuous function. We need to establish that Th is also convex. Choose a number θ such that $0 < \theta < 1$, let $A_h > A_l$, and let $A = \theta A_h + (1-\theta)A_l$. We want to show that $\theta T v_A(A_h) + (1-\theta)Tv_A(A_l) \geq Tv_A(A)$. Let l and n be choices attaining the maximum for A. Since these are feasible, but not necessarily optimal choices at A_h and A_l, we have:

$$\begin{aligned} Tv_A(A_h) &\geq \log(n) + A_h(1-n) - l + z\, v_A(\psi \bar{A} + (1-\psi)A_h + g(l)), \\ Tv_A(A_l) &\geq \log(n) + A_l(1-n) - l + z\, v_A(\psi \bar{A} + (1-\psi)A_l + g(l)). \end{aligned}$$

Using these inequalities, we have:

$$\begin{aligned} &\theta T v_A(A_h) + (1-\theta)Tv_A(A_l) \\ &\qquad \geq \theta \left[\log(n) + A_h(1-n) - l + z\, v_A(\psi \bar{A} + (1-\psi)A_h + g(l))\right] \\ &\qquad\quad + (1-\theta)\left[\log(n) + A_l(1-n) - l + z\, v_A(\psi \bar{A} + (1-\psi)A_l + g(l))\right] \\ &\qquad = \log(n) + A(1-n) - l \\ &\qquad\quad + z\left[\theta v_A(\psi \bar{A} + (1-\psi)A_h + g(l)) + (1-\theta) v_A(\psi \bar{A} + (1-\psi)A_l + g(l))\right] \\ &\qquad \geq \log(n) + A(1-n) - l + z\, v_A(\bar{A} + (1-\psi)\left(A - \bar{A}\right) + g(l)) = Tv_A(A), \end{aligned}$$

which is the required condition. The last inequality follows from the assumed convexity of v_A. The operator T therefore preserves convexity, and thus the fixed point must also be convex. Q.E.D.

Proof of Proposition 4: To prove that $l_A\,(A)$ is a non-decreasing function of A, write the program as:

$$v_A(A) = \sup_{l_A} \left\{ -\log(A) + A - 1 - l_A + z\, v_A(\bar{A} + (1-\psi)\left(A - \bar{A}\right) + g(l_A)) \right\}.$$

Next, let $l_0 = l_A\,(A_0)$ and $l_1 = l_A\,(A_1)$, where $A_1 > A_0$. We want to prove that $l_1 \geq l_0$. To this aim, observe that

$$\begin{aligned} -l_0 + z\, v_A(\bar{A} + (1-\psi)\left(A_0 - \bar{A}\right) + g(l_0)) &\geq -l_0 + z\, v_A(\bar{A} + (1-\psi)\left(A_0 - \bar{A}\right) + g(l_1)) \\ -l_1 + z\, v_A(\bar{A} + (1-\psi)\left(A_1 - \bar{A}\right) + g(l_0)) &\leq -l_1 + z\, v_A(\bar{A} + (1-\psi)\left(A_1 - \bar{A}\right) + g(l_1)) \end{aligned}$$

Subtracting the two equations as before, we get:

$$\begin{aligned} l_1 - l_0 \geq\ & \left(z\, v_A(\bar{A} + (1-\psi)\left(A_0 - \bar{A}\right) + g(l_1)) - z\, v_A(\bar{A} + (1-\psi)\left(A_0 - \bar{A}\right) + g(l_0))\right) \\ & - \left(z\, v_A(\bar{A} + (1-\psi)\left(A_1 - \bar{A}\right) + g(l_1)) - z\, v_A(\bar{A} + (1-\psi)\left(A_1 - \bar{A}\right) + g(l_0))\right) \end{aligned} \tag{19}$$

(19) implies that $l_1 \geq l_0$. To see why, suppose (to derive a contradiction) that $l_1 < l_0$. Then, the left hand-side would be negative, while the right hand-side would be positive, since v_A is increasing and convex. This would contradict the inequality in (19). Therefore, we must have that $l_1 \geq l_0$. Hence, $l_A(A)$ must be non-decreasing in A.

The proof of convergence to the steady state is analogous to the proof of Proposition 2. Consider the equilibrium law of motion $A' = \Gamma(A)$ where

$$\Gamma(A) = \psi \bar{A} + (1-\psi) A + g(l_A(A)).$$

Since g is increasing and l_A is non-decreasing, $\Gamma(A)$ is strictly increasing in A. For a given A_0, the law of motion Γ defines (potentially multiple) optimal sequences of patience $\{A_t\}_{t=0}^{\infty}$. Any such sequence is a monotone sequence on the compact set $[\bar{A}, A_{\max}]$, and must therefore converge. The steady-state expression follows immediately from setting $A = \Gamma(A)$. Q.E.D.

Proof of Proposition 5: The proof is analogous to the proofs of Propositions 3 and 4. The optimal labor supply follows from taking the first-order condition in (8) while respecting the constraints $0 \leq n \leq 1$. Q.E.D.

Sufficient Condition Equilibrium with Constant Wages:

Condition 1 *Assume f to be of the form $f(l_B) = \xi l_B$, where ξ satisfies:*

$$\frac{1 - z(1-\psi)}{z} \leq \xi \log(\gamma) \leq \frac{1 - z(1-\psi)}{z} \frac{1}{z(1-\psi)}. \tag{20}$$

Proposition 6 *Suppose that Condition 1 is satisfied, and that the economy starts out with everyone having the natural patience $B_0 = \bar{B}$ and the steady-state taste for leisure A. Then for sufficiently large $q > 0$ there exists an equilibrium such that for all $t \geq 2$ the proportion of workers and artisans in the population is constant, and the agricultural wage is given by:*

$$\log w_F = \log\left(\alpha \mu^{\alpha-1}\right) = \log(q) + \bar{B}\log(\gamma) - \frac{\bar{l}_B}{z} + \frac{\xi \bar{l}_B \log(\gamma)}{1 - z(1-\psi)}. \tag{21}$$

The equilibrium is characterized by occupational segregation, i.e., from $t \geq 2$ onwards, parents and children in the same dynasty choose the same profession. The taste for leisure remains constant in all dynasties. Worker dynasties do not invest in patience ($l_B = 0$), whereas artisan dynasties invest the maximum amount ($l_B = \bar{l}_B$). The distribution of patience converges to a steady state where the patience of all workers remains at the natural level $\bar{B}$, whereas the patience of all artisans converges to the maximum $B_M = \bar{B} + \xi \bar{l}_B / \psi$.

Proof of Proposition 6: The proposed equilibrium satisfies the following conditions: A positive fraction of the young adults at time $t = 0$ invest in patience (at the level $l_B = \bar{l}_B$) in expectation of their children becoming artisans (at time $t = 2$); the remaining young adults do not invest and set $l_B = 0$; the agricultural wage is constant from time $t = 2$ onwards and adjusts so as to equalize the ex-ante utility of all young adults at time zero; from period $t = 2$ onwards, preferences diverge, and the members of the dynasties that did not invest in the first period prefer to be workers and not to invest in patience, while the members of dynasties that did invest in the first period prefer to be artisans and to invest in patience at the maximum level $l_B = \bar{l}_B$.

We construct the equilibrium in two steps. (i) We derive the equilibrium labor supply μ in agriculture from $t = 2$ onwards (and the corresponding wage) that makes the initial generation just indifferent between investing and not investing, provided that the equilibrium takes the prescribed form. (ii) We show that condition (20) implies that the prescribed occupational choices from period $t = 2$ onwards are indeed optimal.

(i) First notice that since f is linear, conditional on $l_B > 0$ it is (at least weakly) optimal to invest the maximum amount $l_B = \bar{l}_B$. When comparing the utility derived from investing and not investing, we can disregard the utility that the initial generation derives from consumption and leisure because of the separable utility function (this component of utility is the same for all first-generation families). Then, the value of not investing in patience (under the expectation that all future members of the dynasty will choose to be workers) is given by:

$$\tilde{v}_{B,F}\left(\bar{B}\right) = \frac{z}{1-z}\log\left(\alpha\mu^{\alpha-1}\right). \tag{22}$$

This is simply the discounted utility derived from receiving the worker's wage $w_F = \alpha\mu^{\alpha-1}$ from the next generation on. In contrast, the value of investing in patience (under the expectation that all future members of the dynasty will choose to be artisans) is:

$$\tilde{v}_{B,M}\left(\bar{B}\right) = -\bar{l}_B + z\, v_{B,M}\left(\bar{B} + \xi\bar{l}_B\right), \tag{23}$$

where:

$$v_{B,M}\left(B\right) = \log(q) + B\log(\gamma) - \bar{l}_B + z\, v_{B,M}\left(\psi\bar{B} + (1-\psi)B + \xi\bar{l}_B\right).$$

Notice that the artisan's utility depends not just on consumption, but also on the cost of investing $\bar{l}_B$. Solving for $v_{B,M}\left(B\right)$ yields:

$$v_{B,M}\left(B\right) = \frac{\log(q) - \bar{l}_B}{1-z} + \frac{z}{1-z}\frac{\left(\psi\bar{B} + \xi\bar{l}_B\right)\log(\gamma)}{\left(1 - z\left(1-\psi\right)\right)} + \frac{\log\left(\gamma\right)}{1 - z\left(1-\psi\right)}B.$$

Hence,

$$v_{B,M}\left(\bar{B} + \xi\bar{l}_B\right) = \frac{1}{1-z}\left(\log(q) - \bar{l}_B + \frac{\xi\bar{l}_B\log(\gamma)}{1 - z\left(1-\psi\right)} + \log\left(\gamma\right)\bar{B}\right),$$

which can be substituted into (23) to yield:

$$\tilde{v}_{B,M}\left(\bar{B}\right) = -\bar{l}_B + \frac{1}{1-z}\left(\log(q) - \bar{l}_B + \frac{\xi\bar{l}_B\log(\gamma)}{1 - z\left(1-\psi\right)} + \log\left(\gamma\right)B\right).$$

For the first generation to be indifferent between investing and not investing, we must have $\tilde{v_B}_A\left(\bar{B}\right) = \tilde{v}_{B,M}\left(\bar{B}\right)$, which in turn implies (after standard algebra) condition (21) as stated in the proposition:

$$\log(w_F) = \log\left(\alpha\mu^{\alpha-1}\right) = \log(q) + \bar{B}\log(\gamma) - \frac{\bar{l}_B}{z} + \frac{\xi\bar{l}_B\log(\gamma)}{1 - z\left(1-\psi\right)}.$$

In addition, the corresponding μ has to satisfy $\mu < n$ (where n is equilibrium labor supply), so that there is a positive fraction of artisans. This condition can always be met by choosing q sufficiently large.

(ii) We need to ensure that a young adult in period two who is endowed with patience $\bar{B} + \xi \bar{l}_B$ prefers being an artisan to working in agriculture at the flat wage w_F, while the opposite is true for an adult with patience $\bar{B}$. More formally,

$$\log(q) + \bar{B}\log(\gamma) - \frac{\bar{l}_B}{z} + \frac{\xi \bar{l}_B \log(\gamma)}{1 - z(1-\psi)} \leq \log(q) + \left(\bar{B} + \xi \bar{l}_B\right)\log(\gamma),$$
$$\log(q) + \bar{B}\log(\gamma) - \frac{\bar{l}_B}{z} + \frac{\xi \bar{l}_B \log(\gamma)}{1 - z(1-\psi)} \geq \log(q) + \bar{B}\log(\gamma).$$

These inequalities holds if and only if assumption (20) is satisfied. If these inequalities are satisfied, they hold *a fortiori* for all subsequent generations, because patience increases over time in artisan dynasties. Q.E.D.