

Appendix A: Three-period example with logarithmic utility

Suppose that $u(c) = \ln(c)$. The value of a worker $V_0^W(w)$ at $t = 0$ can be represented as

$$V_0^W(w) = \max \{V_{WE}(w), V_{WW}(w), V_{WW}^c(w)\},$$

where $V_{WE}(w)$ is the value of the agent who starts as a worker at $t = 0$ and continues as an entrepreneur at $t = 1$, $V_{WW}(w)$ of the agent who chooses to be a worker in both periods and makes some savings at $t = 0$, and $V_{WW}^c(w)$ is the value of the agent who is a worker in both periods and does not save anything at $t = 0$.¹ When the utility function is logarithmic, the closed form expressions for these values can be found:

$$\begin{aligned} V_{WE}(w) &= (1 + \beta + \beta^2) \ln \frac{w + \phi}{1 + \beta + \beta^2} + (\beta + 2\beta^2) \ln \beta(1 + r) + \beta^2 \ln A - \beta^2 \ln(1 + r) \\ V_{WW}(w) &= (1 + \beta + \beta^2) \ln \frac{w + \phi(1 + \frac{1}{1+r})}{1 + \beta + \beta^2} + (\beta + 2\beta^2) \ln \beta(1 + r) \\ V_{WW}^c(w) &= (1 + \beta + \beta^2) \ln \phi + \ln w + \beta^2 \ln \beta(1 + r) - (\beta + \beta^2) \ln(1 + \beta). \end{aligned} \tag{A-1}$$

Denote by $V_{EE}^s(w)$ the value of the agent who is a safe entrepreneur in both periods. Notice that $V_0^E(w) = V_{EE}^s(w)$ for all $w \geq w_0^H$. Thus the interval of initial wealth levels for which an agent chooses to become a risky entrepreneur at $t = 0$ exists if and only if the crossing of $V_0^W(w)$ and $V_{EE}^s(w)$ occurs to the left of w_0^H . The value $V_{EE}^s(w)$ is given by²

$$V_{EE}^s(w) = (1 + \beta + \beta^2) \ln \frac{w}{1 + \beta + \beta^2} + (\beta + 2\beta^2) \ln \beta A. \tag{A-2}$$

By equalizing each of the expressions in (A-1) with (A-2) and solving for w , we can find the wealth level at which $V_0^W(w)$ crosses $V_{EE}^s(w)$:

$$w_0 = \max \left\{ \frac{\phi}{\left(\frac{A}{1+r}\right)^{\frac{\beta+\beta^2}{1+\beta+\beta^2}} - 1}, \frac{\phi(1 + \frac{1}{1+r})}{\left(\frac{A}{1+r}\right)^{\frac{\beta+2\beta^2}{1+\beta+\beta^2}} - 1}, w_0^c \right\}, \tag{A-3}$$

where w_0^c solves $V_{WW}^c(w_0^c) = V_{EE}^s(w_0^c)$ (the exact formula for w_0^c is too cumbersome, so we did not write it explicitly in (A-3)).

¹Since $u(c)$ satisfies Inada conditions, the agent saves or invest a positive amount independently of his occupation at $t = 1$.

²A combination of Inada conditions and the properties of entrepreneurial technology assures that an agent who remains an entrepreneur in both periods invests positive amounts in his projects.

To find the expression for w_0^H , first notice that

$$w_0^H = \frac{1 + \beta + \beta^2}{\beta^2 A^2} \bar{w}'_1, \quad (\text{A-4})$$

where \bar{w}'_1 is the wealth accumulated by the entrepreneur with initial wealth w_0^H by the end of his life; it is also equal to the payoff of the safe project chosen by the entrepreneur with wealth \bar{w}_1 in the beginning of $t = 1$. The tangent wealth levels \underline{w}_1 and \bar{w}_1 are found from the following system of equations:

$$V_1^{W'}(\underline{w}_1) = V_1^{E'}(\bar{w}_1) = \frac{V_1^E(\bar{w}_1) - V_1^W(\underline{w}_1)}{\bar{w}_1 - \underline{w}_1}. \quad (\text{A-5})$$

Since $u(c)$ satisfies Inada conditions, we can use envelope conditions to simplify the above equations to:

$$(1 + r)u'(\underline{w}'_1) = Au'(\bar{w}'_1) = \frac{u(\bar{w}'_1) - u(\underline{w}'_1)}{\bar{w}'_1/A - \underline{w}'_1/(1 + r) + \phi}. \quad (\text{A-6})$$

Substituting $u(c) = \ln c$ and combining (A-4) and (A-6), we can solve for w_0^H :

$$w_0^H = \frac{1 + \beta + \beta^2}{\beta^2 A} \frac{\phi}{\ln(\frac{A}{1+r})}. \quad (\text{A-7})$$

Thus the interval of initial wealth levels in which agents choose to become risky entrepreneurs is non-empty if and only if

$$w_0 = \max \left\{ \frac{\phi}{\left(\frac{A}{1+r}\right)^{\frac{\beta+\beta^2}{1+\beta+\beta^2}} - 1}, \frac{\phi(1 + \frac{1}{1+r})}{\left(\frac{A}{1+r}\right)^{\frac{\beta+2\beta^2}{1+\beta+\beta^2}} - 1}, w_0^c \right\} < \frac{1 + \beta + \beta^2}{\beta^2 A} \frac{\phi}{\ln(\frac{A}{1+r})} = w_0^H. \quad (\text{A-8})$$

We can notice from (A-8) that $\frac{w_0}{w_0^H} \rightarrow 0$ as $\beta \rightarrow 0$ and $\frac{w_0}{w_0^H} \rightarrow x > 1$ as $\beta \rightarrow +\infty$ as long as $\frac{A}{1+r}$ is bounded from above by a number $R > 1$. Thus for any set of parameter values the risk taking interval must be nonempty if the agents are sufficiently impatient and it disappears if β is sufficiently large and the premium to entrepreneurial activity (relative to saving in a risk free asset) is not too large. In general, we were not able to establish that w_0 and w_0^H have a unique intersection when plotted as functions of β . However, in all the numerical exercises we computed (including the ones presented on Figure 2 in the main text of the paper), the cutoff level of β was unique.

Appendix B: Proofs of Proposition 2 and Lemma 1

The proofs of Proposition 2 and Lemma 1 require a number of intermediate steps. Lemma 2 below states an additional key result which is repeatedly used in the rest of the Appendix. Lemma 3 provides a detailed characterization of the workers' value function $R(w)$. We show that the statements of Proposition 2 and Lemma 1 directly follow from Lemma 3, but the latter holds only under certain conditions on entrepreneurs' value $V_E(w)$. We, therefore, build a recursive argument to verify that $V_E(w)$ satisfies these conditions.

It is convenient to introduce first the following notations. Define the operator T on a space of bounded, strictly increasing, strictly concave and continuously differentiable functions $C(\mathbb{R}_+)$:

$$TF(w) = \max_{a \geq 0} \{u(w + \phi - a) + \beta F((1+r)a)\}. \quad (\text{B-1})$$

Denote by $a(w)$ and $c(w) = w + \phi - a(w)$ the optimal saving and consumption policies associated with this maximization problem.

Lemma 2

Suppose that $F_1(w)$ and $F_2(w)$ (both from $C(\mathbb{R}_+)$) have a unique intersection at $w_0 > 0$ and $F_1(w) > F_2(w)$ for all $w > w_0$.

(i) Then $TF_1(w)$ and $TF_2(w)$ cannot have more than one intersections.

(ii) Suppose also that $TF_1(w)$ and $TF_2(w)$ have an intersection.

Denote by \underline{w}_0 and \bar{w}_0 ($\underline{w}_0 < \bar{w}_0$) the tangent end points with the common tangent line to $F_1(w)$ and $F_2(w)$, and by \underline{w}_1 and \bar{w}_1 ($\underline{w}_1 < \bar{w}_1$) the tangent end points with the common tangent line to $TF_1(w)$ and $TF_2(w)$.

If $\underline{w}_0 > 0$ and $\underline{w}_1 > 0$ then $(1+r)a_1(\underline{w}_1) = \underline{w}_0$ and $(1+r)a_2(\bar{w}_1) = \bar{w}_0$.

Proof of Lemma 2:

(i) First, notice that at any point \tilde{w} , at which $TF_1(\tilde{w}) = TF_2(\tilde{w})$, the optimal saving policies $a_1(\tilde{w})$ and $a_2(\tilde{w})$ should satisfy the following inequalities:

$$F_1((1+r)a_1(\tilde{w})) \geq F_2((1+r)a_1(\tilde{w})) \quad \text{and} \quad F_1((1+r)a_2(\tilde{w})) \leq F_2((1+r)a_2(\tilde{w})), \quad (\text{B-2})$$

which imply that $(1+r)a_1(\tilde{w}) \geq w_0$ and $(1+r)a_2(\tilde{w}) \leq w_0$.

The first statement of Lemma 2 is proven by contradiction. Suppose that $TF_1(w)$ and $TF_2(w)$ have multiple intersections. Take two intersection points, w_1 and w_2 ($w_1 < w_2$) such that $TF_1(w) \neq TF_2(w)$ for all $w \in (w_1, w_2)$. Since $TF_1(w)$ and $TF_2(w)$ are continuously differentiable, there exist $\hat{w}_1 \in (w_1, w_2)$ and $\hat{w}_2 \in (w_1, w_2)$ such that

$$TF_1'(\hat{w}_1) = TF_2'(\hat{w}_2) \quad \text{and} \quad TF_1(\hat{w}_1) = TF_2(\hat{w}_2). \quad (\text{B-3})$$

The first condition in (B-3) implies that $c_1(\hat{w}) = c_2(\hat{w})$. At the same time, strict monotonicity of the policy functions in conjunction with (B-2) imply that $(1+r)a_2(\hat{w}_2) < w_0$ and $(1+r)a_1(\hat{w}_1) > w_0$. Therefore, $F_1((1+r)a_1(\hat{w})) > F_2((1+r)a_2(\hat{w}))$, which suggest that $TF_1(\hat{w}) > TF_2(\hat{w})$, thereby contradicting to the second condition in (B-3). Therefore, $TF_1(w)$ and $TF_2(w)$ cannot have more than one intersection.

(ii) If $\underline{w}_1 > 0$ then the following condition must be satisfied:

$$TF_1'(\bar{w}_1) = TF_2'(\underline{w}_1) = \frac{TF_1(\bar{w}_1) - TF_2(\underline{w}_1)}{\bar{w}_1 - \underline{w}_1}. \quad (\text{B-4})$$

For sake of simplicity introduce the following notations: $\underline{c}_1 = c_1(\bar{w}_1)$, $\underline{c}_2 = c_2(\underline{w}_1)$, $\bar{a}_1 = a_1(\bar{w}_1)$ and $\underline{a}_2 = a_2(\underline{w}_1)$. The first equality in (B-4) implies that $\bar{c}_1 = \underline{c}_1$, and thus

$$\frac{TF_1(\bar{w}_1) - TF_2(\underline{w}_1)}{\bar{w}_1 - \underline{w}_1} = \beta(1+r) \frac{F_1((1+r)\bar{a}_1) - F_2((1+r)\underline{a}_1)}{(1+r)\bar{a}_1 - (1+r)\underline{a}_1}.$$

The envelope theorem and the first order conditions imply that $TF_i'(w) = \beta(1+r)F_i'((1+r)a_i(w))$, $i = 1, 2$. Together with (B-4), this leads to

$$F_1'((1+r)\bar{a}_1) = F_2'((1+r)\underline{a}_2) = \frac{F_1((1+r)\bar{a}_1) - F_2((1+r)\underline{a}_2)}{(1+r)\bar{a}_1 - (1+r)\underline{a}_2}.$$

The above equation may hold only for the end tangent points with the common tangent line to $F_1(w)$ and $F_2(w)$. Since $F_1(w)$ and $F_2(w)$ are strictly concave and have unique intersection, the bundle of such tangent points is also unique. Therefore, $(1+r)a_1(\underline{w}_1) = \underline{w}_0$ and $(1+r)a_2(\bar{w}_1) = \bar{w}_0$. *Q.E.D*

To simplify the notations, in the rest of Appendix B we *denote the value of a worker by $R(w)$ instead of $V_W(w)$* (which allows us to avoid using multiple sub- and super-indices). To prove Proposition 2 and Lemma 1, we use the fact that the worker's decision problem can be represented as the choice of the moment of entry:

$$R(w) = \max\{R_\infty(w), R_1(w), R_2(w), R_3(w), \dots\}, \quad (\text{B-5})$$

where $R_n(w)$, $n \geq 1$, is the value of a worker who plans to remain a worker for n periods (including the current one) and become an entrepreneur in period $n+1$ and $R_\infty(w)$ is the value of the agent who remains a worker forever.

For any $R_0(w) \in C(\mathbb{R}_+)$, define the sequence of functions $\{R_n(w)\}_{n=1}^\infty$ as

$$R_n(w) = TR_{n-1}(w) = T^n R_0(w), \quad n \geq 1, \quad (\text{B-6})$$

and denote by $a_n(w)$ the saving policy which maximizes $R_n(w)$. It is easy to verify that operator $T : C(\mathbb{R}_+) \rightarrow C(\mathbb{R}_+)$ satisfies Blackwell's sufficient conditions for contraction. Thus it has a unique fixed point which can be found as $R_\infty(w) = \lim_{n \rightarrow \infty} T^n R_0(w)$. The following Lemma characterizes some useful properties of the sequence of value functions $\{R_n(w)\}_{n=1}^\infty$ and the corresponding policies $\{a_n(w)\}_{n=1}^\infty$.

Lemma 3

Suppose that the following conditions hold:

- (i) $R_1(w)$ and $R_0(w)$ have at most one intersection and $R_1(w) < R_0(w)$ for sufficiently large $w \geq 0$;
- (ii) $R_0(w) > R_\infty(w)$ for sufficiently large $w \geq 0$;
- (iii) if $\beta(1+r) \leq 1$ then $R_\infty(w)$ and $R_0(w)$ have at most one intersection.¹

Then the sequence of functions $\{R_n(w)\}_{n=1}^\infty$ has the following properties:

- (a) If $\beta(1+r) > 1$ then $R_\infty(w) < \max\{R_1(w), R_2(w), \dots\}$;
- (b) The domain \mathbb{R}_+ can be partitioned into $N+1$ (N might be equal to $+\infty$) intervals by $0 \leq \tilde{w}_N < w_{N-1} < \dots < w_1 < w_0$ so that

$$\max\{R_\infty(w), R_1(w), R_2(w), \dots\} = \begin{cases} R_\infty(w), & w \in [0, \tilde{w}_N) \\ R_N(w), & w \in [\tilde{w}_N, w_N) \\ R_n(w), & w \in [w_{n+1}, w_n), \quad 2 \leq n \leq N-1 \\ R_1(w), & w \in [w_2, +\infty). \end{cases}$$

Note that (a) implies that $\tilde{w}_N = 0$ if $\beta(1+r) > 1$.

- (c) If $\max\{R_\infty(w), R_1(w), R_2(w), \dots\} = R_n(w)$ (where $n = \infty$ or $1 \leq n \leq N$)

$$\max\{R_\infty(w'_n), R_1(w'_n), R_2(w'_n), \dots\} = \begin{cases} R_\infty(w'_n), & n = \infty \\ R_{n-1}(w'_n), & 1 \leq n \leq N, \end{cases}$$
 where $w'_n = (1+r)a_n(w)$ is the next period's wealth level of the agent who maximizes $R_n(w)$ in the current period.

- (d) Denote by $\widehat{R}(w)$ the concave envelope of $\max\{R_\infty(w), R_0(w), R_1(w), R_2(w), \dots\}$. Then for any $w \in [0, w_1)$ such that $\max\{R_\infty(w), R_0(w), R_1(w), R_2(w), \dots\} = \widehat{R}(w) = R_n(w)$ the agent's next period wealth w'_n satisfies

$$\max\{R_\infty(w'_n), R_0(w'_n), R_1(w'_n), R_2(w'_n), \dots\} = \begin{cases} R_\infty(w'_n), & n = \infty \\ \widehat{R}(w'_n), & 1 \leq n \leq N. \end{cases}$$

Lemma 3 lists sufficient conditions under which the value of the worker $R(w)$, his optimal saving policy $a(w)$ and occupational decision can be described using fairly simple rules. Statement (a) says that if $\beta(1+r) > 1$ then any worker will eventually switch to a new (entrepreneurial) technology offering value $R_0(w)$. Statement (b) implies that $\max\{R_\infty(w), R_1(w), R_2(w), \dots\}$ has at most one intersection with $R_0(w)$ and, therefore, the cutoff level of wealth at which occupational switch occurs is unique. Statements (b) and (c) together imply that the savings policy is dynamically consistent: an agent who today plans to remain a worker for n more periods will decide to remain a worker

¹Statement (a) of Lemma 3 suggests that single crossing of $R_0(w)$ and $R_\infty(w)$ is relevant for characterizing $\max\{R_\infty(w), R_1(w), R_2(w), \dots\}$ only if $\beta(1+r) \leq 1$. Thus condition (iii) is imposed only for this case.

for $n - 1$ periods from tomorrow on. They also imply that the worker's wealth profile is monotone: it is (weakly) decreasing if initial worker's wealth falls into $[0, \tilde{w}_N)$ and is strictly increasing if his initial wealth belongs to $[\tilde{w}_N, w_1)$. Finally, (d) guarantees that if $\widehat{R}(w)$ is strictly concave at the agent's initial wealth level, it is also strictly concave at any wealth level the agent might have before he switches to a different technology. Clearly, Lemma 1 in the main text of the paper is a straightforward implication of the results (a)-(c) (with $\tilde{w} = \tilde{w}_N$).

Proof of Lemma 3:

- (a) To prove the first statement of Lemma 3 it suffices to show that there exist $K < +\infty$ such that $R_\infty(w) < \max\{R_1(w), \dots, R_K(w)\}$ for all $w \geq 0$. If $\beta(1+r) > 1$ then $(1+r)a_\infty(w) > w$ for all $w \geq 0$. Thus an agent saving according to $a_\infty(w)$ needs finite amount of time to accumulate any finite wealth level. Denote by \tilde{w}_0 the maximum wealth level at which $R_\infty(\tilde{w}_0) = R_0(\tilde{w}_0)$ (by condition (ii) $\tilde{w}_0 < +\infty$) and let K be the number of periods which are required to accumulate \tilde{w}_0 (or more) starting from zero wealth level. Denote by $\{\widehat{w}_i\}_{i=K}^1$ the sequence of the agent's intermediate wealth levels: $\widehat{w}_K = 0$, $\widehat{w}_{i-1} = (1+r)a_\infty(\widehat{w}_i)$ for all $2 \leq i \leq K$ and $(1+r)a_\infty(\widehat{w}_1) \geq \tilde{w}_0$.

Since $R_0(w) > R_\infty(w)$ for all $w > \tilde{w}_0$ and $(1+r)a_\infty(w) > \tilde{w}_0$ for all $w > \widehat{w}_1$, it follows that

$$\begin{aligned} R_\infty(w) &= u(w + \phi - a_\infty(w)) + \beta R_\infty((1+r)a_\infty(w)) < \\ &< u(w + \phi - a_\infty(w)) + \beta R_0((1+r)a_\infty(w)) \leq \\ &\leq u(w + \phi - a_1(w)) + \beta R_0((1+r)a_1(w)) = R_1(w), \text{ for all } w > \widehat{w}_1. \end{aligned}$$

Next, using $R_\infty(w) < R_1(w)$ for all $w > \widehat{w}_1$ and $(1+r)a_\infty(w) > \widehat{w}_1$ for all $w > \widehat{w}_2$ we can verify that $R_\infty(w) < R_2(w)$ for all $w > \widehat{w}_2$. Applying similar argument K times, we obtain that $R_\infty(w) < R_K(w)$ for all $w > \widehat{w}_K = 0$, which completes the proof of (a).

- (b) Lemma 2 together with assumption (i) imply that for all $n \geq 1$ the value functions $R_n(w)$ and $R_{n-1}(w)$ have at most one intersection and that $R_n(w) < R_{n-1}(w)$ for sufficiently large w . Denote the wealth level at their intersection by w_n and set $w_n = 0$ if $R_n(w) < R_{n-1}(w)$ for all $w \geq 0$. Similarly, assumption (ii) guarantees that $R_\infty(w) < R_n(w)$ for sufficiently large w and (iii) implies that $R_n(w)$ and $R_\infty(w)$ can have at most one intersection if $\beta(1+r) \leq 1$. Denote the largest wealth level at which they intersect by \tilde{w}_n and set $\tilde{w}_n = 0$ if $R_n(w) > R_\infty(w)$ for all $w \geq 0$. Note that if $w_n > \tilde{w}_n$ then $R_n(w) > \max\{R_\infty(w), R_{n-1}(w)\}$ for some $w \geq 0$.

To prove statement (b) of Lemma 3 we show that there exist $N \geq 0$ (possibly, $N = +\infty$) such that

- (I) for all $n \leq N$, $R_n(w) > \max\{R_\infty(w), R_1(w), \dots, R_{n-1}(w)\}$ for some $w \geq 0$ and $\tilde{w}_n < w_n < w_{n-1}$;

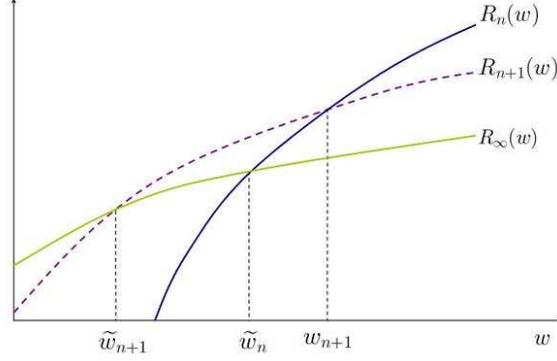


Figure B-1: Illustration to the setup of Claim 1.

(II) for all $n > N$ $R_n(w) \leq \max\{R_\infty(w), R_1(w), \dots, R_N(w)\}$ for all $w \geq 0$.

To verify that (I) is true, we need to establish two additional results:

Claim 1 (Refer to Figure B-1)

Suppose that $R_{n+1}(w) > \max\{R_\infty(w), R_n(w)\}$ for some $w \geq 0$ (and, consequently, $w_{n+1} > \tilde{w}_n > \tilde{w}_{n+1} \geq 0$). Then $(1+r)a_{n+1}(w_{n+1}) > w_{n+1}$.

Proof of Claim 1: Note that the conditions of Claim 1 imply that $R_n(w_{n+1}) > R_\infty(w_{n+1})$. Suppose that $(1+r)a_{n+1}(w_{n+1}) \leq w_{n+1}$. Then

$$\begin{aligned} R_{n+2}(w_{n+1}) &= u(w_{n+1} + \phi - a_{n+2}(w_{n+1})) + \beta R_{n+1}((1+r)a_{n+2}(w_{n+1})) \geq \\ &\geq u(w_{n+1} + \phi - a_{n+1}(w_{n+1})) + \beta R_{n+1}((1+r)a_{n+1}(w_{n+1})) \geq \\ &\geq u(w_{n+1} + \phi - a_{n+1}(w_{n+1})) + \beta R_n((1+r)a_{n+1}(w_{n+1})) = R_{n+1}(w_{n+1}). \end{aligned}$$

The first inequality follows from the fact that the saving policy $a_{n+1}(w)$ is feasible to the agent maximizing $R_{n+2}(w)$; the second inequality is a direct implication of the assumption $(1+r)a_{n+1}(w_{n+1}) \leq w_{n+1}$ and the fact that $R_{n+1}(w) \geq R_n(w)$ for all $w \leq w_{n+1}$. Since $R_{n+1}(w)$ and $R_{n+2}(w)$ have at most one intersection and $R_{n+1}(w) > R_{n+2}(w)$ for sufficiently large w , we conclude that $R_{n+2}(w) \geq R_{n+1}(w) \geq R_n(w)$ for all $w \leq w_{n+1}$.

Applying inductive argument, it is straightforward to show that $R_{n+i}(w) \geq R_{n+1}(w) \geq R_n(w)$ for all $i \geq 2$ and $w \leq w_{n+1}$. This implies that $R_\infty(w_{n+1}) = \lim_{i \rightarrow \infty} T^i R_n(w_{n+1}) \geq R_n(w_{n+1})$, which contradicts to $R_n(w_{n+1}) > R_\infty(w_{n+1})$. Thus the optimal saving at w_{n+1} must be such that $(1+r)a_{n+1}(w_{n+1}) > w_{n+1}$. *Q.E.D*

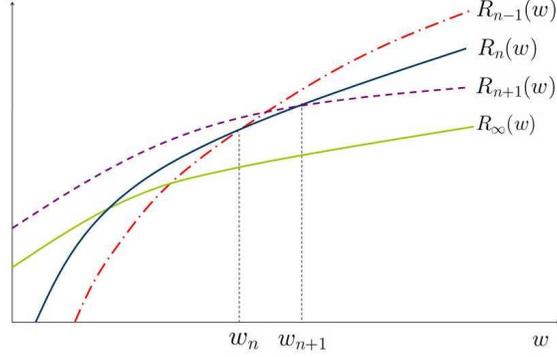


Figure B-2: Illustration to the proof Claim 2.

Claim 2 Suppose that for some $n \geq 1$ $R_n(w) > \max\{R_\infty(w), R_{n-1}(w)\}$ for some $w \geq 0$ and $R_{n+1}(w) > \max\{R_\infty(w), R_n(w)\}$ for some $w \geq 0$. Then $w_{n+1} < w_n$.

Proof of Claim 2: First, notice that Claim 1 implies that $(1+r)a_{n+1}(w_{n+1}) > w_{n+1}$. Suppose that $w_{n+1} \geq w_n$ as it is illustrated on Figure B-2. Since $R_{n-1}(w) > R_n(w)$ for all $w > w_n$ and $w_n \leq w_{n+1}$, it follows that

$$\begin{aligned} R_{n+1}(w_{n+1}) &= u(w_{n+1} + \phi - a_{n+1}(w_{n+1})) + \beta R_n((1+r)a_{n+1}(w_{n+1})) < \\ &< u(w_{n+1} + \phi - a_{n+1}(w_{n+1})) + \beta R_{n-1}((1+r)a_{n+1}(w_{n+1})) \leq \\ &\leq u(w_{n+1} + \phi - a_n(w_{n+1})) + \beta R_{n-1}((1+r)a_n(w_{n+1})) = R_n(w_{n+1}), \end{aligned}$$

which contradicts to $R_{n+1}(w_{n+1}) = R_n(w_{n+1})$. Thus $w_{n+1} < w_n$ must hold. *Q.E.D*

It follows from Claim 2 that if for some $n \geq 1$ it is true that for all $i \leq n$ $R_i(w) > \max\{R_\infty(w), R_1(w), \dots, R_{i-1}(w)\}$ for some $w \geq 0$ (the range of these wealth levels depends on i) then $0 \leq \tilde{w}_n < w_{n-1} < \dots < w_1 < w_0$ and

$$\max\{R_\infty(w), R_1(w), \dots, R_n(w)\} = \begin{cases} \max\{R_\infty(w), R_n(w)\} & \text{if } w \in [0, \tilde{w}_n) \\ R_n(w) & \text{if } w \in [\tilde{w}_n, w_n) \\ R_i(w) & \text{if } w \in [w_{i+1}, w_i), \quad 2 \leq i \leq n-1 \\ R_1(w) & \text{if } w \in [w_2, +\infty). \end{cases}$$

Note that if $\beta(1+r) \leq 1$ then $\max\{R_\infty(w), R_n(w)\} = R_\infty(w)$ for all $w \in [0, \tilde{w}_n)$. At the same time, if $\beta(1+r) > 1$ then $R_\infty(w) < \max\{R_1(w), \dots, R_n(w)\}$ if n is large enough ($n \geq K$) and the first row in the right hand side of the above equality disappears.

Denote by N the smallest n for which $R_{n+1}(w) < \max\{R_\infty(w), R_1(w), \dots, R_n(w)\}$ for all $w \geq 0$. Note that if $\beta(1+r) > 1$ then $N \geq K$ because $R_{n+1}(\tilde{w}_n) > R_\infty(\tilde{w}_n) = R_n(\tilde{w}_n)$ for all $n < K$ (as it was established in the proof of (a)).

In order to verify that the representation of $\max\{R_\infty(w), R_1(w), R_2(w), \dots\}$ described in (a) is correct we remain to verify that (II) holds, i.e. that $R_{N+i}(w) <$

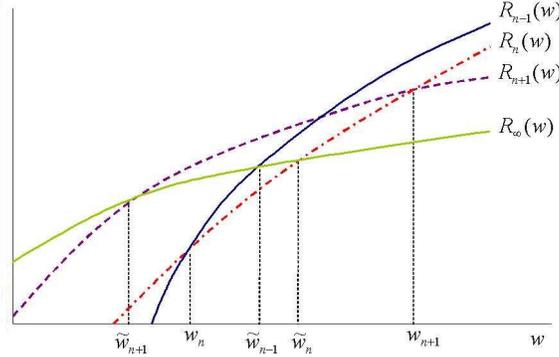


Figure B-3: Illustration to the proof of Claim 3

$\max\{R_{\infty}(w), R_1(w), \dots, R_N(w)\}$ for all $w \geq 0$ for all $i \geq 2$. This can be easily done with the help of another useful result:

Claim 3 *Suppose that for some $n \geq 1$ $R_n(w) \leq \max\{R_{\infty}(w), R_{n-1}(w)\}$ for all $w \geq 0$. Then $R_{n+1}(w) \leq \max\{R_{\infty}(w), R_n(w)\}$ for all $w \geq 0$.*

Proof of Claim 3: Suppose that $R_{n+1}(w) > \max\{R_{\infty}(w), R_n(w)\}$ for some $w \geq 0$. Then Claim 1 implies that $(1+r)a_{n+1}(w_{n+1}) > w_{n+1}$. The combination of $R_n(w) \leq \max\{R_{\infty}(w), R_{n-1}(w)\}$ for all $w \geq 0$ and $R_{n+1}(w) > \max\{R_{\infty}(w), R_n(w)\}$ for some $w \geq 0$ suggests that $w_n < \tilde{w}_{n-1} < \tilde{w}_n < w_{n+1}$ (see Figure B-3). Now we can construct contradiction by using an argument similar to the one that was used in the proof of Claim 2. *Q.E.D*

Applying Claim 3 inductively, we obviously conclude that (II) holds.

- (c) The third result of Lemma 3 directly follows from an intermediate result (B-2) established in the proof of Lemma 2. Since $R_{n+1}(w) = TR_n(w)$ and $R_n(w) = TR_{n-1}(w)$, (B-2) implies that $(1+r)a_n(w) \geq w_n$ for all $w \geq w_{n+1}$ for all $1 \leq n \leq N$. At the same time, since $R_n(w) = TR_{n-1}(w)$ and $R_{n+1}(w) = TR_n(w)$, it follows that $(1+r)a_n(w) \leq w_{n-1}$ for all $w \leq w_n$ for all $1 \leq n \leq N$.

Finally, from (a) we know that the equality $\max\{R_{\infty}(w), R_1(w), R_2(w), \dots\} = R_{\infty}(w)$ can be true for some $w \geq 0$ only if $\beta(1+r) \leq 1$. In this case, $w'_{\infty} = (1+r)a_{\infty}(w) \leq w$ for all $w \geq 0$ and thus $\max\{R_{\infty}(w'_{\infty}), R_1(w'_{\infty}), \dots\} = R_{\infty}(w'_{\infty})$ holds for those w for which $\max\{R_{\infty}(w), R_1(w), R_2(w), \dots\} = R_{\infty}(w)$. This completes the proof of statement (c) of Lemma 3.

- (d) The last statement of Lemma 3 is a direct implication of result (ii) of Lemma 2. Denote by \underline{w}_n and \bar{w}_n the common tangent points to $R_n(w)$ and $R_{n-1}(w)$. Since $R_{n+1}(w) = TR_n(w)$ and $R_n(w) = TR_{n-1}(w)$, it follows from (ii) of Lemma 2 that

$(1+r)a_{n+1}(\underline{w}_{n+1}) = \underline{w}_n$ and $(1+r)a_n(\overline{w}_{n+1}) = \overline{w}_n$. Correspondingly, by the first order and envelope conditions,

$$R'_n(\overline{w}_{n+1}) = R'_{n+1}(\underline{w}_{n+1}) = \beta(1+r)R'_n(\underline{w}_n). \quad (\text{B-7})$$

If $\beta(1+r) > 1$ then (B-7) implies that $R_n(\overline{w}_{n+1}) < R_n(\underline{w}_n)$ and, therefore, $\overline{w}_{n+1} < \underline{w}_n$ (see Figure B-4). Then using (ii) of Lemma 2 again and recalling that $a_n(w)$ is monotone for every $n \geq 1$, we can conclude that if $w \in [\overline{w}_{n+1}, \underline{w}_n)$ (and thus $\widehat{R}(w) = R_n(w)$) then $(1+r)a_n(w) \in [\overline{w}_n, \underline{w}_{n-1})$ and, correspondingly, $\widehat{R}(w'_n) = R_{n-1}(w'_n)$.

Finally, if $\beta(1+r) \leq 1$ then (B-7) implies that $R_n(\overline{w}_{n+1}) \geq R_n(\underline{w}_n)$ and, consequently, $\overline{w}_{n+1} \geq \underline{w}_n$ ($\overline{w}_{n+1} = \overline{w}_n$ if and only if $\beta(1+r) = 1$) (see Figure B-5 for illustration). Thus $\widehat{R}(w) = R(w)$ is possible only either if $w \in [0, \underline{w})$ or, in case of $\beta(1+r) = 1$, also if $w = \overline{w}_n = \underline{w}_{n-1}$. In the former case $w' = (1+r)a_\infty(w) \leq w < \underline{w}$ for all $w < \overline{w}$ and thus $\widehat{R}(w') = R_\infty(w')$. If $\beta(1+r) = 1$ and $w = \overline{w}_n$ for some $n \geq 2$ then, by (ii) of Lemma 2, $w' = \overline{w}_{n-1}$ and thus $\widehat{R}(w') = R_{n-1}(w')$. This completes the proof of (d). *Q.E.D*

Proof of Proposition 2: Note that Lemma 3 immediately implies that the function $\max\{R_\infty(w), R_1(w), R_2(w), \dots\}$ has at most one intersection with $R_0(w)$. Therefore, in order to show that $R(w)$ and $V_E(w)$ have unique intersection, it is enough to verify that the assumptions of Lemma 3 are satisfied for $R_0(w) = V_E(w)$. We do this by using a standard recursive argument and showing that assumptions (i)-(iii) of Lemma 3 are self-generating. Suppose that the assumptions of Lemma 3 hold for some $V_E^n(w)$ (the value function of entrepreneur obtained at the n th step of iteration). Given this $V_E^n(w)$, we construct the value of the worker $R^n(w) = \max\{R_\infty(w), R_1^n(w), R_2^n(w), \dots\}$ (the properties of which are described in Lemma 3) and characterize the value of entrepreneur $V_E^{n+1}(w)$ implied by such $R^n(w)$ and $V_E^n(w)$. Denote by $R_1^{n+1}(w)$ the value of the worker who plans to become an entrepreneur and continue with the value $V_E^{n+1}(w)$ in the next period. To prove Proposition 3, it suffices to show that $V_E^{n+1}(w)$ and $R_1^{n+1}(w)$ satisfy assumptions (i)-(iii) of Lemma 3.

Denote by $a^n(w)$ and $c_R^n(w)$ the optimal saving and consumption policies of the worker with value $R^n(w)$; by $k^n(w)$ and $c_E^n(w)$ – the optimal saving and consumption policies of the entrepreneur whose value function is given by $\widehat{V}_E^n(w)$; and by \underline{w}^n and \overline{w}^n the end tangent points of $R^n(w)$ and $V_E^n(w)$ with their common tangent line (recall that by assumption of the recursive argument $R^n(w)$ and $V_E^n(w)$ have at most one intersection).

As it has been argued in the main text of the paper, risk taking allows entrepreneurs to eliminate kinks in their next period's value $\max\{V_E^n(w), R^n(w)\}$. Denote its concave envelope by $\widehat{V}^n(w)$. Then (ii) of Lemma 2 implies that the value of the entrepreneur $V_E^{n+1}(w)$ is the concave envelope of two value functions, $V_{VE}^n(w)$ and $V_R^n(w)$. The

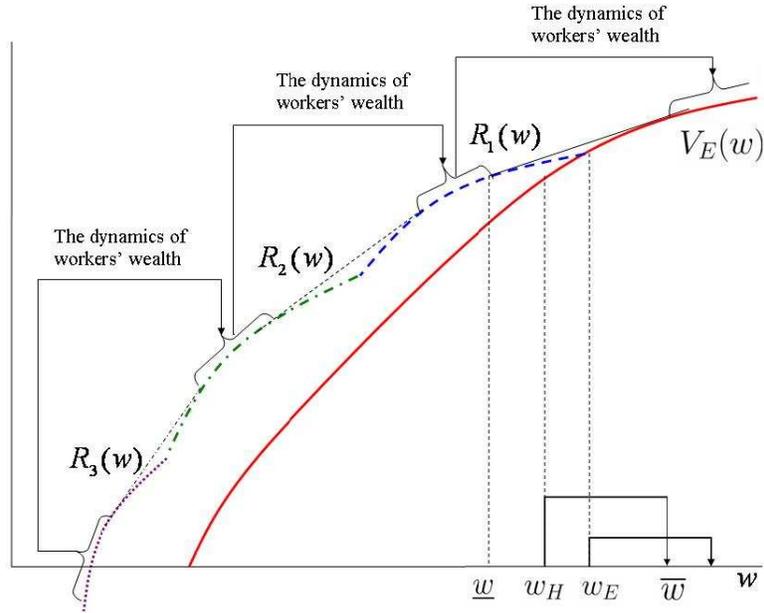


Figure B-4: Workers' value $R(w)$ and the dynamics of workers' wealth for $\beta(1+r) > 1$: $Ak(w_E) > \bar{w}$, no risk taking.

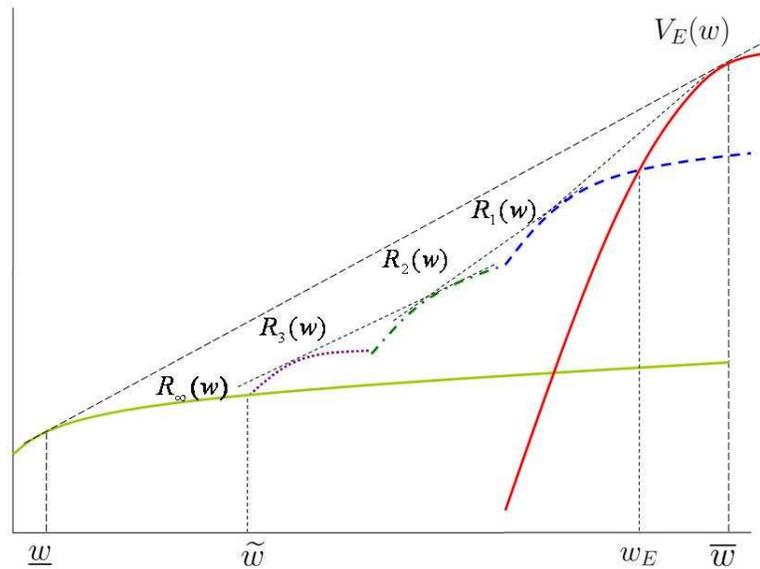


Figure B-5: Workers' value $R(w)$ for $\beta(1+r) < 1$: risk taking is possible.

former stands for the value of an entrepreneur with continuation value $V_E^n(w)$,

$$V_{VE}^n(w) = \max_{k \geq 0} \{u(w - k) + \beta V_E^n(Ak)\}, \quad (\text{B-8})$$

and $V_R^n(w)$ is the value the entrepreneur who becomes a worker in the next period:

$$V_R^n(w) = \max_{k \geq 0} \{u(w - k) + \beta R^n(Ak)\}. \quad (\text{B-9})$$

The proof can be significantly simplified if we choose the “right” initial guess for the entrepreneur’s value. Suppose that we start iterating on $V_E(w)$ by assuming that $V_E^0(w) = V_E^f(w)$ for all $w \geq 0$, where $V_E^f(w)$ is the value of the agent who remains a safe entrepreneur forever (which has been defined in section 3.3 of the main text). Then, using a simple inductive argument, it is straightforward to verify that $V_E^n(w) = V_E^f(w)$ for all sufficiently large w at every step of iteration (the key is to show that any new iteration shifts upward the value of the worker $R^n(w)$ and, correspondingly, raises the end tangent point \bar{w}^n). Knowing this, it is straightforward to verify that $V_E^n(w) > R_1^{n+1}(w)$ and $V_E^n(w) > R_\infty(w)$ for sufficiently large $w \geq 0$ (since $1 + r < A$), which guarantees that the second part of (i) and (ii) of Lemma 3 are satisfied. Thus we remain to check single crossing of $R_1^{n+1}(w)$ and $V_E^{n+1}(w)$ as well as of $R_\infty(w)$ and $V_E^{n+1}(w)$ (if $\beta(1 + r) \leq 1$).

Part I: We start by establishing single crossing of $R_1^{n+1}(w)$ and $V_E^{n+1}(w)$.

Suppose that $R_1^{n+1}(w)$ and $V_E^{n+1}(w)$ have more than one intersection. Since $R_1^{n+1}(0) > V_E^{n+1}(0)$ (due to $\phi > 0$) and $R_1^{n+1}(w) < V_E^{n+1}(w)$ for sufficiently large w (as it was argued in the previous paragraph), there exist w_1 and w_2 such that $R_1^{n+1}(w)$ is strictly concave at w_1 , $V_E^{n+1}(w)$ is strictly concave at w_2 and the following conditions are satisfied:

$$(R_1^{n+1})'(w_1) = (V_E^{n+1})'(w_2) \quad \text{and} \quad R_1^{n+1}(w_1) \geq V_E^{n+1}(w_2). \quad (\text{B-10})$$

We will show now that these two conditions cannot hold simultaneously. Notice that by doing this, we also establish the last claim of Proposition 2, i.e. that the value of the worker is *strictly* flatter than the value of the entrepreneur at their crossing point. The value of the worker at w_1 can be represented as

$$R_1^{n+1}(w_1) = u(c_{R_1}^{n+1}(w_1)) + \beta V_E^{n+1}(w_1'), \quad (\text{B-11})$$

where $w_1' = (1 + r)a_1^{n+1}(w_1)$. Simultaneously, the first order and envelope conditions imply that $(R_1^{n+1})'(w_1) = \beta(1 + r)(V_E^{n+1})'(w_1')$. Similarly, the value of the entrepreneur at w_2 is decomposed as

$$V_E^{n+1}(w_2) = u(c_E^{n+1}(w_2)) + \beta \widehat{V}^n(w_2'), \quad (\text{B-12})$$

where $w_2' = Ak^{n+1}(w_2)$ and $(V_E^{n+1})'(w_2) = \beta A(\widehat{V}^n)'(w_2')$.

The first condition in (B-10) also implies that $c_{R_1}^{n+1}(w_1) = c_E^{n+1}(w_2)$. Therefore, $R_1^{n+1}(w_1) \geq V_E^{n+1}(w_2)$ is possible only if

$$V_E^{n+1}(w_1') \geq \widehat{V}^n(w_2'). \quad (\text{B-13})$$

To show that these inequality cannot hold, consider two separate cases, $w_2 > w_H^n$ and $w_2 < w_L^n$ (where (w_L^n, w_H^n) is the interval within which entrepreneurs invest in risky projects).²

- (i) If $w_2 > w_H^n$ then $w'_2 > \bar{w}^n \geq \bar{w}^{n-1}$ and thus $\widehat{V}^n(w'_2) = V_F(w'_2) = V_E^n(w'_2) = V_E^{n+1}(w'_2)$. Consequently, $(V_E^{n+1})'(w_2) = \beta A(V_E^n)'(w'_2)$, which, in a combination with $(R_1^{n+1})'(w_1) = (V_E^{n+1})'(w_2)$ and $(R_1^{n+1})'(w_1) = \beta(1+r)(V_E^n)'(w'_1)$, implies that $(V_E^n)'(w'_1) > (V_E^n)'(w'_2)$. Since $V_E^n(w)$ is concave, it follows that $V_E^n(w'_1) < V_E^n(w'_2) = \widehat{V}^n(w'_2)$, which contradicts to inequality (B-13).
- (ii) If $w_2 < w_L^n$ then $w'_2 < \underline{w}^n$ and, since $V_E^{n+1}(w)$ is strictly concave at w_2 , $\widehat{V}^n(w'_2) = R^n(w'_2)$ (i.e., the entrepreneur with w_2 will continue as a worker with w'_2 in the following period). To show that (B-13) cannot be satisfied, it is convenient to decompose $V_E^{n+1}(w'_1)$ and $\widehat{V}^n(w'_2)$ into the sum of the utility of optimal consumption at w'_1 and w'_2 and the discounted continuation values at $w''_1 = Ak^n(w'_1)$ and $w''_2 = (1+r)a^n(w'_2)$ respectively. By construction of $V_E^{n+1}(w)$, the entrepreneur at w'_1 continues with $\widehat{V}^n(w''_1)$:

$$V_E^{n+1}(w'_1) = u(\widehat{c}_E(w'_1)) + \beta \widehat{V}^n(w''_1), \quad (\text{B-14})$$

and $(V_E^{n+1})'(w'_1) = \beta A(\widehat{V}^n)'(w''_1)$. At the same time, it follows from (d) of Lemma 3 that the continuation value $\max\{R^n(w), V_E^n(w)\}$ of a worker at w'_2 can be replaced with $\widehat{V}^n(w)$ (i.e., the payoff to worker's optimal saving w''_2 falls into an interval, in which $\widehat{V}^n(w) = \max\{R^n(w), V_E^n(w)\}$ for all w from some neighborhood of w''_2). Therefore,

$$\widehat{V}^n(w'_2) = u(c_R^n(w'_2)) + \beta \widehat{V}^n(w''_2), \quad (\text{B-15})$$

and the marginal condition $(\widehat{V}^n)'(w'_2) = \beta(1+r)(\widehat{V}^n)'(w''_2)$ must be satisfied.

Combining all marginal conditions, we obtain that

$$\begin{aligned} (R_1^{n+1})'(w_1) &= \beta(1+r)(V_E^{n+1})'(w'_1) = \beta^2 A(1+r)(\widehat{V}^n)'(w''_1) \quad \text{and} \\ (V_E^{n+1})'(w_2) &= \beta A(\widehat{V}^n)'(w'_2) = \beta^2 A(1+r)(\widehat{V}^n)'(w''_2), \end{aligned}$$

which suggests that $(\widehat{V}^n)'(w''_1) = (\widehat{V}^n)'(w''_2)$ (since $(R_1^{n+1})'(w_1) = V_E^{n+1}(w_2)$) and thus $\widehat{V}^n(w''_1) = \widehat{V}^n(w''_2)$ (since $\widehat{V}^n(w)$ must be strictly concave in w''_1 and in w''_2). Thus condition (B-13) may hold only if $c_R^n(w'_2) \leq c_E^n(w'_1)$, which is impossible because

$$\begin{aligned} u'(c_R^n(w'_2)) &= \text{/by (B-15)/} = (\widehat{V}^n)'(w'_2) = \text{/by (B-12)/} = \frac{1}{\beta A}(V_E^{n+1})'(w_2) < \\ &< \frac{1}{\beta(1+r)}(R_1^{n+1})'(w_1) = \text{/by (B-11)/} = (V_E^{n+1})'(w'_1) = \text{/by (B-14)/} = u'(c_E^n(w'_1)). \end{aligned}$$

²In principle, it could happen that $w_2 \in (w_L, w_H)$, which would imply that $R(w)$ coincides with $V_E(w)$ inside a part of the interval $[w_L, w_H]$. This could occur only if $\beta(1+r) = 1$. However, simple argument using (ii) of Lemma 2 eliminates this case because it would imply that $w_H = \bar{w}$, which is impossible since $1+r < A$.

Part II: Finally, we need to verify that $V_E^{n+1}(w)$ and $R_\infty(w)$ cannot have multiple intersections if $\beta(1+r) \leq 1$.

Since $\beta(1+r) \leq 1$, the concave envelope of $V_E(w)$ and $R(w)$ coincides with the concave envelope of $V_F(w)$ and $R_\infty(w)$. Thus the entrepreneurs' value function $V_E^{n+1}(w)$ is a concave envelope of $V_F(w)$ and $V_\infty(w) = \max_{k \geq 0} \{u(w-k) + \beta R_\infty(Ak)\}$. Since $R_\infty(0) > V_\infty(0)$, such $V_E^{n+1}(w)$ has a unique intersection with $R_\infty(w)$ if $V_F(w)$ and $R_\infty(w)$, as well as $V_\infty(w)$ and $R_\infty(w)$, cannot have more than one intersection.

For the first pair of the value functions, $V_E^f(w)$ and $R_\infty(w)$, it can be easily shown by contradiction that at any intersection point \hat{w} (such that $R_\infty(\hat{w}) = V_E^f(\hat{w})$) the inequality $R'_\infty(\hat{w}) < V_E^{f'}(\hat{w})$ must hold. Obviously, this property eliminates the possibility of multiple crossing of $R_\infty(w)$ and $V_E^f(w)$. For the second pair of value functions, $V_\infty(w)$ and $R_\infty(w)$, multiple crossing would imply existence of $\hat{w} > 0$ such that $R'_\infty(\hat{w}) = V'_\infty(\hat{w})$ and $R_\infty(\hat{w}) > V_\infty(\hat{w})$. However, these two conditions cannot hold simultaneously because $A > 1+r$: the equality of the slopes at \hat{w} implies that $R_\infty((1+r)a_\infty(\hat{w})) < R_\infty(Ak(\hat{w}))$, and, therefore, $R'_\infty(\hat{w}) = V'_\infty(\hat{w})$ is possible only if $R_\infty(w) < V_\infty(w)$. *Q.E.D*

Proof of Lemma 1:

Both statements of Lemma 1 follow directly from (a)-(c) of Lemma 3 (where $R_0(w) = V_E(w)$). The conditions of Lemma 3 have been verified recursively in the proof of Proposition 2. *Q.E.D.*

Lemma 4 (*Entry and risk taking*)

If entry into entrepreneurship ever occurs (i.e. $(1+r)a(w_E) > w_E$) then $w_E > w_L$.

Proof of Lemma 4:

First of all, notice that $V'_W(w_E) < V'_E(w_E)$ (established in Proposition 2) implies that the worker at w_E has higher current period consumption than the entrepreneur with the same wealth level, i.e. $u(w_E + \phi - a(w_E)) > u(w_E - k(w_E))$, where $a(w)$ and $k(w)$ are the optimal saving policies for the worker and the entrepreneur respectively.

Suppose that $w_E \leq w_L$. By definition of w_L , this implies that $Ak(w_E) \leq \underline{w} < w_E$, i.e. the entrepreneur with wealth w_E invests in a safe project and exits from business in the following period. Thus

$$\begin{aligned} V_E(w_E) &= u(w_E - k(w_E)) + \beta V_W(Ak(w_E)) < u(w_E - k(w_E)) + \beta V_E(w_E) \\ &< u(w_E + \phi - a(w_E)) + \beta V_E(w_E) \\ &< u(w_E + \phi - a(w_E)) + \beta V_E((1+r)a(w_E)) = V_W(w_E), \end{aligned}$$

where the last inequality follows from the assumption of this Proposition ($(1+r)a(w_E) > w_E$). Obviously, the above relationship contradicts to $V_E(w_E) = V_W(w_E)$, and thus the existence of entry into entrepreneurship must imply that $w_E > w_L$. *Q.E.D.*

Appendix C: Assessing the quantitative importance of the risk taking mechanism studied in the paper

1. Restricting the set of available risky projects

The simplicity of the characterization of entrepreneurial risk choice in our paper is driven by the fact that we allow entrepreneurs to invest in projects with two-point distribution of returns. Because of this assumption, entrepreneurs can eliminate all nonconcavities in their continuation value by choosing the projects which put all weight at the tangent points of the concave envelope of the two value functions (the value of the worker $V_W(w)$ and the value of the entrepreneur $V_E(w)$). In this Appendix we analyze how the optimal project choice changes if the set of available entrepreneurial projects that does not include technologies with the two-point returns' distribution. In general, limiting the set of available projects and eliminating the projects with degenerate distribution of returns might affect risk taking incentives for two reasons:

- (i) the projects with non-degenerate distribution of returns may have support on areas where the maximum of the two value functions, $V_W(w)$ and $V_E(w)$, is strictly concave, thereby 'adding risk' where entrepreneurs do not need it;
- (ii) if risk taking opportunities are *very limited*, it might appear that in most cases the distribution of returns would have support in a strictly concave region of $\max\{V_W(w), V_E(w)\}$, which could significantly narrow risk taking wealth interval (or even make it empty).

Unfortunately, eliminating projects with two-point distribution of returns from the set of available projects makes the theoretical characterization of entrepreneurial risk choice impossible due to a number of important technical challenges. In particular, it is clear from (i) above that entrepreneurial risk choice should depend on the functional form of the returns' distribution function and the degrees of concavity of $V_W(w)$ and $V_E(w)$ near their intersection. The shapes of both value functions are, in turn, determined by the optimal risk choice, which significantly complicates any theoretical analysis.

Thus we turn to a series of numerical exercises. First, in order to address (i) above, we limit the set of available projects by allowing entrepreneurs to invest only in the projects with *the uniform distribution* of returns so that risk cannot be restricted to where it is wanted. Second, in order to shed the light on statement (ii) above, we limit the set of available projects even further by restricting the projects' variance.

The main results of our numerical exercise can be summarized as follows:

- (i) The main *qualitative properties* of project choice obtained in our paper *are maintained* if we consider only the projects with the uniform distribution of returns.

Figures 1 and 2 illustrate how the properties of the optimal risk choice adjust if we limit the set of distributions to uniforms for two different set of parameters. The main observations are:

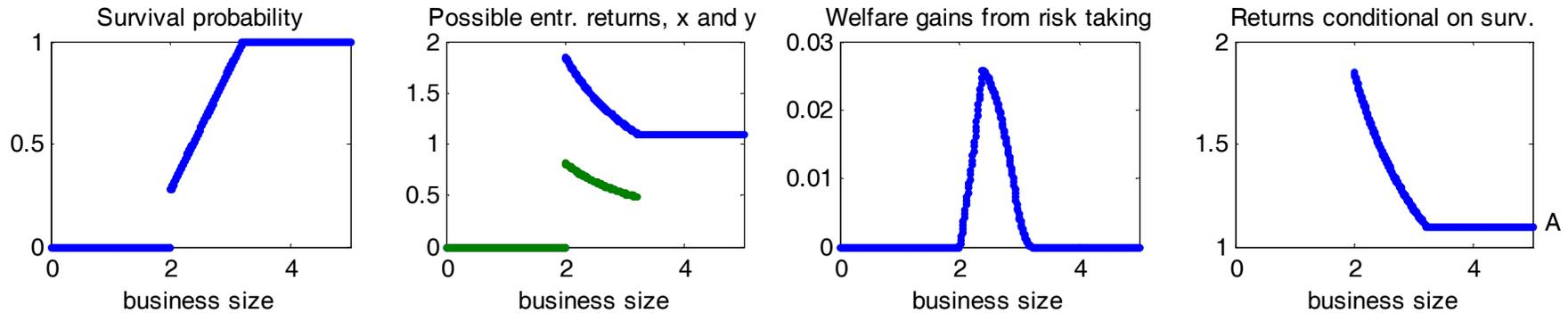
- Only relatively poor entrepreneurs decide to invest in risky projects.
 - The probability of survival of entrepreneurial firm is positively correlated with entrepreneurial wealth.
 - Allowing entrepreneurs to invest only in the projects with uniform distribution of returns still generates substantial welfare gains from risk taking.
- (ii) Even if risk taking opportunities are *very limited*, risk taking does not necessarily become a local phenomenon, *i.e.* the range of wealth within which entrepreneurs make risky investments may be quite large.

Table 1 shows how the range of risk taking interval changes when we restrict the variance of available projects. We can see that even if the level of variance is significantly smaller than the one entrepreneurs would choose in the unconstrained case, the risk taking intervals are very close to the ones we observe when entrepreneurs can invest in any project with the uniform distribution of returns. This is because the possibility of the limited risk taking at the point where the two value functions intersect recursively generates a series of kinks at the higher wealth levels, thereby widening risk taking intervals.

Figure C-1: Optimal project choice from the set of the projects
with uniformly distributed returns, $\beta A > 1$

Parameters: $\beta = 0.95$, $r = 0.042$, $A = 1.1$, $\sigma = 2$, $\varphi = 0.15$ ($\beta(1+r) \approx 0.99$, $\beta A = 1.045$)

1. Choice of (x, p) from the general set of projects $\Omega(A)$ including the projects with two-point-distribution technologies



2. Choice from a set of projects with uniformly distributed returns ($\underline{A} \in [0, A]$ is the lowest possible return of the chosen project)

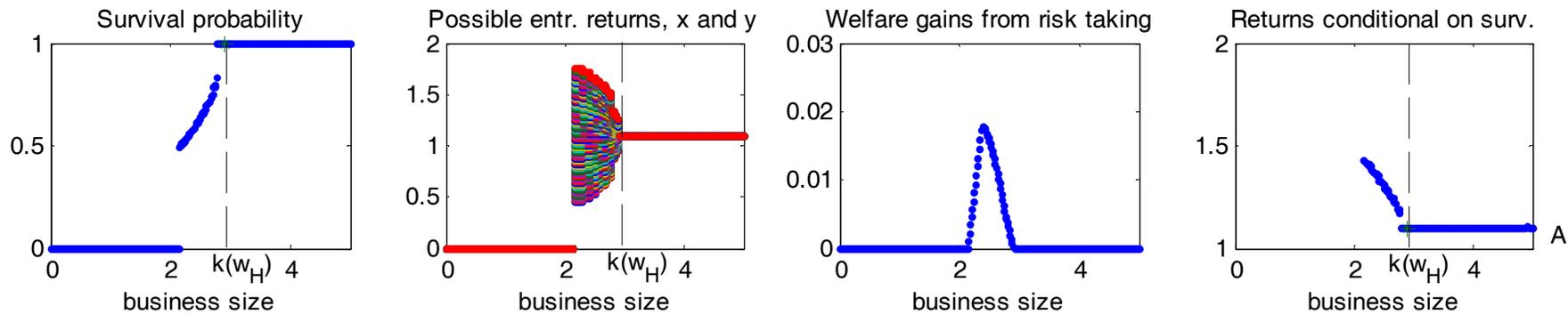
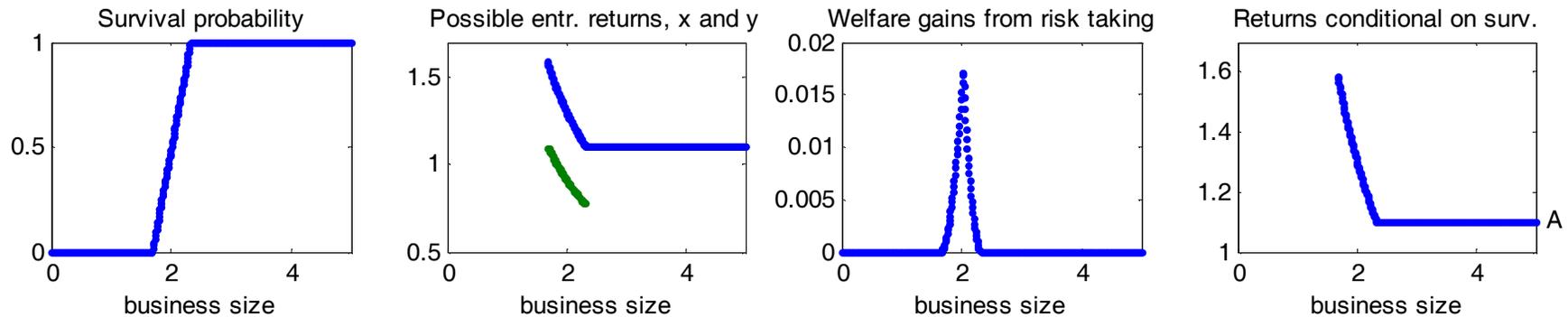


Figure C-2: Optimal project choice from the set of the projects with uniformly distributed returns, $\beta A \approx 1$

Parameters: $\beta = 0.91$, $r = 0.5$, $A = 1.1$, $\sigma = 2$, $\varphi = 0.07$ ($\beta(1+r) \approx 0.95$, $\beta A = 1.001$)

3. Choice of (x, p) from the general set of projects $\Omega(A)$ including the projects with two-point-distribution technologies



4. Choice from a set of projects with uniformly distributed returns ($\underline{A} \in [0, A]$ is the lowest possible return of the chosen project)

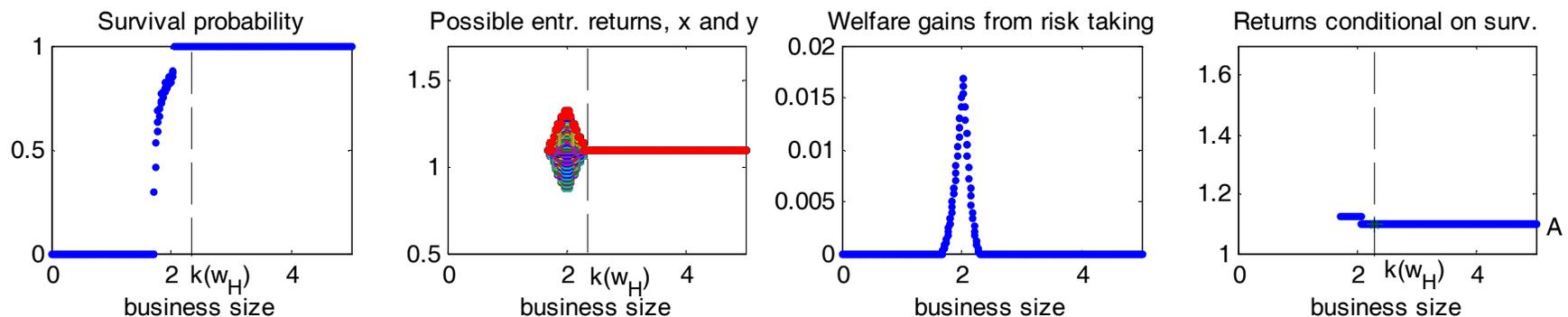


Table C-1: The effects of limiting the projects' allowed variance on the size of risk taking interval

(The allowed returns are in the range $[A-d, A+d]$)

Case 1: $\beta A = 1.045$	
dmax	risk taking interval
A	[2.15, 2.91]
less than $0.15 \cdot A$	empty
$0.20 \cdot A$	[2.31, 2.77]
$0.30 \cdot A$	[2.23, 2.87]
$0.35 \cdot A$	[2.20, 2.90]
if $d_{\max} = A$, the observed d of the chosen risky projects varies from $0.15 \cdot A$ to $0.59 \cdot A$	

Case 2: $\beta A = 1.001$	
dmax	risk taking interval
A	[1.74, 2.31]
less than $0.05 \cdot A$	empty
$0.055 \cdot A$	[1.74, 2.28]
$0.10 \cdot A$	[1.74, 2.29]
$0.15 \cdot A$	[1.74, 2.31]
if $d_{\max} = A$, the observed d of the chosen risky projects varies from $0.055 \cdot A$ to $0.21 \cdot A$	

2. Assessing the aggregate amount of risk taking

The occupational choice model presented in the paper is very stylized and cannot say much about the amount of entrepreneurial risk taking that could occur at the aggregate level. The reason is that, due to the availability of a fully safe entrepreneurial project and since there is no uncertainty in the workers' income, the economy studied in the paper is populated in the long run only by very poor workers and very rich safe entrepreneurs. Here we develop a series of examples illustrating how our model can be extended to a more general setting in which the long run the wealth distribution in the economy is nondegenerate and the agents are endogenously separated into three nonempty groups: workers, safe entrepreneurs, and risky entrepreneurs. Our major goal is to quantitatively evaluate whether the mechanism of entrepreneurial risk taking studied in the paper could potentially be important for understanding the excess volatility of entrepreneurial returns found in the data. However, it is important to point out that the quantitative exercises presented below should be treated only as the first step towards a more rigorous numerical analysis, since the parameters used in these calculations had not been properly calibrated and/or estimated.

2.1. Exogenous death and birth

One easy way to derive endogenous wealth distribution in the occupational choice framework described in the paper is to introduce exogenous death and birth into the model. Suppose that in each period a fraction ξ of the agents dies and their wealth is redistributed between the newborns. With this simple modification, the decision problem of an agent with wealth w is exactly the same as in the main text of the paper, with the only difference that the time discount factor β is adjusted to $\hat{\beta} = \beta(1 - \xi)$.

If risk taking opportunity were not available, all entrepreneurial projects were fully safe, the rate of firms' exit would be equal to the death rate ξ and the standard deviation of the observed entrepreneurial returns would be equal to zero. However, if risk taking is possible, entrants may decide to invest in risky projects, thereby driving the exit rate up and generating some excess volatility in entrepreneurial returns. In particular, if entrants' survival probability is denoted by p^* then in the steady state the exit rate x is equal to $\frac{\xi}{1 - (1 - \xi)(1 - p^*)} > \xi$.¹ That is why we can use some of our simulation results in Section 3.5 in order to evaluate the potential impact of entrepreneurial risk taking on firms' turnover and on the volatility of the observed entrepreneurial returns.

The parameters of the model are chosen as follows. The death rate $\xi = 0.03$

¹In the stationary allocation, the number of exiting incumbents should be equal to the number of surviving entrants, $\xi(1 - x) = (1 - \xi)xp^*$.

implies a 3% exit rate for large businesses and 33 years of working life for the average agent. The workers' interest rate and the rate of entrepreneurial return are set to $r = 0.04$ and $A = 1.10$ to match the returns on a risk-free bond and the average return to S&P500. The time discount factor $\beta = 0.9794$ (implying $\hat{\beta} = 0.95$) is consistent with the average 2.2% growth of consumption of safe entrepreneurs.²

For these parameter values, entering entrepreneurs invest in a risky project with the standard deviation of returns equal to 0.50 and a survival probability $p^* = 0.27$ (see the second column of Table 1 in the main text of the paper). The second row of Table B.2 below reports that such entrants' investment decision raises the firms' exit rate from 0.03 to 0.10. In addition, the high variance of the entrants' payoffs implies that the standard deviation of the projects' returns, measured across all entrepreneurs, increases from zero to 0.16.

The first and the third rows of Table C-2 below illustrate a comparative statics exercise with respect to the rate of entrepreneurial returns A . Consistently with our results in Table 1, an increase in the premium $A - 1 - r$ to entrepreneurial activity induces entrepreneurs to take more risk (as A rises, the standard deviation of entrepreneurial returns goes up). At the same time, an increase in A drives up the entrants' survival rate because it makes entrepreneurship more attractive.

Table C-2: Aggregate risk taking with exogenous death and birth
 $\sigma = 2$, $\xi = 0.03$, $r = 0.04$, $(1 - \xi)\beta(1 + r) = 0.988$, $A = 1.1$, $\phi = 0.15$.

	p	exit rate	std_{RT}	RP_{RT}	std_{ALL}
$A = 1.09$	0.26	0.11	0.38	0.10	0.12
$A = 1.10$	0.27	0.10	0.50	0.15	0.16
$A = 1.11$	0.29	0.09	0.66	0.23	0.21

On one hand, a straightforward interpretation of this numerical exercise is particularly appealing. On the other hand, the simplicity of the simulated model has a number of shortcomings, both qualitative and quantitative. In particular, the model predicts that (i) only entrants may invest in risky projects and (ii) entrepreneurs who fail once never restart new business.³ To fix this drawbacks, one may analyze a more

²The value of the time discount factor is low enough to ensure that there exist a a poverty trap region, within which workers choose a declining wealth profile and never become entrepreneurs ($\hat{\beta}(1 + r) = 0.988 < 1$). This makes it possible to calibrate a rule for wealth distribution among the newborns in order to match the the relative quantities of the workers and entrepreneurs to the ones observed in the data.

³As it is clear from the properties of the workers' value function $V_W(w)$ established in Appendix B, a risky entrepreneur who receives a low return exits into a poverty trap region.

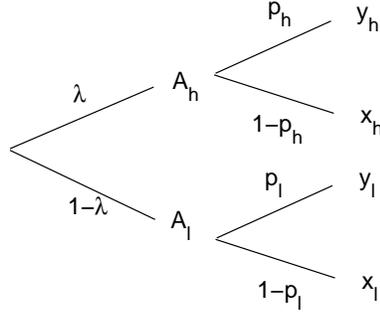


Figure C-3: The structure of available entrepreneurial projects

realistic version of our model, in which workers' wage ϕ is uncertain and a fully safe entrepreneurial project is not available. Introducing these changes may also have important quantitative implications. Additional volatility in A and ϕ reduces the kink in the value function and, therefore, is likely to decrease the amount of excessive risk taking generated by the model. This suggests that simulation results in Table B.2 overestimate potential effects of risk taking. That is why in the next exercise we introduce *ex ante* uncertainty into workers' wage and entrepreneurial returns in order to generate better qualitative predictions and improve our quantitative results.

2.2. Exogenous uncertainty in workers' and entrepreneurs' income

Suppose that a fully safe entrepreneurial project is not available. Instead, the least risky project pays off high return A_h with probability λ and low return A_l with probability $1 - \lambda$. The realizations of A are independent across entrepreneurs and across time. Entrepreneurs may take excessive risk by selecting a project from a set of all compounded lotteries generated by the least risky project. Each project in this set is a lottery over two other projects, $(x_h, p_h) \in \Omega(A_h)$ and $(x_l, p_l) \in \Omega(A_l)$, where $\Omega(A) = \{(x, p) \in [0, A] \times (0, 1)\}$ (see Figure C-3). This means that entrepreneurs have an option of investing in a project that offers the same expected return but higher variance than the least risky project.

Simultaneously with the project choice, entrepreneurs decide about their business size k . Due to the presence of uninsured risk associated with entrepreneurial activity, entrepreneurs may also want to invest in a risk-free bond that offers a rate of return r . Thus the value of an entrepreneur with wealth w is given by

$$\begin{aligned}
 V_E(w) = \max_{k, b, p_h, x_h, p_l, x_l} \{ & u(w - k - b/(1 + r)) \\
 & + \widehat{\beta}\lambda [p_h V(y_h k + b) + (1 - p_h)V(x_h k + b)] \} \\
 & + \widehat{\beta}(1 - \lambda) [p_l V(y_l k + b) + (1 - p_l)V(x_l k + b)], \} \tag{C-1}
 \end{aligned}$$

where $V(w) = \max\{V_E(w), R(w)\}$ and $R(w)$ is the value of becoming a worker.⁴

Workers' decision problem is also subject to uninsured idiosyncratic risk. Suppose that in each period workers' wage ϕ can take one of two values, ϕ_h or ϕ_l ($\phi_h > \phi_l$). The realizations of ϕ are independent both across workers and across time. Assume that ϕ_h occurs with probability δ . After ϕ is drawn, workers decide how much to save in a risk-free bond. Thus the value of a worker with wealth w is defined as

$$V_W(w) = \max_{w'_h, w'_l} \{ \delta [u(w + \phi_h - w'_h / (1 + r)) + \widehat{\beta} V(w'_h)] + (1 - \delta) [u(w + \phi_l - w'_l / (1 + r)) + \widehat{\beta} V(w'_l)] \} \quad (\text{C-2})$$

This framework generates a richer wealth dynamics than the benchmark model described in section 3 in the main text of the paper through two channels – exogenous death and birth and exogenous uncertainty of workers' and entrepreneurs' income. Theoretically, substantial volatility in A and ϕ alone would suffice to derive a non-degenerate wealth distribution and a positive turnover between workers and entrepreneurs. However, our computation suggest that, given the specification of our model, high exit/entry rates among entrepreneurs (about 10% yearly) cannot be solely explained by a variance in A .⁵ At the same time, matching the turnover rate is crucial in our analysis because most of excessive risk is taken by the young entrepreneurs. That is why exogenous death and birth are still present in (C-1) and (C-2) (the time discount factor is equal to $\widehat{\beta} = \beta(1 - \xi)$).

While computing the endogenous wealth distribution, we use the same parameters as in the previous simulation exercise, including the expected values of workers' wage, $E(\phi) = 0.15$, and entrepreneurial returns, $E(A) = 1.10$. The standard deviation of wage earnings, $std(\phi) = 0.04E(\phi)$, has been reported in Abowd and Card (1989); and $std(A) = 0.04$ matches the volatility of the CRSP index of all publicly traded equity (Moskowitz and Vissing-Jorgensen 2002).

⁴If $V_W(w)$ and $V_E(w)$ are concave and single crossing, those entrepreneurs who take excessive risk select such projects that $y_h k + b = y_l k + b = \bar{w}$ and $x_h k + b = x_l k + b = \underline{w}$. As a result, entrepreneurs with high shock A_h are more likely to survive ($p_h > p_l$). The positive correlation between the survival probability and the realization of A is not specific to the timing of the model. It would also be observed if the choice of entrepreneurial investment k were conditional on the realization of entrepreneurial return A .

⁵For the selected parameter values the standard deviation of A would have to be almost three times higher than what we see in the data in order to match a turnover rate of 10%.

Table C-3 Aggregate risk taking under uncertainty in A and ϕ
 $\sigma = 2, \xi = 0.03, r = 0.04, (1 - \xi)\beta(1 + r) = 0.988,$
 $E(A) = 1.1, std(A) = 0.04, E(\phi) = 0.15, std(\phi) = 0.04E(\phi)$

	exit rate	risky, fract.	std_{RT}	RP_{RT}	std_{ALL}	$E(A RT)$
$E(A) = 1.09$	0.092	0.149	0.19	0.035	0.08	1.073
$E(A) = 1.10$	0.091	0.129	0.32	0.083	0.12	1.088
$E(A) = 1.11$	0.084	0.112	0.45	0.142	0.15	1.104

Table C-3 summarizes the simulation results for different values of $E(A)$. Indeed, the comparison of Tables C-2 and C-3 suggests that the model with uninsured risk in A and ϕ produces less risk taking than the model with exogenous death and birth. However, even now the model still predicts a substantial amount of risk taking. In the benchmark case of $E(A) = 1.10$ almost 13% of entrepreneurs bear excessive risk. Among them 71% are entrants and the rest are older entrepreneurs who have experienced low realizations of returns.

Moskowitz and Vissing-Jorgensen (2002) report that the standard deviation of private equity is about 50%. In our model, excessive risk taking raises the standard deviation of observed entrepreneurial returns from 4% to 12% (if measured across all entrepreneurs) or to 32% (if measured only across risk takers). Heaton and Lucas (2004) estimate that entrepreneurs should be compensated with the premium of 10% in order to justify high volatility of their returns observed in the data. In their exercise, entrepreneurs have no option of exiting and taking the outside opportunity. Our model predicts that entrepreneurs bearing excessive risk would require, on average, a premium of 8.3% if the outside opportunity were not available. Though our numbers are not large enough to fully explain the high volatility of entrepreneurial returns, they are definitely quite substantial.

It remains to point out one last interesting implication of our model. The last column of Table B.3 reports that the average project returns measured across entrepreneurs bearing excessive risk is even smaller than the average returns $E(A)$ measured across all entrepreneurs. The reason is that older entrepreneurs may reenter into risk taking interval only if they receive a low realization of expected returns A_t . For example, if $E(A) = 1.10$, entrants constitute 71% of all risk takers, while 29% are older entrepreneurs with low realization $A_t = 1.06$. Thus the average return across risk takers is equal to $E(A|RT) = 0.71 \cdot 1.1 + 0.29 \cdot 1.06 \approx 1.088$. Even though such selection mechanism is specific to the structure of compound lotteries available for entrepreneurial choice, it seems to be an interesting implication of our model, especially in the light of the private equity premium puzzle.