Entrepreneurship, Human Capital, and the Misallocation of Talent

CRETE 2023

Alexandros Loukas

Washington University in St. Louis

July 11, 2023

Introduction and Motivating Facts

Entrepreneurship and Human Capital: an Understudied Relationship

- Entrepreneurship is widely recognized as a key driver of economic development for more than a century. *Who* becomes an entrepreneur is of paramount significance.

- Occupational choice determines and is being determined by macroeconomic outcomes. In the literature, it usually boils down to differences in *entrepreneurial ability/talent*; *wealth* and *access to credit*; *risk aversion*; or all of the above.

▷ Relatively scarce macro research on differences stemming from conventional human capital as proxied by educational attainment (Schultz-Becker-Mincer view). Focus on the rate of entrepreneurship along the dimension of education.

▶ Is it possible that the entrepreneurship-human capital nexus is connected to long-run economic performance?

Main Research Questions

- *Is the allocation of human capital between entrepreneurs and workers a key determinant of aggregate productivity and income? How pervasive are its implications for macro-development?*

- We already know the importance of aggregate human capital for development. This paper argues that its allocation between occupational groups is also crucial.
- The goal is to shed some light on yet another source of cross-country differences in output per worker, which are mainly due to differences in **aggregate TFP**.

In Search of a (new) Motivating Fact

▷ First pass: is there a strong cross-country correlation between the (mean-adjusted) rate of entrepreneurship for higher educated individuals and output per worker?

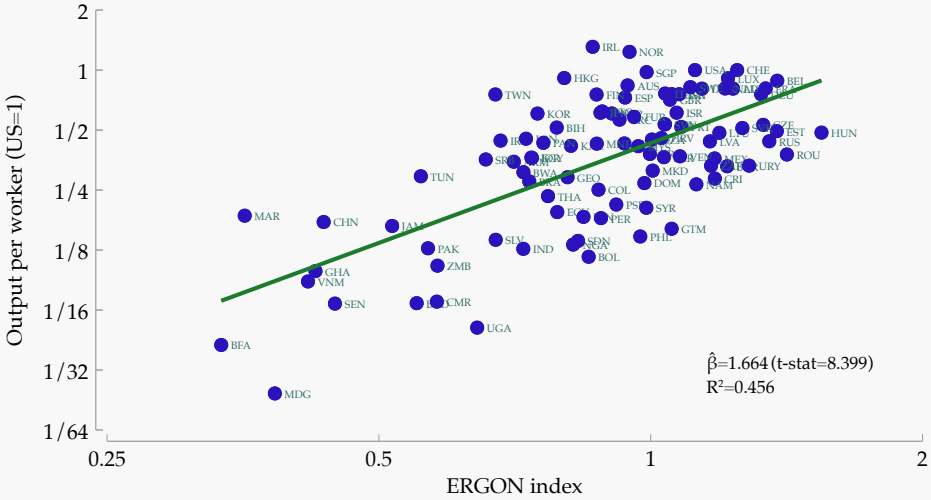
- **Data:** *Global Entrepreneurship Monitor (GEM)* adult population surveys over $2009 - 2019$. The richest and largest (more than 100 countries) ongoing study on entrepreneurship in the world. Merge with PWT 10.01 dataset.

▶ ERGON index (Entrepreneurship Rate of Graduates Over National rate):

$$\frac{ER(\text{educ} \geq \overline{S})}{ER(\text{total})} = \frac{\mu(\mathcal{E} \,|\, \text{educ} \geq \overline{S})/\mu(\Omega \,|\, \text{educ} \geq \overline{S})}{\mu(\mathcal{E})/\mu(\Omega)}$$

$\overline{S}$ = ISCED level 5, i.e., individuals with *at least* short-cycle tertiary education.
($\mathcal{E}$:= set of entrepreneurs; Ω := set of labor force; μ := measure)

Output per Worker vs ERGON index



– ERGON index (GEM): Pooled estimate for 2009-2019 according to available survey data.
– Output per worker (PWT): geometric average between first-last GEM survey year.

more

A Closer Look for the case of the US

- *Dataset:* **Survey of Consumer Finances (SCF)** over $1989 - 2019$.

- *Advantages: i)* Many nationally-representative samples that adequately capture the full distribution of economic, demographic, and occupational characteristics; *ii)* An abundance of business-related questions; *iii)* Relative over-sampling of the wealthy results in more observations of households associated with entrepreneurship.

- *Disadvantage:* Repeated cross-sections, not panel data.

▶ *Proper statistical analysis*: **repeated imputation inference (RII)**. Formal theory is well-understood; see *Rubin (1987)*, *Little (1992)*, *Van Buuren (2018)*. details

▶ Entrepreneurs := **Self-Employed Business Owners/Managers (SEBO)** details
Identification scheme expands on *Cagetti and De Nardi (2006)*.

Facts about U.S. Entrepreneurs: Education and Business Outcomes

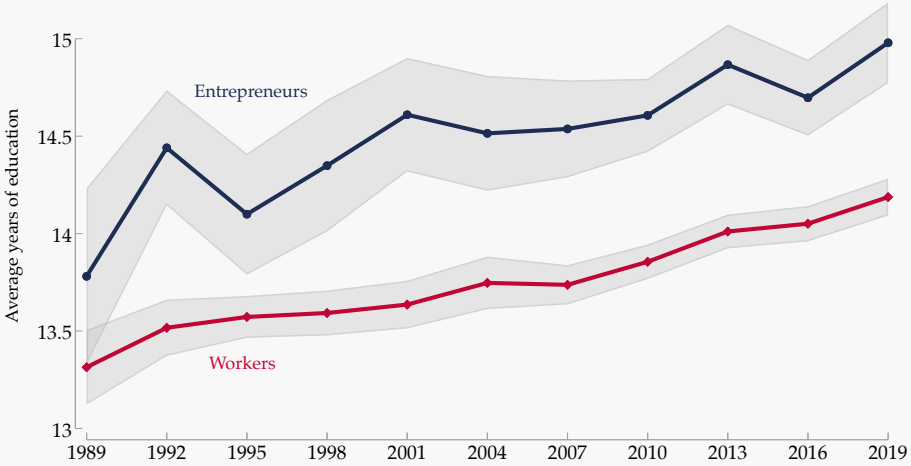
Fact 1: *Entrepreneurs are more educated than workers on average*
- slightly less than one additional year of schooling, increasing over time;
- this is a substantial difference even by international standards. details

Fact 2: *The rate of entrepreneurship by educational attainment is approximately U-shaped*
- asymmetric relationship, prolonged right branch;
- left branch falling over time, almost strictly increasing pattern in recent years.

Fact 3: *Average entrepreneurial outcomes are strongly increasing – not U-shaped – in the educational level of the firm owner/manager*
- e.g., business profits, firm employment size, net firm value per employee.

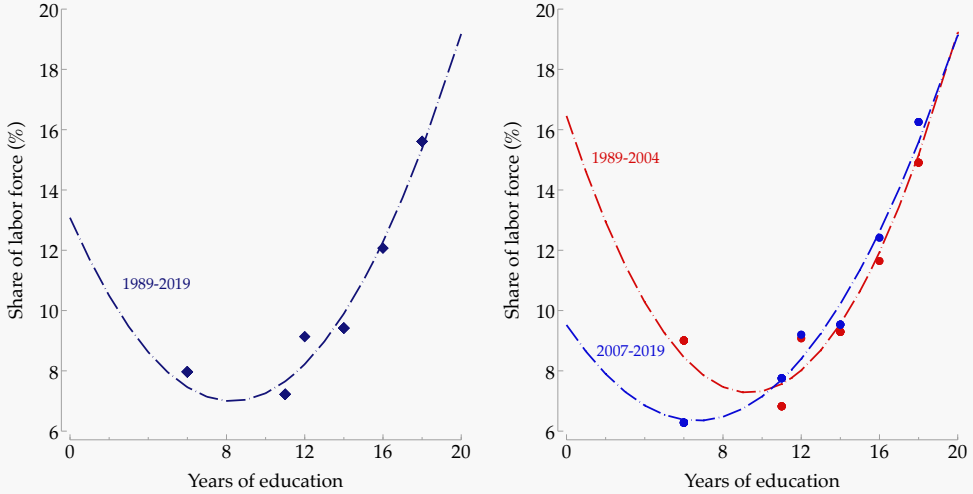
Reduced-form evidence: Descriptive results hold after controlling for a very wide range of factors and under conservative RII bootstrapped standard errors. details

Average Years of Education: Entrepreneurs vs Workers



– Shaded areas represent 95% RII confidence intervals for the sample mean.

Entrepreneurship Rate by Educational Attainment



– Fitted (weighted) quadratic regression line using full set of points; bin scatter to reduce visual clutter.

Average Entrepreneurial Outcomes by Educational Attainment

	Years of Education (pooled 1989–2019 Sample)					
Variable	[0,8]	(8,12)	12	(12, 16)	16	(16, 21]
Total business income (> 0)	53,302	64,625	82,502	93,122	188,941	260,210
Total business income	38,831	45,332	54,886	64,102	129,745	195,047
Sole proprietorship income	26,014	33,849	38,948	42,098	59,181	100,865
Total business income per hour (> 0)	23.4	35.1	54.9	64.5	114.5	229.0
Total business income per hour	17.1	24.5	37.0	42.0	78.6	170.5
Sole proprietorship income per hour	12.5	19.9	28.1	33.3	34.2	60.2
Firm employment size (> 0)	5.4	6.3	8.2	16.8	49.1	80.9
Firm employment size	3.4	4.1	5.7	10.9	35.5	62.1
Firm net value per employee	85,725	90,667	137,249	140,997	209,560	233,104
Main business ownership share	0.92	0.91	0.91	0.89	0.82	0.77
Potential experience	35.1	33.4	30.7	27.8	26.9	27.5
Self-employed experience	14.6	13.6	12.7	11.0	11.1	12.8
Prior worker experience	20.5	19.8	18.1	16.8	15.7	14.8

Theoretical & Quantitative Contributions

- Rationalize empirical findings within a DGE framework. I propose an elaboration of a heterogeneous agent model with endogenous *schooling* and *occupational choice*. Human capital can be a factor of production for both workers and entrepreneurs.
- Theoretical predictions align with empirics under weak restrictions; quite flexibile in practice. Relevant parameters can be identified and estimated from surveys.

▷ Quantitative exploration: potentially sizeable losses due to **inadequate complementarity between entrepreneurial talent and human capital**.

▶ First-order consequences for macro-development: *ceteris paribus* shifts in the entrepreneurship-education nexus generate up to about -40% TFP vis-à-vis the US.

▶ Far-reaching ramifications when considering variations in average years of schooling, which can alone account for up to -70% output per worker.

Some Related Literature

- **Empirical:** There is some empirical literature on entrepreneurship-education, see surveys by *Van der Sluis et al. (2003, 2004)*, but most authors do not deliberate or examine nonlinearities. Studies reporting some U-shaped relationship: *Blanchflower (2000)*, *Hamilton (2000)*, *Hipple (2010)*.

- Notable exception (empirics and theory): *Poschke (2013)*. Evidence for U-shaped relationship using the NLSY79. Roy-type model together with labor market search; completely different approach from this paper. Somewhat related: *Poschke (2018)*.

- **Entrepreneurs & Human Capital:** *Bhattacharya et al. (2013)*, *Mestieri et. al (2017)*.

- **Macro & Entrepreneurship:** What I mean by "established" DGE framework: *Quadrini (2000)*, *Cagetti and De Nardi (2006)*, *Kitao (2008)*, *Buera et al. (2011)*, *Buera and Shin (2013)*, *Midrigan and Xu (2014)*, *Moll (2014)*, among others.

The Model Economy

Standard Setup in Continuous Time with 3D Heterogeneity

- Infinite-horizon economy set in continuous time under the usual conditions for $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. The sample space is the set $\Omega = (\mathcal{A} \times \mathcal{Z} \times \mathcal{H}) \subset \mathbb{R}^3_+$.

- Measure-one continuum of infinitely lived agents, heterogeneous in terms of **assets**, $a \in \mathcal{A}$; **entrepreneurial ability**, $z \in \mathcal{Z}$; and **learning attitude**, $\theta \in \Theta$.

▷ *Endogenous states*:
– **human capital** $(h \in \mathcal{H})$, determined via choice of years of schooling, $S(\theta) \geq 0$.
– assets/wealth (a_t), determined via forward-looking saving behavior, $\forall t > S$.

▷ *Exogenous state*: entrepreneurial ability (z_t) folllows an independent Itô process.

▷ Earned income $\mathcal{Y}_t(h, a, z) = \max\{\text{labinc}_t, \text{profit}_t\}$ is determined *endogenously* by occupational choice: **entrepreneur/producer** vs **worker/wage-earner**.

Preferences, Budget Constraints, Intertemporal Optimization

- Inelastic labor supply; standard preferences from consumption flows for $t \geq S$:

$$V_t(h,a,z) = \max_{\substack{c_t \in \mathcal{C} \\ S \geq 0}} \int_0^S e^{-\rho t}\phi(B,\theta)\,dt \;+\; \mathbb{E}_S\left[\int_S^\infty e^{-\rho t}u(c_t)\,dt\;|\mathcal{F}_S\right] \tag{1}$$

$$da_t = (\mathcal{Y}_t(h,a,z) + r_t a_t - c_t)\,dt; \quad a_t \geq 0,\ \forall t > S \tag{2}$$

$$dz_t = \widetilde{\mu}_z(z,t)\,dt + \widetilde{\sigma}_z(z,t)dW_t; \quad W_t \perp \Theta, \forall t \geq S \tag{3}$$

$$h = \exp\left(\psi(S)\right); S = 0,...,20;\ \psi'(S) > 0, \forall t \geq S \tag{4}$$

$\phi(.)$ is flow utility from attending school, e.g., *Schultz (1963)*, *Bils and Klenow (2000)*, with $\partial\phi(B,\theta)/\partial\theta > 0$ and $\partial\phi(B,\theta)/\partial B > 0$. Higher learning ability (θ) or better aggregate schooling "quality" (B) makes pursuing education more fulfilling.

$\Longrightarrow$ State of the economy is the endogenous joint distribution $G_t(h,a,z)$, $t \in [S,\infty)$.

Labor and Financial Markets; Entrepreneurs and Production

☐ The labor market is perfectly competitive. Labor supply/demand denominated in *units of human capital*. Perfect substitutability $\Rightarrow$ *labor income* $= w_t h_t, \forall h \in \mathcal{H}$.

☐ Financial markets are fully liquid, but *incomplete*. Perfectly competitive financial intermediaries receive deposits from savers and create capital loans for firms. ZPF $\Longrightarrow R_t = r_t + \delta$. Common *borrowing constraint*: $a_t \geq 0, \ \forall t \geq 0$.

- Entrepreneurs hire k units of *physical capital* and ℓ units of *human capital* to maximize profits and produce a homogeneous consumption good (numéraire):

$$y(z, h, k, \ell) = \zeta(z, h) \left(k^{\alpha} l^{1-\alpha}\right)^{1-\nu} \tag{5}$$

- Due to limited contract enforceability, entrepreneurs face a *collateral constraint*:

$$k_t \leq \frac{1}{1-\vartheta} a_t, \quad \vartheta \in [0, 1] \tag{6}$$

Entrepreneurs and Intratemporal Technology Choice

- Hypothesis: there exists a **disembodied technology** that coalesces entrepreneurial talent (z) and human capital (h) to form the fundamental "effective productivity" $\zeta(z,h)$. Positive complementarity between z and h $(\omega > 0)$.

▶ Entrepreneurs face a costly intratemporal **technology adoption choice**:

$$\zeta(z,h) := \begin{cases} z & \text{at no cost} \\ zh^{\omega} & \text{at cost } \kappa \cdot y(), \kappa \in (0,1) \end{cases} \tag{7}$$

▷ *Ceteris paribus*, **more productive** = **more talented** or **more educated**. Entrepreneurial human capital as an *additional factor of production*.

∗ A complementary view on the importance of **organization capital** in the production process, e.g., *Prescott and Visscher (1980)*, *Atkenson and Kehoe (2005)*.

Recursive Formulation: Stationary HJB and KF Equations

$$\begin{aligned}
\rho V &= \max_{c \in \mathcal{C}} \; u(c) + \; V_a \left(\mathcal{Y} + ra - c\right) + V_z \widetilde{\mu}_z(z) + \frac{1}{2} V_{zz} \, \widetilde{\sigma}_z^2(z) \\
&= \max_{c \in \mathcal{C}} \; u(c) + \mathcal{A} \, V(h, a, z) \qquad \text{(HJB equation)} \qquad (8) \\
0 &= -\frac{\partial}{\partial a}[\widetilde{s} \, g] - \frac{\partial}{\partial z}[\widetilde{\mu}_z(z) \, g] + \frac{1}{2} \frac{\partial^2}{\partial z^2}[\widetilde{\sigma}_z^2(z) \, g] + g_h(h) \\
0 &= \; \mathcal{B} \, g(h, a, z) + g_h(h) \qquad \text{(KF equation)} \qquad (9)
\end{aligned}$$

plus *boundary conditions*: TVC; SCBC for $\underline{a}$; NBC for $(\underline{z}, \overline{z})$; BC such that $\mathcal{B} = \mathcal{A}^*$.

boundary conditions

■ Market clearing conditions: capital/labor demand = asset/human capital supply.

details

• System of weakly coupled, nonlinear, 2nd order PDEs. Classical solutions are not guaranteed to hold because of non-convexities $\Longrightarrow$ use powerful theory of constrained *viscosity solutions*.

∗ Existence and uniqueness of a **Mean Field Game** equilibrium without common noise: ✓

Theoretical Findings

Entrepreneurship Rate by Human Capital: it's (almost) all about ω

Proposition: For any sequence of equilibrium prices $\{w_t, r_t\}_{t\in\mathbb{R}_+}$, joint distributions $\{G_t(h,a,z)\}_{t\in\mathbb{R}_+}$, adoption cost $\kappa \in [0,1]$, and collateral constraint $\vartheta \in [0,1]$, there exists a non-increasing measurable function $\chi(\vartheta) : [0,1] \mapsto [\nu, \alpha + \nu - \alpha\nu]$, such that the entrepreneurship rate by human capital is:

- strictly decreasing $\forall h \in \mathcal{H}$ *almost surely*, if and only if $\omega < \chi(\vartheta)$;
- strictly decreasing $\forall h < \hbar(\kappa,\omega)$ and strictly increasing $\forall h \geq \hbar(\kappa,\omega)$ *almost surely*, if and only if $\omega > \chi(\vartheta)$;
- strictly decreasing $\forall h < \hbar(\kappa,\omega)$ *almost surely*, and anything goes $\forall h \geq \hbar(\kappa,\omega)$, if and only if $\nu < \omega \leq \chi(\vartheta)$.

NOTE: $\hbar(\kappa,\omega) = \left(\frac{1}{1-\kappa}\right)^{1/\omega}$

some identification/estimation advantages

Entrepreneurial Outcomes by Human Capital

Proposition: For any sequence of equilibrium prices $\{w_t, r_t\}_{t\in\mathbb{R}_+}$, joint distributions $\{G_t(h,a,z)\}_{t\in\mathbb{R}_+}$, adoption cost $\kappa \in [0,1]$, and collateral constraint $\vartheta \in [0,1]$, the expected value of any entrepreneurial outcome X (net production plans, profits, capital and labor demands) is strictly increasing in human capital *almost surely*.

That is, $\frac{\partial \mathbb{E}[X|h,\mathcal{E}]}{\partial h} > 0 \; a.s. \; \forall h \in \mathcal{H}$.

- Result is guaranteed to hold irregardless of the value of ω; why?

– Without technology adoption, self-selection in terms of z prevails.
– With technology adoption, human capital induces first-order stochastic dominance over *effective productivities*: $\mathbb{E}\left[\zeta(z,h'')\right] \geq \mathbb{E}\left[\zeta(z,h')\right], \forall h'' \geq h'$.

Production-Side Aggregation

▶ The (net) production side of the economy aggregates and total output p.w. becomes:

$$Y = \underbrace{\mu(\mathcal{E})^{\nu}}_{\text{TFP extensive margin}} \underbrace{\frac{\mathbb{E}\left[(z\tilde{h})^{\frac{1}{\nu}}\ (\hat{r}(h,a,z)+\delta)^{-\frac{\hat{\alpha}}{\nu}} \,\middle|\, \mathcal{E}\right]^{\hat{\alpha}+\nu}}{\mathbb{E}\left[(z\tilde{h})^{\frac{1}{\nu}}\ (\hat{r}(h,a,z)+\delta)^{-\frac{(\hat{\alpha}+\nu)}{\nu}} \,\middle|\, \mathcal{E}\right]^{\hat{\alpha}}}}_{\text{TFP intensive margin}} K^{\hat{\alpha}} H^{1-\hat{\alpha}} \tag{10}$$

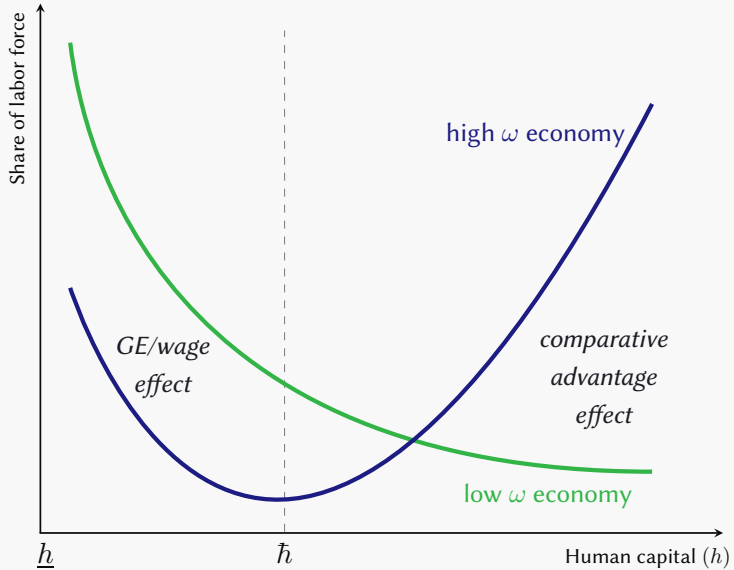
$$Y_* = \mu(\mathcal{E}_*)^{\nu}\ \mathbb{E}\left[(z\tilde{h})^{\frac{1}{\nu}} \,\middle|\, \mathcal{E}_*\right]^{\nu} K_*^{\hat{\alpha}} H_*^{1-\hat{\alpha}} \qquad (\text{if } \vartheta \to 1) \tag{11}$$

▷ Development accounting decomposition à la *Hall and Jones (1999)*, *Caselli (2005)*:

$$Y = \underbrace{A(h,a,z,\mathcal{E})^{\frac{1}{1-\hat{\alpha}}}}_{\text{TFP contribution}} \left(\frac{K}{Y}\right)^{\frac{\hat{\alpha}}{1-\hat{\alpha}}} H^{\frac{1-\hat{\alpha}-\nu}{1-\hat{\alpha}}} \qquad \hat{\alpha} := \alpha(1-\nu) \tag{12}$$

Notation: $\tilde{h} := \underset{h\le\hbar}{1} + \underset{h>\hbar}{1}\left(\frac{h}{\hbar}\right)^{\omega}$

Entrepreneurship Rate along h: High vs Low Complementarity



Quantitative Exploration

Calibration Strategy and Numerical Solution

- Length of one period = one year to allow for internal consistency with U.S. aggregate, survey, and firm-level data. Target sample averages for $1989 - 2019$.
  - Reduce degrees of freedom by preassigning values to conventional parameters. details
  - Estimate parameters that can be inferred from data in the empirical section. details
  - The model is then disciplined via the joint calibration of only six parameters that are relatively nonstandard, and thus of principal interest. details

▶ **Numerical solution:** implicit upwind finite difference scheme that is stable, consistent, and monotone in the sense of *Barles and Souganidis (1991)*; yields unique viscosity solutions to HJB and KFE. I follow and expand on **Achdou et al. (2022)**.

Calibration Results: Targeted Moments, Parameters

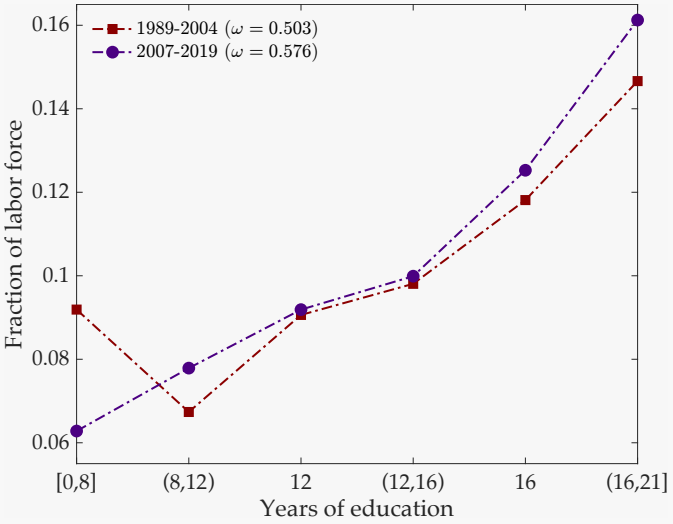
Targeted Moments	Data	Model	Source
Entrepreneurship rate in the labor force	0.105	0.105	SCF
Share of total income received by top 20%	0.499	0.492	CPS ASEC
Dispersion of (log) firm TFP	0.200	0.200	*Haltiwanger (2011)*
Firm entry/exit rate	0.110	0.110	BDS
External finance-to-output ratio	1.650	1.650	GFDD/World Bank
Real rate of return on capital	0.040	0.040	Standard

Calibrated Parameters	Value	Assigned & Estimated	Value
Span-of-control parameter (ν)	0.232	Elasticity of output wrt capital (α)	0.360
Complementarity between z and h (ω)	0.537	Capital depreciation rate (δ)	0.060
Dispersion of entrep. ability shock (σ_z)	0.210	Relative risk aversion coefficient (γ)	1.500
Autocorrelation of entrep. ability z ($e^{-\varphi_z}$)	0.908	Mean (log) entrep. ability (μ_z)	0.000
Collateral constraint (ϑ)	0.870	$\log h \sim \mathcal{N}(\mu_h, v_h)$	$1.337, 0.051$
Subjective discount rate (ρ)	0.088	Technology adoption cost (κ)	0.418

Calibration Results: Some Untargeted Moments

Untargeted Moments	Data	Model	Source
Fraction of wealth owned by entrepreneurs	0.437	0.441	SCF
Fraction of entrepreneurs in top wealth decile	0.394	0.401	SCF
Ratio of mean assets; entrepreneur to worker	6.876	6.753	SCF
Share of total income received by the middle quintile	0.148	0.146	CPS ASEC
Job destruction rate: deaths	0.047	0.041	BDS
Fraction of entrepreneurs with $16+$ years of education	0.438	0.438	SCF
Fraction of entrepreneurs with $12-$ years education	0.329	0.329	SCF
Firm size distribution:			
Employment size: $1-9$	0.769	0.775	BDS
Employment size: $10-19$	0.115	0.129	BDS
Employment size: $20-99$	0.096	0.093	BDS
Employment size: $100+$	0.020	0.003	BDS
Entrepreneurship rate by educational attainment:			
Years of education: $[0-8]$	0.080	0.080	SCF
Years of education: $(8-12)$	0.072	0.072	SCF
Years of education: 12	0.091	0.091	SCF
Years of education: $(12-16)$	0.094	0.099	SCF
Years of education: 16	0.121	0.121	SCF
Years of education: $(16, 21]$	0.156	0.156	SCF

Higher Complementarity (ω) Over Time



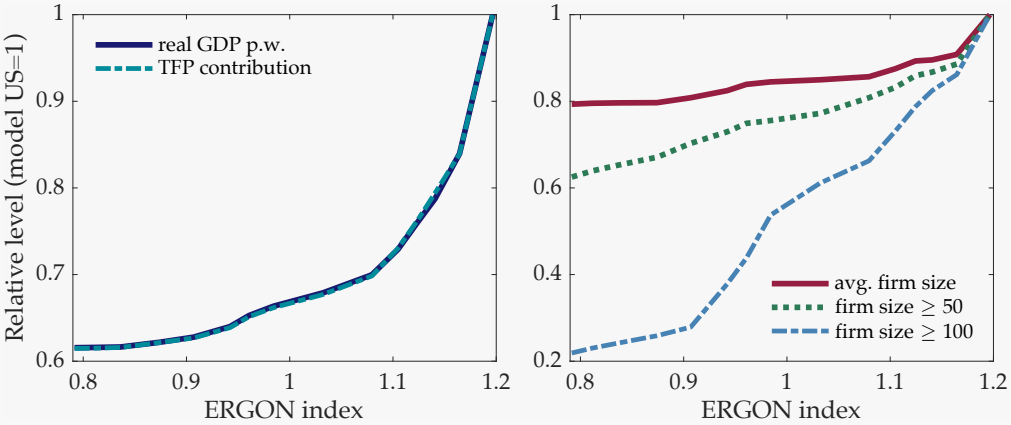
- vary ω accordingly to estimated $\hat{\hbar}$ for the two consecutive subsamples

- small changes in ω can have large aggregate effects

- real GDP p.w. $+12\%$, almost entirely due to aggregate TFP;

- wage pass-through almost 1-to-1; essentially no change in r.

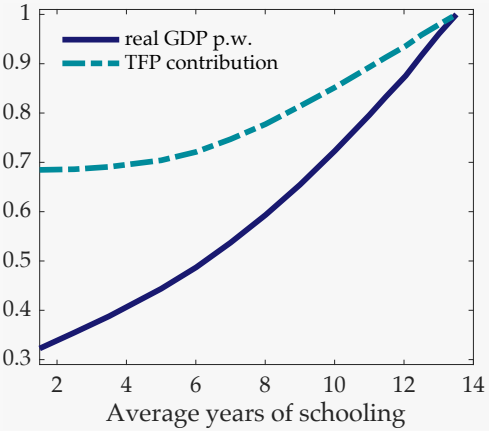
Implications for Macro-Development (I)

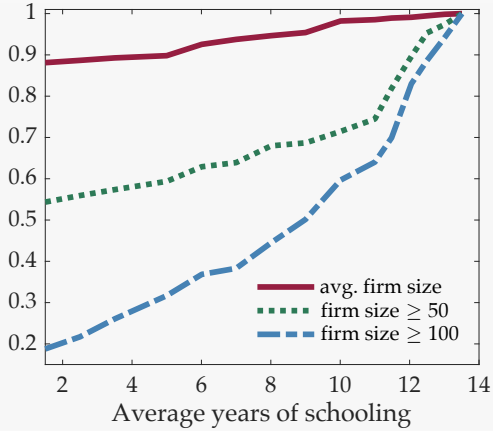


– Model output: stationary equilibria obtained by varying ω from 0.576 (U.S. 2007-2019) to 0.310.

related

Implications for Macro-Development (II)





– Model output: stationary equilibria obtained by varying B from 1.000 (US 2007-2019) to 0.139.

More Implications for Macro-Development (in progress)

- How prolonged is the impact of financial frictions in the presence of a different entrepreneurship-human capital nexus?
  - Economies with higher ERGON index (relatively more highly educated entrepreneurs) benefit **more** from relaxing the collateral constraint.

- Development accounting: How much of the cross-country differences in output p.w. and TFP can the model explain if we vary *i)* only ω; *ii)* ω and μ_h; *iii)* ω, μ_h, and θ. Target *i)* ERGON index; *ii)* ERGON index and average y.o.s.; *iii)* ERGON index, average y.o.s., external finance-to-GDP ratio.

- Transitions and some dynamics (MIT shocks).

Conclusion

Taking Stock

- In this setup, the allocation of human capital between workers and entrepreneurs is interconnected with aggregate TFP and income.

- Higher entrepreneurship rate for more educated individuals is mapped to enhanced productive capacities for firms, which in turn determines long-run performance.

- Lower entrepreneurship rate for less educated individuals is also mapped to favorable macroeconomic outcomes through general equilibrium.

- The role of aggregate human capital can be even more important than commonly thought, as it may contribute directly to TFP (and capital) formation.

- A small step closer to better understand cross-country income differences.

Thank you for your time and attention

(enjoy the summer break!)

Appendix / Additional Slides

Output per worker vs ERGON index: OLS with Additional Controls

Covariates	Dependent variable: output p.w./US (log)		
	(1)	(2)	(3)
ERGON index (log)	1.664^{***}	1.149^{***}	0.527^{***}
	(0.198)	(0.230)	(0.201)
Entrepreneurship rate (log)		-0.468^{***}	-0.270^{***}
		(0.116)	(0.095)
PWT human capital index (log)			2.040^{***}
			(0.233)
Observations	96	96	89
Adjusted R^2	0.456	0.531	0.735

Note: Robust standard errors in parentheses; *** $p < 0.01$.

back to lecture

Some Statistical Considerations using the SCF (I)

- Data will involve some missing responses, either because of unanswered questions or for confidentiality/non-disclosure reasons. The SCF staff imputes missing values using a statistical procedure called *multiple imputation*; see Kennickell (2017).
- Imputations stored as five successive implicates of each data record: $N^m = 5 \times N$. It is critical not to treat all five implicates as independent observations in analysis.
- Datasets also come with 999 bootstrap replicate draws and respepective weights for each variable to simulate the complex sampling scheme and replicate process.

back to the dataset slide

Some Statistical Considerations using the SCF (II)

Throughout this study:

- All sample estimates are calculated under RII using the appropriate SCF sampling weights, as well as all 5 SCF implicates for every observation and all 999 SCF bootstrap replicate draws and weights.

- Same applies to covariance matrices, but there's even more. Bootstrapped std. errors should account for both *within variability* – the sampling variance within each implicate, and *between variability* – the variance contributed across implicates.

- The equality of means/proportions is assessed with a weighted, two-sided, unequal-variance hypothesis test (Welch's t-test).

back to the dataset slide

More on Proper Statistical Inference using the SCF (I)

- To obtain the point estimate for any parameter, compute the average point estimate *across* the five implicates. For instance, the correct sample mean of a variable should equal the average of the means of the five implicates, or the correct $\hat{\beta}$ coefficient vector of an OLS regression should be computed as the average of the estimated $\hat{\beta}$ vectors after running the model five times using each of the implicates.

- For the multivariate case, let $\widehat{\mathbf{Q}}$ be a $k \times 1$ vector of (properly weighted) estimates of the parameter vector $\mathbf{Q}$ that would have been obtained if no data were missing. Each imputed data set $m = 1, \ldots, M$ admits an estimate of $\widehat{\mathbf{Q}}$, denoted $\widehat{\mathbf{Q}}_m$, along with a standard error $\sqrt{\mathbf{W}_m}$. The RII estimate of $\mathbf{Q}$ based on M imputed datasets is,

$$\bar{\mathbf{Q}} = \frac{1}{M} \sum_{m=1}^{M} \widehat{\mathbf{Q}}_m$$

back to the dataset slide

More on Proper Statistical Inference using the SCF (II)

- Within variability is given by the sample variance within the imputed data sets,

$$\bar{\mathbf{W}} = \frac{1}{M-1} \sum_{m=1}^{M} \widehat{\mathbf{W}}_m$$

- Between variability is given by the sample variance of the five parameters computed within each implicate,

$$\mathbf{B} = \frac{1}{M-1} \sum_{m=1}^{M} \left(\widehat{\mathbf{Q}}_m - \bar{\mathbf{Q}}\right) \left(\widehat{\mathbf{Q}}_m - \bar{\mathbf{Q}}\right)'$$

The estimate of the total covariance matrix of the parameter vector is given by

$$\mathbf{V} = \bar{\mathbf{W}} + \left(1 + \frac{1}{M}\right) \mathbf{B}$$

back to the dataset slide

Who are the Entrepreneurs in the data?

Identification inspired by *Cagetti and De Nardi (2006)*, but with some differences; e.g., I take into account both the respondent (R) and the spouse/partner (S/P) instead of just the arbitrarily defined "household head". Consider labor force participants only.

► **Four conditions** (jointly) to be classified as an entrepreneur:

1. Either R or S/P engages in a form of self-employment as their primary occupation;
2. Owns or shares ownership in at least one privately-held business;
3. Has an active management role in at least one business;
4. The net value of actively managed businesses in greater than zero.

▷ I refer to entrepreneurs as **Self-Employed Business Owners/Managers**, or **SEBO**.

back to lecture

Comparisons using the Barro-Lee Dataset to Gain Perspective

▷ Consult the Barro-Lee educational attainment dataset for population aged 25-64 over the period $1995 - 2015$. Compare average years of schooling.

- Ireland, Denmark, Sweden, Germany, Australia, Japan, **vs**
  Hungary, Poland, Lithuania, Latvia, Kazakhstan, Ukraine,
  all pairwise differences $\leq \mathbf{0.9}$ years of schooling.

- Tunisia, Egypt, Namibia, **vs**
  Republic of the Congo, Syria, and Cameroon,
  all pairwise differences $\leq \mathbf{0.5}$ years of schooling.

back to lecture

From Descriptive to Reduced-Form Evidence

• Differences likely to stem from sources aside from education, e.g., experience, health, risk tolerance, demographics. Apparent impact of y.o.s could be fairly biased upward.

▶ Consider regressions (RII & weighted) of the form:

$$y_i = \beta_0 + \beta_1 \cdot educ_i + \beta_2 \cdot educ_i^2 + X_i\boldsymbol{\gamma} + \Gamma_s + \Gamma_o + \Gamma_s \times \Gamma_o + \Gamma_t + \varepsilon_i$$

vector X_i includes a rich set of controls; sector fixed effects (Γ_s); occupation fixed effects (Γ_o); their interactions $(\Gamma_s \times \Gamma_o)$; year fixed effects (Γ_t).

$\Rightarrow$ Do the descriptive results still hold after controlling for a wide range of factors and under conservative RII bootstrapped standard errors? YES.

back to lecture

RII-Probit Regressions: Probability of Being an Entrepreneur

	Dependent variable: SEBO = 1						
COVARIATES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Education	0.047^{***}	0.028^{***}	0.017^{***}	-0.080^{***}	-0.036^{***}	-0.037^{***}	-0.039^{***}
Education^2 ($\times 100$)				0.347^{***}	0.246^{***}	0.248^{***}	0.191^{***}
Potential experience					0.020^{***}	0.033^{***}	
Potential experience^2 ($\times 100$)						-0.022^{***}	
Self-Employed experience							0.073^{***}
Worker experience							-0.002^{**}
ADDITIONAL CONTROLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Annual labor supply (log)			0.018	0.018	0.086^{***}	0.076^{***}	0.125^{***}
Ever received inheritance $(0/1)$			0.245^{***}	0.244^{***}	0.156^{***}	0.156^{***}	0.125^{***}
Expects to receive inheritance $(0/1)$			-0.010	-0.009	0.080^{***}	0.078^{***}	0.073^{***}
Risk willingness $(1-4)$			0.133^{***}	0.136^{***}	0.168^{***}	0.166^{***}	0.174^{***}
Health $(1-4)$		0.130^{***}	0.119^{***}	0.116^{***}	0.153^{***}	0.155^{***}	0.132^{***}
Married $(0/1)$		0.499^{***}	0.491^{***}	0.490^{***}	0.474^{***}	0.457^{***}	0.511^{***}
Male $(0/1)$		0.106^{***}	0.088^{***}	0.078^{***}	0.083^{***}	0.093^{***}	0.100^{***}
Black $(0/1)$		-0.433^{***}	-0.393^{***}	-0.391^{***}	-0.346^{***}	-0.353^{***}	-0.302^{***}
Hispanic $(0/1)$		-0.417^{***}	-0.366^{***}	-0.414^{***}	-0.291^{***}	-0.302^{***}	-0.254^{***}
Other non-white $(0/1)$		-0.083^{***}	-0.047	-0.056	0.011	0.010	0.010
McFadden's adjusted R^2	0.117	0.152	0.160	0.163	0.186	0.187	0.331
Observations	40,413	40,413	40,413	40,413	40,413	40,413	40,413

RII-OLS Regressions: SEBO Hourly Business Income (> 0)

	Dependent Variable: log hourly business income (> 0)						
COVARIATES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Education	0.079^{***}	0.076^{***}	0.075^{***}	-0.042	-0.040	-0.033	-0.041
Education^2 $(\times 100)$				0.417^{***}	0.393^{***}	0.346^{***}	0.368^{***}
Worker experience	0.007^{***}	0.008^{***}	0.009^{***}	0.009^{***}	0.006	0.003	0.002
Worker experience^2 $(\times 100)$					0.007	0.014	0.016
Self-empl. experience	0.018^{***}	0.017^{***}	0.019^{***}	0.018^{***}	0.046^{***}	0.039^{***}	0.028^{***}
Self-empl. experience^2 $(\times 100)$					-0.070^{***}	-0.058^{***}	-0.043^{***}
Business employment size (log)						0.283^{***}	0.176^{***}
Net business value (log)							0.167^{***}
ADDITIONAL CONTROLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ever received inheritance $(0/1)$			-0.081^{*}	-0.080^{*}	-0.089^{**}	-0.072^{*}	-0.045
Expects to receive inheritance $(0/1)$			-0.035	-0.032	-0.041	-0.008	0.014
Inherited any business $(0/1)$			0.221^{**}	0.235^{**}	0.229^{**}	0.181^{**}	0.051
Risk willingness $(1-4)$			0.073^{***}	0.079^{***}	0.080^{***}	0.062^{***}	0.027
Health $(1-4)$		0.140^{***}	0.137^{***}	0.135^{***}	0.122^{***}	0.107^{***}	0.097^{**}
Male $(0/1)$		0.106^{**}	0.104^{**}	0.089^{**}	0.088^{*}	0.014	-0.065
Black $(0/1)$		-0.240^{***}	-0.250^{**}	-0.244^{***}	-0.239^{***}	-0.276^{***}	-0.216^{**}
Hispanic $(0/1)$		-0.003	-0.007	-0.044	-0.055	-0.114	-0.122
Other non-white $(0/1)$		0.078	0.070	0.057	0.047	0.047	0.046
Adjusted R^2	0.149	0.154	0.158	0.159	0.165	0.215	0.249
Observations	$7,810$	$7,810$	$7,810$	$7,810$	$7,810$	$7,810$	$7,810$

More Facts about U.S. Entrepreneurs: Wealth, Income, Demographics

Fact 4: *Entrepreneurs form a small group of working population* figure
- hovering around 10.5% of the labor force; relatively stable over time.

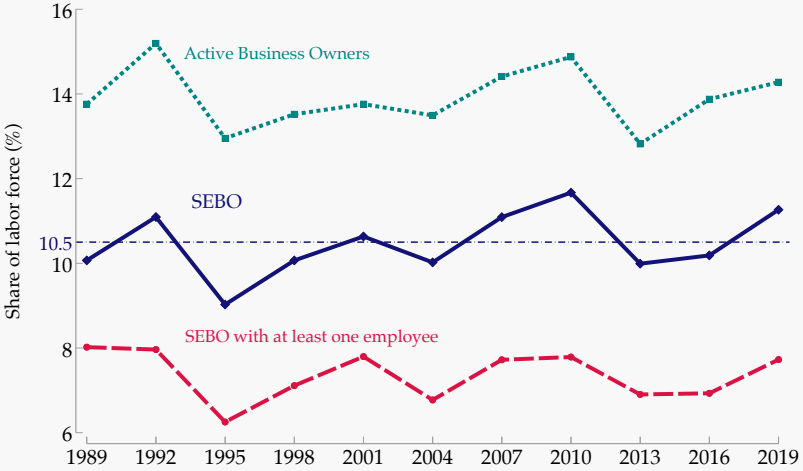
Fact 5: *Entrepreneurs hold a large amount of wealth* figure
- about 44% of total net worth in the labor force;
- dramatically overrepresented in higher percentiles of the wealth distribution.

Fact 6: *Entrepreneurs earn a considerable amount of income* figure
- at least 23% of total income in the labor force;
- significantly overrepresented in higher percentiles of the income distribution.

Additional facts about differences in economic and demographic characteristics across entrepreneurs and workers in the U.S. labor force. table

back to lecture

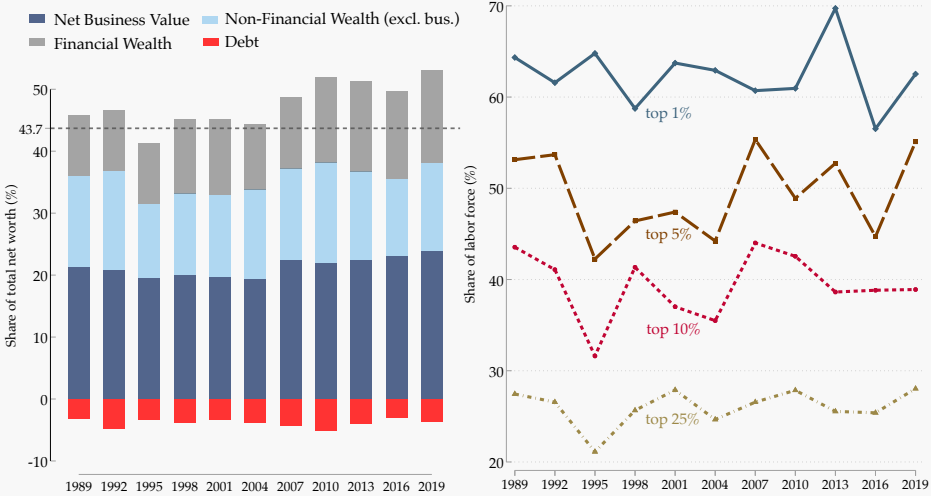
U.S. Entrepreneurship Rate



- SEBO = preferred definition
- results largely in line with U.S. Census/BLS data
- compare with *Hipple (2010)* using CPS: 10.4% average over 1990 – 2009.

back to the lecture

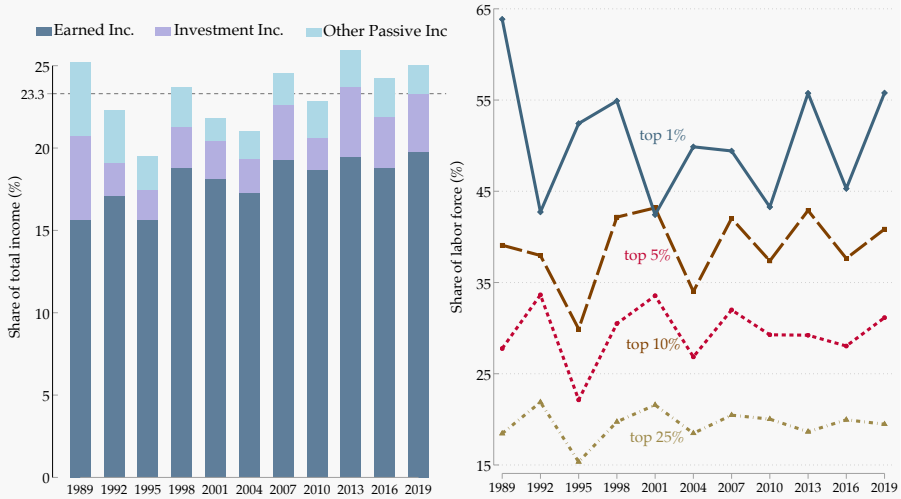
Systematic Wealth Concentration by Entrepreneurs



– Share of U.S. Wealth Held by SEBO (Left panel) ; Fraction of SEBO in Selected Wealth Percentiles (Right Panel)

back to the lecture

Systematic Income Concentration by Entrepreneurs



– Share of U.S. Income Held by SEBO (Left panel) ; Fraction of SEBO in Selected Income Percentiles (Right Panel)

back to the lecture

Descriptive Statistics: Entrepreneurs vs Workers

	Pooled 1989–2004 Sample			Pooled 2007–2019 Sample		
Variable	Workers	SEBO	p-value (diff.)	Workers	SEBO	p-value (diff.)
Education	13.6	14.3	$<.01$	13.9	14.8	$<.01$
Potential experience	22.3	26.5	$<.01$	24.2	30.7	$<.01$
Worker experience	21.2	15.7	$<.01$	23.2	17.4	$<.01$
Self-emp. experience	1.1	10.8	$<.01$	1.0	13.3	$<.01$
Annual labor supply	2,142	2,418	$<.01$	2,109	2,230	$<.01$
Health (1-4)	3.19	3.31	$<.01$	3.08	3.20	$<.01$
Risk willingness (1-4)	1.87	2.14	$<.01$	1.88	2.17	$<.01$
Unemployed S/P (0/1)	0.25	0.22	$<.01$	0.26	0.23	$<.01$
Male	0.68	0.77	$<.01$	0.65	0.73	$<.01$
Married	0.65	0.82	$<.01$	0.63	0.82	$<.01$
White/Caucasian	0.74	0.89	$<.01$	0.64	0.79	$<.01$
Black/African-American	0.12	0.03	$<.01$	0.14	0.05	$<.01$
Hispanic/Latino	0.09	0.03	$<.01$	0.12	0.06	$<.01$
Other non-white	0.05	0.05	.96	0.10	0.10	.79

back to the lecture

Recursive Formulation: Boundary Conditions

$$V_a(h,\underline{a},z) \geq u'\left(\mathcal{Y}(h,\underline{a},z)\right) \quad \forall z,h \qquad \text{(SCBC for } \underline{a}=0) \tag{13}$$

$$V_z(h,a,\underline{z}) = V_z(h,a,\bar{z}) = 0 \quad \forall a,h \qquad \text{(NBC for } \underline{z} \text{ and } \bar{z}) \tag{14}$$

$$\lim_{t\to\infty} e^{-\rho t} a_t \, V_a(h,a,z) = 0 \quad \forall h,a,z \qquad \text{(TVC)} \tag{15}$$

$$1 = \iiint_{\Omega} g(h,a,z)\, dh\, da\, dz;\ g(h,a,z) \geq 0 \qquad \text{(proper density)} \tag{16}$$

$$0 = -V\left(\widetilde{s}\, g\right)\Big|_{\mathcal{A}} + V\left(-\widetilde{\mu}_z\, g + \frac{1}{2}\frac{\partial}{\partial z}(\widetilde{\sigma}_z^2\, g)\right)\Big|_{\mathcal{Z}} \tag{17}$$

(necessary BC for $\mathcal{B} = \mathcal{A}^*$)

where $V_x := \frac{\partial}{\partial x} V(h,a,z)$; $\widetilde{s} := \mathcal{Y}(h,a,z) + ra - \widetilde{c}$; $g := g(h,a,z) = \frac{\partial^3}{\partial h \partial a \partial z} G(h,a,z)$

back to model

Stationary Competitive Equilibrium

A **S.C.E** consists of a set of $\mathcal{F}_t$-adapted allocation processes $\{\widehat{c}_t, \widehat{s}_t, \widehat{\ell}_t, \widehat{k}_t, \mathcal{Y}_t\}_{t=s}^{\infty}$, constant price processes $\{w, r\}_{t=s}^{\infty}$, and a stationary joint distribution $G(h, a, z)$, such that, for all $t \geq s$:

1. Given prices, the allocations solve the stochastic control problem (8)–(15);
2. All markets clear:

$$\int_{\Omega} a \; dG(h,a,z) = \int_{\Omega_{\mathcal{E}}} \widehat{k}(h,a,z) \; dG(h,a,z) \qquad \text{[asset/physical capital market]}$$

$$\int_{\Omega_{\mathcal{W}}} h \; dG(h,a,z) = \int_{\Omega_{\mathcal{E}}} \widehat{\ell}(h,a,z) \; dG(h,a,z) \qquad \text{[labor/human capital market]}$$

$$\int_{\Omega_{\mathcal{E}}} \widehat{y}(h,a,z) \; dG(h,a,z) = \int_{\Omega} \widetilde{c}(h,a,z) \; dG(h,a,z) + \delta K(r,w) \qquad \text{[goods market (net))]}$$

3. The joint density of human capital, wealth, and entrepreneurial ability satisfies the Kolmogorov-Forward equation (9) with boundary conditions (14)–(15).

back to model

Entrepreneurs: Static Optimization (I)

Factor demands depend on each entrepreneur's state vector (h_e, a, z) and can be expressed as:

$$\widehat{\ell}(h_e,a,z) = (z\widetilde{h})^{\frac{1}{\nu}} \left(\frac{\alpha(1-\nu)}{\widehat{r}_t+\delta}\right)^{\frac{\alpha(1-\nu)}{\nu}} \left(\frac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{1-\alpha(1-\nu)}{\nu}} \tag{18}$$

$$\widehat{k}(h_e,a,z) = (z\widetilde{h})^{\frac{1}{\nu}} \left(\frac{\alpha(1-\nu)}{\widehat{r}_t+\delta}\right)^{\frac{1-(1-\alpha)(1-\nu)}{\nu}} \left(\frac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{(1-\alpha)(1-\nu)}{\nu}} \tag{19}$$

The entrepreneur-specific *shadow interest rate* is $\widehat{r}_t(h_e,a,z) = r_t + \lambda(h_e,a,z)$, where $\lambda(.) \geq 0$ is the Lagrange multiplier on the collateral constraint, and corresponds to:

$$\widehat{r}_t(h_e,a,z) = \begin{cases} r_t \quad \textit{if } a > (1-\vartheta)\,(z\widetilde{h})^{\frac{1}{\nu}} \left(\dfrac{\alpha(1-\nu)}{r_t+\delta}\right)^{\frac{1-(1-\alpha)(1-\nu)}{\nu}} \left(\dfrac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{(1-\alpha)(1-\nu)}{\nu}} \\ \left(\dfrac{(z\widetilde{h})^{\frac{1}{\nu}}}{\frac{1}{1-\vartheta}a}\right)^{\frac{\nu}{1-(1-\alpha)(1-\nu)}} \left(\dfrac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{(1-\alpha)(1-\nu)}{1-(1-\alpha)(1-\nu)}} \alpha(1-\nu)-\delta \quad \textit{otherwise} \end{cases} \tag{20}$$

Identification/Estimation of ω and κ, hence $\hbar$

- Sharp theoretical prediction; quite flexible in practice.

– Can accommodate various shapes and forms of nonlinear relationships.

▷ Global minimum property of $\hbar$ offers an empirical advantage:

– Given a mapping between years of schooling and human capital, $\hbar$ is an identifiable parameter that can be easily calibrated or estimated using standard methods.

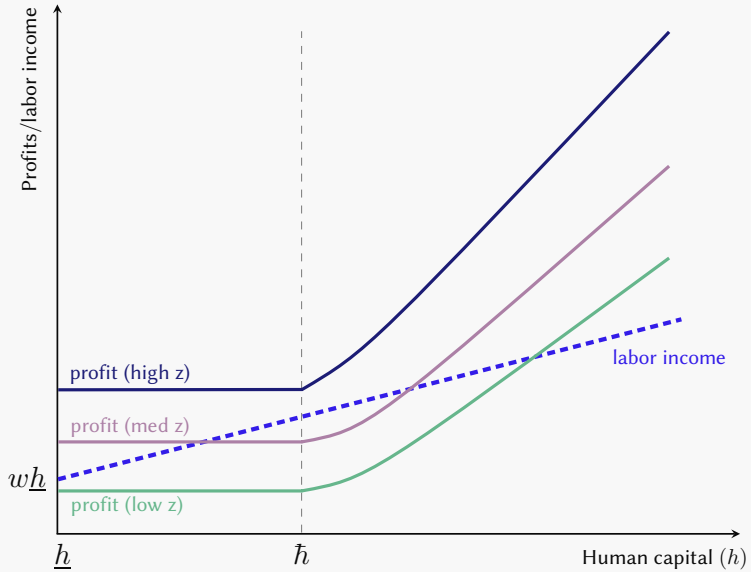
▷ The elasticity ω can also be calibrated or estimated using micro data.

– Holding other parameters fixed, the ERGON index we saw before (or similar) is a *sufficient statistic* for ω:

$$\frac{\mu(\mathcal{E}, h > \overline{h})/\mu(\Omega, h > \overline{h})}{\mu(\mathcal{E})/\mu(\Omega)}, \quad \text{for some } \overline{h} > \hbar$$

back to Proposition1

Occupational Choice: Profit Functions vs Labor Income when $\omega > \chi(\theta)$



Quantitative Model: Preassigned Parameters

☐ Set of preassigned parameters: $\{\alpha, \gamma, \delta, \mu_z, \varphi_h\}$.

- Values for the first three are fairly canonical in the macro-development literature. $\gamma = 1.5;\ \delta = 0.06;\ \alpha = 0.36.$

☐ Ornstein-Uhlenbeck (OU) process in natural logs for entrepreneurial ability:

$$d \log z_t = \varphi_z(\mu_z - \log z_t)\, dt + \sigma_z\, dW_t^z \tag{21}$$

$$dz_t = \left[\varphi_z(\mu_z - \log z_t) + \frac{1}{2}\sigma_z^2\right] z_t\, dt + \sigma_z z_t\, dW_t^z \tag{22}$$

→ The (marginal) stationary distribution is log-normal: $\log z_t \sim \mathcal{N}(\mu_z, \frac{\sigma_z^2}{2\varphi_z})$. Normalization: $\mu_z = 0$, which simply implies a zero-mean process in logs.

back to calibration strategy

Quantitative Model: Estimated Parameters

☐ Three parameters can be estimated using the employed data: $\{\mu_h, \sigma_h, \hbar\}$.

- Specify a bijective map between years of schooling and units of human capital. Draw upon standard practice, e.g., *Hall and Jones (1999)*, *PWT* hc index:

$$h_i = e^{\psi(S_i)},\ \psi(S) = \begin{cases} 0.134 \cdot S & \text{if } S \leq 4 \\ 0.134 \cdot 4 + 0.101 \cdot (S-4) & \text{if } 4 < S \leq 8 \\ 0.134 \cdot 4 + 0.101 \cdot 4 + 0.068 \cdot (S-8) & \text{if } 8 < S \end{cases}$$

▷ Use this transformation along with MLE in the pooled SCF data for $1989 - 2019$: $\log h \sim \mathcal{N}(1.338, 0.0513)$.

▷ Estimate κ and thus $\hbar$ from RII Probit regressions: 9th year of schooling ($\hbar = h_9$).

back to calibration strategy

Quantitative Model: Calibrated Parameters

- Six parameters to be jointly calibrated for the model to best fit six relevant moments in the data: $\{\nu, \omega, \sigma_z, \varphi_z, \vartheta, \rho\}$.
  - $\nu \models$ fraction of entrepreneurs in the labor force
  - $\omega \models$ share of total income received by top 20%
  - $\sigma_z \models$ dispersion of (log) TFP
  - $\varphi_z \models$ firm entry rate
  - $\vartheta \models$ External finance-to-output ratio
  - $\rho \models$ real rate of return on capital

back to calibration strategy