Income Distribution

Ping Wang
Department of Economics
Washington University in St. Louis

April 2019
A. Introduction

- **Stylized facts (U.S. over the past 4 or 5 decades):**
  - wage inequality increased sharply: 90%-10% ratio rose by over 40%, documented by Katz-Autor (1999)
  - despite an increase in skill premium/between-group inequality, the majority of the increase in wage inequality is *residual*, due to unobserved characteristics of workers in the same education and demographic group

- While the literature provides adequate explanation on the between-group inequality, it is largely failed in explaining the within-the-skilled-group inequality, with only a few attempts including, Aghion (2000), Violante (2002), Jovanovic (2009) and Tang and Wang (2014)


- Inequality is also associated with geographic stratification, particularly within municipals and to some degree across different regions
  - *Banabou (1996)* offers a simple framework for human capital stratification
  - *Acemoglu-Dell (2009)* provide useful decomposition of wage inequalities
Recently, Piketty (2014) emphasizes a sharp rise in inequality particularly between the super rich (top 1%) and the low or even middle income groups
  ○ historical data: Piketty (2014)
  ○ new data: tax administrative data (no top coding), wealth data

- Wealth inequality:
  ○ survey: De Nardi (2015)
B. Education Provision, Growth and Inequality: Glomm-Ravikumar (1992)

- Different from the representative-agent framework developed by Lucas (1988), this paper allows for human capital heterogeneity, which enables a clean study of the issues of growth vs. distribution as well as private vs. public education

1. The Model

- 2-period lived agents, who work when young and consume when old (endogenous labor-leisure trade-off, with altruism)
- Preferences: \( V_t = \ln n_t + \ln c_{t+1} + \ln e_{t+1} \), that is, agent of generation-\( t \) cares leisure, consumption and the offspring’s quality of education
- Human Capital:
  - distribution: \( G_t(h) \sim \text{log normal} (\mu_t, \sigma_t^2) \)
  - evolution: \( h_{t+1} = \vartheta h_t^\delta (1-n_t)^\beta e_t^\gamma \), \( \delta, \beta, \gamma \in (0,1) \) (Lucas: \( \gamma = 0, \delta = \beta = 1 \))
- CRS production: output = \( h_{t+1} \)
Two educational systems:

○ public education: \[ E_{t+1} = \tau_{t+1} H_{t+1}, \quad H_{t+1} = \int h_{t+1} dG_{t+1}(h_{t+1}) \] 
  (income tax)  (mean income)

○ private education: \[ e_{t+1} = h_{t+1} - c_{t+1} \]

2. Optimization and Equilibrium

a. Public Education:

• Individual optimization:

\[
\max_{n,c} \ln n_t + \ln c_{t+1} + \ln E_{t+1} \\
\text{s.t.} \quad c_{t+1} = (1 - \tau_{t+1})h_{t+1} \\
\quad h_{t+1} = \theta (1 - n_t)^\beta E_t^\gamma h_t^\delta \\
\Rightarrow \max_{n_t} \ln n_t + \ln[(1 - \tau_{t+1})\theta E_t^\gamma h_t^\delta] + \beta \ln(1 - n_t) + \ln E_{t+1}
\]

• FOC: \[ 1 - n_t = \frac{\beta}{1 + \beta} \]
- Government optimization:

\[
\max_{\tau} \ln[(1 - \tau_{t+1})h_{t+1}] + \ln \tau_{t+1}H_{t+1} \quad (\therefore n_t = \frac{1}{1 + \beta} \text{ fixed})
\]

\[
\Rightarrow \max_{\tau} \ln(1 - \tau) + \ln \tau
\]

- FOC: \( \tau = 1/2 \)

- Equilibrium:
  - human capital evolution: 
    \[
    h_{t+1} = \theta \left( \frac{\beta}{1 + \beta} \right)^{\beta} \left( \frac{1}{2} \right)^{\gamma} H_{t+1}^\gamma h_{t+1}^\delta = A H_{t+1}^\gamma h_{t+1}^\delta
    \]
  - aggregate human capital: 
    \[
    H_t = \exp[\mu_t + \frac{\sigma_t^2}{2}]
    \]
    - mean: \( \mu_{t+1} = \ln A + \gamma \ln H_t + \delta \mu_t, \) or, \( \mu_{t+1} = \ln A + (\gamma + \delta) \mu_t + \frac{\gamma \sigma_t^2}{2} \)
    - variance (inequality measure): \( \sigma_{t+1}^2 = \delta^2 \sigma_t^2 \)
b. Private Education

- **Individual optimization**

\[
\max_{n_t, c_{t+1}, e_{t+1}} \ln n_t + \ln c_{t+1} + \ln e_{t+1} \\
\text{s.t.} \quad h_{t+1} = \Theta(1 - n_t)^\beta e^\gamma h_t^\delta \\
\quad c_{t+1} = h_{t+1} - e_{t+1} \\
\quad \Rightarrow \max_{n_t, c_{t+1}, e_{t+1}} \ln n_t + \ln[\Theta(1 - n_t)^\beta e^\gamma h_t^\delta - e_{t+1}] + \ln e_{t+1}
\]

- **FOCs:** \( c_{t+1} \equiv e_{t+1} = \frac{1}{2} h_{t+1} ; \quad 1 - n_t = \frac{\beta}{1 + \beta} > \frac{\beta}{1 + \beta} \) (free-rider in public education)

- **Equilibrium:**
  - \( h_{t+1} = \Theta(\frac{\beta}{1 + \beta})^\beta (\frac{1}{2})^\gamma h_t^{\gamma + \delta} \equiv B h_t^{\gamma + \delta} \quad (B > A) \)
  - \( \mu_{t+1} = \ln B + (\gamma + \delta) \mu_t \)
  - \( \sigma_{t+1} = (\gamma + \delta)^2 \sigma_t^2 \)
3. **Growth vs. Inequality**

- **Inequality:**
  - Public education: inequality $\downarrow$ over time
  - Private education: inequality may decline (or rise) over time if $\delta + \gamma < (or >) 1$

- **Is inequality harmful for growth?**
  - Public education: $H_{t+1} = AH_t^\gamma \delta \exp[-\frac{1}{2} \delta (1-\delta) \sigma_i^2] \Rightarrow d(H_{t+1}) / d\sigma_i^2 < 0$
  - Private education: $H_{t+1} = BH_t^\gamma \delta \exp[\frac{1}{2} (\gamma + \delta) (\gamma + \delta - 1) \sigma_i^2]$
    
    $\Rightarrow d(H_{t+1}) / d\sigma_i^2 < (or >) 0$ if $\delta + \gamma < (or >) 1$

- **Kuznets curve:** the correlation between growth and inequality is consistent with the Kuznets curve under private education
4. Political Economy and Institutional Choice: Public vs. Private Education

- Mechanism: majority voting by the old (political economy) – ignore $n_t$ (decision by the young)
- Value functions:
  - Public education: $V^{old}(public) = 2\ln(\frac{1}{2}) + \ln h + \mu + \frac{\sigma^2}{2}$
  - Private education: $V^{old}(private) = 2\ln(\frac{1}{2}) + 2\ln h$
- Median voter’s decision:
  - $V^{old}(public) - V^{old}(private) = [\mu - \ln h(\text{median})] + \frac{\sigma^2}{2} = \frac{\sigma^2}{2} > 0$
    (ex ante mean $\mu = \text{median} < \text{ex post mean} = \mu + \frac{\sigma^2}{2}$, because log normal distribution has a long tail)
  - outcome: select public education system (U.S.: 86%-public education)
- Problem: under public education, the declined income inequality is inconsistent with the real world observation
C. General Purpose Technology and Between/Within-Group Inequality: Aghion (2000)

- Stylized facts in U.S. & U.K: within-group inequality started before between-group inequality
- Equipment price and skill premium – Krusell et al. (2000 Econometrica):

\[ y_t = A_t \left( K_S^\alpha \left[ \mu u^\sigma + (1 - \mu) (\lambda k_e^\rho + (1 - \lambda) S_t^\rho) \right] \right)^\frac{\sigma}{1 - \alpha} \]

under \( \frac{1}{1 - \sigma} > \frac{1}{1 - \rho} \) (stronger complementarity between \( k_e \) and \( S \)),

equipment price \( \downarrow \Rightarrow \frac{W_s}{W_u} \uparrow \)

1. Between-Group Inequality

- General purpose technology (GPT) experimentation and adoption require skilled labor
- Production: \( y = \left[ \int_0^1 A(i)^a x(i)^a di \right]^{1/a} \), \( A(i) = \begin{cases} 1 & \text{if sector } i \text{ uses old GPT} \\ \gamma > 1 & \text{if sector } i \text{ uses new GPT} \end{cases} \)

- Skilled Labor: \( L_s(t) = L[1-(1-s)e^{-\beta t}] \)
  - \( \beta \) = speed of exogenous skill acquisition
  - 1 = \( n_0 \) (old GPT) + \( n_1 \) (experimenting new) + \( n_2 \) (new)

- Arrival of new GPT:
  \( \lambda(n_2) = \begin{cases} \lambda_0 & \text{if } n_2 \leq \bar{n} \\ \lambda_0 + \Delta & \text{if } n_2 \geq \bar{n} \end{cases} \)
  where \( \lambda_0 \) is small, \( \Delta \) is large and \( \lambda_1 \) is the arrival of successful experimentation

- Population dynamics:
  - \( \dot{n}_1 = \lambda(n_2)n_0 - \lambda_1 n_1 \)
  - \( \dot{n}_2 = \lambda_1 n_1 \)
• Early stage (A): \( n_1 + n_2 \) is too small to absorb \( L_s \) \( \Rightarrow \) integrated labor market with wage equalization, i.e.,

\[
(1 - n_2)x_0 + n_1L_1 + n_2x_2 = L
\]

• Later stage (B): \( L_s \) is fully absorbed by \( n_1 \) and \( n_2 \) \( \Rightarrow \) segmented labor market with \( n_1L_1 + n_2x_2 = L_s \) and \( (1 - n_2)x_0 = L_u \)

2. Within-Group Inequality

• Machine lasts exactly two periods (with no depreciation within the two periods)
• Only a random fraction (\( \sigma \)) of workers get chance to adopt new GPT (crucial to create within-group heterogeneity)
• Continual adoption of new GPT yields higher productivity due to learning (at rate \( \tau \))
• By experience, learning of old GPT is more efficient (at rate \( \eta > \tau \))
• Production
  ○ new GPT: $v_t = A_t x_{ot}^{1-\alpha}$
  ○ old GPT: $z_t = A_{t-1} [(1+\eta)x_{1t}]^{1-\alpha}$

• Technology evolution: $A_t = (1+\gamma)A_{t-1}$

• Labor and Population Identity:
  ○ $n_{ij}$ (transition from $i$ to $j$) with $i, j = 0$ (new) or 1 (old)
  ○ $x_0 = (1+\tau)n_{00} + n_{10}$
  ○ $x_1 = n_{01} + n_{11}$
  ○ $n_{00} + n_{10} + n_{01} + n_{11} = 1$

• Adaptability Constraints: $\dot{n}_{00} \leq \sigma(n_{00} + n_{10})$ and $\dot{n}_{10} \leq \sigma(n_{01} + n_{11})$

• Steady-State Transition: $n_{10} = n_{01}$

• Consumption Efficiency: $u(c) = \sum \beta^t \ln c \Rightarrow 1 + r = \frac{1}{\beta} \frac{c_{t+1}}{c_t} = \frac{1}{\beta} (1+\gamma)$
• Labor Demand:
  \[
  \frac{w_0}{w_1} = \frac{1 + \gamma}{(1 + \eta)^{1-\alpha}} \left( \frac{x_0}{x_1} \right)^{-\alpha}
  \]
  \[w_{00} = (1 + \tau) w_0 \ ; \ w_{10} = w_0 \ ; \ w_{01} = w_{11} = w_1\]

• Labor Supply:
  ○ value functions:
    - \[\nu_i = \frac{w_i}{w_i} + \beta \{ \sigma \max(\nu_{00}, \nu_1) + (1 - \sigma)\nu_1 \}\]
    - \[\nu_1 = \frac{w_1}{w_1} + \beta \{ \sigma \max(\nu_{10}, \nu_1) + (1 - \sigma)\nu_1 \}\]
  ○ cases:
    - when \(\nu_{10} < \nu_1\), labor supply decision \(\Rightarrow x_0/x_1 = 0\)
    - when \(\nu_{10} > \nu_1\), labor supply decision \(\Rightarrow x_0/x_1 = \chi\)
    - when \(\nu_{10} = \nu_1\) \(\left(\frac{w_0}{w_1} = \Omega\right)\)
      \[\nu_1 = \frac{w_1}{w_1} + \beta \sigma v_1 + (1 + \sigma)\nu_1 \ ; \ w_1 = \sigma (1 - \beta) \nu_1,\]
      \[w_0 = \sigma [\nu_1 - \beta v_{00}], \ w_{00} = (1 - \sigma) v_{00} - (1 - \sigma) \nu_1\]
Labor Market Equilibrium

\[ L^d = L^s \Rightarrow \frac{w_0}{w_1} = \frac{1 + \gamma}{(1 + \eta)^{1-\alpha}} \left[ \frac{1 - \sigma}{\sigma(1 + \sigma\tau)} \right]^\alpha \equiv \Phi(\gamma, \sigma, \eta, \tau) \]

Wage inequality within the skilled group:

- \[ \max \left\{ \frac{w_{00}}{w_0}, \frac{w_{00}}{w_1} \right\} \]
- \[ \max \left\{ \frac{w_{00}}{w_0}, \frac{w_{00}}{w_0} \right\} \]
- \[ = (1 + \tau) \max \{ 1, \Phi \} \]
- in general, within-group inequality rises when GPT size (\( \gamma \)) \( \uparrow \), GPT learning (\( \tau \)) \( \uparrow \), and monopoly rent \( \uparrow \) (\( \sigma \) \( \downarrow \) or \( \eta \) \( \downarrow \))

Timing: even at the early stage (A) when skill premium is zero, within-group inequality can arise already

Problem: the underlying force driving within-group inequality is rather ad hoc
D. Skill Transferability and Residual Wage Inequality: Violante (2002)

- Stylized facts (US over the past 4 or 5 decades):
  - wage inequality increased sharply: 90%-10% ratio rose by over 40%, documented by Katz-Autor (1999)
  - despite an increase in skill premium/between-group inequality, the majority of the increase in wage inequality is residual, due to unobserved characteristics of workers in the same education and demographic group

- Previous studies on wage inequality focus on ex ante fixed innate ability
  - such as Acemoglu (1999), Caselli (1999), Aghion (2000), and Galor-Moav (2000)
  - counterfactually high persistency in inequality: Gottschalk-Moffitt (1994) find temporary components are as large as permanent ones

- Violante (2002) takes a deeper look at the data, finding that increased earning variability is due to:
  - more frequent job separation for a given turnover rate
  - more volatile dynamics of wages on the job and between jobs

- The above observations motivate the construction of a theory of inequality focusing on the accumulation and the transferability of specific human capital

- Key driving force: technology differences across machines of different vintages
1. The Basic Structure and Results

- Technology frontier advances at rate $\gamma > 0$
- Each machine has two periods of productive life and does not depreciate after the first period (as in Aghion 2000)
- A machine $M_j$ of age $j$ matched with worker of skill $z$ produces output:
  \[ y_j = (1 + \gamma)^{-\theta j} z \]
- Matching surplus sharing rule: $\xi$ to worker and $1-\xi$ to firm
- Value functions:
  - value of employed:
    - with machine $M_0$: $V_0 = w_0 + \beta \max\{V_1, U\}$
    - with machine $M_j$: $V_1 = w_1 + \beta U$
  - value of unemployed: $U = \alpha V_0 + (1 - \alpha)V_1$
  where \( \beta = \) productivity-adjusted discount factor
  \( \alpha = \) probability of meeting a new machine
- Separation decision for workers on new technologies: $\chi = \{0,1\}$
  - by construction, $w_0 > w_1$; thus, $U > V_1$
  - so if $\chi = 1$, we must have equal fractions of idle $M_0$ and $M_1$, i.e., $\alpha = 1/2$
- Wage inequality $\text{var}(\ln(w)) = [(\theta\ln(1+\gamma)/2)^2 - ((\theta\gamma)/2)^2$, depending exclusively on the technology differences across machines of different vintages ($\gamma$)
2. Generalization: Vintage Human Capital

- A worker on \( M_j \) may move on \( M_{j'} \) with cumulated skills determined by the transferability process: \( z_{jj'} = (1+\gamma)^{\tau(j'-j+1)} \) (following the adaptation structure in Aghion 2000)
  - the transferability of specific human capital is measured by \( \tau \)
  - equilibrium skill levels:
    - \( z_{01} = 1 \)
    - \( z_{00} = z_{11} = (1+\gamma)^{-\tau} \)
    - \( z_{10} = (1+\gamma)^{-2\tau} \)
- Productivity-adjusted wage: \( w_{ij} = (1+\gamma)^{-\theta_j} \)
- Value functions: change to \( V_{ij} \) based on \( w_{ij} \)
- Worker’s separation decision:
  - \( \tau \leq \theta \Rightarrow \chi = 1 \) for all \( \gamma \)
  - \( \tau > \theta \Rightarrow \chi = 1 \) for \( \gamma > \gamma_c \)
- Wage inequality: \( \text{var}(\ln(w)) \propto (\theta \gamma)^2 \text{var}(j) + \text{var}(\ln(z)) - 2\theta \gamma \text{cov}(\ln(z),j) \)
  - higher \( \gamma \) increases \( \text{var}(\ln(z)) \) and \( \text{cov}(\ln(z),j) \), raising \( \text{var}(\ln(w)) \) if \( \chi = 0 \)
  - the effect of \( \gamma \) on \( \text{var}(\ln(w)) \) is ambiguous if \( \chi = 1 \)
3. Calibration

- Observation: residual wage inequality
Parameterization

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Moment to match (yearly average)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_L = .036$</td>
<td>growth of rel. price of equipment ($&lt; 1974$)</td>
<td>Krusell et al. [2000]</td>
</tr>
<tr>
<td>$\gamma_H = .048$</td>
<td>growth of rel. price of equipment ($&gt; 1974$)</td>
<td>Krusell et al. [2000]</td>
</tr>
<tr>
<td>$\theta = .7$</td>
<td>growth of real average wage $= .024$</td>
<td>Murphy and Welch [1992]</td>
</tr>
<tr>
<td>$\beta = .964$</td>
<td>rate of return on capital $= .05$</td>
<td>Cooley [1995]</td>
</tr>
<tr>
<td>$\kappa = 5$</td>
<td>labor share $= .68$</td>
<td>Cooley [1995]</td>
</tr>
<tr>
<td>$J = 28$</td>
<td>average age of equipment $= 7.7$</td>
<td>Bureau of Economic Analysis [1994]</td>
</tr>
<tr>
<td>$\lambda = .345$</td>
<td>wage growth within job $= .03$</td>
<td>Topel [1991]</td>
</tr>
<tr>
<td>$\tau = 1.90$</td>
<td>wage loss upon layoff $= .23$</td>
<td>Jacobson et al. [1993], Topel [1991]</td>
</tr>
<tr>
<td>$Z = 20$</td>
<td>transitory residual wage variance $= .053$</td>
<td>CPS data, Gottschalk and Moffitt [1994]</td>
</tr>
<tr>
<td>$\delta = .05$</td>
<td>separation rate from employment $= .166$</td>
<td>Blanchard and Diamond [1990]</td>
</tr>
</tbody>
</table>
Fitness of the Model

<table>
<thead>
<tr>
<th></th>
<th>Variance of log wages</th>
<th>Variance of technologies</th>
<th>Variance of skills</th>
<th>Covariance component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DATA</td>
<td>MODEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_L = .035$</td>
<td>.053</td>
<td>.053</td>
<td>.008</td>
<td>.085</td>
</tr>
<tr>
<td>$\gamma_H = .048$</td>
<td>.089</td>
<td>.085</td>
<td>.014</td>
<td>.145</td>
</tr>
<tr>
<td></td>
<td>Average age of capital</td>
<td>Average skill level</td>
<td>Wage growth within-job</td>
<td>Wage loss upon layoff</td>
</tr>
<tr>
<td>$\gamma_L = .035$</td>
<td>7.700</td>
<td>11.086</td>
<td>.030</td>
<td>$-.230$</td>
</tr>
<tr>
<td>$\gamma_H = .048$</td>
<td>7.448</td>
<td>8.595</td>
<td>.044</td>
<td>$-.305$</td>
</tr>
</tbody>
</table>

4. Open Issues

- firm-specific technologies
- occupational mobility
- general vs. specific human capital
E. Human Capital Stratification

- In reality, households are stratified in various degrees by race, income, education and other socioeconomic indicators
- The Dissimilarity index (Duncan-Duncan 1955): using the 2000 Census data, Peng and Wang (2005) show highly stratified top 30 MSAs in the US:

<table>
<thead>
<tr>
<th>Metropolitan Statistical Area (MSA)</th>
<th>Dissimilarity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>D C - B a l t i m o r e , D e t r o i t</td>
<td>0.70 or higher</td>
</tr>
<tr>
<td>M i l w a u k e e , C l e v e l a n d , S t . L o u i s , N e w Y o r k</td>
<td>0.60 - 0.69</td>
</tr>
<tr>
<td>P h i l a d e l p h i a , C i n c i n n a t i , C h i c a g o , I n d i a n a p o l i s</td>
<td></td>
</tr>
<tr>
<td>P i t t s b u r g h , A t l a n t a , K a n s a s C i t y</td>
<td>0.50 - 0.59</td>
</tr>
<tr>
<td>H o u s t o n , B o s t o n , L o s A n g e l e s</td>
<td></td>
</tr>
<tr>
<td>T a m p a , S a n A n t o n i o , P h o e n i x , M i n n e a p o l i s</td>
<td></td>
</tr>
<tr>
<td>S a n D i e g o , N o r f o l k , S a n F r a n c i s c o</td>
<td>0.40 - 0.49</td>
</tr>
<tr>
<td>M i a m i , D e n v e r , S a c r a m e n t o , O r l a n d o</td>
<td></td>
</tr>
<tr>
<td>D a l l a s , S e a t t l e , P o r t l a n d</td>
<td>0.39 or lower</td>
</tr>
</tbody>
</table>
It has been shown that since 1980, racial segregation in the U.S. has declined while economic segregation has risen.

Human capital and housing are believed the two primary sources of economic segregation.


Interactions
- Local positive spillovers – in human capital evolution
- Global positive spillovers – in goods production

Human Capital and Education
- Human capital evolution: $ h_{t+1}^i = \phi^i ((1-u_t^i)h_t^i)^\delta (E_t^i)^{1-\delta} $
- Public education: $ E_t^i = \tau_t^i \int y_t^i dG_t^i(y_t^i) $

Output: $ y_{t+1}^i = A(H_t)^\alpha (h_t^i)^{1-\alpha} $

Combining the above relationships $ \Rightarrow h_{t+1}^i = B^i (h_t^i)^\delta (H_t)^{\alpha(1-\delta)} (L_t^i)^{(1-\alpha)(1-\delta)} $, where $L_t^i$ is a “local” human capital aggregator.
2. Segregated vs. Integrated Equilibrium

- Segregated equilibrium features locational clustering by human capital/income
- Integrated equilibrium features mixture of groups with different human capital/income
- Two fundamental forces:
  - complementarity between $L^i$ and $h_i$ => segregation ( assortative matching)
  - complementarity between $H$ and $h^i$ => integration (homogenizing)

3. Results

- Co-existence of segregated and integrated equilibria
- Integration lowers inequality as compared to segregation
- Integration lowers growth in SR but raises it in LR, because $H$ has a larger scale effect in the long run
F. Income Inequality Across Space and Time: Acemoglu-Dell (2009)

- Stylized fact: large cross-country and within-country differences in per capita income
- Potential causes of such disparities:
  - differences in *human capital*
  - differences in technological know-how
  - differences in production efficiency due to various institutions and organizations

1. The Model

- Measure of inequality (municipal m in country j) by the Theil index:
  
  \[
  T = \sum_{j=1}^{J} \frac{L_j}{L} \frac{y_j}{y} \left( \ln \frac{y_j}{y} \right) + \sum_{j=1}^{J} \frac{L_j}{L} \frac{y_j}{y} \left( \sum_{m=1}^{M_j} \frac{L_{jm}}{L_j} \frac{y_{jm}}{y_j} T_{jm} + \sum_{m=1}^{M_j} \frac{L_{jm} y_{jm}}{L_j y_j} \ln \left( \frac{y_{jm}}{y_j} \right) \right)
  \]

  where \( T_{jm} = \sum_{i=1}^{L_{jm}} \frac{y_{jmi}}{L_{jm} y_{jm}} \ln \left( \frac{y_{jmi}}{y_{jm}} \right) \) is the within-municipal m Theil index in country j

- Alternative measures: mean log deviation, variance/coefficient of variation, gini coefficient
Wage inequality

<table>
<thead>
<tr>
<th></th>
<th>90/10</th>
<th>Theil index</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Between</td>
<td>Within</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Country</td>
<td>Country</td>
<td></td>
</tr>
<tr>
<td>Municipals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>actual pop weights</td>
<td>34.2</td>
<td>0.250</td>
<td>0.544</td>
<td></td>
</tr>
<tr>
<td>equal pop weights</td>
<td>28.6</td>
<td>0.285</td>
<td>0.622</td>
<td></td>
</tr>
<tr>
<td>Regions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>actual pop weights</td>
<td>36.7</td>
<td>0.203</td>
<td>0.529</td>
<td></td>
</tr>
<tr>
<td>equal pop weights</td>
<td>32.7</td>
<td>0.139</td>
<td>0.615</td>
<td></td>
</tr>
</tbody>
</table>

- more within than between country inequalities
- more inequality using municipal than region data
Decomposition of wage inequality measured by Theil index

<table>
<thead>
<tr>
<th></th>
<th>Overall Inequality</th>
<th>Residual Inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Municipals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>actual pop weights</td>
<td>0.265</td>
<td>0.067</td>
</tr>
<tr>
<td>equal pop weights</td>
<td>0.301</td>
<td>0.105</td>
</tr>
<tr>
<td>U.S.</td>
<td></td>
<td>0.050</td>
</tr>
</tbody>
</table>

- “residual” within-the-skilled-group inequalities account for a large portion of overall inequalities
- within-municipal disparities are most important for wage inequalities
- between-country disparities are important only for “non-residual” between-skilled-and-unskilled-group inequalities
- between-municipal disparities are never important
G. The Battle between the Top 1% and the Remaining 99%: Pikety (2014)

- **Income inequality**

  ![Graph: Top 1% Share of Total Pre-Tax Income (1913-2012)]

- **Wealth inequality**
  - U.S. Wealth Inequality: [https://www.youtube.com/watch?v=QPKKQnijnsM](https://www.youtube.com/watch?v=QPKKQnijnsM)
  - Capital In The 21st Century:
    - BBC: [https://www.youtube.com/watch?v=HL-YUTFqtuI](https://www.youtube.com/watch?v=HL-YUTFqtuI)
    - ABC: [https://www.youtube.com/watch?v=I05wLUuvQGM](https://www.youtube.com/watch?v=I05wLUuvQGM)
Methodological issues:

- Piketty: $r$ measures return to capital, $g$ measures return to labor, so $r > g$ implies widened inequality
- Krusell-Smith (2015): Piketty’s $r > g$ theory works only with the unconventional definition of capital-output in terms of net capital (net of depreciation) and NNP
- Weil (2015): market value of tradeable assets are incomplete measures for productive capital and wealth, missing
  - value of human capital
  - transfer wealth
  - these omitted types of wealth are distributed more equally than tradeable assets
H. Wealth Inequality: De Nardi (2015)

- Cagetti-De Nardi (2006): over the past 3 decades in the U.S., top 1% own 1/3 of national wealth, top 5% more than 1/2 (see also an older literature led by Wolff 1992, 1998)
- Can typical models predict such a high concentration of wealth?

1. The Bewley (1977) Model of Permanent Income

- Infinitely lived agents with time-additive preferences:
  \[
  E \left\{ \sum_{t=1}^{\infty} \beta^t u(c_t) \right\}
  \]
  - \( u \) takes a CRRA form
  - Labor endowment subject to an idiosyncratic labor productivity shock \( z \), taking finite number of values and following a first-order Markov process with transition matrix \( \Gamma(z) \)
  - A single asset \( a \) that may be used to insure against labor income risk
  - Production of a single good \( Y \) using \( K \) and \( L \) under a CRS technology
- **Household’s problem:**

\[
V(x) = \max_{(c,a')} \left\{ u(c) + \beta E \left[ V(a', z') | x \right] \right\}
\]

subject to

\[
c + a' = (1 + r)a + zw
\]

- \(a = \text{net borrowing limit}\)
- state \(x = (a, z)\)

- In a stationary equilibrium, the distribution of people with \((a, z)\) is constant
- Quantitative analysis by Aiyagari (1994): log(labor earning) follows AR(1) with autocorrelation = 0.6 and std dev of the innovations = 0.2

<table>
<thead>
<tr>
<th>% wealth in top</th>
<th>Gini</th>
<th>1%</th>
<th>5%</th>
<th>20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. data, 1989 SCF</td>
<td>.78</td>
<td>29</td>
<td>53</td>
<td>80</td>
</tr>
<tr>
<td>Aiyagari Baseline</td>
<td>.38</td>
<td>3.2</td>
<td>12.2</td>
<td>41.0</td>
</tr>
</tbody>
</table>

- wealth inequality largely underestimated compared to the 1989 Survey of Consumer Finance (not much improved even doubling std dev)

- Agents live for at most N periods, subject to survival probability $s_t$ of surviving up to t conditional on surviving at t-1
- Lifetime utility: $E \left\{ \sum_{t=1}^{N} \beta^t \left( \prod_{j=1}^{t} s_j \right) u(c_t) \right\}$
- Labor endowment is now age-specific: $e(z, t)$
  - again, z is Markov with transition $\Gamma(z)$
- No annuity, so people self-insure against earning risk and long life
- Those die prematurely leave accidental bequests
- Same production technology as in Bewley
- Household’s problem:

$$V(a, z, t) = \max_{(c, a')} \left\{ u(c) + \beta s_{t+1} E \left[ v(a', z', t + 1) | z \right] \right\}$$

subject to:

$$c + a' = (1 + r)a + e(z, t)w + T + b_t$$

- $c \geq 0$, $a' \geq a$ and $a' \geq 0$ if $t = N$
- $T =$ lump-sum redistributed accidental bequests
- $b =$ social security payments to the retired
- **Stationary equilibrium:** similar to Bewley, with periodically balanced bequest transfers and government budget

- **Quantitative results:**

<table>
<thead>
<tr>
<th>Transfer wealth ratio</th>
<th>Wealth Gini</th>
<th>1%</th>
<th>5%</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>Percentage with negative or zero wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989 U.S. data</td>
<td>.60</td>
<td>.78</td>
<td>29</td>
<td>53</td>
<td>80</td>
<td>93</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.8–15.0</td>
</tr>
<tr>
<td>A basic overlapping-generations Bewley model</td>
<td>.67</td>
<td>.67</td>
<td>7</td>
<td>27</td>
<td>69</td>
<td>90</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

- improved, but still far off for the top 1 or 5% wealth distribution

3. **Wealth Distribution in Variations of the Bewley Model**

- **Benhabib-Bisin (2015):** with intergenerational transmission and redistributive fiscal policy, the stationary wealth distribution is Pareto, driven critically by capital income and estate taxes

- **Benhabib-Bisin-Zhu (2016):** capital income shocks more important than labor income shocks

- Household’s value:
  \[ V(a, t) = \max_{c,a'} \left\{ u(c) + s_t \beta E_t V(a', t + 1) + (1 - s_t) \phi(b(a')) \right\} \]

  - value from leaving bequest by providing a worm glow (enjoyment of giving a la Andreoni (1989):
    \[ \phi(b(a')) = \phi_1 \left( 1 + \frac{b(a')}{\phi_2} \right)^{1-\sigma} \]
  - overall bequest motive: \( \phi_1 \)
  - bequest luxuriousness \( \phi_2 \)

- Two intergenerational linages:
  - human capital: inheritance in labor productivity
  - bequests
**Quantitative results**

<table>
<thead>
<tr>
<th>Transfer wealth ratio</th>
<th>Wealth Gini</th>
<th>1%</th>
<th>5%</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>Percentage with negative or zero wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989 U.S. data</td>
<td></td>
<td>29</td>
<td>53</td>
<td>80</td>
<td>93</td>
<td>98</td>
<td>5.8–15.0</td>
</tr>
<tr>
<td>No intergenerational links, equal bequests to all</td>
<td>.60</td>
<td>.78</td>
<td>29</td>
<td>53</td>
<td>80</td>
<td>93</td>
<td>98</td>
</tr>
<tr>
<td>No intergenerational links, unequal bequests to children</td>
<td>.67</td>
<td>.67</td>
<td>7</td>
<td>27</td>
<td>69</td>
<td>90</td>
<td>98</td>
</tr>
<tr>
<td>One link: parent’s bequest motive</td>
<td>.38</td>
<td>.68</td>
<td>7</td>
<td>27</td>
<td>69</td>
<td>91</td>
<td>99</td>
</tr>
<tr>
<td>Both links: parent’s bequest motive and productivity inheritance</td>
<td>.55</td>
<td>.74</td>
<td>14</td>
<td>37</td>
<td>76</td>
<td>95</td>
<td>100</td>
</tr>
</tbody>
</table>

- unequal bequests do not matter
- both intergenerational links matter to top group wealth distribution

- Agents are altruistic and face uncertainty about death time
- Occupational choice: workers vs. entrepreneurs
  - entrepreneurial production with working capital $k$ and ability $\theta$:
    \[ f(k) = \theta k^\nu + (1 - \delta)k \]
  - working capital subject to borrowing constraints, so $k = a + b(a)$, with borrowing $b$ depending on asset collateral $a$
- Quantitative findings:

<table>
<thead>
<tr>
<th>Wealth Gini</th>
<th>Fraction of entrepreneurs</th>
<th>Percentage wealth in the top 1%</th>
<th>5%</th>
<th>20%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.78</td>
<td>10%</td>
<td>29</td>
<td>53</td>
<td>80</td>
</tr>
<tr>
<td>Baseline model with entrepreneurs</td>
<td>0.8</td>
<td>7.50%</td>
<td>31</td>
<td>60</td>
<td>83</td>
</tr>
</tbody>
</table>

- over-estimation in top 5% wealth share especially under a smaller share of entrepreneurs

- Even the best fit model stated above is off, not to mention its ad hoc modeling strategy
- Can we do better? A potential new avenue is to consider heterogeneous financial knowledge
- Education and lifecycle income profile:
- **Lifecycle wealth profile:**

- **Fraction of financial knowledgeable and fraction of using financial advisors**
Financial knowledge => high return $R$, but with unit cost $\pi$

With saving $s$, wealth $a = Rs$

Household optimization:

$$\max_{a,R} u(y - \pi R - a/R) + \beta u(a)$$

- with log utility, wealth-income ratio is:
  $$\frac{a^*}{y} = \frac{y}{(2 + \frac{1}{\beta})^2 \pi}$$
  - increasing in $y$
  - decreasing in $\pi$

Model the evolution of financial knowledge:

$$f_{t+1} = (1 - \delta)f_t + i_t$$

Cash on hand: $x_t = a_t + y_t - oop_t$ (oop = out of pocket expenditure)

Wealth evolution: $a_{t+1} = \tilde{R}_\kappa(f_{t+1})(x_t + tr_t - c_t - \pi(i_t) - c_dI(\kappa_t > 0))$ where $\kappa =$ fraction of wealth in sophisticated financial asset and $\tilde{R}_\kappa(f_{t+1}) = (1 - \kappa_t)\bar{R} + \kappa_t\tilde{R}(f_t)$

Income process

$$\log y_{e,t} = g_y, e(t) + \mu_{y,t} + \nu_{y,t}$$

$$\mu_{y,t} = \rho_{y,e}\mu_{y,t-1} + \varepsilon_{y,t}$$

$$\varepsilon_{y,t} \sim N(0, \sigma_{y,\varepsilon}^2), \nu_{y,t} \sim N(0, \sigma_{y,v}^2)$$
• Out of pocket expenditure process:

\[
\log oop_{e,t} = g_{o,e}(t) + \mu_{o,t} + \nu_{o,t}
\]

\[
\mu_{o,t} = \rho_{o,e}\mu_{o,t-1} + \varepsilon_{o,t}
\]

\[
\varepsilon_{o,t} \sim N(0, \sigma_{o,\varepsilon}^2), \quad \nu_{o,t} \sim N(0, \sigma_{o,v}^2)
\]

• Bellman equation:

\[
V_d(s_t) = \max_{c_t, i_t, \kappa_t} n_{e,t} u(c_t/n_{e,t}) + \beta p_{e,t} \int_{\varepsilon} \int_{\eta_y} \int_{\eta_o} V(s_{t+1}) dF_e(\eta_o) dF_e(\eta_y) dF(\varepsilon)
\]

\[
a_{t+1} = \tilde{R}_\kappa(f_{t+1})(a_t + y_{e,t} + oop_{e,t} + tr_t - c_t - \pi(i_t) - c_d I(\kappa_t > 0))
\]

\[
f_{t+1} = (1 - \delta)f_t + i_t
\]

\[
\tilde{R}_\kappa(f_{t+1}) = (1 - \kappa_t)\overline{R} + \kappa_t \tilde{R}(f_t).
\]
Calibration results (using Tauchen 1986 discretization of the two processes):
  - decomposition of wealth inequality

- importance of financial knowledge: accounting for 30-40% of wealth inequality of the retired, even more important than replacement rate, demographics and health mortality factors