# Human Capital Investments and Expectations about Career and Family

Sebastian Valet, Johannes Walter

July 7, 2020

# Summary I
Research questions and design

- What do students believe about the consequences of their education choices?
- How do students sort into majors?
- Novel: what role do family variables play in such choices?

- Survey with undergraduate students at NYU on perceptions about consequences of educational choices
- Specifically: choice of a major
- Follow-up survey after six years

# Summary II
Results

- Students believe in importance of consequences for own earnings and family life
- Particularly women, major choice also corresponds to different rates and timing of marriage and fertility
- Belief about marriage market "return" to higher earning majors
- Ex-ante beliefs are systematically related to educational choices and ex-post realized outcomes

# Model I
Human capital investment under uncertainty

- Expected utility for human capital choice at time $\tau$:

$$E_{i,\tau}(V_k) = \sum_{t=\tau+1}^{T} \beta^{t-\tau} \int u_t(X) \, dG_{i,\tau}(X|k,t)$$

- with discount rate *beta* and outcome $X$ for all subsequent periods given a human capital investment $k$

- $G_{i,\tau}(X|k,t)$ is the belief distribution about the outcome given human capital investments $k$

# Model II

- Survey design elicits beliefs $G_{i,\tau}(X|k,t)$ about the choice of a major
- Belief distrubtions have four characteristics:
  - reflect individual *uncertainty*
  - are *heterogenous*
  - can be *incorrect*
  - can evolve over time due to *learning*
- Natural limitation: elicitation of degree of uncertainty ask Jogibär if put here; also how do they elicit?

# Model III
### Different effects of human capital choices

- Ex-ante individual differences in beliefs

$$\Delta_{G,i}(k, k') = G_i(X|k, t) - G_i(X|k', t)$$

- Ex-post individual differences in potential outcomes

$$\Delta_{F,i}(k, k') = F_i(X|k, t) - F_i(X|k', t)$$

- Ex-post individual differences realized outcomes

$$\Delta_H(k, k') = H(X|k, t) - H(X|k', t)$$

with $H(X|k, t) = \frac{1}{M_k} \sum_{t \in \Omega_k} F_i(k = k^*, t)$

# Current Population Characteristics I

- Earnings, employment, and marriage data for the US population using the 2009

- Not suited for causal inference; needs not reflect the student's beliefs

- Data from older cohort; includes not only high-ability participants

- But data is suited to document that career and family outcomes differ by educational choices in observational data

# Current Population Characteristics I

- Earnings, employment, and marriage data for the US population using the 2009
- Not suited for causal inference; needs not reflect the student's beliefs
- Data from older cohort; includes not only high-ability participants
- But data is suited to document that career and family outcomes differ by educational choices in observational data

# Current Population Characteristics II

Table 2: Descriptive Statistics of 2009 ACS Data

| | Age 23 | | Age 30 | | Age 45 | |
|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female |
| **Earnings (in $10,000s)** | | | | | | |
| Science/Business | 3.33 | 3.22 | 6.74 | 5.48+++ | 11.61 | 7.46+++ |
| | (2.15) | (2.19) | (4.81) | (3.15) | (9.79) | (6.49) |
| Humanities | 2.51 | 2.57 | 5.40 | 4.47+++ | 9.07 | 5.93+++ |
| | (1.33) | (1.88) | (4.20) | (2.71) | (8.48) | (5.67) |
| No Degree | 2.54 | 2.15+++ | 4.21 | 3.08+++ | 5.70 | 3.88+++ |
| | (1.52) | (1.41) | (2.50) | (1.59) | (4.13) | (2.57) |
| p-value[a] | 0 | 0 | 0 | 0 | 0 | 0 |
| **Spousal Earnings (in $10,000s)** | | | | | | |
| Science/Business | 3.41 | 4.75+++ | 5.26 | 8.25+++ | 7.44 | 12.68+++ |
| | (2.09) | (3.11) | (3.44) | (5.79) | (6.69) | (10.15) |
| Humanities | 2.27 | 3.49+++ | 4.30 | 6.66+++ | 5.71 | 9.85+++ |
| | (1.33) | (1.93) | (2.61) | (5.64) | (4.72) | (9.42) |
| No Degree | 2.21 | 3.50+++ | 3.24 | 4.82+++ | 3.76 | 6.36+++ |
| | (1.13) | (1.93) | (1.86) | (2.92) | (2.59) | (4.81) |
| p-value | 0 | 0.003 | 0 | 0 | 0 | 0 |
| **Full-time Employed (%)** | | | | | | |
| Science/Business | 38.5 | 42.4+++ | 80.86 | 64.40+++ | 82.68 | 58.42+++ |
| Humanities | 30.9 | 36.2+++ | 72.96 | 57.92+++ | 75.86 | 52.07+++ |
| No Degree | 40.1 | 34.4+++ | 66.53 | 46.51+++ | 67.88 | 52.44+++ |
| p-value | 0 | 0 | 0 | 0 | 0 | 0 |
| **Married (%)** | | | | | | |
| Science/Business | 8.2 | 15.9+++ | 61.72 | 67.49+++ | 80.79 | 76.14+++ |
| Humanities | 11.5 | 15.3+++ | 55.7 | 64.94+++ | 76.58 | 74.51+ |
| No Degree | 15.2 | 26.4+++ | 54.86 | 59.29+++ | 69.3 | 69.65 |
| p-value | 0 | 0 | 0 | 0 | 0 | 0 |

Earnings and spousal earnings shown in $10,000s.
Mean (standard deviation) shown for the continuous outcomes.
+++, ++, + gender differences statistically significant at the 1, 5, and 10% levels,
respectively. Symbols denoted on female column.
[a] p-value of a F-test of the joint equality of means across majors. p-value of zero
implies p-vlaue < 0.001.

# Earnings Beliefs: Earnings Levels

- male students believe to earn more than female students at each age
- all students believe to see rapid growth in earnings
- students believe to see substantially smaller earning growth if they don't major in science/business
- Perceived gender gap is largest in science/business and at later stages

# Current Population Characteristics II

Table 2: Descriptive Statistics of 2009 ACS Data

| | Age 23 | | Age 30 | | Age 45 | |
|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female |
| **Earnings (in \$10,000s)** | | | | | | |
| Science/Business | 3.33 | 3.22 | 6.74 | 5.48+++ | 11.61 | 7.46+++ |
| | (2.15) | (2.19) | (4.81) | (3.15) | (9.79) | (6.49) |
| Humanities | 2.51 | 2.57 | 5.40 | 4.47+++ | 9.07 | 5.93+++ |
| | (1.33) | (1.88) | (4.20) | (2.71) | (8.48) | (5.67) |
| No Degree | 2.54 | 2.15+++ | 4.21 | 3.08+++ | 5.70 | 3.88+++ |
| | (1.52) | (1.41) | (2.50) | (1.59) | (4.13) | (2.57) |
| p-value$^a$ | 0 | 0 | 0 | 0 | 0 | 0 |
| **Spousal Earnings (in \$10,000s)** | | | | | | |
| Science/Business | 3.41 | 4.75+++ | 5.26 | 8.25+++ | 7.44 | 12.68+++ |
| | (2.09) | (3.11) | (3.44) | (5.79) | (6.69) | (10.15) |
| Humanities | 2.27 | 3.49+++ | 4.30 | 6.66+++ | 5.71 | 9.85+++ |
| | (1.33) | (1.93) | (2.61) | (5.64) | (4.72) | (9.42) |
| No Degree | 2.21 | 3.50+++ | 3.24 | 4.82+++ | 3.76 | 6.36+++ |
| | (1.13) | (1.93) | (1.86) | (2.92) | (2.59) | (4.81) |
| p-value | 0 | 0.003 | 0 | 0 | 0 | 0 |
| **Full-time Employed (%)** | | | | | | |
| Science/Business | 38.5 | 42.4+++ | 80.86 | 64.40+++ | 82.68 | 58.42+++ |
| Humanities | 30.9 | 36.2+++ | 72.96 | 57.92+++ | 75.86 | 52.07+++ |
| No Degree | 40.1 | 34.4+++ | 66.53 | 46.51+++ | 67.88 | 52.44+++ |
| p-value | 0 | 0 | 0 | 0 | 0 | 0 |
| **Married (%)** | | | | | | |
| Science/Business | 8.2 | 15.9+++ | 61.72 | 67.49+++ | 80.79 | 76.14+++ |
| Humanities | 11.5 | 15.3+++ | 55.7 | 64.94+++ | 76.58 | 74.51+ |
| No Degree | 15.2 | 26.4+++ | 54.86 | 59.29+++ | 69.3 | 69.65 |
| p-value | 0 | 0 | 0 | 0 | 0 | 0 |

Earnings and spousal earnings shown in \$10,000s.
Mean (standard deviation) shown for the continuous outcomes.
+++, ++, + gender differences statistically significant at the 1, 5, and 10% levels,
respectively. Symbols denoted on female column.
$^a$ p-value of a F-test of the joint equality of means across majors. p-value of zero
implies p-vlaue < 0.001.

# Earnings Beliefs: Earnings Levels

Table 3: Self Earnings

|  | Age 23 | | Age 30 | | Age 45 | |
|---|---|---|---|---|---|---|
|  | Male | Female | Male | Female | Male | Female |
| **Panel A: Levels (in 10,000s of dollars)** | | | | | | |
| Science/Business | 5.93 | 5.39 | 13.74 | 10.86++ | 19.00 | 13.81+++ |
|  | (7.32) | (4.66) | (16.61) | (9.31) | (22.38) | (14.12) |
| Humanities | 4.71 | 3.94 | 6.87 | 6.86 | 11.03 | 9.60 |
|  | (7.38) | (3.51) | (5.51) | (7.4) | (13.53) | (11.75) |
| No Degree | 3.50 | 2.45++ | 5.07 | 3.27++ | 8.97 | 5.86+++ |
|  | (7.54) | (1.16) | (11.0) | (4.56) | (15.95) | (10.22) |
| Overall | 5.60 | 4.68+ | 12.95 | 9.21+++ | 18.44 | 12.33+++ |
|  | (7.36) | (3.81) | (16.35) | (8.45) | (22.52) | (13.90) |
| | | | | | | |
| **Panel B: Individual Log Differences** | | | | | | |
| Sci/Business versus. Humanities | .267*** | .304*** | .523*** | .425***++ | .446*** | .347***+ |
|  | (.019) | (.017) | (.048) | (.025) | (.051) | (.026) |
| Graduate versus. No Degree | .594*** | .642*** | 1.022*** | 1.038*** | .829*** | .833*** |
|  | (.033) | (.026) | (.055) | (.037) | (.054) | (.038) |

Panel A shows the mean and standard deviations of expected earnings (in $10,000s). +++, ++, + denote gender differences are statistically different at the 1, 5, and 10% levels, respectively.
Panel B shows the avg. log differences and standard deviations in parentheses. ***, **, * denote the means are statistically different from zero at the 1, 5, and 10% levels, respectively. +++, ++, + (shown on the female column) denote gender differences are statistically different at the 1, 5, and 10% levels, respectively.

# Earnings Growth

Table 4: Earnings growth beliefs

| | Age 23-30 | | Age 30-45 | |
| --- | --- | --- | --- | --- |
| | Male | Female | Male | Female |
| **Panel A: Levels (in %)** | | | | |
| Science/Business | .67 | .63 | .25 | .19 |
| | (.72) | (.65) | (.47) | (.54) |
| Humanities | .41 | .51+ | .32 | .27 |
| | (.56) | (.53) | (.45) | (.52) |
| No Degree | .23 | .21 | .47 | .43 |
| | (.78) | (.55) | (.74) | (.58) |
| Overall | .66 | .6 | .29 | .23 |
| | (.73) | (.58) | (.48) | (.52) |
| **Panel B: Individual differences** | | | | |
| Sci/Business versus. Humanities | .26*** | .12***+++ | -.08* | -.08*** |
| | (.05) | (.03) | (.04) | (.03) |
| Graduate versus. No Degree | .42*** | .39*** | -.19*** | -.2*** |
| | (.06) | (.03) | (.06) | (.03) |

Panel A shows the mean and standard dev of beliefs about earnings growth (in %).
+++, ++, + denote gender differences are statistically different at the 1, 5, and 10%
levels, respectively.
Panel B shows average log differences and standard deviations in parentheses.
***, **, * denote means are statistically different from zero at the 1, 5, and 10%
levels, respectively. +++, ++, + (shown on the female column) denote gender differences
are statistically different at the 1, 5, and 10% levels, respectively.

# Earnings Uncertainty

Table 5: Age 30 Earnings Uncertainty - Std deviations from fitting a Beta Distribution

|  | Male | Female |
|---|---|---|
| **Panel A: Levels (in $10,000)** | | |
| Science/Business | 9.17 | 9.49 |
|  | (1.44) | (2.48) |
| Humanities | 10.34 | 10.01 |
|  | (27.44) | (2.32) |
| No Degree | 14.73 | 15.27 |
|  | (7.34) | (7.53) |
| Overall | 9.71 | 9.68 |
|  | (2.02) | (2.01) |
| | | |
| **Panel B: Individual differences** | | |
| Sci/Business versus. Humanities | -.11*** | -.057***+++ |
|  | (.014) | (.012) |
| Graduate versus. No Degree | -.305*** | -.335*** |
|  | (.052) | (.043) |

Panel A shows the mean and std dev of age 30 earnings uncertainty beliefs (in $10,000). Uncertainty is the standard deviation of the individual-specific (beta-) fitted earnings distribution.
+++, ++, + denote gender differences statistically different at the 1, 5, and 10% levels, respectively.
Panel B shows average log differences and standard deviations in parentheses. ***, **, * denote means are statistically diff from 0 at the 1, 5, and 10% levels, respectively. +++, ++, + (shown on female column) denote gender differences are statistically different at the 1, 5, and 10% levels, respectively.

# Beliefs about Marriage

Table 6: Beliefs about Marriage

| Prob Marriage: | Age 23 | | Age 30 | | Age 45 | |
|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female |
| **Panel A: Levels** (0-1 scale) | | | | | | |
| Science/Business | .148 | .167 | .593 | .594 | .804 | .778 |
| | (.207) | (.214) | (.286) | (.271) | (.248) | (.253) |
| Humanities | .152 | .182 | .601 | .66++ | .797 | .800 |
| | (.214) | (.229) | (.291) | (.268) | (.253) | (.246) |
| No Degree | .153 | .221+++ | .535 | .605++ | .727 | .743 |
| | (.219) | (.26) | (.329) | (.29) | (.302) | (.287) |
| Overall | .149 | .179 | .589 | .634+ | .797 | .793 |
| | (.213) | (.225) | (.288) | (.266) | (.25) | (.242) |
| **Panel B: Individual Log Differences** | | | | | | |
| Sci/Business versus. Humanities | -.008 | -.096* | -.024 | -.147***++ | .013 | -.020 |
| | (.046) | (.053) | (.042) | (.039) | (.014) | (.024) |
| Graduate versus. No Degree | .075 | -.192**+ | .354*** | .127**++ | .317*** | .161*** |
| | (.099) | (.091) | (.11) | (.054) | (.09) | (.054) |

Panel A shows the mean and standard deviations of marriage beliefs. +++, ++, + denote gender diffs are
statistically significant at the 1, 5, and 10% levels, respectively.
Panel B shows the average log differences and standard deviations in parentheses.
***, **, * denote the means are statistically different from zero at the 1, 5, and 10% levels,
respectively. +++, ++, + (shown on the female column) denote gender differences are statistically
significant at the 1, 5, and 10% levels, respectively.

# Beliefs about Potential Spousal Earnings

Table 7: Beliefs about Potential Spousal Earnings, Conditional on Own Major (and Own Age)

| | Age 23 | | Age 30 | | Age 45 | |
|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female |
| **Panel A: Levels (in 10,000s of dollars)** | | | | | | |
| Science/Business | 5.06 | 5.74+ | 9.00 | 10.76++ | 11.29 | 13.68+ |
| | (4.12) | (3.92) | (7.72) | (9.14) | (13.25) | (13.67) |
| Humanities | 4.52 | 4.75 | 7.05 | 7.86 | 8.02 | 11.07+++ |
| | (7.35) | (3.75) | (8.93) | (7.69) | (7.95) | (12.90) |
| No Degree | 4.58 | 3.46 | 4.57 | 5.54 | 6.25 | 7.76 |
| | (11.99) | (2.26) | (5.56) | (9.11) | (9.89) | (12.03) |
| Overall | 5.02 | 5.30 | 8.42 | 9.74+ | 10.77 | 12.73 |
| | (5.90) | (3.88) | (7.60) | (8.91) | (13.20) | (13.61) |
| | | | | | | |
| **Panel B: Individual Log Differences** | | | | | | |
| Sci/Business versus. Humanities | .185*** | .198*** | .282*** | .292*** | .241*** | .221*** |
| | (.019) | (.015) | (.044) | (.024) | (.04) | (.026) |
| Graduate versus. No Degree | .432*** | .481*** | .687*** | .741*** | .587*** | .632*** |
| | (.048) | (.028) | (.05) | (.041) | (.054) | (.039) |

Panel A shows the mean and standard dev of beliefs about spouse's expected earnings (in $10,000s) conditional on own major.
+++, ++, + denote gender differences are statistically different at the 1, 5, and 10% levels, respectively.
Panel B shows avg. log differences and standard deviations in parentheses. ***, **, * denote means are statistically different from zero at the 1, 5, and 10% levels, respectively. +++, ++, + (shown on the female column) denote gender differences are statistically different at the 1, 5, and 10% levels, respectively.