Demand for Post-compulsory Education: The Choice Between Academic and Vocational Tracks

Cristina Lopez-Mayan

CEMFI

Royal Economic Society Fifth PhD Presentation Meeting

London 2010
Objective (I)

Interest in understanding the **schooling decisions** of Spanish youth following compulsory education with a focus on **Academic-Vocational choice**
Interest in understanding the **schooling decisions** of Spanish youth following compulsory education with a focus on **Academic-Vocational choice**

Standard **human capital model**: schooling choices depend on expected life-cycle earnings
Objective (I)

Interest in understanding the **schooling decisions** of Spanish youth following compulsory education with a focus on **Academic-Vocational choice**

Standard **human capital model**: schooling choices depend on expected life-cycle earnings

**Data**: new rich microdataset on educational histories of Spanish youth (*ETEFIL: Encuesta de Transición Educativa-Formativa e Inserción Laboral 2005*)
Objective (I)

Interest in understanding the **schooling decisions** of Spanish youth following **compulsory education** with a focus on **Academic-Vocational choice**

Standard **human capital model**: schooling choices depend on expected life-cycle earnings

**Data**: new rich microdataset on educational histories of Spanish youth (*ETEFIL: Encuesta de Transición Educativa-Formativa e Inserción Laboral 2005*)

**Structural approach**: sequential model of schooling choices taylored to these data
Interest in understanding the **schooling decisions** of Spanish youth following compulsory education with a focus on **Academic-Vocational choice**

Standard **human capital model**: schooling choices depend on expected life-cycle earnings

**Data**: new rich microdataset on educational histories of Spanish youth *(ETEFIL: Encuesta de Transición Educativa-Formativa e Inserción Laboral 2005)*

**Structural approach**: sequential model of schooling choices tailored to these data

**Final purpose**: answer questions of potential interest (effect on dropout rates and track decision of changes in life-cycle earnings, relative wages, unemployment rates)
Objective (II)

Post-compulsory education in Spain

- **Compulsory Education (ESO)** (up to age of 16)
  - Academic High School *(Bachillerato)* (2 years)
  - Vocational High School *(FP medio)* (1 or 2 years)
  - Vocational College *(FP superior)*
  - University
Contributions

1. Structural model on schooling decisions reflecting academic and vocational tracks in post-compulsory education

2. Options following completion of each track linked to expected life-time earnings

- Estimation methodology:
  - Unobserved heterogeneity (finite-mixture approach) + finite horizon model

Other papers:

- Structural approach: Eckstein and Wolpin (1999); Keane and Wolpin (1997); Arcidiacono (2004, 2005)
The database (I)

- *Encuesta de Transición Educativa-Formativa e Inserción Laboral 2005 (ETEFIL)*: new microdataset produced by the Spanish Statistics Institute in 2005
- Representative sample of people who completed *compulsory education* in 2000/01
- **Retrospective interview:** People are asked about their schooling and labour market participation since 2000/01 until the moment of the interview (May-July 2005) ⇒ 4 years
- Data cover high school education and first years following high school completion
- Sample size: **7750 individuals**
The database (II)

**Personal characteristics:** gender, date of birth, nationality, parents’ education, region of residence

**Education:**
- Compulsory: age when finishing, private/semi-private/public school
- Post-compulsory (each year): decision of leaving or attending school, kind of track, grade and program attended, whether she finished that year, private/public school, decision following high school completion

**Wages:** Information on net monthly wage for individuals with a past (for at least 6 months) or current job

Appendix I
The model (I)

- While the individual does not complete high school: decision among academic/vocational track and labor market

- Individual problem finishes with decision following high school diploma (terminal decision: university, vocational college, labor market,...)
  - Options following completion of each track linked to their expected present value of life-cycle earnings (terminal values)
  - If no HS diploma: life-cycle earnings of compulsory schooling diploma

- **Expectations on life-cycle earnings:** youth observes wages of workers with similar characteristics to infer life-time profiles

  Annual wages from the Spanish Wage Structure Survey 2002
The model (II)

- Obtaining a HS diploma requires to accumulate passed grades:
  \[ n_{t+1}^k = n_t^k + c_t^k \]

- Uncertainty on passing a grade: \( p(c_t^k = 1 \mid X_t, d_t = k) \)

- Potentially, individuals finish compulsory schooling with differences in preferences and in the ability to progress in each track

- Source of persistence in decisions and individual self-selection not observed by the econometrician

- Population heterogeneity accounted with \( M \) types: discrete support with \( \pi_m \) the proportion of type \( m \) individuals
The model (III)

\[ V_t(S_t) = \max_{\{d_t\}_{t=1}^{T_i}} \mathbb{E} \left[ \sum_{\tau=t}^{T_i} \beta^{\tau-t} U_t(d_t) \mid S_t \right] \]

\[ V_t(S_t) = \max_k \{ V_t^k(S_t) \} \quad k = \{ AHS, VHS1, VHS2, LM \} \]

\[ V_t^k(S_t) = U_t^k + \beta \mathbb{E} \left( V_{t+1}(S_{t+1}) \mid S_t, d_t = k, c_t^k \right) \]

- Current utility: consumption value of schooling or labor market wage, utility shocks (type I extreme value distribution)

- Expected discounted utility:
  - Expectation taken over shocks and probability of passing
  - Depends on current state variables and decision, life-cycle earnings attached to each terminal option
The model (IV)

Current value of school attendance:

\[ \overline{educ}^{AHS}_t = \mu^{AHS} + \beta_1^{AHS} n_t^{AHS} + \beta_2^{AHS} I(d_{t-1} = LM) \]

\[ \overline{educ}^j_t = \mu^VHS + \beta_1^{VHS} n_t^j + \beta_2^{VHS} I(d_{t-1} = LM) \]

with \( j = \{VHS1, VHS2\} \)

Probability of passing:

\[ p(c_t^k = 1 \mid X_t, d_t = k) = \frac{e^{x'_{k,t} \gamma_{t}^{k}}}{1 + e^{x'_{k,t} \gamma_{t}^{k}}} \]

\( x_{k,t} = \{\text{constant, } n_t^k, \text{ female, finishing compulsory schooling with delay}\} \)
Estimated model used to simulate schooling histories

Exercise: increasing 10% the annual wage earned at each age \( t \)

Five scenarios, each one corresponding to an increase in the wage corresponding to: university, vocational college, vocational HS, academic HS and compulsory schooling
### Quantifying the impact of expected life-time earnings (II)

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>$w_t^{VHS}$</th>
<th>$w_t^{VC}$</th>
<th>$w_t^{UN}$</th>
<th>$w_t^{AHS}$</th>
<th>$w_t^{CompDip}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>With a HS diploma</td>
<td>80.03</td>
<td>87.46</td>
<td>85.14</td>
<td>80.77</td>
<td>82.92</td>
<td>72.19</td>
</tr>
<tr>
<td>VHS diploma</td>
<td>19.86</td>
<td>42.52</td>
<td>32.98</td>
<td>11.82</td>
<td>21.30</td>
<td>15.73</td>
</tr>
<tr>
<td>Only attend academic HS</td>
<td>76.68</td>
<td>56.83</td>
<td>65.00</td>
<td>86.22</td>
<td>76.40</td>
<td>75.34</td>
</tr>
<tr>
<td>Only attend vocational HS</td>
<td>14.17</td>
<td><strong>34.50</strong></td>
<td>24.49</td>
<td>7.45</td>
<td>14.12</td>
<td><strong>11.54</strong></td>
</tr>
<tr>
<td>Never attend post-compulsory ed.</td>
<td>3.35</td>
<td><strong>1.06</strong></td>
<td>2.05</td>
<td><strong>1.37</strong></td>
<td><strong>1.83</strong></td>
<td><strong>10.32</strong></td>
</tr>
<tr>
<td>Switch between AHS and VHS</td>
<td>5.79</td>
<td>7.61</td>
<td>8.46</td>
<td>4.97</td>
<td>7.65</td>
<td>2.80</td>
</tr>
</tbody>
</table>

|                          |          |            |            |            |            |            |
| High school completion rate | 82.80 | **88.39** | 86.92      | 81.89      | **84.46**  | 80.50    |
| Academic HS completion rate | 77.75 | 78.01 | 77.67 | 78.11 | 77.63 | **77.86** |
| Vocational HS completion rate | 79.64 | **88.30** | 85.20 | 76.92 | 81.15 | **79.21** |

7750 simulated individuals
Concluding remarks

- Structural model of schooling decisions following compulsory education to analyse the effect of expected life-time earnings
  - Academic and Vocational tracks
  - Options following completion of each track linked to expected life-time earnings
- Estimated using a new Spanish microdataset on schooling histories
- Exercises show how changes in life-cycle earnings affect schooling choices
- Elements to be incorporated: unemployment, unobserved ability
- Potential to show the effects of unemployment shocks or different wage structures on schooling (track) choices
- Model not specific for Spain
Observed decisions of Spanish youth (ETEFIL 2005)

- Post-compulsory decisions:
  - Attend academic track: 80%
  - Attend vocational track: 20%
  - Never attend school: 5%

- Few people (5%) moving between academic and vocational tracks

- Following an academic diploma: 73% university and 19% vocational college

- Following a vocational diploma: 78% labor market

- Dropout rate in post-compulsory education: 20%
  - 21% in academic track
  - 33% in vocational track
Descriptive statistics

Choice distribution of individuals who start in each school-year without having completed high school education in the previous year

<table>
<thead>
<tr>
<th>School-year</th>
<th>Academic</th>
<th>Vocational 1</th>
<th>Vocational 2</th>
<th>Labor Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001/02</td>
<td>79.57</td>
<td>3.66</td>
<td>11.11</td>
<td>5.66</td>
</tr>
<tr>
<td>2002/03</td>
<td>74.72</td>
<td>2.58</td>
<td>11.41</td>
<td>11.29</td>
</tr>
<tr>
<td>2003/04</td>
<td>51.26</td>
<td>2.56</td>
<td>11.49</td>
<td>35.69</td>
</tr>
<tr>
<td>2004/05</td>
<td>27.63</td>
<td>1.06</td>
<td>6.48</td>
<td>64.83</td>
</tr>
</tbody>
</table>

Population weights

Decision following high school completion (%)

<table>
<thead>
<tr>
<th></th>
<th>UNIV</th>
<th>VC</th>
<th>VHS</th>
<th>AHS</th>
<th>LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic high school diploma (83%)</td>
<td>72.64</td>
<td>18.97</td>
<td>1.52</td>
<td>-</td>
<td>6.87</td>
</tr>
<tr>
<td>Vocational high school diploma (17%)</td>
<td>-</td>
<td>7.50</td>
<td>7.72</td>
<td>7.04</td>
<td>77.74</td>
</tr>
</tbody>
</table>

7750 individuals (population weights). UNIV: University, VC: Vocational college, AHS: Academic high school, VHS: Vocational high school, LM: Labor market
Wages in the model

\textbf{Wages in } U_{t}^{LM}: \\
- Subsample of individuals who not attend school and with data on wages
- Each period $t$, log of real annual wages regressed on female and region dummies. Estimated coefficients to impute $wage_t$

\textbf{Expected life-cycle earnings:} \\
- Spanish Wage Structure Survey 2002: real gross annual wages of Spanish workers aged 25-50
- Regression of logwages on female, region, age and age squared, by schooling level
- Coefficients to compute present values of life-cycle earnings for UNIV, VC, AHS, VHS, compulsory schooling
- Terminal values of UNIV, VC, AHS, VHS, weighted by probability of graduation
## Relative premium

<table>
<thead>
<tr>
<th></th>
<th>$EW_{\text{UNAHS}}^{\text{UN}}$</th>
<th>$EW_{\text{VCAHS}}^{\text{VC}}$</th>
<th>$EW_{\text{VHSAHS}}^{\text{VHS}}$</th>
<th>$EW_{\text{VHSAHS}}^{\text{LM}}$</th>
<th>$EW_{\text{VHSAHS}}^{\text{VC}}$</th>
<th>$EW_{\text{VHSVHS}}^{\text{AHS}}$</th>
<th>$EW_{\text{VHSVHS}}^{\text{VHS}}$</th>
<th>$EW_{\text{VHSVHS}}^{\text{LM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.56</td>
<td>1.38</td>
<td>1.35</td>
<td>1.43</td>
<td>1.37</td>
<td>1.37</td>
<td>1.31</td>
<td>1.37</td>
</tr>
<tr>
<td>$\Delta 10% \text{VHS}$</td>
<td>1.56</td>
<td>1.38</td>
<td>1.43</td>
<td>1.43</td>
<td>1.41</td>
<td>1.43</td>
<td>1.44</td>
<td>1.51</td>
</tr>
<tr>
<td>$\Delta 10% \text{VC}$</td>
<td>1.56</td>
<td>1.49</td>
<td>1.35</td>
<td>1.43</td>
<td>1.46</td>
<td>1.37</td>
<td>1.31</td>
<td>1.37</td>
</tr>
<tr>
<td>$\Delta 10% \text{Comp}$</td>
<td>1.42</td>
<td>1.25</td>
<td>1.23</td>
<td>1.3</td>
<td>1.24</td>
<td>1.25</td>
<td>1.19</td>
<td>1.24</td>
</tr>
<tr>
<td>$\Delta 10% \text{Univ}$</td>
<td>1.66</td>
<td>1.38</td>
<td>1.35</td>
<td>1.43</td>
<td>1.37</td>
<td>1.37</td>
<td>1.31</td>
<td>1.37</td>
</tr>
<tr>
<td>$\Delta 10% \text{AHS}$</td>
<td>1.61</td>
<td>1.41</td>
<td>1.4</td>
<td>1.57</td>
<td>1.37</td>
<td>1.45</td>
<td>1.31</td>
<td>1.37</td>
</tr>
</tbody>
</table>

$W_{\text{CompDip}} = 1$

### Variation in the relative premium with respect to baseline

<table>
<thead>
<tr>
<th></th>
<th>$EW_{\text{UNAHS}}^{\text{UN}}$</th>
<th>$EW_{\text{VCAHS}}^{\text{VC}}$</th>
<th>$EW_{\text{VHSAHS}}^{\text{VHS}}$</th>
<th>$EW_{\text{VHSAHS}}^{\text{LM}}$</th>
<th>$EW_{\text{VHSAHS}}^{\text{VC}}$</th>
<th>$EW_{\text{VHSVHS}}^{\text{AHS}}$</th>
<th>$EW_{\text{VHSVHS}}^{\text{VHS}}$</th>
<th>$EW_{\text{VHSVHS}}^{\text{LM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta 10% \text{VHS}$</td>
<td>-</td>
<td>-</td>
<td>0.08</td>
<td>-</td>
<td>0.04</td>
<td>0.06</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>$\Delta 10% \text{VC}$</td>
<td>-</td>
<td>0.11</td>
<td>-</td>
<td>-</td>
<td>0.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta 10% \text{Comp}$</td>
<td>-0.14</td>
<td>-0.13</td>
<td>-0.12</td>
<td>-0.13</td>
<td>-0.12</td>
<td>-0.12</td>
<td>-0.12</td>
<td>-0.12</td>
</tr>
<tr>
<td>$\Delta 10% \text{Univ}$</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta 10% \text{AHS}$</td>
<td>0.05</td>
<td>0.03</td>
<td>0.06</td>
<td>0.14</td>
<td>-</td>
<td>0.08</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Variation in the relative premium with respect to baseline
Log-likelihood function:

\[
L(\theta_{u,m}, \theta_c) = \sum_{i=1}^{N} \sum_{t=1}^{T_i-1} p(c_{it} = 1 \mid X_{it}, \theta_c, d_{it}) + \sum_{i=1}^{N} \log \sum_{m=1}^{M} \pi_m \prod_{t=1}^{T_i} P(d_{it} \mid X_{it}, \theta_{u,m}, \theta_c)
\]
Appendix

Solution and estimation (II)

Outer algorithm:
Step 1: $\hat{\theta}_c$ and guess for $\hat{P}^0$, $\theta_{u,m}$, $\pi_m$, $\forall$ $m$
Step 2: Construct the alternative-specific value functions using Hotz-Miller's representation in terms of $\hat{P}^0$
Step 3: Construct the pseudo finite-mixture likelihood function

$$Q(\theta_{u,m}, \pi_m, \hat{P}^0, \hat{\theta}_c) = \sum_{i=1}^{N} \log \sum_{m=1}^{M} \pi_m \prod_{t=1}^{T_i} P(d_{it} \mid X_{it}, \theta_{u,m}, \hat{\theta}_c, \hat{P}^0)$$

Step 4: Inner algorithm: Maximization
Step 5: Update CCPs: $\hat{P}^1 \equiv \hat{P}(d_t = k \mid X_t, \hat{\theta}_{u,m}^1, \hat{\theta}_c, \hat{P}^0)$
Step 6: Repeat until convergence in probability space

Result: $\hat{\theta}_{u,m}$ and $\hat{\pi}_m$, $\forall$ $m$

Cristina Lopez-Mayan (CEMFI)
## Model fit (I)

Choice distribution of individuals who start in each period without having completed high school education

<table>
<thead>
<tr>
<th>Period</th>
<th>AHS</th>
<th>VHS</th>
<th>LM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Predicted</td>
<td>Actual</td>
</tr>
<tr>
<td>1</td>
<td>79.57</td>
<td>81.43</td>
<td>14.77</td>
</tr>
<tr>
<td>2</td>
<td>74.72</td>
<td>75.36</td>
<td>13.99</td>
</tr>
<tr>
<td>3</td>
<td>51.26</td>
<td>55.41</td>
<td>13.05</td>
</tr>
<tr>
<td>4</td>
<td>27.63</td>
<td>18.35</td>
<td>7.55</td>
</tr>
</tbody>
</table>

AHS: Academic high school, VHS: Vocational high school, LM: labor market

7750 simulated individuals

### Completion rates (%)

<table>
<thead>
<tr>
<th></th>
<th>AHS</th>
<th>VHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Predicted</td>
</tr>
<tr>
<td>Full sample</td>
<td>78.62</td>
<td>77.75</td>
</tr>
<tr>
<td>Males</td>
<td>75.16</td>
<td>74.42</td>
</tr>
<tr>
<td>Females</td>
<td>81.30</td>
<td>80.24</td>
</tr>
</tbody>
</table>

Graduated over enrolled individuals

7750 simulated individuals
## Model fit (II)

<table>
<thead>
<tr>
<th>Category</th>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals with a high school diploma</td>
<td>76.00</td>
<td>80.03</td>
</tr>
<tr>
<td>Individuals with a vocational HS diploma</td>
<td>17.31</td>
<td>19.86</td>
</tr>
<tr>
<td>Individuals who only attend academic high school</td>
<td>75.24</td>
<td>76.68</td>
</tr>
<tr>
<td>Individuals who only attend vocational high school</td>
<td>14.99</td>
<td>14.17</td>
</tr>
<tr>
<td>Individuals who never attend post-compulsory education</td>
<td>5.08</td>
<td>3.35</td>
</tr>
<tr>
<td>Individuals who switch between academic and vocational HS</td>
<td>4.69</td>
<td>5.79</td>
</tr>
</tbody>
</table>

7750 simulated individuals