Age at immigration and the education outcomes of children

Miles Corak
Graduate School of Public and International Affairs

Université d’Ottawa | University of Ottawa
Narrowing the Objectives
How does migration influence the well being of children?

1. There are a number of different ways to answer the question depending upon the counterfactual

Would the child have been better off if:

• The parents had decided not to migrate?
• The parents had migrated to a different country?
• The parents had migrated at a different point in the child’s life cycle?
• The parents had given birth in the new country?
Narrowing the Objectives
How does migration influence the well being of children?

1. There are a number of different ways to answer the question depending upon the counterfactual

2. We focus on how age at arrival in the new country “influences” the long run outcomes of the child
   – Would the same child have done better if he or she had arrived at a different point in his or her life cycle?
   – We don’t observe the same child in two different states: we cannot rely on a randomized experiment; and we must use observational data potentially subject to selection on unobservables
   – The need for an “identification strategy”
High school graduation and age at arrival: For men a threshold at 8 and for women at 7 years
The “Critical Period Hypothesis”
As a way to structure the descriptive analysis and frame the causal question

1. **There is a distinct threshold beyond which outcomes are inferior**
   - Language acquisition reflects distinct patterns in brain development with windows of opportunity opening and then closing
   - Once the window of opportunity has passed, language acquisition is more difficult and not attained to the same degree
   - These thresholds are associated with the onset of puberty
The “Critical Period Hypothesis”
As a way to structure the descriptive analysis and frame the causal question

1. **There is a distinct threshold beyond which outcomes are inferior**
   - Language acquisition reflects distinct patterns in brain development with windows of opportunity opening and then closing
   - Once the window of opportunity has passed, language acquisition is more difficult and not attained to the same degree
   - These thresholds are associated with the onset of puberty

2. **Implications for empirical analysis**
   - There should be distinct breaks in outcomes, at ages 12 or 15 – though the exact definition of the thresholds is not clear in the literature
   - An age dependent relationship would suggest that other processes are also at work, biological or social
   - It is not entirely clear if discrete breaks represent “Critical Social Periods” as opposed to “Critical Learning Periods,” telling us more about the host country than about the children
The Critical Period Hypothesis as a Regression Discontinuity Design

1. *A regression discontinuity design*
   - If the relationship between age and outcomes is deterministic then groups on either side of the threshold are comparable: one is offered a “treatment”; the other serves as control
   - There is no selection bias and the causal impact of the intervention is the difference in outcomes between those just below the threshold and those just above

2. *A regression discontinuity design?*
   - A need to be agnostic: theory is not a perfect guide
   - Thresholds could be “fuzzy” rather than “sharp”
   - For the purpose of description let the data tell us where the threshold is
   - Then define a binary treatment that can be used in instrumental variables estimation
The Data is from the 2006 Canadian Census
All 35 to 55 year olds who came as children

1. **“One-in-five” Census sample**
   - Everyone is in the census, but one-in-five individuals respond to a long questionnaire
   - Respondents to the “long” form represent 20% of the population
   - Detailed socio-economic information, including immigrant status and year of arrival

2. **All individuals 35 to 55 who are immigrants and arrived between 0 and 20 years of age**
   - An adult population of about 10.2 million individuals represented by a sample of 2.1 million individuals
   - Of these about 470,000 are immigrants, and we end up using about 137,000 observations on individuals who arrived in the country between 0 and 20 years of age
   - When appropriately weighted these individuals represent just over 700,000 Canadians
A flavour of the data and the adult outcomes
The use of group averages for the time being

- Sample sizes of about 55,000 for each gender and about 3,000 at each Age at Arrival
- There are large increases in the proportion of the sample at 18, 19 and 20 years of age
  - These are excluded as they may reflect independent migrants
  - CRC defines a child as someone under 18
- Possibility of selection effects
  - More altruistic parents may decide to migrate when the children are young in anticipation of the impact the decision may have on their children
  - Their children will perform better than a randomly selected group, making causal inferences across Age at Arrival thresholds suspect
Three related cautions in offering a causal interpretation

1. **Self-selection bias if parents make migration decisions on the basis of the impact Age at Arrival will have on their children**
   - Unobserved family characteristics may vary across the Age at Arrival distribution so that children who arrive earlier are raised in a family environment more favourable to their long run success.
   - The results overstate the relative success rates during the early years, compared to a randomly selected group of individuals.

2. **Heterogeneity along other dimensions such as country of origin**
   - Immigrants come from a large number of source countries, and therefore previous exposure to English and French could vary.
   - If those arriving at younger ages are disproportionately from English or French countries the results will be biased.

3. **Age at Arrival is not the same thing as Age at First Exposure**
   - Immigrants come from a large number of source countries, and therefore previous exposure to English and French could vary.
   - If those arriving at younger ages are disproportionately from English or French countries the results will be biased.
All countries of origin were classified according to the linguistic distance from English adapting Chiswick and Miller (2005)

<table>
<thead>
<tr>
<th>“Low” Score: countries with the highest linguistic distance from English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greece</td>
<td>Anguilla</td>
</tr>
<tr>
<td>Palestine/West Bank/Gaza</td>
<td>Antigua and Barbuda</td>
</tr>
<tr>
<td>Peoples’ Republic of China</td>
<td>Bahamas</td>
</tr>
<tr>
<td>Macau</td>
<td>Barbados</td>
</tr>
<tr>
<td>Japan</td>
<td>Bermuda</td>
</tr>
<tr>
<td>North Korea</td>
<td>Cayman Islands</td>
</tr>
<tr>
<td>South Korea</td>
<td>Dominica</td>
</tr>
<tr>
<td>Taiwan</td>
<td>Grenada</td>
</tr>
<tr>
<td>Brunei Darussalam</td>
<td>Jamaica</td>
</tr>
<tr>
<td>Laos</td>
<td>Montserrat</td>
</tr>
<tr>
<td>Vietnam</td>
<td>Saint Kitts and Nevis</td>
</tr>
<tr>
<td>Fiji</td>
<td>Saint Lucia</td>
</tr>
<tr>
<td>Hong Kong and India are categorized separately</td>
<td>Saint Vincent and the Grenadines</td>
</tr>
<tr>
<td></td>
<td>Trinidad and Tobago</td>
</tr>
<tr>
<td></td>
<td>British Virgin Islands</td>
</tr>
<tr>
<td></td>
<td>United States and United Kingdom are categorized separately</td>
</tr>
</tbody>
</table>

All countries of origin were classified according to the linguistic distance from English according to Chiswick and Miller (2005)

<table>
<thead>
<tr>
<th>“High” Score: lowest linguistic distance from English</th>
<th>“Medium “ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costa Rica</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>El Salvador</td>
<td>Czech Republic</td>
</tr>
<tr>
<td>Guatemala</td>
<td>Slovakia</td>
</tr>
<tr>
<td>Honduras</td>
<td>Hungary</td>
</tr>
<tr>
<td>Mexico</td>
<td>Poland</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>Estonia</td>
</tr>
<tr>
<td>Panama</td>
<td>Finland</td>
</tr>
<tr>
<td>Cuba</td>
<td>Bosnia and Herzegovina</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>Croatia</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>Macedonia</td>
</tr>
<tr>
<td>Argentina</td>
<td>Serbia and Montenegro</td>
</tr>
<tr>
<td>Bolivia</td>
<td>Slovenia</td>
</tr>
<tr>
<td>Brazil</td>
<td>Yugoslavia n.o.s.</td>
</tr>
<tr>
<td>Chile</td>
<td>Eretria</td>
</tr>
<tr>
<td>Colombia</td>
<td>Afghanistan</td>
</tr>
<tr>
<td>Ecuador</td>
<td>Iran</td>
</tr>
<tr>
<td>Sweden</td>
<td>Turkey</td>
</tr>
<tr>
<td>Italy and Portugal are categorized separately</td>
<td>Phillipines</td>
</tr>
</tbody>
</table>

France:
- Saint Pierre and Miquelon
- Guadeloupe
- Haiti
- Martinique
- France
- Réunion
- Gabon
Frequency distribution of the data (men) by linguistic distance from English in country of origin

Sample Frequency by Language and Region of Origin

Age at arrival

Source countries in which English is spoken do not show a critical period (men)
Critical periods are most evident for those children (boys) from Italy and Portugal.
Critical periods are also evident for those children (boys) from countries farthest in linguistic distance from English.
Proportion of sample (men) with a non-official language as a mother tongue by linguistic distance.
Table 2
Least squares regression results for most likely thresholds of Age at Arrival impacts on
not having graduated from High School: Men and Women

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th></th>
<th>Men</th>
<th></th>
<th>Women</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T = 7</td>
<td>T = 7</td>
<td>T = 8</td>
<td>T = 10</td>
<td>T = 10</td>
<td>T = 8</td>
</tr>
<tr>
<td>Age at arrival</td>
<td>0.000105</td>
<td>-0.00183</td>
<td>-0.000805</td>
<td>0.00223</td>
<td>0.00124</td>
<td>-0.000532</td>
</tr>
<tr>
<td>T</td>
<td>-0.0138</td>
<td>-0.00805</td>
<td>0.00250</td>
<td>-0.0123</td>
<td>-0.0100</td>
<td>-0.0107</td>
</tr>
<tr>
<td>T x Age at Arrival</td>
<td><strong>0.0112</strong></td>
<td><strong>0.0106</strong></td>
<td><strong>0.0127</strong></td>
<td><strong>0.0168</strong></td>
<td><strong>0.0152</strong></td>
<td><strong>0.0163</strong></td>
</tr>
<tr>
<td>Mother Tongue</td>
<td>0.112</td>
<td>0.143</td>
<td>0.112</td>
<td>0.220</td>
<td>0.112</td>
<td>0.220</td>
</tr>
<tr>
<td>Region of Origin</td>
<td>none</td>
<td>none</td>
<td>eleven</td>
<td>none</td>
<td>none</td>
<td>eleven</td>
</tr>
<tr>
<td>Constant</td>
<td><strong>0.150</strong></td>
<td><strong>0.0930</strong></td>
<td><strong>0.0592</strong></td>
<td><strong>0.131</strong></td>
<td><strong>0.0725</strong></td>
<td><strong>0.0218</strong></td>
</tr>
<tr>
<td>R-squared adjusted</td>
<td>0.0980</td>
<td>0.284</td>
<td>0.790</td>
<td>0.168</td>
<td>0.358</td>
<td>0.733</td>
</tr>
<tr>
<td>F test – p value</td>
<td>0.0010</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>RMS Error</td>
<td>0.0985</td>
<td>0.0878</td>
<td>0.0475</td>
<td>0.0977</td>
<td>0.0858</td>
<td>0.0553</td>
</tr>
</tbody>
</table>

Note: Derivation by author from Statistics Canada, Census 2006 using analytical files described in the text.
T refers to the threshold used in the estimation of a piece-wise linear least squares model. Results presented are those for models with highest adjusted R-squared among all possible thresholds from 5 to 12 years. Sample size is 198 observations by Age at Arrival from zero to 17 for each gender, representing 50,253 men and 51,631 women. All regressions use analytical weights. Standard errors are robust to heteroscedasticity, with shading indicating the marginal significance level: **less than or equal to 0.05**; greater than 0.05 and less than or equal to 0.10; **greater than 0.10**. Region of origin controls refer to indicator variables for the eleven regions categorized by linguistic distance from English and French as described in the text, with the omitted category for the sake of the regression analysis being the group of English speaking countries other than the UK or the US.
High school graduation and age at arrival:
For men a threshold at 8 and for women at 7 years.
An empirical model with the impact of age at arrival on labour market outcomes working through High School Graduation

\[ Y_i = \alpha + \beta \text{NoHighSchool}_i + \gamma X_i + \varepsilon_i \]  \hspace{0.5cm} (1)

\[ \text{NoHighSchool}_i = a + b Z_i + c X_i + e_i \]  \hspace{0.5cm} (2)

where \( Y \) is

- natural logarithm of Annual Earnings
- Receipt of Government Transfers

\( Z \) is an instrument motivated by the Critical Period Hypothesis

- Arriving after \( T \) years of age from a Non English speaking country, specifically: \( \max(0, \text{age at arrival}-T) \times 1\{\text{Non English}\} \)
Some tentative conclusions and next steps for analysis

1. There may be a “critical” period with respect to education outcomes
   - But it occurs before the onset of puberty reflecting other transitions children must go through as they progress from “learning to read” toward “reading to learn”
   - Need a closer look with the micro-data to correct for cohort effects
   - Related to language of the country of origin but not clear if this is a critical “learning” effect or a critical “social” effect – reflecting destination country education system
Some tentative conclusions and next steps for analysis

1. *There may be a “critical” period with respect to education outcomes*

2. *What are the causal relationships between outcomes related to labour market independence, age at arrival and education?*
   - Education and earnings – use the critical period as an “instrument” for education in order to estimate the impact of earnings, and the independent impact of age at arrival
   - Education and receipt of transfers
Some tentative conclusions and next steps for analysis

1. There may be a “critical” period with respect to education outcomes

2. What are the causal relationships between outcomes related to labour market independence, age at arrival and education?

3. Also need to examine other outcomes indicative of social and economic success.
Some tentative conclusions and next steps for analysis

1. *There may be a “critical” period with respect to education outcomes*

2. *What are the causal relationships between outcomes related to labour market independence, age at arrival and education?*

3. *Also need to examine other outcomes indicative of social and economic success.*

4. *Comparative studies with a number of other countries*
   - Canada, US, UK, Australia?
   - Similar source countries?
Age at immigration and the education outcomes of children

Miles Corak
Graduate School of Public and International Affairs

Université d’Ottawa | University of Ottawa
A draft outline of the upcoming paper

1. **Motivation:** the importance of studying the impact of migration on children
   - Rights of the child, offering voice
   - Relates directly to social inclusion and goals of public policy

2. **Why Canada?**

3. **Theory**
   - Narrow the question to the impact of age of arrival on adult outcomes
   - Draw implications from the Critical Period Hypothesis
   - Regression discontinuity as method for description and as an identification strategy

4. **Data:** 2006 Canadian Census

5. **Results pertaining to education**

6. **Results pertaining to other socio-economic outcomes**

7. **Conclusion**
Canada may be an interesting case to study because the children of immigrants have relatively successful outcomes

1. **OECD based comparisons from the PISA show success**
   - second generation Canadians as well as immigrant children perform relatively well in assessments at age 15 in international comparisons

2. **Little indication of the type of social unrest that has marked other countries**
   - The second generation are overall an advantaged and engaged segment of society, with higher education and just as good labour market outcomes
   - Toronto is a city in which close to the majority have not been born in Canada, second only to Sydney, though there are concerns with particular ethnic groups
   - Exception may be Quebec where there is a strong sense of immigrants needing to “accommodate” to the mainstream, rather than the mainstream being redefined
Labour market developments among male immigrants have deteriorated ...

Reasons for the deterioration

- About 1/3 due to the changing mix of language ability and country of origin
- Another 1/3 to a decline in the value of foreign labour market experience, particularly for men from non traditional countries
- The remainder is due to adverse labour market conditions at the time of entry
For those who are married they are more likely to be married to immigrants and particularly to other second generation immigrants.