Leveraging Facebook's Advertising Platform to Estimate Migration Stocks

Emilio Zagheni[1] **Ingmar Weber**[2]
Krishna Gummadi[3]

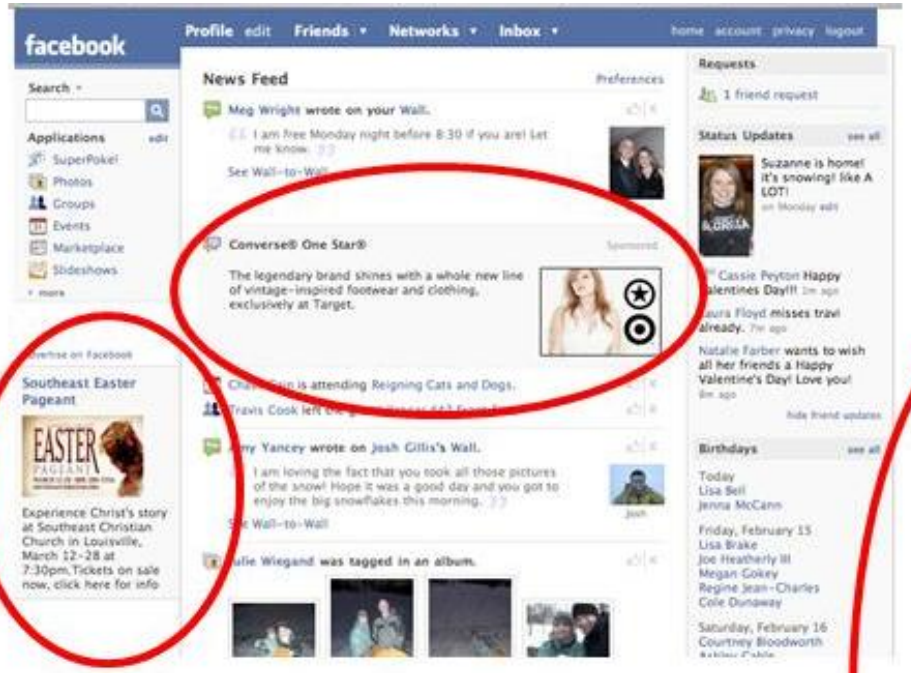
[1]University of Washington, Seattle

[2]Qatar Computing Research Institute

[3]Max Planck Institute for Software Systems

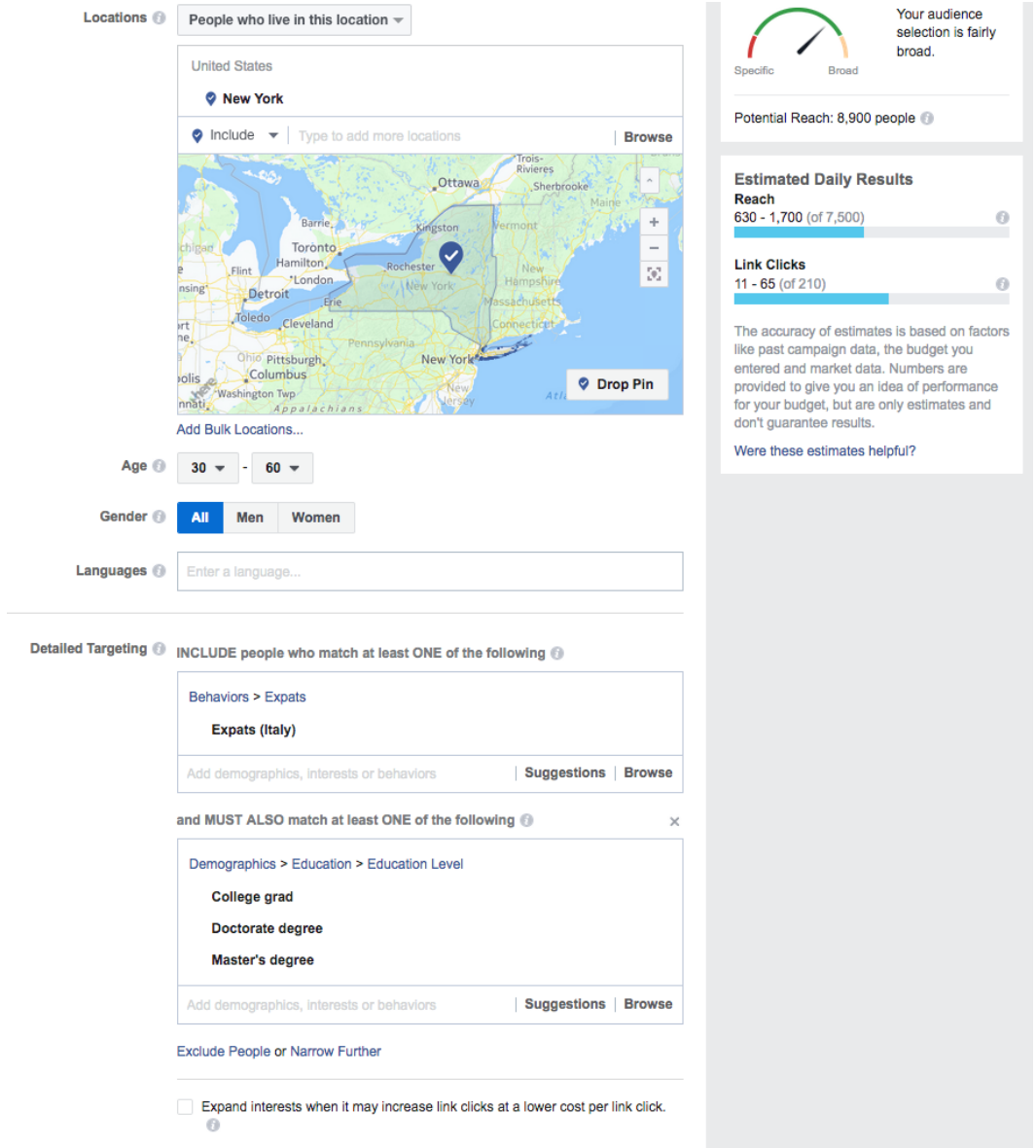
IUSSP International Population Conference 2017
Cape Town, South Africa

Online (targeted) advertising: A booming trend



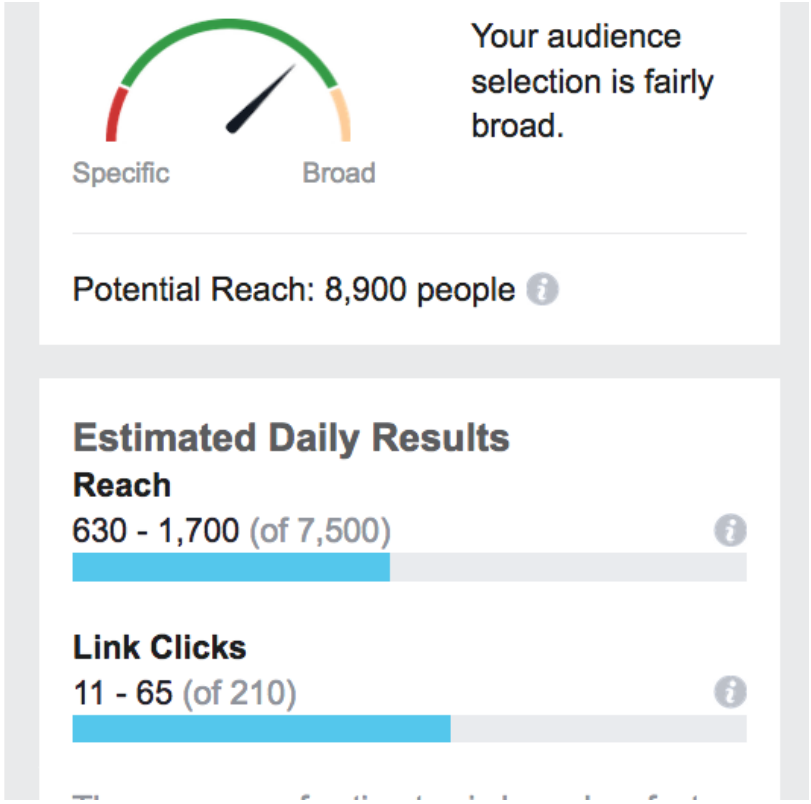
Reach the right people.
Instead of creating an advertisement and hoping that it reaches the right customers, you can create a Facebook Social Ad and target it precisely to the audience you choose. The ads can also be shown to users whose friends have recently engaged with your Facebook Page or engaged with your website through Facebook Beacon. Social Ads are more likely to influence users when they appear next to a story about a friend's interaction with your business.

Targeting a demographic group on Facebook



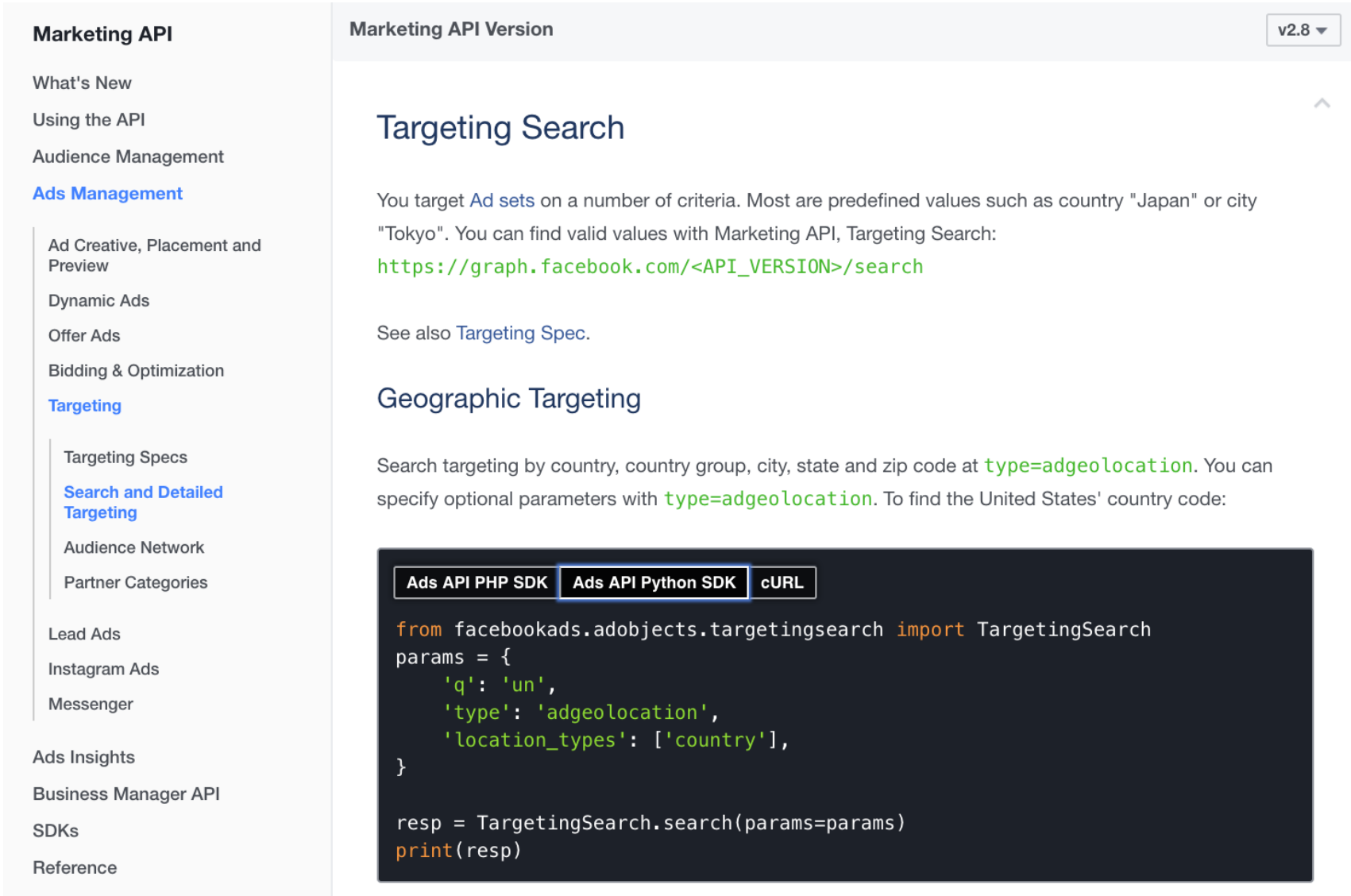
http://www.facebook.com/business

Targeting a demographic group on Facebook



http://www.facebook.com/business

We can access the data in a programmatic way



Marketing API

- What's New
- Using the API
- Audience Management
- Ads Management
  - Ad Creative, Placement and Preview
  - Dynamic Ads
  - Offer Ads
  - Bidding & Optimization
  - Targeting
    - Targeting Specs
    - Search and Detailed Targeting
    - Audience Network
    - Partner Categories
  - Lead Ads
  - Instagram Ads
  - Messenger
- Ads Insights
- Business Manager API
- SDKs
- Reference

Marketing API Version v2.8

Targeting Search

You target Ad sets on a number of criteria. Most are predefined values such as country "Japan" or city "Tokyo". You can find valid values with Marketing API, Targeting Search:
`https://graph.facebook.com/<API_VERSION>/search`

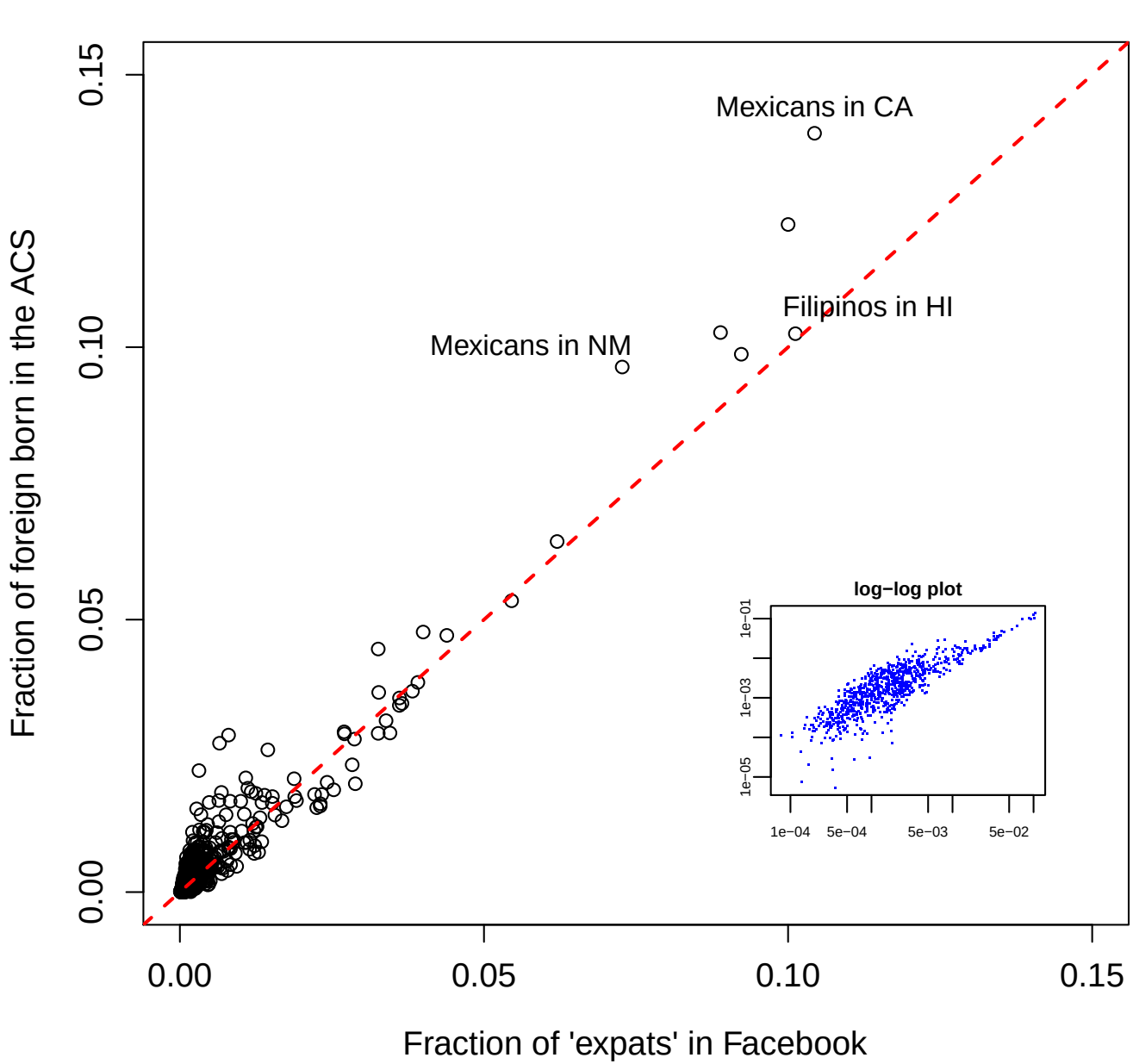
See also Targeting Spec.

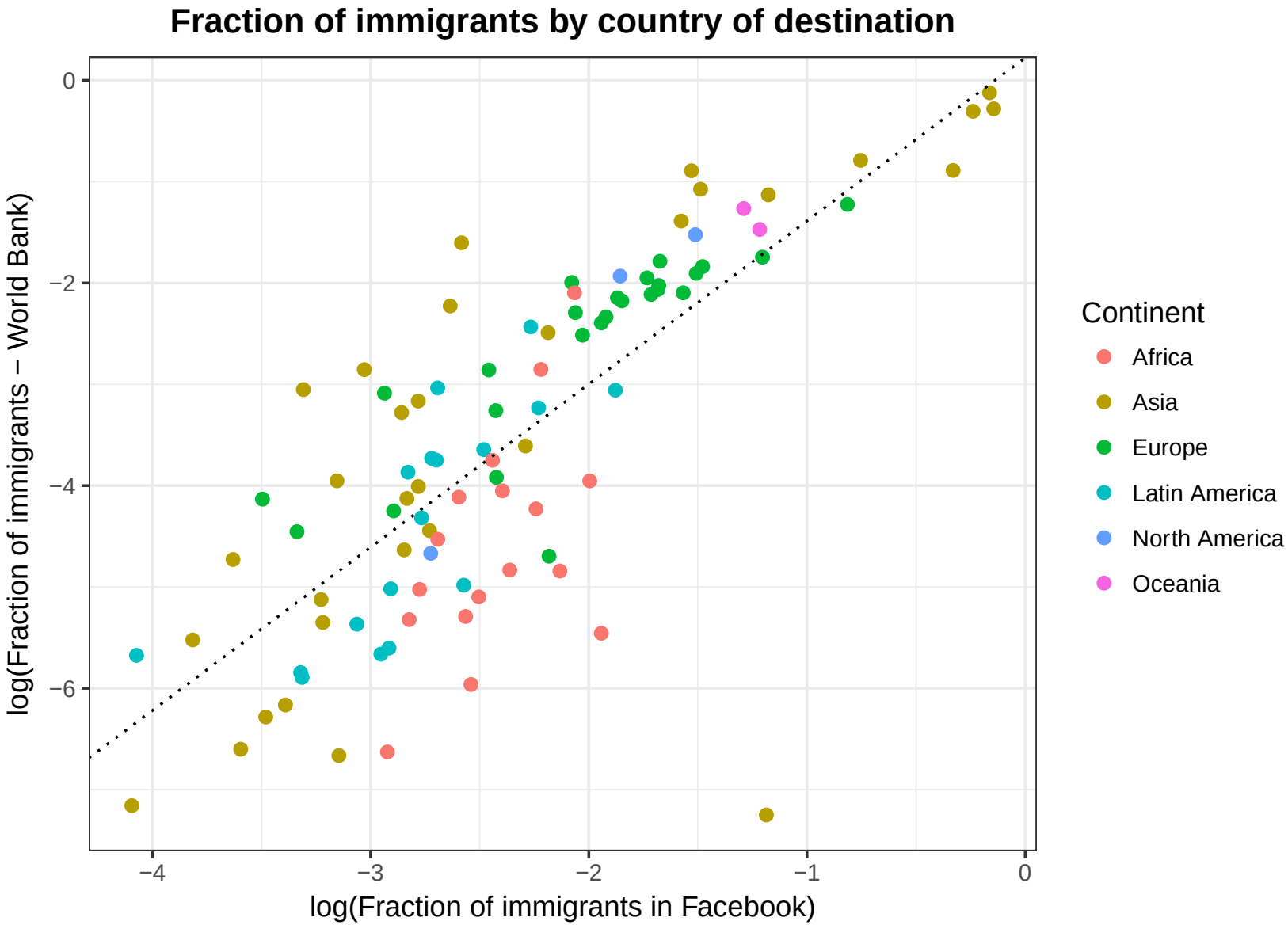
Geographic Targeting

Search targeting by country, country group, city, state and zip code at `type=adgeolocation`. You can specify optional parameters with `type=adgeolocation`. To find the United States' country code:

Ads API PHP SDK | Ads API Python SDK | cURL

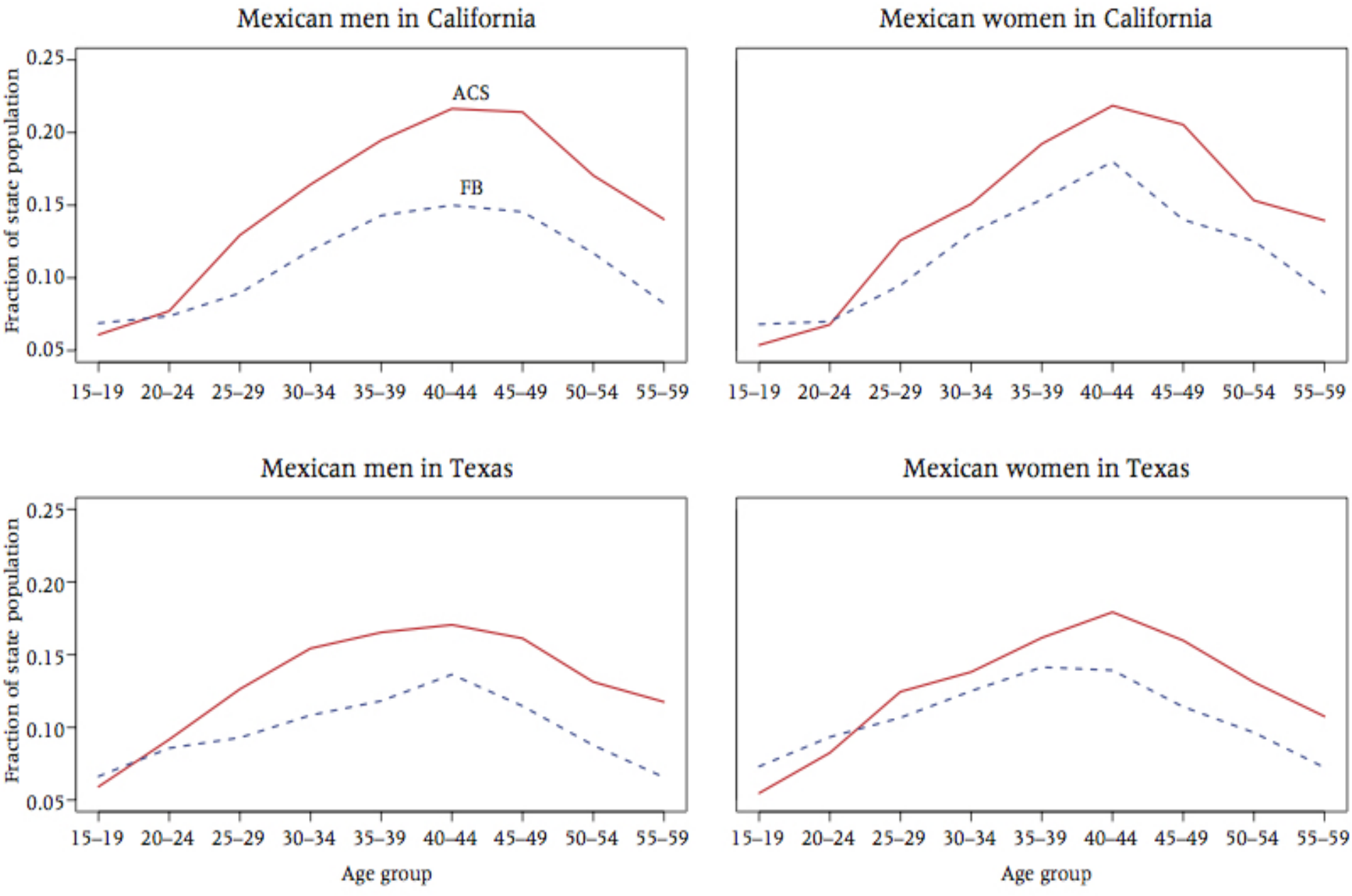
```
from facebookads.adobjects.targetingsearch import TargetingSearch
params = {
    'q': 'un',
    'type': 'adgeolocation',
    'location_types': ['country'],
}

resp = TargetingSearch.search(params=params)
print(resp)
```


Migrants to US states for different countries of origin
Mexicans in CA
Filipinos in HI
Mexicans in NM
log−log plot
Fraction of foreign born in the ACS
Fraction of 'expats' in Facebook
0.00
0.05
0.10
0.15
1e−01
1e−03
1e−05
1e−04
5e−04
5e−03
5e−02


Fraction of immigrants by country of destination
log(Fraction of immigrants – World Bank)
log(Fraction of immigrants in Facebook)
0
−2
−4
−6
−4
−3
−2
−1
0
Continent
Africa
Asia
Europe
Latin America
North America
Oceania

Profiles by age and sex

FIGURE 3 Facebook and ACS profiles of stocks of migrants by age and sex for Mexicans in California and in Texas



Evaluating patterns in the bias

$$
\begin{aligned}
log(\text{ACS foreign born pop}_{ij}^{z}) = \beta_0 &+ \beta_1 log(\text{Facebook expats}_{ij}^{z}) + \\
&+ \beta_2 \mathbb{1}(\text{Origin 1}) + \cdots + \beta_{30} \mathbb{1}(\text{Origin 29}) + \\
&+ \beta_{31} \mathbb{1}(\text{Age group 1}) + \cdots + \beta_{38} \mathbb{1}(\text{Age group 8}) + \\
&+ \epsilon_{ij}^{z}
\end{aligned}
$$

	log(Foreign-born Population - ACS)
log(FB Expats Population)	0.744*** (0.005)
Austria	0.420*** (0.083)
Canada	0.200*** (0.051)
China	1.132*** (0.050)
France	0.013 (0.056)
Germany	0.879*** (0.050)
Greece	1.443*** (0.064)
Hungary	0.264*** (0.078)
India	0.648*** (0.051)
Indonesia	−0.223*** (0.065)
Ireland	0.193** (0.064)
Israel	0.077 (0.063)
Italy	0.051 (0.057)
Japan	0.538*** (0.052)
Malaysia	0.159* (0.068)
Mexico	0.540*** (0.052)
Nepal	−0.018 (0.062)
Philippines	0.098 (0.051)
Poland	0.526*** (0.060)
Portugal	0.479*** (0.067)
Puerto Rico	0.136* (0.053)
Romania	0.174** (0.059)
Russia	1.069*** (0.052)
Singapore	0.367*** (0.075)
South Korea	0.811*** (0.051)
Spain	0.041 (0.060)
Turkey	0.044 (0.060)
UAE	0.376*** (0.099)
UK	−0.634*** (0.055)
Vietnam	0.301*** (0.052)
Age group (20-24)	−0.483*** (0.032)
Age group (25-29)	−0.291*** (0.032)
Age group (30-34)	−0.010 (0.031)
Age group (35-39)	0.094** (0.031)
Age group (40-44)	0.301*** (0.031)
Age group (45-49)	0.309*** (0.031)
Age group (50-54)	0.460*** (0.031)
Age group (55-59)	0.519*** (0.031)
Constant	1.374*** (0.052)
N	13,328
R^2	0.698
Adjusted R^2	0.697
Residual Std. Error	0.813 (df = 13289)
F Statistic	807.060*** (df = 38; 13289)

*p < .05; **p < .01; ***p < .001

Age pattern

Age group (20-24)	-0.483^{***} (0.032)
Age group (25-29)	-0.291^{***} (0.032)
Age group (30-34)	-0.010 (0.031)
Age group (35-39)	0.094^{**} (0.031)
Age group (40-44)	0.301^{***} (0.031)
Age group (45-49)	0.309^{***} (0.031)
Age group (50-54)	0.460^{***} (0.031)
Age group (55-59)	0.519^{***} (0.031)

Predictive capacity

- Goal: Predict the total number of foreign born from country i living in US state j (e.g, what is the stock of Mexicans in CA, Italians in NY?)

Predictive capacity

- Goal: Predict the total number of foreign born from country i living in US state j (e.g, what is the stock of Mexicans in CA, Italians in NY?)
- Split the data into training set (80% of US states) and test set (remaining 20% of US states)
- Estimate a model with no age & country of origin disaggregation and a model with disaggregation

Predictive capacity (continued)

The average out-of-sample Mean Absolute Percentage Error (MAPE) for total number of foreign born from country i living in state j:

- MAPE with no disaggregation by age-origin = 56%
- MAPE with disaggregation by age-origin = 37%

Predictive capacity (continued)

The average out-of-sample Mean Absolute Percentage Error (MAPE) for total number of foreign born from country i living in state j:

- MAPE with no disaggregation by age-origin = 56%
- MAPE with disaggregation by age-origin = 37%

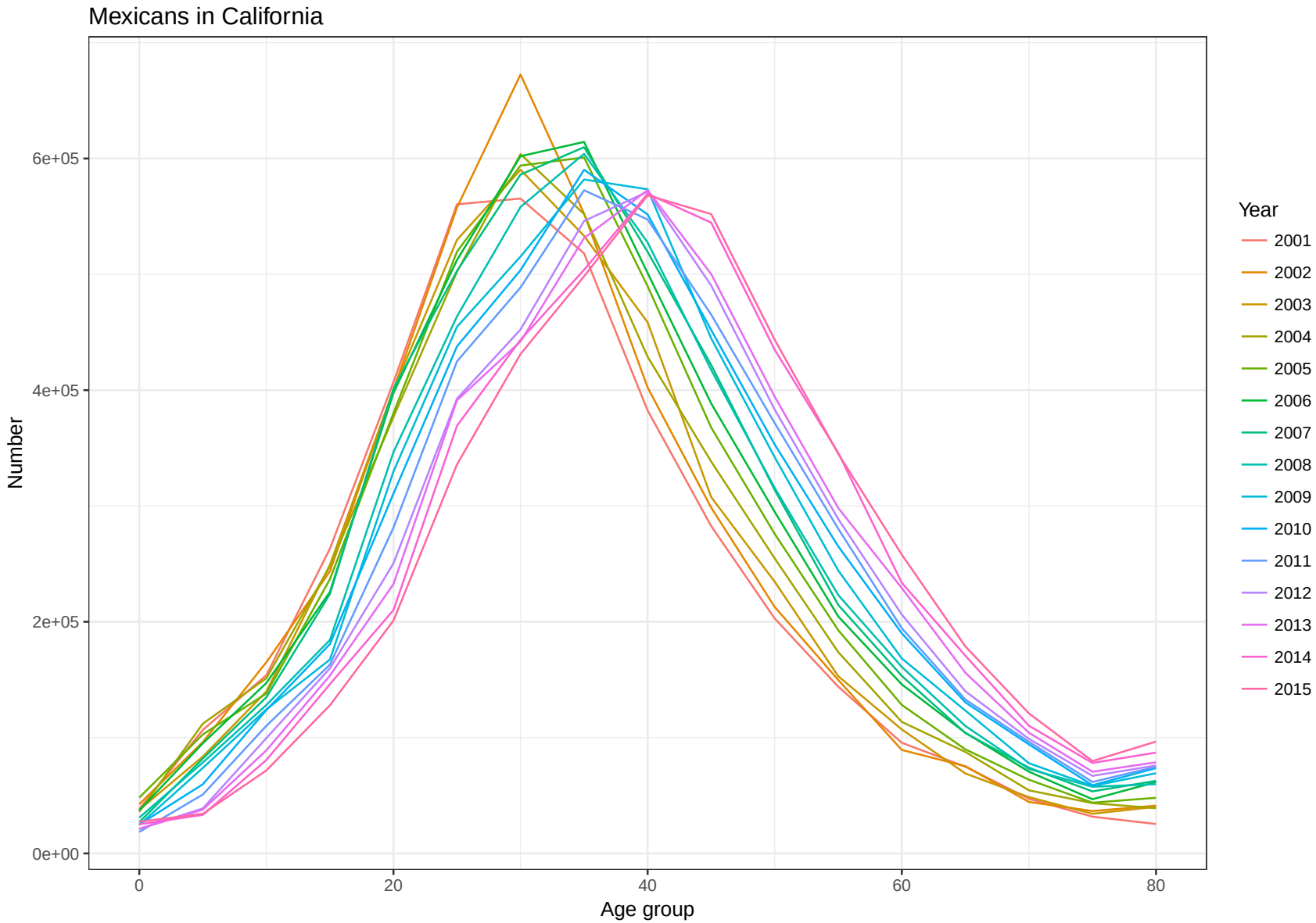
$\Rightarrow$ Indication that accounting for biases for subgroups of the population, in a way analogous to post-stratification, helps with predictions

Next steps

With K. Polimis, M. Alexander, I. Weber and F. Billari

- Combining Facebook data with time series analysis of ACS data
  - Facebook provides timely, but biased data, with no constraints on age patterns
  - A Lee-Carter type of model for time series of migration stocks adds constraints on demographic patterns

ACS time series



The road ahead: exciting + difficult = rewarding

The road ahead: exciting + difficult = rewarding

- Developing methods is key

The road ahead: exciting + difficult = rewarding

- Developing methods is key
- Eventually digital data will become just another type of data and best practices for how to use them will become common knowledge, similarly to how to use and interpret vital statistics and life tables

The road ahead: exciting + difficult = rewarding

- Developing methods is key
- Eventually digital data will become just another type of data and best practices for how to use them will become common knowledge, similarly to how to use and interpret vital statistics and life tables
- In the meanwhile, getting there will be a fun process...

Thank you for your attention