# Predicting Donations to an International Health Organization

Carolyn Cotterman
UC Berkeley, Biostatistics Group

Introduction

Watsi (www.watsi.org) posts photos and descriptions of people in poor countries who are in need of a medical treatment that they cannot afford. Each patient profile remains on the site until the patient's treatment is fully funded by donors accessing the site. Using Python's webpage parsing and machine-learning tools, I explore the determinants of donation rate. In particular, do patients who smile in their photo get funded more quickly?

Data Exploration

Who gets money fastest?

Who's smiling?

Time trends?
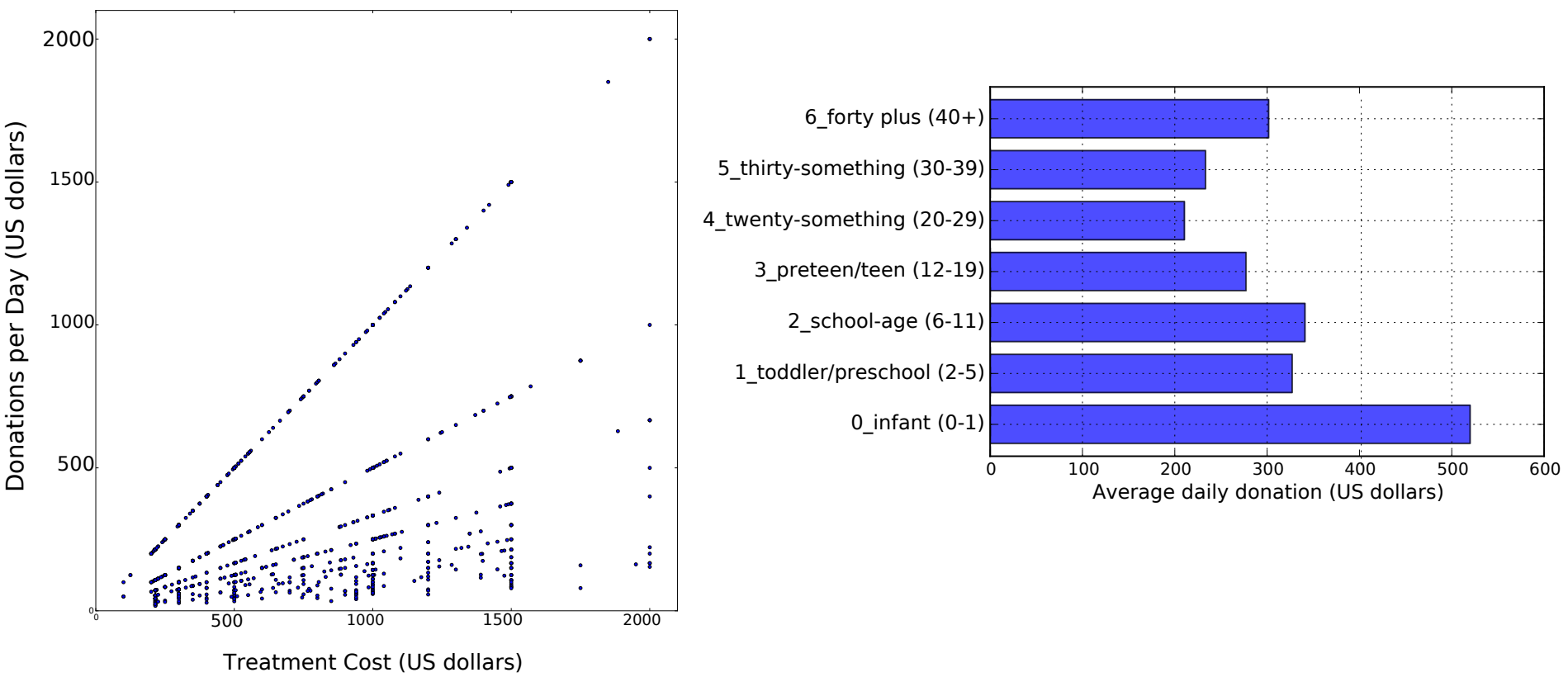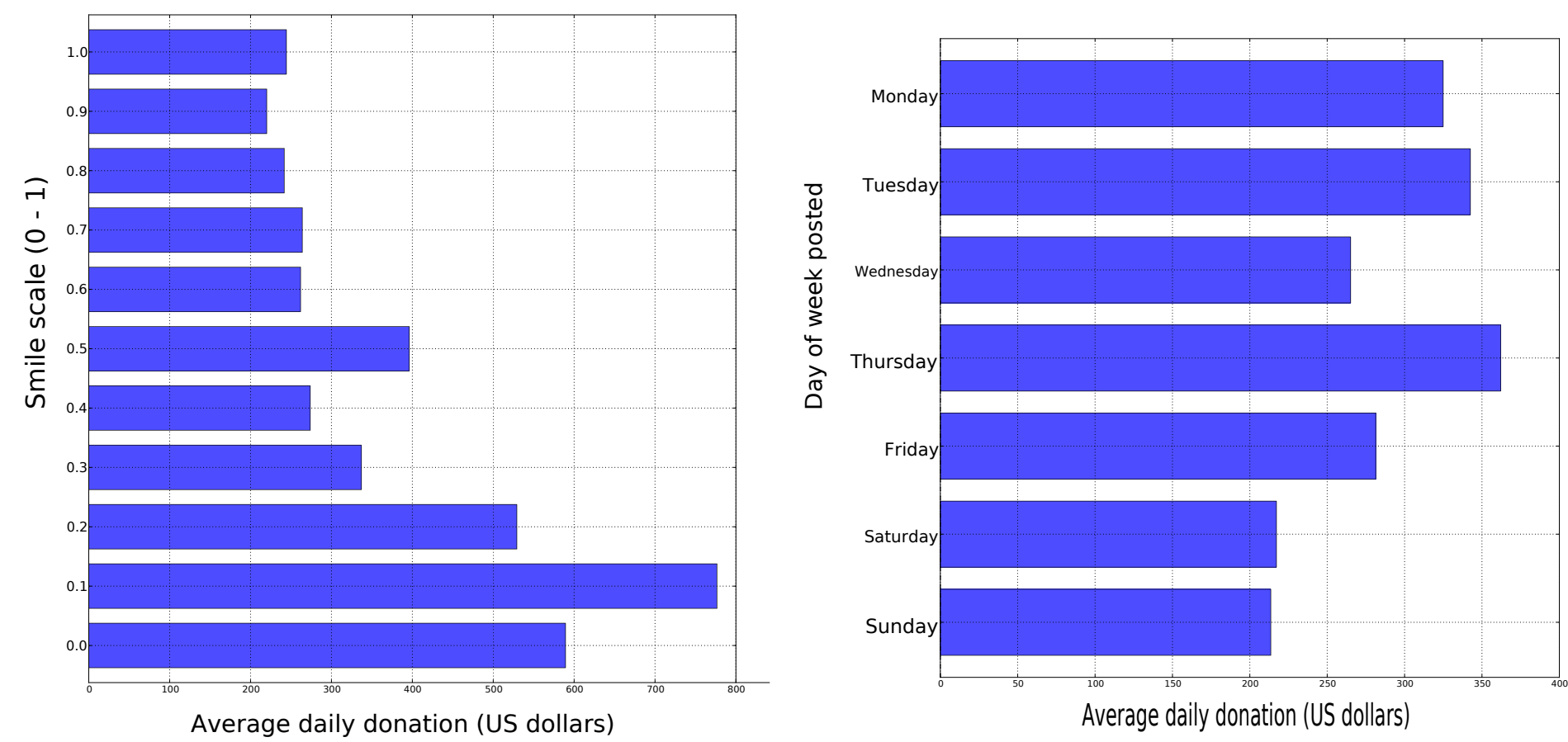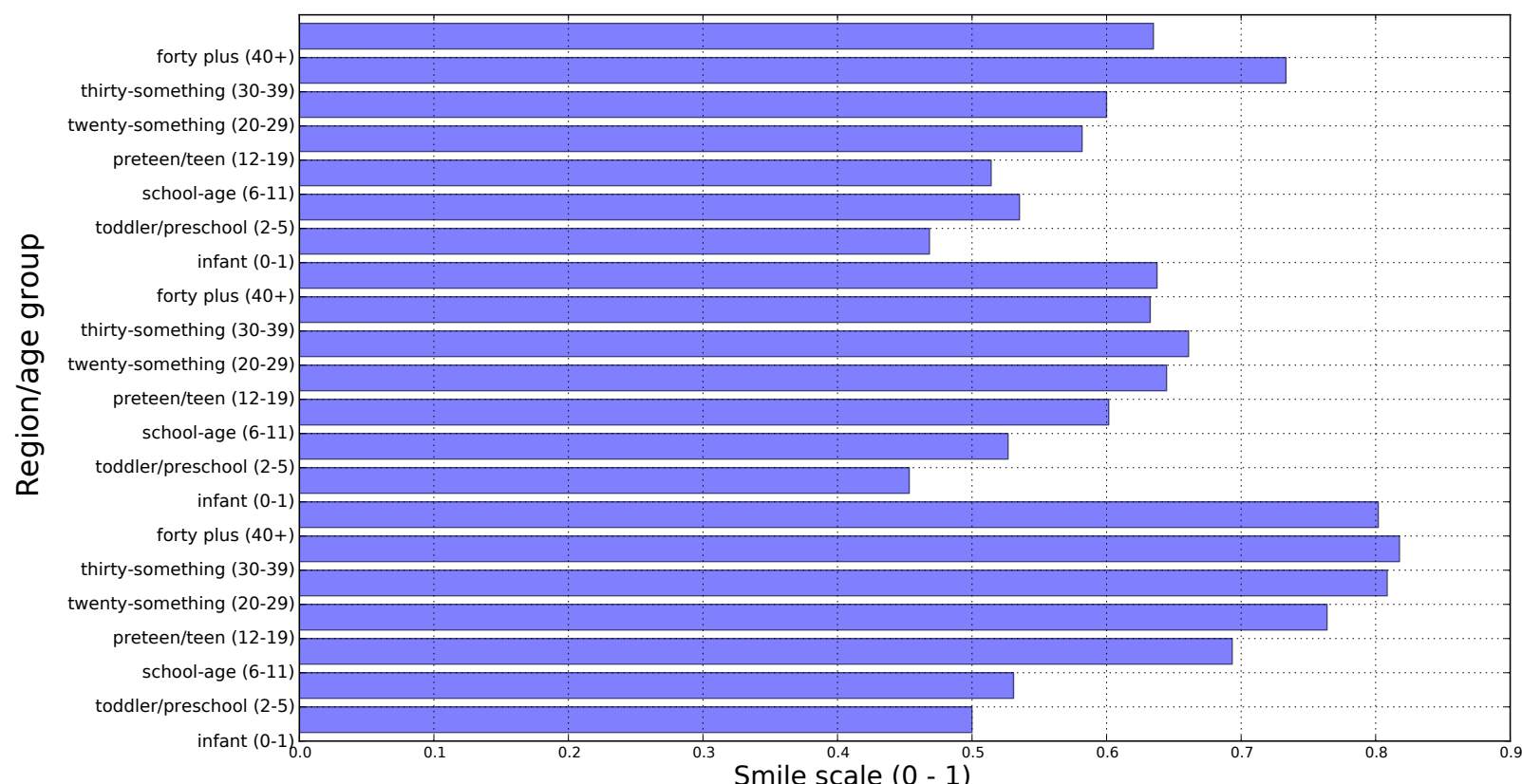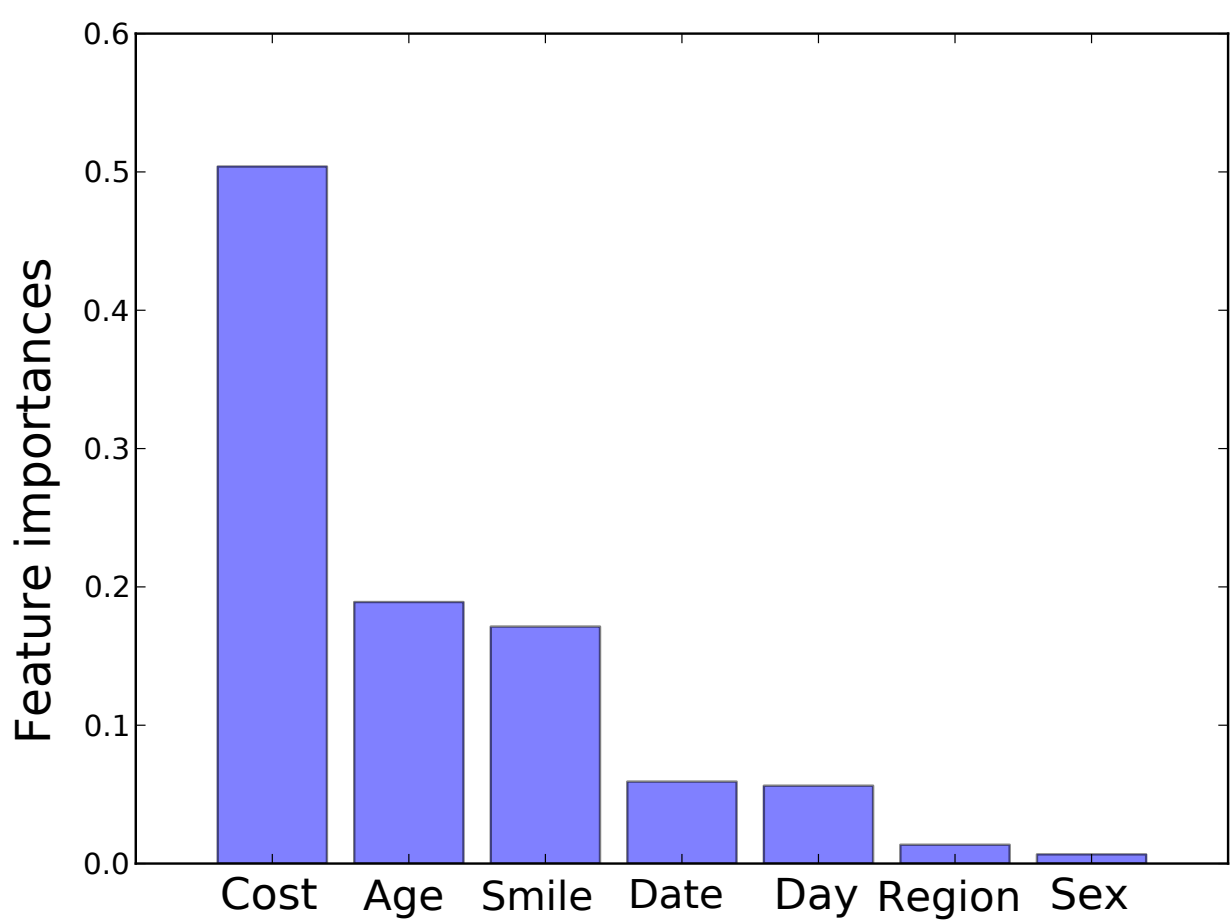

Measures of variable importance

\# Observational unit = patient
\# Outcome measure = mean dollars donated per day to patient
\# Predictors / controls:
    Total cost of treatment
    Patient age
    Patient gender (surmised)
    Region (Latin America vs. Asia vs. Africa)
    Facial expression in photo (0-1 scale, 0 is unhappy)
    Time trend
    Day-of-the-week

Random Forest

Linear Regression

```
                        OLS Regression Results
========================================================================
Dep. Variable:        dollars_per_day   R-squared:            0.254
Model:                            OLS   Adj. R-squared:       0.251
Method:                 Least Squares   F-statistic:          82.50
Date:             Thu, 12 Dec 2013      Prob (F-statistic):   2.71e-60
Time:                      03:38:02     Log-Likelihood:       -6824.2
No. Observations:               973     AIC:                  1.366e+04
Df Residuals:                   968     BIC:                  1.368e+04
Df Model:                         4
========================================================================
                coef    std err      t      P>|t|    [95.0% Conf. Int.]
------------------------------------------------------------------------
Intercept    164.0544    35.007    4.686    0.000    95.356   232.753
weekend_post -106.7809   24.661   -4.330    0.000   -155.175  -58.386
treat_cost     0.3176     0.020   16.021    0.000     0.279    0.356
patient_age   -1.1834     0.518   -2.285    0.023    -2.200   -0.167
smile_scale  -72.0287    41.767   -1.725    0.085  -153.992    9.935
========================================================================
Omnibus:              278.654   Durbin-Watson:           1.280
Prob(Omnibus):          0.000   Jarque-Bera (JB):      781.732
Skew:                   1.449   Prob(JB):             1.77e-170
Kurtosis:               6.299   Cond. No.              5.07e+03
========================================================================
```
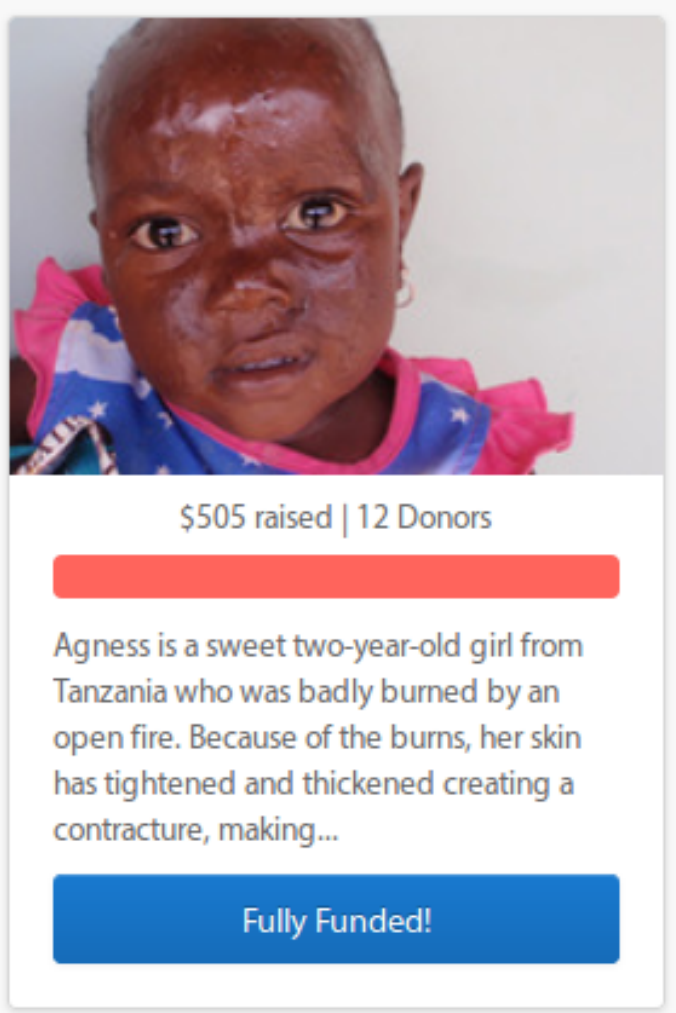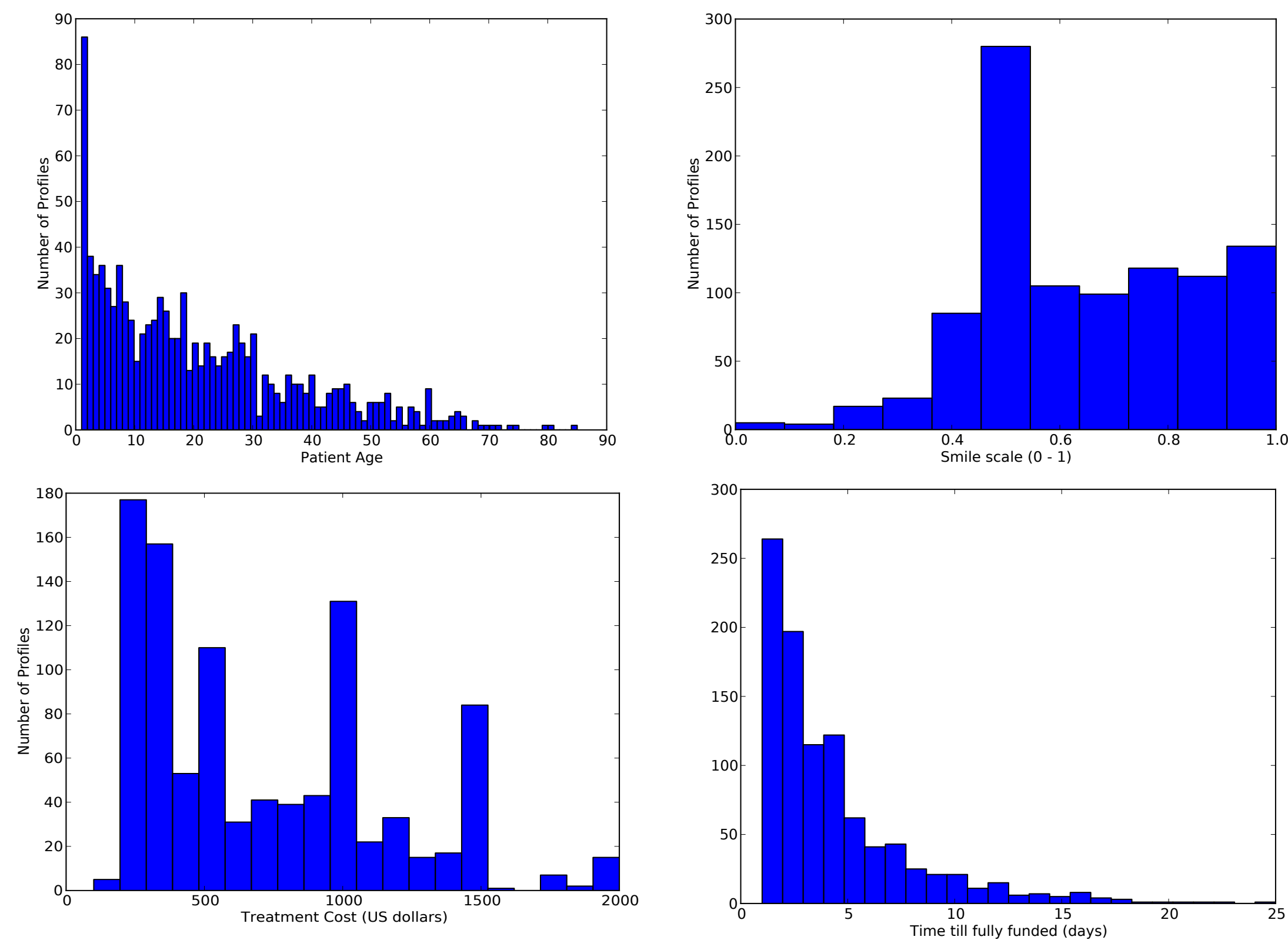
Inverse Probability of Treatment Weighted (IPTW)

Point estimate for average treatment effect of smiling: $34 per day
Non-parametric bootstrap indicates no statistical significance.

Extensions

\# Explore additional covariates (e.g., text processing)
\# Employ more robust statistcal procedures
\# Determine optimal patient mix for maximizing total donations
   (across patients for Watsi as a whole)