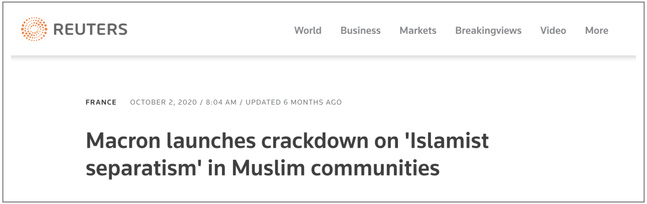
The Social Integration of International Migrants: Evidence from the Networks of Syrians in Germany

Michael Bailey (Facebook)
Drew Johnston (Harvard)
Martin Koenen (Harvard)
Theresa Kuchler (NYU Stern, CEPR, NBER)
Dominic Russel (Harvard)
Johannes Stroebel (NYU Stern, CEPR, NBER)

Background / Motivation

- Over 70 million people alive today have been forcibly displaced from their home country
- Big question: how can host countries help migrants integrate into their new communities?
  - Some prior work on job market integration
  - Little study of *social* integration


REUTERS World Business Markets Breakingviews Video More

FRANCE OCTOBER 2, 2020 / 8:04 AM / UPDATED 6 MONTHS AGO

Macron launches crackdown on 'Islamist separatism' in Muslim communities

Our Contribution

- We answer five questions:

Our Contribution

- We answer five questions:
    1. How can we measure migrants' social integration?

Our Contribution

- We answer five questions:
    1. How can we measure migrants' social integration?
    2. How much does integration vary across space?

Our Contribution

- We answer five questions:
  1. How can we measure migrants' social integration?
  2. How much does integration vary across space?
  3. What makes a place good at integrating migrants?

Our Contribution

- We answer five questions:
  1. How can we measure migrants' social integration?
  2. How much does integration vary across space?
  3. What makes a place good at integrating migrants?
  4. Can regional policies influence integration?

Our Contribution

- We answer five questions:
  1. How can we measure migrants' social integration?
  2. How much does integration vary across space?
  3. What makes a place good at integrating migrants?
  4. Can regional policies influence integration?
  5. Does exposure to migrants affect natives' attitudes?

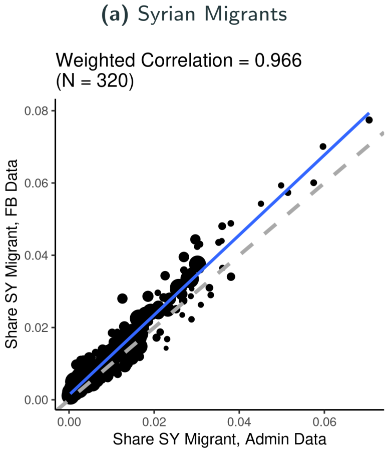
Our Approach

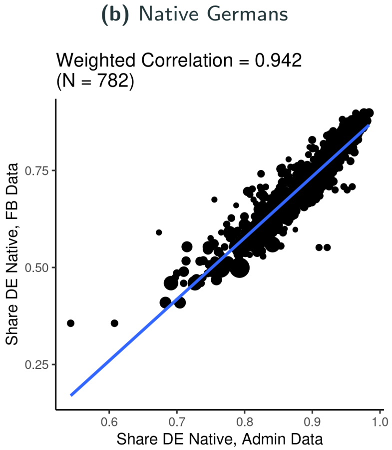
- We focus on Syrian refugees in Germany
  - $\approx$900k migrants, largest refugee population in Europe
  - Almost all arrived after Syrian Civil War
- Work with de-identified data from Facebook

Sample Construction (1)

- Active Facebook users aged 18+ in Germany
- We split into Syrians/Natives/Others using:
  1. Past and present location signals
  2. Self reported hometown/high school
  3. Language usage
- Sample size = 350k Syrians, 18m Germans

Sample Construction (2)





- Each dot is a gender x age bucket x region bucket

Sample Construction - Primary Integration Measures

1. **Friendships to nearby German natives**

Sample Construction - Primary Integration Measures

1. **Friendships to nearby German natives**

2. **German language usage**

Sample Construction - Primary Integration Measures

1. **Friendships to nearby German natives**
2. **German language usage**
3. **Local groups with native Germans**
   - Local soccer clubs, volunteering groups, etc

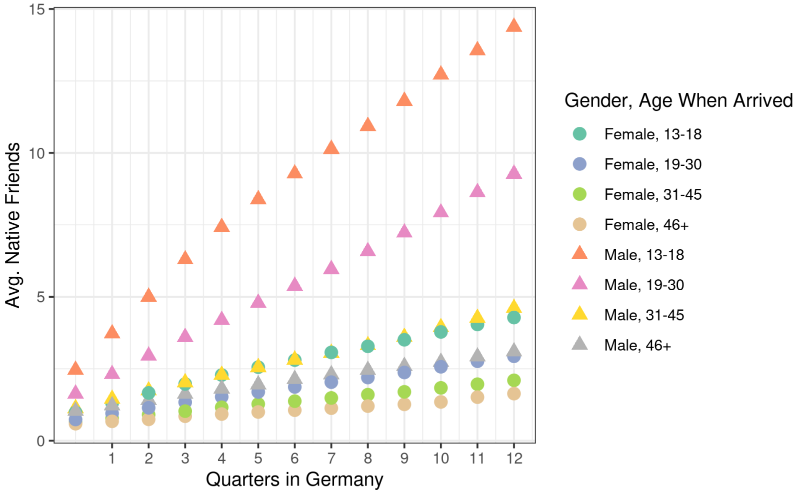
Sample Summary (1)

Panel (a): Syrian Migrant Sample

	Mean	SD	P10	P25	P50	P75	P90	P99
Age	32.90	10.26	22	25	31	38	48	66
Female (0/100)	32.07	46.68	0	0	0	100	100	100
DE College (0/100)	7.92	27.00	0	0	0	0	0	100
N Friends	347.89	385.84	62	117	226	423	751	2431
N Groups	104.55	137.09	8	22	56	129	256	831
Qs Since 1st on FB in DE	20.30	8.04	7	15	23	25	28	36
N Local Native Friends	5.03	12.24	0	0	1	4	13	87
N Local Syrian Friends	14.99	17.43	1	4	9	20	36	103
Produces DE Content (0/100)	30.40	46.00	0	0	0	100	100	100
N Local Native Groups	0.55	1.41	0	0	0	0	2	9

- Syrian migrant sample is (correctly) young, male
- Relatively low levels of integration
  - Highly correlated within individuals
  - Matches SOEP survey (regular contact with 6 Germans on average)

Sample Summary (2)

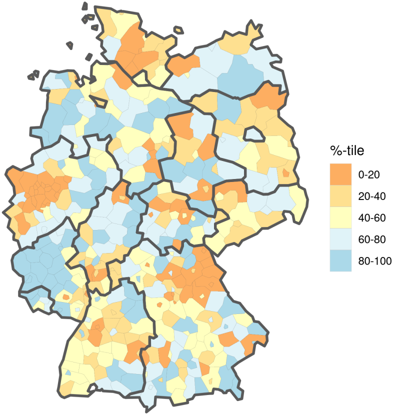


- Male, younger migrants better integrated

Measuring Local Integration

- Large data allow us to measure *county-level* integration
- Use average integration outcomes of SY migrants
  - Will focus on friendships to local German natives
- Ensure we capture real-world patterns by residualizing on (small) spatial differences in FB usage among natives
  - No differences in migrant usage

Regional Estimates of Integration - N Local Native Friends



- Top decile has 2x as many friends as bottom (3.9 vs 7.9)
- High reliability in split-sample
- Matches SOEP survey data on migrant friendships

SOEP correlation

What drives regional variation in integration outcomes?

- Three possible drivers:
  1. Differences in migrant observables
  2. Differences in migrant *un*-observables
  3. Effects of place
- Able to rule out possibility 1 directly
  - No large differences in observables across places
- We will separate possibilities 2 and 3 by looking at the (few) migrants who move between counties

Measuring the Effects of Place Using Movers (1)

- Consider a migrant who moves from a "low integration" place to a "high integration" place. If...

Measuring the Effects of Place Using Movers (1)

- Consider a migrant who moves from a "low integration" place to a "high integration" place. If...
  - ...place differences are from **migrant characteristics**, movers' behavior *WILL NOT* change to stayer levels

Measuring the Effects of Place Using Movers (1)

- Consider a migrant who moves from a "low integration" place to a "high integration" place. If...
  - ...place differences are from **migrant characteristics**, movers' behavior *WILL NOT* change to stayer levels
  - ...place differences are from **place effects**, movers' behavior *WILL* change to stayer level

Measuring the Effects of Place Using Movers (1)

- Consider a migrant who moves from a "low integration" place to a "high integration" place. If...
  - ...place differences are from **migrant characteristics**, movers' behavior *WILL NOT* change to stayer levels
  - ...place differences are from **place effects**, movers' behavior *WILL* change to stayer level

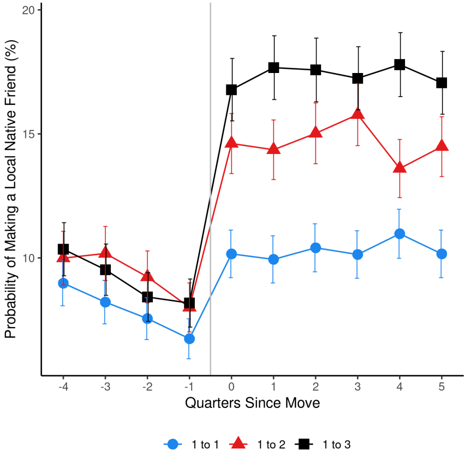
- Intuition follows number of recent movers papers [Card et al., 2013, Finkelstein et al., 2016, Finkelstein et al., 2019, Chetty and Hendren, 2018]

Measuring the Effects of Place Using Movers (2)

- *Sample*: Migrants who move to a non-neighboring county
- *Outcome*: Do they make a local native friend in a quarter?
- First: Group counties by integration outcome terciles
- Then: Study migrants moving from one tercile to another
  - $\rightarrow$ Can measure changes in integration around moves
  - $\rightarrow$ Do people who move to a "better" area integrate more?

Measuring the Effects of Place Using Movers (3)

Movers From Bottom Integration Tercile



- Suggestive evidence for place-based effects

Measuring the Effects of Place Using Movers (4)

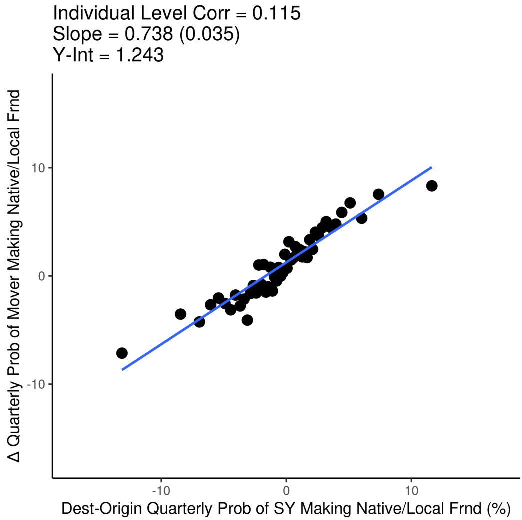
- Now: Model migrants' integration as sum of individual unobservables and place-based effects
  - These place-based effects can vary with observables
- When a migrant moves only place-based effect changes
- Movers then let us estimate the share of variation driven by place-based effects

Measuring the Effects of Place Using Movers (5)

$$y_{i,t}^{\Delta} = \alpha_0 + \alpha_1 x_{i,t}^{\Delta} + \xi_t + \epsilon_{i,t}$$

- $y_{i,t}^{\Delta}$ = change in friending after moving
  - $\rightarrow$ The change in a Syrian's probability of making a local native friend in each of the 4 quarters after vs before moving
- $x_{i,t}^{\Delta}$ = change in friending if user *i* adapted perfectly
  - $\rightarrow$ The difference in average friending between Syrians in the origin and destination who match the mover on demographics
- ξ_t = Quarter of move fixed effect
- α_1 identifies share of variation due to place effects

Measuring the Effects of Place Using Movers (6)



- Suggests 74% of variation is due to place-based effects

Measuring the Effects of Place Using Movers (7)

- **Takeaway: Local environments have strong effects on migrants' integration**

- Our estimates are probably a lower bound
  - We can't capture any place-based effects a person can bring with them (language, education, etc)

What makes environments better or worse for integration?

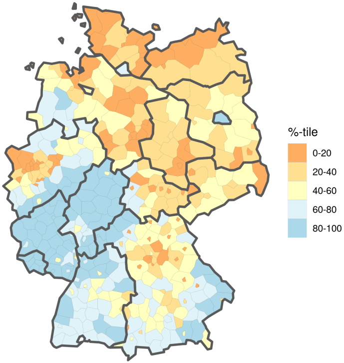
General and Relative Friendliness (1)

$$\underbrace{NLocalFriends_j^{SY \to DE}}_{Friending\ \ Integration} = \underbrace{NLocalFriends_j^{DE \to DE}}_{General\ \ Friendliness} \times \underbrace{\frac{\frac{NLocalFriends_j^{DE \to SY}}{NLocalFriends_j^{DE \to DE}}}{\frac{NSyr_j}{NGer_j}}}_{Relative\ \ Friendliness}$$

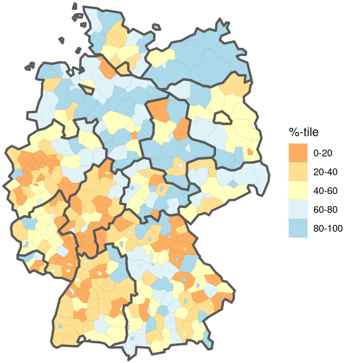
- **General Friendliness**: How many friends do natives have?
- **Relative Friendliness**: Do natives befriend Syrians in proportion to their local population share?
- Both components are strongly correlated with economic outcomes among Syrians
  - However, this distinction is important for policy

General and Relative Friendliness (2)

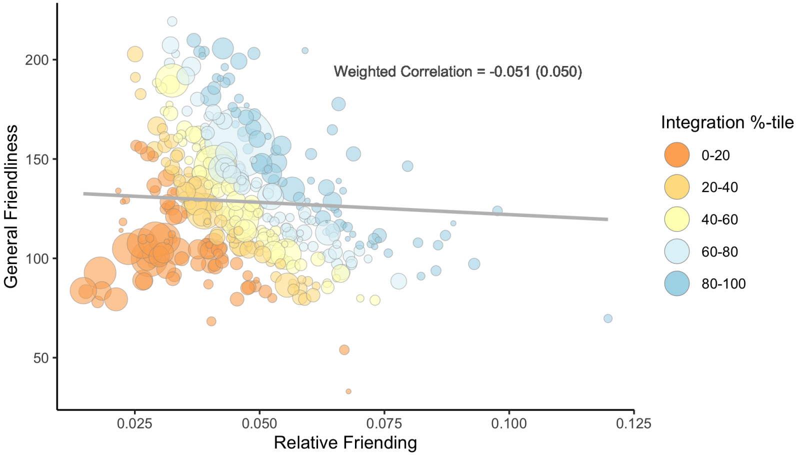
(a) General Friendliness



(b) Relative Friending



General and Relative Friendliness (3)



To What Extent are Native Behaviors Place-Based? (1)

- Why do general/relative friendliness differ across places?
  - Could be characteristics of natives (e.g. preferences)
  - Could be institutions
- Unlike the migrant case, there is no initial random assignment across places
- Size of place-based effects is important for policy

Correlates of Native Behavior

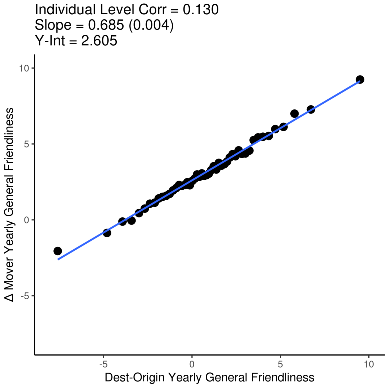
To What Extent are Native Behaviors Place-Based? (2)

$$y_{i,t}^{\Delta} = \alpha_0 + \alpha_1 x_{i,t}^{\Delta} + \xi_t + \epsilon_{i,t}$$

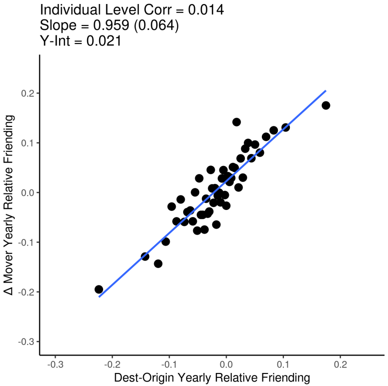
- This design is similar to the design we used for migrants
- $y_{i,t}^{\Delta}$ = change in the native's behavior post-move
  - $\rightarrow$ The change in the native's level of general/relative friending in the year following their move, relative to the year before
- $x_{i,t}^{\Delta}$ = change in behavior if the native adapted perfectly
  - $\rightarrow$ Difference in general/relative friending between observably identical natives in the destination and origin
- α_1 identifies share of variation due to place effects

To What Extent are Native Behaviors Place-Based? (3)

(a) General Friendliness



(b) Relative Friending



- Large slopes → large role for place-based factors

Why Do Places Differ? (1)

- Something about places seems to determine integration.
- But what is it?
  - Civic programs?
  - Geography?
  - Government policies?

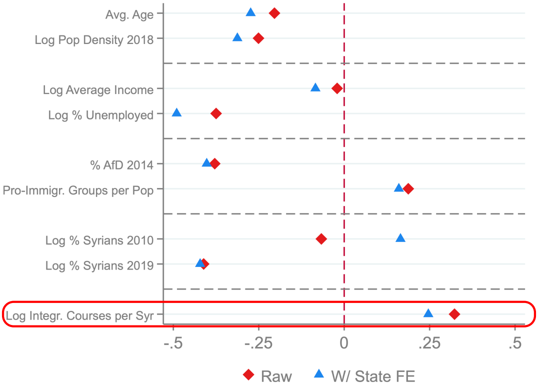
Why Do Places Differ? (2)

County-Level Univariate Correlations with Friending Integration

- Correlations give some sense, but are they are causal?
- Do these affect general friendliness or relative friendliness?

Why Do Places Differ? (2)

County-Level Univariate Correlations with Friending Integration



- Correlations give some sense, but are they are causal?
- Do these affect general friendliness or relative friendliness?

Causal Effects of Regional Policy (1)

- Integration courses are the most common policy
- Can teach language skills, culture, civics
- Teachers in the courses need experience teaching German as a second language
  - Relatively few have these qualifications
- Areas varied in the availability of potential teachers in 2015-2016

  → Instrument for course availability using unemployed specialized teachers

Causal Effects of Regional Policy (2)

Integration Courses and Teacher Unemployment Rates

	Log Integration Courses per Syrian 2015-19			
Log Unemp. General Schools Teachers 2014 per Syrian	0.088 (0.05)			
Log Unemp. Vocat. School Teachers 2014 per Syrian		0.084 (0.05)		
Log Unemp. Driving and Sports Teachers 2014 per Syrian			0.052 (0.06)	
Log Unemp. Other School Teachers 2014 per Syrian				0.229*** (0.05)
Control Covariates	x	x	x	x
Control Log General Unemployment Rate	x	x	x	x
F-statistic	2.37	3.67	0.94	20.97
N	390	367	388	390
R-Squared	0.349	0.354	0.347	0.379

- Unemployment among German as a second language teachers predicts integration course completion
- Strong F-stat given county-level regression

Causal Effects of Regional Policy (3)

Table 1: IV Estimates - Measures of Integration and Integration Courses

	Integration	General Friendliness	Relative Friending	Language	Employ. / Training
Log Integration Courses per Syrian	1.698***	0.204	1.389***	0.193***	0.891***
	(0.33)	(0.21)	(0.25)	(0.07)	(0.15)
Control Covariates	x	x	x	x	x
Control Log General Unemployment Rate	x	x	x	x	x
N	390	390	390	390	384

- Integration courses tend to improve language acquisition, employment outcomes, and relative friending
- No large impact on general friendliness
- We tend to find IV estimates > OLS estimates
    - Marginal courses tend to be in low-integration areas
    - Women more likely to forgo courses if supply limited

Individual-level Determinants of Friending

- We've seen that:
    - Migrants in different regions vary in their integration
    - Migrant characteristics do not explain these differences
    - Characteristics of place matter more than characteristics of its residents
- But what explains heterogeneity *within* a place?
    - Previous contact may shape attitudes for natives
    - ...But not clear how wide-reaching these effects are.

High School Cohort Analysis (1)

- School entry cutoffs cause quasi-random variation in contact
- Students born before or after the cutoff are placed into cohorts with different demographics
- Around the cutoff there is quasi-random variation in an individual's social network
- We consider neighboring cohorts in a school, where one year has a Syrian and one does not

High School Cohort Analysis (2)

$$Y_i = \alpha_1 SyrianInCohort_s + \xi_{t,L} + \gamma_s + \epsilon_{i,t}.$$

- Here, Y_i is a social outcome, $SyrianInCohort_s$ is an indicator if an individual has a Syrian in their class, $\xi_{t,L}$ is a county-by-year FE, and γ_s is a school FE.

High School Cohort Analysis (3)

	Syrian Friends		Syrian Friends (Excluding Classmates)		Syrian Friends (Excluding Syrian Classmates and their Friends)	
Syrian in Cohort	0.020***	0.020***	0.005***	0.005***	0.005***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Syrian in Cohort x Standardized Cohort Size		-0.007***		-0.003***		-0.003***
		(0.001)		(0.001)		(0.001)
School FE	X	X	X	X	X	X
Birth Year x County FE	X	X	X	X	X	X
N	115,625	115,625	115,625	115,625	115,625	115,625
Mean in Control Cohort	0.054	0.054	0.029	0.029	0.027	0.027

- Germans exposed to a Syrian make more Syrian friends
    - Even friends in totally different settings
- Points to a shift in attitudes
- Effect is larger in smaller cohorts

Summary

Summary

1. We present the most systematic measurements of the social integration of Syrians in Germany

Summary

1. We present the most systematic measurements of the social integration of Syrians in Germany
2. Integration levels are generally quite low, but with a lot of variation across people and places

Summary

1. We present the most systematic measurements of the social integration of Syrians in Germany
2. Integration levels are generally quite low, but with a lot of variation across people and places
3. These differences seem to be largely driven by the characteristics of the place, not the people living there

Summary

1. We present the most systematic measurements of the social integration of Syrians in Germany
2. Integration levels are generally quite low, but with a lot of variation across people and places
3. These differences seem to be largely driven by the characteristics of the place, not the people living there
4. These local institutions can also be changed with policy, such as integration courses

Summary

1. We present the most systematic measurements of the social integration of Syrians in Germany
2. Integration levels are generally quite low, but with a lot of variation across people and places
3. These differences seem to be largely driven by the characteristics of the place, not the people living there
4. These local institutions can also be changed with policy, such as integration courses
5. Contact between Syrians and native Germans can improve integration in the long run

References

Card, D., Heining, J., and Kline, P. (2013).
Workplace heterogeneity and the rise of west german wage inequality.
The Quarterly journal of economics, 128(3):967–1015.

Chetty, R. and Hendren, N. (2018).
The impacts of neighborhoods on intergenerational mobility ii: County-level estimates.
The Quarterly Journal of Economics, 133(3):1163–1228.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016).
Sources of geographic variation in health care: Evidence from patient migration.
The Quarterly Journal of Economics, 131(4):1681–1726.

Finkelstein, A., Gentzkow, M., and Williams, H. L. (2019).
Place-based drivers of mortality: Evidence from migration.
Technical report, National Bureau of Economic Research.

Appendix

Appendix - Multivariate Sample Summary

	Facebook Sample N Local Native Friends			
Age 25 - 34	-1.012***	-0.894***	-0.873***	-1.148***
	(0.053)	(0.052)	(0.052)	(0.129)
Age 35 - 44	-2.963***	-3.019***	-2.941***	-2.375***
	(0.062)	(0.061)	(0.061)	(0.158)
Age 45 - 54	-4.012***	-4.102***	-4.147***	-4.765***
	(0.080)	(0.079)	(0.079)	(0.184)
Age 55+	-4.548***	-4.531***	-4.586***	-7.226***
	(0.100)	(0.098)	(0.099)	(0.241)
Female	-3.676***	-3.610***	-3.225***	-3.267***
	(0.043)	(0.042)	(0.045)	(0.090)
Household Member in DE 1+ Year Prior	-0.377***	-0.290**	-0.352***	
	(0.100)	(0.099)	(0.099)	
Non-Household Family in DE 1+ Year Prior	0.524***	0.621***	0.421***	
	(0.091)	(0.089)	(0.089)	
Quarters Since DE FEs	X	X	X	X
Prev Quarters in NUTS3 FEs	X	X	X	X
Personal Usage Controls	X	X	X	X
County / State FEs		X	X	X
Log (1 + Total Outside Germany Friends)			X	X
Log (1 + Total Other Groups)			X	X
Log (1 + Total Content Produced Past Year)			X	X
Household FE				X
N	349,072	349,072	349,072	84,216
R-Squared	0.132	0.160	0.165	0.658
Sample Mean	5.029	5.029	5.029	4.195

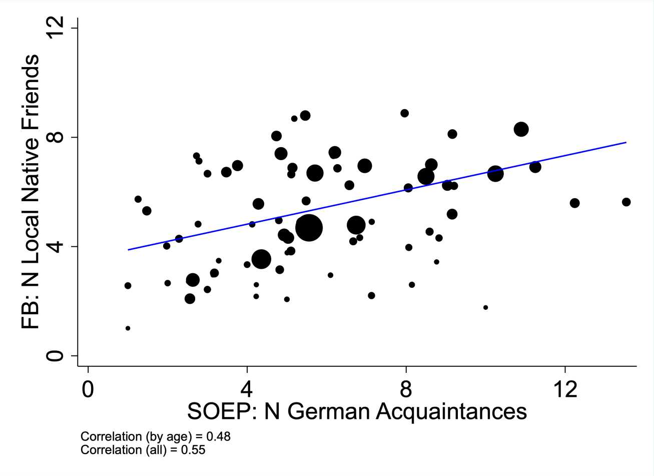
Return

Appendix - Multivariate Sample Summary

	N Local SY Friends		General Friendliness		Relative Friending		In Pro Imm. Group (0/100)	
Age 25 - 34	-0.073***	-0.073***	-19.097***	-14.407***	-0.059***	-0.061***	0.359***	0.146***
	(0.000)	(0.000)	(0.098)	(0.092)	(0.001)	(0.001)	(0.018)	(0.018)
Age 35 - 44	-0.116***	-0.114***	-55.586***	-52.328***	-0.081***	-0.080***	0.951***	0.858***
	(0.000)	(0.000)	(0.103)	(0.097)	(0.001)	(0.001)	(0.018)	(0.018)
Age 45 - 54	-0.132***	-0.131***	-62.533***	-62.415***	-0.098***	-0.095***	1.116***	1.152***
	(0.000)	(0.000)	(0.108)	(0.102)	(0.001)	(0.001)	(0.019)	(0.019)
Age 55+	-0.139***	-0.141***	-82.666***	-84.728***	-0.098***	-0.095***	2.105***	2.157***
	(0.000)	(0.000)	(0.108)	(0.102)	(0.001)	(0.001)	(0.020)	(0.020)
Female	-0.015***	-0.015***	-19.519***	-18.725***	-0.008***	-0.009***	0.882***	0.843***
	(0.000)	(0.000)	(0.056)	(0.053)	(0.001)	(0.001)	(0.010)	(0.010)
Has College	0.006***	0.006***	4.131***	7.619***	-0.000	-0.002***	1.931***	1.788***
	(0.000)	(0.000)	(0.060)	(0.056)	(0.001)	(0.001)	(0.011)	(0.011)
Prev Quarters in NUTS3 FEs	X	X	X	X	X	X	X	X
Personal Usage Controls	X	X	X	X	X	X	X	X
County FEs		X		X		X		X
N	17,768,822	17,768,822	17,768,822	17,768,822	17,515,164	17,515,164	17,768,141	17,768,141
R-Squared	0.020	0.031	0.170	0.263	0.001	0.002	0.035	0.042
Sample Mean	0.086	0.086	122.510	122.510	0.074	0.074	4.835	4.835

Return

Regional Estimates of Integration - SOEP Validation



- Our estimates match state-level survey results

Return