Presence of language-learning opportunities abroad and migration to Germany

Matthias Huber Silke Uebelmesser

University of Jena, Germany

International Forum on Migration Statistics
OECD, Paris, January 2018

Funded by the German Science Foundation (DFG, grant number UE 124/2-1)
– see also www.fiwi.uni-jena.de/dfg –

Motivation

- **Benefits of migrants' proficiency in the language of the destination country**
  - Labour market outcomes: earnings (Dustmann and Soest 2001; Chiswick and Miller 1995) and employment probability (Dustmann and Fabbri 2003) increase
  - Social integration: probability of intermarriage increases and the likelihood of living in an ethnic enclave decreases (Bleakley and Chin 2010)
- **Language proficiency as determinant of migration flows**
  - But linguistic distance has been found to have a negative impact due to higher costs of language acquisition (Adsera and Pytlikova 2015; Belot and Hatton 2012)

Language learning and migration decisions

- Focusing on linguistic distance neglects actual language acquisition – despite higher costs.
- **Children-age language learning**
  - Decision determined by factors outside the learner's direct control: language skills $\rightarrow$ migration decision
  - Compulsory foreign language learning positively related to migration flows within the EU (Fenoll and Kuehn 2016)
- **Adult-age language learning**
  - Decision determined by learner's migration (intention): migration decision $\rightarrow$ language learning
  - Migration as determinant of adult-age language learning (Uebelmesser and Weingarten 2017)

Research Question: What is the effect of the presence of language learning opportunities for adults abroad on migration to Germany?

Procedure:

- Basic panel regressions
- Robustness checks (especially to address reverse causality concerns)

Results "in a nutshell":

- Language learning opportunities (here: **Goethe Institutes**) are positively correlated with migration to Germany.
- There is evidence that this is a causal effect from language learning opportunities to migration.

Data

Goethe-Institut (GI)

- Main actor in promoting German culture and language worldwide.
- Institutes worldwide offer
  - Language services: courses and standardized exams.
  - Information on the German culture and society: (cultural) events and libraries.
- Funded by the German government; language services by fees.
- Annual reports of the GI provide information about institutes.

▸ Map

Data

From its annual reports, we constructed three datasets.

1. Dataset about the regional distribution of the GI from 1965-2014, the opening and closing years of all institutes and whether they provide language services.
   $\rightarrow$ 2014: 137 institutes in 86 countries (plus 12 in Germany)
2. Dataset about language learning in GI placed in countries all over the world. We report numbers on course registrations (1990-2014), sold course units (1972-1989 and 1997-2014) and exam participation (1986-2014).
   $\rightarrow$ 2014: 229,702 registrations and 17,113,040 sold course units, 287,630 exams.
3. Dataset about information on language course participation at GI in Germany (1966-2015).
   $\rightarrow$ 2014: 13,459 European registrations and 20,397 non-European registrations from about 200 countries.

$\Rightarrow$ In 2014, almost 1.5m people migrated to Germany.

Data

For this study here, we use "Dataset 1" about the number of institutes (including openings and closings).

- Our sample is a balanced panel of 77 countries from 1968 – 2014.
- In 2014,
  - 51 countries had at least one GI with the number of institutes in these countries amounting to 86.
  - 152,600 registrations took place.
  - Almost 550,000 migrants from these countries with a GI came to Germany (and 621,000 from all countries in our sample).

Data

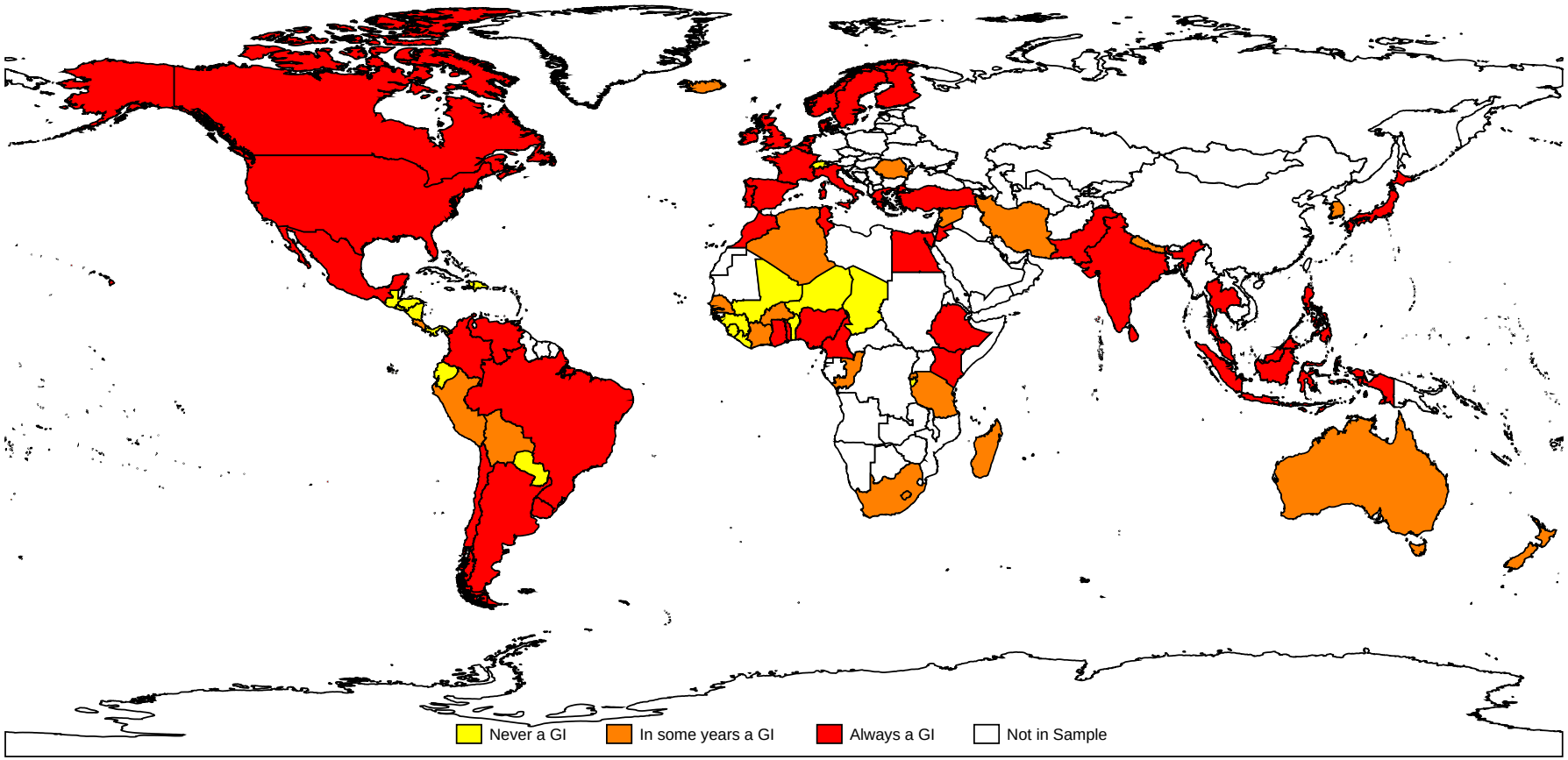

Figure 1: The presence of GI (our sample)

Data

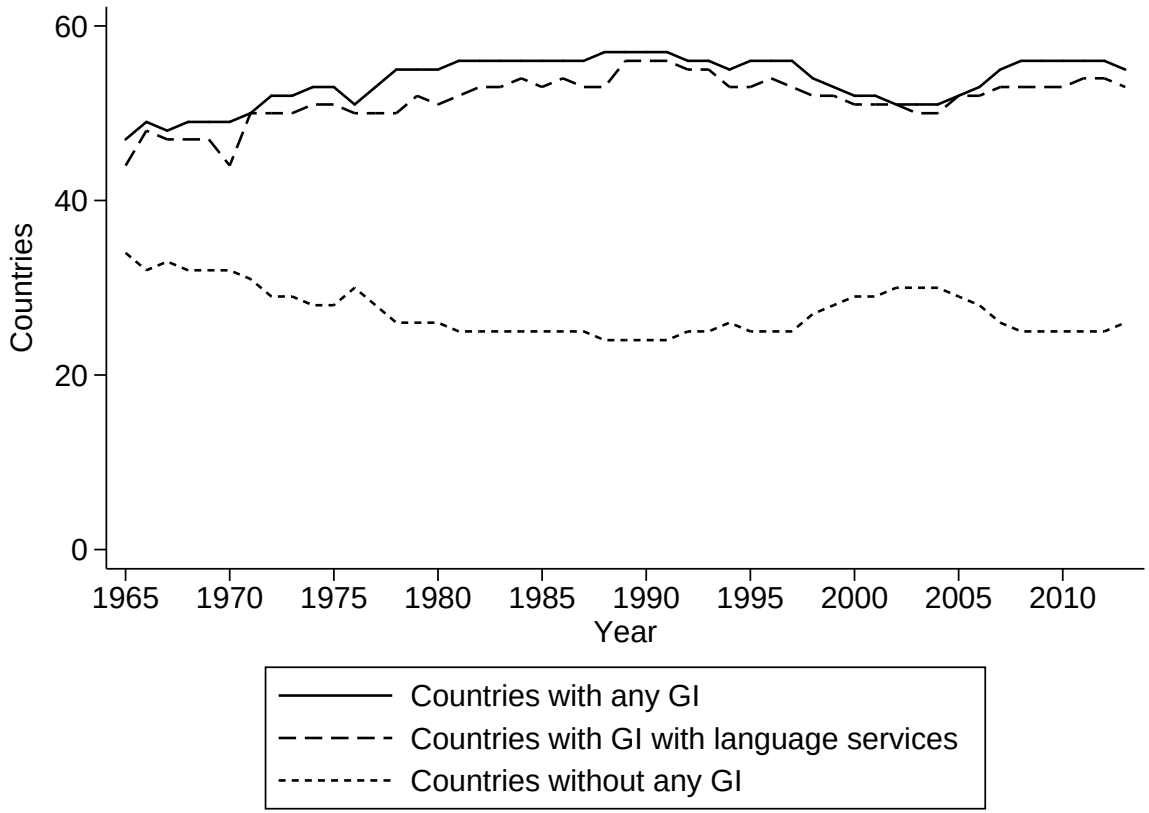

Figure 2: Number of countries with GI (our sample)

Data

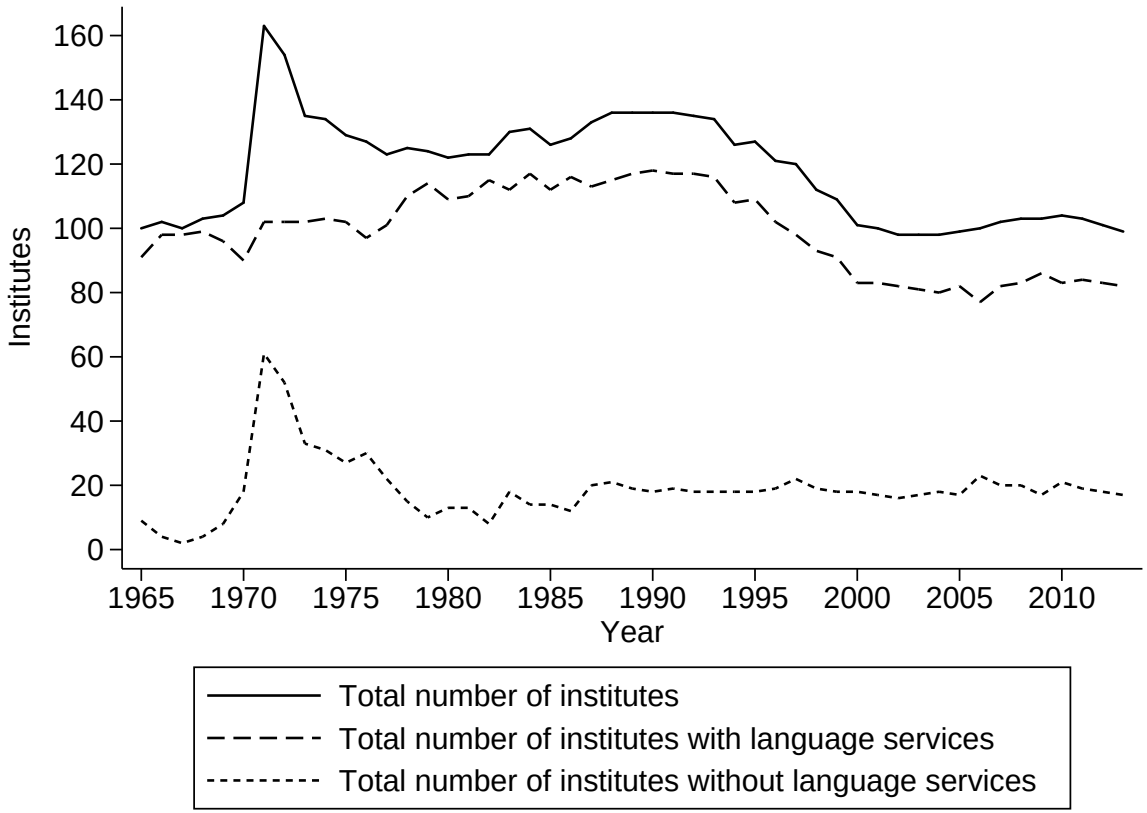

Figure 3: Number of institutes (our sample)

Data

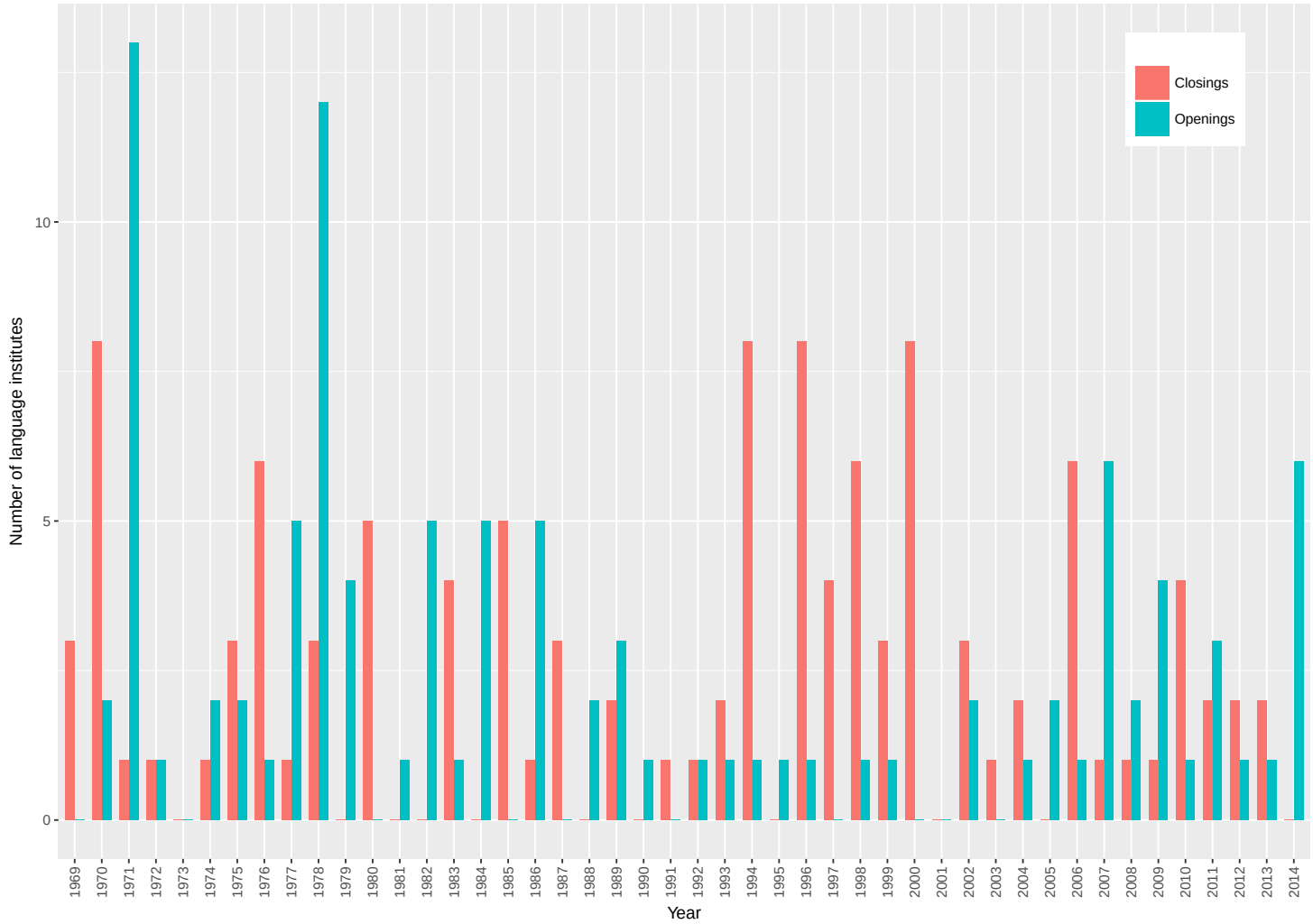

Figure 4: Openings and closings

▸ All, ▸ Language Institutes, ▸ Language Institutes (jr), ▸ Unbalanced Panel

Estimation Strategy

Fixed-effects model

$$y_{jt} = \alpha' GI_{jt} + \boldsymbol{\beta'} \boldsymbol{x_{jt}} + \boldsymbol{\phi'_t} \boldsymbol{d_t} + \boldsymbol{\phi'_j} \boldsymbol{d_j} + \boldsymbol{\phi'_j} \boldsymbol{d_{jT}} + \eta_{jt} \qquad (1)$$

y_{jt}	log of migration rate (annual migr. inflows (Destatis)/ population size of origin country (PWT))
GI_{jt}	number of (language) institutes
$\boldsymbol{x_{jt}}$	vector of control variables:
	log GDP/capita (PWT), EU, log population (PWT), conflicts (UCDP), log bilateral trade flows (Destatis) log migrant stock
$\boldsymbol{d_t}$, $\boldsymbol{d_j}$, $\boldsymbol{d_{jT}}$	time, origin-country and origin-country-10-year FE

▸ Summary Statistics

- Balanced panel dataset with 77 countries from 1968 – 2014.
- Regressions weighted by population size.

Results: basic specifications

DV: log migration rate	(1)	(2)	(3)	(4)
Number of language institutes	0.0427**	0.0537***	0.0558***	0.0685***
	(0.0175)	(0.0179)	(0.0184)	(0.0184)
log GDP per capita		-0.343***	-0.306***	-0.154*
		(0.0866)	(0.0931)	(0.0792)
EU member		0.454***	0.441***	0.443***
		(0.120)	(0.123)	(0.103)
log population		0.608*	0.537*	0.308
		(0.317)	(0.314)	(0.289)
Conflict			0.0602**	0.0525**
			(0.0257)	(0.0248)
log (Exports+Imports)			-0.0355	-0.0852*
			(0.0508)	(0.0435)
log (Migrant Stock / Population), lag=1				0.646***
				(0.0611)
Constant	-10.98***	-18.50***	-16.91***	-7.200
	(0.0910)	(5.406)	(5.592)	(5.112)
Observations	3,619	3,619	3,619	3,619
Adjusted R-squared	0.967	0.968	0.969	0.973
Year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes
Countries	77	77	77	77
Years	1968-2014	1968-2014	1968-2014	1968-2014

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Observations are weighted by population size.

Results: robustness checks I - GI institutes

DV: log migration rate	(1)	(2)	(3)	(4)	(5)
Number of institutes	0.0282**				
	(0.0124)				
Number of institutes without language services		-0.000272			
		(0.0114)			
Number of language institutes		0.0796***	0.0594***	0.0576**	
		(0.0198)	(0.0223)	(0.0230)	
Number of language institutes, lag=1			0.0469**	0.0439*	
			(0.0214)	(0.0225)	
Number of language institutes, lag=2				0.00794	
				(0.0194)	
log (Number of GI with language services per 1m inhabitants)					0.0299**
					(0.0143)
Observations	3,619	3,619	3,542	3,465	3,619
Adjusted R-squared	0.973	0.974	0.973	0.973	0.973
Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes	Yes
Other controls included	Yes	Yes	Yes	Yes	Yes
Countries	77	77	77	77	77
Years	1968-2014	1968-2014	1968-2014	1968-2014	1968-2014

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; observations are weighted by population size of the origin.

Results: robustness checks II – interaction effects

DV: log migration rate	(1)	(2) economic distance	(3) geographic distance	(4) linguistic distance	(5) EU	(6) conflict
Language institutes	0.0658***	0.0383**	0.117***	0.0204	0.0613***	0.0679***
	(0.0181)	(0.0151)	(0.0357)	(0.0214)	(0.0195)	(0.0182)
Language institutes *...		0.177***	-0.0615	0.0969***	0.0432	-0.0313**
		(0.0529)	(0.0409)	(0.0343)	(0.0419)	(0.0147)
Observations	3,619	3,619	3,619	3,619	3,619	3,619
Adjusted R-squared	0.973	0.974	0.973	0.973	0.973	0.973
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country*10year FE	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Countries	77	77	77	77	77	77
Years	1968-2014	1968-2014	1968-2014	1968-2014	1968-2014	1968-2014

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; observations are weighted by population size of the origin.

Possible issues

(a) Reverse causality
(b) Multi-lateral resistance (MLR)
(c) Omitted variable bias

(a) Reverse Causality: Switzerland

- Decision by Federal Foreign Office and GI to open (and close) institutes.
- Biased estimation possible: decision potentially not exogenous to migration to Germany.

$\rightarrow$ But: decision exogenous to migration flows to Switzerland.

$\Rightarrow$ Analysis with DV: log migration rate to Switzerland...

- to assess the relevance of reverse causality
- to study the language effect vs. information effect of the GI

by using variation of languages within Switzerland.

▸ Correlation, ▸ Rel. Size

Reverse Causality: Switzerland

DV: log migration rate	(1) Germany	(2) Switzerland (non-German-speak.)	(3) Switzerland (German-speak.)
Number of language institutes	0.0305*	0.0268	0.0702***
	(0.0167)	(0.0204)	(0.0265)
Observations	1,771	1,771	1,771
Adjusted R-squared	0.978	0.983	0.969
Year FE	Yes	Yes	Yes
Country Fe	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes
Other controls	Yes	Yes	Yes
Countries	77	77	77
Years	1992-2014	1992-2014	1992-2014

*** p<0.01, ** p<0.05, * p<0.1.

Shorter time-period ▸ Robust, ▸ Conclusion

(b) MLR: common correlated effects estimator

- Migration decisions not only influenced by the chosen destination country's attractiveness, but also by the attractiveness of other (alternative) destinations.

$\Rightarrow$ CCE estimator (Pesaran 2006) controls for multilateral resistance of migration (Bertoli, Fernandez-Huertas Moraga 2013).

$$y_{jt} = \alpha' GI_{jt} + \boldsymbol{\beta'} \boldsymbol{x_{jt}} + \boldsymbol{\phi'_t} \boldsymbol{d_t} + \boldsymbol{\phi'_j} \boldsymbol{d_j} + \boldsymbol{\phi'_j} \boldsymbol{d_{jT}} + \boldsymbol{\lambda'_j} \tilde{\boldsymbol{z}}_{\boldsymbol{t}} + \eta_{jt} \quad (2)$$

y_{jt}	log of migration rate
GI_{jt}	number of (language) institutes
$\boldsymbol{x_{jt}}$	vector of independent variables
$\boldsymbol{d_t}$, $\boldsymbol{d_j}$ and $\boldsymbol{d_{jT}}$	time, origin-country and origin-country-10-year FE

and the cross-sectional averages of independent and dependent variables weighted with ω_{jt} (population) interacted with country dummies λ_j

$$\tilde{\boldsymbol{z}}_{\boldsymbol{t}} = \frac{1}{\sum_j \omega_{jt}} \left(\sum_j \omega_{jt} y_{jt}, \sum_j \omega_{jt} x_{jt} \right)$$

Robustness checks: CCE estimator

DV: log migration rate	(1)	(2)	(3)	(4)
Number of language institutes	0.0680***	0.0776***	0.0723***	0.0683***
	(0.0121)	(0.0118)	(0.0116)	(0.0116)
Observations	3,619	3,619	3,619	3,619
Adjusted R-squared	0.974	0.981	0.983	0.984
Year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes
Countries	77	77	77	77
Years	1968-2014	1968-2014	1968-2014	1968-2014
CCE-test (p-value)	0	0	0	0

Observations are weighted by population size of the origin country; results are estimated with the CCE-estimator (Pesaran 2006); the CCE-test is a F-test on the joint significance of the cross-sectional averages of all dependent and independent variables interacted with country dummies. Standard errors in parentheses.

(c) Omitted variable bias

As next steps, we want to include information about possible other influences on migration that might affect the number of institutes as well:

- Language learning at schools:
  - Compulsory language learning at schools in Europe (data used by Fenoll and Kuehn 2016)
  - German schools abroad ("Auslandsschulen")
  - International Association of German Teachers ("IDV")
  - Percentage of pupils learning German (Eurostat)
- Language learning at universities:
  - Institutes of German studies abroad ("Germanistik")
  - German language courses
- Other influences:
  - Branches of chambers of commerce ("AHK")
  - German Academic Exchange Service ("DAAD")

Conclusion

- The number of institutes provided by the GI is positively correlated with migration rates to Germany.
- GI also affect migration flows to the German-speaking part of Switzerland, but not to the French- and Italian-speaking part.

$\Rightarrow$ Causal effect from language learning opportunities to migration flows.

$\Rightarrow$ The relationship is due to language learning and not due to other effects (like information effect) coming with the GI.

So:

- Language learning shapes international migration flows beyond linguistic properties, like linguistic distance.
- Contrary to children language learning, adult language learning is within reach of the policy-makers in the destination country.

Thank you for your attention!

For more information, see www.fiwi.uni-jena.de/dfg

Theoretical considerations

- Individuals choose the destination country which maximizes expected utility.
- Cost-benefit analysis on the basis of origin and destination country characteristics.
  - Benefits: monetary (e.g. wage, social security,...) and non-monetary (safety, partner, culture)
  - Costs: monetary (e.g. transportation, visa, temporary unemployment) and non-monetary (e.g. leaving family and friends, social integration)
- Language skills are costly, but increase expected income in destination country.
- On an aggregate level, language learning opportunities can be expected to play a role in the migration decision.

Data

Figure 5: Countries with GI in 2013

Go back

Data

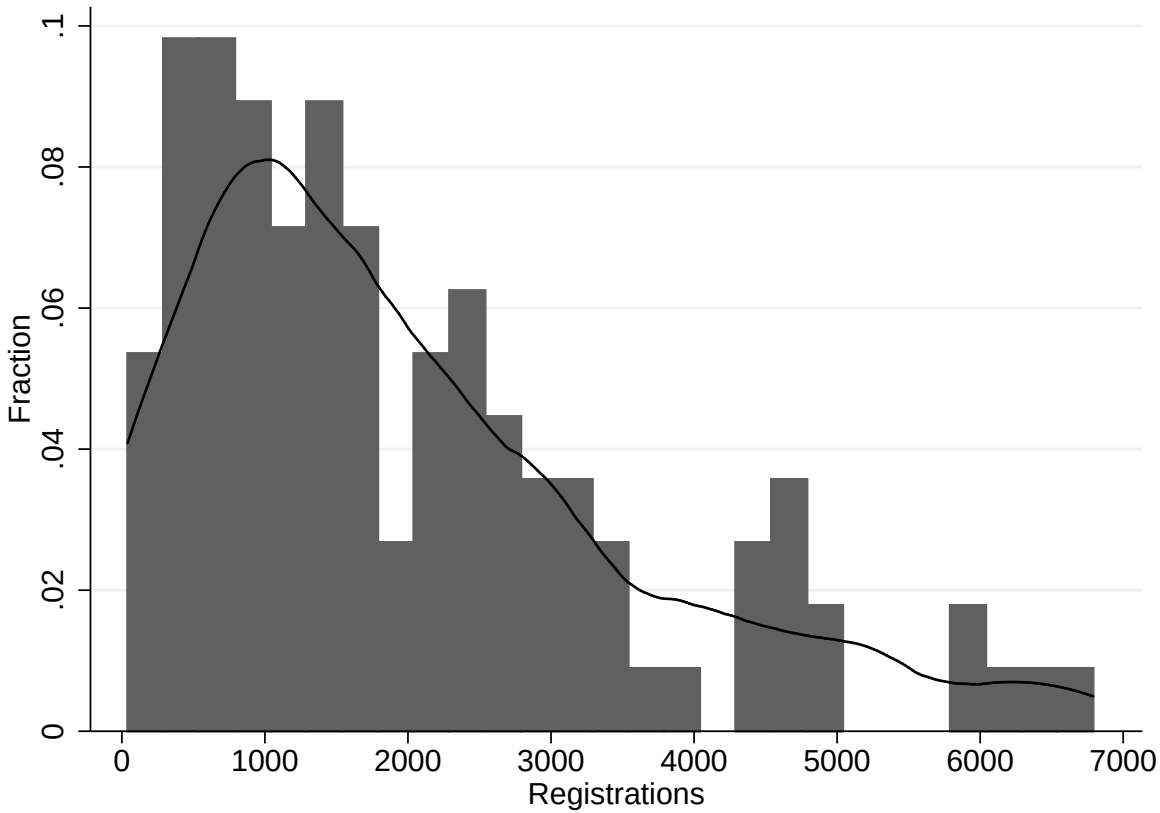

Figure 6: Distribution of institutes (our sample), 2013

Data

Summary statistics

Variable	Obs		Mean	Std. Dev.	Min	Max
Migration rate (emigration to Germany/pop.)	3619	Overall	0.00020	0.00061	5.48e-07	0.01084
Emigration to Germany	3619	Overall	4622.816	15894.45	3	251520
Number of institutes	3619	Overall	1.50898	1.99367	0	14
		Between		1.87117	0	7.93617
		Within		0.71969	-4.00166	9.08345
Number of language institutes	3619	Overall	1.26085	1.58470	0	9
		Between		1.48791	0	6.59575
		Within		0.57056	-3.65405	5.47361
Number of language institutes per 1m inhabit.	3619	Overall	0.07530	0.11899	0	0.78797
GDP per capita	3619	Overall	10062.91	11494.58	142.3924	65104.98
EU member	3619	Overall	0.11246	0.31598	0	1
Population in 1m	3619	Overall	41.31132	111.2261	0.20153	1295.292
Conflict	3619	Overall	0.23985	0.52199	0	2
Migrant stock/population	3619	Overall	0.00175	0.00511	2.86e-06	0.04581
Exports + Imports	3619	Overall	6.81e+09	1.94e+10	1818000	1.67e+11

▸ Variation (at least one GI), ▸ Go back

Data and estimation strategy

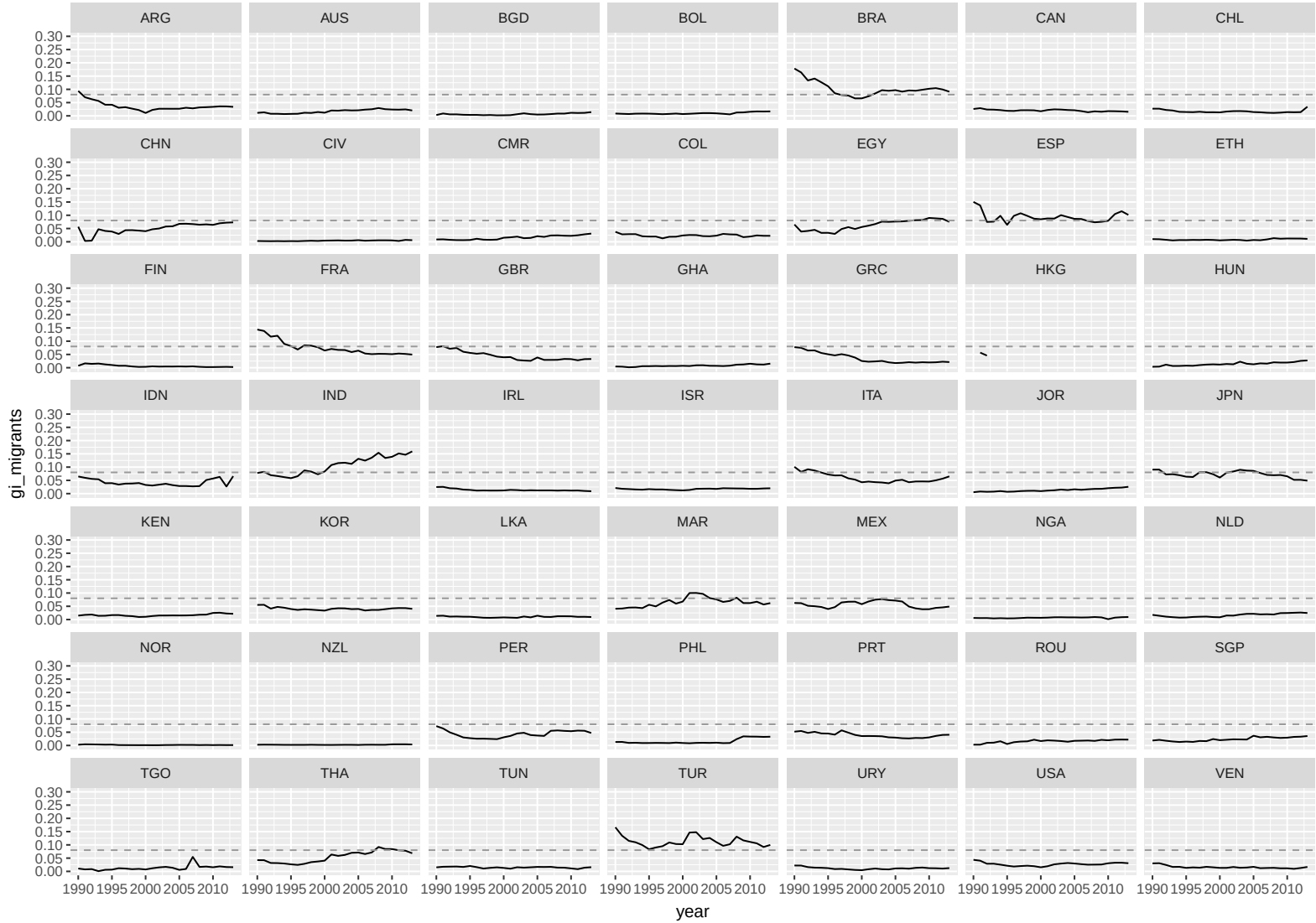

Figure 7: Share of course participants with migration intention on total migration (proxied), 2013

Data and estimation strategy

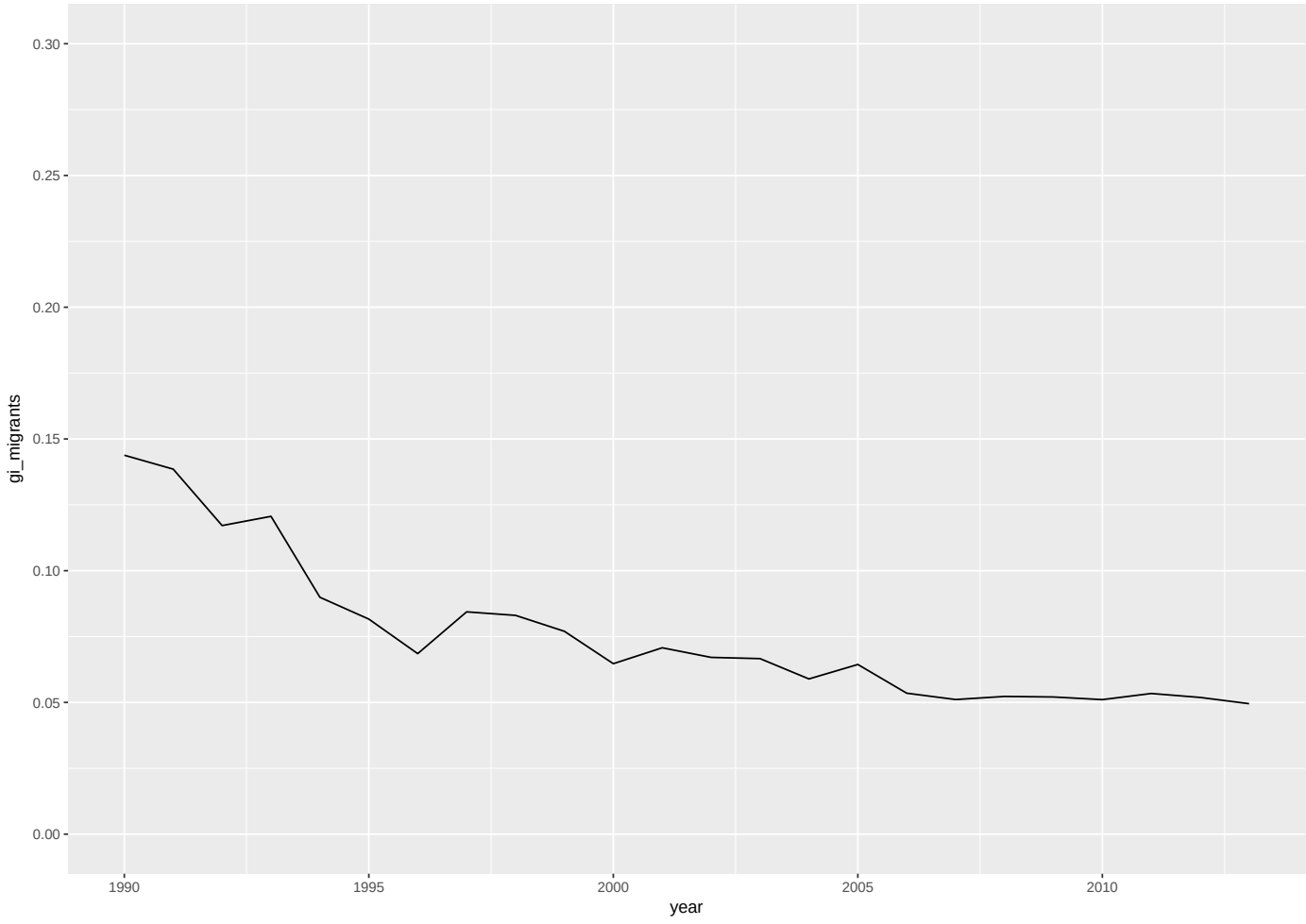

Figure 8: Share of course participants with migration intention on total migration (proxied), France

Results: robustness checks - unweighted

DV: log migration rate	(1)	(2)	(3)	(4)
Number of language institutes	0.0589***	0.0642***	0.0705***	0.0788***
	(0.0208)	(0.0204)	(0.0197)	(0.0185)
log GDP per capita		-0.534***	-0.454***	-0.236***
		(0.0751)	(0.0781)	(0.0619)
EU member		0.294*	0.312**	0.373***
		(0.152)	(0.151)	(0.113)
log population		-0.723**	-0.733**	-0.368
		(0.301)	(0.296)	(0.258)
Conflict			0.0883***	0.0910***
			(0.0229)	(0.0220)
log (Exports+Imports)			-0.0806**	-0.0739**
			(0.0354)	(0.0316)
log (Migrant Stock / Population), lag=1				0.673***
				(0.0448)
Constant	-10.95***	5.669	6.772	5.037
	(0.0835)	(5.355)	(5.352)	(4.667)
Observations	3,619	3,619	3,619	3,619
Adjusted R-squared	0.954	0.956	0.956	0.964
Year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes
Countries	77	77	77	77
Years	1968-2014	1968-2014	1968-2014	1968-2014

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Data

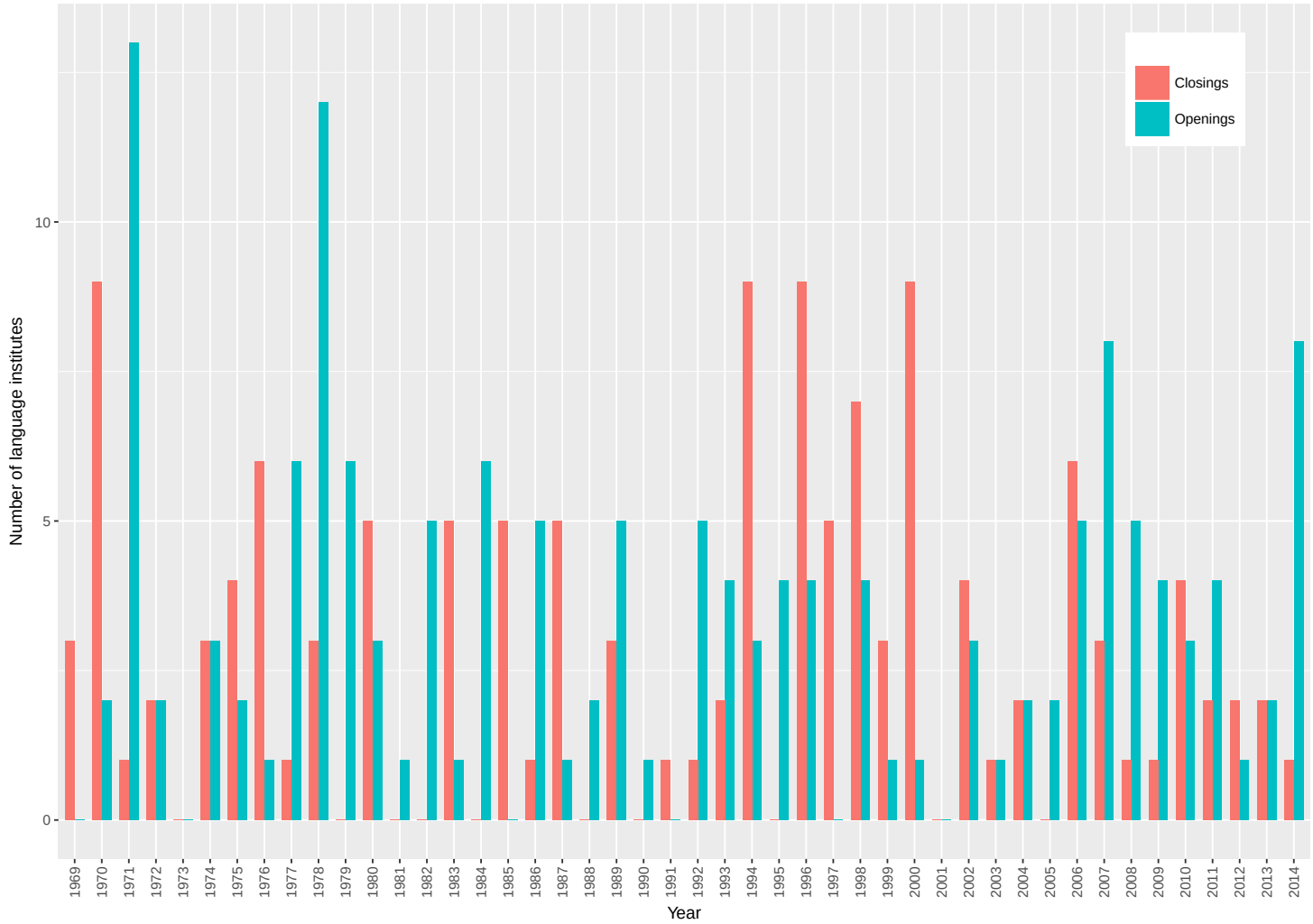

Figure 9: Openings and closings

▸ Go Back

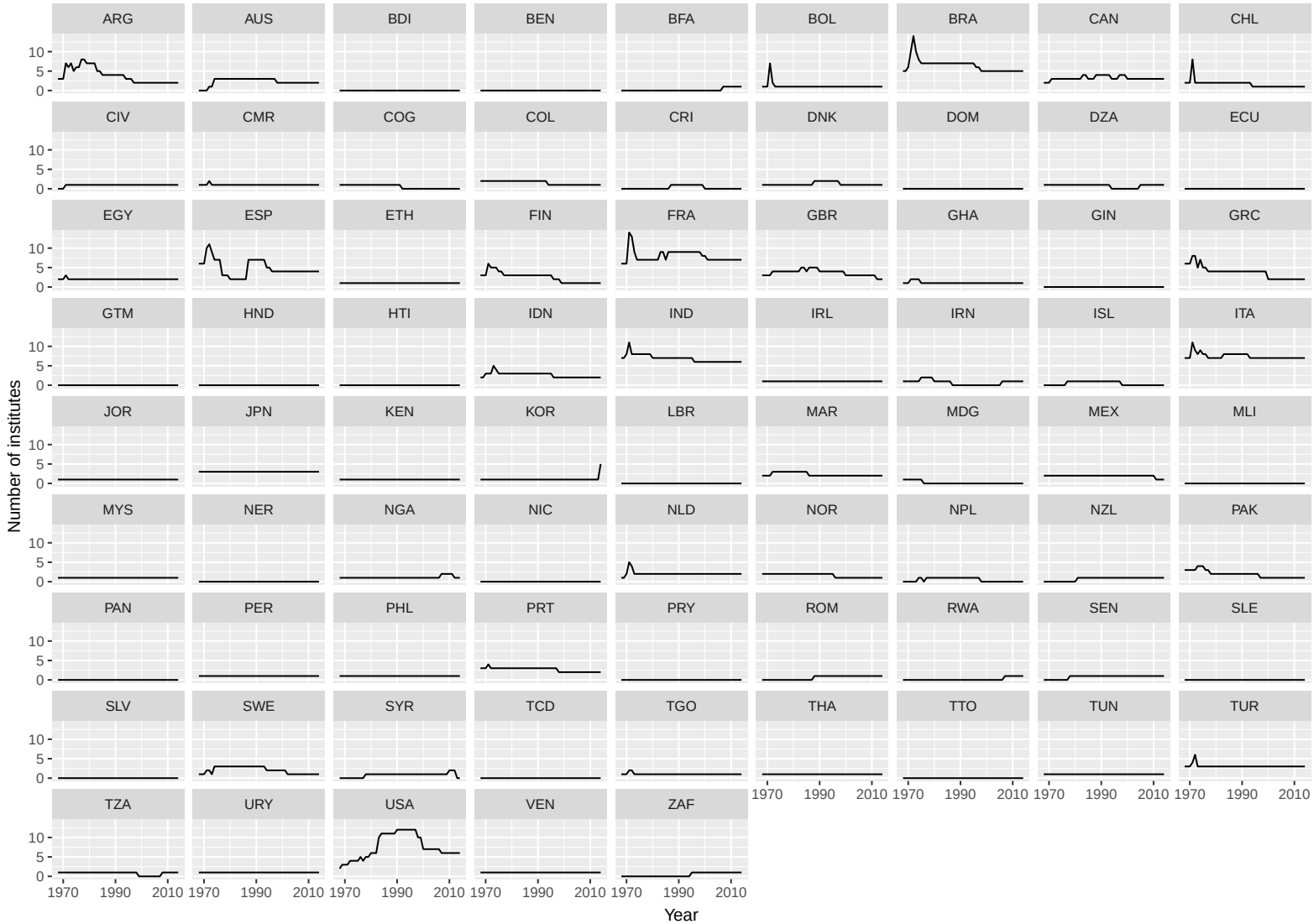

Figure 10: Numbers of all institutes, by origin countries

Go back

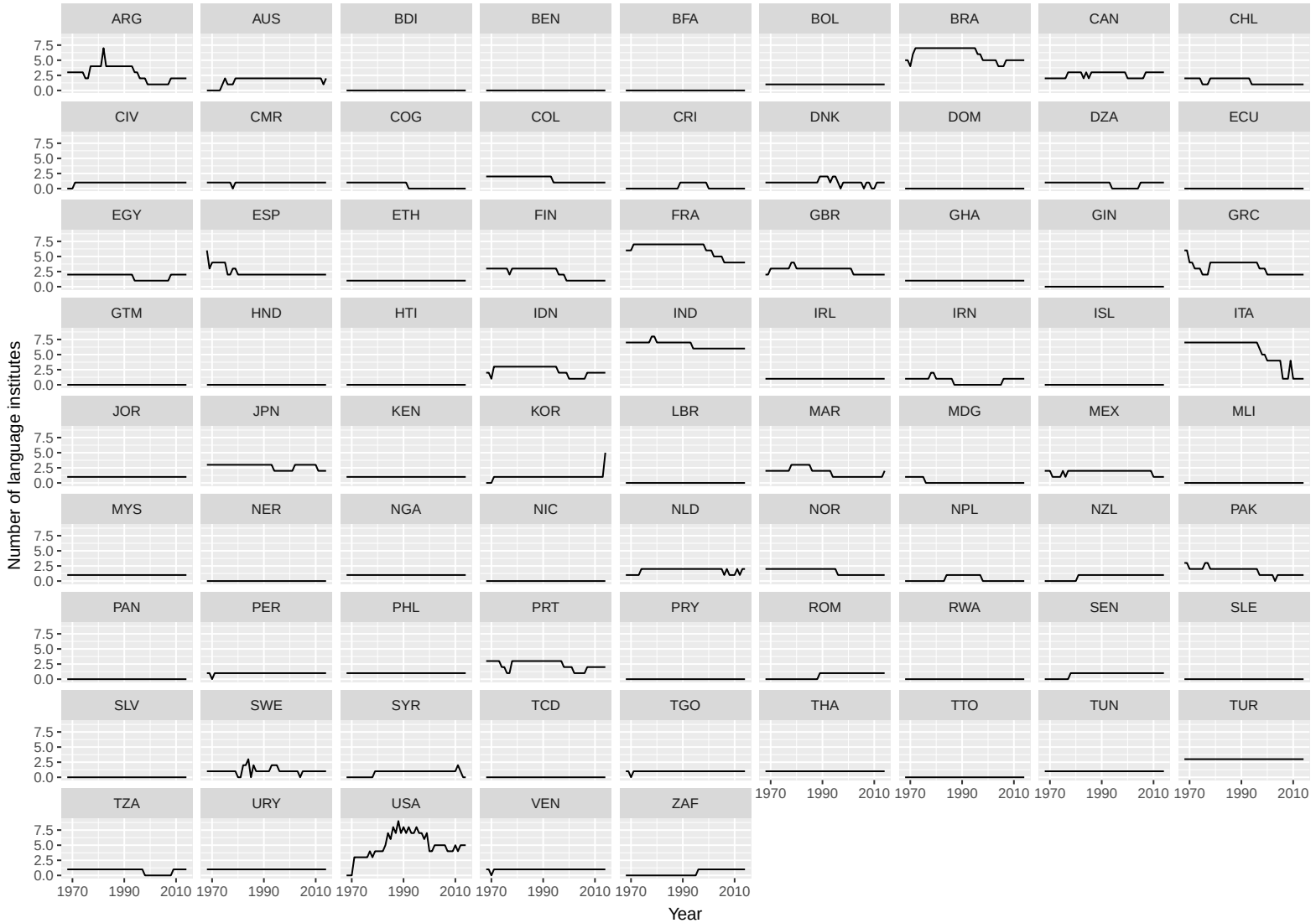

Figure 11: Numbers of language institutes, by origin countries

▸ Go back

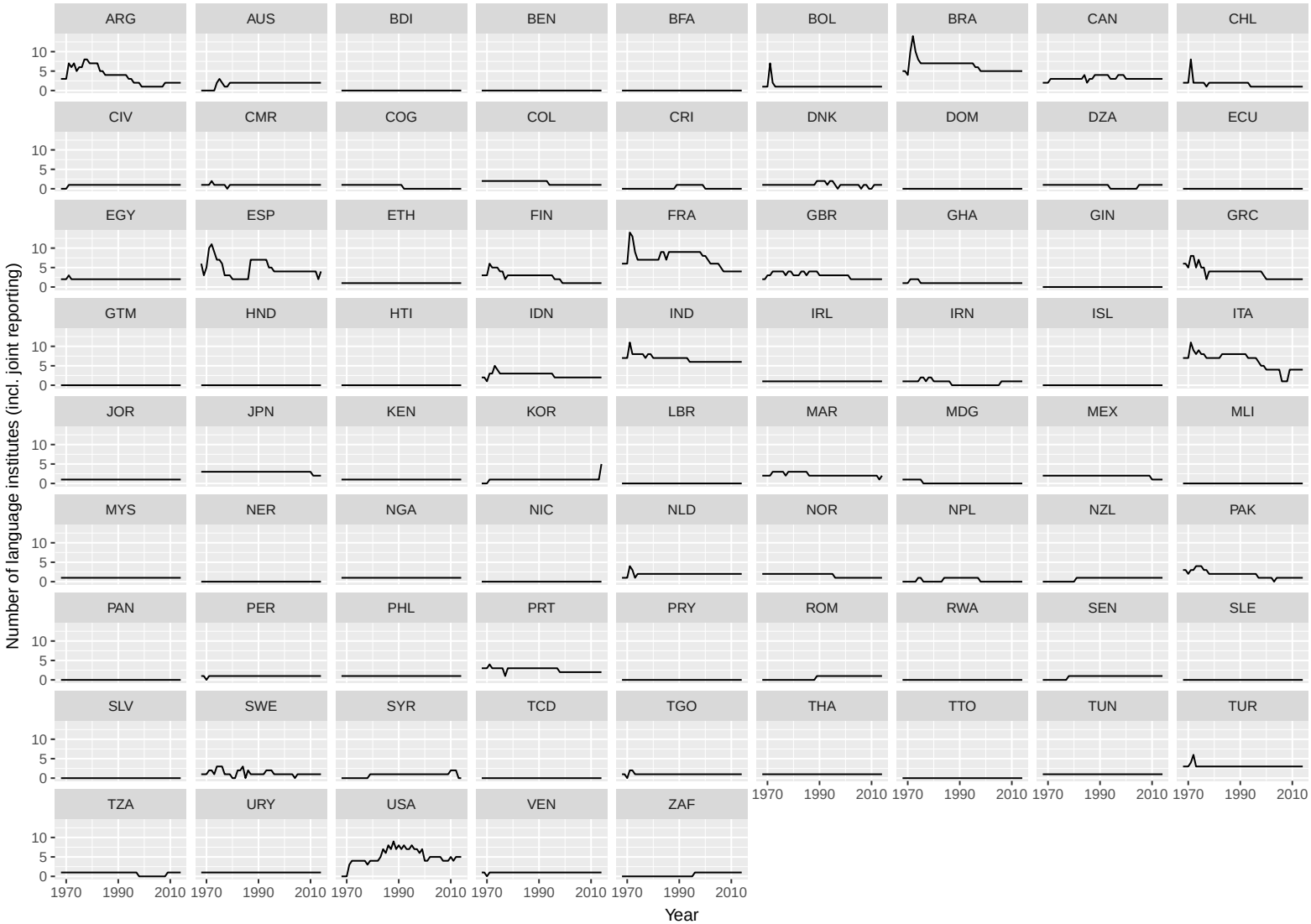

Figure 12: Numbers of language institutes, by origin countries language institutes (courses assumed to take place in case of joint reporting)

▸ Go back

Correlation between migration to Switzerland and to Germany

Figure 13: Correlation, "within" countries with and without GI

▸ Go back

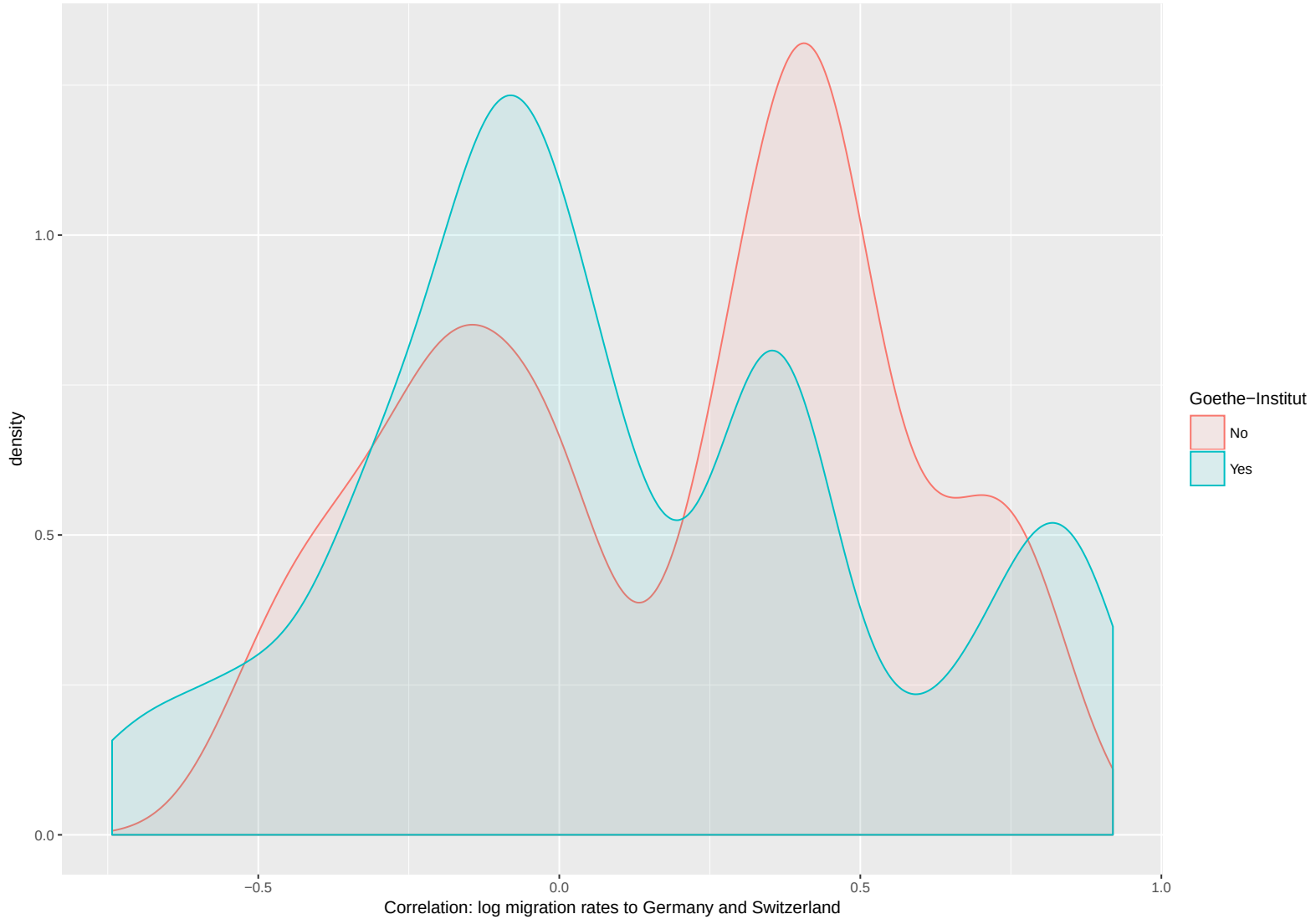

Figure 14: Correlation (density), countries with and without GI

Go back

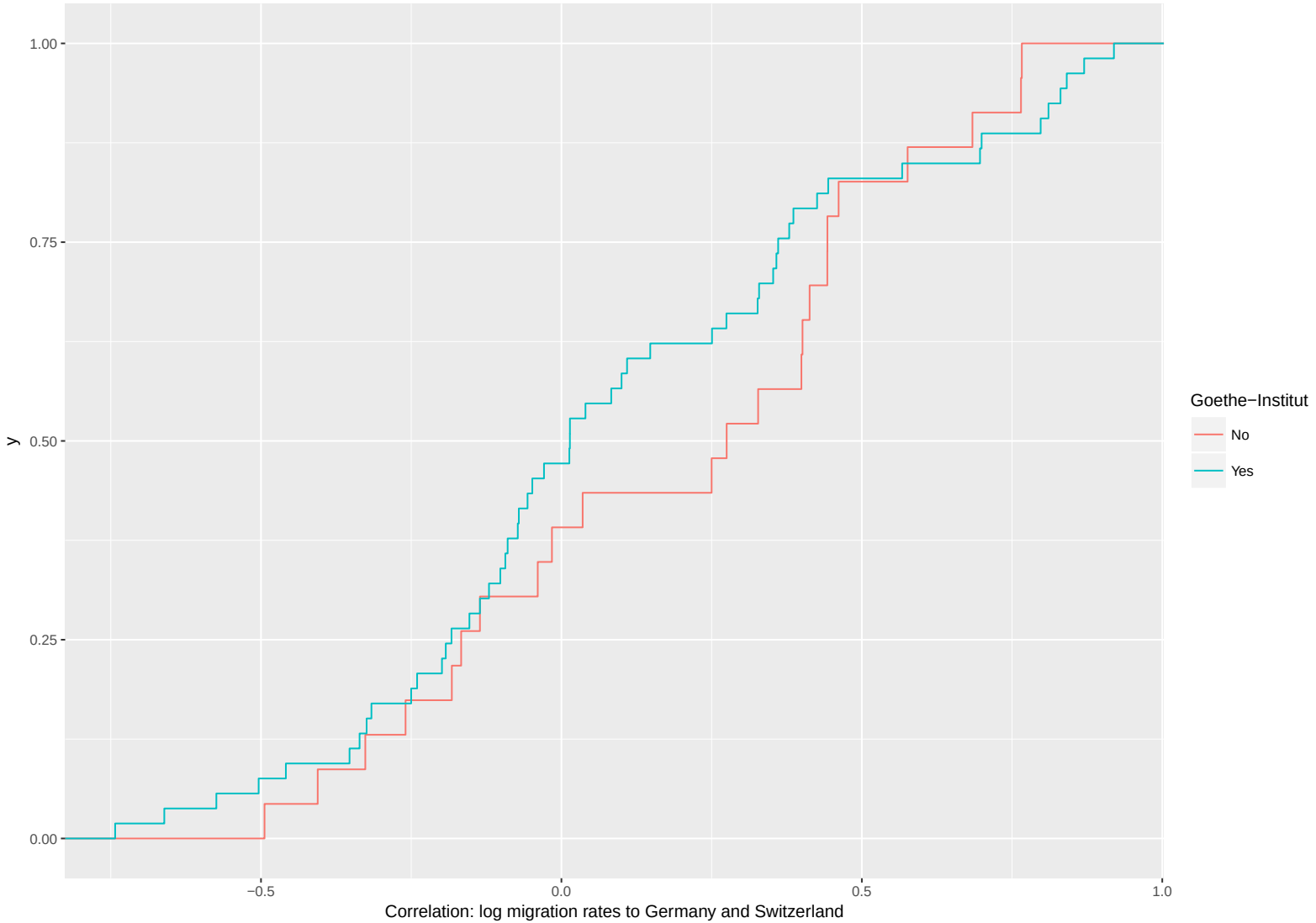

Figure 15: Correlation (cdf), countries with and without GI

Go back

Flows to Switzerland relative to Germany

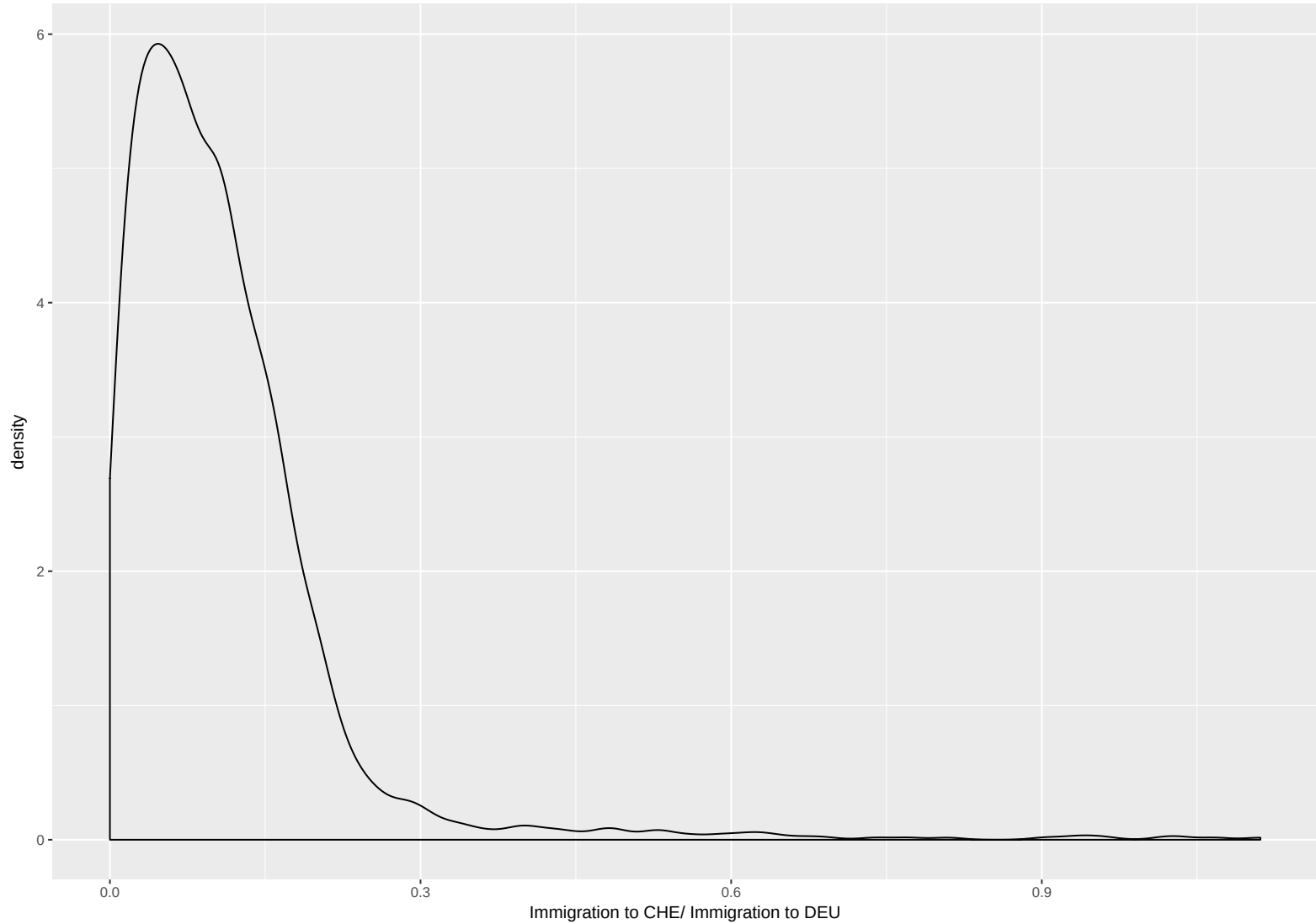

Figure 16: Swiss/German inflows ▸ Go back

Flows to Switzerland relative to Germany

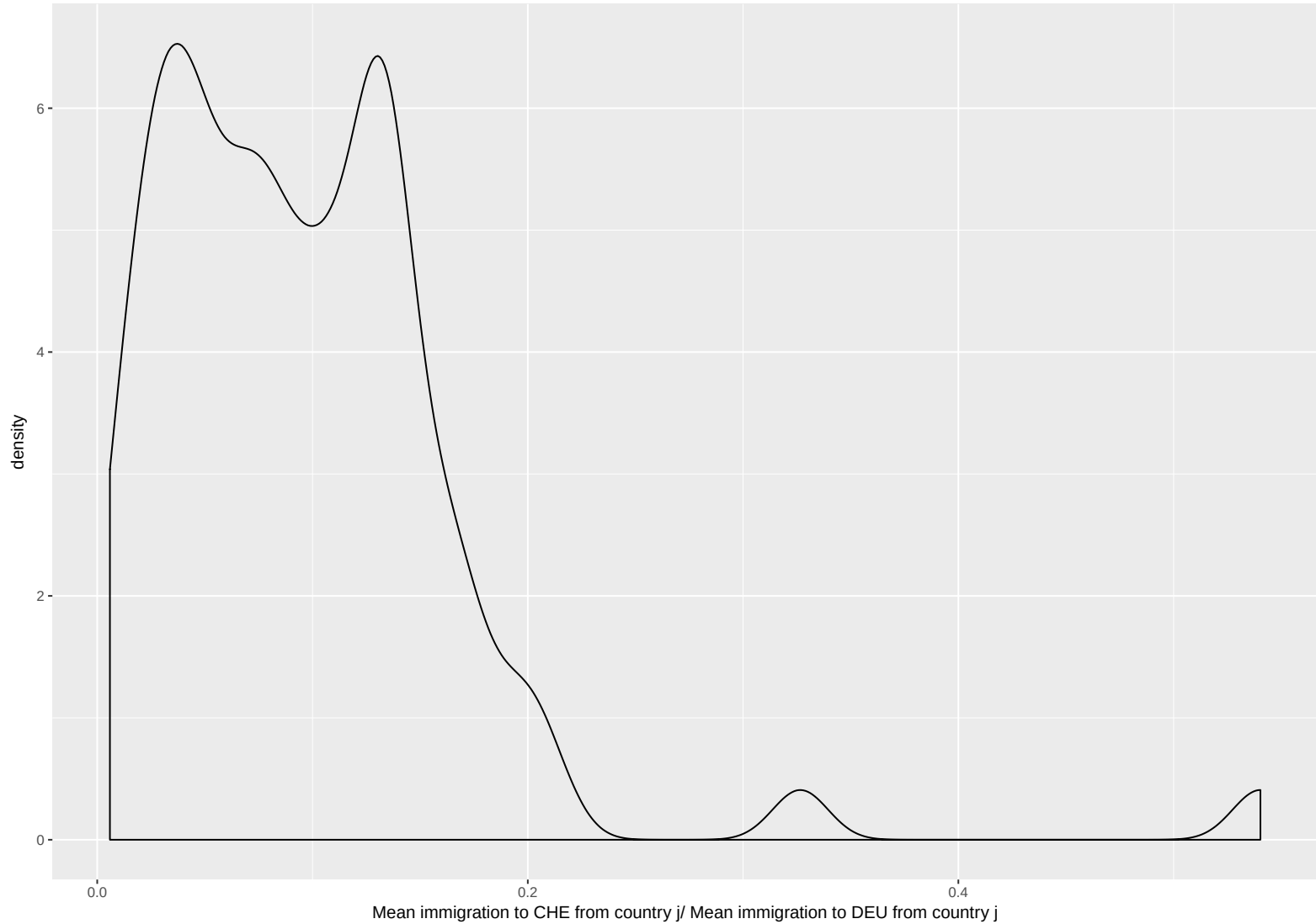

Figure 17: Swiss/German inflows, by origin countries (1992-2014)

Go Back

Data

Summary statistics (countries with at least one GI)

Variable	Obs		Mean	Std. Dev.	Min	Max
Number of institutes	2773 (59 countries)	Overall	1.96935	2.06901	0	14
		Between		1.91458	0.17021	7.93617
		Within		0.82221	-3.54129	9.54382
Number of language institutes	2632 (56 countries)	Overall	1.73366	1.62277	0	9
		Between		1.49150	0.17021	6.59575
		Within		0.66907	-3.18123	5.94643

▸ Go back

Reverse Causality: Switzerland

DV: log migration rate	(1) Germany	(2) Switzerland (non-German-speak.)	(3) Switzerland (German-speak.)	(4) Switzerland (German-speak.)
Number of language institutes	0.0305*	0.0268	0.0702***	0.0671*
	(0.0167)	(0.0204)	(0.0265)	(0.0389)
Observations	1,771	1,771	1,771	1,472
Adjusted R-squared	0.978	0.983	0.969	0.959
Year FE	Yes	Yes	Yes	Yes
Country Fe	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes
Countries	77	77	77	64
Years	1992-2014	1992-2014	1992-2014	1992-2014

*** p<0.01, ** p<0.05, * p<0.1. In (4), countries are excluded with significantly (at least on the 10 %-level) related migration flows to Germany and to the German-speaking part of Switzerland that have a variation in the number of language institutes in the period 1992 – 2014.

Shorter time-period ▸ Go back