– ONLINE APPENDIX –

Labor Supply and Automation Innovation: Evidence from an Allocation Policy*

Alexander M. Danzer [abcd] Carsten Feuerbaum [ah] Fabian Gaessler [efgh]

[a] KU Eichstaett-Ingolstadt [b] IZA Bonn

[c] CReAM, University College London [d] CESifo, Munich

[e] Universitat Pompeu Fabra [f] UPF Barcelona School of Management

[g] Barcelona School of Economics [h] Max Planck Institute for Innovation and Competition, Munich

February, 2024

*Email: alexander.danzer@ku.de, feuerbaum.c@gmail.com and fabian.gaessler@upf.edu.

Content

A Institutional Details

The ethnic German labor inflows result from a complex data generation process. In the following, we detail the decision-making of governmental authorities and immigrants throughout the application and allocation process, differentiating between the periods before and after the migrant allocation policy was introduced.

A.1 Ethnic German Inflows in the Pre-Allocation Period

Government

Before 1990, the German government admitted ethnic Germans without any quotas in an attempt to take responsibility for a population that suffered from discrimination and political persecution in the USSR.[1] With Soviet emigration restrictions crumbling in the late 1980s, significant numbers of ethnic Germans applied for emigration to Germany.[2]

In 1990, the German government introduced the *Aussiedleraufnahmegesetz* (Ethnic German Reception Act) to regulate the total number of admitted ethnic Germans. From 1993 onward, 225,000 ethnic Germans per year were admitted (see Appendix Figure A-1), while the number of applications quickly exceeded 550,000. This led to queuing in the countries of origin, and a constant stream of migrants over the years to come.[3]

Upon arrival, all ethnic Germans were sent to the central reception center Friedland for registration.[4] Historically, the Friedland center had been established by the British military administration in 1945 close to the American and Soviet occupational zones to handle the arriving refugee treks and exiled persons from the lost former eastern German territories.

Having received their German passports, ethnic Germans could choose their destination of residence in Germany. Their choice was recorded in the Friedland reception center, entering our data.

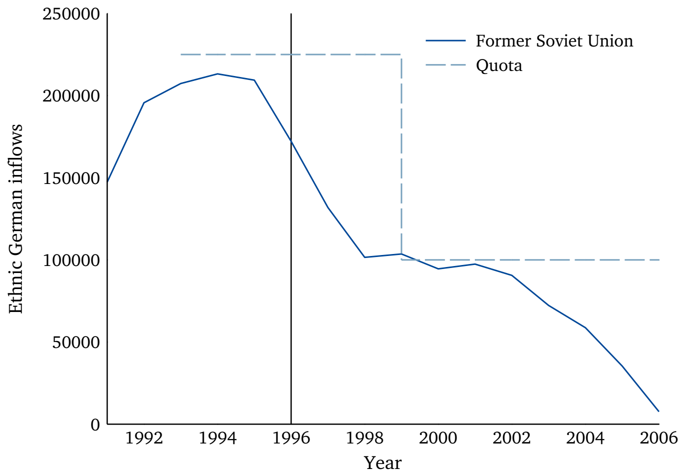
[1]The incoming ethnic Germans were descendants of German-speaking emigrants to Eastern Europe in the 18th and 19th centuries (Bade, 1990). Many ancestors had followed a resettlement offer by Russian Empress Catherine the Great in 1763 (granting land and religious freedoms).

[2]Between the 1950s and 1980s, few ethnic Germans could leave the USSR and its satellite states for political reasons. Under Gorbachev's government, the new law "On the procedure for entering and leaving the USSR" was passed, which eased the emigration of ethnic Germans substantially, so an early wave of several ten thousand ethnic Germans left the USSR between 1986 and 1990.

[3]As the German government covered travel costs (air tickets), financial constraints were no selection criterion for applications.

[4]For a few high-inflow years, two additional sub-centers were set up in Unna (North-Rhine-Westfalia) and Nuremberg (Bavaria).

Figure A-1: Ethnic German inflows by arrival year



Notes: Annual ethnic German inflows from the former Soviet Union to Germany. More than 95 percent of the incoming ethnic Germans came from the former Soviet Union during the analysis period from 1991 to 2006. Figure based on data from the German Federal Office of Administration (*Bundesverwaltungsamt*).

Ethnic Germans

According to the *Aussiedleraufnahmegesetz* (Ethnic German Reception Act), ethnic Germans had to apply for naturalization in their home countries and wait for approval of their application. To this end, they had to register at the German embassies in the Former Soviet Union, and document German ancestry (through German language proficiency) and personal experience of discrimination due to their German nationality.[5] Once granted the status of ethnic German (*Spätaussiedler*), individuals could apply for emigration to Germany. The emigration request could include family members and children even if these did not have the status of ethnic Germans themselves (so-called *Einbeziehung*).[6]

Many ethnic Germans tended to settle in regions near the Friedland reception center, specifically targeting areas with manufacturing industries and those with rural or provincial settlement structure (Friedrichs, 2022; Panagiotidis, 2021). This pattern can be explained by the limited information ethnic Germans initially had about Germany and its different regions. No ethnic German had been born in Germany, and close to none had ever been to Germany.[7] Having received their education and training in the Soviet Union, ethnic Germans were unacquainted with the industrial structure or specific features of the German labor market. As a result, a small number of regions received exceptionally high per-capita inflows during the pre-allocation period (see Appendix Table A-1).

[5]Note that German nationality was ascribed by Soviet authorities not based on citizenship, but ethnicity.

[6]Unlike the US, Germany had an implicit ethnic, cultural-linguistic nationality concept. This implies that children of ethnic Germans will automatically hold German nationality.

[7]The predominant residential areas of ethnic Germans in Siberia and Kazakhstan were several thousand kilometers away. They lived in rural or peri-urban settlements that were significantly poorer than the destination areas in Germany.

Table A-1: Characteristics of the ten regions with the highest per-capita inflows

Region	State	Distance to entry port	Sector share of manuf.	Per-capita inflows: Pre-allocation period	Per-capita inflows: Allocation period	Change in %
Cloppenburg	NS	239km	0.3699	0.0539	0.0040	−93
Paderborn	NW	101km	0.4139	0.0389	0.0202	−48
Gummersbach	NW	240km	0.4674	0.0368	0.0139	−62
Höxter	NW	58km	0.3521	0.0356	0.0168	−53
Wolfsburg	NS	114km	0.5895	0.0354	0.0050	−86
Korbach	HE	82km	0.3773	0.0341	0.0044	−87
Soest	NW	148km	0.4043	0.0333	0.0205	−38
Minden	NW	127km	0.3894	0.0327	0.0195	−40
Gütersloh	NW	138km	0.5137	0.0309	0.0199	−36
Detmold	NW	80km	0.4137	0.0300	0.0175	−41
Mean region	–	229km	0.3548	0.0150	0.0110	−26

Notes: The table presents selected characteristics of the ten regions with the highest per-capita inflows in the pre-allocation period. NS: Lower Saxony, NW: North Rhine-Westphalia, HE: Hesse. Entry port refers to the Friedland reception center. Sector share of manufacturing in 1991.

A.2 Ethnic German Inflows in the Allocation Period

Government

With the introduction of the allocation rule, Germany enacted annual caps on immigrant inflows as well as a regional distribution mechanism, which bound newly arriving ethnic Germans to a designated place of residence for three years. Immigrants arriving at the central reception center were distributed across states according to a quota. This quota followed the *Königssteiner Schlüssel*, which combines population and economic criteria (legally enshrined in the *Bundesvertriebenengesetz BVFG*). More specifically, the quota is based on population size (with weight 1/3) and tax revenues (with weight 2/3). All states in West Germany except Bavaria and Rhineland-Palatinate implemented the allocation policy in March 1996, with Lower Saxony following in April 1997 and Hesse in January 2002.

Within states, slightly different distribution schemes were used to allocate ethnic Germans across municipalities (according to the so-called *Landesaufnahmegesetze* or *Durchführungsverordnungen*). Most states chose an internal distribution scheme according to population (Saarland, Schleswig-Holstein) or a combination of population and area (Baden-Wurttemberg, North-Rhine-Westfalia).[8] Further aspects taken into account by some states were temporary housing shortages/offers, particularly in the social housing market (Piopiunik and Ruhose, 2017; Panagiotidis, 2021; Haug and Sauer, 2007), and political interventions. Specifically, a group of regions that had experienced substantial inflows in

[8] The state-specific allocation parameters and weights used each year are not documented. Given that certain states and regions irregularly updated these parameters (Jahn and Steinhardt, 2023), determining the precise distribution scheme from regional characteristics proves challenging.

the pre-allocation period successfully requested a more even distribution of ethnic Germans: the *Gifhorn Declaration for the Integration of Ethnic German Immigrants* was signed by the counties Cloppenburg, Emsland, Gifhorn, Nienburg/Weser, Osnabrück, Salzgitter, and Wolfsburg in 1995 (Haug and Sauer, 2007; Jahn and Steinhardt, 2016).[9]

The distribution scheme was closely followed in the allocation period (Dietz, 2006). In a survey by Haug and Sauer (2007), the representatives of municipalities agreed that the distribution scheme had created a relatively even distribution of ethnic Germans and balanced the social welfare expenditure burden across regions.[10] Indeed, the per-capita inflows became more equal between regions (see Appendix Figure A-2).

Before the ethnic Germans departed from the Friedland reception center, their designated place of residence (state and county) was recorded, entering our data.[11]

Ethnic Germans

As in the pre-allocation period, ethnic Germans still had to apply in their home countries for emigration. However, the immigration approval now included one mandatory destination state, based on the central allocation according to the *Königssteiner Schlüssel*. Ethnic Germans could express a preferred state of allocation when applying for emigration, which they did in about 50% of cases. According to a representative survey among ethnic Germans (Haug and Sauer, 2007), the most important reason (88%) for stating a preference was to live close to family members/relatives who had previously emigrated. Only 10% stated that they expected job opportunities. While a majority of individuals with a preference were allocated to their chosen state, this was conditional upon the compatibility with the quota-based allocation limits. Notwithstanding, survey evidence indicates that immigrants ultimately had no control over their designated place of residence.

Discrepancies between allocated numbers and actual residential uptake are challenging to quantify as naturalization made ethnic Germans statistically indistinguishable from native-born Germans. Nonetheless, it is widely argued that the percentage of ethnic Germans who did not settle in their designated place of residence was very low (Dietz, 2006). The primary reason is that migrants who chose not to live in their designated place of residence forfeited their social welfare (*Sozialhilfe*) and unemployment benefits. Indeed, only ten percent of ethnic Germans did not receive any benefits, including those who found employment immediately, suggesting an upper bound of non-compliance. Consistent with a

[9]Indeed, several regions that experienced a high per-capita inflow during the pre-allocation period received fewer ethnic Germans afterward (see Appendix Figure A-3).

[10]The states paid municipalities financial support (so-called *Integrationspauschalen*) for each immigrant for the first two years.

[11]Note that the data also include family members (e.g., partners, parents-in-law, children) even if they were not legally certified ethnic Germans. These "tied movers" were distributed in the allocation process together with their ethnic German family members, without being legally subject to the individual allocation.

low share of non-compliance, Haug and Sauer (2007) find a deviation of below 0.5% between the recorded allocation locations and placement statistics for a subset of states and years.[12]

Likewise, evidence suggests low inter-regional mobility of ethnic Germans after the binding allocation period of three years ended. First, ethnic Germans often relied on publicly administered social housing, which made inter-regional relocations unattainable (Heller et al., 2002). Second, larger families of ethnic Germans tended to acquire residential property soon after immigration, contributing to their limited geographic mobility (Fuchs, 1999). A panel study by the BiB, conducted among ethnic Germans in the 1990s, reveals very low inter-regional mobility (Worbs et al., 2003). Consistent with this, survey results suggest that ten years after the allocation, more than two-thirds of ethnic Germans still lived in their designated place of residence (Haug and Sauer, 2007).[13]

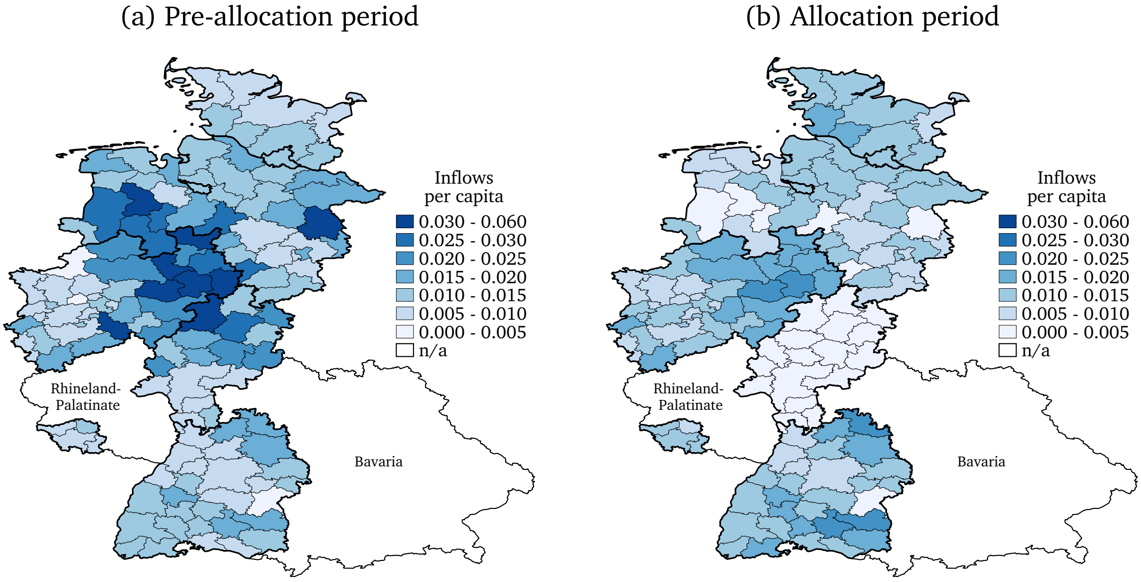
To summarize, evidence from multiple secondary sources supports the notion of high compliance with the allocation policy, indicating that the allocated numbers closely mirror actual inflows.[14]

[12]According to the authors, some of the deviations may still be due to other reasons, such as delayed registrations and hospital admissions.

[13]This is probably still an underestimation, as the survey by Haug and Sauer (2007) includes disproportionately few ethnic Germans residing in residential property.

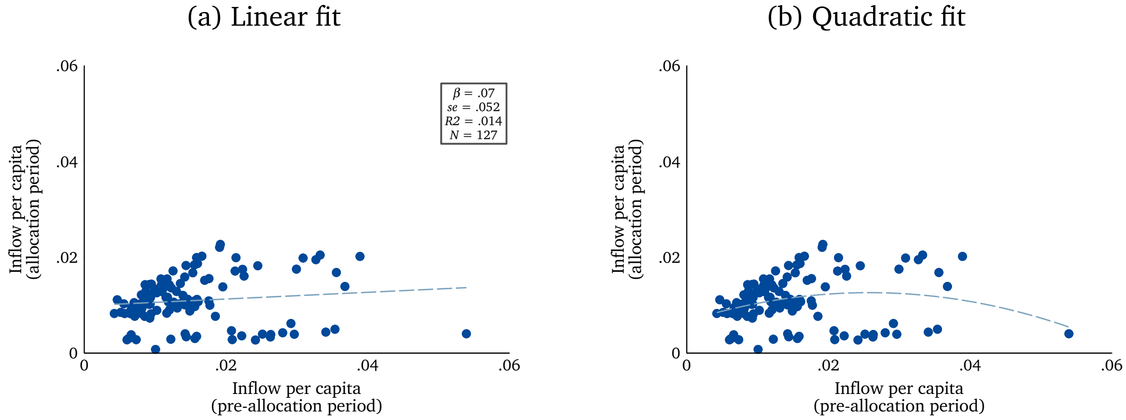
[14]Personal conversations with Barbara Dietz, an expert on ethnic German immigration, further corroborate this.

Figure A-2: Cumulative per-capita inflows across regions in pre- and allocation period



Notes: The figure illustrates the geographical distribution of the cumulative per-capita inflows in the pre-allocation and the allocation period. Per-capita inflows are defined as the cumulative inflow of ethnic Germans in the allocation (pre-allocation) period divided by the mean population during the pre-allocation period.

Figure A-3: Relationship between per-capita inflows in the pre-allocation and allocation period



Notes: The figure illustrates the relationship between per-capita inflows in the pre-allocation and the allocation period. Observations are at the level of the region, each scatter representing a region.

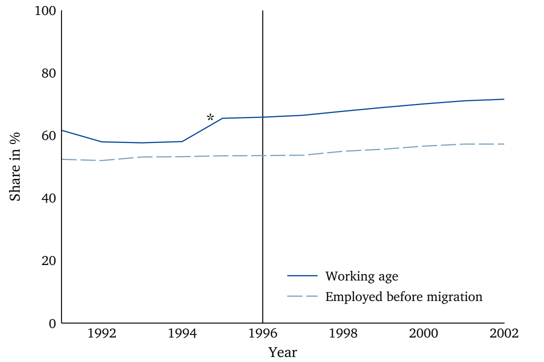
B Further Descriptives and Results

B.1 Ethnic German Inflows

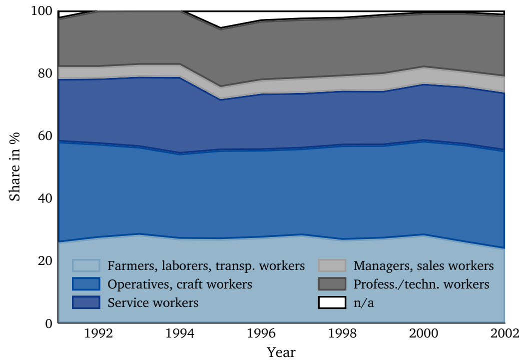
Characteristics of ethnic Germans

Figure B-1: Occupations and demographics of incoming ethnic Germans

(a) Employment before migration and working age by arrival year



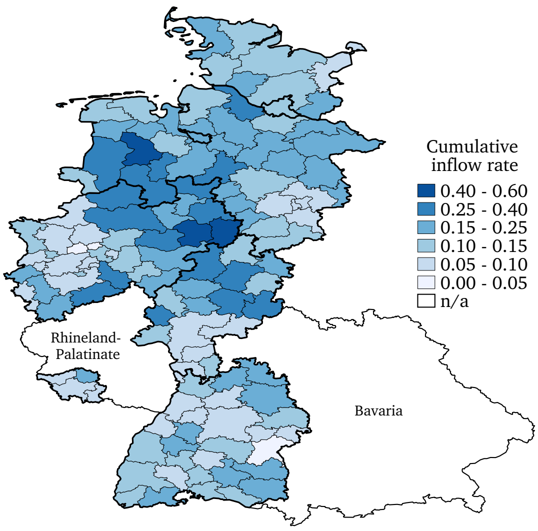
(b) Last occupation in country of origin by arrival year



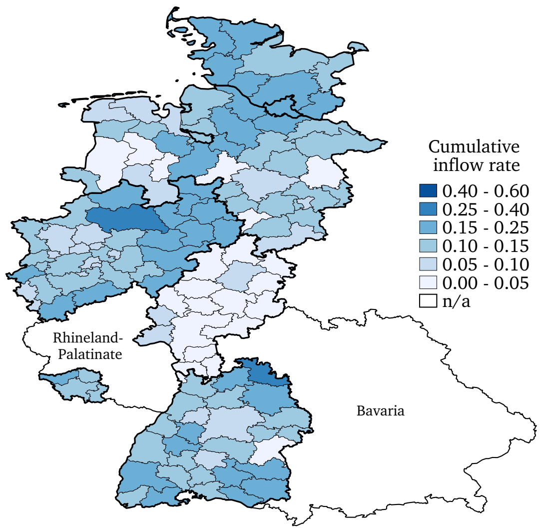
Notes: Top figure: shares with respect to employment before migration and working age by arrival year. * refers to a change in the definition of working age (age between 18-64 (1991-1994) and 15-64 (1995-2002)). Bottom figure: occupation shares of the last occupation in the country of origin of incoming ethnic Germans by arrival year. Own calculations based on data from Glitz (2012). Original data from the *Jahresstatistik für Aussiedler*, published annually by *Bundesverwaltungsamt*.

Figure B-2: Cumulative inflow rate across regions in pre- and allocation period

(a) Pre-allocation period



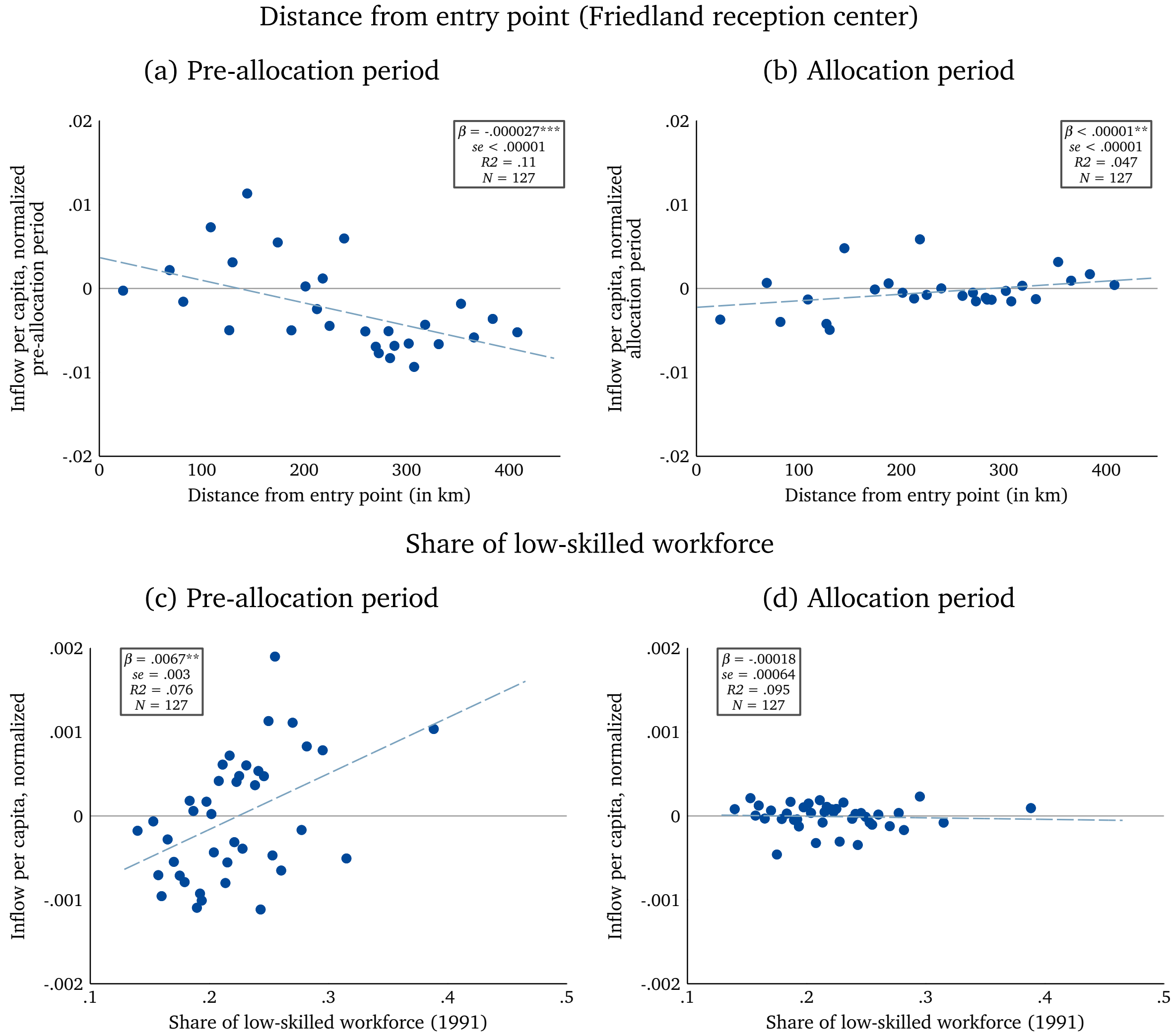
(b) Allocation period



Notes: The figure illustrates the geographical distribution of the region-specific cumulative inflow rate in the pre-allocation and the allocation period. The cumulative inflow rate is defined as the cumulative inflow of ethnic Germans in the allocation (pre-allocation) period divided by the average stock of the low-skilled workforce during the pre-allocation period.

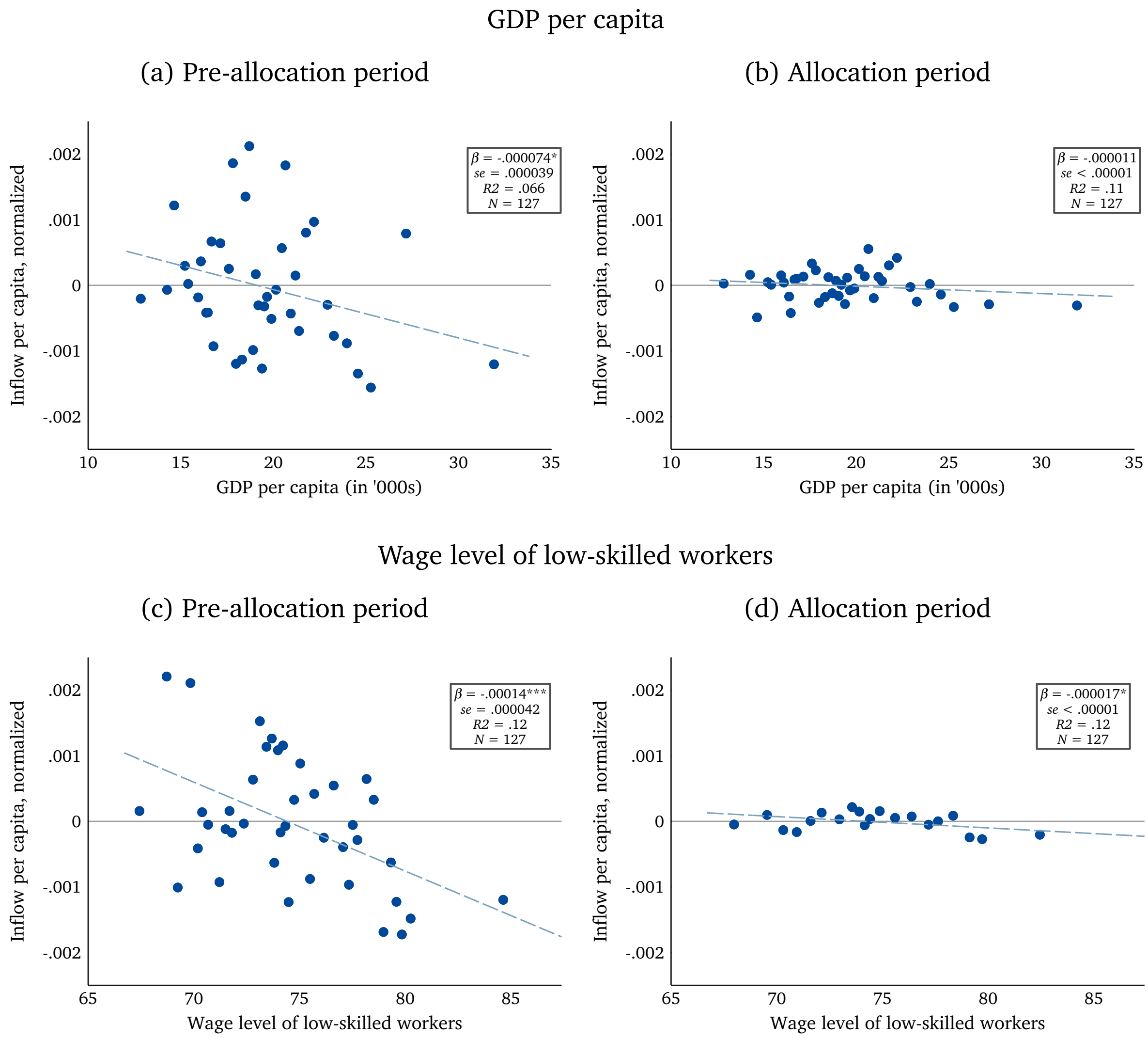
Sorting before and during the allocation period

Figure B-3: Differences in per-capita inflows by distance and share of low-skilled workforce



Notes: The figure illustrates the relationship between the distribution of regional differences in per-capita inflows and regional characteristics, in the pre-allocation and the allocation period. Regional differences in inflows are normalized. In Panel (a) and (b), the regional characteristic is the distance from the entry point (Friedland reception center). In Panel (c) and (d), the regional characteristic is the share of the low-skilled workforce as of 1991. Observations are at the level of the region, aggregated into 40 bins representing 2.5 percentiles each.

Figure B-4: Differences in per-capita inflows by region characteristics I

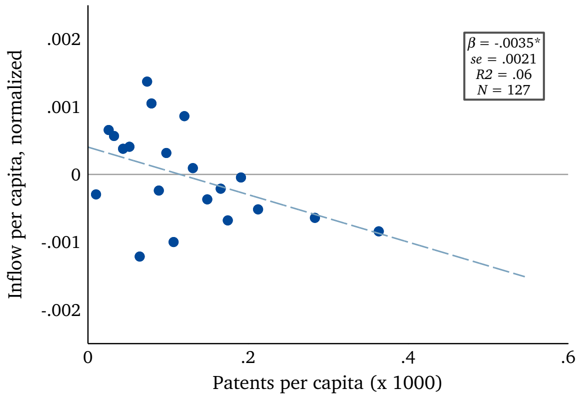


Notes: The figure illustrates the relationship between the distribution of regional differences in per-capita inflows and regional characteristics in the pre-allocation and the allocation period. Regional differences in inflows are normalized. In Panel (a) and (b), the regional characteristic is the GDP per capita as of 1991. In Panel (c) and (d), the regional characteristic is the wage level of low-skilled workers (farmers, laborers, and transport workers) as of 1995 (as no earlier data are available). Observations are at the level of the region, aggregated into 40 bins representing 2.5 percentiles each.

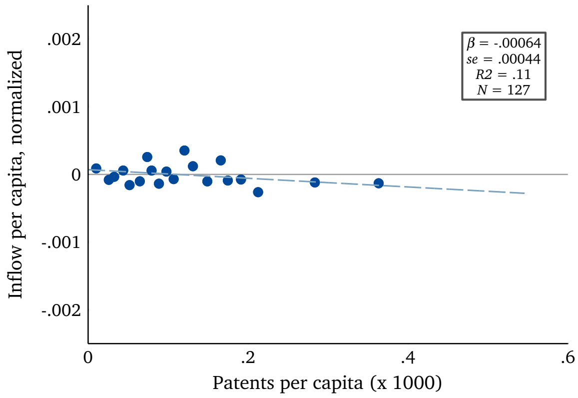
Figure B-5: Differences in per-capita inflows by region characteristics II

Patents per capita

(a) Pre-allocation period

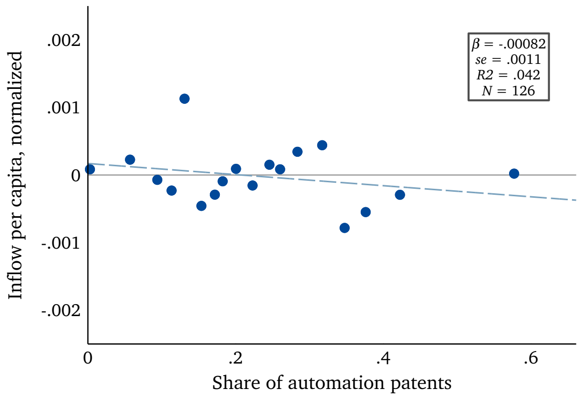


(b) Allocation period

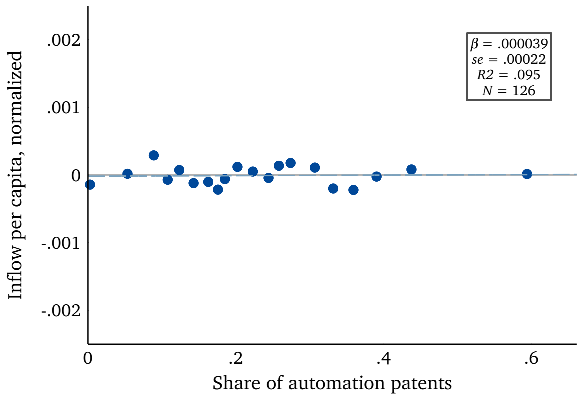


Share of automation patents

(c) Pre-allocation period



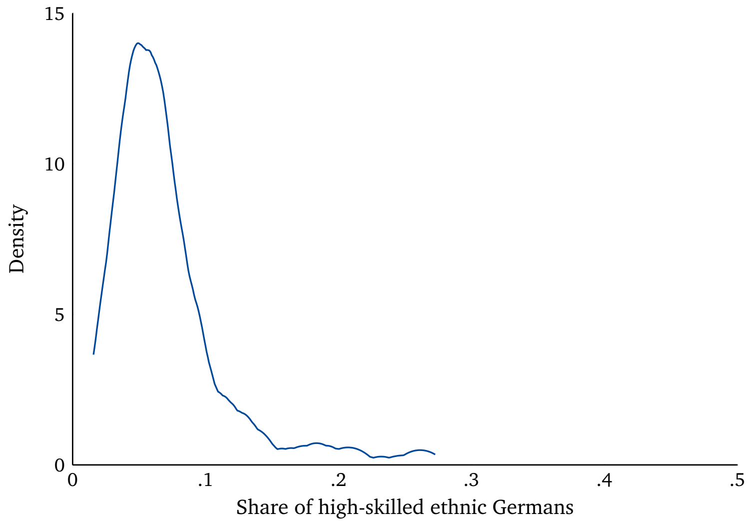
(d) Allocation period



Notes: The figure illustrates the relationship between the distribution of regional differences in per-capita inflows and regional characteristics in the pre-allocation and the allocation period. Regional differences in inflows are normalized. In Panel (a) and (b), the regional characteristic is the GDP per capita as of 1991. In Panel (c) and (d), the regional characteristic is the share of automation patents as of 1991. Observations are at the level of the region, aggregated into 40 bins representing 2.5 percentiles each.

Sorting of high-skilled ethnic Germans

Figure B-6: Distribution of the regional share of high-skilled Ethnic Germans



Notes: The figure illustrates the distribution of the regional share of high-skilled Ethnic Germans during the allocation period, with skill level determined by education. Skill information of ethnic Germans extrapolated from a small random sample of ethnic Germans identified in social benefits data. Observations are at the level of the region.

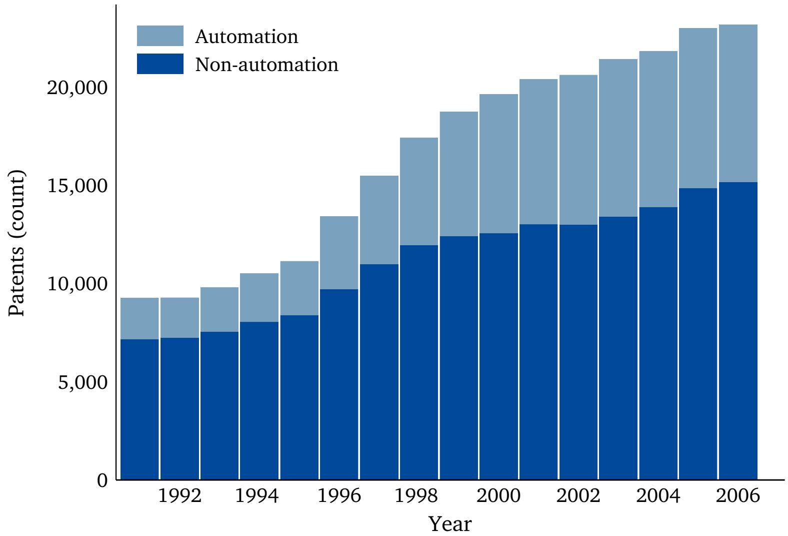
Table B-1: Correlation between the share of high-skilled ethnic Germans during the allocation period and region characteristics in the pre-allocation period

	(1)	(2)	(3)	(4)
Dep. Var.:	Share of high-skilled ethnic Germans			
Wage level (professional and technical workers)	0.001			
	(0.001)			
GDP per capita (in '000s)		0.000		
		(0.001)		
Share of automation patents			0.061	
			(0.050)	
Share of high-skilled workforce				−0.008
				(0.191)
Observations	125	125	125	125
R-squared	0.005	0.001	0.012	0.000

Notes: The table presents the results of linear regression analyses. In all columns, the dependent variable is the share of high-skilled ethnic Germans. Skill information of ethnic Germans extrapolated from a small random sample of ethnic Germans identified in social benefits data. The independent variables are the pre-allocation means of the wage level (professional and technical workers) in column (1), GDP per capita in column (2), the share of automation patents in column (3) and the share of the pre-existing high-skilled workforce in column (4). The regressions are estimated at the region level and weighted by the regional population in 1991. Standard errors are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

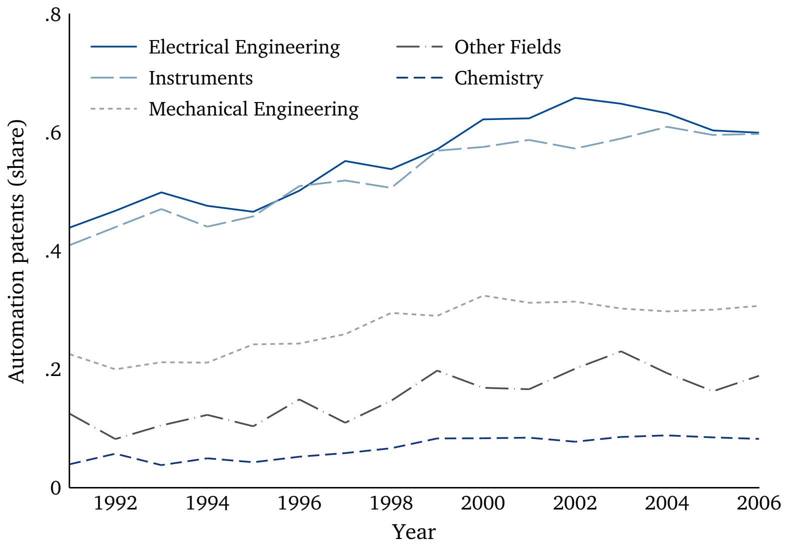
B.2 Automation Patents

Figure B-7: Number of automation and non-automation patents by year



Notes: The figure illustrates the annual number of patent applications filed at the European Patent Office with at least one inventor located in one of the sample regions and with a priority date between 1991 and 2006. Patents are classified into automation and non-automation according to the method outlined in Appendix D.2. Source: PATSTAT, own calculations.

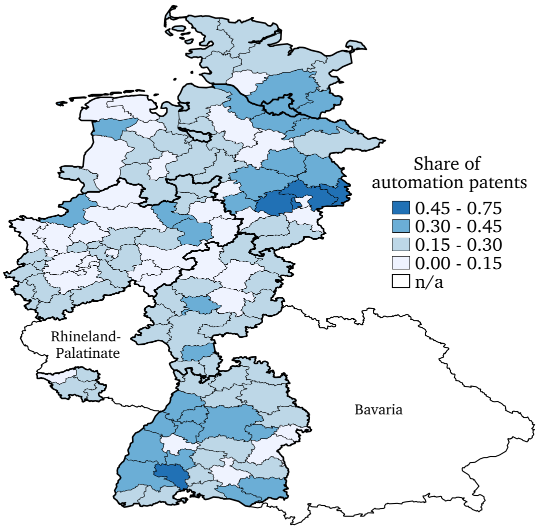
Figure B-8: Share of automation patents by technology area



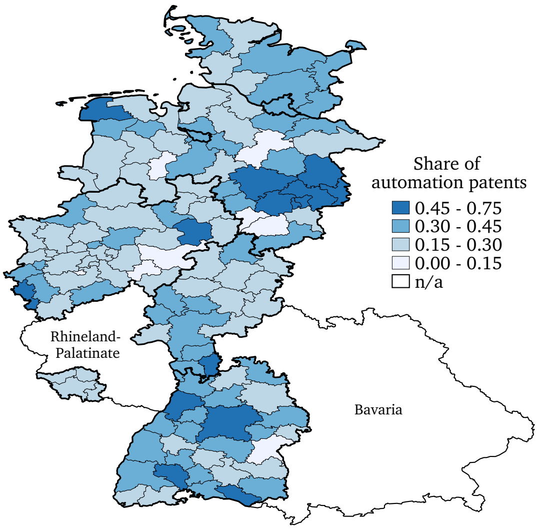
Notes: The figure illustrates the annual share of automation patents by technology area. For this purpose, we map the patents' International Patent Classification (IPC) codes to five main technology areas (Schmoch, 2008). Based on all patent applications filed at the European Patent Office with at least one inventor located in one of the sample regions and with a priority date between 1991 and 2006. Patents are classified into automation and non-automation according to the method outlined in Appendix D.2. Source: PATSTAT, own calculations.

Figure B-9: Regional share of automation patents in pre- and allocation period

(a) Pre-allocation period



(b) Allocation period



Notes: The figure illustrates the regional share of automation patents in the pre-allocation and the allocation period. Based on all patent applications filed at the European Patent Office with at least one inventor located in one of the sample regions and with a priority date between 1991 and 2006. Patents are classified into automation and non-automation according to the method outlined in Appendix D.2. Source: PATSTAT, own calculations.

B.3 Migration and Residential Growth Effects

Table B-2: Effect of labor inflows on migration during the allocation period

Dep. Var.:	Inflow rate$_{pop}$		Outflow rate$_{pop}$	Residential growth rate
	West Germans	East Germans	All (ethnic) Germans	All (ethnic) Germans
Inflow rate$_{pop}$	0.267	−0.208	−1.083*	1.181***
	(0.586)	(0.127)	(0.575)	(0.318)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1256	1256	1256	1256
R-squared	0.634	0.764	0.766	0.429
Within R-squared	0.000	0.003	0.004	0.006
Dep. Var. mean	0.002	0.002	-0.001	0.002

Notes: The table presents the results of linear regression analyses. The dependent variable is the annual (population-based) inflow rate of West Germans (column 1) and East Germans (column 2), the annual (population-based) outflow rate of all Germans, including ethnic Germans, (column 3), and the annual residential growth rate of all Germans, including ethnic Germans (column 4). The dependent variable is the percentage change in the annual growth rate of all residential. The treatment variable is continuous: the annual inflow rate, modified by measuring the annual inflow of ethnic Germans relative to the regional population in that year. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

B.4 Labor Market Effects

Table B-3: Effect of labor inflows on wage growth during the allocation period

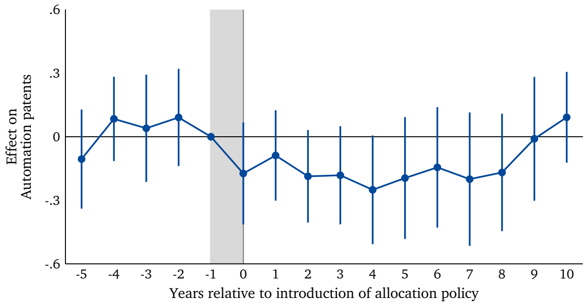
	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Wage growth					
	Average	Occ I	Occ II	Occ III	Occ IV	Occ V
Inflow rate$_{\text{total workforce}}$	−0.059	−0.321	−0.302	0.194	0.040	0.053
	(0.336)	(0.630)	(0.610)	(0.596)	(0.941)	(0.608)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	626	626	626	626	626	626
R-squared	0.496	0.206	0.144	0.292	0.188	0.327
Within R-squared	0.000	0.001	0.000	0.000	0.000	0.000

Notes: The table presents the results of linear regression analyses. The dependent variable is the change in the log wage averaged over all occupation groups in column (1), in occupation group I (farmers, laborers, and transport workers) in column (2), in occupation group II (operatives, craft workers) in column (3), in occupation group III (service workers) in column (4), in log wages in occupation group IV (managers, sales workers) in column (5), and in occupation group V (professional and technical workers) in column (6). The wage outcomes refer to individuals who were already present in 1995 in the IAB data. Imputation of top censored wages follows Gartner (2004). The treatment variable is continuous: the inflow rate during the allocation period interacted with a dummy (Allocation$_{t\geq 0}$) that indicates the allocation period. The inflow rate is modified by measuring the annual inflow of ethnic Germans relative to the regional workforce in that year. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

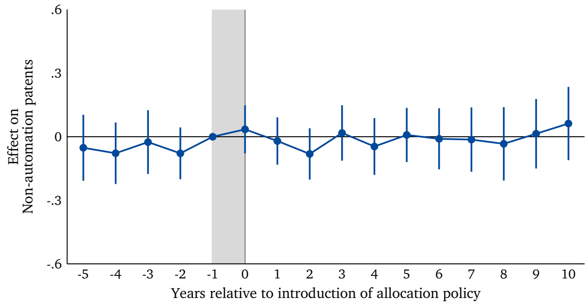
B.5 Main Results

Figure B-10: Effect of labor inflows on (non-)automation patents – event study with binary treatment

(a) Dep. Var.: Automation patents

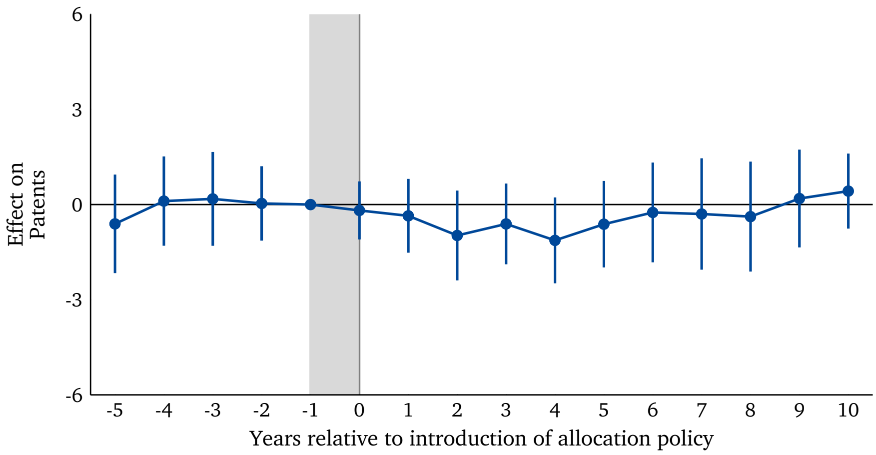


(b) Dep. Var.: Non-automation patents



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on Poisson pseudo-maximum-likelihood regressions as specified in Equation 1. The dependent variable in panel (a) is the count of automation patents. The dependent variable in panel (b) is the count of non-automation patents. In both panels, the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with leads and lags. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The coefficients can be interpreted as semi-elasticities. Region and time fixed effects included. Standard errors clustered at the regional level. The coefficients correspond to those reported in Appendix Table B-7.

Figure B-11: Effect of labor supply on patents – event study with continuous treatment



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on a Poisson pseudo-maximum-likelihood regression as specified in Equation 1. The dependent variable is the count of patents. The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The coefficients can be interpreted as semi-elasticities. Region and time fixed effects included. Standard errors clustered at the regional level. The coefficients correspond to those reported in Appendix Table B-6.

Table B-4: Effect of labor inflows on the share of automation patents – event study with continuous treatment

Dep. Var.:	Automation patents / patents (1)	(2)
Cum. inflow rate		
× Allocation$_{t-5}$	0.111	−0.048
	(0.230)	(0.181)
× Allocation$_{t-4}$	0.286	0.271
	(0.184)	(0.187)
× Allocation$_{t-3}$	−0.006	−0.023
	(0.197)	(0.203)
× Allocation$_{t-2}$	0.239	0.221
	(0.158)	(0.159)
× Allocation$_{t-1}$	0.000	0.000
	(.)	(.)
× Allocation$_{t}$	−0.026	−0.012
	(0.211)	(0.211)
× Allocation$_{t+1}$	0.111	0.132
	(0.192)	(0.192)
× Allocation$_{t+2}$	−0.256	−0.254
	(0.238)	(0.238)
× Allocation$_{t+3}$	−0.476*	−0.497*
	(0.253)	(0.253)
× Allocation$_{t+4}$	−0.483*	−0.546**
	(0.266)	(0.261)
× Allocation$_{t+5}$	−0.470*	−0.427
	(0.282)	(0.273)
× Allocation$_{t+6}$	−0.393*	−0.359
	(0.226)	(0.218)
× Allocation$_{t+7}$	−0.224	−0.199
	(0.230)	(0.224)
× Allocation$_{t+8}$	−0.342	−0.329
	(0.207)	(0.206)
× Allocation$_{t+9}$	−0.205	−0.215
	(0.330)	(0.330)
× Allocation$_{t+10}$	−0.112	−0.146
	(0.217)	(0.214)
Region fixed effects	No	Yes
Year fixed effects	Yes	Yes
Observations	2029	2029
R-squared	0.118	0.688
Within R-squared	0.017	0.034
Dep. Var. mean	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 1. The dependent variable is the share of automation patents. The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-5: Effect of labor inflows on the share of automation patents – event study with binary treatment

Dep. Var.:	Automation patents / patents	
	(1)	(2)
1(Cum. inflow rate>p50)		
× Allocation$_{t-5}$	−0.007	−0.014
	(0.021)	(0.020)
× Allocation$_{t-4}$	0.032	0.031
	(0.021)	(0.020)
× Allocation$_{t-3}$	0.005	0.003
	(0.023)	(0.023)
× Allocation$_{t-2}$	0.015	0.013
	(0.019)	(0.020)
× Allocation$_{t-1}$	0.000	0.000
	(.)	(.)
× Allocation$_{t}$	−0.015	−0.015
	(0.019)	(0.019)
× Allocation$_{t+1}$	0.017	0.018
	(0.020)	(0.020)
× Allocation$_{t+2}$	−0.031	−0.031
	(0.022)	(0.022)
× Allocation$_{t+3}$	−0.045*	−0.047*
	(0.026)	(0.026)
× Allocation$_{t+4}$	−0.064**	−0.068***
	(0.025)	(0.025)
× Allocation$_{t+5}$	−0.061***	−0.059**
	(0.023)	(0.023)
× Allocation$_{t+6}$	−0.042*	−0.041*
	(0.023)	(0.023)
× Allocation$_{t+7}$	−0.042**	−0.042**
	(0.021)	(0.021)
× Allocation$_{t+8}$	−0.038*	−0.038*
	(0.022)	(0.022)
× Allocation$_{t+9}$	−0.026	−0.027
	(0.033)	(0.033)
× Allocation$_{t+10}$	−0.010	−0.021
	(0.027)	(0.026)
Region fixed effects	No	Yes
Year fixed effects	Yes	Yes
Observations	2029	2029
R-squared	0.119	0.690
Within R-squared	0.018	0.040
Dep. Var. mean	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 1. The dependent variable is the share of automation patents. The treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-6: Effect of labor inflows on (non-)automation patents – event study with continuous treatment

Dep. Var.:	(1) Automation patents	(2) Non-automation patents	(3) All patents
Cum. inflow rate			
× Allocation$_{t-5}$	−1.775	−0.198	−0.608
	(1.339)	(0.809)	(0.793)
× Allocation$_{t-4}$	0.017	0.232	0.109
	(1.164)	(0.768)	(0.718)
× Allocation$_{t-3}$	−0.331	0.391	0.178
	(1.282)	(0.857)	(0.755)
× Allocation$_{t-2}$	0.687	−0.128	0.036
	(1.036)	(0.671)	(0.598)
× Allocation$_{t-1}$	0.000	0.000	0.000
	(.)	(.)	(.)
× Allocation$_{t}$	−1.552*	0.280	−0.185
	(0.906)	(0.576)	(0.467)
× Allocation$_{t+1}$	−1.745	0.116	−0.355
	(1.122)	(0.587)	(0.594)
× Allocation$_{t+2}$	−2.544*	−0.406	−0.975
	(1.328)	(0.669)	(0.722)
× Allocation$_{t+3}$	−3.173**	0.429	−0.611
	(1.250)	(0.609)	(0.649)
× Allocation$_{t+4}$	−3.647***	−0.084	−1.129
	(1.415)	(0.638)	(0.691)
× Allocation$_{t+5}$	−2.483*	0.120	−0.619
	(1.458)	(0.717)	(0.695)
× Allocation$_{t+6}$	−2.045	0.468	−0.249
	(1.424)	(0.712)	(0.802)
× Allocation$_{t+7}$	−2.222	0.479	−0.297
	(1.776)	(0.764)	(0.896)
× Allocation$_{t+8}$	−2.480*	0.479	−0.380
	(1.421)	(0.845)	(0.884)
× Allocation$_{t+9}$	−0.865	0.547	0.190
	(1.415)	(0.830)	(0.786)
× Allocation$_{t+10}$	−0.194	0.579	0.424
	(0.789)	(0.718)	(0.604)
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	2032	2032	2032
χ^2	84.686	96.602	106.238
Dep. Var. mean	23.744	48.471	72.215

Notes: The table presents the results of Poisson pseudo-maximum-likelihood regression analyses as defined in Equation 1. The dependent variable is the count of automation patents in column (1), the count of non-automation patents in column (2), and the count of all patents in column (3). The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The coefficients can be interpreted as semi-elasticities. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

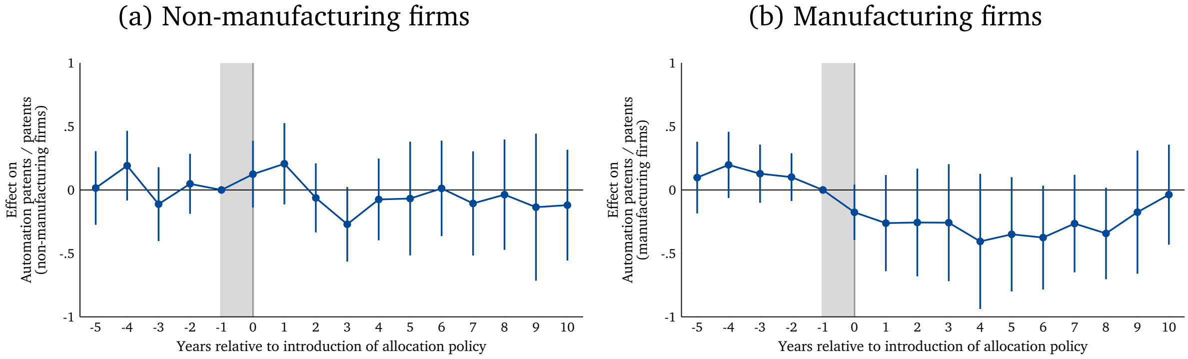
Table B-7: Effect of labor inflows on (non-)automation patents – event study with binary treatment

Dep. Var.:	(1) Automation patents	(2) Non-automation patents	(3) All patents
1(Cum. inflow rate>p50)			
× $\text{Allocation}_{\text{t-5}}$	−0.105	−0.052	−0.062
	(0.119)	(0.079)	(0.074)
× $\text{Allocation}_{\text{t-4}}$	0.084	−0.078	−0.042
	(0.102)	(0.074)	(0.064)
× $\text{Allocation}_{\text{t-3}}$	0.040	−0.025	−0.009
	(0.129)	(0.077)	(0.071)
× $\text{Allocation}_{\text{t-2}}$	0.091	−0.078	−0.039
	(0.117)	(0.062)	(0.057)
× $\text{Allocation}_{\text{t-1}}$	0.000	0.000	0.000
	(.)	(.)	(.)
× $\text{Allocation}_{\text{t}}$	−0.173	0.035	−0.020
	(0.123)	(0.058)	(0.061)
× $\text{Allocation}_{\text{t+1}}$	−0.088	−0.020	−0.038
	(0.109)	(0.057)	(0.057)
× $\text{Allocation}_{\text{t+2}}$	−0.187*	−0.081	−0.116*
	(0.111)	(0.062)	(0.060)
× $\text{Allocation}_{\text{t+3}}$	−0.182	0.018	−0.051
	(0.118)	(0.067)	(0.065)
× $\text{Allocation}_{\text{t+4}}$	−0.251*	−0.046	−0.118*
	(0.131)	(0.068)	(0.068)
× $\text{Allocation}_{\text{t+5}}$	−0.195	0.008	−0.063
	(0.147)	(0.065)	(0.070)
× $\text{Allocation}_{\text{t+6}}$	−0.144	−0.010	−0.059
	(0.145)	(0.074)	(0.080)
× $\text{Allocation}_{\text{t+7}}$	−0.200	−0.013	−0.077
	(0.161)	(0.077)	(0.084)
× $\text{Allocation}_{\text{t+8}}$	−0.169	−0.033	−0.082
	(0.141)	(0.088)	(0.088)
× $\text{Allocation}_{\text{t+9}}$	−0.010	0.014	0.005
	(0.149)	(0.084)	(0.081)
× $\text{Allocation}_{\text{t+10}}$	0.092	0.062	0.071
	(0.109)	(0.088)	(0.074)
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	2032	2032	2032
χ^2	82.544	99.427	112.743
Dep. Var. mean	23.744	48.471	72.215

Notes: The table presents the results of Poisson pseudo-maximum-likelihood regression analyses as defined in Equation 1. The dependent variable is the count of automation patents in column (1), the count of non-automation patents in column (2), and the count of all patents in column (3). The treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with leads and lags. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The coefficients can be interpreted as semi-elasticities. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

B.6 Heterogeneity Results

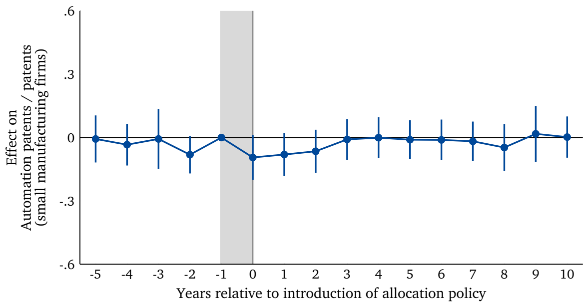
Figure B-12: Effect of labor inflows on the share of automation patents by (non-)manufacturing firms – event study with continuous treatment



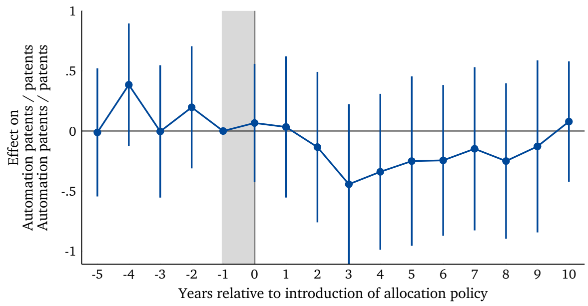
Notes: The figure displays event study estimates and the 95 percent confidence intervals based on linear regressions as specified in Equation 1. The dependent variable in panel (a) is the share of automation patents held by non-manufacturing firms. The dependent variable in panel (b) is the share of automation patents held by manufacturing firms. The classification in (non-)manufacturing is based on the respective firm's major sector code, as reported in ORBIS IP. In both panels, the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. Region and time fixed effects included. Standard errors clustered at the regional level. The coefficients correspond to those reported in Appendix Table B-8.

Figure B-13: Effect of labor inflows on the share of automation patents by firm and patent characteristics – event study with continuous treatment

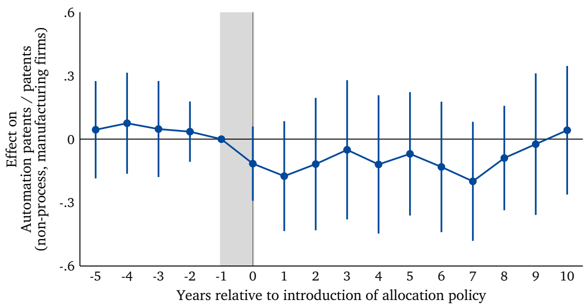
(a) Small firms



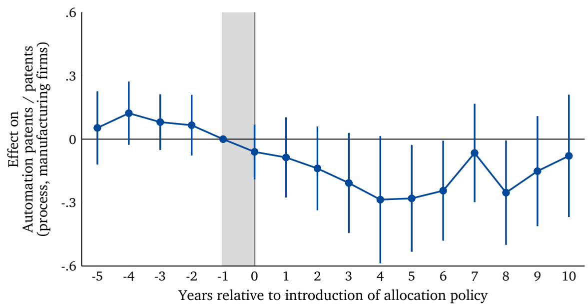
(b) Large firms



(c) Product patents



(d) Process patents



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on linear regressions as specified in Equation 1. The dependent variable is the share of automation patents held by small manufacturing firms in panel (a), the share of automation patents held by large manufacturing firms in panel (b), the share of product-related automation patents held by manufacturing firms in panel (c), and the share of process-related automation patents held by manufacturing firms in panel (d). Firm size follows the classification of the European Commission for small and medium-sized enterprises (max. 249 employees and max. 50 million EUR annual turnover). The distinction between process and product patents is text-based and follows Ganglmair and Reimers (2019). In all panels, the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. Region and time fixed effects included. Standard errors clustered at the regional level. The coefficients correspond to those reported in Appendix Table B-9.

Table B-8: Effect of labor inflows on the share of automation patents by (non-)manufacturing firms – event study with continuous treatment

Dep. Var.:	(1) Automation patents / patents	(2)
	Non-manufacturing firms	Manufacturing firms
Cum. inflow rate		
× Allocation$_{t-5}$	0.015	0.098
	(0.147)	(0.143)
× Allocation$_{t-4}$	0.192	0.198
	(0.139)	(0.132)
× Allocation$_{t-3}$	−0.112	0.128
	(0.147)	(0.116)
× Allocation$_{t-2}$	0.048	0.101
	(0.119)	(0.095)
× Allocation$_{t-1}$	0.000	0.000
	(.)	(.)
× Allocation$_{t}$	0.124	−0.176
	(0.133)	(0.110)
× Allocation$_{t+1}$	0.207	−0.261
	(0.162)	(0.192)
× Allocation$_{t+2}$	−0.062	−0.256
	(0.137)	(0.215)
× Allocation$_{t+3}$	−0.270*	−0.258
	(0.148)	(0.233)
× Allocation$_{t+4}$	−0.075	−0.406
	(0.163)	(0.269)
× Allocation$_{t+5}$	−0.068	−0.349
	(0.226)	(0.228)
× Allocation$_{t+6}$	0.012	−0.375*
	(0.190)	(0.207)
× Allocation$_{t+7}$	−0.106	−0.265
	(0.208)	(0.194)
× Allocation$_{t+8}$	−0.037	−0.343*
	(0.220)	(0.182)
× Allocation$_{t+9}$	−0.136	−0.175
	(0.293)	(0.245)
× Allocation$_{t+10}$	−0.120	−0.037
	(0.221)	(0.199)
Region fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	2029	2029
R-squared	0.406	0.729
Within R-squared	0.021	0.041
Dep. Var. mean	0.092	0.135

Notes: The table presents the results of linear regression analyses as defined in Equation 1. The dependent variable in column (1) is the share of automation patents held by non-manufacturing firms. The dependent variable in column (2) is the share of automation patents held by manufacturing firms. The classification in (non-)manufacturing is based on the respective firm's major sector code, as reported in ORBIS IP. The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-9: Effect of labor inflows on the share of automation patents by firm and patent characteristics – event study with continuous treatment

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents (manuf. firms)			
	Small firms	Large firms	Non-process pat.	Process pat.
Cum. inflow rate				
× Allocation$_{t-5}$	−0.007	0.111	0.045	0.053
	(0.056)	(0.131)	(0.116)	(0.088)
× Allocation$_{t-4}$	−0.034	0.236*	0.075	0.123
	(0.050)	(0.123)	(0.121)	(0.076)
× Allocation$_{t-3}$	−0.006	0.137	0.048	0.081
	(0.072)	(0.102)	(0.115)	(0.067)
× Allocation$_{t-2}$	−0.081*	0.195**	0.035	0.066
	(0.045)	(0.091)	(0.072)	(0.073)
× Allocation$_{t-1}$	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
× Allocation$_{t}$	−0.094*	−0.060	−0.116	−0.060
	(0.054)	(0.106)	(0.089)	(0.065)
× Allocation$_{t+1}$	−0.080	−0.143	−0.175	−0.086
	(0.052)	(0.184)	(0.131)	(0.096)
× Allocation$_{t+2}$	−0.065	−0.177	−0.118	−0.138
	(0.052)	(0.204)	(0.158)	(0.100)
× Allocation$_{t+3}$	−0.009	−0.232	−0.050	−0.207*
	(0.049)	(0.237)	(0.166)	(0.120)
× Allocation$_{t+4}$	−0.001	−0.350	−0.120	−0.286*
	(0.049)	(0.274)	(0.165)	(0.152)
× Allocation$_{t+5}$	−0.010	−0.287	−0.069	−0.280**
	(0.047)	(0.235)	(0.148)	(0.128)
× Allocation$_{t+6}$	−0.011	−0.334	−0.131	−0.244**
	(0.049)	(0.217)	(0.156)	(0.120)
× Allocation$_{t+7}$	−0.018	−0.208	−0.199	−0.065
	(0.047)	(0.198)	(0.142)	(0.118)
× Allocation$_{t+8}$	−0.047	−0.275	−0.089	−0.253**
	(0.056)	(0.207)	(0.125)	(0.125)
× Allocation$_{t+9}$	0.018	−0.169	−0.024	−0.151
	(0.067)	(0.233)	(0.169)	(0.132)
× Allocation$_{t+10}$	0.002	−0.023	0.042	−0.079
	(0.049)	(0.205)	(0.154)	(0.146)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.207	0.750	0.560	0.659
Within R-squared	0.015	0.043	0.020	0.043
Dep. Var. mean	0.016	0.117	0.072	0.064

Notes: The table presents the results of linear regression analyses as defined in Equation 1. The dependent variable is the share of automation patents held by small manufacturing firms in column (1), the share of automation patents held by large manufacturing firms in column (2), the share of product-related automation patents held by manufacturing firms in column (3), and the share of process-related automation patents held by manufacturing firms in column (4). Firm size follows the classification of the European Commission for small and medium-sized enterprises (max. 249 employees and max. 50 million EUR annual turnover). The distinction between process and product patents is text-based and follows Ganglmair and Reimers (2019). The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-10: Effect of labor inflows on the share of automation patents by firm and patent characteristics – DiD with continuous treatment

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents (non-manuf. firms)			
	Small firms	Large firms	Non-process pat.	Process pat.
Cum. inflow rate × Allocation$_{t\geq 0}$	−0.116	−0.004	−0.037	−0.014
	(0.151)	(0.318)	(0.074)	(0.084)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1904	1939	2029	2029
R-squared	0.211	0.460	0.257	0.412
Within R-squared	0.000	0.000	0.003	0.000
Dep. Var. mean	0.204	0.215	0.048	0.044

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents held by small non-manufacturing firms in column (1), the share of automation patents held by large non-manufacturing firms in column (2), the share of product-related automation patents held by non-manufacturing firms in column (3), and the share of process-related automation patents held by non-manufacturing firms in column (4). Firm size follows the classification of the European Commission for small and medium-sized enterprises (max. 249 employees and max. 50 million EUR annual turnover). The distinction between process and product patents is text-based and follows Ganglmair and Reimers (2019). The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-11: Effect of labor inflows on the share of automation patents by firm and patent characteristics – DiD with binary treatment

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents (manuf. firms)			
	Small firms	Large firms	Non-process pat.	Process pat.
1(Cum. inflow rate>p50) × Allocation$_{t\geq 0}$	−0.001	−0.026	−0.009	−0.019**
	(0.003)	(0.019)	(0.013)	(0.009)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.195	0.742	0.552	0.649
Within R-squared	0.000	0.013	0.003	0.016
Dep. Var. mean	0.016	0.117	0.072	0.064

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents held by small manufacturing firms in column (1), the share of automation patents held by large manufacturing firms in column (2), the share of product-related automation patents held by manufacturing firms in column (3), and the share of process-related automation patents held by manufacturing firms in column (4). Firm size follows the classification of the European Commission for small and medium-sized enterprises (max. 249 employees and max. 50 million EUR annual turnover). The distinction between process and product patents is text-based and follows Ganglmair and Reimers (2019). The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-12: Effect of labor inflows on the share of automation patents in slack and tight labor market regions – event study with continuous treatment

	(1)	(2)
Dep. Var.:	Automation patents / patents (manufacturing firms)	
Labor markets:	Slack	Tight
Cum. inflow rate		
× Allocation$_{t-5}$	0.429*	−0.044
	(0.223)	(0.171)
× Allocation$_{t-4}$	0.268	0.157
	(0.268)	(0.153)
× Allocation$_{t-3}$	0.172	0.084
	(0.300)	(0.115)
× Allocation$_{t-2}$	−0.165	0.100
	(0.202)	(0.096)
× Allocation$_{t-1}$	0.000	0.000
	(.)	(.)
× Allocation$_{t}$	−0.032	−0.304**
	(0.207)	(0.135)
× Allocation$_{t+1}$	−0.282	−0.388
	(0.216)	(0.240)
× Allocation$_{t+2}$	−0.093	−0.511*
	(0.203)	(0.271)
× Allocation$_{t+3}$	0.021	−0.489*
	(0.234)	(0.279)
× Allocation$_{t+4}$	0.021	−0.622*
	(0.329)	(0.322)
× Allocation$_{t+5}$	0.091	−0.639**
	(0.331)	(0.292)
× Allocation$_{t+6}$	−0.134	−0.574**
	(0.295)	(0.289)
× Allocation$_{t+7}$	0.105	−0.436
	(0.333)	(0.275)
× Allocation$_{t+8}$	−0.004	−0.644***
	(0.288)	(0.241)
× Allocation$_{t+9}$	0.221	−0.491
	(0.298)	(0.343)
× Allocation$_{t+10}$	0.116	−0.224
	(0.338)	(0.246)
Region fixed effects	Yes	
Year fixed effects	Yes	
Observations	2029	
R-squared	0.734	
Within R-squared	0.060	
Dep. Var. mean	0.135	

Notes: The table presents the results of linear regression analyses as defined in Equation 1. The dependent variable is the share of automation patents held by manufacturing firms. The classification in (non-)manufacturing is based on the respective firm's major sector code, as reported in ORBIS IP. The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. We introduce further interactions with a binary variable that indicates regions with a tight labor market: an unemployment rate in the last pre-allocation year that is higher than the median rate. The estimates in column (1) present the baseline coefficients (with the additional interaction being zero). The estimates in column (2) present the sum of the baseline and interaction coefficients indicating regions with a tight labor market. The coefficients for $l = -1$ are normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-13: Effect of labor inflows on the share of automation patents in slack and tight labor market regions – DiD with binary treatment

	(1)	(2)	(3)
Dep. Var.:	Automation patents / patents (manuf. firms)		
Labor markets:	Slack	Tight	All
1(Cum. inflow rate>p50) × Allocation$_{t\geq0}$	0.025	−0.068**	0.024
	(0.016)	(0.027)	(0.016)
1(Cum. inflow rate>p50) × Allocation$_{t\geq0}$ × Tight labor market			−0.091***
			(0.030)
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	1021	1008	2029
R-squared	0.690	0.762	0.728
Within R-squared	0.006	0.077	0.039
Dep. Var. mean	0.135	0.135	0.135

Notes: The table presents the results of linear regression analyses as defined in Equation 1. The dependent variable is the share of automation patents held by manufacturing firms. The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. We introduce further interactions with a binary variable that indicates regions with a tight labor market: an unemployment rate in the last pre-allocation year that is higher than the median rate. The estimates in column (1) present the baseline coefficients (with the additional interaction being zero). The estimates in column (2) present the sum of the baseline and interaction coefficients indicating regions with a tight labor market. The coefficients for $l = -1$ are normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B-14: Effect of labor inflows on the share of automation patents in high and low wage regions – DiD with continuous treatment

	(1)	(2)	(3)
Dep. Var.:	Automation patents / patents (manuf. firms)		
Wage level:	Low	High	All
Cum. inflow rate × Allocation$_{t\geq0}$	−0.204*	−0.490*	−0.211*
	(0.122)	(0.265)	(0.114)
Cum. inflow rate × Allocation$_{t\geq0}$ × High wage level			−0.275
			(0.276)
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	1021	1008	2029
R-squared	0.687	0.750	0.724
Within R-squared	0.008	0.036	0.023
Dep. Var. mean	0.131	0.140	0.135

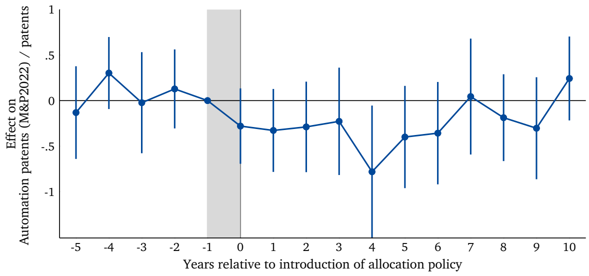
Notes: The table presents the results of linear regression analyses as defined in Equation 2. In all columns, the dependent variable is the share of automation patents held by manufacturing firms, and the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. The sample in column (1) consists of "low wage" regions with a low-skilled wage level in the last pre-allocation year that is lower or equal to the median rate. The sample in column (2) consists of "high wage" regions with a low-skilled wage level in the last pre-allocation year that is above the median rate. In column (3), we use the full sample but introduce a further interaction of the treatment variable with a dummy that indicates "high wage" regions. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

C Robustness Checks

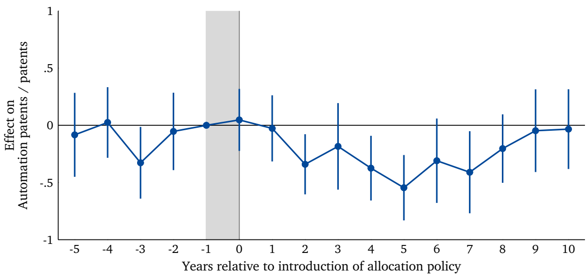
C.1 Alternative Measures of Automation Innovation

Figure C-1: Effect of labor inflows on the share of automation patents – event study with continuous treatment – alternative sets of automation keywords

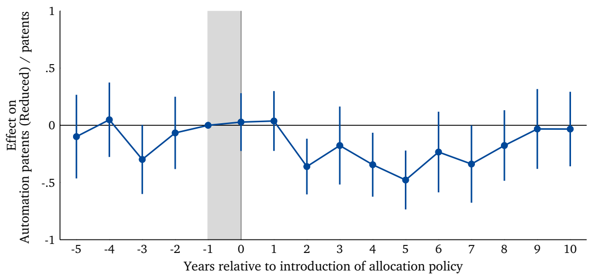
(a) Mann and Püttmann (2023) classification



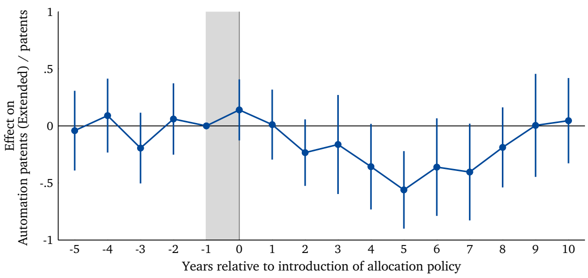
(b) Set of automation keywords



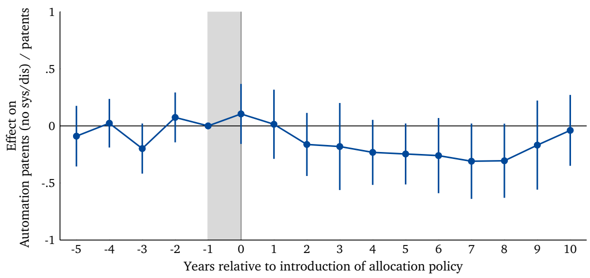
(c) Reduced set of automation keywords



(d) Expanded set of automation keywords



(e) Set w/o "system" and "display"

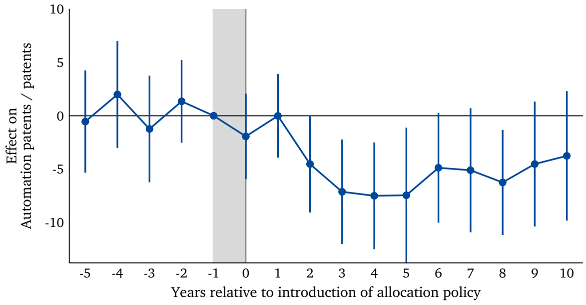


Notes: The figure displays event study estimates and the 95 percent confidence intervals based on linear regressions as specified in Equation 1. In all panels, the dependent variable is the share of automation patents, and the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The estimates in each panel are obtained using a different classification of automation patents. Panel (a) uses the classification used by Mann and Püttmann (2023). Note that these estimates are based on the subsample of patents with a US patent family member. Panels (b)-(e) use a keywords-based classification. The standard set includes the following keywords: "automat", "execut", "detect", "input", "system", "display", "output", and "inform". The reduced set excludes "detect" and "display". The expanded set further includes "signal" and "sensor". The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. Region and time fixed effects included. Standard errors clustered at the regional level.

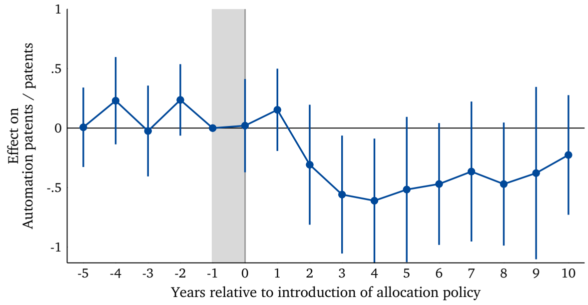
C.2 Alternative Operationalizations of the Cumulative Inflow Rate

Figure C-2: Effect of labor inflows on the share of automation patents – event study with continuous treatment – different inflow rate operationalizations

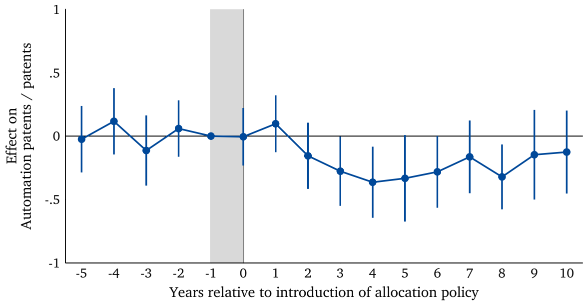
(a) Inflow rate based on population



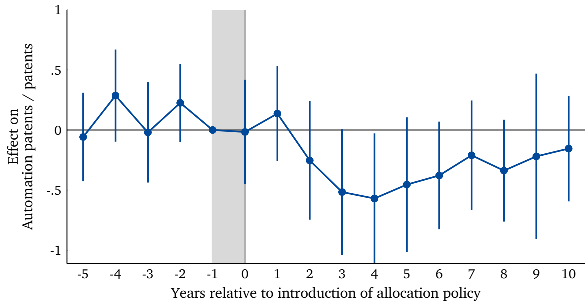
(b) Inflows excl. high-skilled migrants



(c) Inflow rate based on education-based low-skilled workforce



(d) Inflow rate winsorized



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on linear regressions as specified in Equation 1. In all panels, the dependent variable is the share of automation patents, and the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The estimates in each panel are obtained using a different operationalization of the inflow rate. In Panel (a), the rate's denominator is the population instead of the low-skilled workforce. In Panel (b), the rate's numerator excludes high-skilled ethnic Germans. In Panel (c), the rate's denominator is the low-skilled workforce as determined by education. In Panel (d), the inflow rate is winsorized at the 5th and 95th percentile. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. Region and time fixed effects included. Standard errors clustered at the regional level.

Table C-1: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – inflow rate based on population

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate$_{pop}$ × Allocation$_{t\geq0}$	−5.890***	−4.367***		
	(1.518)	(1.141)		
1(Cum. inflow rate$_{pop}$>p50) × Allocation$_{t\geq0}$			−0.050***	−0.048***
			(0.017)	(0.016)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.117	0.681	0.112	0.683
Within R-squared	0.016	0.013	0.011	0.020
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. The inflow rate's denominator is the population instead of the low-skilled workforce. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-2: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – inflows excl. high-skilled migrants

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate$_{\text{no high-skilled}}$ × Allocation$_{t\geq0}$	−0.456***	−0.394**		
	(0.173)	(0.163)		
1(Cum. inflow rate$_{\text{no high-skilled}}$>p50) × Allocation$_{t\geq0}$			−0.049***	−0.047***
			(0.017)	(0.017)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.112	0.682	0.112	0.683
Within R-squared	0.011	0.017	0.010	0.019
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. The inflow rate's numerator excludes high-skilled ethnic Germans. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-3: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – inflow rate based on education-based low-skilled workforce

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate$_{\text{education}}$ × Allocation$_{t\geq0}$	−0.266***	−0.193***		
	(0.086)	(0.072)		
1(Cum. inflow rate$_{\text{education}}$>p50) × Allocation$_{t\geq0}$			−0.046***	−0.039**
			(0.016)	(0.016)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.115	0.680	0.111	0.681
Within R-squared	0.013	0.009	0.009	0.014
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. The inflow rate's denominator is the low-skilled workforce as determined by education. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-4: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – winsorized inflow rate

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate$^{\text{winsorized}}$ × Allocation$_{t\geq0}$	−0.392**	−0.335**		
	(0.164)	(0.157)		
1(Cum. inflow rate$^{\text{winsorized}}$>p50) × Allocation$_{t\geq0}$			−0.044**	−0.042**
			(0.018)	(0.018)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.110	0.681	0.110	0.682
Within R-squared	0.008	0.012	0.008	0.015
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. The inflow rate is winsorized at the 5th and 95th percentile. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-5: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – alternative reference year of low-skilled workforce

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate$_{1991}$ × Allocation$_{t\geq 0}$	−0.307**			
	(0.141)			
1(Cum. inflow rate$_{1991}$>p50) × Allocation$_{t\geq 0}$		−0.044**		
		(0.017)		
Cum. inflow rate$_{t-1}$ × Allocation$_{t\geq 0}$			−0.329**	
			(0.149)	
1(Cum. inflow rate$_{t-1}$>p50) × Allocation$_{t\geq 0}$				−0.040**
				(0.018)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.681	0.682	0.680	0.681
Within R-squared	0.012	0.016	0.011	0.014
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

C.3 Accounting for Pre-allocation Labor Inflows

Table C-6: Effect of labor supply on the share of automation innovation – DiD model with continuous and binary treatment – pre-allocation inflow controls

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate × $\text{Allocation}_{t\geq0}$	−0.343**	−0.287*		
	(0.160)	(0.167)		
Cum. inflow rate_{pre} × $\text{Allocation}_{t\geq0}$	−0.004	−0.053		
	(0.083)	(0.076)		
1(Cum. inflow rate>p50) × $\text{Allocation}_{t\geq0}$			−0.042*	−0.041*
			(0.021)	(0.022)
1(Cum. inflow rate_{pre}>p50) × $\text{Allocation}_{t\geq0}$			−0.003	−0.002
			(0.018)	(0.019)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.131	0.681	0.137	0.682
Within R-squared	0.031	0.012	0.038	0.015
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. All specifications further include the corresponding treatment variables based on the *pre-allocation* cumulative inflow rate. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-7: Effect of labor supply on the share of automation innovation – DiD model with continuous – regions with high pre-allocation inflows excluded

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Regions below the ... pctile of Cum. inflow rate_{pre}:	90 pctile	75 pctile	25 pctile	10 pctile
Cum. inflow rate × $\text{Allocation}_{t\geq0}$	−0.478***	−0.627***	−0.584**	−0.814**
	(0.172)	(0.210)	(0.249)	(0.320)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1821	1518	496	192
R-squared	0.688	0.704	0.817	0.851
Within R-squared	0.020	0.029	0.018	0.032
Dep. Var. mean	0.272	0.278	0.289	0.248

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In all columns, the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. Subsamples include only regions below the indicated percentile of the *pre-allocation* cumulative inflow rate. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-8: Effect of labor supply on the share of automation innovation – DiD model with continuous and binary treatment – regions in proximity to entry port excluded

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Regions in distance from entry port:	>150km	>300km	>150km	>300km
Cum. inflow rate × $\text{Allocation}_{t\geq0}$	−0.336*	−0.618**		
	(0.197)	(0.269)		
1(Cum. inflow rate>p50) × $\text{Allocation}_{t\geq0}$			−0.042**	−0.045*
			(0.021)	(0.025)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1485	621	1485	621
R-squared	0.645	0.616	0.647	0.613
Within R-squared	0.011	0.034	0.015	0.029
Dep. Var. mean	0.266	0.275	0.266	0.275

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. Subsamples include regions more than 150 (300) km away from the entry port (Friedland reception center). The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

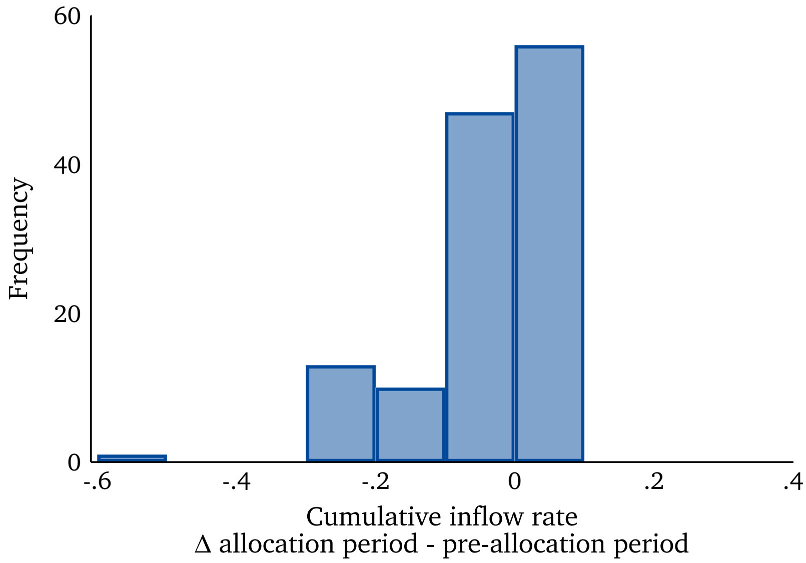
Table C-9: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – Gifhorn regions excluded

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate × $\text{Allocation}_{t\geq0}$	−0.264	−0.354**		
	(0.179)	(0.152)		
1(Cum. inflow rate>p50) × $\text{Allocation}_{t\geq0}$			−0.036*	−0.043**
			(0.019)	(0.018)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	No	Yes	No	Yes
Observations	1933	1933	1933	1933
R-squared	0.090	0.687	0.090	0.688
Within R-squared	0.090	0.014	0.090	0.017
Dep. Var. mean	0.268	0.268	0.268	0.268

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. Subsamples exclude the so-called Gifhorn regions. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

C.4 Accounting for Relative Change in Labor Inflows

Figure C-3: Distribution of the relative change in the region-specific cumulative inflow rate from the pre-allocation to the allocation period



Notes: The figure illustrates the distribution of the relative change in the region-specific cumulative inflow rate from the pre-allocation to the allocation period. The (pre-allocation) cumulative inflow rate is defined as the cumulative inflow of ethnic Germans in the allocation (pre-allocation) period divided by the average stock of the low-skilled workforce during the pre-allocation period. Observations are at the level of the region. Appendix Figure B-2 illustrates the geographical distribution of the cumulative inflow rate.

Table C-10: Effect of labor supply on the share of automation innovation – DiD model with continuous treatment – regions with a (highly) negative change in labor inflows excluded

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Regions with Δ allocation period - pre-allocation period:	> -0.2	> -0.1	> -0.05	> 0
Cum. inflow rate × Allocation$_{t\geq 0}$	−0.403**	−0.432**	−0.587***	−0.620**
	(0.161)	(0.181)	(0.201)	(0.291)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1805	1645	1326	894
R-squared	0.685	0.691	0.730	0.749
Within R-squared	0.017	0.019	0.032	0.030
Dep. Var. mean	0.277	0.278	0.277	0.292

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In all columns, the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. Subsamples include only regions with a change in labor inflows from the pre-allocation to the allocation period larger than the indicated threshold. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

C.5 Accounting for Other Concurrent Labor Inflows

Table C-11: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – other migration

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate × $\text{Allocation}_{t\geq0}$	−0.349**	−0.351**		
	(0.160)	(0.160)		
1(Cum. inflow rate>p50) × $\text{Allocation}_{t\geq0}$			−0.046**	−0.046**
			(0.019)	(0.020)
West German inflow rate	0.004	0.003	−0.004	−0.005
	(0.059)	(0.059)	(0.058)	(0.058)
East German inflow rate	0.015	0.019	0.011	0.016
	(0.195)	(0.197)	(0.178)	(0.179)
Foreign inflow rate		0.021		0.026
		(0.033)		(0.034)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1903	1903	1903	1903
R-squared	0.685	0.685	0.687	0.687
Within R-squared	0.013	0.013	0.017	0.017
Dep. Var. mean	0.272	0.272	0.272	0.272

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq0}$) that indicates whether the allocation policy was introduced in that year or earlier. Inflow rates of East Germans, West Germans, and foreigners based on administrative migration statistics from the Federal Statistical Office (*Wanderungsstatistik*). Information on foreign inflows is from INKAR online and from Glitz (2012). The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

C.6 Accounting for Differences in Industry and Skill Composition

Table C-12: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – industry composition trends

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate × Allocation$_{t\geq 0}$	−0.264**	−0.304**		
	(0.128)	(0.142)		
1(Cum. inflow rate>p50) × Allocation$_{t\geq 0}$			−0.033**	−0.042**
			(0.014)	(0.021)
Industry share: Agriculture (ext.)	−0.098		−0.114	
	(1.347)		(1.357)	
Industry share: Mining (ext.)	−0.690		−0.564	
	(0.585)		(0.562)	
Industry share: Manufacturing (ext.)	−0.511		−0.464	
	(0.506)		(0.507)	
Industry share: Energy/utilities (ext.)	−0.506		−0.588	
	(0.917)		(0.948)	
Industry share: Construction (ext.)	−0.756		−0.634	
	(0.621)		(0.601)	
Industry share: Services (ext.)	−0.338		−0.310	
	(0.381)		(0.387)	
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Industry share trends	No	Yes	No	Yes
Observations	1997	1997	1997	1997
R-squared	0.686	0.698	0.686	0.699
Within R-squared	0.014	0.052	0.016	0.054
Dep. Var. mean	0.267	0.267	0.267	0.267

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. To measure industry composition, we calculate for each region the share of employees working in one of six industry categories: Agriculture, Mining, Manufacturing, Energy and Utilities, Construction, and Services. In Columns (1) and (3), we extrapolate the industry share trends from the pre-allocation period (1990 to 1995) for the allocation period. In Columns (2) and (4), we take the industry shares averaged over the pre-allocation period and interact these with year dummies. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-13: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – skill composition trends

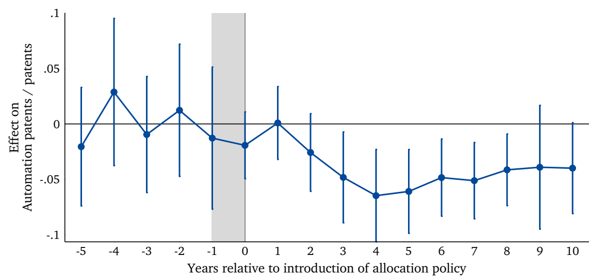
	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate × $\text{Allocation}_{t\geq 0}$	−0.320**	−0.257*		
	(0.158)	(0.143)		
1(Cum. inflow rate>p50) × $\text{Allocation}_{t\geq 0}$			−0.042**	−0.038**
			(0.019)	(0.019)
Skill share: high (ext.)	−0.484		−0.968	
	(2.244)		(2.216)	
Skill share: medium (ext.)	−0.848		−1.256	
	(2.365)		(2.296)	
Skill share: low (ext.)	−0.789		−1.243	
	(2.410)		(2.342)	
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Skill share trends	No	Yes	No	Yes
Observations	2029	2029	2029	2029
R-squared	0.681	0.690	0.682	0.692
Within R-squared	0.014	0.042	0.017	0.046
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. To measure skill composition, we calculate the share of employees with high, medium, and low skills for each region according to an education-based classification. In Columns (1) and (3), we extrapolate the skill share trends from the pre-allocation period (1990 to 1995) for the allocation period. In Columns (2) and (4), we take the skill shares averaged over the pre-allocation period and interact these with year dummies. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

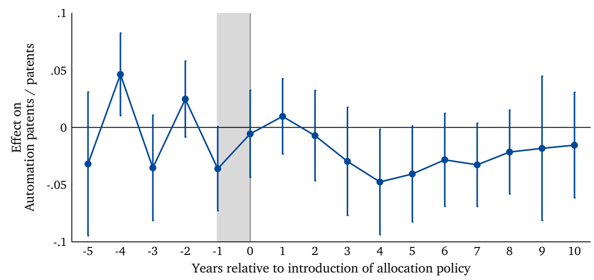
C.7 Alternative Estimation Techniques

Figure C-4: Effect of labor inflows on the share of automation patents – event study with binary treatment – alternative estimators

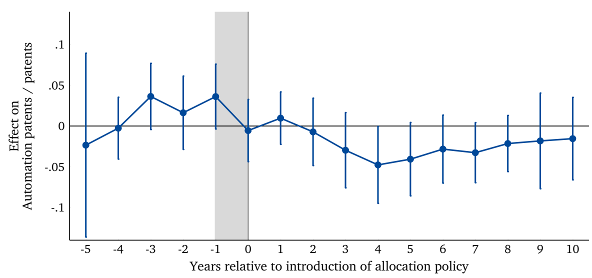
(a) Borusyak et al. (2021)



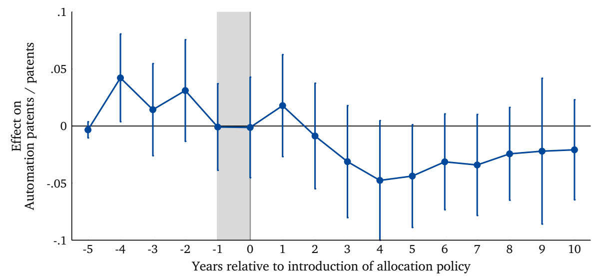
(b) Callaway and Sant'Anna (2021)



(c) De Chaisemartin and d'Haultfoeuille (2020)



(d) Sun and Abraham (2021)



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on linear regressions as specified in Equation 1. In all four panels, the dependent variable is the share of automation patents, and the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with leads and lags. The estimates in each panel are obtained using a different estimator. Panel (a) uses the estimator introduced by Borusyak et al. (2021), using the Stata package *did_imputation*. Panel (b) uses the estimator introduced by Callaway and Sant'Anna (2021), using the Stata package *csdid*. Panel (c) uses the estimator introduced by De Chaisemartin and d'Haultfoeuille (2020), using the Stata package *did_multiplegt*. Panel (d) uses the estimator introduced by Sun and Abraham (2021), using the Stata package *eventstudyinteract*. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. Region and time fixed effects included. Standard errors clustered at the regional level.

C.8 Alternative Sample Definitions and Subsample Analysis

Table C-14: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – alternative weighting

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Cum. inflow rate × $Allocation_{t\geq 0}$	−0.267	−0.372**		
	(0.195)	(0.177)		
1(Cum. inflow rate>p50) × $Allocation_{t\geq 0}$			−0.040*	−0.049**
			(0.022)	(0.021)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	No	Yes	No	Yes
Observations	2029	2029	2029	2029
R-squared	0.091	0.703	0.092	0.705
Within R-squared	0.091	0.017	0.092	0.023
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($Allocation_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($Allocation_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. The regressions are estimated at the region-year level and weighted by the regional GDP in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-15: Effect of labor inflows on the share of automation patents – DiD model with continuous and binary treatment – regions with low or high innovative capacity excluded

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
Excluding regions with ... patent activities:	Low		High	
Cum. inflow rate × $Allocation_{t\geq 0}$	−0.330**		−0.147	
	(0.153)		(0.120)	
1(Cum. inflow rate>p50) × $Allocation_{t\geq 0}$		−0.044**		−0.019
		(0.018)		(0.014)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1840	1840	1837	1837
R-squared	0.723	0.725	0.625	0.626
Within R-squared	0.014	0.020	0.002	0.003
Dep. Var. mean	0.268	0.268	0.267	0.267

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (3), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($Allocation_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (2) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($Allocation_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. Regions with patent activities in the pre-allocation period below the 10th (low) and above the 90th percentile (high) excluded. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-16: Effect of labor supply on the share of automation innovation – DiD model with continuous treatment – state-excluding subsamples

	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
State excluded:	BW	NS	NW	HE/SL/SH
Cum. inflow rate × Allocation$_{t\geq 0}$	−0.341*	−0.336**	−0.443**	−0.211
	(0.181)	(0.155)	(0.201)	(0.140)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1581	1486	1453	1567
R-squared	0.648	0.702	0.638	0.714
Within R-squared	0.013	0.017	0.016	0.004
Dep. Var. mean	0.253	0.264	0.284	0.274

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In all columns, the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. Subsamples exclude regions in the indicated state(s) (BW: Baden-Württemberg, NS: Lower Saxony, NW: North Rhine-Westphalia, HE/SL/SH: Hesse, Saarland, Schleswig-Holstein). The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table C-17: Effect of labor supply on the share of automation innovation – DiD model with binary treatment – state-excluding subsamples

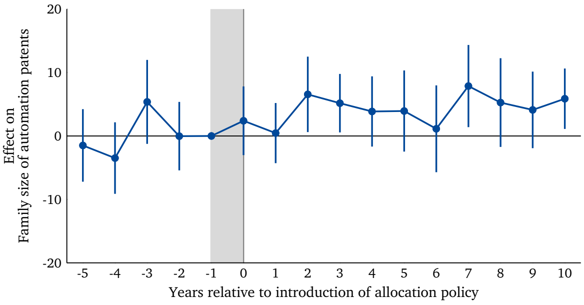
	(1)	(2)	(3)	(4)
Dep. Var.:	Automation patents / patents			
State excluded:	BW	NS	NW	HE/SL/SH
1(Cum. inflow rate>p50) × Allocation$_{t\geq 0}$	−0.046**	−0.044**	−0.050**	−0.029**
	(0.022)	(0.019)	(0.024)	(0.012)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1581	1486	1453	1567
R-squared	0.649	0.703	0.638	0.715
Within R-squared	0.017	0.022	0.017	0.007
Dep. Var. mean	0.253	0.264	0.284	0.274

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In all columns, the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. Subsamples exclude regions in the indicated state(s) (BW: Baden-Württemberg, NS: Lower Saxony, NW: North Rhine-Westphalia, HE/SL/SH: Hesse, Saarland, Schleswig-Holstein). The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

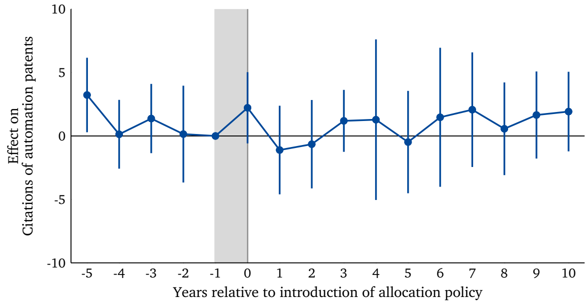
C.9 Examining Changes in Patent Quality

Figure C-5: Effect of labor inflows on the quality of automation patents – event study with continuous treatment

(a) Quality proxy: patent family size



(b) Quality proxy: citations



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on a linear regression as specified in Equation 1. In both panels, the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. In Panel (a), the dependent variable is the average patent family size of automation patents. In Panel (b), the dependent variable is the average number of US citations (received within 5 years after filing) of automation patents. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. Region and time fixed effects included. Standard errors clustered at the regional level. The coefficients correspond to those reported in Appendix Table C-19.

Table C-18: Effect of labor inflows on the share of automation patents by patent characteristics – DiD with continuous treatment

	(1)	(2)	(3)	(4)
Dep. Var.:	Family size		Citations	
Cum. inflow rate × $\text{Allocation}_{t\geq 0}$	1.501		−2.412	
	(1.363)		(1.645)	
1(Cum. inflow rate>p50) × $\text{Allocation}_{t\geq 0}$		0.432**		−0.186
		(0.184)		(0.132)
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1956	1956	1956	1956
R-squared	0.387	0.390	0.397	0.395
Within R-squared	0.001	0.006	0.022	0.019
Dep. Var. mean	5.169	5.169	1.123	1.123

Notes: The table presents the results of linear regression analyses as defined in Equation 2. In columns (1) and (2), the dependent variable is the average patent family size of automation patents. In columns (3) and (4), the dependent variable is the average number of US citations (received within 5 years after filing) of automation patents. In columns (1) and (3), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (2) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy ($\text{Allocation}_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

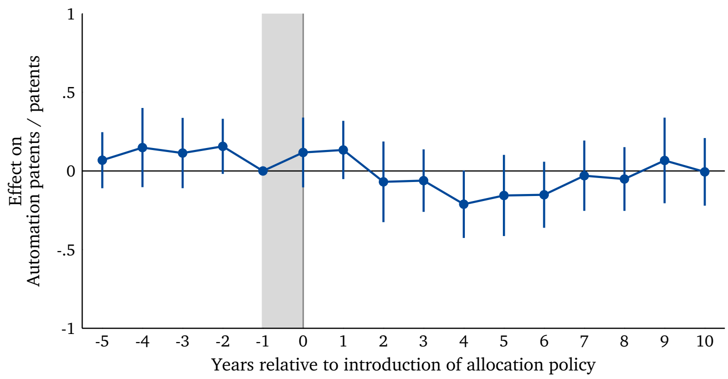
Table C-19: Effect of labor inflows on the share of automation patents by patent characteristics – event study with continuous treatment

Dep. Var.: Automation patent	(1) Patent family size	(2) Citations (US, 5yrs)
Cum. inflow rate		
× Allocation$_{t-5}$	−1.491	3.230**
	(2.889)	(1.483)
× Allocation$_{t-4}$	−3.484	0.134
	(2.849)	(1.370)
× Allocation$_{t-3}$	5.373	1.374
	(3.343)	(1.379)
× Allocation$_{t-2}$	−0.027	0.148
	(2.726)	(1.928)
× Allocation$_{t-1}$	0.000	0.000
	(.)	(.)
× Allocation$_{t}$	2.396	2.222
	(2.736)	(1.418)
× Allocation$_{t+1}$	0.453	−1.108
	(2.396)	(1.764)
× Allocation$_{t+2}$	6.558**	−0.648
	(3.005)	(1.760)
× Allocation$_{t+3}$	5.170**	1.191
	(2.328)	(1.233)
× Allocation$_{t+4}$	3.864	1.284
	(2.797)	(3.198)
× Allocation$_{t+5}$	3.934	−0.481
	(3.231)	(2.038)
× Allocation$_{t+6}$	1.132	1.473
	(3.458)	(2.769)
× Allocation$_{t+7}$	7.869**	2.073
	(3.273)	(2.282)
× Allocation$_{t+8}$	5.266	0.568
	(3.532)	(1.846)
× Allocation$_{t+9}$	4.109	1.653
	(3.049)	(1.734)
× Allocation$_{t+10}$	5.872**	1.922
	(2.407)	(1.584)
Region fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	1956	1956
R-squared	0.400	0.422
Within R-squared	0.023	0.063
Dep. Var. mean	5.169	1.123

Notes: The table presents the results of linear regression analyses as defined in Equation 1. In column (1), the dependent variable is the average patent family size of automation patents. In column (2), the dependent variable is the average number of US citations (received within 5 years after filing) of automation patents. The treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. These time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

C.10 Analysis Based on Pre-allocation Labor Inflows

Figure C-6: Effect of labor inflows on the share of automation patents – event study with continuous treatment – inflow rate from pre-allocation period



Notes: The figure displays event study estimates and the 95 percent confidence intervals based on a linear regression as specified in Equation 1. The dependent variable is the share of automation patents, and the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with leads and lags. The cumulative inflow rate is modified in so far as the numerator is replaced with the endogenous cumulative inflows from the pre-allocation period. The time dummies indicate that the allocation policy is introduced l years away, with $l \in \{-5; +10\}$. The coefficient for $l = -1$ is normalized to zero. Region and time fixed effects included. Standard errors clustered at the regional level.

Table C-20: Effect of labor inflows on the share of automation patents – DiD with continuous treatment – inflow rate from pre-allocation period

	(1)	(2)	(3)	(4)
Dep. Var.:		Automation patents / patents		
Cum. inflow rate^{Pre} × Allocation$_{t\geq 0}$	−0.086	−0.135**		
	(0.075)	(0.064)		
1(Cum. inflow rate^{Pre}>p50) × Allocation$_{t\geq 0}$			−0.017	−0.019
			(0.016)	(0.015)
Region fixed effects	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2029	2029	2029	2029
R-squared	0.122	0.679	0.129	0.678
Within R-squared	0.021	0.005	0.029	0.003
Dep. Var. mean	0.269	0.269	0.269	0.269

Notes: The table presents the results of linear regression analyses as defined in Equation 2. The dependent variable is the share of automation patents. In columns (1) and (2), the treatment variable is continuous: the cumulative inflow rate during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. In columns (3) and (4), the treatment variable is binary: a dummy indicating regions with above-average cumulative inflows of ethnic Germans during the allocation period, interacted with a dummy (Allocation$_{t\geq 0}$) that indicates whether the allocation policy was introduced in that year or earlier. The cumulative inflow rate is modified in so far as the numerator is replaced with the endogenous cumulative inflows from the pre-allocation period. The regressions are estimated at the region-year level and weighted by the regional population in 1991. Standard errors, clustered at the regional level, are in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

D Technical Appendix

D.1 Data

Sample

We assign NUTS3 regions to labor market regions using a reference file from the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). With Baden-Württemberg and North Rhine-Westphalia our sample includes the most innovative and most populous federal states. No regional data on ethnic German inflows were recorded for the state of Bavaria (Glitz, 2012). We exclude the labor market region of Ulm as it is partly in the state of Bavaria, which did not implement the placement policy. We merge the labor market regions "Osterode" and "Goettingen" to make the patent data compatible with the regional data on occupation groups and skill groups.

We do not analyze the years before 1991, because the selection of ethnic Germans was different before the collapse of the Soviet Union (1991). Also, regional data on migrant inflows are not systematically available before 1991. After 2006 the number of incoming ethnic Germans was negligible.

For further details on the assignment of labor market regions to the analysis sample, see Table D-2 in this Appendix.

Details on the patent data

Patent Texts. The overwhelming majority of EP patents has a publication in the English language. If not, we draw on English publications within the same DOCDB patent family from patent offices where English is the official language (e.g., patent publications from the USPTO and the UKIPO). The English abstract is available for more than 99.86 percent of the patents in our sample. Patent applications with a missing English abstract are excluded from the construction of the regional measures of automation innovation.

Identification of Corporate Applicants. We identify corporate vs. non-corporate entities with the sector categorization provided by PATSTAT. We consider applicants as corporate entities if they are labeled as either "Company", "Company Gov Non-Profit" or "Company Hospital". Non-corporate entities are labeled as "hospital", "individual", "governmental non-profit university", "governmental non-profit" or "university". In the very rare case of missing applicant entity, we assume that the applicant entity is corporate.

Firm-level Data from Orbis. The record linkage between patent applicant names and corporate accounts is part of Orbis IP. About 5 percent of patents with a corporate applicant remain unmatched, and another 8 percent of patents are linked to corporate accounts with no information on the number of employees or annual turnover.

Table D-1: Regional characteristics

Variable	Source and description
Ethnic inflows	Glitz (2012) and Piopiunik and Ruhose (2017), original data from 1991 to 2001 from the admission center and from 2002 to 2006 from Bundesarbeitsgemeinschaft Evangelische Jugendsozialarbeit e.V., Jugendmigrationsdienste, based on statistics collected in Friedland.
Population	1992 and 1994-2008: Working Group Regional Accounts VGRdL. 1991 and 1993: Glitz (2012), original data from the German Statistical Office.
GDP	Working Group Regional Accounts VGRdL. The statistical offices of all German states and the Federal Statistical Office are members of the Working Group Regional Accounts VGRdL. We impute the 1993 values using the 1992 values and the 1993 national growth rate. We impute the 1991 values using the 1992 values and the 1992 national growth rate.
Workforce	Federal Employment Agency. The dependent workforce.
Unemployment rate	Federal Employment Agency. Based on the dependent workforce.
Skill groups	Establishment History Panel from the Institute for Employment (customized analysis, based on the population of establishments with at least one employee subject to social security in Germany). The shares of high skilled employees (with university degree or applied university degree), medium skilled employees (with school degree and vocational education but no higher degree) and the share of engineers and scientists (employees with a degree from a university or a university of applied sciences and with specific occupation classifications). From 1999 onwards, the data set also covers marginally employed persons.
Occupation groups	Establishment History Panel from the Institute for Employment (customized analysis, based on the population of establishments with at least one employee subject to social security in Germany). The shares of 12 different occupation groups (agricultural, unskilled manual, unskilled services, unskilled commercial and admin., skilled manual, skilled services, skilled commercial and admin., technicians, semiprofessions, engineers, professions, managers). From 1999 onwards, the data set also covers marginally employed persons.

Table D-2: Analysis sample

Federal State of Germany	Labor market regions	Analysis sample	Pre-allocation period	Allocation period			
		Region-year observations	Region-year observations	Region-year observations	Start year	End year	Implementation date
Baden-Wuerttemberg	28	448	140	308	1996	2006	01.03.1996
North Rhine-Westphalia	36	576	180	396	1996	2006	01.03.1996
Schleswig-Holstein	8	128	40	88	1996	2006	01.03.1996
incl. Hamburg[*]					1996	2006	01.03.1996
Lower Saxony	34	544	204	340	1997	2006	07.04.1997
incl. Bremen[**]					1997	2006	01.03.1996
Saarland	4	64	20	44	1996	2006	11.03.1996
Hesse	17	272	187	85	2002	2006	01.01.2002
Total	127	2032	771	1261			

Notes: Analysis sample and the implementation of the assigned place of residence act. Regional level: labor market region. We have merged the regions "Osterode" and "Goettingen" into one region to make the data compatible with the regional employment data. We conservatively exclude the region "Ulm" because it is partly located in the state of Bavaria which did not implement the placement policy. [*]The region "Hamburg" contains the state of Hamburg, three counties in Schleswig-Holstein and one county in Lower Saxony. Since the counties in Schleswig-Holstein and the city of Hamburg are dominant, we follow Glitz (2012) and include this region from 1996 onward when Hamburg and Schleswig-Holstein adopted the placement policy. [**]The regions "Bremen" and "Bremerhaven" are partly in the state of Bremen and contain counties that are in Lower Saxony. We conservatively include these regions only from 1997 onward when Lower Saxony implemented the allocation policy. The region "Mannheim" contains also one county in Hesse. We conservatively include this region only from 2002 onward when Hesse implemented the allocation policy. For the construction of the state fixed effects and year-by-state fixed effects, we assign the region "Hamburg" to the state of Schleswig-Holstein, the regions "Bremen" and "Bremerhaven" to the state of Lower Saxony and the region "Mannheim" to the state of Hesse. Data on ethnic German inflows are not available for the region "Hannover" (Lower Saxony) for 2002-2006 and for regions within the state of Hesse for 1992-1994. To merge the inflow data with the patent data, we account for territorial reforms in the regions using historical files of changes between the various NUTS versions from Eurostat. There were only two territorial reforms in the labor market regions of the allocation states: in 2001, “Hannover, Kreisfreie Stadt" and “Hannover, Landkreis" were united into “Region Hannover.” Likewise in 2009, “Aachen, Kreisfreie Stadt" and “Aachen, Kreis" were united into “Städteregion Aachen." We do not analyze the years before 1991, because the selection of ethnic Germans was different before the collapse of the Soviet Union (1991). Also, regional data on migrant inflows are not systematically available before 1991. After 2006 the number of incoming ethnic Germans was negligible.

D.2 Classification of Automation Patents

We classify EP patents into automation and non-automation with natural language processing. This classification eercise unfolds in several steps: data loading, text pre-processing, feature extraction, model training, and prediction. The following *python* libraries are used to support these operations: *pandas*, *numpy*, *nltk*, *sklearn*, and *gensim*.

We start by loading our training and testing data sets. Our training data originates from Mann and Püttmann (2023) and consists of the EP patents in our sample with a US patent counterpart previously classified as either 'automation' or 'non-automation' by Mann and Püttmann (2023). The EP patents' titles and abstracts provide the text input for our machine-learning model, while the previous classification informs the model about the target output. The testing data consists of all EP patents in our sample.

In the pre-processing step, we transform the patent raw text to prepare them for analysis. We convert the text to lowercase, strip extraneous spaces, and remove punctuation, HTML tags, numbers, and other non-word characters. We further remove English stopwords and lemmatize the remaining words, reducing them to their root form.

Following pre-processing, we extract features from the processed text. To this end, we convert the cleaned text into numerical representations using both Tf-Idf (Term frequency-Inverse document frequency) and Word2Vec. These techniques effectively capture the important characteristics of the text and convert them into numerical vectors that our machine-learning model can understand.

Having extracted features from the text data, we train two different classification models: Logistic Regression and Naive Bayes, using the Tf-Idf vectors. We then evaluate these models, looking at a variety of metrics such as the confusion matrix, classification report, and AUC-ROC curve to assess performance.

In the final step, we use the better-performing model to classify all patents in our test data into 'automation' or 'non-automation.' We save our trained model using *pickle* for replication.

Table D-3: Examples of automation patents

EP application number	EP19940115956
Title	Method for controlling dryers in brick factories
Assignee	INNOVATHERM Prof Dr Leisenberg GmbH
Abstract	According to a method for controlling dryers in the ceramic industry, which are operated with an external and/or internal heating and are operated in conjunction with a tunnel furnace, and in which, by means of an empirical or mathematical model of the drying process, the state of drying-out of the chambers is approximately determined, the variation of the air condition values and/or of the convection capacity of the individual chambers is **automat**ically selected or calculated as a function of the products, the available drying time and the available furnace waste heat, from the point of view of a minimum total energy or energy cost expenditure, and is started as a drying programme. This mode of operation achieves an **automat**ic optimisation of the heat consumption in a targeted manner, using the prescribed boundary conditions, and, as a result, the costs of energy and personnel are reduced in a concern without large investment in terms of installation being necessary.
Label	Automation patent
Classification score	0.998
EP application number	EP19990942756
Title	Method for **automat**ically controlling and selecting the bodies of slaughtered poultry
Assignee	CSB SYST SOFTWARE ENTWICKLUNG, Csb-System Software-Entwicklung & Unternehmensberatung AG
Abstract	The invention describes a method for **automat**ically controlling and selecting the bodies of slaughtered poultry. The invention aims at providing a very easy and cost-effective method for **automat**ically controlling and selecting the bodies of slaughtered poultry. According to the invention, said aim is achieved in that the body of the slaughtered poultry to be controlled is selected by conducting color analysis of the light reflected from the visible surface parts, said light being detected as a diffuse color mixing light eliminating spatial contours using a measuring technique, wherein the measured value is used for selection as integrating value for the totality of visible surface parts.
Label	Automation patent
Classification score	0.867

Notes: Source: PATSTAT.

References

Bade, K. J. (1990). Aussiedler-Rückwanderer über Generationen hinweg. In K. J. Bade (Ed.), *Neue Heimat im Westen. Vertriebene. Flüchtlinge. Aussiedler*. Münster: Westfaelischer Heimatbund.

Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. arXiv preprint arXiv:2108.12419.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics 225*(2), 200–230.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–96.

Dietz, B. (2006). Aussiedler in Germany: From Smooth Adaptation to Tough Integration. In L. Lucassen, D. Feldmann, and J. Oltmer (Eds.), *Paths of Integration. Migrants in Western Europe (1880-2004)*, pp. 116–136. Amsterdam: Amsterdam University Press.

Friedrichs, N. (2022). Integration gelungen?: Lebenswelten und gesellschaftliche Teilhabe von (Spät-) Aussiedlerinnen und (Spät-) Aussiedlern. *SVRStudie 2022-1*.

Fuchs, M. (1999). Identifikation und integration. die subjektive perspektive von aussiedlern 15 jahre nach der einreise. In: Wendt, Hartmut (Hg.): Zuwanderung nach Deutschland. Prozesse und Herausforderungen. Materialien zur Bevölkerungswissenschaft, Heft 94, 18-38.

Ganglmair, B. and I. Reimers (2019). Visibility of Technology and Cumulative Innovation: Evidence from Trade Secrets Laws. ZEW-Centre for European Economic Research Discussion Paper 19-035.

Gartner, H. (2004). Die Imputation von Löhnen oberhalb der Beitragsbemessungsgrenze in der IAB-Beschäftigten-Stichprobe. *Methodenreport. Nürnberg*.

Glitz, A. (2012). The Labor Market Impact of Immigration: A Quasi-experiment Exploiting Immigrant Location Rules in Germany. *Journal of Labor Economics 30*(1), 175–213.

Haug, S. and L. Sauer (2007). Zuwanderung und Integration von (Spät-) Aussiedlern: Ermittlung und Bewertung der Auswirkungen des Wohnortzuweisungsgesetzes. Forschungsbericht 3. Bundesamt für Migration und Flüchtlinge. Nürnberg.

Heller, W., H.-J. Bürkner, and H.-J. Hofmann (2002). Migration, segregation und integration von aussiedlern-ursachen, zusammenhänge und probleme. In: Heller, Hartmut (Hg.): Neue Heimat Deutschland. Aspekte der Zuwanderung, Akkulturation und emotionalen Bindung. Erlangen: Universitätsbund Erlangen-Nürnberg, 79-108.

Jahn, V. and M. F. Steinhardt (2016). Innovation and Immigration – Insights from a Placement Policy. *Economics Letters 146*, 116–119.

Jahn, V. and M. F. Steinhardt (2023). Immigration and New Firm Formation – Evidence from a Quasi-Experimental Setting in Germany. *International Migration Review forthcoming*, 01979183231165496.

Mann, K. and L. Püttmann (2023). Benign Effects of Automation: New Evidence from Patent Texts. *The Review of Economics and Statistics 105*(3), 562–579.

Panagiotidis, J. (2021). *Postsowjetische Migration in Deutschland*. Bundeszentrale für politische Bildung.

Piopiunik, M. and J. Ruhose (2017). Immigration, Regional Conditions, and Crime: Evidence from an Allocation Policy in Germany. *European Economic Review 92*, 258–282.

Schmoch, U. (2008). Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation (WIPO).

Sun, L. and S. Abraham (2021). Estimating Dynamic Treatment effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics 225*(2), 175–199.

Worbs, S., E. Bund, M. Kohls, and C. B. von Gostomski (2003). Segregation, regionale mobilität und soziale integration von aussiedlern. in: Bade, Klaus/Oltmer, Jochen (Hg): Aussiedler: deutsche Einwanderer aus Osteuropa, IMIS-Schriften Band 8. Göttingen, 107-126.