Top Flight: Who migrates in response to top tax rates?

Arun Advani
University of Warwick

César Poux
LSE

Andy Summers
LSE

April 29, 2024

HMRC Disclaimer

This work contains statistical data from HMRC which is Crown Copyright. The research datasets used may not exactly reproduce HMRC aggregates. The use of HMRC statistical data in this work does not imply the endorsement of HMRC in relation to the interpretation or analysis of the information.

Motivation

- Concern about migration response to taxes at the top, and whether it constrains tax authorities.
- Existing evidence focused on specific sub-populations:
  - special tax regimes for foreigners (Kleven, Landais, Saez & Schultz, 2014 ; Advani, Burgherr & Summers, 2024)
  - specific occupations (Kleven, Landais & Saez, 2013; Akcigit, Baslandze & Stantcheva, 2016).
- Lack of 'systematic evidence on the mobility elasticities of the broader population' (Kleven, Landais, Muñoz & Stantcheva, 2020) needed to draw conclusions from a revenue perspective.

Who responds to tax reforms by migrating, and what are the consequences of this migration on population in both the short and long run?

This paper

- We leverage two large tax reforms affecting top earners – one in the UK and one in France – to determine the scale of the migration response.
  - We look at all top earners, not limited by industry/occupation/nationality
- We examine the heterogeneity in responses across different dimensions
  - By native/foreigner, income level, ex-ante probability of emigration
- We build a model that rationalises these results, develop testable predictions that allow us to distinguish alternative mechanisms, and use structural estimation to infer the long-run stock and migration elasticities.
- (To come: we translate migration responses into short- and long -run effects on tax revenue).

Main results

1. Small, precisely estimated semi-elasticity for natives with a point estimate of **-0.01** (-0.03 to 0.01)
2. Moderate semi-elasticity for migrants: **-0.20** (-0.25 to -0.14).
3. Tax rate increases migration rate only for top two deciles by baseline migration probability.
4. Semi-elasticity is declining in income.
5. Immigration effects are small in absolute terms, but significant.
6. Many foreigners who respond to tax would otherwise have stayed for many years.
    - Implying large stock elasticity: 7.6 for individuals with incomes £150-225k from arrival.

Contribution to the literature

- **Estimating the magnitude of international migration responses to tax**
  - Have evidence for specific occupations (Kleven, Landais & Saez, 2013; Akcigit, Baslandze & Stantcheva, 2016; Agrawal & Tester, 2024)
  - ...and special regimes for foreigners (Kleven, Landais, Saez & Schultz, 2014; Bassetto & Ippedico, 2023; Giarola, Marie, Cörvers & Schmeets, 2023; Advani, Burgherr & Summers, 2024).
  - We provide evidence from a broad based tax reform, affecting all top incomes.
- **Heterogeneity of migration responses**
- **Revenue implications of migration responses**

Contribution to the literature

- **Estimating the magnitude of international migration responses to tax**
- **Heterogeneity of migration responses**
  - Existing evidence that natives more responsive than foreigners (Kleven, Landais & Saez, 2013; Akcigit, Baslandze & Stantcheva, 2016).
  - We provide evidence on natives vs foreigners in a general setting.
  - We show important heterogeneity among foreigners – both by tenure and income level – which affects interpretation of standard elasticities.
- **Revenue implications of migration responses**

Contribution to the literature

- **Estimating the magnitude of international migration responses to tax**
- **Heterogeneity of migration responses**
- **Revenue implications of migration responses**
  - Significant public finance impact of intra-national migration (Agrawal, Foremny & Martinez-Toledano, 2023; Agrawal & Tester, 2024; Brülhart, Gruber, Krapf & Schmidheiny, 2022; Rauh, 2022; Young & Varner, 2011; Young, 2017)
  - ...and more recently of international migration (Advani, Burgherr & Summers, 2024; Jakobsen, Kleven, Kolsrud, Landais & Muñoz, 2024).
  - We study revenue effect of international migration, in both short- and long-run, to establish the degree of the constraint on public finance.
  - We highlight implications of observed heterogeneity in migration rates for how to design foreign worker regimes.

Context and data

Data

- Administrative tax data from HMRC, covering the universe of income tax return filers (mandatory for annual income $>$ £100,000) from 2004 to 2018.
- Observe:
  - UK income (including breakdown into components and industry), capital gains, and tax paid.
  - (Some) personal characteristics: sex, age, native/foreign status + country they came from if foreign, date of first arrival in the UK for foreigners.

Measurement of migration variables

The 2010 top tax rate increase in the UK

- Income tax in the UK: progressive, includes labour and capital income (different rate for dividends up to 2016).
- **The reform:** top marginal rate goes **from 40%** in tax year 2009/2010 **to 50%** in 2010/2011.
  - Also an increase in marginal rate between roughly 100-120k, from 40% to 60%.
- Largest reform in rates since 20 years prior.
- **Treatment:** taxpayers earning £165,000-225,000 contemporaneously.
- **Control:** taxpayers earning £120,000-135,000 contemporaneously.
- Separating natives and foreigners, and also by baseline emigration probability.

Overall migration response to tax increases

Estimation of emigration elasticity – UK reform

Classic IV difference-in-difference approach:

$$E_{i,t} = \alpha X_{i,t} + \log(1 - \tau_i)\beta + \gamma_t + \varepsilon_{i,t} \tag{1}$$

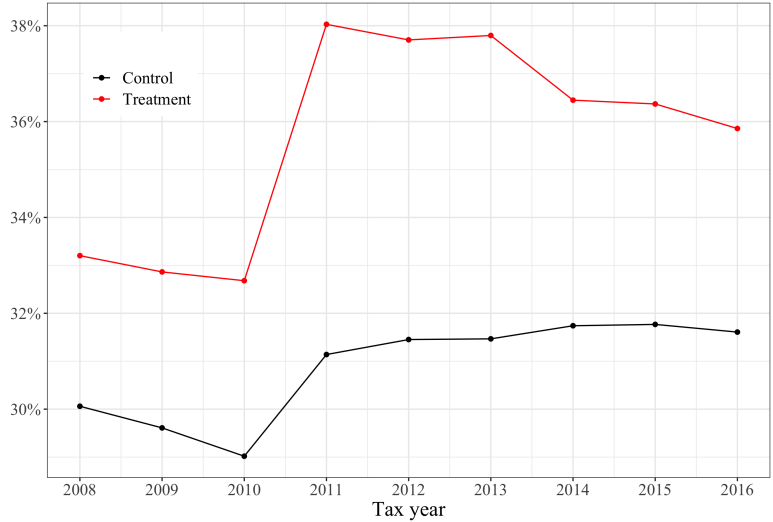
where $E_{i,t}$ the emigration dummy, $X_{i,t}$ the individual level controls, γ_t the year fixed effect. β captures the semi-elasticity: pp change in emigration wrt 1% change in the log net-of-tax rate. We instrument the log net of tax rate with static DiD estimator: event$\times$treatment ($\delta D_i + \gamma_t$).

Reduced form equation:

$$E_{i,t} = \alpha X_{i,t} + \sum_{\substack{t=2004 \\ t \neq 2009}}^{2016} \beta_t D_i . T_t + \delta D_i + \gamma_t + \varepsilon_{i,t} \tag{2}$$

D_i is the treatment dummy and β_t are the coefficients of interest.
X_i includes: time since arrival (as three categories), main source of income, and sex.

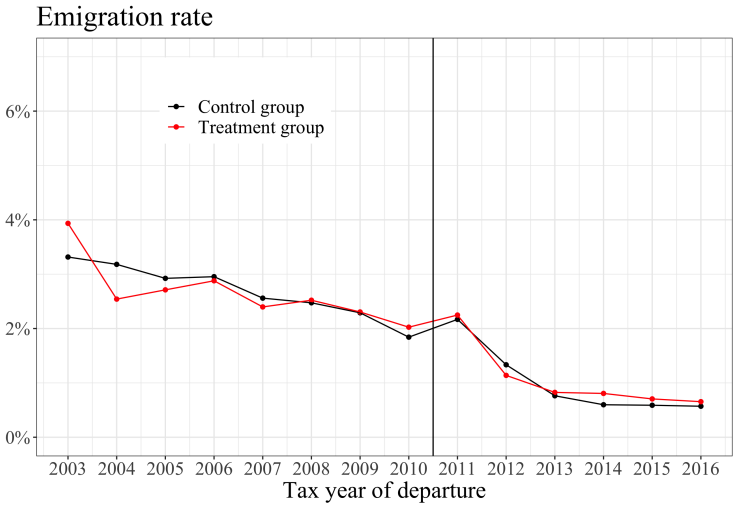
First stage of UK reform: ATR evolution



Notes: This figure shows observed average tax rate (ATR) in control and treatment group from 2008 to 2016. Control group includes taxpayers earning £165,000-225,000 contemporaneously, treatment group taxpayers earning £120,000-135,000 contemporaneously.

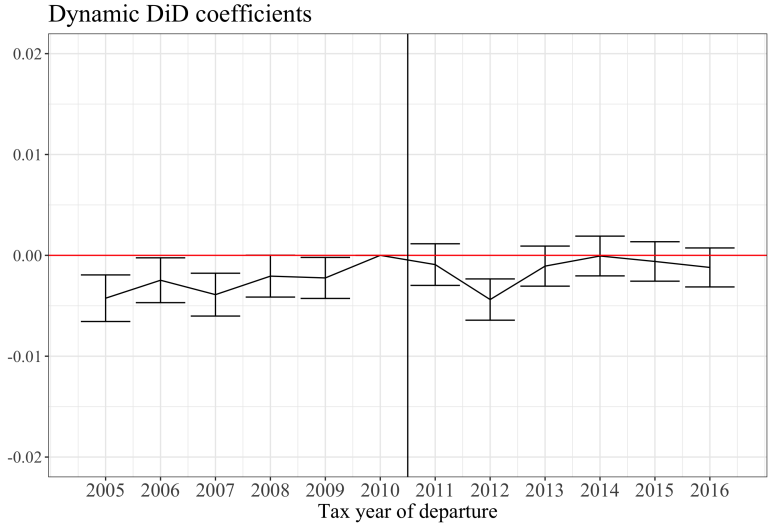
Source: Authors' calculations based on HMRC administrative datasets. ATR by income

No large effect of tax reform on migration of natives (1/2)



Notes: This figure shows the share of native emigrants in the control and treatment group between 2003 and 2016. Shares are computed after rounding number of emigrants and population size to the nearest 50. Control group includes foreign taxpayers earning £165,000-225,000 contemporaneously, treatment group includes foreign taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

No large effect of tax reform on migration of natives (2/2)

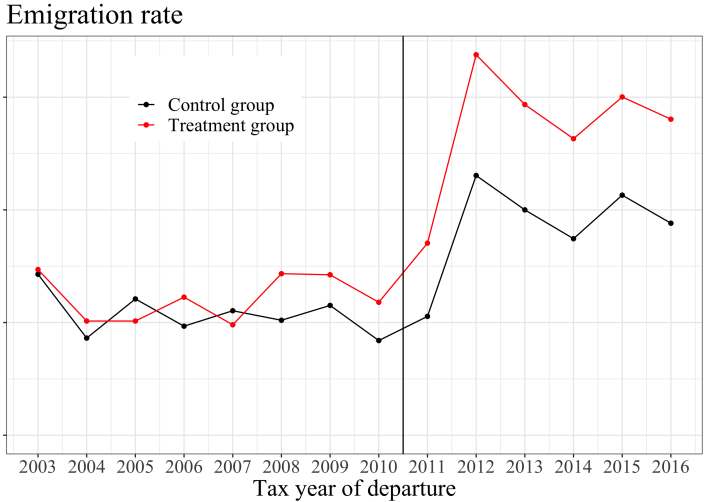


Notes: This figure shows shows the coefficients and associated standard errors from Equation 2 for natives. Control group includes foreign taxpayers earning £165,000-225,000 contemporaneously, treatment group includes foreign taxpayers earning £120,000-135,000 contemporaneously.

Source: Authors' calculations based on HMRC administrative datasets.

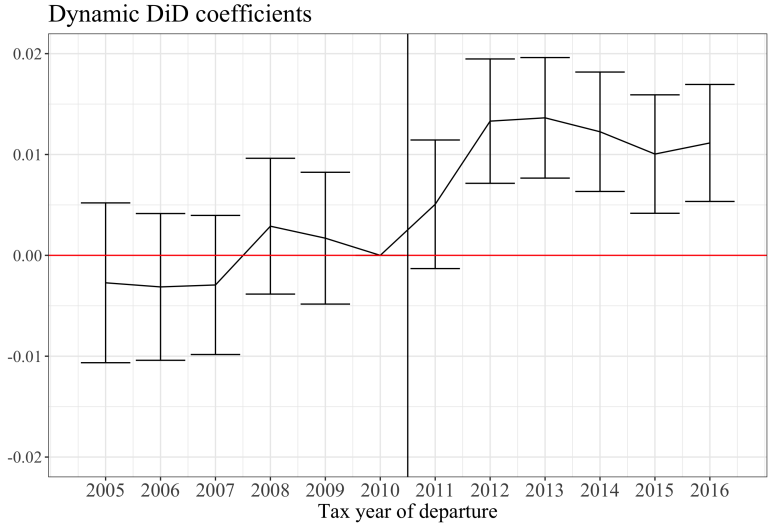
Native response - French reform

Clear response by foreigners (1/2)



Notes: This figure shows the share of foreigner emigrants in the control and treatment group between 2003 and 2016. Shares are computed after rounding number of emigrants and population size to the nearest 50. Control group includes foreign taxpayers earning £165,000-225,000 contemporaneously, treatment group includes foreign taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Clear response by foreigners (2/2)



Notes: This figure shows shows the coefficients and associated standard errors from Equation 2 for foreigners. Control group includes foreign taxpayers earning £165,000-225,000 contemporaneously, treatment group includes foreign taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Semi-elasticity estimates

	First stage: log net-of-average tax rate	Second stage: reduced form emigration rate	2SLS: semi-elasticity
Panel A: Natives			
Coefficient	-0.0501^{***}	0.00070	-0.0140
	(0.0053)	(0.00061)	(0.0123)
N	1,676,880	1,658,251	-
Panel B: Foreigners			
Coefficient	-0.0599^{***}	0.0118^{***}	-0.1963^{***}
	(0.0050)	(0.0013)	(0.0268)
N	354,816	358,540	-
Time FE	Yes	Yes	Yes
Controls	-	Yes	Yes

Notes: IV estimates of the semi-elasticity of the emigration rate with respect to the net-of-average-tax rate, exploiting the 2011 increase in the top tax rates in the UK. First-stage estimate captures the effect of the reform on the net-of-average-tax rate and is computed at the aggregate level. Reduced-form estimate shows the effect of the reform on the emigration rate estimated with Equation 2. 2SLS estimate of the migration semi-elasticity η is the percentage-point change in the emigration rate in response to a 1% increase in the net-of-average-tax rate, obtained from estimating Equation 1. Standard error for the 2SLS coefficient are obtained using the Delta-Method.
Source: Authors' calculations based on HMRC administrative datasets.

Who responds? Heterogeneity analysis.

Why does heterogeneity matter?

1. Can bias our interpretation of migration elasticity.
   - Semi-elasticity always well-defined for any group g, and aggregates nicely.
   - If there is heterogeneity in both semi-elasticity ($\tilde{E}_i$) and ex-ante emigration probability (E_i), aggregate elasticity is not a weighted average of group-level elasticities.

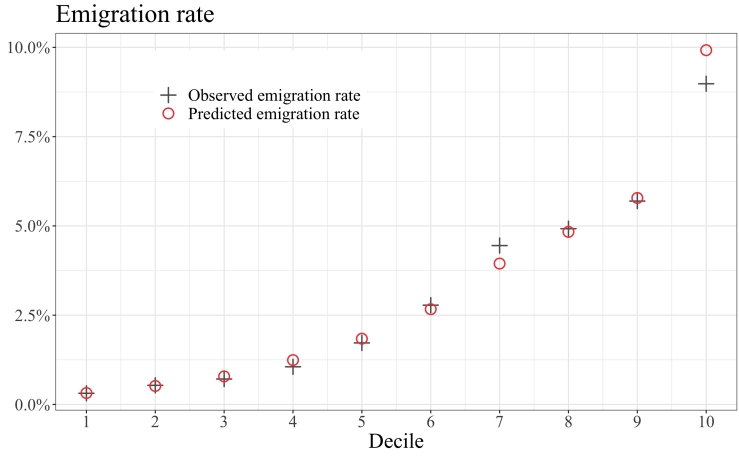
$$\hat{\eta} = \frac{\sum_i \tilde{E}_g}{\sum_g E_g} = \sum_g \Big[\underbrace{\frac{\tilde{E}_g}{E_g}}_{\eta_g} \times \underbrace{\frac{E_g}{\sum_j E_j}}_{w_g} \Big] \neq \sum_g \frac{\tilde{E}_g}{E_g}$$

   - Aggregate elasticity biased towards the elasticity for groups with highest baseline emigration
2. Potentially wider impacts of migration on the economy, depending on who migrates
   - business ownership (Jakobsen et al., 2024)
   - entrepreneurship (Hanson, Kerr & Turner, 2018; Kerr, 2019; Kerr & Kerr, 2020)
   - patenting (Mayda, Orefice & Santoni, 2022)

Three key facts

1. Natives hardly respond, while modest response for migrants.
2. Tax-induced mobility only among those with highest ex-ante probability of leaving.
3. Semi-elasticity is declining in income at high levels of income.

Predicting the baseline probability of leaving

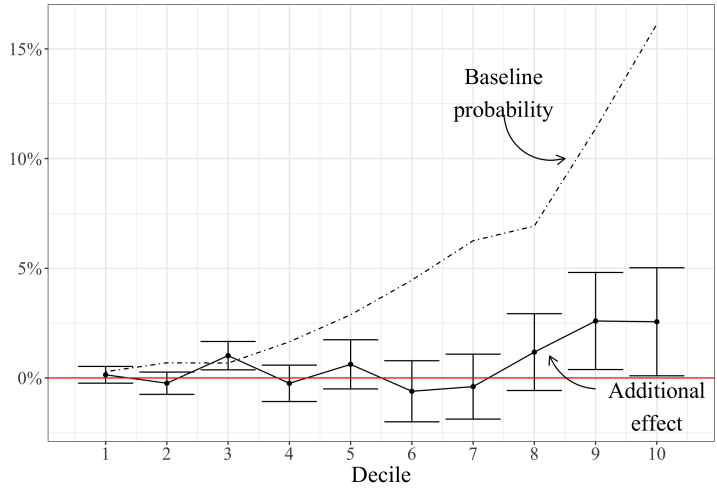


Notes: Points show the effective and predicted emigration rate for each decile by predicted baseline probability of leaving. Each cell corresponds to the average of 50 iterations where the sample is trained on a randomly chosen half of the sample and cross-validated on the other half. The sample is made of 184,946 individuals $\times$ year observations of foreigners earning over £97,500 annually in the period running from 2007 to 2009, and the outcome variable is the dummy taking the value 1 if an individual will be an immigrant in the following year. The variables used to train the model are: income (and squared income), age, sex, time since arrival (and squared and cubed), the main source of income, industry when it's available, and country of origin grouped by main regions.

Source: Authors' calculations based on HMRC administrative datasets.

Emigration rate and age for natives | Methodology

Response concentrated among people with a high baseline Pr(leaving)



Notes: This figure shows the coefficients and associated standard error from a pooled DiD model (see previous slide) Each coefficient is associated with a population in the k-decile by estimated probability of leaving using a random forest model trained on four years of data. The dotted line represents the baseline probability of leaving in the treatment group post period.

Source: Authors' calculations based on HMRC administrative datasets.

Prediction with 4 years of training | DiD by tercile | Specification

2012 PIT tax reform in France

- Income tax in France in 2012: progressive, includes earnings but excludes dividends and gains.
- **The reform**
  - Changes in rates. Freeze of thresholds. **Top rate from 41% to 45%** above €150,000, with additional exceptional contribution' for individuals with labour income over €1,000,000.
  - Changes in tax base: **inclusion of dividend income and other savings income**, tightening of rules on tax deductions.
- **Treatment:** UK taxpayers of French nationality earning over £100,000 contemporaneously.
- **Control:** UK taxpayers of German, Dutch or Belgian nationality earning over £100,000 contemporaneously.
- All EU members faced the same immigration rules: free movement of individuals for the entire period considered. None of the countries in the control group experienced tax changes affecting top earners during the period considered (2006-2017).

Estimating the migration response by income

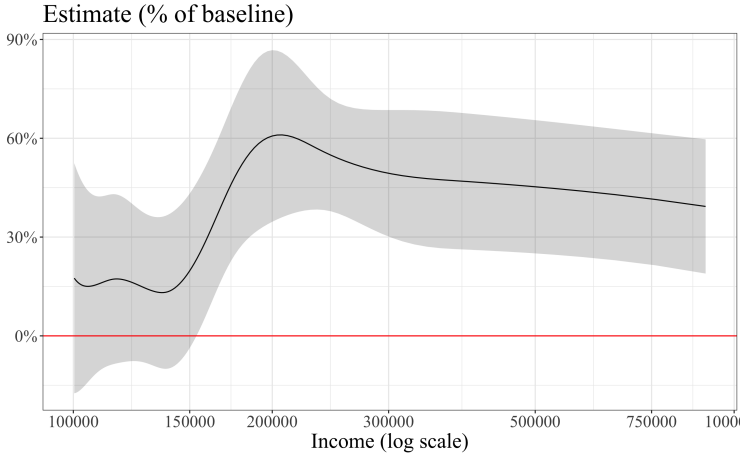
- French reform set-up allows us to get treatment heterogeneity by income.
- Use general additive model (GAM) to predict local emigration rate on the income distribution for treatment and control group, pre- and post-period.
- Three years of sample before and after the reform.
- A local DiD estimator is computed by taking the local value of the double difference following:

$$\mathbb{P}(X_{g,t}|y) = \alpha_y \mathbb{1}(Treat = 1) + \delta_y \mathbb{1}(Post = 1) + \beta_y \mathbb{1}(Post = 1) \times \mathbb{1}(Treat = 1) + \varepsilon_{y,g,t}$$

With $\mathbb{P}(X_{g,t}|y)$ the probability that an individual will have arrived the previous year from country g conditional on income, g the group, t the time, and $\varepsilon_{y,g,t}$ the income-specific error term.

- Inference from assuming independence of error terms.

Emigration response flat at higher income levels, despite rising EATR

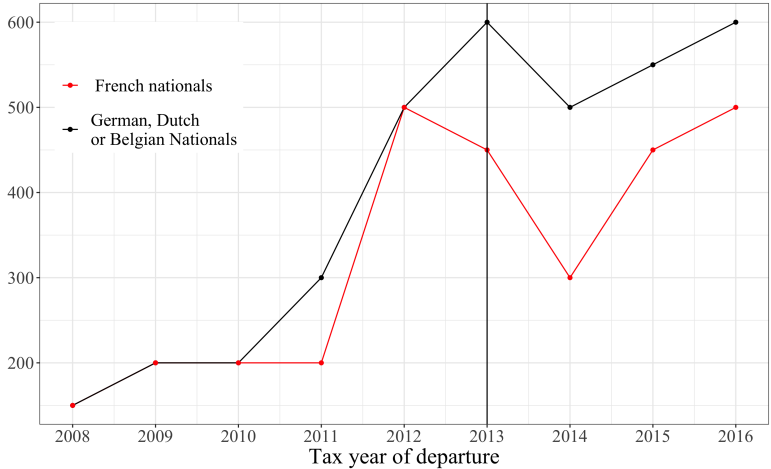


Notes: This figure shows the local value of the double difference between the coefficients of the local share of immigrants in the treatment and control group pre- and post-reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously. Immigrants are those who have arrived in the UK for the first time in the year before they are observed.

Source: Authors' calculations based on HMRC administrative datasets. Treatment and control

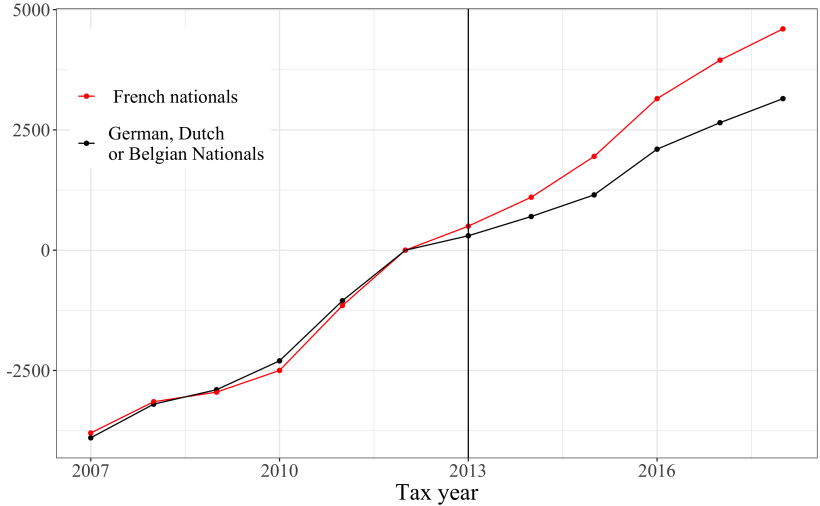
Immigration and short-term stock response

Immigration response to French reform



Notes: This figure shows the number of emigrants from the control and treatment group between 2008 and 2016. Values are rounded to the nearest 50. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously. Emigrants are those who have left the country in the year after they are observed.
Source: Authors' calculations based on HMRC administrative datasets.

Effect on stock: short to medium term evidence



Notes: This figure shows the stock of taxpayers from the control and treatment group between 2006 and 2016. Values are rounded to the nearest 50. Treatment is made French UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian UK residents earning over 100k contemporaneously. The values have been shifted to be equal to 0 in the year 2012.
Source: Authors' calculations based on HMRC administrative datasets.

Effect on stock: short to medium term evidence

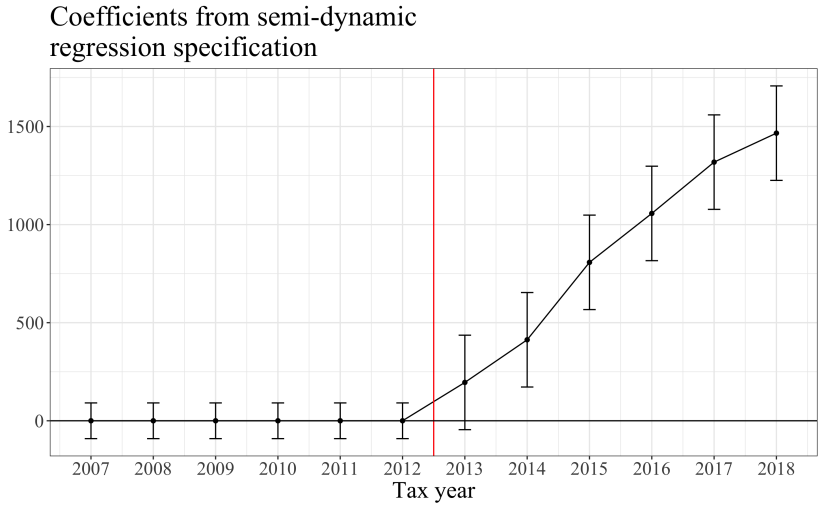
We use the absence of pre-trend to estimate a semi-dynamic diff in diff specification

$$N_{g,t} = \gamma_t + \alpha \mathbb{1}(g = FRA) + \sum_{t=2013}^{2017} \beta_t \mathbb{1}(T = t) \times \mathbb{1}(g = FRA) + \varepsilon_{g,t}$$

With $N_{g,t}$ the size of group $g \in \{FRA, CONTROL\}$ in year t.

- After 5 years there are an additional 1300 French individuals in the treatment, 14% more than the counterfactual.

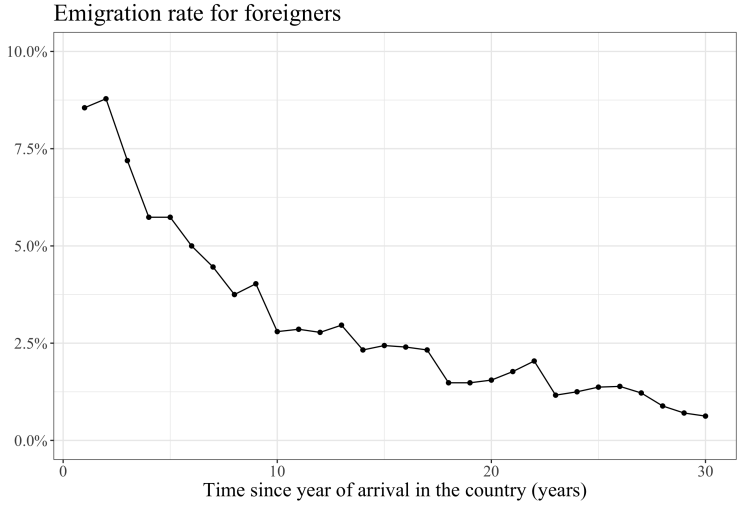
Semi-dynamic DiD stock effect:



Notes: This figure shows the coefficients and associated standard errors from the semi-dynamic DiD specification (see previous slide). Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

The long-run effect of tax-induced migration

Emigration rates decline with time spent in the country



Notes: This figure shows the emigration rate by time spent since first year of arrival in the country. The sample is made of all foreigners earning over £97.5k observed in the years 2008 to 2010 inclusive and for which date of first arrival in the UK is known. Shares are computed after rounding number of emigrants and bin size to the nearest 50.
Source: Authors' calculations based on HMRC administrative datasets.

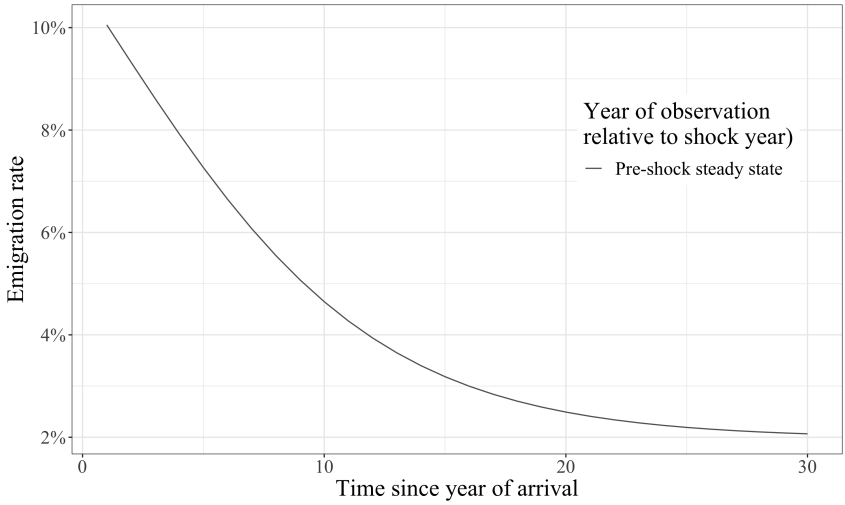
Why does emigration decline with tenure?

Two possible mechanisms

1. Two types: migrants arrive with either high or low probability of leaving.
   - as high types leave, composition changes, so aggregate migration rate falls
   - tax rise affects only high types: we saw no response for low types
2. Adaptive preferences: migrants arrive with common Poisson probability of leaving, and this probability declines.
   - can microfound this via a model where migrants build connections and take root the longer they stay, so the fixed cost of leaving rises. Model
   - tax rise increases the probability of leaving

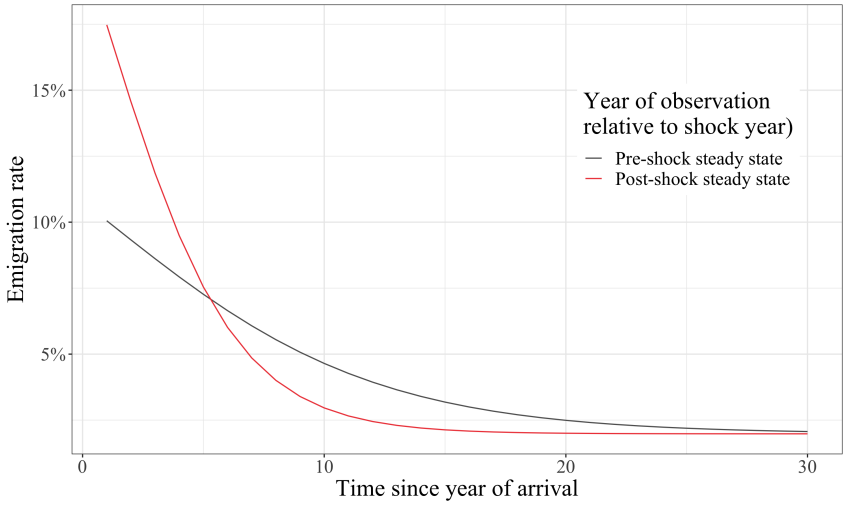
Either way, the migration response is concentrated among individuals with high ex-ante probability of leaving.

Two types scenario - Observing the shock over several years



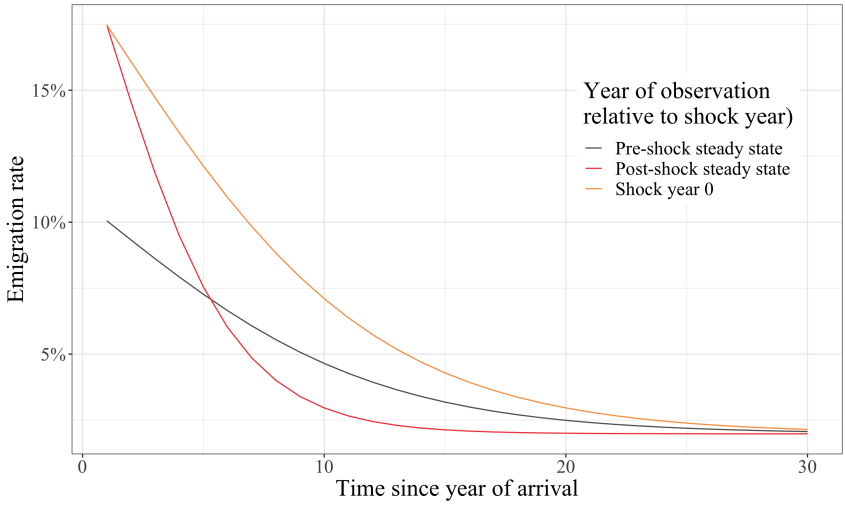
Notes: Emigration rate by year of observation relative to the shock (color of the curve) and cohort of arrival (*x* axis, in number of years since arrival in the year of observation).
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline. Details

Two types scenario - Observing the shock over several years



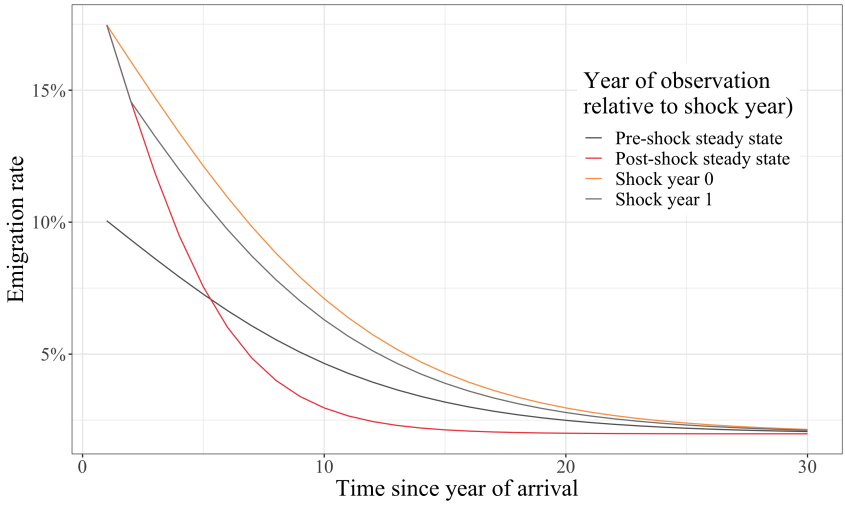
Notes: Emigration rate by year of observation relative to the shock (color of the curve) and cohort of arrival (*x* axis, in number of years since arrival in the year of observation).
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline. Details

Two types scenario - Observing the shock over several years



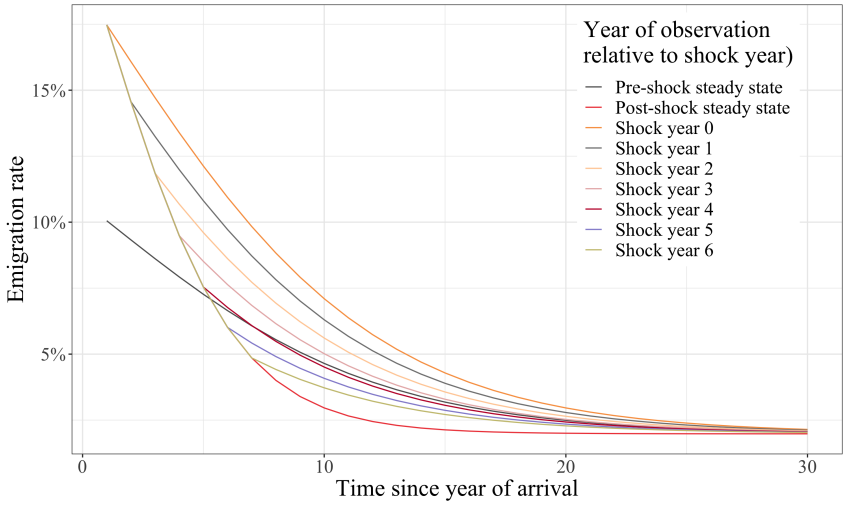
Notes: Emigration rate by year of observation relative to the shock (color of the curve) and cohort of arrival (*x* axis, in number of years since arrival in the year of observation).
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline. Details

Two types scenario - Observing the shock over several years



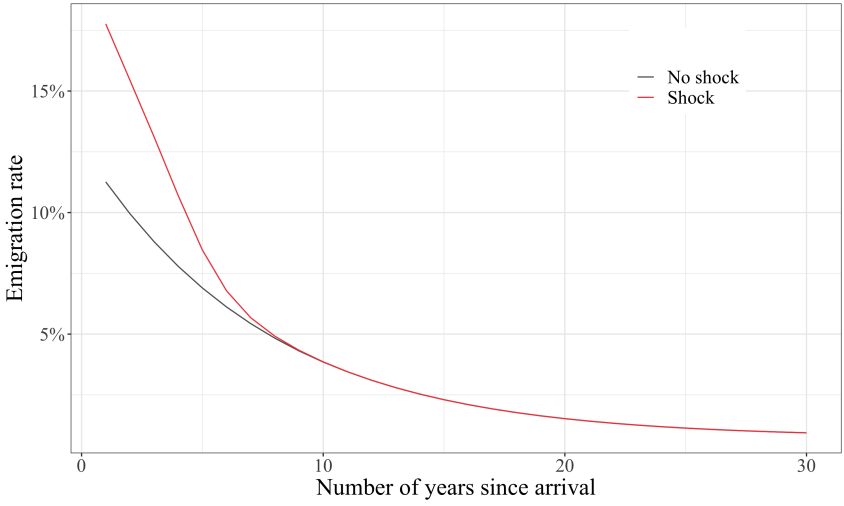
Notes: Emigration rate by year of observation relative to the shock (color of the curve) and cohort of arrival (*x* axis, in number of years since arrival in the year of observation).
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline. Details

Two types scenario - Observing the shock over several years



Notes: Emigration rate by year of observation relative to the shock (color of the curve) and cohort of arrival (*x* axis, in number of years since arrival in the year of observation).
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline. Details

Adaptive scenario - Effect of shock on emigration rate

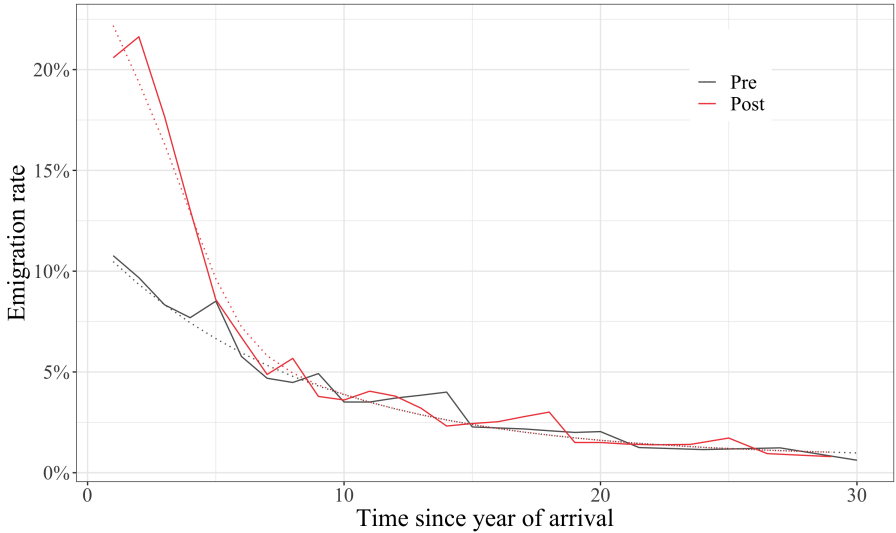


Notes: Emigration rate by year of observation relative to the shock (color of the curve) and cohort of arrival (x axis, in number of years since arrival in the year of observation).
Source: Simulated data based on a population with an initial emigration rate of 12%, $\lambda = 0.14$, a terminal emigration rate of 1%, a shock of magnitude 70% and an inflection point at 4.5 years. Details

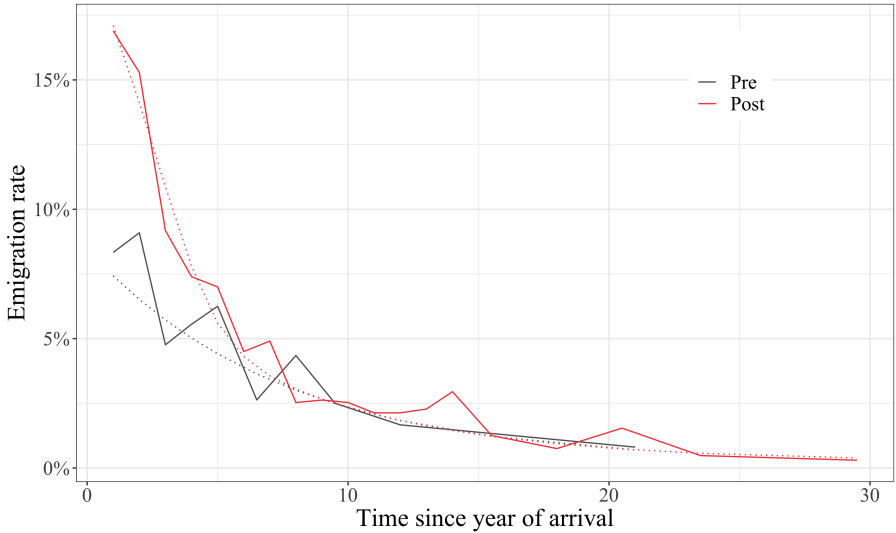
Structural estimation

- Descriptively, see no evidence of post-reform emigration rate falling below pre-reform after some years.
  - Suggestive that adaptive preferences more likely than two types model.
- To formally test, and study long run implications, we estimate the parameters in a model which nexts both mechanisms.
- Estimate via Simulated Method of Moments
  - baseline specification: 120 moments - emigration rate by tenure and time since reform.
  - five model parameters.
  - simulate moments under different parameter combinations, and select those which minimise absolute error with observed moments.
  - estimate the parameters separately at three different bands of income: £120,000-£150,000, £150,000-£225,000 and over £225,000.
- Best fit model puts all weight on a single type, and supports adaptive preferences.

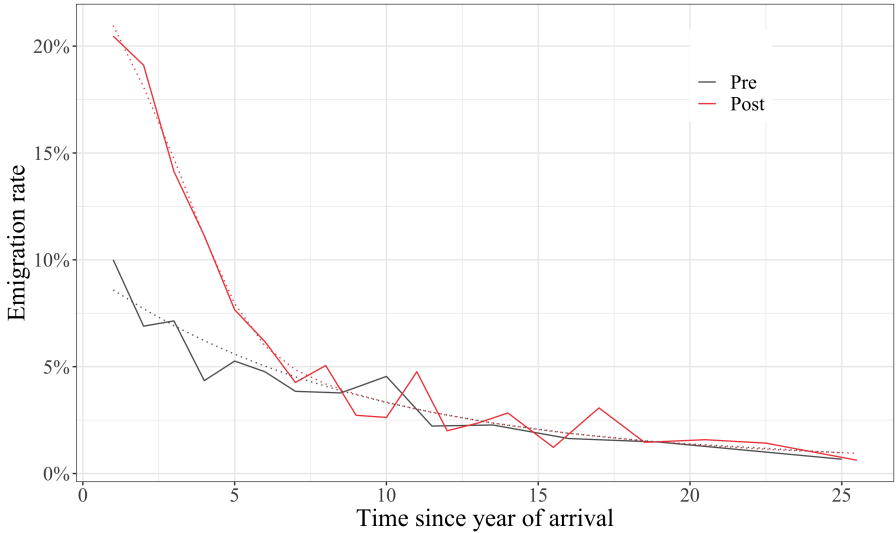
Fit for different sub-populations: Full population



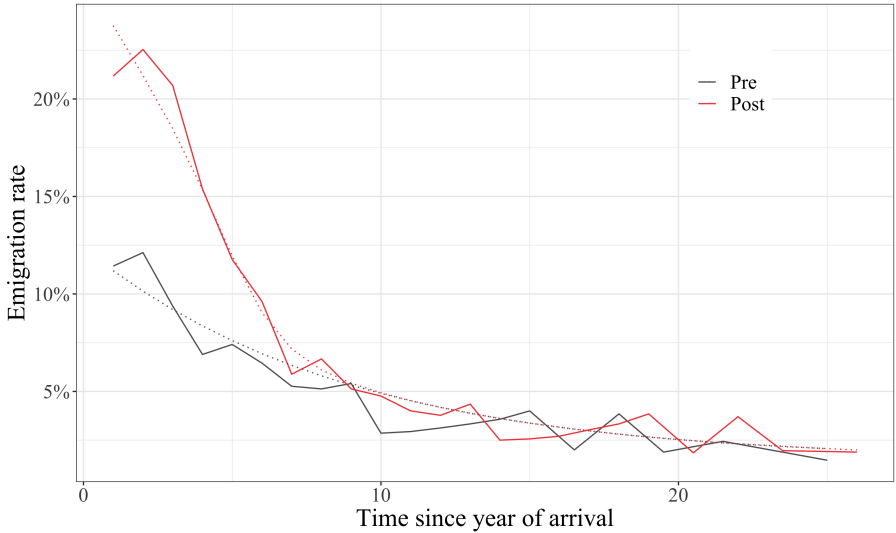
Fit for different sub-populations: 120k-150k



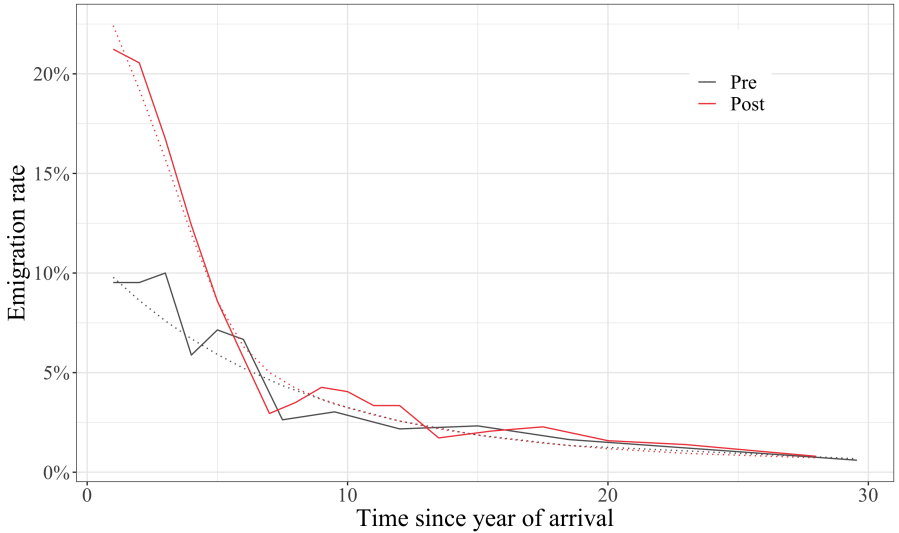
Fit for different sub-populations: 150k-225k



Fit for different sub-populations: 225k+



Fit for different sub-populations: 165k-225k (treatment group)



Structural estimation results

Parameter	Income range		
	120-150k	150-225k	225k+
Baseline parameters			
Baseline emigration rate	0.030	0.034	0.054
Rate results			
Aggregate semi-elasticity	0.49	0.45	0.28
Aggregate elasticity	16.52	13.20	5.16
Average elasticity	7.29	6.12	2.92
Stock results			
Total stock elasticity	9.19	7.60	3.76

Notes: Paremeters of three income groups based on the results of the structural estimation.
Source: Authors' calculations based on HMRC administrative datasets. Dynamic effects

Robustness: structural vs. reduced form

Comparing reduced form estimates in the structural and quasi-experimental context.

	Parametric estimation	Parametric scaled for first stage	Quasi-experimental (QE)	p-value for equal coefficients
Short stayers (<5 years)	0.0639	0.0330	0.0345	0.81
Medium stayers (5-14 years)	0.0015	0.0007	0.0007	0.85
Long stayers (>15 years)	0.0000	0.0000	0.0018	0.24

Notes: This table compare the short term emigration estimates from the structural estimation to the ones from the quasi-experimental approach on the same sample of migrants earning between £165,000 and £225,000 annually, for three different length of stay groups. The first stage is re-scaled so that point estimates are comparable. Indeed, the quasi-expermiental estimation's difference-in-differences approach excludes the effect of the personal allowance removal, which means that the first stage is lower.
Source: Authors' calculations based on HMRC administrative datasets.

Conclusion

Conclusion

- From a fiscal standpoint, migration response to tax is small.
  - Most top earners are natives, even in very international countries, who have very low semi-elasticities.
  - Even among migrants, most are not very responsive.
- Importantly, semi-elasticity declines with income.
  - Those paying the most tax are least responsive.
- Most responsive group are recent arrivers.
  - Motivates special impatriate regimes, so this group are less affected by wider top tax rate.

Thank you!

Work in progress – feedback is most welcome.

Email: a.advani.1@warwick.ac.uk

Appendix

UK reform - Balance table

	Natives		*Migrants*	
	Control	Treatment	Control	Treatment
Main source of income				
Employment	62%	61%	65%	66%
Investment	8%	8%	8%	7%
Owner-manager	5%	6%	3%	2%
Partner	13%	15%	13%	14%
Pension	5%	4%	4%	2%
Self-Employment	7%	7%	8%	8%
Male	80%	84%	76%	81%
Share of investment income	14%	15%	12%	11%

Notes: This table compares the characteristics of treatment and control groups in 2009. The control group is made of people individuals between 120k and 135k and the treatment group of individuals earning between 165k and 225k. For natives, there are 74,400 93,500 individuals in the treatment and control group respectively (rounded to the nearest 100). Those values are 11,900 and 16,900 for foreigners.
Source: Authors' calculations based on HMRC administrative datasets.

ATE by decile of baseline probability of leaving

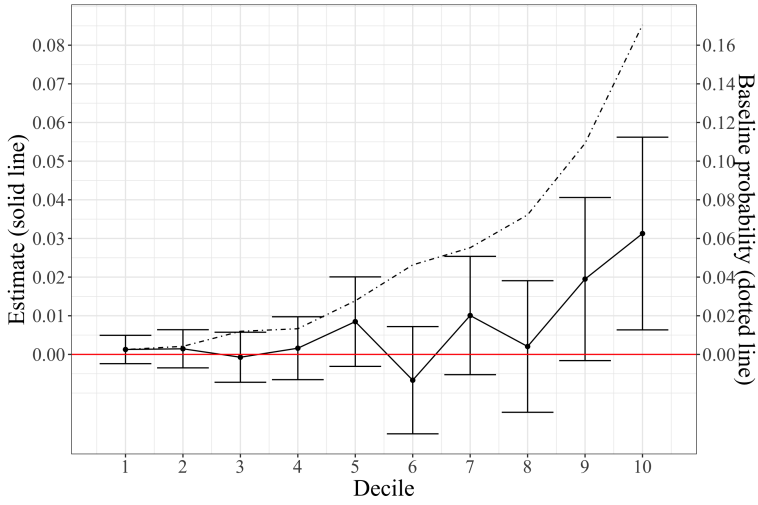
- We divide each year $\times$ treatment status group of observations into deciles based on our predicted probability of leaving.
- For each decile, we run an individual-level pooled DiD OLS regression:

$$E_{i,t} = \alpha X_{i,t} + \sum_{\substack{t=2006 \\ t \neq 2009}}^{2013} \beta D_i . T_t + \delta D_i + \gamma_t + \varepsilon_{i,t} \quad (3)$$

- We use the predicted emigration rate from the random forest algorithm as control, encompassing the effect of all variables on emigration.
- We plot the values of the coefficient of the treatment effect for each decile.
- We also plot a dynamic DiD specification to control for pre-trends that could drive the results.

Back

Heterogeneity by baseline - 4 years of data training



Notes: This figure shows the coefficients and associated standard error from a pooled DiD model. Each coefficient is associated with a population in the k-decile by estimated probability of leaving using a random forest model trained on three years of data. The dotted line represents the baseline probability of leaving in the treatment group post period.

Source: Authors' calculations based on HMRC administrative datasets.

Back

Measurement

Emigration

- Departure year is year t if individual is resident in year t but absent from records or filing as non resident in $t+1$ and $t+2$.
- Income is measured in year $t-1$ to get full year of income.

Immigration

- Immigration defined as *first year of arrival in the UK*.
- Making use of multiple sources: Migrant Worker Scan (MWS), analysis of National Insurance Number (NINo) prefixes and tax filings (SA 109).
- Where immigration year is t, income is measured in year $t+1$ to get full year of income.

Back

Predicting the baseline probability of leaving

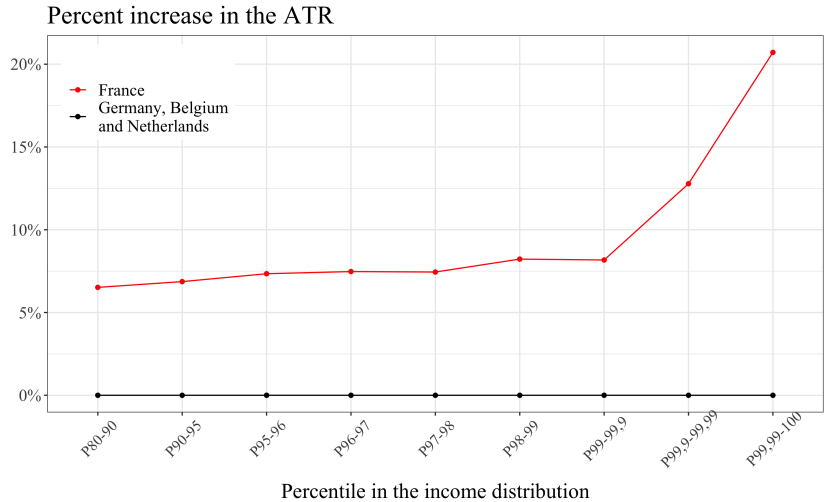
- Individual probability of leaving can never be observed at the individual level.
- We train a random forest model to predict the probability of leaving using individual characteristics on pre-reform years.
- The performance of our algorithm is tested by binning out-of-sample individuals by quantile of predicted baseline probability and comparing the group average with the predicted average.

Back

Do some natives respond?

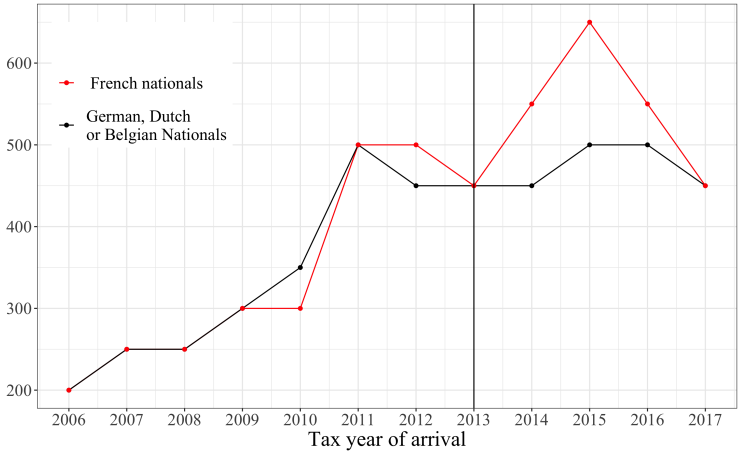
- Findings from UK reform suggest no meaningful or significant response from the native population as a whole.
- However, difficult to distinguish from this approach if there is a small minority who respond but cannot be separated from noise in this aggregate results.
- To get at this, we use the French reform set-up to look at arrivals from France (emigration of French natives) in response to Hollande reforms.

First stage for the French reform



Notes: Percentage increase in the ATR after the Hollande reform. Each point represents the ratio of the new tax average rate and a counterfactual measure in the absence of reform for each quantile in the population. There are no reforms in the control group so all values are equal to 0.
Source: Bozio, Fabre, Goupille and Lafféter (2012) IPP note n°2 (2012).

Some natives do respond: emigration from French reform (observed as French nationals vs controls arriving in UK)

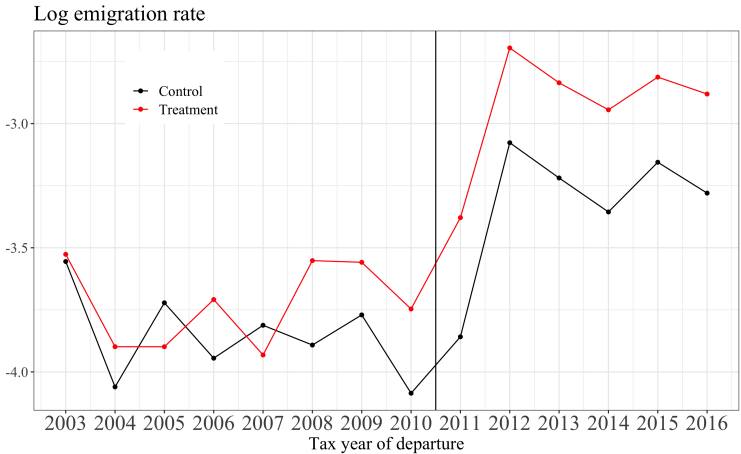


Notes: This figure shows the number of immigrants from the control and treatment group between 2006 and 2016. Values are rounded to the nearest 50. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously. Immigrants are those who have arrived in the UK for the first time in the year before they are observed.

Source: Authors' calculations based on HMRC administrative datasets

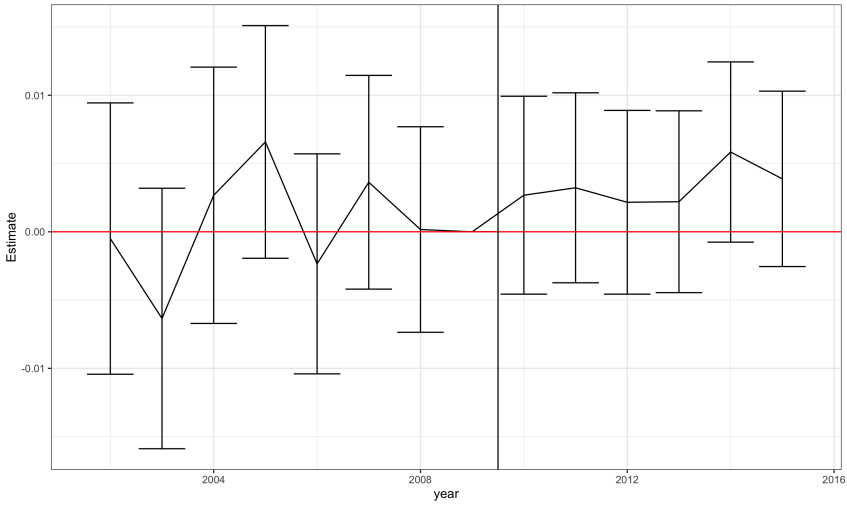
Back

Robustness - log specification DiD



Notes: This figure shows the log share of foreigner emigrants in the control and treatment group between 2003 and 2016. Shares are computed after rounding number of emigrants and population size to the nearest 50. Control group includes foreign taxpayers earning £165,000-225,000 contemporaneously, treatment group includes foreign taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

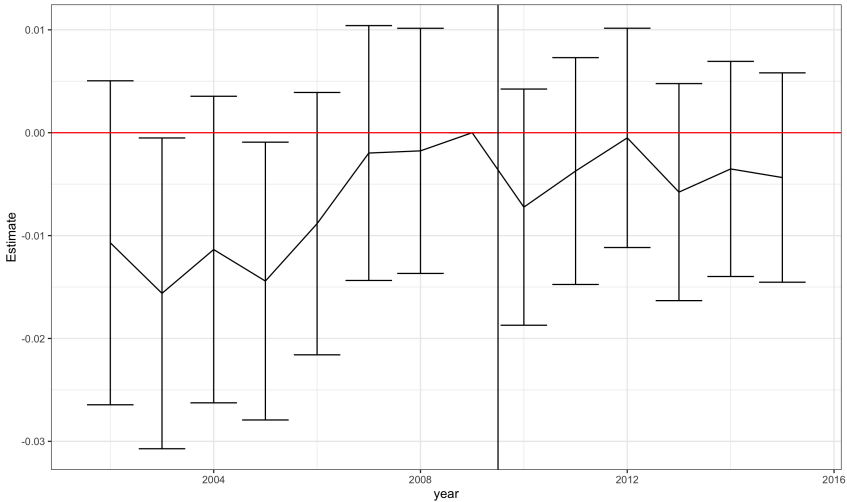
No response from lowest tercile



Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from the lower tercile of predicted emigration rate.

Source: Authors' calculations based on HMRC administrative datasets. Back

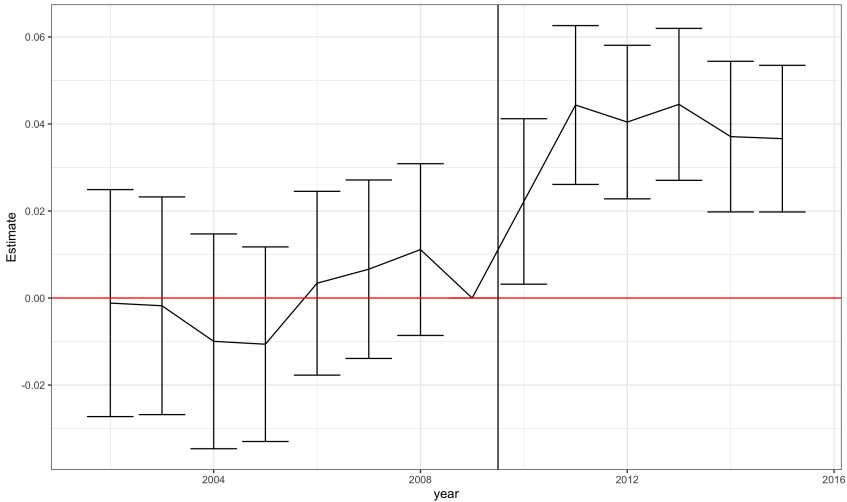
No response from middle tercile



Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from the middle tercile of predicted emigration rate.

Source: Authors' calculations based on HMRC administrative datasets. Back

Response from highest tercile

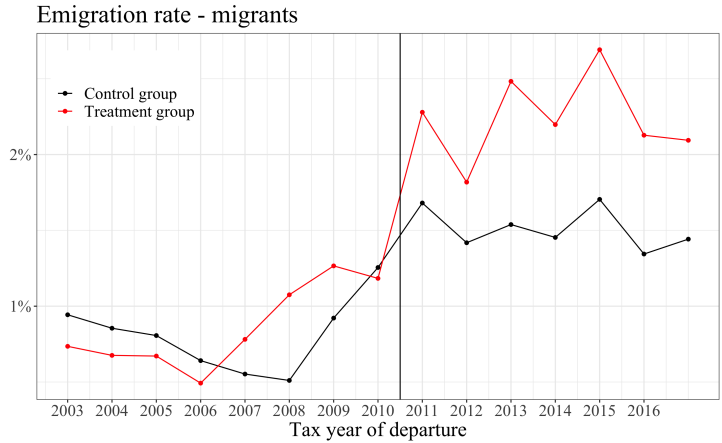


Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from the higher tercile of predicted emigration rate.

Source: Authors' calculations based on HMRC administrative datasets.

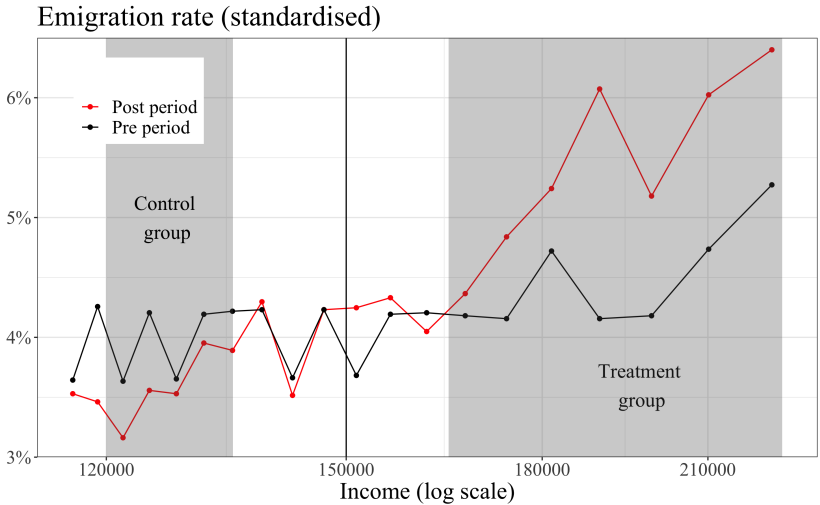
Back

Robustness - lower-bound emigration measure



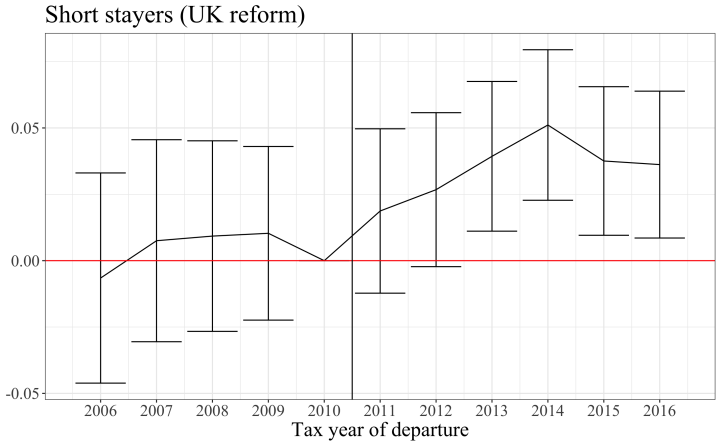
Notes: This figure shows the share of foreigner emigrants in the control and treatment group between 2003 and 2016, where emigration is defined as individuals who are resident filing taxpayers in t and non-resident filing taxpayers in $t+1$. Shares are computed after rounding number of emigrants and population size to the nearest 50. Control group includes foreign taxpayers earning £165,000-225,000 contemporaneously, treatment group includes foreign taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Treatment intensity and income: UK reform



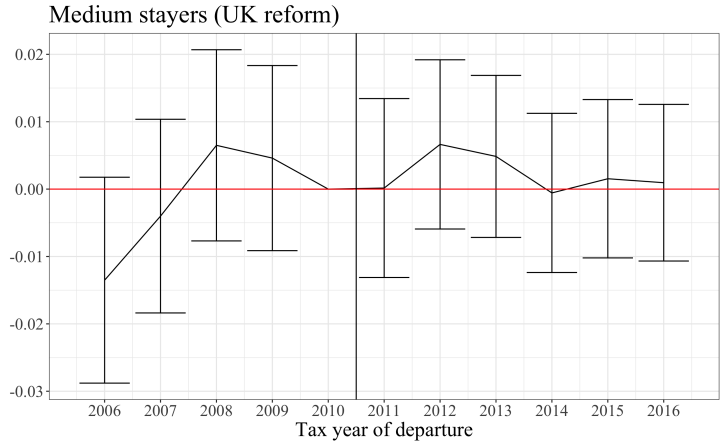
Notes: This figure shows the average emigration rate by income bin computed for three years pre and post reform (2008-2010 and 2011-2013). Values in the control group are standardised so that they match the treatment group's value for bin at £150k.
Source: Authors' calculations based on HMRC administrative datasets.

DiD by length of stay UK - short stayers



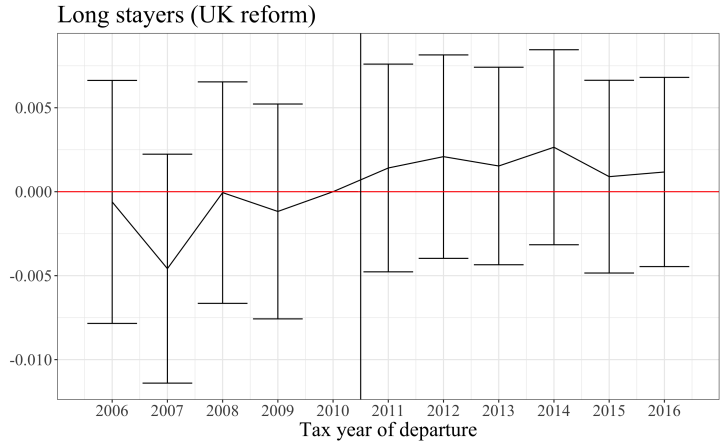
Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from individuals staying for 4 years or less. Control group includes native taxpayers earning £165,000-225,000 contemporaneously, treatment group includes native taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

DiD by length of stay UK - medium stayers



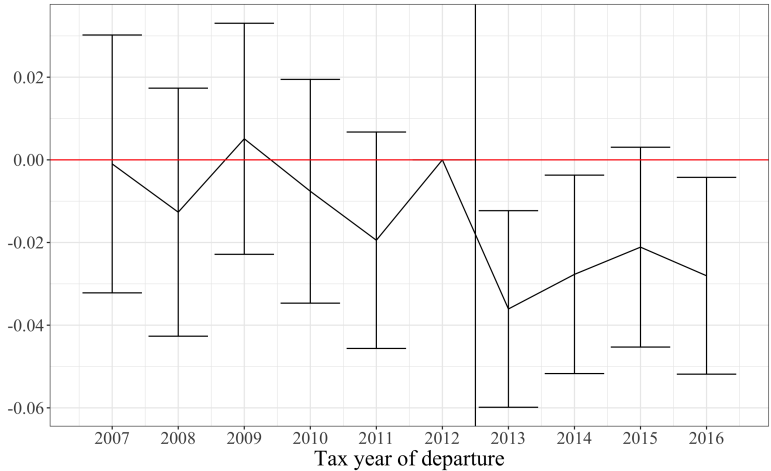
Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from individuals staying from 5 to 14 years. Control group includes native taxpayers earning £165,000-225,000 contemporaneously, treatment group includes native taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

DiD by length of stay UK - long stayers



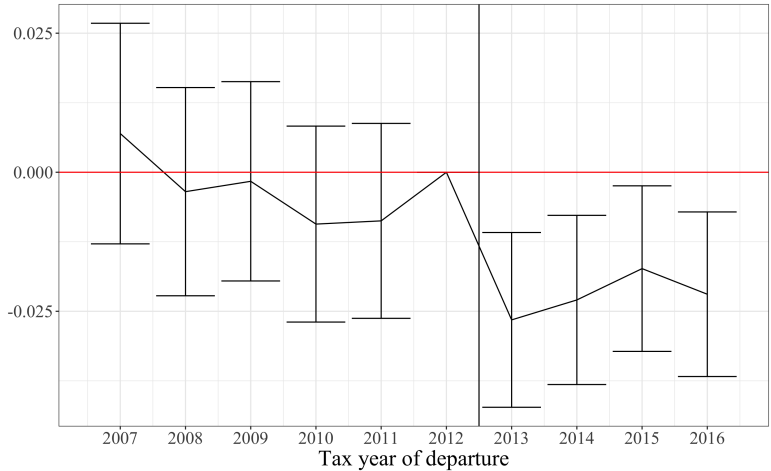
Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from individuals staying for 15 years or more. Control group includes native taxpayers earning £165,000-225,000 contemporaneously, treatment group includes native taxpayers earning £120,000-135,000 contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

DiD by length of stay France - short stayers



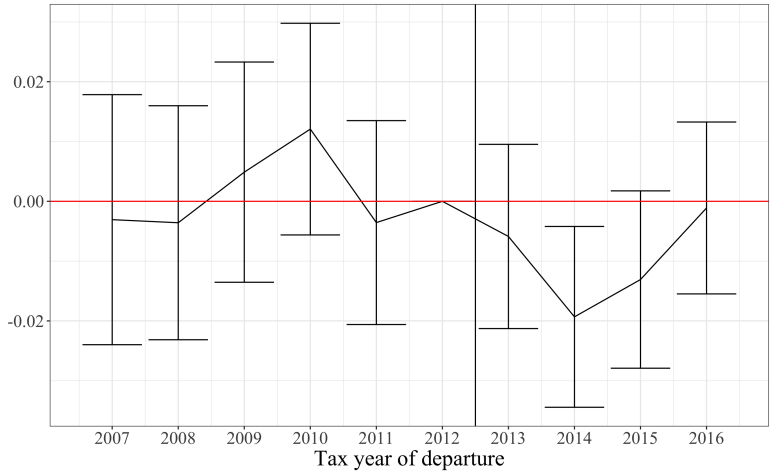
Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from individuals staying for 4 years or less. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

DiD by length of stay France - medium stayers



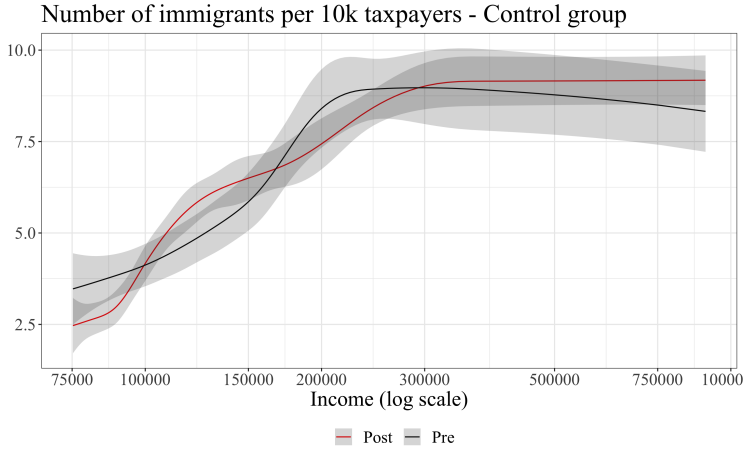
Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from individuals staying from 5 to 14 years. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

DiD by length of stay France - long stayers



Notes: This figure shows the coefficients and associated standard error from a dynamic DiD model regressing the probability of emigration on our instrument and controling for the baseline probability of emigration. The population is from individuals staying for 15 years or more. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Emigration response heteregoneity by income: control

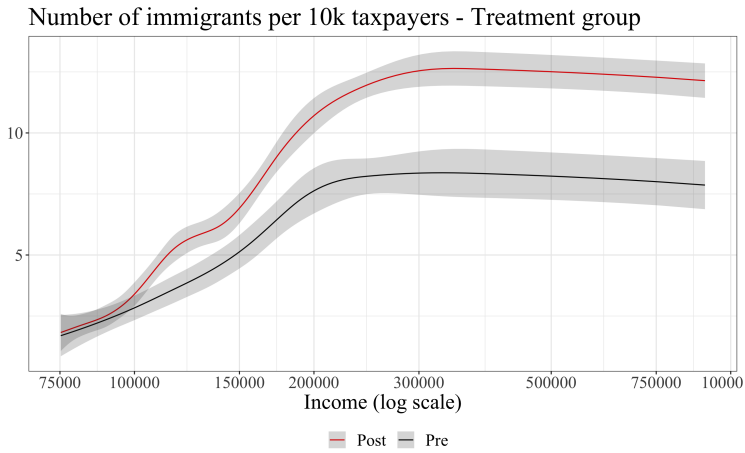


Notes: This figure shows the local value of the coefficients of the local share of immigrants in the control group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.

Source: Authors' calculations based on HMRC administrative datasets.

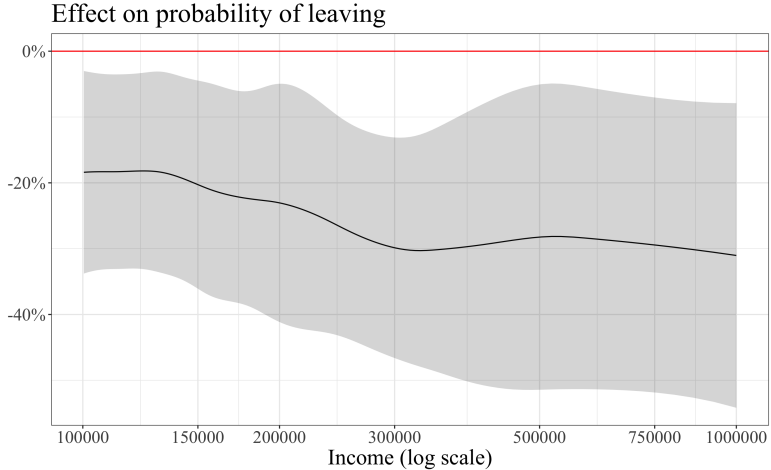
back

Emigration response heteregoneity by income: treatment



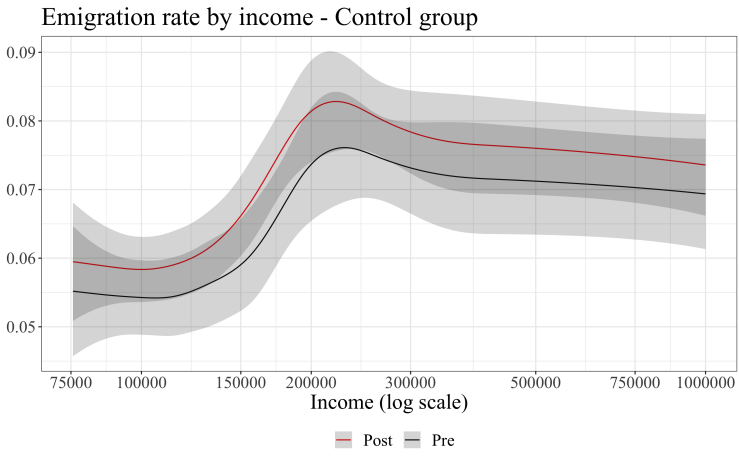
Notes: This figure shows the local value of the coefficients of the local share of immigrants in the treatment group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Treatment intensity and income – French reform (immigration response: departures from UK to France)



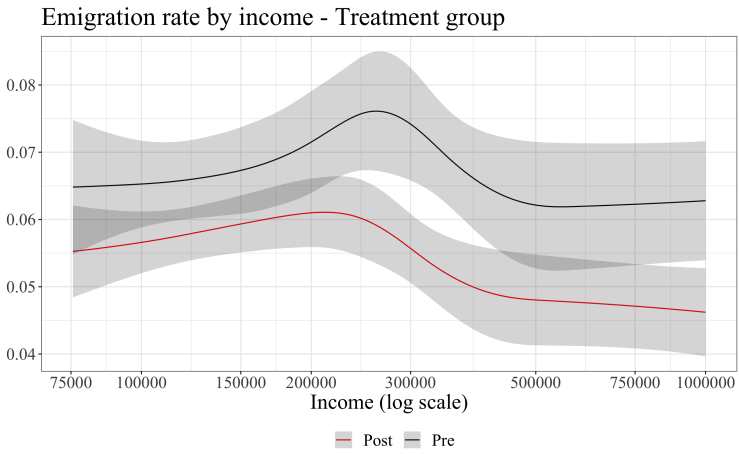
Notes: This figure shows the local value of the double difference between the coefficients of the local share of emigrants in the treatment and control group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Immigration response heteregoneity by income: control



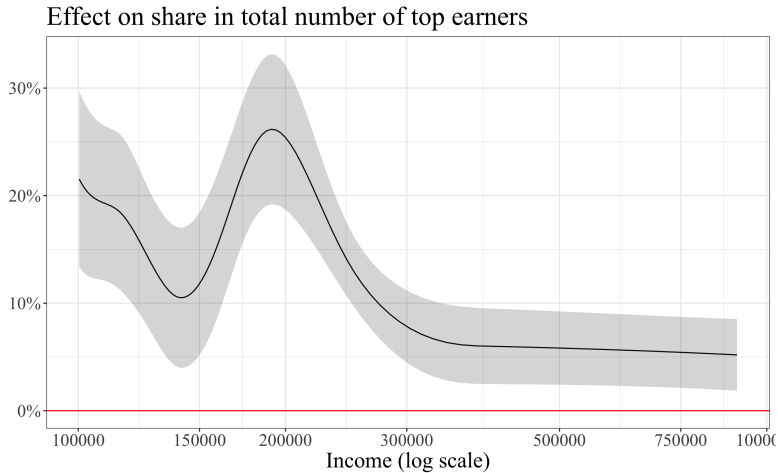
Notes: This figure shows the local value of the coefficients of the local share of emigrants in the control group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Immigration response heteregoneity by income: treatment



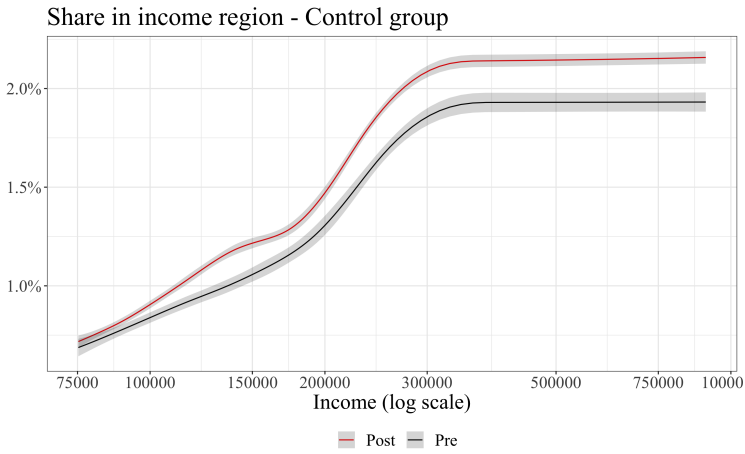
Notes: This figure shows the local value of the coefficients of the local share of emigrants in the treatment group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Treatment intensity and income – French reform (stock response)



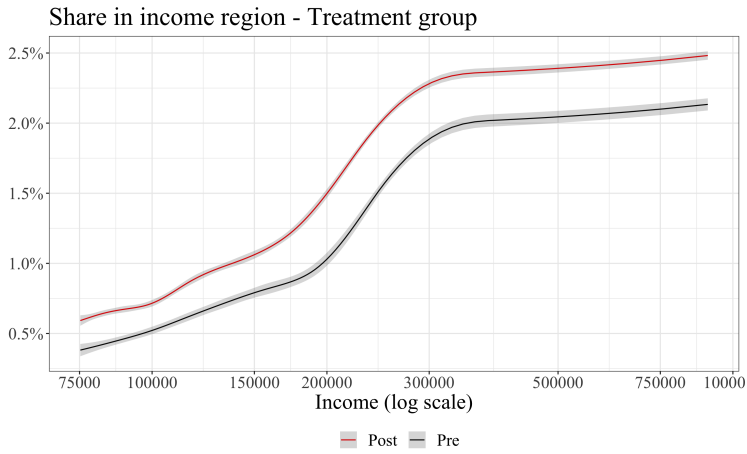
Notes: This figure shows the local value of the double difference between the coefficients of the local share of foreigners in the treatment and control group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Stock response heteregoneity by income: control



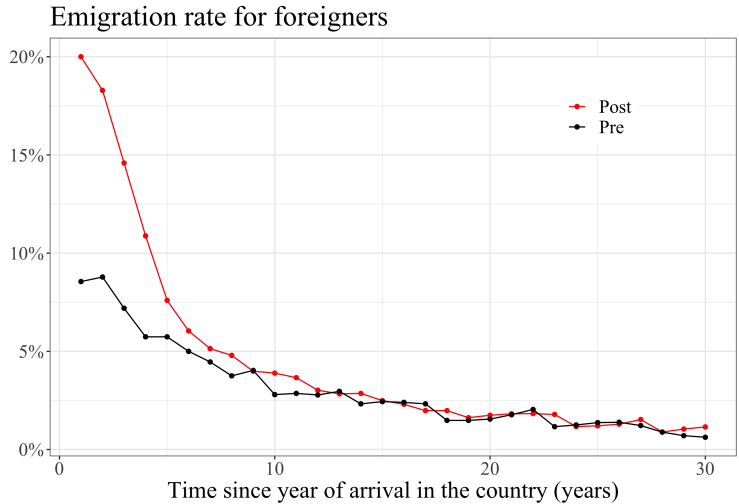
Notes: This figure shows the local value of the coefficients of the local share of foreigners in the control group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Stock response heteregoneity by income: treatment



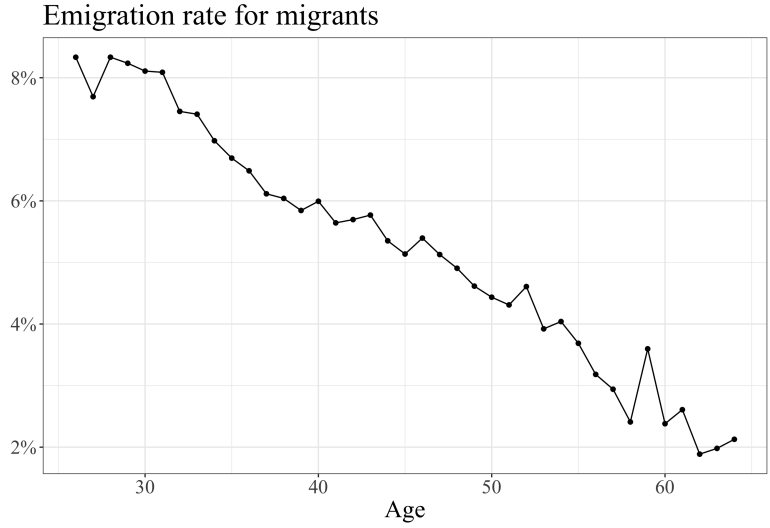
Notes: This figure shows the local value of the coefficients of the local share of foreigners in the treatment group pre and post reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian national UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Emigration rate and time spent, pre and post reform



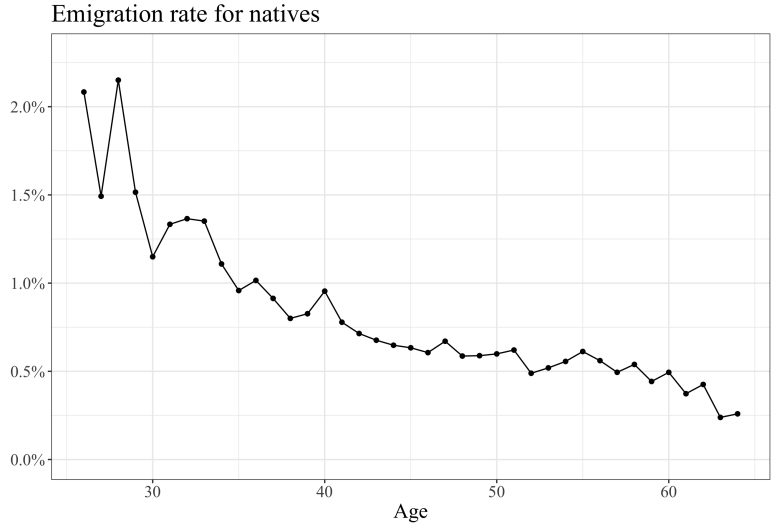
Notes: This figure shows the emigration rate by time spent since first year of arrival in the country. The sample is made of all foreigners earning over £97.5k observed in the years 2008 to 2010 inclusive for the pre period, and post 2011 for the post period, and for which date of first arrival in the UK is known. Shares are computed after rounding number of emigrants and bin size to the nearest 50.
Source: Authors' calculations based on HMRC administrative datasets.

Emigration rate and age - foreigners



Notes: This figure shows the emigration rate by age. The sample is made of all foreigners earning over 97.5£. Shares are computed after rounding number of emigrants and bin size to the nearest 50.
Source: Authors' calculations based on HMRC administrative datasets.

The determinants of migration among native top earners: age

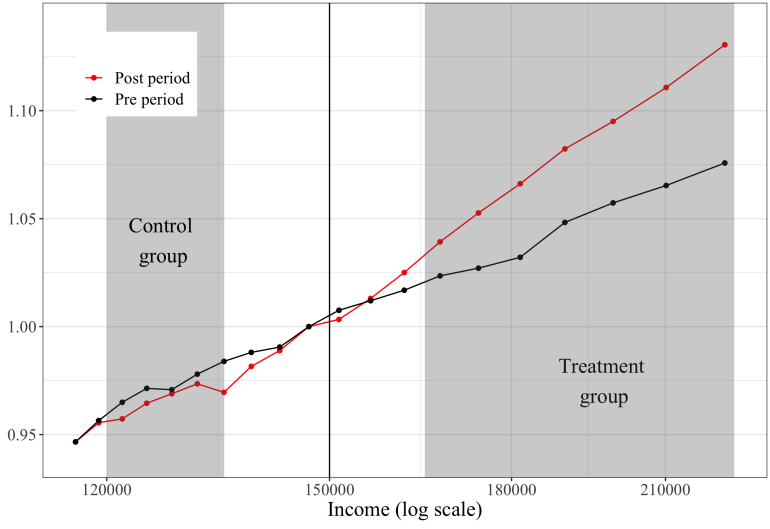


Notes: this figure shows the emigration rate by age. The sample is made of all natives earning over £. Shares are computed after rounding number of emigrants and bin size to the nearest 50.

Source: Authors' calculations based on HMRC administrative datasets.

Back

First stage of UK reform: ATR by income pre and post reform

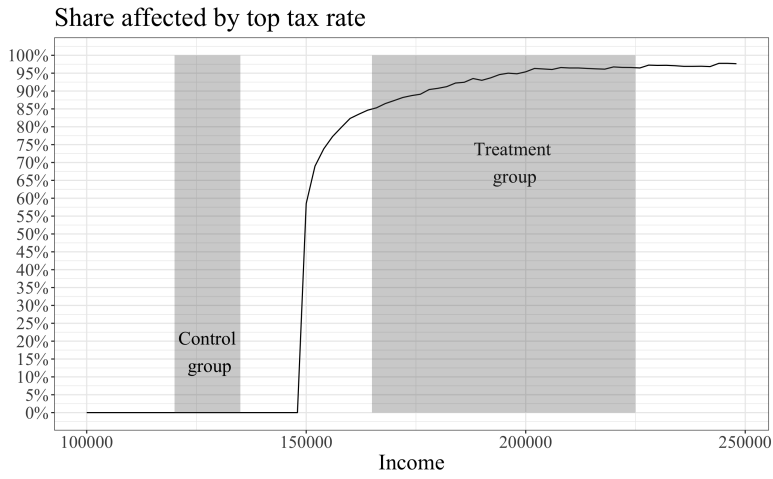


Notes: This figure shows the average ATR by income bin computed for three years pre and post reform (2008-2010 and 2011-2013). Results are normalised by ATR at £150k for both periods.

Source: Authors' calculations based on HMRC administrative datasets.

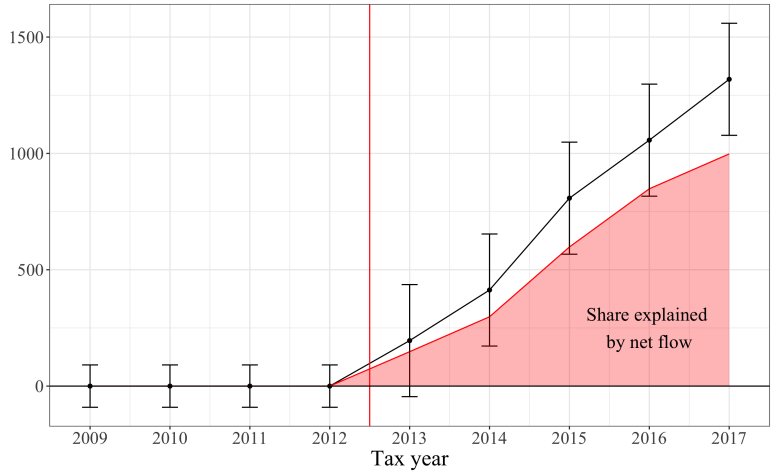
back

Treatment intensity



Notes: This figure shows the share of taxpayers affected by the 50% top marginal tax rate by income. The population is made of all individuals who file a tax return, and the reason why it is not 0% before the threshold and 100% just after is because of tax deductions.
Source: Authors' calculations based on HMRC administrative datasets.

Semi-dynamic DiD stock effect - share explained by net flow



Notes: This figure shows the coefficients and associated standard errors from the semi-dynamic DiD specification (see previous slide). The red area corresponds to the share of the treatment effect that can be attributed by the change in immigration and emigration flow. It is obtained by adding the difference in number of immigrants and emigrants after the reform. Treatment is made of French national UK residents earning over 100k contemporaneously and control is made of German, Dutch and Belgian nationals UK residents earning over 100k contemporaneously.
Source: Authors' calculations based on HMRC administrative datasets.

Effect not scaled by first stage

We present the results from our structural estimation for three different income group. In this table, results are shown scaled to a 1% increase of the net-of-t.

	120-150k	150-225k	225k+
Total population	17.433	16.290	12.791
Baseline emigration rate	0.03	0.034	0.053
Aggregate semi-elasticity	0.01	0.02	0.03
Max semi-elasticity	0.10	0.12	0.13
Aggregate elasticity	0.44	0.58	0.56
Average elasticity	0.19	0.27	0.31
Max elasticity	1.22	1.44	1.19
Share responding	0.27	0.32	0.41
Long run stock semi-elasticity	0.12	0.15	0.12
Long run elasticity	0.26	0.36	0.45
Total stock elasticity	0.24	0.34	0.40

Model Appendix

A simple model of location choice

- Two countries S (for stay) and M (for move). For individual i and $J \in \{S, M\}$

$$U_{J,i,t} = \log\left((1-\tau_J)y^J_{i,t}\right) + v_J(X_{i,t}) + \varepsilon_{J,i,t}$$

- with $U_{J,i,t}$ the utility for i in country J and at time t. t is the time since arriving in country S, $(1-\tau)$ is the net of tax rate on income $y^J_{i,t}$, $v_J(X_{i,t})$ is the utility derived from staying in country J in time t and $\varepsilon_{J,i,t}$ is the error term.
- We define $U_{i,t}$ the utility differential:

$$U_{i,t} = \left[\tilde{y}^M_i - \tilde{y}^S_i\right] + \left[v_M(X_{i,t}) - v_s(X_{i,t})\right] + \left[\varepsilon_{M,i,t} - \varepsilon_{S,i,t}\right] = U(X_{i,t}) + \varepsilon_{i,t}$$

Back

A simple model of location choice

- The decision to move to country M in t is made if

$$U_{i,t} > C_{i,t} \iff \varepsilon_{i,t} > C_{i,t} - U(X_{i,t})$$

  Where $C_{i,t}$ is the fixed moving cost.
- If we remain agnostic on the relationship between the different variables, we can write, for individual with characteristics X (dropping subscript t):

$$P\Big(M|X\Big) = P\Big(\varepsilon_X > L(X)\Big)$$

  Where $L(X) = C_{i,t} - U(X_{i,t})$ is the net cost of moving, and ε_X the random variable that is the error term of for individual with characteristics X

Emigration rate

- Conditional on X, the associated empirical moment to $P\Big(M|X\Big)$ is the emigration rate E_X, defined as:

$$E_X = \int_{L(X)}^{+\infty} f_X(t)dt = 1 - F_X(L(X)) \tag{4}$$

  Where F_X is the CDF of ε_X
- The difficulty is that this value depends not only on the net cost of moving $L(X)$, but also on the type-specific distribution of the error term

Response to a tax shock

- Let's consider the response to change in the log net-of-tax rate.

$$\begin{aligned}\frac{\partial E_X(\tau_S)}{\partial(1-\tau_S)} &= \frac{\partial(1-F_X(L(X,\tau_S))}{\partial(1-\tau_S)} \\ &= -f\big(L(X,\tau_S)\big) \times \frac{\partial L(X,\tau_S)}{\partial(1-\tau_S)}\end{aligned}$$

- Since $\frac{\partial L(X)}{\partial \log(1-\tau_S)} = -1$ in our specification, it follows that

$$\frac{\partial E_X(\tau_S)}{\partial \log(1-\tau_S)} = f\big(L(X,\tau_S)\big)$$

- The magnitude of the shock is entirely determined by the value of the PDF in $L(X)$

Shape of F_X

- The shape of the error term plays a crucial role, as the pair of values $F_x(L(X), f(L(x))$ conjointly determines the emigration rate and the sensitivity of the individual to income shocks (including tax shocks). In particular the local emigration elasticity is given by the value:

$$\eta_X = \frac{\partial \log\left(E_X(\tau_S)\right)}{\partial \log(1-\tau_S)} = \frac{f\big(L(X,\tau_S)\big)}{1-F\big(L(X,\tau_S)\big)} \tag{5}$$

- Assumptions of constant elasticity or constant semi-elasticity imply respectively that ε_X follows an exponential or uniform distribution.
- $1 - F_X(L(X))$ represents mass of individuals for which staying is more costly than leaving, and $f_X(L(X))$ represents the mass of marginal individuals who are indifferent between M and S.

Estimating the parameters

- Because this general framework allows both the net cost and the distribution to be type specific, it is very difficult to estimate without making assumptions on the functional form and the distribution of parameters over time.
- If we consider a population Ω associated with a measure W of 1 (the weight on each value of X), and assume the (F_X) family is "continuous" (i.e. $\forall Y \in \mathbb{R}, F_X(Y)$ is continuous in X), the aggregate semi-elasticity $\tilde{E}_\Omega$:

$$\tilde{E}_\Omega = \int_\Omega f_X\big(L(X)\big)\, W(X)\, dX \tag{6}$$

- We can show that when (F_X) and L are well behaved, $\tilde{E}_\Omega$ converges towards $f(X^*)$ where X^* is the barycentre of Ω as Ω 'shrinks' towards X^*.

ATE and LATE

- In practice, we can't observe all dimensions of X, and even for the variables we can observe, power would be lacking to estimate $f_X(L(X))$ for all X.
- When Ω is wide, it is difficult to infer LATE from $\tilde{E}_\Omega$: **need to find a way to "shrink down"** Ω
- We assume that conditional on observed level of income, (F_X) can be considered locally constant around each observed probability of leaving. (Stronger version of this is F is always the same conditional on observed level of income).
- In this case, we can show that for the subset Ω'

$$\tilde{E}_{\Omega'} = \int_{F(L_1)}^{F(L_2)} w \circ F^{-1}(L) dL \tag{7}$$

Where w is the measure on the domain of F obtained by the application of Ω' on $\mathbb{R}$, L_1 and L_2 are the infimum and supremum of $L([\Omega'])$.

ATE and LATE

- For all individuals that have a similar probability of emigration (bounded between $1 - F(L(2))$ and $1 - F(L(1))$, this means that the semi-elasticity can be measured locally as the weighted integral of the inverse of the PDF around the point at which it is evaluated.
- This assures that by binning individuals into bins by baseline probability of leaving, we create estimators that converge towards a value as L_1 and L_2 converge toward each other. In practice, there is a trade-off between precision of the estimation (more bins) and power of the estimator (more individuals in each bin).
- In our reduced-form estimates, we have estimated values with 10 bins.

The role of *t*

- Our results show that individuals who have been in the country for shorter periods of time are more likely to emigrate.
- This is compatible with two scenarios:
  - There are different types: some people arrive with the intention of staying for a limited period of time and have a high probability of leaving while other come to settle indefinitely. As the high type attrite, the observed emigration rate goes down.
  - All foreigners arrive with a strong probability of leaving, but as they remain longer, their preference for staying increases and their preference for leaving decreases.
- Learning which of these two scenarios is true matters not only for tax purposes but important question in economics of migration (Dustman and Gorlach 2016, Adda Dustman and Gorlach 2020).
- It is usually impossible to distinguish between these scenarios, but the tax reforms provides an exogenous shock in the net cost of migration and a unique opportunity to look into those patterns.

Two types scenario

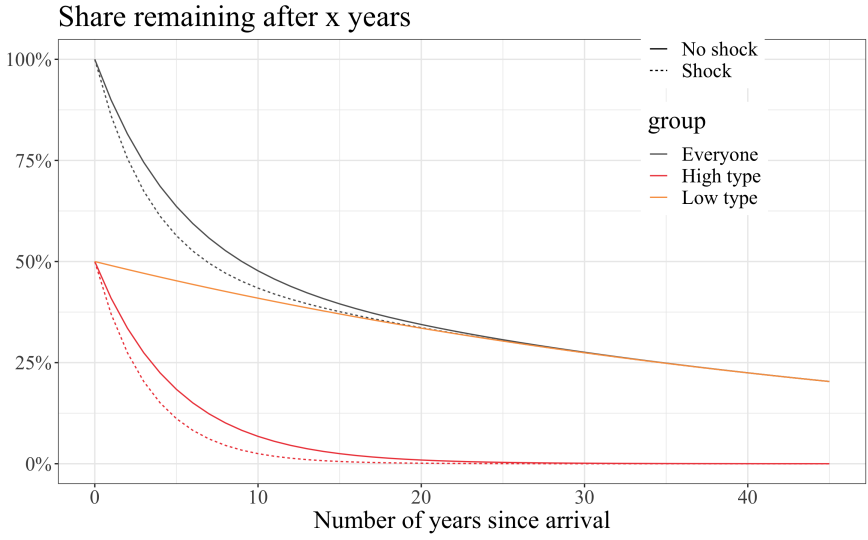
- Foreigners arrive in the country with either a “high” yearly probability of emigration (type H; proability of leaving p_H) or low (type L, probability of leaving p_L). Types are unobservable
- The probability of leaving is constant over time, implying an exponential distribution for the share of remainers after t years.
- If the share of foreigners remaining after t years is denoted by S_t, $S_0 = 1$ and s_h is the share of foreigners from the high type, we have:

$$S_t = s_h e^{-p_h \cdot t} + (1 - s_h) e^{-p_l \cdot t}$$

- We consider a multiplicative shock $1 + \gamma$ on the probability of emigration of the high group.

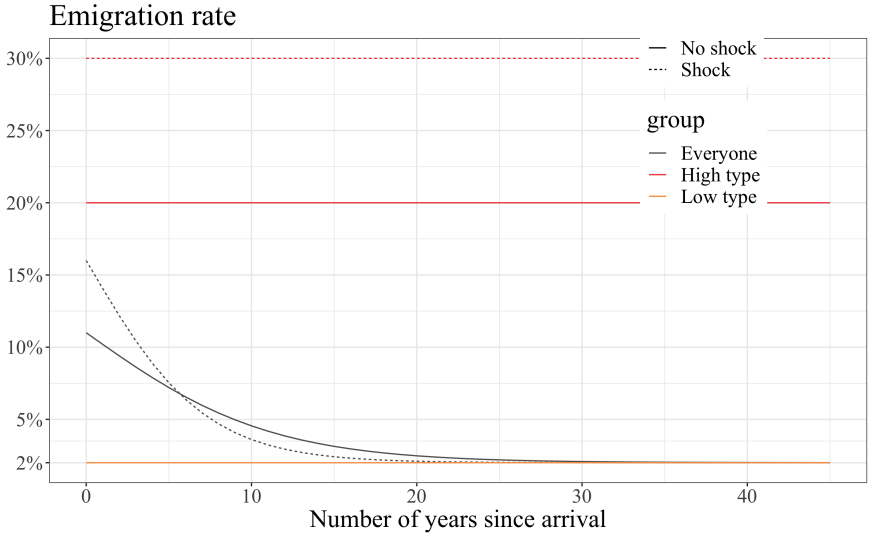
Back

Two types scenarios - Effect of shock on survival rate



Notes: This figure shows the survival rate of each type and in total as a function of the time spent in the country.
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline.

Two types scenario - Effect of shock on emigration rate



Notes: This figure shows the emigration rate of each type and in total as a function of the time spent in the country.
Source: Simulated data based on a population with half and half types, a high type emigration rate of 10% and a low type emigration rate of 2%. The shock is 50% of the baseline.

Adaptive preferences scenario

- We now model the type where foreigners take roots, and preferences for staying increase over time.
- let's denote S the survival function. S has to satisfy:

$$S(t+dt) = S(t) \times (1 - p(t)dt) \tag{8}$$

  Where p is the instant probability of emigrating in time t.
- Rearranging and solving the differential equation yields:

$$S(t) = \exp\left(- \int_0^t p(z)dz \right)$$

- If we model p as $p(t) = p_0 \cdot e^{(-\lambda t)} + p_{fin}$ with p_0 the probability of leaving in $t = 0$, λ the parameter that control the speed at which it decreases and p_{fin} the long run probability of leaving, we get:

$$S(t) = e^{\frac{-p_0}{\lambda}} \exp\left(\frac{p_0}{\lambda} e^{-\lambda t} + t \cdot p_{fin} \right)$$

Back

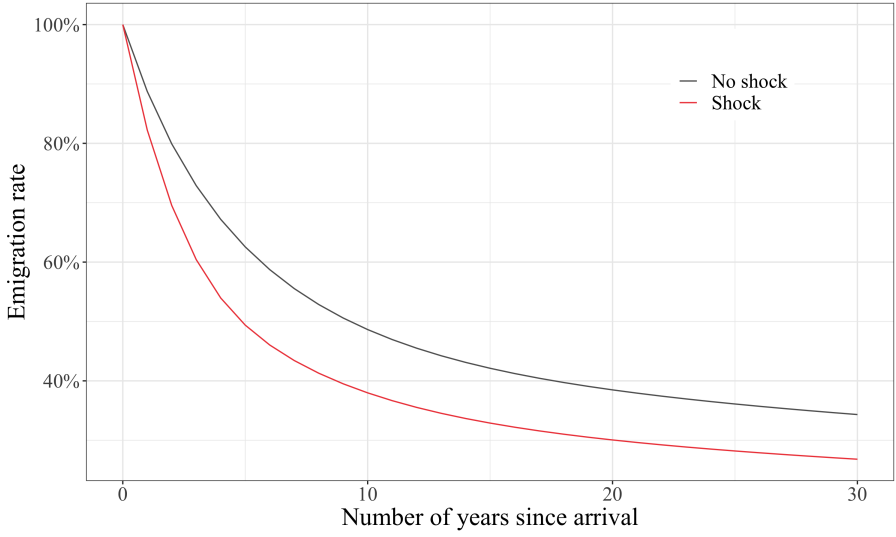
Adaptive scenario: shock

- We model a multiplicative shock $(1 + \gamma(t))$. We know from our results that γ is a decreasing function of t.
- A logistic-style is the best fit with our empirical results, so we define:

$$\gamma(t) = \gamma_0 \frac{e^{-k(t-t^*)}}{1 + e^{-k(t-t^*)}}$$

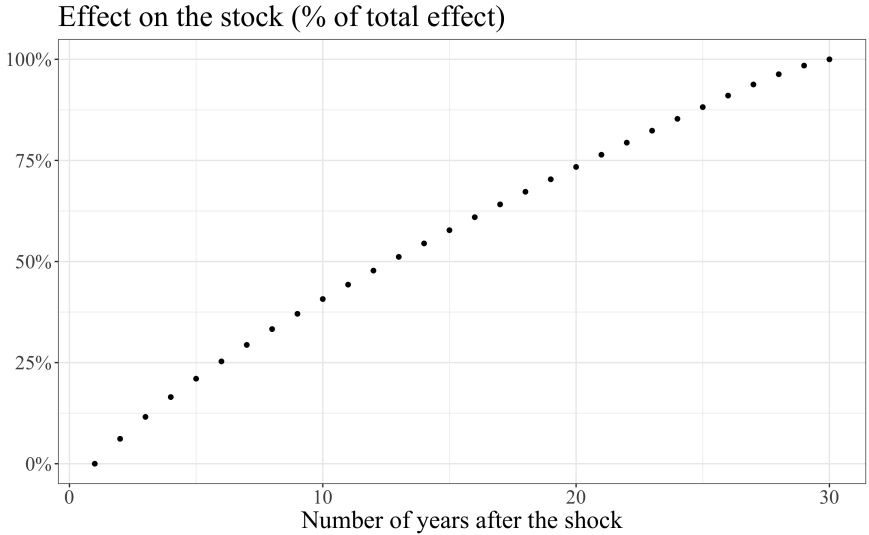
With γ_0 the intensity of the shock in $t = 0$ and t^* the inflexion point of the logitistic function (where the effect starts to fade out).

Adaptive scenario - Effect of shock on survival rate



Notes: This figure shows the survival rate as a function of the time spent in the country.
Source: Simulated data based on a population with an initial emigration rate of 12%, $\lambda = 0.14$, a terminal emigration rate of 1%, a shock of magnitude 70% and an inflexion point at 4.5 years.

Dynamic effect of a tax shock on the stock of foreigners

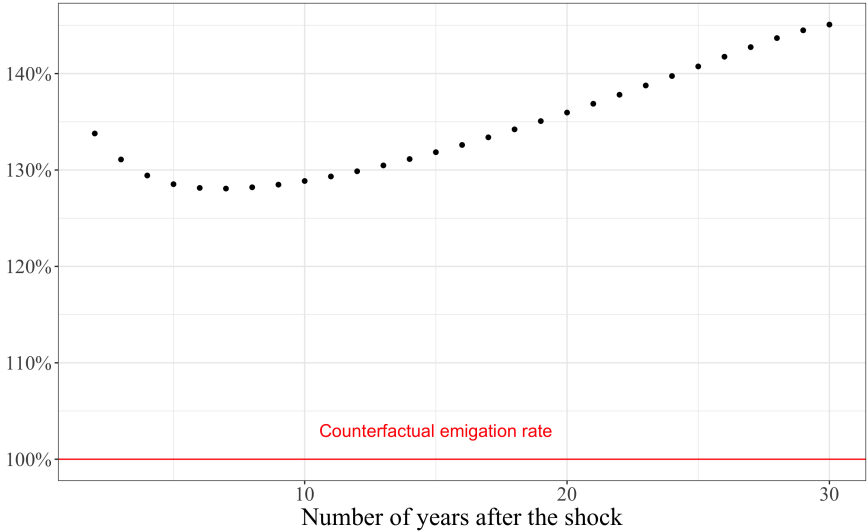


Notes: Impact of the tax change on the stock (in % of the total effect at 30 years) based on the estimated parameters in the stuctural estimation approach.

Source: Authors' calculations based on HMRC administrative datasets

Back

Dynamic effect of a tax shock on the emigration rate



Notes: Impact of the tax change on the aggregate emigration rate (in % of the initial emigration rate) based on the estimated parameters in the stuctural estimation approach.