BANK OF ENGLAND

Staff Working Paper No. 729

The macroeconomic determinants of migration

John Lewis and Matt Swannell

May 2018

BANK OF ENGLAND

Staff Working Paper No. 729

The macroeconomic determinants of migration

John Lewis[(1)] and Matt Swannell[(2)]

Abstract

We estimate a gravity model of the determinants of migration flows using pairwise data from around 160 origin countries to 35 advanced economy destinations over the period 1990–2013. When we interact the various explanatory variables with freedom of movement we find that the elasticities of migration with respect to macroeconomic variables are not constant across country pairs. Under freedom of movement, the response to macroeconomic variables is stronger, and the response to distance and historical migrant stocks is weaker. However, the elasticity with regard to linguistic and historical variables does remain constant. Alongside macro variables commonly used in the literature, we also find a significant role for expected GDP growth. Migration flows are higher to destinations with stronger expected GDP growth, and from origins with weaker expected GDP growth. In addition, greater labour market flexibility in destination countries is associated with higher inward migration.

Key words: Migration, macroeconomics, common correlated effects, multilateral resistance.

JEL classification: F22, E00, C23.

(1) Bank of England. Email: john.lewis@bankofengland.co.uk
(2) Bank of England. Email: matt.swannell@bankofengland.co.uk

The authors wish to thank Neeltje van Horen, Rebecca Freeman, Mike Joyce, Tommaso Aquilante and Patrick Schneider for useful comments. The views expressed are those of the authors and not necessarily those of the Bank of England.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/Working-papers

Publications and Design Team, Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 7601 4030 email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

Just as goods and capital can move across international borders, so can people. And as with trade and financial flows, migration can reflect many influences including linguistic factors, historical linkages, distance, the institutional arrangements, as well as the state of the macroeconomy.

This paper studies the macroeconomic determinants of bilateral migration flows through the lens of the gravity model. Our primary methodological innovation is to relax a standard assumption in the literature and allow the elasticity of each variable to vary according to whether there is a free movement agreement between a given country pair.

We find that the sensitivity of migration flows to many macroeconomic variables is significantly higher under freedom of movement. By contrast, the elasticity of almost all of the linguistic, historical and cultural variables we consider appears to be uniform across both cases. However, the coefficients on distance and legacy migrant stocks are smaller under free movement, suggesting the relative roles of space and historical stocks in determining migrant flows are smaller in such cases.

Our results have implications for understanding the effect of free movement itself. The traditional approach of capturing this via a dummy variable effectively imposes the property that freedom of movement boosts migration across any given country pair by a uniform scalar amount. By contrast, when interaction terms are included, the effect of freedom of movement varies with the underlying conditions between the origin and destination. We present robust evidence that there is substantial variation in the estimated size of this effect across country pairs.

In addition, we innovate with respect the rest of the literature by looking at the impact of expected future economic growth on migration decisions. To our knowledge this represents the first paper to explicitly consider the role of expected future economic growth alongside current GDP per capita, and we find a significant role for this. The pull factor of stronger expected growth in the destination country exerts a positive effect on migration flows, and does so with a common elasticity regardless of freedom of movement arrangements. By contrast the push of weaker expected growth in the origin country is only statistically significant for pairs where freedom of movement operates.

Our paper builds on the existing literature analysing the determinants of migration, and contributes to several strands of the literature.

Several papers have used large bilateral datasets covering multiple origins and destinations to explore the determinants of migration. Mayda (2010) and Ortega and Peri (2013) focussed largely on the role of institutional and policy variables. In a detailed investigation of the linguistic determinants of migration, Adserà and Pytliková (2015) (on whose dataset we build), also considered the role of GDP per capita and unemployment in driving migration flows. Consistent with the intuition that workers tend to move to places where job and income prospects are better, the common finding of these papers is that lower GDP and higher unemployment in origin countries are associated with larger migration outflows, and higher GDP and lower unemployment in destination countries are associated with larger inflows. Beine, Bourgeon and Bricogne (2013), analysed the role of a variety of macroeconomic factors and found a significant role for (un)employment rates at destination, though not at the origin. Relatedly, Keita (2016) analyses the influence of real exchange rates in the destination of economy on migration, and found that an appreciation of the destination currency 10% in real terms, yields a 19% increase in migration flows that country.

A related strand of the literature focusses on migration flows into our out of a single country. The inherent tradeoff in such studies is sacrificing cross-country generality for greater detail on inward flows and/or additional explanatory variables. For example, Bertoli, Brücker and Fernández-Huertas Moraga (2013) exploit the higher frequency nature of Spanish data to test for the role of multilateral resistance, or Yang (2017) combines migration data with remittances data to explore Filipino emigrants responses to exchange rate shocks, or **?** utilising richer data on policy variables to analyse inflows into the US. See also Forte and Portes (2017) for the UK. In studies with a single destination, one cannot distinguish between pairwise fixed effects and origin fixed effects, since for each origin country there is only a single pair. This eliminates the ability to test for the role of time-invariant pairwise factors such as language or distance. And more particularly for our research question, it also limits the amount of variation in freedom of movement available to the researcher, since in many cases, the bulk of variation occurs across country pairs, rather than stemming from changes over time within a country pair. In what follows we investigate both sources of variation in our dataset, and find that they exert similar effects on the estimated elasticities.

A common feature of both strands is the assumption that the elasticity of migration with respect to macroeconomic, linguistic or other variables is uniform. Ortega and Peri (2013) do investigate whether the coefficient with respect to destination GDP differs between a Europe-only sample, versus a broader sample and find a significant difference. But their split is based on geography rather than institutional arrangements[2]. Our paper tests this assumption of a uniform elasticity regardless of freedom of movement, and finds strong evidence that for macroeconomic variables, it does not hold. The relative magnitude of this disparity is substantial. We find, for example, that the elasticity of migration flows with respect to destination GDP is around three times higher under freedom of movement, than without it. [3]

Our paper also provides a first attempt to estimate the effect of expected future income on migration decisions. To our knowledge, there is only a single paper which attempts to test for the influence of future economic prospects. Bertoli, Brücker and Fernández-Huertas Moraga (2013) use the 10-year bond yield as a proxy for crisis risk, as a means of capturing future economic prospects. In our paper, we use a direct measure of expected future income- i.e the IMF's forecast economic growth- for both origin and destination country. We view this as a more direct measure of future income prospects than bond yields, and one which isn't affected by liquidity premia, term premia and other financial market developments orthogonal to future growth. Whilst the inclusion of bond yields may serve as a good proxy for overall crisis risk which is to the countries in their sample over the crisis period Bertoli, Brücker and Fernández-Huertas Moraga study, it

[2] In their classification country pairs which are on the European continent but didn't have free movement- e.g Poland and France - would be included in the "European" group. And others outside which do have free movement- e.g Australia and New Zealand- would not be included. We therefore their measure as only an approximation to freedom of movement status

[3] This is higher than the disparity found by Ortega and Peri (2013), which i) reflects the fact that their measure of geographic measure is only a proxy for freedom of movement and ii) We can estimate the elasticity for pairs where no freedom of movement exists, whereas their alternative is a pooled regression in which combines all country pairs.

may be a less good proxy for growth prospects in more tranquil times, over the much broader and longer sample period that we use in this paper.

Our paper also contributes to the literature on migration and labour market institutions. Migali (2017) looks at migration between EU countries over the period 2001-11. Our paper builds on this by greatly expanding the time and country coverage, with a dataset around 40 times the size. Relatedly, Cigagna and Sulis (2015) look at the effect of employment protection on migration for a broader set of countries, but they don't consider the role of origin-specific macroeconomic factors. Our paper adds to both of these by explicitly considering the interaction of such measures with freedom of movement arrangements, though we find that the elasticity of migration with respect to labour market flexibility appears to be uniform with respect to these.

The remainder of the paper is organised as follows. Section 2, details our estimation approach; Section 3 gives details of our dataset; Section 4 presents our empirical results, and Section 5 concludes.

2 Estimation Approach

As is common practice in the literature, we derive our empirical specification from the Randomised Utility Model (RUM) of migration.[4] The utility that an individual from origin country o makes gets from moving to a given destination country $d \in D$ can be expressed as the utility an individual gets from two deterministic components: the intrinsic attractiveness of the destination country V_{odt} and the associated cost of locating there C_{odt}, plus an individual-specific stochastic component ϵ_{iodt}.

$$U_{iodt} = V_{odt} - C_{odt} + \epsilon_{iodt} \tag{2.1}$$

The stochastic component's distributional assumption determines the probability an individual will leave origin country O to settle in destination country D. Beine, Bertoli and Moraga (2015) et al show that under the assumption that the stochastic term follows an independent and identically

[4]See Beine, Bertoli and Moraga (2015) for a full exposition of the model

distributed Extreme Value Type-1 distribution (McFadden, 1974) the probability of an individual emigrating from country o to country d, and so the corresponding emigration flow, can be expressed as a multiplicative function of: the size of the origin country; the attractiveness of the destination; and accessibility of the destination. So it follows that the ratio of those who are expected to emigrate to a given destination to those who choose to remain in the origin nation can be expressed in terms of the relative attractiveness of the two countries and the accessability of the destination country.

Therefore, the dependant variable is the emigration rate from origin country o to destination country d at time t. This is equal to the flow M divided by the stock of people who remain O_{odt}. Due to data limitations on this stock of non-migrants, we following the standard proxy used in the literature – i.e. the total population, P_{ot}. As Beine, Bertoli and Moraga (2015) show, the assumption that there is no cost to remaining in country o implies:

$$\frac{M_{odt}}{M_{oot}} \approx \frac{M_{odt}}{P_{ot}} = m_{odt} = e^{V_{odt} - V_{oot} - C_{odt}} \frac{\eta_{odt}}{\eta_{oot}} \tag{2.2}$$

The assumption that the stochastic component follows an independent and identically distributed Extreme Value Type-1 distribution is analytically convenient, but it does not allow for correlations across destinations. As Bertoli and Fernández-Huertas Moraga (2012) argue, this seems too restrictive. There may be unobserved determinants of migration - for example, cultural proximity - that mean migrants from country O get, conditioned on observable determinants, systematically higher utility from migrating to some nations rather than others. Therefore, it is important to relax the assumption of independence between the stochastic components of utility when working with aggregate economic data.

Having relaxed this assumption, Bertoli and Fernández-Huertas Moraga (2012) et al show that if the set of potential destination countries D can be split into subsets - or nests - of nations that share either observed or unobserved characteristics that can have a heterogenous impact on the utility that an individual gets from settling there than the migration flow from o to d in year t, M_{odt} can be expressed as:

$$M_{odt} = s_j \frac{e^{(V_{odt}-C_{odt})/\tau}\left(\sum_{l\in b(k)} e^{(V_{olt}-C_{olt})/\tau}\right)^{\tau-1}}{\sum_q \left(\sum_{l\in b_q} e^{(V_{olt}-C_{olt})/\tau}\right)^{\tau}} \eta_{od} \qquad (2.3)$$

The parameter τ in 2.3 - often referred to as the dissimilarity parameter - is inversely proportional to correlation of the stochastic component of utility across alternative destination countries. This measures of the substitutability of alternative destinations when individuals make their migration decision. Bertoli, Brücker and Fernández-Huertas Moraga (2013) refer to this as multilateral resistance. All else equal, a higher degree of substitutability between destinations implies a greater sensitivity of migration decisions to changes in their relative attractiveness.

Recent work has demonstrated the importance of considering the attractiveness of alternative destinations when modelling migration flows. A rise in the attractiveness of one location increase migration both via inducing those who would otherwise have remained in their country of origin to migrate, and also by diverting migrants from other destinations. In an important paper, Bertoli and Fernández-Huertas Moraga (2013) extend the 'multilateral resistance' concept of Anderson and van Wincoop (2003) into the analysis of migration and demonstrate that failure to account for this can induce a bias into the estimated coefficients, resulting in bigger coefficients on output and lower coefficients on migration policies. In a single destination context, Bertoli, Brücker and Fernández-Huertas Moraga (2013), find these divergence effects to be significant, accounting for up to 78% of the observed increase in migration flows to Germany.

The empirical estimation must therefore control for the relative attractiveness of other destination countries. One approach is to control for the effect of multi-lateral resistance by including an origin-time fixed effect (Ortega and Peri (2013)). But, underlying this specification is the rather strong assumption that for a given origin country, all destinations are equally viable substitutes for one another.

We therefore implement an alternative approach of grouping countries into 'nests' of alternative destinations, , pioneered by Bertoli, Brücker and Fernández-Huertas Moraga (2013). In addition to our standard regressors, we include the within-nest weighted average of all destination specific variables, where the within-nest weight is equal to the inverse of the distance between the destination and origin countries, which embodies the intuition that geographically closer countries are likely to be closer substitutes for a given destination. Econometrically speaking, this is equivalent to applying the Common Correlated Effects (CCE) estimator proposed by Pesaran (2006). In the same spirit as Bertoli, Brücker and Fernández-Huertas Moraga (2013) we group destinations into 6 nests based on geographic proximity and cultural similarity.[5] The PPML estimator will be consistent when appropriate 'nests' of substitutable destinations are imposed. Accounting for multilateral resistance means that our empirical estimates represent the true structural parameters underpinning V_{odt} and C_{odt} in 2.3; rather than the structural parameters divided by the dissimilarity term, τ.

Estimation of multiplicative models of migration flows in the literature, typically follows one of two specifications: either log-linearising the equation or alternatively using the Poisson Pseudo Maximum Likelihood (PPML) estimator of Santos-Silva and Tenreyro (2006). We utilise the latter as it has two distinct advantages: (i) it doesn't require log-linearisation of a potentially heteroscedastic error term and so the consistency of the estimator is preserved and (ii) it can be computed for cases where the dependent variables take the value 0, which is often the case in bilateral migration dataset.[6] We include three sets of fixed effects unless otherwise stated- origin fixed effects, destination fixed effects and year fixed effects.

3 Dataset

We combine data from a wide variety of sources. We have over 100,000 bilateral migration observations over our sample period, but our regression sample reduces to around 65,000 observations, because coverage of our other variables (detailed below) is not as extensive as for migration flows themselves. Nevertheless, in terms of sheer size the dataset used in our re-

[5]Country groups are (i) Western European EU members: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Ireland, Luxembourg, Netherlands, Portugal, Spain and the United Kingdom; (ii) EFTA members: Iceland, Norway and Switzerland; (iii) Central Europe: Czech Republic, Hungary, Poland, Slovakia; (iv) North America: Canada and USA; (v) Pacific Rim Advanced Economies: Australia, New Zealand and Japan and (vi) A residual group comprising the other destinations: Chile, South Korea, Mexico and Turkey.

[6]In our dataset around 7% of observations are zero

gressions is one of the largest bilateral dataset of its kind so far assembled in the literature. Full details on definitions, sources and rationale for the inclusion of variables are detailed in the remainder of this section.

3.1 Demographic Variables

Our starting point for migrant flows is Adserà and Pytliková (2015) comprehensive dataset on bilateral migration, which they gathered based on data from individual national statistical authorities.[7] These data cover flows to 30 OECD destination nations and 223 origin nations[8] over the period 1980 to 2010. We augment their dataset using the OECD's International Migration Database OECD (2017) along two dimensions. First, we extend the time dimension to include the years 2011 to 2014 and second, we include Chile as an additional destination country.

In this measure, a migrant is therefore defined as a person who moves legally from an origin country defined by either where they are currently residing or hold citizenship, into another country where they were neither born nor hold citizenship.[9] As a result this excludes both illegal migration flows and migrants moving back to their country of birth/citizenship. Such return migrants are likely to be motivated by quite different factors to other migrants and so should be studied as a distinct phenomenon (Mesnard, 2004; Dustmann and Kirchkamp, 2003; Dustmann, 2002). Similarly, empirical evidence indicates the composition and determinants of illegal migration flows are different Bratsburg (1995), which suggests even if comprehensive bilateral data on such flows were available, they too should be analysed separately.

[7]This is available online from the online version of the Economic Journal under "Supplementary Materials" for the article

[8]These are: Australia, Austria, Belgium, Canada , Czech Republic, Germany, Denmark, Finland, France, Greece, Hungary, Ireland, Iceland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, Norway, New Zealand, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland, Turkey, United Kingdom and the United States

[9]The definition of country of origin varies from destination to destination. Some record inflows by country by residency, others by birth or citizenship. This may mean that for a small proportion of migrants, the "origin" country does not correspond to the place where they were previously living. For example, a UK citizen living in the US moving to Canada would count as a migrant "from" the UK, rather than Canada. But in practice, such third country migration cases are likely represent only a tiny fraction of migration flows between a country pair

We also make use of Özden et al.'s (2011) dataset on migrant stock in each destination country by their country of origin on a decennial basis which draws primarily on the United Nations Population Division's Global Migration database. We include the migrant stock in 1990 – i.e. directly prior to the start of our estimation sample – as a time-invariant variable in our specification. As Beine, Bertoli and Moraga (2015) note, including contemporaneous stocks to proxy network effects leads to an endogeneity problem, and including lags of flows are invalidated by the possibility of serial correlation in the error term.

An additional advantage is that this variable captures the effect of other time-invariant factors without having to resort to pair-specific fixed effects. This allows us to retain other time-invariant pair-specific variables in our specification such as distance or freedom of movement which is useful, because much of the variation in freedom of movement occurs between different country pairs rather than within pairs over time. We return to this latter point in our results section.

Finally, our figures for population required for the normalisation of migration flows are taken from the IMF's April 2017 World Economic Outlook (International Monetary Fund, 2017).

3.2 Macroeconomic variables

Ideally one would like to capture income prospects with real wages, but limitations in country coverage of wage data mean that, like the bulk of the existing literature, we have to proxy this by using real output per head. GDP per capita and population estimates are taken from the IMF's April 2017 World Economic Outlook . The IMF only provide estimates of PPP-weighted GDP per capita in current US dollars, so we deflate these using the IMF's US GDP deflator to derive a constant price PPP adjusted measure of income per capita.

To allow for the influence of expected future income alongside current GDP, we calculate the cumulative real GDP growth between years t and $t + 5$ as reported in the year $t - 1$ 'Fall World Economic Outlook'[10]. For unemployment, we utilise the ILO unemployment rates from the World Bank's

[10] Data is taken from the IMF's Historical WEO Database (2017b)

statistics database World Bank (2017).

We also consider the role of the real exchange rate.[11] This is defined in the "European" style, such that a rise indicates a real exchange rate appreciation. [12]

3.3 Linguistic Factors

Our choice of linguistic variables is informed by recent advances in the literatures on the effect of language on trade and migration flows. We use Melitz and Toubal's (2014) continuous variables on the commonality of native and spoken languages which they show outperforms simple binary measures of common languages in explaining trade flows. The common native language variable measures the probability that a randomly chosen citizen from the origin shares a common native language with a randomly chosen citizen from the destination. Similarly, the common spoken language variable records the probability that such a pair has at least one language in which they are both mutually proficient. [13] Instead of including both variables directly as Melitz and Toubal do, we include the common native language and a second variable which is the difference between the two. This latter variable captures the marginal effect of an overlap in a common but non-native language.

Adserà and Pytliková's (2015) showed that alongside commonality of language, including the relative similarly of languages in origin and destination countries has significant explanatory power for migration flows. Pick-

[11]If migrants are motivated by either the desire to build up savings for use on return (Dustmann and Görlach, 2016)or to send remittances home, the higher the relative purchasing power of money earned in the destination economy but spent in the origin economy the more attractive migration would be (see, for example Mishra and Spilimbergo, 2011). Our measure is based obtained dividing market exchange GDP by PPP GDP which has the advantage of using identical price indices use to calculate real GDP and real exchange rates are

[12]Keita (2016) uses a Consumer Price Index (CPI) based Real Exchange Rate. Since CPIs are not comparable in levels across countries, CPI based real exchange rates cannot be compared in levels across country pairs. In Keita's setup this is fine because pairwise fixed effects are used and hence only the variation in real exchange rates within country pairs is used in estimation. But because our preferred estimation approach doesn't use pairwise fixed effects, we cannot use such real exchange rate indices.

[13]In this measure each person has only one native language, but may have additional spoken languages, so the common spoken language score between a country pair will always be greater than or equal to the common native language score.

ing up this theme, we use the same measure as Melitz and Toubal, which is an index of similarities of words with identical meanings, developed by the *Automated Similarly Judgement Project (ASJP)*, because this affords us slightly greater data coverage. This yields a discrete variable which with six categories between zero (no similarity) and one (common language)[14]

3.4 Historical, Geographic, and Other Determinants

We include a variety of time-invariant, country pair specific variables to capture the influence of non-economic factors.

Our data on historical and geographical variables is taken from the CEPII GeoDist database (Mayer and Zignago, 2011). As in other models of trade and migration flows, distance between country pairs[15]

Alongside distance, we also use this database as the source for dummy variables measuring contiguity (which equals one if two countries share a land border; zero otherwise), colonial links (equal to one if the countries have ever had a colonial link) and common country (equal to one if the country pair were ever part of the same country[16])

We measure the flexibility of the labour market using two OECD's indicators of employment protection (OECD, 2017). The first records the strictness of the procedures which govern dismissal and the costs associated with laying off workers and the second records the prevalence of temporary contracts within the labour market. These data run from 0 to 6, where a lower value corresponds to greater flexibility. Migali (2017) finds evidence in support of the view that migration flows are higher into destinations with more deregulated labour markets, which is attributed to the greater ease with

[14]Since this variable can only be calculated over two languages, we use the single most widely spoken language in for each country. We also experimented with Adserà and Pytliková's proximity variables, and in all cases the coefficients on the macro variables- our primary focus in the paper- remained very similar. See appendix for details

[15] As measured by the great circle distance between largest city in each country. As a robustness check, we also experimented with alternative distance measures in the GeoDist database based on distance between capital cities, and population weighted average distance between major cities obtaining very similar results. See appendix for details.

[16]This is equal to one if countries were ever part of the same state for a prolonged period. This captures cases where two states were part of the same empire (e.g. Bosnia-Herzegovina and the Czech Republic under the Hapsburg empire), or were part of a country which broke up (e.g. Czech Republic and Slovakia after the dissolution of Czechoslovakia)

which new arrivals can enter employment.

3.5 Free movement arrangements

We create a dummy which takes the value of one if a country pair has a free movement agreement permitting nationals of the origin country to work in the destination without any additional approval. We identify four different such arrangements in the span of our dataset.

First, the European Economic Area (EEA)[17] ordinarily gives citizens the right to live and work in another member state. In addition, Switzerland has been included through bilateral arrangement since 2002. Following the accession of New Member States between 2004 and 2014 some countries implemented transitory controls delaying the implementation of free movement with respect to those states. In such cases, we only code our free movement dummy as a one if the transitory restrictions have been lifted.[18]

Second, the Nordic Passport Union between Denmark, Finland, Iceland, Norway and Sweden has allowed nationals of one member state the right to reside and work in another since its inception in the 1950s, predating any rights conferred by the EEA arrangements above.

Third, the Trans Tasman Travel Arrangement, has permitted citizens of New Zealand and Australia the right to work in each other's countries since 1973.

Fourth, the South American Mercosur bloc[19] has a freedom of movement provision. In the context of our sample, this means that nationals of member states are able to live and work freely in Chile, the only Mercosur destination country covered in our dataset.

We believe this represents a better measure than the simple dummy for EU membership which is typically used in the literature.[20] Most importantly, it

[17] This consists of the European Union, plus Norway and Iceland

[18] Where a country pair's status changes mid-year, we code this is a one if for the majority of days in the year in question free movement was in operation between a given pair.

[19] These are Argentina, Brazil, Paraguay and Uruguay from 1991 onwards, and subsequently Bolivia, Chile, (both 1996), Peru (2003), Colombia and Ecuador (both 2004).

[20] Cigagna and Sulis (2015) use a common currency dummy variable, which therefore excludes a number of EEA country pairs and records a zero value for all pairs, prior to 1999 Keita (2016) uses the year from which both origin and destination were signatories of the Maastricht Treaty, which also excludes certain EEA pairs and doesn't take into account transitionary controls on migration for newer member states. Mayda (2010), whose sample

has a broader coverage, encompassing four rather than one free movement arrangements. In addition, it accords directly with the mobility provisions in place- capturing both the mobility provisions that extend to Switzerland, Norway and Iceland; as well as the restrictions on movement associated with the accession of new members after the millennium.

4 Results

Table 1 below, shows how we obtained our baseline specification.

The first regression represents the standard approach in the literature, using a simple dummy to capture the effect of free movement. The variables are all significant with the exception of the real exchange rate, common border, and origin country expected growth and those that are significant have the expected signs. Looking at the origin country variables- higher unemployment and weaker expected growth are associated with higher outflows. Origin GDP and its square are both significant, indicating a non-linear relationship between origin GDP and migration flows. On the destination side, higher output, higher expected growth, lower unemployment and greater flexibility of labour markets are all associated with higher inflows. Country pairs with more similar languages, and greater overlap in spoken and native languages tend to have larger flows between them. Colonial links between countries also increase migration flows. Interestingly, the 1990 migrant stock is also strongly significant, suggesting that the presence of an larger initial community is associated with higher migration flows. This could be capturing either network effects (i.e. migrants tend to prefer destinations with an established community from their own origin country), or simply picking up other unobserved time-invariant characteristics between a given country pair (i.e. factors outside of our regression specification that made a destination country more attractive to migrants from a given origin country prior to the sample remain in operation during the sample). Finally, distance exerts a drag on flows, with more distant country pairs

only runs up to 1995, and Adserà and Pytliková (2015) do not include a freedom of movement dummy of any kind. Ortega and Peri (2013) use a combination of the Maastricht dummy, a common currency dummy and an index which records the direction of migration restrictions. Only Migali (2017) looking only intra-EU mobility uses a free movement dummy which is defined in the same way as ours.

Table 1: Estimation Results

	(1) Common coeffs		(2) Split coeffs		(3) Baseline	
ninflow						
Free movement	1.135***	(0.208)	-4.972	(3.128)	-5.810*	(3.102)
Origin GDP per capita	1.311***	(0.411)	1.187***	(0.438)	1.226***	(0.438)
Origin GDP per capita squared	-0.251***	(0.072)	-0.211***	(0.077)	-0.216***	(0.077)
Destination GDP per capita	4.029***	(1.276)	3.678***	(1.259)	3.848***	(1.277)
Origin unemployment rate	0.237***	(0.073)	0.194**	(0.092)	0.184**	(0.081)
Destination unemployment rate	-0.241*	(0.137)	-0.236	(0.151)	-0.231*	(0.134)
Origin expected growth	-0.092	(0.340)	0.166	(0.345)	0.155	(0.360)
Destination expected growth	2.117**	(0.967)	2.044*	(1.062)	2.070*	(1.079)
Collective dismissal	-0.646***	(0.142)	-0.590***	(0.147)	-0.606***	(0.137)
Temporary employment	-0.140	(0.088)	-0.204**	(0.096)	-0.182**	(0.085)
Distance	-0.534***	(0.067)	-0.604***	(0.063)	-0.616***	(0.063)
Shared spoken language	1.437***	(0.298)	1.297***	(0.333)	1.468***	(0.291)
Common native language	0.640**	(0.322)	0.753**	(0.352)	0.640**	(0.315)
Linguistic similarity	0.845***	(0.246)	0.941***	(0.280)	0.949***	(0.254)
Contiguous border	-0.267*	(0.150)	0.375**	(0.166)		
1990 migrant stock	0.413***	(0.026)	0.442***	(0.026)	0.449***	(0.026)
Colonial link	0.299**	(0.130)	0.242*	(0.130)	0.279**	(0.118)
Real exchange rate	-0.096	(0.142)	-0.011	(0.151)		
Formerly same country	0.074	(0.205)	-0.053	(0.217)		
FM*Origin GDP per capita			2.455	(1.630)	2.570*	(1.495)
FM*Destination GDP per capita			1.147**	(0.519)	1.315***	(0.475)
FM*Origin unemployment rate			-0.034	(0.131)		
FM*Destination unemployment rate			-0.011	(0.155)		
FM*Origin expected growth			-3.194***	(0.842)	-3.180***	(0.910)
FM*Destination expected growth			-0.875	(1.999)	-1.324	(1.965)
FM*Origin GDP per capita squared			-0.554**	(0.265)	-0.590**	(0.233)
FM*Collective dismissal			-0.147	(0.148)		
FM*Temporary employment			-0.012	(0.071)		
FM*Distance			0.327**	(0.145)	0.345***	(0.118)
FM*Shared native language			0.053	(0.505)		
FM*Common native language			-0.946	(0.873)		
FM*Linguistic similarity			0.093	(0.510)		
FM*Contiguous border			-0.427*	(0.243)		
FM*1990 migrant stock			-0.175***	(0.030)	-0.189***	(0.029)
FM*Colonial link			0.254	(0.240)		
FM*Real exchange rate			0.116	(0.281)		
FM*Formerly same country			0.412	(0.307)		
N	66390		66390		66422	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; standard errors in brackets

tending to have lower migration flows between them. The coefficient on the free movement dummy variable is positive and significant, indicating that free movement arrangements increase migration flows.

But this specification effectively assumes two things- i) A freedom of movement agreement between any given pair of countries will always have an identical scalar effect on migration flows. ii) The elasticities of migration with respect to a given variable are the same regardless of whether freedom of movement exists between them.

Our second regression puts those assumptions to the test by including a set of interaction terms between the dummy variable and each of the other explanatory variables. The coefficients on the interaction terms show the *marginal* effect on the elasticity of a free movement agreement, and so the elasticity under free movement is given by sum of the coefficient on the variable plus the coefficient on the interaction term.

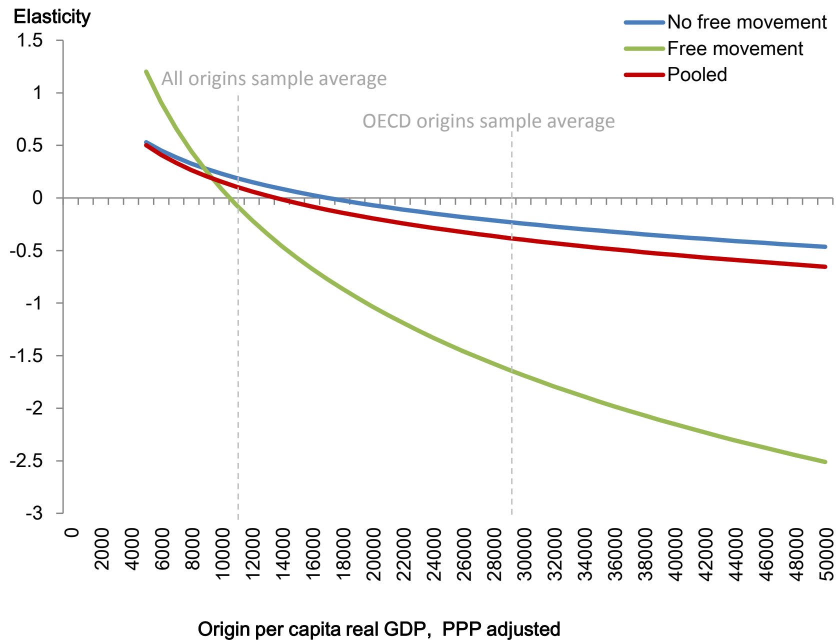
Five of these terms are significant, indicating that interaction effects are important and that assumptions above do not hold. Under free movement, the response to destination GDP and destination growth are all stronger. On the other hand, the drag exerted by distance is smaller, as is the influence of pre-sample migrant stocks.

We then drop the real exchange rate variables (as they are always insignificant), as well as any interaction terms which are insignificant for both origin and destination. After doing this, the contiguity variable then becomes insignificant, so we also drop this variable. This results in regression (3), in the rightmost column of the table. This constitutes our baseline for the rest of the analysis and we discuss the estimates in more detail below.

We begin with the macroeconomic variables. The overall elasticity of migration with respect to origin GDP is not straightforward to read from our results table, since the relationship is non-linear. Omitting other terms and the interaction effects, our model posits that $ln(m_{odt}) = \alpha + \beta_1 ln(GDP_{ot} + \beta_2[ln(GDP_{ot}]^2 + \ldots$. The overall elasticity $\frac{\partial ln(m_{odt})}{\partial ln(GDP_{odt})}$ is obtained by differentiating this expression with respect to the log of origin GDP and is thus given by: $\beta_1 + 2\beta_2$.[21]. The chart below plots this relationship:

[21]Under free movement, the elasticity is given by the above expression, plus the interaction coefficient on GDP plus two times the interaction coefficient on GDP squared

Figure 1: The effect of origin GDP on migration flows



In common with Clemens (2014) and Forte and Portes (2017) we find that at very low levels of income the elasticity is positive- i.e. rising incomes increase emigration rates, but at higher levels of income, the elasticity is negative- implying higher incomes reduce emigration. The difference between the blue and green lines makes plain the sizeable difference in the elasticities across the two groups. At the mean OECD income in our dataset, just under $ 30,000 , the elasticity is -1.6 under free movement, implying that a 1% rise in income reduces emigration by about 1.7%, which is much higher than when there is no freedom of movement. The red dotted line shows the elasticity when a common set of coefficients is estimated for both groups (taken from regression (1)). The estimated response is fairly close to the no free movement case, which reflects the fact that the bulk of observations in our sample (95%) fall into this category. This illustrates starkly how failure to allow for appropriate interaction effects can lead to serious errors in estimating elasticities.

The coefficient on destination GDP per capita is 3.8, rising to 5.2 when the interaction effect is included.

Turning to the role of expected growth, we find no significant "push" factor exerted by weaker expected growth in the origin economy, but there is a significant “pull" from expected growth in the destination economy- for every 1pp rise in expected growth, immigration flows increase by 2.1%. Looking at the interaction terms, the destination expected growth one is insignificant, suggesting a uniform elasticity; but the origin expected growth term is significantly negative. That means under freedom of movement a 1% fall in expected GDP growth produces a 3.2% rise in emigration.

Both the origin and destination unemployment rates are significant A 1% rise in the origin unemployment rate leads to a 0.18 % rise in the emigration, and corresponding 1% rise in destination unemployment leads to a 0.23% fall in inward migration. Interestingly, neither interaction term is significant, implying that freedom of movement doesn't increase the sensitivity of migration decisions to unemployment.

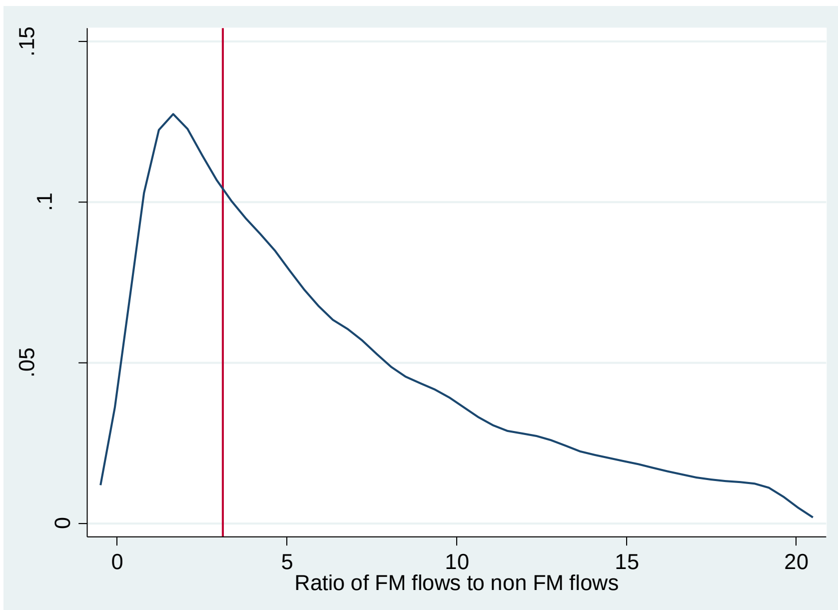
Turning to the non-economic determinants, we find a positive role for common native language, though the effect of overlap in non-native language is much higher. Similarly, we find a significant positive effect of linguistic similarity. Strikingly, regression (2) shows that none of the interaction terms for these variables are significant, implying that the elasticity of migration flows are constant across both groups. Similarly, colonial links boost migration flows, but do so in uniform manner.

In keeping with most of the literature, distance exerts a significant drag on migration. The coefficient of -0.62 implies that a doubling of distance reduces migration flows by just over a quarter. But the interaction coefficient is positive, which means under free movement the sensitivity to distance more than halves. There is a similar story for the role of pre-sample migrant stocks. The coefficient of 0.449 means that a doubling of the pre-existing migrant stocks is associated with 45% more migration. But the interaction effect is negative, and so the effect of stocks under free movement is almost halved. That suggests that under freedom of movement, migration flows are less inert, and less sensitive to spatial considerations.

Finally, we consider the effect of free movement. This cannot be read directly from table. In a model where dummy variables are also interacted with regressors, the coefficient on the dummy variable has no direct interpretation, since the effect of moving from the control to the treatment group is also a function of the interaction terms. Formally speaking the effect is given by the coefficient on the dummy variable plus the sum of the differences in coefficients times the values of the regressors for the country pair in question. We perform this for each country pair, and then estimate a kernel density function which is shown below:

The modal value is just under 2, but there is a quite a large spread in the estimated effect across country pairs and the kernel density function has a pronounced positive skew. For comparison, the vertical line shows the magnitude implied by dummy variable in the “single elasticity" estimates from regression (1).[22] Evidently, failure to account for differential elasticities can mask substantial variation in the estimated magnitude of response.

Figure 2: The effect of free movement



4.1 Alternative Estimation Strategies

Table 2 below presents the results using alternative estimators.

[22]The coefficient on the dummy variable in regression (1) is 1.135, the vertical line corresponds to $e^{1.14} = 3.11$

Table 2: Alternative Estimators

	(1) Baseline		(2) No CCE		(3) Dyadic FEs	
main						
Free movement	-5.810*	(3.102)	-5.124	(3.205)	-1.679	(3.385)
Origin GDP per capita	1.226***	(0.438)	1.109**	(0.450)	1.296***	(0.431)
Origin GDP per capita squared	-0.216***	(0.077)	-0.198**	(0.080)	-0.243***	(0.075)
Destination GDP per capita	3.848***	(1.277)	2.548**	(1.016)	3.622***	(1.274)
Origin unemployment rate	0.184**	(0.081)	0.197**	(0.084)	0.156**	(0.076)
Destination unemployment rate	-0.231*	(0.134)	-0.243*	(0.127)	-0.260**	(0.130)
Origin expected growth	0.155	(0.360)	0.150	(0.381)	-0.087	(0.332)
Destination expected growth	2.070*	(1.079)	2.419**	(1.042)	1.750	(1.107)
Collective dismissal	-0.606***	(0.137)	-0.610***	(0.145)	-0.625***	(0.122)
Temporary employment	-0.182**	(0.085)	-0.186**	(0.073)	-0.112	(0.082)
Distance	-0.616***	(0.063)	-0.609***	(0.062)		
Shared spoken language	1.468***	(0.291)	1.462***	(0.292)		
Common native language	0.640**	(0.315)	0.571*	(0.312)		
Linguistic similarity	0.949***	(0.254)	0.989***	(0.254)		
1990 migrant stock	0.449***	(0.026)	0.451***	(0.027)		
Colonial link	0.279**	(0.118)	0.316**	(0.123)		
FM*Origin GDP per capita	2.570*	(1.495)	1.845	(1.451)	0.792	(2.074)
FM*Destination GDP per capita	1.315***	(0.475)	1.485***	(0.395)	0.890*	(0.518)
FM*Origin expected growth	-3.180***	(0.910)	-3.193***	(0.905)	-2.775***	(0.738)
FM*Destination expected growth	-1.324	(1.965)	-0.623	(2.011)	1.828	(1.440)
FM*Origin GDP per capita squared	-0.590**	(0.233)	-0.481**	(0.224)	-0.282	(0.322)
FM*Distance	0.345***	(0.118)	0.306**	(0.125)		
FM*1990 migrant stock	-0.189***	(0.029)	-0.195***	(0.031)		
N	66422		66422		66089	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; standard errors in brackets

Regression (2) shows the consequences of failing to control for the role of alternative destinations by omitting the CCE terms. Whilst the bulk of coefficients are similar and not statistically significantly different across the two approaches, the response to destination GDP per capita is around a third lower when one fails to account for multilateral resistance. This difference appears reasonable for similar reasons to those outlined in Bertoli and Fernández-Huertas Moraga (2013). If we assume that GDP growth between different destination countries in a nest is positively correlated, then when GDP rises in one country, it is likely to rise in others too. If not controlled for, then the estimate of the effect of destination GDP also capture the partially offsetting effect of rising GDP in alternative destinations, which would bias coefficients downward.

Unlike Bertoli and Fernández-Huertas Moraga we do not find a significant difference in the coefficients in origin GDP per capita across the two specifications.[23] This probably reflects the fact that origin and destination

[23] A cross-equation restriction test of coefficient equality for each of the four origin GDP coefficient pairs failed to reject the hypothesis of equality in all four cases

income per heads, are not, in general significantly correlated across country pairs. In Bertoli and Fernández-Huertas Moraga, there was a single destination country (Spain), and the origin countries were largely either European neighbours or Spanish speaking countries where they may be deeper trade and financial linkages, and so the correlations between origin and destination GDP might be significant. By contrast, our data contains a much larger and more diverse set of destinations (35) and origins (223) so the average GDP correlation across origin-destination pairs is likely to be lower. Similarly, we find only a very small difference in the kernel density function for the effect of freedom of movement (see appendix). This is probably also explained via a similar logic. In Bertoli and Fernández-Huertas Moraga's the variation in freedom of movement comes entirely from changes over time in the visa status of a particular origin country, many of which were based on EU-wide policies, and hence they were strongly correlated with similar changes in other destinations, and in a single destination analysis any other time-invariant differences between countries (e.g. one origin always had a free movement arrangement, another never did) would be eliminated by those fixed effects. By contrast, in our dataset, with a much more diverse set of countries, and the bulk of variation coming between country pairs rather than changes within them over time, means that our freedom of movement variable is relatively speaking subject to much less correlation across destinations.

Regression (3) estimates the equation using dyadic fixed effects. Many papers in the literature, especially those focussing on macroeconomic variables, use such pairwise fixed effects to remove the effect of any other unobserved pair-specific factors that may bias results. This requires us to remove any time invariant variables (i.e. our linguistic, historical and geographic ones), but acts as a valuable robustness check confirming that our key results on the non-uniformity of macroeconomic elasticities are not driven by failure to adequately purge the for pairwise effects. A comparison of regressions (1) and (3) makes plain that the coefficients on our macroeconomic variables and the interaction terms are very similar across the two specifications.[24]

We now investigate whether there are any significant differences in the in-

[24]The coefficient on the freedom of movement dummy variable does change, but in a model where interaction effects are included this has no direct interpretation

teraction depending on the source of the variation in the free movement dummy. We create new a dummy variable, C, equal to one if the FM dummy always takes the same value in each year for a given dyad, and is equal to zero if the FM dummy changes within the sample period for a given dyad. This allows us to split out the effects of cross sectional variation in the FM dummy (i.e. comparing country pairs which always had free movement vs those which never had free movement) and the effects of time series variation in the FM dummy (i.e. looking what happened “before" and “after" free movement was introduced). This creates a third set of variables, which is the interaction terms, interacted again with our new dummy. If any of these coefficients from this third set are significant, that would indicate differences in the effect of free movement across the two types.

Table 3: Cross section vs time series variation

	(1) Baseline		(2) C interactions		(3) C dummy	
ninflow						
Free movement	-5.810*	(3.102)	-7.718**	(3.097)	-6.991**	(3.035)
Origin GDP per capita	1.226***	(0.438)	1.208***	(0.433)	1.174***	(0.435)
Origin GDP per capita squared	-0.216***	(0.077)	-0.218***	(0.075)	-0.203***	(0.077)
Destination GDP per capita	3.848***	(1.277)	3.977***	(1.297)	3.918***	(1.290)
Origin unemployment rate	0.184**	(0.081)	0.181**	(0.081)	0.183**	(0.081)
Destination unemployment rate	-0.231*	(0.134)	-0.227*	(0.134)	-0.225*	(0.134)
Origin expected growth	0.155	(0.360)	0.151	(0.358)	0.153	(0.364)
Destination expected growth	2.070*	(1.079)	1.749	(1.095)	1.928*	(1.075)
Collective dismissal	-0.606***	(0.137)	-0.632***	(0.134)	-0.605***	(0.136)
Temporary employment	-0.182**	(0.085)	-0.189**	(0.086)	-0.189**	(0.085)
Distance	-0.616***	(0.063)	-0.611***	(0.069)	-0.571***	(0.065)
Shared spoken language	1.468***	(0.291)	1.528***	(0.287)	1.525***	(0.287)
Common native language	0.640**	(0.315)	0.717**	(0.313)	0.699**	(0.313)
Linguistic similarity	0.949***	(0.254)	0.877***	(0.252)	0.883***	(0.253)
1990 migrant stock	0.449***	(0.026)	0.443***	(0.026)	0.450***	(0.026)
Colonial link	0.279**	(0.118)	0.296***	(0.113)	0.281**	(0.116)
FM*Origin GDP per capita	2.570*	(1.495)	5.116**	(2.600)	3.164**	(1.466)
FM*Destination GDP per capita	1.315***	(0.475)	0.471	(0.937)	1.435***	(0.473)
FM*Origin expected growth	-3.180***	(0.910)	-8.916***	(3.454)	-3.020***	(0.888)
FM*Destination expected growth	-1.324	(1.965)	5.100	(3.214)	-0.546	(1.945)
FM*Origin GDP per capita squared	-0.590**	(0.233)	-0.863**	(0.371)	-0.681***	(0.228)
FM*Distance	0.345***	(0.118)	0.296	(0.181)	0.259**	(0.114)
FM*1990 migrant stock	-0.189***	(0.029)	-0.253***	(0.091)	-0.178***	(0.028)
Changes in dyad FM status			0.399*	(0.204)	0.576***	(0.147)
C*FM*Destination GDP per capita			0.902	(0.939)		
C*FM*Origin GDP per capita			-1.715	(2.149)		
C*FM*Origin GDP per capita squared			0.170	(0.298)		
C*FM*Destination expected growth			-5.595	(3.446)		
C*FM*Origin expected growth			6.451*	(3.468)		
C*FM*Distance			-0.003	(0.204)		
C*FM*1990 migrant stock			0.089	(0.091)		
N	66422		66422		66422	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; standard errors in brackets

Regression (2) presents the results: only one of these new interaction terms is individually significant and an F-test[25] strongly confirms that they are jointly insignificant. This new "constant freedom of movement" dummy is significant with a value of 0.3, and remains so even when the new interaction terms are dropped in regression (3). That suggests that there is a small additional boost to migration flows when restrictions are lifted, beyond what would be predicted by our baseline. But the coefficients on the other variables are very similar in specification (3) therefore we conclude our results are highly robust to allowing for a differential impact from cross-sectional vs time series variation.

[25]Testing the null hypothesis that these variables are jointly insignificant, it returns a p-value of 0.44

5 Conclusions

We estimate a gravity model of the determinants of migration flows. When we interact our explanatory variables with a freedom of movement dummy, we find that the elasticities of migration with respect to macroeconomic variables are not constant across country pairs. Under freedom of movement, the response to macroeconomic variables is stronger, and the response to distance and historical migrant stocks is weaker. The elasticity with regard to most linguistic and historical variables does remain constant. Alongside more traditional macro variables, we find a significant role for expected GDP growth- weaker origin growth and stronger destination growth are found to boost migration flows. In addition, greater labour market flexibility in destination countries is associated with higher inward migration.

These results challenge the widespread assumption that elasticities are constant across all country pairs. Additionally, the effect of freedom of movement itself is not uniform either. We find that for a range of macroeconomic variables, the elasticity is larger in absolute size under free movement, implying that migration flows are more sensitive to macro factors in such cases.

Interestingly however, the role of linguistic and historical factors appear to be the same across the two groups. The only two time invariant variables which did exhibit differences were distance and legacy migrant stock. The coefficients on latter suggest that network effects and previous historical patterns may be less important in shaping flows of people under free movement.

Perhaps more puzzling is the apparent reduction in the drag exerted by distance under free movement, since the factors which distance is said to proxy[26]– transport costs, ease of communication, psychic costs of separation, information costs, opportunity costs– would seem to be largely unaffected by changes to freedom of movement.

More generally, the results demonstrate that the issue of parameter constancy, rarely considered in the gravity literature, might be important. The large number of observations typically available in dyadic datasets make this possibility, this could be a fruitful avenue for future empirical work.

[26]See Greenwood (2016) for a discussion of these

Appendix: Additional Results

Figure A1: The effect of free movement with and without CCE estimation

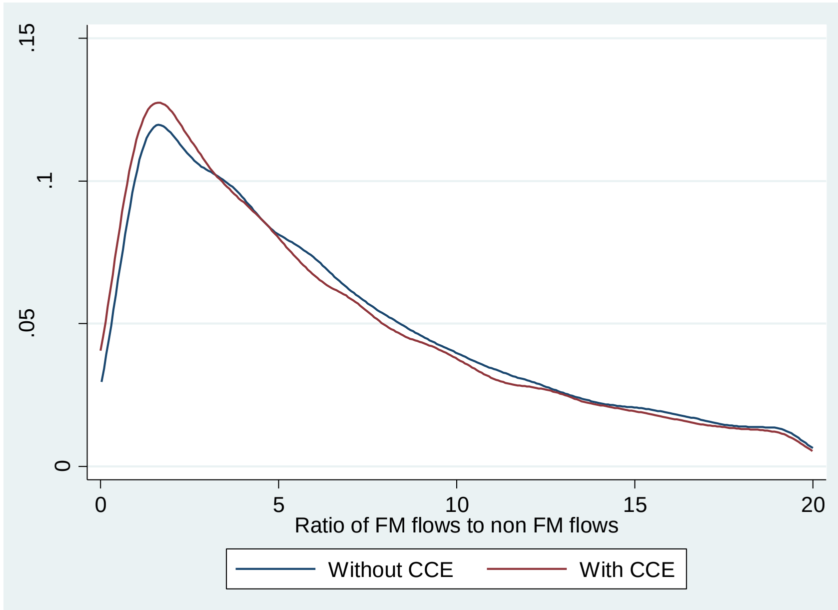


Table A1: Robustness to alternative distance measures

	(1)		(2)		(3)	
	Baseline		Weighted distance 1		Weighted distance 2	
ninflow						
Free movement	-5.810*	(3.102)	-5.851*	(3.251)	-5.773*	(3.161)
Origin GDP per capita	1.226***	(0.438)	1.223***	(0.437)	1.215***	(0.437)
Destination GDP per capita	3.848***	(1.277)	3.728***	(1.273)	3.715***	(1.272)
Origin unemployment rate	0.184**	(0.081)	0.185**	(0.080)	0.184**	(0.080)
Destination unemployment rate	-0.231*	(0.134)	-0.240*	(0.134)	-0.240*	(0.134)
Origin expected growth	0.155	(0.360)	0.153	(0.360)	0.151	(0.360)
Destination expected growth	2.070*	(1.079)	2.163**	(1.087)	2.161**	(1.085)
Origin GDP per capita squared	-0.216***	(0.077)	-0.216***	(0.077)	-0.215***	(0.077)
Collective dismissal	-0.606***	(0.137)	-0.602***	(0.137)	-0.606***	(0.137)
Temporary employment	-0.182**	(0.085)	-0.182**	(0.085)	-0.183**	(0.085)
Distance	-0.616***	(0.063)	-0.672***	(0.071)	-0.651***	(0.068)
Shared spoken language	1.468***	(0.291)	1.513***	(0.296)	1.514***	(0.295)
Common native language	0.640**	(0.315)	0.676**	(0.318)	0.689**	(0.316)
Linguistic similarity	0.949***	(0.254)	0.941***	(0.256)	0.937***	(0.255)
1990 migrant stock	0.449***	(0.026)	0.443***	(0.027)	0.441***	(0.027)
Colonial link	0.279**	(0.118)	0.280**	(0.119)	0.276**	(0.119)
FM*Origin GDP per capita	2.570*	(1.495)	2.341	(1.523)	2.372	(1.521)
FM*Destination GDP per capita	1.315***	(0.475)	1.306***	(0.480)	1.310***	(0.478)
FM*Origin expected growth	-3.180***	(0.910)	-3.197***	(0.912)	-3.188***	(0.915)
FM*Destination expected growth	-1.324	(1.965)	-1.483	(1.976)	-1.424	(1.975)
FM*Origin GDP per capita squared	-0.590**	(0.233)	-0.558**	(0.238)	-0.563**	(0.238)
FM*Distance	0.345***	(0.118)	0.402***	(0.143)	0.386***	(0.129)
FM*1990 migrant stock	-0.189***	(0.029)	-0.182***	(0.029)	-0.182***	(0.029)
N	66422		66422		66422	

Our baseline specification uses the distance between the capitals of each nation; regressions (2) and (3) uses the distance weighted by population of the 25 most populated cities where the CES parameter is set to 1 and -1 respectively. See Mayer and Zignago (2011) for more details.

Table A2: Robustness to alternative language measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Dyen 1st	Dyen Maj	Dyen	Max distance	Index All	Index Maj
ninflow							
Free movement	-5.810*	-7.620**	-5.978*	-7.950**	-6.287**	-4.955	-6.716**
	(3.102)	(3.612)	(3.477)	(3.432)	(3.032)	(3.167)	(3.056)
Origin GDP per capita	1.226***	0.941**	1.275**	1.838**	1.357***	1.224***	1.250***
	(0.438)	(0.425)	(0.521)	(0.817)	(0.439)	(0.431)	(0.430)
Destination GDP per capita	3.848***	4.806***	4.247***	4.940***	3.966***	3.773***	3.908***
	(1.277)	(1.450)	(1.455)	(1.519)	(1.280)	(1.280)	(1.274)
Origin unemployment rate	0.184**	0.302***	0.289***	0.369***	0.180**	0.183**	0.185**
	(0.081)	(0.076)	(0.075)	(0.088)	(0.086)	(0.080)	(0.081)
Destination unemployment rate	-0.231*	-0.228	-0.205	-0.322*	-0.276**	-0.235*	-0.228*
	(0.134)	(0.159)	(0.149)	(0.172)	(0.139)	(0.135)	(0.134)
Origin expected growth	0.155	-0.239	0.177	-0.300	0.107	0.136	0.165
	(0.360)	(0.409)	(0.442)	(0.968)	(0.384)	(0.360)	(0.358)
Destination expected growth	2.070*	1.994	1.553	3.176**	2.069*	2.113*	1.995*
	(1.079)	(1.399)	(1.184)	(1.481)	(1.225)	(1.084)	(1.081)
Origin GDP per capita squared	-0.216***	-0.152	-0.208**	-0.251*	-0.257***	-0.217***	-0.221***
	(0.077)	(0.093)	(0.092)	(0.149)	(0.078)	(0.076)	(0.076)
Collective dismissal	-0.606***	-0.654***	-0.565***	-0.618***	-0.618***	-0.602***	-0.612***
	(0.137)	(0.155)	(0.146)	(0.214)	(0.144)	(0.137)	(0.137)
Temporary employment	-0.182**	-0.312***	-0.200**	-0.298***	-0.178**	-0.178**	-0.182**
	(0.085)	(0.093)	(0.092)	(0.105)	(0.087)	(0.085)	(0.085)
Distance	-0.616***	-0.713***	-0.579***	-0.600***	-0.621***	-0.574***	-0.647***
	(0.063)	(0.081)	(0.073)	(0.079)	(0.065)	(0.063)	(0.064)
Shared spoken language	1.468***	1.135***	1.640***	0.431	1.280***	1.603***	1.417***
	(0.291)	(0.336)	(0.314)	(0.389)	(0.302)	(0.278)	(0.286)
Common native language	0.640**	0.628*	1.042***	-0.439	0.116	1.155***	0.568*
	(0.315)	(0.326)	(0.257)	(0.467)	(0.375)	(0.231)	(0.293)
Linguistic similarity	0.949***	0.001***	0.001***	0.002***	-0.014***	0.679***	1.231***
	(0.254)	(0.000)	(0.000)	(0.000)	(0.003)	(0.184)	(0.266)
1990 migrant stock	0.449***	0.440***	0.441***	0.482***	0.449***	0.437***	0.456***
	(0.026)	(0.034)	(0.031)	(0.031)	(0.027)	(0.027)	(0.026)
Colonial link	0.279**	0.480***	0.406***	0.594***	0.357***	0.215*	0.314***
	(0.118)	(0.145)	(0.121)	(0.148)	(0.115)	(0.115)	(0.115)
FM*Origin GDP per capita	2.570*	2.199	1.595	2.021	2.664*	2.181	2.656*
	(1.495)	(1.655)	(1.673)	(1.739)	(1.453)	(1.571)	(1.470)
FM*Destination GDP per capita	1.315***	1.277**	1.269**	1.699***	1.378***	1.313***	1.501***
	(0.475)	(0.557)	(0.524)	(0.630)	(0.469)	(0.464)	(0.471)
FM*Origin expected growth	-3.180***	-3.020***	-3.412***	-2.991**	-3.087***	-3.301***	-3.178***
	(0.910)	(0.983)	(0.993)	(1.169)	(0.925)	(0.910)	(0.921)
FM*Destination expected growth	-1.324	-0.178	0.053	-1.247	-1.167	-1.276	-1.332
	(1.965)	(2.084)	(2.055)	(2.064)	(1.969)	(1.943)	(1.945)
FM*Origin GDP per capita squared	-0.590**	-0.536**	-0.444*	-0.511*	-0.599***	-0.542**	-0.605***
	(0.233)	(0.258)	(0.261)	(0.279)	(0.228)	(0.244)	(0.230)
FM*Distance	0.345***	0.615***	0.558***	0.558***	0.342***	0.330***	0.374***
	(0.118)	(0.134)	(0.123)	(0.139)	(0.108)	(0.122)	(0.121)
FM*1990 migrant stock	-0.189***	-0.164***	-0.166***	-0.190***	-0.186***	-0.180***	-0.203***
	(0.029)	(0.034)	(0.032)	(0.034)	(0.029)	(0.028)	(0.030)
N	66422	33780	45716	23966	62548	66422	66422

Dyen Our baseline specification uses the measure of Melitz and Toubal (2014). Regressions (2) to (4) use the linguistic proximity measures of Dyen, Kruskal and Black (1992) for first, closest, and major languages respectively across a country pair. Regressions (5) to (7) use the

References

Adserà, Alícia, and Mariola Pytliková. 2015. "The Role of Language in Shaping International Migration." *Economic Journal*, 125(Feature Issue): F49–F81.

Anderson, James, and Eric van Wincoop. 2003. "Gravity with Gravitas: A solution to the border puzzle." *American Economic Review*, 93(1): 170–192.

Beine, Michel, Pauline Bourgeon, and Jean-charles Bricogne. 2013. "Aggregate Fluctuations and International Migration." *mimeo*.

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga. 2015. "A Practitioners' Guide to Gravity Models of International Migration." *The World Economy*, 39(4): 496–512.

Bertoli, Simone, and Jesus Fernández-Huertas Moraga. 2012. "Visa Policies, Networks and the Cliff at the Border." *IZA Discussion Paper*, 7094.

Bertoli, Simone, and Jesus Fernández-Huertas Moraga. 2013. "Multilateral Resistance to Migration." *Journal of Development Economics*, 102: 79–100.

Bertoli, Simone, Herbert Brücker, and Jesus Fernández-Huertas Moraga. 2013. "The European Crisis and Migration to Germany: Expectations and the Diversion of Migration Flows." *IZA Discussion Paper*, 7170.

Bratsburg, Berndt. 1995. "Legal versus Illegal U.S. Immigration and Source Country Characteristics." *Southern Economic Journal*, 61(3): 715–727.

Cigagna, Claudia, and Giovanni Sulis. 2015. "On the Potential Interaction Between Labour Market Institutions and Immigration Policies." IZA Discussion Papers 9016.

Clemens, Michael. 2014. "International Handbook on Migration and Economic Development." , ed. Robert Lucas, Chapter Does Development Reduce Migration?, 152–185. Edward Elgar.

Dustmann, Christian. 2002. "The Optimal Migration Duration and Activity Choice after Re-Migration." *Journal of Development Economics*, 4(15): 351–372.

Dustmann, Christian, and Joseph-Simon Görlach. 2016. "The economics of temporary migrations." *Journal of Economic Literature*, 54(1): 98–136.

Dustmann, Christian, and Oliver Kirchkamp. 2003. "Return Migration, Wage Differentials and the Optimal Migration Duration." *European Economic Review*, 2(1): 242–262.

Dyen, Isidore, Joseph Kruskal, and Paul Black. 1992. "An Indo-European classification: A lexicostatistical experiment." *Transactions of the American Philosophical Society*, 82(5).

Forte, Giuseppe, and Jonathan Portes. 2017. "Macroeconomic Determinands of International Migration to the UK." *IZA Discussion Paper*, 10802.

Greenwood, Michael. 2016. "International Handbook of Migration and Population Distribution." , ed. Michael White, Chapter Perspectives on Migration Theory: Economics, 31–40. Springer.

International Monetary Fund. 2017. "World Economic Outlook." IMF.

Keita, Sekou. 2016. "Bilateral Real Exchange Rates and Migration." *Applied Economics*, 48(31): 2937–2951.

Mayda, Anna Maria. 2010. "International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows." *Journal of Population Economics*, 23: 1249–1274.

Mayer, Thierry, and Soledad Zignago. 2011. "Notes on the CEPII's distances measures: the GeoDist Database." *CEPII Working Paper*, , (2011-25).

McFadden, Daniel. 1974. "Frontier in Econometrics." , ed. Paul Zarembka, Chapter Conditional logit analysis of qualitative choice behaviour, 105–142. New York: Academic Press.

Melitz, Jacques, and Farid Toubal. 2014. "Native Language, Spoken Language, Translation and Trade." *Journal of International Economics*, 93: 351–363.

Mesnard, Alice. 2004. "Temporary Migration and Capital Market Imperfections." *Oxford Economic Papers*, 2(1): 242–262.

Migali, Silvia. 2017. "Migration and Institutions: Evidence from Internal EU Mobility." *The World Econonmy*, Forthcoming.

Mishra, Prachi, and Antonio Spilimbergo. 2011. "Exchange Rates and Wages in an Integrated World." *American Economic Journal: Macroeconomics*, 3(4): 53–84.

OECD. 2017. "International Migration Database." https://www.imf.org/external/pubs/ft/weo/2017/01/weodata/index.aspx, last accessed 18 August 2017.

Ortega, Francesc, and Giovanni Peri. 2013. "The Effect of income immigration policies on international migration." *Migration Studies*, 1(1): 47–74.

Özden, Çağlar, Christopher Parsons, Maurice Schiff, and Terrie Walmsley. 2011. "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960-2000." *World Bank Economic Review*, 25(1): 12–56.

Pesaran, Hashem. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica*, 74(4): 967–1012.

Santos-Silva, and Silvana Tenreyro. 2006. "The Log of Gravity." *Review of Economic Statistics*, 88(4).

World Bank. 2017. "World Bank Open Data." `http://data.worldbank.org/indicator/SL.UEM.TOTL.ZS` Last accessed 18 August 2017.

Yang, Dean. 2017. "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *Economic Journal*, 118(528): 591–630.