University of Sussex
Sussex Centre for Migration Research

Quantifying the trade-migration nexus of the enlarged EU

A Comedy of Errors or Much Ado About Nothing?

Sussex Migration Working Paper no. 27

Chris Parsons
Sussex Centre for Migration Research

Abstract

Gravity models, as applied in the growing number of immigrant-link studies, have revealed a robust and positive relationship between immigration and bilateral trade flows. This paper attempts to further the immigrant-link literature by applying, for the first time, this methodology to the European Union. Specifically the paper seeks to evaluate the robustness of the findings from this strand of literature; to quantify the impact upon EU-15 bilateral trade flows, of East-West European immigration, and to identify the underlying mechanisms underpinning this relationship. The results indicate that Eastern European immigrants exert a positive influence on both EU-15 imports and exports. It is predicted that a 10% rise in Eastern European immigration will increase EU-15 imports from these countries by 1.4% and EU-15 exports by 1.2%. These results indicate that immigrants' demand for native products outweighs the increase in trade associated with immigrants forming business-links between European trading partners.

1. Introduction and Background

On May 1st 2004, the European Union, originally consisting of a core of six nation states, undertook its largest expansion to date, swelling in size from 15[1] to 25 members. This, the seventh stage of the expansion process, began in March 1998; and The Treaty of Accession 2003 was signed on April 16th, guaranteeing membership in the Union for the ten accession countries[2]. Further expansion is possible, and of the remaining potential candidates, Romania and Bulgaria hope to join in 2007, Croatia was approved candidature in March 2004, while the Former Yugoslav Republic of Macedonia applied for EU membership in March 2004. Turkey awaits the European Council's verdict as to whether it is a viable candidate country. The current EU zone has increased in terms of geographical coverage by 34%, its' population has swelled by 105 million, and nine official languages are now additionally spoken.

Perhaps the main arguments for and against the further expansion of the EU, are the positive economic gains involved, and those negative forces commonly associated with increased East-West intra-European migration. The majority of any potential economic benefits will likely accrue to the newer Eastern states, those starting from a far lower base. Baldwin et al (1997) cautiously estimated the EU-15 to gain €10 billion with the accession-10 gaining €23 billion. Similarly the Directorate General for Economic and Financial Affairs (Free 2001), estimated a growth in GDP of accession countries of between 1.3% and 2.1% with the members of the EU-15 realising a 0.7% GDP gain. Increased bilateral trade, post-enlargement, is believed to constitute a significant part of this economic gain. In 2000 the EU had a trade surplus of €17 billion with the accession-10[3], and this is likely to grow significantly in the future.

Immigration happens for many reasons. In the context of intra-EU, East-West migration, this will likely occur due to immigrants evading civil unrest, searching for superior economic rewards, and attempting to reunite their families. Intra-EU immigration displays a steady and increasing long-run trend. Between 1965 and 1990 the share of migrants in Western Europe relative to the share of the native regional population, rose 2.5 % (Nelson et al 2000). Boeri et al (2000) predict an initial influx, post May 1st, of approximately 335,000 migrants into the EU, with an additional 220,000 moving annually. This trend has raised concerns however. The great disparities (in wages, employment opportunities etc), which exist between the old and new members of the Union have fuelled fears of waves of Eastern Europeans emigrating to more traditional European nations; 'stealing' employment, straining the public purses, lowering wages and competing for already sparse cheap accommodation. The latest Eurobarometer results (Spring 2004), a measure of public opinion throughout EU countries; shows that although UK residents believe freedom of movement of labour to be the most significant advantage of EU membership, 54% of UK residents believe the EU plays a negative role in immigration matters, 14% higher than the EU average.

These key issues are not mutually exclusive however, and are perhaps ironically, the greatest forces for further EU-cohesion. European citizens are subject to a free movement across Europe, and furthermore:

> "*International trade is one of the most expedient economic factors in pushing economies into integration*" (Paas 2002a: 1)

The immigrant-link literature brings these two issues together in attempting to evaluate the impact of immigration on bilateral trade flows. This paper contributes to the immigrant-link literature by implementing a gravity model to analyse these effects in relation to EU-15 trade, and immigration from EU-expansion[4] countries. Previous research indicates that typically positive immigrant-links are formed between trading

1 Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, UK.

2 The Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovenia and Slovakia.

3 European Commission's own estimates refer to http://europa.eu.int.

4 Throughout this paper all those newly acceded and potential EU entrant states will be referred to as EU-expansion countries

nations, which, either by the lowering of transaction costs, or via preference effects, boost bilateral trade volumes. Positive results, which in light of previous work are anticipated, would further evidence the robust findings of the immigrant literature, and lend credence to further EU-expansion.

Section 2 provides a comprehensive review of the literature to date, highlighting the disparities between the various approaches, and the subsequent influence these have on the results obtained. Section 3 offers a discussion of the theoretical issues involved, a background to the theory, and commentary on the variables commonly used in gravity models. Section 4 describes the estimated empirical model. Section 5 provides a discussion of the data sources used, of the limitations of the data collected, and details of the analysed panel. Section 6 analyses the results in light of comparable studies, and quantifies the effects of immigration on EU bilateral trade flows, additionally investigating the robustness of the results obtained. Lastly, in Section 7, conclusions are drawn and suggestions are made for further research in the field.

2. Literature review

It would be an unenviable task to specify the creator of gravity modelling since the underlying principles are derived from Isaac Newton's (1687) 'Law of Universal Gravitation', that was not invented, but instead discovered; since it is a binding law that governs nature. It was centuries until these principles were directly applied in international economics, despite their earlier use by experts in other subject areas. Early academic synthesises of the 'gravity principles' and migration include: Carey (1858), who observed the influence of gravitational forces in the social sciences; and the geographer Ravenstein (1885) who formulated a number of 'laws of migration' having first categorised migrants according to their various motivations for seeking alternative residence. These laws laid the foundations for his gravity principles that still largely hold true today. Early applications include Zipf (1946) who brought Newton's gravitational principles into the realm of the social sciences by successfully testing the underlying predictions of the gravity model in relation to migration between major cities; and Lowry (1966) whose model draws together gravity and wage determination principles into a single gravity model predicting migration levels. Gravity models have since been used to model numerous, 'social interactions', Head (2000).

The application of Gravity models in economics however, is attributed to Tinbergen (1962) and Poyhonen (1963) who first applied the gravity principles to examine international trade flows. Specifically, in this regard, gravity modelling applies Newton's Law of Gravity to provide an empirically tractable framework; positing a log-linear relationship between trade volumes, trading distances and the importing and exporting countries' GDP. Trade is treated as analogous to the attractive force between two particles. Trade volumes are therefore predicted to be an increasing function of the size of countries, but a decreasing function of the distance between them. Linnemann (1966), continuing Poyhonen's work, is commonly cited as the first to provide a detailed and thorough economic application of gravity modelling. He utilised a more expansive set of variables, and applied a Walrasian general equilibrium model.

In recent years there has been a significant revival of interest in these models for a number of reasons. Firstly, gravity models have experienced remarkable empirical success in predicting bilateral trade flows in many geographical regions worldwide. Moreover, gravity modelling lends itself to explain many additional complexities including: currency unions (Frankel et al 2000, 2002), regional trade agreements (Cernat et al 2003), commonality of language (Hutchinson 2002), the effects of foreign aid on trade levels (Helliwell 1999) and the effects of immigration on trade flows; the focus of this paper. Secondly, whereas once significant criticism was levelled at gravity models for not having a solid theoretical underpinning, now these foundations have been provided from a variety of sources[5]. Lastly, in today's world of increasing inter-disciplinary co-operation there is a growing desire to:

> "*Treat countries or regions as physically placed at particular locations rather than as disembodied constructs*" (Frankel 1997: 49)

Gravity models provide both the geographer and the economist with a flexible tool for accomplishing this.

It was Gould (1994) that first introduced using gravity modelling to investigate the affects of immigration upon bilateral trade flows. This represented a marked improvement over previous models that treated immigration as equivalent to labour force growth; and thus allowed the effects of immigration upon (bilateral) trade flows to be quantified for the first time. This strand of thought therefore represents a recent addition to the economics literature. In his seminal paper Gould asserts his motivation as the documented coincident movements between immigration and

[5] Readers should refer to theory section for the relevant discussion.

trade. Specifically he notes that immigrants are often able to establish themselves in the labour market into which they emigrate, by finding a niche. This - drawing on case-study evidence from Korea (Min 1990) - Gould believes, pertains to positive human-capital externalities that increase the propensity to trade between the two countries among which migration has taken place. It was this insight that proved the catalyst for subsequent research in the field.

To date, to the authors' knowledge, there have been only eight studies applying the gravity model to investigate the mechanisms underlying the immigration trade nexus - or to be precise, in the case of Raulch and Trindade (2002), the effects of the presence of immigrants on trade. All these papers provide evidence of a statistically significant relationship between immigration and trade, supporting the notion that immigrants facilitate trade creation. These include Gould's' original study, and the two papers by Dunlevy and Hutchinson, who investigated the effects of immigration on imports and exports separately. The studies reviewed here are: Gould (1994), Head and Ries (1998), Dunlevy and Hutchinson (1999 & 2001), Girma & Yu (2002), Combes et al (2002), Rauch and Trindade (2002), and Wagner, Head and Ries (2002). There exist significant disparities between these studies in terms of econometric specification, geographical coverage, the period under study, and regression technique. The results from these studies thus vary quite significantly with different aspects of the immigration-trade nexus being identified. This section aims to describe the mechanisms via which immigration is hypothesised to affect bilateral trade flows, and to report the main findings of the literature in light of these immigrant-links. The main similarities and differences between the studies will be highlighted, and discussion will be offered as to how the inconsistencies in the methodologies utilised give rise to the variations in the results obtained.

A number of authors have provided explanations as to the underlying mechanisms via which immigrants influence trade, though this list is by no means exhaustive. Gould originally postulated but two key mechanisms, though additional linkages have subsequently been referred to. Firstly, Gould's immigrant-preference hypothesis states migrants are likely to have a demand bias for home products, because of preferences they have developed when growing up in their native land. This direct linkage will unambiguously raise host country imports. Secondly, Gould's immigrant-link hypothesis states immigration will likely lower the transaction costs of trade between two trading nations. These costs may represent a serious impediment to trade (McCallum 1995). Immigrants possess additional knowledge of home markets, which assuming heterogeneous products between the two countries, may lower the costs of obtaining foreign market information. If immigrants have previously established business contacts in their domestic economy both the costs of enforcing and of negotiating contracts may be reduced. Immigrants' fluency in their native language may also lower those costs to trade of communication barriers. Crucially this second effect will affect both imports and exports of the host country. This, Wagner et al (2002) refer to as the information-hypothesis; the lowering of barriers to trade due to the additional knowledge possessed by immigrants. This includes naivety of business opportunities, knowledge of local laws and culture, and higher levels of trust between home nation exporters and immigrants. Both hypotheses are important, since they are both associated with higher volumes of trade. The immigrant-link hypothesis may be considered more important though as it crucially influences both imports and exports.

Rauch and Trindade hypothesise that the presence of ethnic Chinese workers, makes it easier to enforce contracts within business networks, created by immigrants and exporters from the home nation. This notion is not too dissimilar to Wagner et al's suggestion of more trusting relationships between migrant workers and businesses in the home nation, and paralleling the information-hypothesis is believed to facilitate trade creation in terms of both imports and exports. Subsequent authors note the existence of additional immigrant-link effects. For example Dunlevy and Hutchinson describe ambiguous trade diverging effects. These additional mechanisms all broadly fall under the umbrella headings of Gould's original hypotheses; and thus for the purposes of this paper we confine our attentions solely to the preference and information hypotheses.

The literature consistently attempts to exploit the differing predictions of the two hypotheses. Disparities between the magnitudes of the import and export elasticities, (with respect to immigration), provide an indication as to the dominance of one of the prevailing hypotheses. If imports are more strongly influenced by immigration, one would expect immigrant preference effects to dominate. Conversely if exports are influenced more significantly by immigration, this would evidence the prevalence of the immigrant-link effects. Indeed if immigration affects both imports and exports significantly this could evidence both hypotheses.

It is not difficult to envisage additional mechanisms via which immigrants may affect bilateral trade flows however; mechanisms absent from any of the literature to date. Remittances, for example, would likely exert influence on any model of bilateral trade and immigration. Consider an immigrant who sends remittances home to help support their family. If immigration takes place to a specific country it seems likely that their family may also have a bias for those countries' products. For example a Mexican, who aspires to working in the U.S. to share in the American dream, may well have a preference for American products and thus their family may well too. The increase in the families' disposable income from the addition of remittances would therefore likely raise imports of the home nation from the host nation, though this effect is likely to be small. The increased disposable income may also enable additional funds to be invested in domestic family firms boosting the home nations exports and thus possibly the imports of the host nation. Though encapsulating similar effects on bilateral trade flows to the immigrant-link hypothesis (i.e. effects both imports and exports); this mechanism does not work through the lowering of transaction costs via information gains, and thus remains separate. As the literature omits remittances, the analysis is not undermined; rather the inclusion of remittances would identify additional dynamic processes.

The literature correctly recognizes that immigrants will likely have a bias for home products, but fails to recognize that the immigrant-preference hypothesis will likely extend beyond the borders of the home country to neighbouring regions. For example, a migrant from France will no doubt have a preference for French-made items, but would also likely have a preference for other European products. They may prefer Belgium chocolate to chocolate produced in the U.K. or German wine over that produced in Australia. This effect, though small, will likely account for some of the additional host nation imports from the home region. The feasibility of testing for either of these additional hypotheses may be called into question though. It is likely that due to the data constraints faced by the authors that these mechanisms were omitted in their entirety from the literature. The very existence of additional mechanisms underlying the immigrant-trade relationship however, ensures that it is not as easy to distinguish between the two main hypotheses as the immigrant-link literature would imply. It is probable that there exist many different effects underlying the analysis. To differentiate between them accurately, within the confines of the data available is virtually impossible. For this reason, the focus of this paper is to investigate the magnitudes of the estimated elasticities and use these to state the influence of one hypothesis over the other.

All of the studies under review estimate, either directly or indirectly, immigrant-link effects. Further, each examines additional complexities underlying the immigrant-trade nexus. Gould utilises data for the U.S. and 47 of its trading partners, for the years 1970-86. He additionally examines the effects of different commodity groups, and the length of stay and skill levels of migrants. Head and Ries, the first paper to estimate immigrant-link effects utilising the base line gravity model, look at Canada between 1980-92, utilising data for Canada and 136 of its trading partners. Additionally motivated by the 1976 change to the Canadian immigration laws (i.e. the introduction of the business immigration programme), the authors examine different classifications of immigrants. Dunlevy and Hutchinson, in two separate papers, first investigate U.S. imports over the period 1870-1910, and in their later study, estimate U.S. exports over the same period; using data for the U.S. and 17 trading partners. Additionally they investigate immigrant-link effects on specific commodities, and the effects of a common language; whilst distinguishing between trade with old, new, and non-European countries. Girma and Yu for the period 1981-1993, examine UK trade with 48 trading partners. Studying the UK also provided these authors with the opportunity to distinguish between commonwealth and non-commonwealth trading partners. They hypothesise that former colonies have more anglicised institutions, culture and law, which would facilitate greater degrees of trade creation. Wagner, and Head and Ries - examining Canada for the second time, use data from 1992-95 to investigate Canadian provincial trade with 160 countries. Additionally these authors investigate the influence of a common language between trading partners, and further introduce a random encounter model specification. The last two studies to date, though related, differ fundamentally from those studies already cited, and thus receive less attention for the purposes of this paper. Rauch and Trindade, rather than investigating the effects of immigration on bilateral trade flows directly, instead examine what effect different concentrations of ethnic Chinese workers from home nations have on the host nation's trade levels. Altogether 63 countries are investigated for both 1980 and 1990. Combes et al examine trade and immigrant movements between 93 French 'departments' in 1993. Therefore, in stark contrast to the other studies reviewed here Combes et al quantify intra-

national trade flows, as opposed to international flows.

The multitude of results from those studies reviewed, though insightful, do not all lend themselves to direct comparison, as a whole range of estimates are obtained. The studies taken as a whole do provide evidence of a definitive and robust relationship between immigration and international trade for a wide variety of specifications. Gould reports a positive coefficient on his immigrant information variable. The largest coefficients he reports are on consumer-manufactured exports, with the smallest coefficients on aggregate and producer imports. This, Gould believes, is due to the homogenous nature of these products such that additional information about foreign markets does not substantially benefit trade flows. Generally, Gould finds immigration impacts upon exports the most, and Gould cites this as evidence in favour of the dominance of the immigrant-link hypothesis. Gould's inclusion of variables for the relative skill level of immigrants and two variables for the length of stay of immigrants (length of stay and its square) provide scope for additional analysis. Gould concludes that relative skill levels of immigrants have but a weak effect on immigrant-links. Immigrant-link effects increase over time at an ever-decreasing rate for bilateral imports, though this effect is delayed some years before any increase in immigrant-effects on export flows is detected.

Girma and Yu also observe a smaller coefficient for imports than exports, evidencing the immigrant-link hypothesis over any preference or taste effects. Interestingly they fail to find a significant effect of immigration into the UK from commonwealth countries. This, Girma and Yu attribute to a trade-diverging effect during the period under study, a 'trade-substitution' effect (Diaz-Alejandro 1970). Due to this peculiarity in the results obtained from Girma and Yu's paper, for immigration from commonwealth countries, these estimates are omitted from any further analysis. Combes et al study, as noted already, is fundamentally different to the others. These authors look at intra-national migration; all the immigrant-link effects brought about by migrants possessing country specific information should be negated. All French people should have a reasonable knowledge of French laws, culture, business opportunities etc. However, contrary to expectations they find that immigrants within a country can also positively influence trade. Moreover they find French migrants exert greatest influence on exports, providing further support for the immigrant-information hypothesis.

Head and Ries, in contrast to the results described so far, observe that immigration exerts strongest influence on imports. Preference effects dominate in this instance. Further, by disaggregating the immigrant stock variable, the authors are able to conclude that independent migrants have more effect on bilateral trade than either family or refugee immigrants. This is hardly surprising, since it is the independents that are more intensively screened by immigration officials. This group is only allowed entry if they are able to demonstrate that they have attained the minimum entry conditions. They are also those migrants with the closest ties to their home nations.

Similarly, Dunlevy and Hutchinson's first study, reports a strong pro-import immigrant effect. The authors fail to report any pro-trade effect for new Europe however, a very surprising result as newer European nationals should have more divergent tastes relative to the average American, whose ancestry almost certainly derives from older European nations. This may be due to immigrants failing to raise enough capital such that the formation of business links was prevented. Turning to their estimation by commodity groupings, the authors correctly hypothesise that immigrant-preference effects should be stronger on consumer goods and processed foodstuffs, than for other categories of expenditure. The paper provides strong evidence that this is the case, lending additional weight to the presence of the immigrant-preference hypothesis. Similar results are obtained when the data is estimated by individual commodities. Correspondingly, in their follow-up study, Dunlevy and Hutchinson find a significant pro-export immigrant effect, though this is significantly smaller than their previously reported pro-import effect. This pro-export immigrant effect had a large variance across countries however, and was significantly larger for specific countries than on the aggregate. Taken together, these results show that taste effects are most significant. Wagner, Head and Ries also find evidence that the import elasticities with respect to immigration are greater than their estimates for exports.

Finally, Rauch and Trindade find a definitive positive relationship between the concentrations of ethnic Chinese residents, and trading volumes between country pairs. Reinforcing the conclusions of both Gould, and Dunlevy and Hutchinson, they find that the positive effect of additional knowledge possessed by migrants is most pronounced in the trade of differentiated goods. Furthermore, Rauch and Trindade find support for their enforcement hypothesis as ethnic Chinese workers are found to also promote trade in homogenous products.

The greatest disparities between the results in the papers under review, other than the different conclusions drawn with regard to the underlying hypotheses, is between the magnitudes of the export and import trade elasticities with respect to immigration. Wagner et al (2002) provide a useful summary of the import and export elasticities (with respect to immigration) observed in the papers under review, when the full sample is estimated. This facilitates a direct comparison of the studies to date, though Wagner et al make a few additional calculations to estimate the elasticities from the Gould, and Rauch and Trindade papers[6]. Table 1 in Annex 1 replicates Wagner et al's comparison table, additionally containing their results.

The broad range of elasticities is easily observed from Table 1. Except for the studies by Gould and Rauch and Trindade, all use log-linear estimations i.e. they estimate a constant elasticity relationship by regressing the log of bilateral exports/imports on the log of the stock of immigrants (plus additional variables). This specification is useful in terms of facilitating comparisons, and reading off coefficients as elasticities, but may not be an accurate description of reality. Gould originally specified a function to incorporate decreasing marginal returns to immigrants. This may in part explain the particularly low elasticities obtained by this author.

As Wagner et al (2002) correctly note, the range in those estimates obtained can also be partly explained by the alternative estimation techniques utilised throughout the various papers. While Head and Ries, Dunlevy and Hutchinson, Wagner et al, Girma and Yu, Rauch and Trindade, and Combes et al, use cross-sectional data; Gould use time series estimation. Cross-sectional studies are likely to bias estimates upwards due to unobserved effects, while conversely, fixed effect estimation may bias estimates downwards due to, the loss of informational content and larger measurement errors. Additional explanations include the multitude of samples sizes implemented, the different sample periods under investigation, and alternative econometric specifications.

Causality is additionally an issue of key concern, for if causation does not run from immigrants to trade then the analyses of all the studies under review is undermined.

> "*If the magnitude of the immigration effect is systematically related to factors that theory indicates it should be related to, then there is evidence that the measured effects are not simply spurious but instead represent a causal relationship*" (Wagner et al., 2002: 511)

This certainly seems to be the case in light of the previous discussion and the results reported in Table 1. For example, as Wagner et al note, the information hypothesis would suggest that immigrants should utilise their knowledge in areas where the returns are greatest. This intuition is corroborated by both Gould and Dunlevy and Hutchinson's studies', who find that largest estimated coefficients on consumer manufactured goods and finished and semi-finished products respectively. Furthermore, Gould uses Granger-causality tests to investigate his assumption that causality runs from immigration to trade. His results corroborate his intuition. Wagner et al provide many additional examples of how immigrant-link effects vary consistently with theory providing ample evidence against merely a spurious relationship between immigration and trade.

3. Theory

It is vital in any applied empirical work to carefully choose the correct type of model applicable for the subject matter; to include all the relevant explanatory variables, to control for various factors, and to utilise the correct functional form. Failures to do this will likely bias results, and in the case of omitted variables that simultaneously encapsulate the effects of both trade and immigration, bias results upwards (Wagner et al 2002). The gravity model has consistently been appropriately applied to international trade, in a wide variety of contexts, and therefore an augmented gravity model was chosen for the purposes of this paper. For decades these models were criticised for a lack of theoretical justification. In recent years a whole plethora of alternative theoretical derivations have been formulated however, these rooted in many competing strands of economic theory. This section discusses the evolution of the theoretical debate regarding the justifications for gravity modelling; and though not offering a detailed account of each paper, will provide background as to the main schools of thought. It will provide

[6] Wagner et al take a single year 1986, and calculate approximate elasticities for Gould's study. Similarly, for Rauch and Trindade's study, which implements an alternative estimation procedure – they estimate the log of aggregated exports and imports as the dependent variable and thus their main explanatory variable, rather than immigrant stock, is instead the sum of the two stocks of ethnic Chinese workers in the two countries, relative to the total populations as a whole - Wagner, Head and Ries estimate a single elasticity. Readers should refer to Wagner et al (2002) for further details of these calculations.

discussion as to why particular variables are important, derive the basic gravity model of trade, and identify the reduced form to be estimated, a variant of the basic gravity approach - augmented to include additional dimensions necessary for the purposes of this paper.

Paas (2002a) provides a categorisation of the various theoretical backgrounds to the gravity model; those based in regional science or economic geography, those whose foundations lie in microeconomics, predominantly based on utility maximisation, and those that derive from trade theory. Krugman (1991a, 1991b) provides theoretical justification for how geographical proximity can lead to production agglomeration in the regionalisation process, thus biasing international trade flows. In a similar vein, Frankel (1997) provides evidence that geographical proximity, in part, explains regionalisation. Writers adopting the more formal economic approach to theorising the justifications for the gravity model; those continuing the work of Tinbergen and Linnemann, notably include Anderson (1979) and Bergstrand (1985). The idea behind these models is that consumers maximize their utility subject to a binding budget constraint. The point of maximization, which indicates the optimal allocation of resources, can be identified by assuming various preferences that underlie the curvature of the individual's indifference curve. It is therefore necessary to assume product differentiation. Anderson, utilising a Linear Expenditure System (Stone 1954) assumes a CES (constant elasticity of substitution) set of preferences, for a country's goods, which are uniform across importing countries. Bergstrand introduced a more flexible and realistic utility function, including additional price variables and preference effects. Gravity models based on trade theory can be distinguished by the way in which product specialisation is obtained in equilibrium according to Evenett and Keller (1998). Either those based on technology differences (Ricardian Model), differences in factor endowments (Hecksher-Ohlin Model), or increasing returns to scale (New Trade Theories). Seminal papers include Helpman and Krugman (1985), and Helpman (1987). As the Hechsher-Ohlin theories has traditionally been viewed as suitable for providing explanations for North-South trade (Frankel 1997) it does not seem unreasonable to suppose that gravity models would be appropriate for evaluating Western-Eastern EU trade; additionally:

> "(gravity models) *are also valid when exploring the changes in international trade patterns during transition and EU eastward enlargement processes*" (Paas, 2002a: 1)

It is the belief of the author, that in spite of the many theoretical justifications for the gravity model, the underlying fact that it works so well empirically, as applied to international trade, is most likely because it is based on an underlying law of nature; which is appropriately applied to an alternative setting. Indeed it would be extremely difficult to imagine another instance where a natural law has been so appropriately applied with such efficacy, outside of the realms of the natural sciences; other than the statement being a tautology. Imports and exports for example, could be said to equate to an equal and opposite force (thus reflecting Newton's third Law of Motion), but this is true by definition. Deordorff (1998: 7) believes such models to merely confirm a 'fact of life', not evidencing one theory over another. Typically gravity models as applied in international economics have incredible predictive power, typically having an R-squared of between 70-95% (Paas 2002a). Eichengreen and Irwin (1998: 34) conclude "few aggregate relationships are as robust". Gravity models could almost be a justification for themselves. Their use continued for decades prior to them being formally derived. Indeed the only reason for theory being formulated was that they showed such potential, but lacked serious economic reasoning. The theoretical foundations for these models were merely an afterthought.

It is worth noting that despite the multitude of formal theoretical justifications for gravity models, many studies to date are only barely related to the theory, instead inserting variables into their functional forms in an ad-hoc fashion. Frankel (1997) believes this to be an acceptable practice, as many of the theoretical justifications to date omit important aspects of bilateral trade. Helpman (1987), for example does not include distance in his model whatsoever, and thus does not create a true foundation for the gravity model. Frankel cites geographical features as a common omission from gravity models. Indeed the majority of immigrant-link gravity models studies, those detailed in the literature review, though claiming to be based in theory are commonly only loosely based on theory typically simply including variables that the theory suggests are important. This paper in light of the preceding comments, adopts an ad-hoc stance to estimation, and draws upon both the gravity model and immigrant-link literature when identifying variables of relevance.

Here Head (2000) is followed in deriving the basic gravity model as applied to trade. Newton's original 'Law of Gravitation' is, formally:

$$F_{ij} = G \frac{M_i M_j}{D_{ij}^2} \quad (1)$$

Where:

F_{ij} = Attractive force between two bodies, i and j.

G = Gravitational constant, a constant of proportionality.

M_i, M_j = Masses of bodies i and j respectively.

D_{ij} = Distance between i and j.

Head notes (1) as is typically estimated in the social sciences:

$$F_{ij} = G \frac{M_i^{\alpha} M_j^{\beta}}{D_{ij}^{\theta}} \quad (2)$$

[7]

Where additionally:

F_{ij} = Flow from country i to j

G = Gravitational constant, a constant of proportionality

M_i, M_j = Respective (economic) masses of the two countries

D_{ij} = Measure of distance between country i and j

Head continues by equating the national incomes of the countries i and j with the forces of supply and demand, with distance reflecting a measure of those costs imposed by longer trading distances. It follows that if; s_{ij}, the share of M_j spent on goods from country i, which lies between 0 and 1, increases if country i produces a wider variety of goods (denoted by n), or a higher average quality of goods (denoted μ); and decreases in the distance between the trading pair, then formally:

$$s_{ij} = \frac{g(\mu_i, n_i, D_{ij})}{\sum_l g(\mu_l, n_l, D_{ij})} \quad (3)$$

Drawing on the Dixit and Stiglitz (1977) model of monopolistic competition between differentiated but symmetric firms, such that $\mu_i = 1$, n_i is proportional to M_i, and trade is a power function of distance, Head notes (3) can be rewritten as:

$$s_{ij} = M_i D_{ij}^{-\theta} R_j \quad (4)$$

Where:

$$R_j = \frac{1}{\left(\sum_l \left(Y_l / D_{lj}^{-\theta}\right)\right)} \quad (5)$$

Rearranging Head obtains:

$$F_{ij} = R_j \frac{M_i M_j}{D_{ij}^{\theta}} \quad (6)$$

Where:

F_{ij} = Flow of Trade between country i and country j

R_j = Gravitational constant, a constant of proportionality

M_i, M_j = GDPs of the importing and exporting country respectively

D_{ij} = Measure of distance between country i and country j

Equation (6) constitutes the base-line gravity model as applied in international trade. This predicts the volume of trade (imports and exports) to be proportionally related to the size of the two trading nations, but inversely related to the distance between them. These two variables thus constitute the 'basic' gravity model of international trade that takes account of the 'gravitational' attractive force between two (trading) masses. Due to the additive property of logarithms, Head suggests natural logs can be taken throughout (6) yielding the following functional form suitable for estimation:

(7)

$$\ln F_{ij} = \alpha \ln M_i + \beta \ln M_j - \theta \ln D_{ij} + \rho \ln R_j + \varepsilon_{ij}$$

Gross Domestic Product and population are both commonly used to measure a countries economic mass, and it is important to include measures for both countries i and j. A countries' land area can also be used to proxy for economic mass, this more a measure of the abundance of natural resources belonging to a country rather than of it's economic stature. Although this measure seems intuitive, for the U.S. Russia, China, and Brazil all have large factor endowments of natural resources, this measure could capture the effect that larger countries are more self-sufficient and thus could trade less than they otherwise would. Factor endowments can also be used to control for bilateral trading size, but Leamer (1974), concludes this measure to be inferior to the standard GDP and population measures.

In terms of gravity models as applied to international trade, the distance between two economic masses (countries) represents a proxy measure for transportation costs, historical ties, the time taken for the for delivery of a shipment of goods, and additionally, a measure of transaction costs. In general, larger distances generate higher communication barriers,

[7] Equation (2) is equivalent to Newton's Law when α = β = 1 and θ = 2.

linguistically, culturally as well as legally. Leamer and Stern (1970) believe that any estimation of bilateral trade omitting transport costs is meaningless.

The main difference between equations (1) and (6) is the replacement of the gravitational constant with R_j, – a measure for the remoteness of a country from world markets. Head therefore believes this to be a core variable of the 'basic' gravity model and offers a poignant example to why this variable should be included. Consider trade between two pairs of countries, Australia and New Zealand, and Austria and Portugal. Both of these pairs have similar GDP products, differing by approximately 20% less in the case of Australia and New Zealand; with capital cities approximately the same distance apart. Without the inclusion of the remoteness variable the model would predict that the European pair would trade slightly more. In reality, trade between the two southern hemisphere nations is about nine times larger than between their northern counterparts. Just as the gravitational constant is often assumed constant in calculations using Newton's original equation when applied in international trade R_j is typically assumed constant. There is no justification for this however. One would anticipate the greater the accessibility of a trading nation has to the world markets, i.e. the less remote it is, the more it will trade with localised sources of goods, diverting trade away from more distant trading partners. Head criticises Helliwell's (1998) measure of remoteness believing it to give too large a weight to small distant countries, instead proposing the reciprocal of this measure (i.e. equation (5)), as this formulation gives less weight to those small countries very distant from world markets.

A consensus of modern literary opinion recommends additional variables suitable for international trade gravity models, what Frankel (1997) refers to as the 'full' gravity model. Some authors include an additional estimate for per capita GDP in bilateral trade gravity models. Other specifications omit per capita income measures altogether i.e. Krugman-Helpman type models. Frankel demonstrates that including GNP and per capita GNP in a bilateral trade gravity model is equivalent to having GNP and population but prefers to estimate a measure of per capita GDP as it is more insightful. There are numerous arguments for the inclusion of a function of per capita income into a gravity model. It can be used as a measure of relative income, or the level of development of an economy – a proxy for the standard of living. Frankel believes per capita GDP to indicate the level of development of a country, not simply because it is wealthier in absolute terms, but also because of global trends in trade liberalisation. In today's political economic climate of WTO agreements and IMF conditionalities it is not unreasonable to suppose that poorer countries, by-and-large, have higher tariff barriers. These barriers will likely be correlated with per capita incomes. Frankel additionally suggests that part of the reason poorer nations have higher barriers to trade is because lower-income countries attempt to raise a larger fraction of their outgoings via tariffs and quotas whilst richer countries can impose, more effectively, indirect and direct taxes to raise revenues. Head (2000) notes opposed to these forces is the fact that in more developed nations the service sector accounts for a higher proportion of national income, resulting in a lower trade volume in goods.

Frankel suggests that the product of the per capita incomes of trading nations be included in bilateral trade gravity models; for the underlying theoretical justifications make differing predictions with regard to the expected sign on its' coefficient. Frankel argues that if the factors of production in the Hecksher-Ohlin model are capital and labour then these type of models would make the prediction that those nations with dissimilar capital-labour ratios, proxied here by incomes per capita, would trade proportionately more with one-another, relative to those with more similar income levels. Conversely the basic predictions of the gravity model suggest the opposite is true, i.e. that nations with more similar per capita incomes will trade relatively more with those with more dissimilar per capita incomes. Therefore a negative coefficient on this variable would evidence the Hecksher-Ohlin theories, while a positive coefficient would provide support for the basic predictions of the gravity model.

One would typically assume, ceteris paribus, that neighbouring nations would trade more with one another. This can be investigated with the inclusion of a dummy variable indicating whether or not countries share a common border, which captures the effects of cultural proximity and ease of trading. However, one would expect the inclusion of a variable for distance should include these border effects. Despite this intuition a number of authors choose to include a border dummy variable, which is typically of considerable significance (Head and Ries 1998). This fact may reflect the inaccurate calculations of the great circle formula[8], biasing upwards true trading distance when calculating the true distance between two points. A dummy variable is also commonly used to account for the state of landlockedness of a country.

[8] Readers should refer to data section

Language effects, and linguistic links, what Linnemman (1966) refers to as psychic links, are also commonly cited as significant in gravity models of bilateral trade, and immigrant-link studies alike. A higher proportion of residents in one nation that speak the language of one of its trading partners, ceteris parabus, the higher the volume of trade we would expect to observe between them. This is what previous research suggests being the case (Rauch and Trindade 2002), though Wagner et al's results evidence the contrary. A language variable could be deemed appropriate for estimation therefore.

The key variable of interest in any immigrant-link study is the stock of immigrants. In this paper this refers to the stock of immigrants from each of the EU-expansion countries residing within each EU-15 country. Of primary interest are the coefficients on this variable, for both import and export estimations; a comparison of which will highlight the prevalence of the relevant hypotheses. Most of the immigrant-link studies assume constant marginal returns from immigration, though this may not be a realistic description of reality. Gould first addressed this issue, by incorporating a function for decreasing returns to migrants. This reasoning, though intuitive, was not derived from theory. Wagner et al subsequently provided a theoretical justification for this, by implementing a random encounter specification. Whilst in the U.S. and Canada the proportion of immigrants to the native population is relatively high, the proportion of immigrants in the EU relative to the 370 million residents is small. Therefore it is assumed for the purposes of this paper, that the returns to immigrants are relatively constant.

The list of variables in the preceding discussion is by no means exhaustive. Many examples of different applications of the gravity model (as applied in international economics) were provided in the literature review. To list all the additional variables included in other authors regressions would be a hapless task. It is not the purpose of this paper to investigate the effects of regional trade agreements, a common currency, or institutional qualities; nor any other extraneous factors. Moreover, due to the heterogeneity between the theoretical justifications drawn upon in the immigrant-link literature there exists an extremely broad range of optional variables that could be estimated. These include price indices (Head and Ries 1998), trade barriers (Gould 1994) and terms of trade (Dunlevy and Hutchinson 1999). The theoretical stance adopted in this paper is somewhat ad-hoc and simplistic in approach. As ''full' a gravity model as the data permits will be estimated, in conjunction with an additional variable for the stock of migrants from each accession country in each of the EU-15 nations. As these gravity variables have consistently been proven to exert a significant influence on bilateral trade, it is the belief of the author that by omitting additional variables, the paper will remain focused on the key variable of interest yielding higher quality results.

Both the flows of imports and exports, to and from the EU-15 nations, are to be estimated. Equation (3) augmented, though suitable for these estimations if we were to treat the EU as a single entity, is not suitable for the estimation of bilateral trade between the Western and Eastern European countries over time. It is assumed that the reduced form below is appropriate for estimation in light of the requirement of an additional dimension when estimating is:

$$Y_{ijt} = f\left(X_{ijt} M_{jit}\right) \quad (9)$$

Where:

Y_{ijt} = Flow of goods (imports or exports) from country i to country j at time t

X_{ijt} = Vector of variables which exert influence on bilateral flows between country i and country j at time t

M_{jit} = Stock of migrants from country j in country i at time t

4. Empirical model

In light of the preceding theoretical discussion, this section aims to specify the functional form estimated, additionally offering explanation as to what variables were deemed suitable for estimation. The focus of this paper is to quantify the effects of EU-expansion countries' immigration on the EU-15 countries bilateral trade flows. This is accomplished by testing for the 'core' gravity variables, in conjunction with a variable for the stock of immigrants (additionally introducing dummy variables), therefore abstracting from other additional variables. In doing so this paper closely follows the intuition of Wagner et al (2002), though it adopts an alternative (and simplified) functional form and estimation technique. Specifically it replaces a two-stage Heckman procedure, with a standard pooled OLS and fixed effects estimation. In the absence of excessive numbers of zeros, the Heckman procedure was judged unnecessary. Indeed Wagner et al did not correctly utilise the technique; failing to use additional variables when calculating the Mills ratio in their first stage regression, thus casting doubt on their results. Despite these remarks there remains one significant advantage with their technique.

These authors, when studying Canada, exploited the significant variation in migrant stocks between observations of Canadian provinces, to quantify the effects of immigration from other countries on Canadian provincial trade. Notably, they used country level fixed effects to account for the unobservable effects which influence both trade and immigration, between Canada and its individual trading partners. As noted, cross-sectional studies may bias estimates upwards due to unobserved effects whilst fixed effects estimation may imply the opposite bias due to a loss in informational content (Griliches 1986). Using observations and fixed effects at different levels, Wagner et al believe, gives the model the advantages of fixed effects specification i.e. the minimising of the influence of unobservable characteristics (which are assumed to occur at the national level), without needing to rely solely on time series variation. A, 'best of both worlds', scenario. When applying this intuition in the context of the EU it is necessary to make an additional assumption therefore. In Canada though each province is governed by its own rules there exists a set of national or federal laws that apply to each state in addition. Similarly, in the EU, each individual country has its own laws but each is subject to EU law that is formulated in Brussels. Crucially, it is assumed that that these two situations are comparable. For estimation purposes the EU-15 countries are treated as equivalent to Canadian provinces. Observations for EU-15 trade volumes and numbers of EU-expansion migrants are used in conjunction with fixed effects for EU-expansion countries, to benefit from the insight offered by Wagner et al. These provincial observations allow us to investigate whether greater stocks of immigrants residing in a particular 'province' of the EU facilitate greater trade with those provinces. This paper represents the first time, to the authors knowledge, an application of an immigrant-link model to two sets of countries simultaneously. Even though in effects we are treating the EU as single nation state consisting of 15 provinces.

Initially pooled OLS is used to investigate the underlying patterns in the data. Pooled OLS is a simple econometric specification for investigating the independent affects on the dependent variable of any particular explanatory variable; holding constant the influence from other explanatory factors. This procedure is valid even if explanatory variables are correlated with some of dependent variables. This specification therefore allows the affects of immigration on EU-15 trade levels to be identified independently from the other explanatory variables; and thus is deemed suitable for the purposes of this paper. The additional use of fixed effects minimises the influence of unobserved effects by removing time-constant and unobserved effects, prior to estimation. Here this refers to the removal of unobserved effects from the EU-expansion countries. The correct use of fixed effects, over a random effects specification, was confirmed by the use of a Hausman procedure. The estimation of fixed effects also facilitates the separate investigation of 'within' and 'between' country variation.

It is particularly important to control for the size of countries in the context of this project since there will likely be large disparities between the Western and Eastern European states. Wagner et al, and Head and Ries, calculate their measure of economic mass as the product of the trading pair's national income, as a proportion of world income. This gives additional weight therefore, to those trading links that represent a higher proportion of world trade. Instead the simpler product of the trading partners GDPs was used as a measure of total bilateral economic mass (Baltagi 2003) in conjunction with a variable for the product of per capita incomes (Frankel 1997). The product of trading pair's national income has consistently been demonstrated to be of key significance in these models. The estimation of per capita income, as opposed to a measure of population was chosen as this additionally indicates the prevalence of one underlying theory over another. This variable may also be correlated with the level of trade protection of a country. This is particularly important since the main variable of interest omitted from our analysis is a measure of trade barriers. A priori one would expect ceteris parabus for the coefficient on the product of trading nations GDPs to be positive, larger countries are anticipated to trade relatively more with one another. Head (2000) notes this is typically found to be close to 1, though comments that the coefficients on per capita incomes can vary significantly. A priori, we can only speculate on the likely sign on this variables' coefficient. In light of the preceding theoretical discussion this is hypothesised to be negative for the H-O theories are generally believed to provide a realistic theoretical underpinning for North-South trade (Frankel 1997).

In terms of the distance variable, it is hypothesised that the further each EU-15 country is from the EU-expansion countries the less the pair will trade due to higher cultural, legal, transportation and communication barriers. The remoteness variable as prescribed by Head (2000) – as discussed in the theory section, though theoretically correct, is unsuitable for the purposes of this project, due to the limited number of countries in the data. Calculating the 'distance from world markets' on this basis is

nonsensical. Instead what is applicable, is how remote each EU-expansion country is from the main Western European markets. Ceteris Parabus one would expect the more remote an EU-expansion country is from an EU incumbent nation, the more the incumbent nation will trade with alternative 'third country' options (Wagner et al 2002). Formally:

$$R_f = \frac{1}{\sum_p \left(Y_p / D_{fp} \right)}$$

Where:

R_f = Remoteness of country f from European Markets.

Y_p = National Income of European 'province' p.

D_{fp} = Distance from country f from province p.

Formulating the measure of remoteness in this manner also circumvents the need to calculate a countries' distance from itself[9]. However, this measure is time invariant and will therefore drop out from estimation under fixed effects specifications.

In keeping with the immigrant-link studies to date, the natural log of the stock of immigrants from each EU-expansion country residing in each EU-15 country was added to the estimated equation. In light of the literature on immigrant-links, it is hypothesised that greater numbers of immigrants facilitates greater degrees of trade creation, in terms of both imports and exports, though it is difficult a priori to predict which trade elasticity with respect to immigration will be larger.

There are many opportunities for the inclusion of dummy variables, as outlined in the theory section. Many were considered unsuitable for the purposes of this project however. A simple border dummy variable, taking the values of 1 if the trading pair borders each other and 0 otherwise, could easily have been constructed. So few EU-15 'provinces' share a common border with EU-expansion countries however that this dummy variable will unlikely capture any cultural or proximity effects. Instead it is hypothesised that this will be captured be the distance variable and the error term. Similarly, a dummy variable could be constructed for the landlocked status of a country; but only Luxembourg and Austria are landlocked, and these will only represent a small proportion of total East-West European trade. While Wagner et al (2002) and Rauch and Trindade (2002) calculate their language commonality variable on the basis of the probability that any two individuals chosen at random from any trading pair will be able to speak the same language; this proved impossible due to data constraints. Indeed Europe has numerous languages and dialects (22510). To conceive of a dummy variable for each language is meaningless. Instead a dummy variable which took a value of 1 if the there existed any official common language between two trading nations and 0 would be more suitable. There are few such pairings however. Bearing in mind the limited number of overall pairings of countries involved in this study (225), it is especially important not to 'ask of the data too many questions' Frankel (1997); as the results obtained will become increasingly spurious. Technically this involves the excessive loss of degrees of freedom. For these reasons it was decided to omit dummy variables for language, landlockedness and borders, and instead focus upon dummy variables of more likely and obvious significance.

A set of year dummies is used in order to correct for the possibility that results may capture the simultaneous growth in immigrant populations and trade (Wagner at al 2002), i.e. to reflect the fact that variables may have different distributions in successive years. Technically this allows for the intercept to vary over time to account for this. Regional dummy variables representing successive waves of EU expansion were also included to capture the joint effects of EU-expansion countries acceding to the EU. Lastly a set of country of origin dummies, a la Gould, were included to estimate the effects of immigration independently of variations in time. Technically, not all dummy variables can be estimated simultaneously due to a dummy variable trap (Gujarati 1995), a situation of perfect multicollinearity. Before estimation the regional dummy variable for Macedonia was dropped along with the year dummy for 1994 and the country of origin dummy for Macedonia. It is believed that the accession of Macedonia will have less of an impact than other waves, and as is usual the first year dummy variable is treated as the base year for comparison. Table 2 provides a summary as to which countries constitute each wave.

[9] A dubious practice necessary when i = f, under previous formulations for example: $R_f = 1/\sum_i [(Y_i / Y_w) / D_{if}]$ from Wagner et al (2002). In this circumstance Nitsch (2000) suggests this internal distance $D_{ff} = (AREA_f / \Pi)^{1/2}$.

[10] European Commission's own estimate.

Table 2. Waves of EU expansion

Wave number	Countries involved
1	Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovenia and Slovakia
2	Romania, Bulgaria
3	Croatia
4	Macedonia
5	Turkey

The functional form estimated in this paper, similar to that of Wagner et al (2002) builds upon equation (6). It is additionally augmented for the inclusion of an alternative formulation of the remoteness variable, and an extra dimension to account for variations over time. The specific functional form estimated was:

Without country fixed effects:

$$\ln T_{pft} = \beta_1 \ln(MIG_{fpt}) + \beta_2 \ln(Y_{pt} * Y_{ft}) + \beta_3 \ln(D_{pf}) + \beta_4 \ln(R_f) + \beta_5 \ln((Y_{pt}/POP_{pt}) * (Y_{ft}/POP_{ft})) + \beta\lambda_{pft} + \varepsilon_{pft}$$

And with country fixed effects:

$$\ln T_{pft} = \beta_1 \ln(MIG_{fpt}) + \beta_2 \ln(Y_{pt} * Y_{ft}) + \beta_3 \ln(D_{pf}) + \beta_4 \ln(R_f) + \beta_5 \ln((Y_{pt}/POP_{pt}) * (Y_{ft}/POP_{ft})) + \beta\lambda_{pft} + FE_f + \varepsilon_{pft}$$

Where:

$\ln T_{pft}$ = either volume of exports from EU province p to country f or for estimation for imports the volume of imports from EU-expansion country f to EU province p.

$\ln(MIGf_{pt})$ = the natural log of the number of immigrants from EU-expansion country f, in each EU province (country) p.

$\ln(Y_{pt}*Y_{ft})$ = natural log of the product of trading pairs GDP

$\ln(D_{pf})$ = natural log Distance between country p and province f

$\ln(R_f)$ = natural log of measure of remoteness from Western European markets of country f

$\ln((Y_{pt}/POP_{pt})*(Y_{ft}/POP_{ft})$ = natural log of the product of trading pairs per capita incomes.

β = row vector of coefficients for λ_{pft}

λ_{pft}= vector of dummy variables for years, regions, country of origin and the constant

FE_f =fixed effects for country f

ε_{pft}= error term

5. Data

This section aims to address all the relevant data considerations for the estimated panel; to specify the sources, the units of measurement, and the constraints that the data imposed. The estimated panel, spanning 8 years, comprised of 1800 observations, which summarized trade (and the other core variables) between the 225 EU-15 and 15 EU-expansion country pairings[11]. The final panel covers the period 1994-2001, and contained complete data for all of the EU-15 countries and 15 EU-expansion countries, (with the exception of very few missing values). A panel spanning a longer period of time would have been favoured - observations every five years, for example would likely demonstrate greater variations between the observations. However, the utilisation of a short-fat panel circumvents some of the econometric problems associated with longer time-series studies. It is the belief of the author that the data used should represent a fairly typical period in EU history, fairly representative of the EU as a whole. Indeed if these are 'normal' years then any estimates should provide reasonable predictions as to future European immigrant-link effects.

The key variable of interest in the model, as in any immigrant-link study, is the stock of immigrants. Here this refers to the number of immigrants from each of the EU-expansion countries residing in each EU-15 country. The accuracy of this data is of critical importance. No data was able to be collected for the total number of migrants from EU-15 countries residing in expansion countries however. The data is therefore suitable for studying the bilateral effects of east-west EU migration, but will fall short of being able to explain the complementary migratory flows in the opposite direction. The technique commonly used when collecting data on the stock of immigrants is to collect data from the censuses of individual countries and to supplement this data with information on the flows of immigrants from various countries – data more commonly collated and easier to access. Typically, a migrant stock rule is implemented to calculate the annual stocks of migrants. A preliminary regression is then run to calculate the rate of attrition due to death and departures, which is commonly assumed constant. An example of this technique is provided by Head and Ries (1998) which is reproduced here for convenience:

(9) Migrant stock rule:

$$S_t = (1-\delta)S_{t-1} + F_{t-1}$$

[11] Summary statistics can be found in Appendix A.

(10) Annual stock formula:

$$S_t = \frac{S_{baseyear}}{(1-\delta)^{baseyear-t}} - \sum_{i=1}^{baseyear-t} \frac{F_{t-1+i}}{(1-\delta)^i}$$

(11) Regression to estimate rate of attrition:

$$\hat{S}_t = (1-\delta)^5 \hat{S}_{t-5} + (1-\delta)^4 F_{t-5} + (1-\delta)^3 F_{t-4} + (1-\delta)^2 F_{t-3} + (1-\delta) F_{t-2} + F_{t-1} + u_t$$

Where:

S_t = Immigrant Stock at time t

F_t = Flow of migrant at time t

U_t = error term

ς = Rate of attrition

This technique proved prohibitively time consuming within the constraints of the project however. It was thus necessary to obtain a 'ready-made' and complete migrants stock data series. There is a definite lack of availability of such series. The only institution that collates reliable data on European migration is Eurostat, who recently discontinued their main series due to the fact that they could no longer guarantee its accuracy. The only available way to model this variable was to use a proxy. The two series available most suited for use as a proxy; were the stock of immigrants in the EU-15 by 'workers by citizenship', and by 'acquisition of citizenship'. The former definition no doubt included a greater percentage of the overall stock of immigrants; as this excludes only those immigrants residing in the EU-15 that are not working, as opposed to excluding all of those immigrants other than those which have taken up a foreign nationality. This proxy could not be considered however, as the data contained far too many zeros or missing values, with no means of being able to distinguish the two. This would have prevented the application of meaningful econometric techniques. The rather inferior 'acquisition of citizenship' series therefore had to be estimated. It is assumed that over time the proportion of those migrants granted citizenship is approximately fixed. This variable should be closely correlated to the overall stock of migrants but the estimated coefficients may be significantly less than we would otherwise expect to observe. We return to this issue in the results section when discussion is offered as to the robustness of this proxy.

The final problem involving the stock of immigrant's variable was that even with the inferior measure of immigrants, some of the values needed to be imputed. Fortunately, there were not many instances where this was necessary, with one notable exception, Ireland. By and large the imputed values were calculated as the mean for years either side of the missing value. Where missing values occurred at the ends of the dataset linear trends were assumed. Only for Ireland was the data too poor to impute the necessary values. In this instance the average of the other EU-14 countries was taken. It is the belief of the author that treating Ireland as equivalent to the average of the other EU countries should not bias results too heavily, even though Ireland has a slightly higher than EU-15 average per capita income.

Distance (calculated in kilometres), commonly taken in models of international trade as the geographical proximity of capital cities between trading nations, is typically measured using the Great Circle formula, which is formally provided by Head (2000) as:

(12)
$$D_{ij} = 3962.6 \arccos([\sin(Y_i)\sin(Y_j)] + [\cos(Y_i)\cos(Y_j)\cos(X_i - X_j)])$$

Notes:

Where: X is longitude in Radians when multiplied by 57.3 to convert, and Y is latitude in radians when multiplied by -57.3.

Where: D_{ij} = Distance between capital cities of country i and country j.

This in itself can be quite misleading, rarely are capital cities situated in the centre of a country. Indeed there is no guarantee that any single capital city will be the most significant city in terms of trade. Boisso and Ferrantino (1997) however, find that gravity models calculated on the basis of using those cities of greatest trading significance changes the overall results of gravity models very little. The formula calculates the shortest distance between two points whilst (implicitly) assuming the earth to be perfectly spherical, when in fact it is elliptical. Even if all goods were transported by air this calculation would yield slightly inaccurate results as the formula ignores the fact that the majority of flying bypass the North Pole (Head 2000). Nor does the formula take into consideration impassable geographical features. If anything the formula will underestimate the true distances involved. Using the Great Circle distances does however guarantee consistency, and has been widely used in international trade literature. It is the most reliable measure of geographical proximity available within the constraints of the project.

All income data was collected from the World Development Indicators 2004. It makes little difference whether one implements GDP or GNP (Linnemann 1966), though different definitions of GDP have been reported to influence results (Gros and Consiarz, 1996, Cornett and Iversen 1998). Arguments have arisen between those that argue that GDP (MER – market exchange rates) to be the correct measure, and those that cite GDP (PPP – purchasing power parity) as more suitable. However in view of the results of Paas (2001, 2002b) who finds in relation to both Baltic and European countries, GDP (PPP) to be a superior measure (in terms of a superior R^2 measure), here GDP data in real terms, in equivalent purchasing power parity prices, is used.

All trade data was collected from the IMF's Direction of Trade Statistics; all imports are valued for c.i.f. (cost, insurance, freight) with all exports measured f.o.b. (free on board). Both import and export data were recorded in the same units as the GDP data to ensure consistency. The only problem presented by the trade data was for Belgium and Luxembourg for the first half of the 1990s when these countries reported their statistics jointly. In order to impute these values the average of the relative contributions from both Belgium and Luxembourg was taken over the period 1995-2001. This average was then used to calculate the share of the two countries import and export volumes for 1994 and 1995. As Luxembourg's contribution to the total was always relatively small it is predicted that this calculation will have a negligible effect on the results.

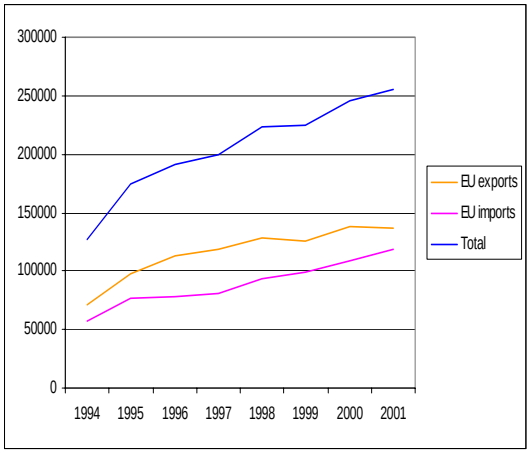
The constraints of the data did prohibit some interesting avenues of inquiry. No data was available on different commodity groups, or remittances. Neither was information available for the skill level of migrants, the average length of migrants' residence abroad, on different types of migrant, nor any detailed percentages on the languages spoken in each country within Europe. A lack of data for languages, prevented a detailed common language variable from being constructed. As all of the data was obtained from reliable sources it is hoped that good results will be obtained.

6. Results

This section aims to highlight the main findings of this paper, and to evaluate them in conjunction with previous studies. In total, four log-linear equations were estimated; one each for imports and exports, for pooled OLS; both with and without fixed effects. The main advantage of using a log-linear specification is that one may effectively read off the coefficients as elasticities. In light of the ad-hoc way the model was formulated, and the inconsistencies and lack of standardization found in the immigrant-link literature, the results obtained are evaluated with respect to both the immigrant-link and base-line gravity model literature. In particular, where applicable, comparisons will be drawn between Paas (2002a), and Africano-Silva et al (2001), who both implemented a gravity model, when examining a similar sample of countries to this paper, whilst omitting immigrant-link effects. And Wagner et al (2002) whose estimated equations bear the closest resemblance to the functional from estimated here. Each of the variables estimated will be discussed in turn, before attentions are turned toward the key variable of interest; the stock of immigrant's. Table 3 in Annex 2 summarizes the results (including robust standard errors) obtained from estimation.

Overall the model fits the data well, and a sufficient number of observations were estimated. Due to the infancy of some of countries under investigation a number of entries were left blank in the dataset however. This is the best way for most econometric packages to reconcile these data entries. All the estimated coefficients are statistically significant, of reasonable magnitudes, and of the predicted sign. This is not surprising given the pedigree of gravity models. The overall coefficient of determination ranges from 0.768 to 0.8769. The explanatory variables account for between nearly 77% and 88% of the variation in EU-15 bilateral import and export trade volumes. Though this is significantly lower than in Gould's original study where 99% of the variations in imports and exports were explained, it represents about the average among the immigrant-link studies, despite estimating significantly fewer variables. The 'within' and 'between' coefficient of variation results from the fixed effects specification imply, somewhat unexpectedly, that both variations between countries, and over time are important for determining EU-15 bilateral trade flows. A priori, one would expect that the 'within' estimate would be somewhat smaller, as indeed Girma and Yu (2002) do, indicating that most of the variation arises from the 'between' country differences. Here intertemporal variations are also important. This could be due to the growing importance of the EU-expansion states throughout the previous decade. At the beginning of this period these countries traded far less with the EU-15. Over the decade intra-European trade rapidly expanded (see Chart 1 below) and the EU-expansion countries rapidly developed.

Chart 1. EU total trade with EU-expansion countries, 1994-2001, (US$ mn)



Source: IMF Direction of Trade Statistics

The coefficients on all variables are significant to 99% confidence, excepting the product of per capita incomes in the pooled export equation, which is accurate with only 95% confidence. This is in contrast to some of the immigrant-link studies where a number of gravity variables are reported as insignificant. Income per capita for example is found to be insignificant in both Dunlevy and Hutchinson's studies, and national income is insignificant in some of Wagner et al's (2002) estimations. In a typical bilateral gravity model one would expect all of these variables to be very significant however, as these variables are effectively the fundamental tests of the gravity principles. As the measure of remoteness used relates to a specific EU-expansion country, this variable is dropped under the fixed effects specification, as it is effectively just a multiple of an EU-expansion country's dummy variable. The standard errors are generally smaller under the fixed effects specification, and taken together, the results indicate a high degree of homogeneity between immigrants from different EU-expansion countries; as the coefficients from the pooled OLS and fixed effects specifications are almost identical.

It is important before commenting on the results to note that although:

> "*gravity models have strong power in explaining trade pattern and testing hypotheses…..the modeling results are not very reliable if we want to estimate the level of trade flows in absolute terms*" (Paas 2002a: 13)

If true, this casts doubt on the magnitude of the estimates obtained; and then it would seem likely the most reliable conclusions, which can be drawn from the study, are in relation to the signs on the estimated coefficients, and the relative size of the trade elasticities with respect to immigration. This provides another justification for evaluating the results from this paper in relation to both strands of the relevant literature; i.e. to ensure as fair an evaluation as is feasible.

Trading distances have a significant and strong affect on both the exports and the imports of EU-15 nations. A 10% increase in the distance between trading pairs is estimated to decrease imports and exports by approximately 14%. This initially seems reasonable as distance; a proxy for transportation costs is one of the most significant barriers to trade. The size of the coefficient could be cause for concern however. Over time, the average cost of transporting goods should decrease, as better trade links are created. A coefficient on the log of distance greater than one signifies that the reverse is true. It is not unusual to find such a large influence exerted by distance on bilateral trade flows among the immigrant-link papers however. Wagner et al estimate a distance coefficient of between -1.05 and -1.85, Head and Ries estimate this to lie between -1.027 and -1.474. This is one of the clearest disparities between the bilateral gravity literature and the immigrant-link literature. In general the effects of the distance variable are significantly larger in immigrant-link studies. Paas (2002a) however, estimates the distance coefficient to lie between -0.293 and -1.368. Similarly, Africano-Silva et al (2001) find their distance coefficient to range from -0.979 and -1.124, for the same sample of countries as are estimated in this paper. Those gravity model studies investigating EU expansion therefore find similar results to those obtained here, providing support for this paper's results. Indeed, in light of the use of Great Circle formula likely underestimating the true distances involved this effect will likely be even greater. These results together imply that transportation costs are extremely important with regard to West-East European trade. In light of Paas's comments it could simply be the case that the magnitudes of the estimates are inaccurate. The results do confirm Leamer and Levinsohn (1994) assertion that the influence of distance on bilateral trade flows to be one of the 'clearest and robust findings in economics', however.

Typically the measure for remoteness, as implemented in either traditional gravity models, or immigrant-link studies, represents a countries' distance or remoteness from world markets. One would expect, under this definition that the further a country is from world markets, the more one of its trading partners will prefer to trade with closer alternatives. Here an alternative remoteness variable was used however, assigned not to the EU-15 trading 'provinces' but rather to the EU-expansion countries. This measure therefore represents how remote each EU-

expansion country is from European markets. This denies a direct comparison with previous studies. The further each of these countries is from European markets the less one would expect each of the EU-15 countries to trade with these nations. This is what the model suggests. Indeed the effects are particularly strong; a 10% increase in the remoteness of an EU-expansion country will decrease imports to the EU-15 countries by 5.4% and exports from the EU-15 by nearly 17%. This does not seem unreasonable as the majority of EU-15 countries trade between themselves.

The coefficient on the product of the national incomes of trading pairs is both highly significant and large in magnitude. Trade in exports increases almost exactly in proportion with bilateral economic mass, though trade increases more than proportionately in imports. A positive coefficient was anticipated as wealthier countries will likely have developed superior infrastructures relative to poorer countries, facilitating better-quality trade routes. This finding (for imports) contradicts Frankel (1997), who notes that this coefficient is often less than one. This he believes evidences the fact that smaller economies are generally more open to trade, as they are more dependent on it. Poignant examples include Hong Kong and Singapore. Here the opposite holds true. The larger EU-15 economies tend to be more open to (importing) trade from the EU-expansion states, than their smaller European counterparts. This could simply indicate that these states are better exploiting business opportunities available there. Paas (2002a) omits a measure of total national income instead using a measure for population, though Africano-Silva (2001) estimates values of between 0.779 and 0.811. Comparison is difficult with Wagner et al, and Head and Ries who use an alternative measure which accounts for total bilateral economic size as a proportion of world income. These two studies do however both report estimates for national income to both be greater than one, despite this additional weighting. Therefore, although larger than one may expect, this result does not jeopardise the overall conclusions from the study.

The results from the product of trading pair's per capita incomes are particularly interesting. As outlined in the theory section the sign on this coefficient is indicative of the prevalence of the underlying theories underpinning the gravity model as applied in international trade. Interestingly, the coefficient for this variable is negative for the estimation of the import equations, but positive otherwise. This indicates that the Hecksher-Ohlin theories hold in relation to imports though not for exports; when the predictions of the basic gravity model are of greater significance. EU-15 countries export relatively more to those EU-expansion countries with the most dissimilar per capita outputs (incomes). This may be evidence of EU-15 countries trying to exploit those export trading relationships where the wealth disparities (and thus potential profits) between the trading nations are greatest. Conversely, the EU-15 countries import relatively more from those EU-expansion countries with more similar levels of per capita output (income). This may evidence the fact that generally nations prefer to trade with countries with similar tastes (as determined by incomes), so that they can enjoy greater variety in consumption.

Dunlevy and Hutchinson (1999, 2001) include relative income to facilitate the testing of the Linder (1961) Hypothesis. This states that richer countries as measured by per capita incomes have similar preferences and products though the latter will be differentiated. These countries will therefore trade relatively more with one another. The finding that EU-15 countries import from those EU-expansion countries with most similar per capita income levels would provide some support for the Linder Hypothesis. The Linder hypothesis has previously been tested via the inclusion of the absolute difference in per capita incomes as an explanatory variable however, as demonstrated early by Thursby and Thursby (1987). Therefore the product of per capita GDPs as estimated here, though not the exact theoretical measure for testing the Linder Hypothesis, does provide some evidence for this notion.

Many immigrant-link studies omit any per capita income variable, and of those few that do utilise one, a measure for a single country's per capita income is commonly reported; as opposed to the product of the trading pair's per capita incomes. The exception to this is Rauch and Trindade (2002), who report a coefficient of between 0.177 and 0.284 from their 1990 conservative estimate – that estimate deemed to be best for comparison. Thus the estimated coefficient from the export equations seems to fit with the previous immigrant-link literature well, though on the imports side, the result confounds previous findings. Paas (2002a) estimates both sets of trading countries per capita incomes' separately, and further does not disaggregate imports and exports. Due to the additive nature of logarithm's we may simply sum the trade elasticities with respect to both per capita incomes from Paas's study, and compare them with the sum of the trade elasticities from this paper. In 2000, for example, Paas's (2002a) results suggest a combined per capita income coefficient of -0.278. The sum of per capita coefficients from both imports and exports from this paper equals -

0.373, not too dissimilar. These negative per capita income coefficients, Paas believes, is due to the expanding economic relations between nations at very different levels of development. If Paas had further disaggregated the trade data, it could be that the import elasticities with respect to per capita income were significantly negative, with the corresponding export elasticities smaller, though positive. This would yield consistent results to those obtained in this paper. The very fact that such a similar study though omitting immigrant-link effects, yields similar results adds credence to those estimates obtained in this paper.

Of principle interest is the coefficient on the immigrant stock variable. The estimates on this coefficient demonstrate that immigration into the EU-15 exerts a positive influence on both export and import bilateral trade flows. The magnitude of the import elasticity with respect to trade is larger for imports than for exports however. This indicates that immigrants from EU-expansion countries residing in EU-15 countries exert a greater influence on EU-15 imports than EU-15 exports. Ceteris parabus a 10% increase in the stock of immigrants from EU-expansion countries residing in the EU-15, will increase the imports of the EU-15 by 1.4%, and their exports by 1.2%. If it is believed that immigrant-preference effects exert greatest influence on import elasticities, while immigrant-information effects influence export elasticities most significantly, then these results would indicate the dominance of preference effects stimulating trade. This key finding is in line with Head and Ries (1998), Dunlevy and Hutchinson (1999, 2001), and Wagner et al (2002). As taste effects do not influence exports, the results are most likely indicative of the presence of (a combination of) both the immigrant-preference and the immigrant-information hypotheses. Further distinguishing between the two would require additional data on differentiated products and immigrant heterogeneity however.

If indeed preference effects dominate information effects this could be because the majority of Eastern European migrants largely fill vacant low skilled low wage positions, such as working as agricultural labourers; as opposed to finding 'niches' in western European markets, which they are able to exploit. Eastern European migrants are significantly less wealthy than their Western counterparts; it may be that they do not have enough initial capital to take advantage of business opportunities even if the opportunities present themselves. The start-up costs for such ventures are likely prohibitively expensive. If this line of reasoning holds true, those Eastern European migrants that do manage to start businesses will likely be those that have migrated long ago and have managed to save the necessary start-up funds; or those that became very wealthy prior to emigrating. An additional explanation could be that such business links have already been largely exploited by western Europeans, leaving few immigrant-link type business opportunities.

As previously noted, one must take care when evaluating the magnitudes of estimates obtained from gravity models. However, as all the immigrant-link studies to date all implement gravity models, a comparison between these limited studies may still prove insightful. The import elasticities with respect to immigration, from those immigrant-link studies to date, range from 0.02 to 0.47. The export elasticities with respect to immigration range from 0.01 to 0.47. These results are summarized in Table 4 (see Annex 3) – a duplicate of Table 1 with the results from this paper additionally reported.

The range in these results was attributed, in the literature review, to the different samples, time periods and empirical estimations implemented by the various authors. These estimates vary in sensible directions depending on the underlying specifications and samples. The 'mid-range' immigrant stock coefficient estimate, obtained in this paper, from the implementation of both fixed effects and a longitudinal panel, seem quite reasonable in light of the previous literature; especially relative to those two European immigrant-link studies to date - arguably those most similar in terms of geographical coverage and time period, that both estimate very similar trade elasticities. However, the estimates from both these papers predict a stronger pro-export effect, the diametrically opposed conclusion. This could simply be because for the UK and France migrants exert greatest influence on exports, which are outweighed by more significant import effects from the other EU-15 countries. Alternatively it could be due to the proxy used for the stock of migrants.

Sensitivity Analysis

A number of diagnostic tests were used to investigate the feasibility of the estimated model, the results of which are reported in Appendix B. This section aims to report these findings before discussing the feasibility of the acquisition of citizenship variable as a proxy for the total stock of immigrants, and of the functional form estimated.

For each of the regressions F-tests, testing the likelihood that all the coefficients are simultaneously zero were calculated. In every instance these were passed with high levels of

confidence. Although when the fitted values from the regressions were plotted against the residuals no definitive patterns indicative of heteroskedasticity were observed, diagnostic tests confirmed its' presence. To correct for this failure which has the effect of biasing the estimated standard errors, the model was re-estimated using robust standard errors. Both import and export equations were also tested for multicollinearity. No multicollinearity was discovered however, as was expected, as the R^2 is high and the coefficients are all highly significant.

Of far greater concern were the results from Ramsey's reset test, which showed with no degree of uncertainty the use an incorrect functional form, or alternatively an omitted variable bias. This is potentially a very serous problem, as this weakness can influence both the estimates of the coefficients and the standard errors, potentially casting doubt on all of the results obtained. Unfortunately all of the immigrant-link studies to date fail to report the results from their diagnostic test which impedes comparison. Firstly, no data was collected on trade barriers i.e. tariffs and quotas. These contribute to the friction that impedes the smooth flow of trade and constitute one of the key variables Linnemann (1966) suggested should be included in any gravity model of trade. Simply no data was available, and there was no other option left but to estimate the model omitting this potentially crucial aspect of trade. The only immigrant-link study to date to include such a measure was Gould. Indeed among this class of studies this was the only one that obtained an extremely high R^2. This might be indicative of the fact that this variable should be included, i.e. as then an extra 10-15% higher R^2 is obtained. Only the per capita income variable may capture some of this effect. Secondly it must be considered how severe this bias is likely to be. All of the estimated coefficients are highly significant, the coefficient of determination is fairly high and most, if not all, of the magnitudes of the estimates seem feasible. It may be that the bias is not that great, but there is no way of knowing without additional data, which was unavailable. Thirdly, one may expect a priori, that the Ramsey test result was to have been expected, as the formulation of the model was ad-hoc. In light of the results obtained, when compared to previous estimates, the omission of trade barriers and the ad-hoc formulation of the model; it is the belief of the author that the bias associated with this test is not that serious.

The second potentially serious problem with the estimation is with the use of the stock of migrants by acquisition of citizenship as a proxy for the total stock of migrants. There will likely be a significant disparity between these figures. The estimates obtained seem very reasonable with respect to the previous immigrant-link studies though. One explanation could be that it is specifically those migrants that have acquired citizenship that primarily form immigrant-links and exert an influence on native exports. Alternatively, it could be that the magnitude on this variable is significantly lower than would be the case if the total stock of migrants was estimated. This could potentially raise the estimates of both the import and export trade elasticities with respect to immigration quite significantly, if the proxy represents only a small fraction of the overall stock of migrants. Indeed it could be the case that those migrants that have obtained citizenship will be those that form specific types of immigrant-links. For example, it is more likely that those migrants that have resided longest abroad, and have acquired citizenship, will be those that are more likely to form business links to their native countries. If this is true then the estimation including the total stock of migrants would yield much higher import elasticities. If the converse were true, then one would expect to observe relatively higher export elasticities, paralleling those European immigrant-link studies to date. Ideally to investigate this latter problem one would collect detailed data either from a country with very reliable immigration data, or alternatively procure the total stock of migrants together with the stock of migrants by acquisition of citizenship for the same to see what fraction of the total this constituted. This data was unavailable, and thus it proved difficult to investigate the feasibility of this variable.

Several authors of immigrant-link studies additionally estimate a dynamic version of their model by including the lagged dependent variable as an explanatory variable (Gould (1994), Head and Ries (1998), Girma and Yu (2002) and Wagner et al (2002)). The addition of this variable will likely capture any momentum in trading (Girma and Yu 2002). Eichengreen and Irwin (1996) argue that this variable should always be included in any gravity model of international trade. Though fixed effects were implemented to capture some of the effects of omitted variables it seems appropriate to estimate the model making an inclusion for this variable (the results of which are reported in Appendix C). Surprisingly, in contrast to the immigrant-link studies that include the lagged dependent variable, its inclusion here makes very little difference to the results overall. There is very little auto regression. This could of course be symptomatic of too short a data series, but further interpretation is difficult without additional data.

7. Conclusion

Motivated predominantly by Gould's (1994) original immigrant-link study, and Wagner et al (2002), this paper aimed to quantify the effects of East-West European immigration on EU-15 bilateral trade flows. In keeping with the literature, and due to its proven pedigree in international economics, a gravity model was implemented to analyze these effects. This represents, to the author's knowledge, the first time this methodology had been applied to either to the entire EU or to two sets of countries simultaneously before.

In light of the range of theoretical justifications for the gravity model, and the multitude of variables estimated, among both the immigrant-link, and base-line, gravity literature, this paper adopted an ad-hoc stance to estimation. Specifically, the model remained as simple as possible and tested core gravity variables, in conjunction with a variable for the stock of immigrants. This simple specification, it was hypothesised, should yield high quality and significant results, such that the effects from immigration on EU-15 trade could be accurately quantified. The most reliable bilateral data available were obtained for evaluation. Equations for both imports and exports were estimated, with a new variable for 'remoteness' being utilised, one more suited to the paper, within the constraints of the data available.

The model worked surprisingly well, with both a high R^2 and very significant coefficients. The results indicated that immigration from Eastern European countries positively affects both the imports and the exports of EU-15 nations. In light of Paas's comments on the inaccuracies of the results of gravity models this is perhaps the most significant conclusion that can be drawn from the paper. A 10% rise in immigration from EU-expansion countries into the EU-15, is predicted to increase EU-15 imports by 1.4% and EU-15 exports by 1.2%. In terms of the mechanisms underlying the immigration-trade nexus investigated in this paper, these results suggest that immigrant-taste effects are more important than immigrant-link effects, in terms of generating EU-15 bilateral trade. The results from the estimation of the dynamic model indicated a surprising lack of auto regression in the model, though this could be accounted for by the short period under study.

The greatest constraint on the project was with the lack of data available, which ensured that more conventional testing of immigrant-links was prohibited. Reliable data for many of the countries in question was simply unavailable. Specifically it would have been very interesting to have obtained data on trade barriers. This unavailability also prevented the proxy for the stock of immigrant's variable being subjected to a thorough sensitivity analysis. This could have serious repercussions for the results of this project; though the most likely effect is a scaling up of the magnitudes of the trade elasticities with respect to immigration, as the proxy was a fraction of the total stock of immigrants. With more time, and better access to data, the most reliable stock of immigrant's variable would have been constructed from census, and flow data, from individual countries in question.

Immigration remains a topical issue of growing importance. This paper was confined to the affects of immigrant-links, there are many additional avenues open for future research within this framework. Additional data would facilitate the investigation of the joint effects of immigration and remittances that the literature has shown to be of importance (Harrison 2003). Data on the stock of EU-15 immigrants residing in EU-expansion countries would facilitate investigation as to the opposite immigrant-link effects, to those examined here. Data on the length of stay, and permanence of immigrant's, would mean one could look at the effects of temporary migration, an issue of growing importance since the inclusion of mode 4 of GATS - to allow for the freedom of movement of temporary workers. Indeed, if one were to successfully obtain these variables, one could look beyond an immigrant-link study, and gravity modelling, to model the wider effects of immigration; using a General Equilibrium model for example. Looking at the effects of intra-EU rural-urban migration would be a particularly interesting line of inquiry.

For now, this paper would concur with Paas, and say there is a good potential for trade between Eastern and Western European states. Immigration will likely boost this further in the future. On the basis of expanding EU-15 bilateral trade via increased immigrant flows this paper would support the notion for continued expansion of the EU.

Bibliography

Africano-Silva, A.P. and Teles, P. (2001), 'EU Enlargement and Trade Adjustments', The ETSG Conference – Brussels, 14 to 16 of September.

Anderson, J. E. (1979), 'A Theoretical Foundation for the Gravity Equation', American Economic Review, Vol. 69 (1), pp. 106-116.

Baldwin, R. François, J.F. and Portes, R. (1997), 'The Costs and Benefits of Eastern Enlargement', Economic Policy 24.

Baltagi, B.H., Egger P. and Pfaffermayr, M. (2003), 'A generalized design for bilateral trade flow models', Economics Letters, September vol. 80, issue. 3, pp. 391-397(7) Elsevier Science.

Bergstrand, J. H. 'The Gravity Equation in International Trade - Some Microeconomic Foundations and Empirical Evidence', Review of Economics and Statistics,Vol.67, 1985, pp.474-481.

Boeri, T. and Brücker, H. (2000), 'The Impact of Eastern Enlargement on Employment and Labour Markets in the EU Member States', a study made for the Directorate General for Employment and Social Affairs, European Integration Consortium, Berlin and Milano.

Boisso, D. and Ferrantino, M. (1997) 'Economic Distance, Cultural Distance, and Openness in International Trade: Empirical Puzzles', Journal of Economic Integration 4, 12: 456-484.

Carey, H. C. (1888), 'Principles of Social Science'. Philadelphia.

Cernat, L. (2003), 'Assessing south-south regional integration: same issues, many metrics: The trade impacts of seven Regional Trade Agreements on developing countries', Conference on Trade and Development (UNCTAD).

Combes, P.-P., Lafourcade, M. and Mayer, T. (May 16th 2002 draught), 'Can business and social networks explain the border effect puzzle?'

Cornett, A. P. and Iversen, S. P. (1998) 'The Baltic State in a European and Baltic Perspective: Trade and Transition in the Baltic Rim' Materials of the Conference "Global-Local Interplay in the Baltic Sea Region", Pärnu, Estonia October 1-4.

Deadorff A. V. 'Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World', in Frankel J.A., ed. The Regionalization of the World Economy, The University of Chicago Press, Chicago, London, 1998, pp. 7-31.

Diaz-Alejandro C.F. (1970), 'Essays on the Economic History of the Argentine Republic', Yale University Press: New Haven, CT.

Dixit, A. K. and Stiglitz J. E. (1977): 'Monopolistic Competition and Optimum Product Diversity', American Economic Review, 67, 297-308.

Dunlevy, J. A. and Hutchinson, W. K. (1999), 'The Impact of immigration on American import trade in the late nineteenth and early twentieth centuries', Journal of Economic History, pp. 1043-62.

Dunlevy, J. A. and Hutchinson, W. K. (2001), 'The pro-trade effects of immigration on American exports during the period 1870 to 1910', Vandablit University Working Paper 01-W25.

Eichengreen, B. and Irwin, D. A. (1997), 'The role of history in bilateral trade flows. Revised version published in Frankel J. A. (ed.), 'The regionalization of the World Economy', Chicago: University of Chicago Press. 1998.

Evenett, S. J., Keller, W. (1998), 'On Theories Explaining the Success of the Gravity Equation' NBER Working Paper, No W6529, Cambridge Mass.: National Bureau of Economic Research,

Frankel, J. A. (1997), 'Regional Trading Blocks in the World Economic system', Washington DC: Institute for International Economics.

Frankel, J.A. and Rose A. K. (2000), 'Estimating the Effect of Currency Unions on Trade and Output', NBER Working Paper No. w7857.

Frankel, J. and Rose, A.K. (2002), 'An estimate of the effect of common currencies on trade and growth', The Quarterly Journal of Economics 117, 61-75

Free A. (2001) 'The economic impact of enlargement', Enlargement Papers. No. 4. June 2001. European Commission. Brussels.

Girma, S. and Yu, Z. (2002), 'The link between immigration and trade: evidence from the UK', Weltwirtschaftliches Archive.

Griliches, Z. (1986), 'Economic data issues', in Griliches Z. and Intriligator M. D. (ed.), Handbook of Econometrics, Volume III', Elsevier Science Publications BV, Amsterdam.

Gould, D. 1994, 'Immigrant Links to Home Country: Empirical Implications for US Bi-lateral Trade Flows', Review of Economics and Statistics, 76: 302-316.

Gros, D. and Gonciarz, A. (1996), 'A Note on the Trade Potential of Central and Eastern Europe', European Journal of Political Economy, Vol.12, pp. 709-721.

Gujarati, D. N. (1995), 'Basic Econometrics', McGraw-Hill International Editions

Harrison A., Britton T., and Swanson A. (2003), 'Working Abroad – the benefits flowing from nationals working in other economies', OECD, presented to the Round Table on Sustainable Development on November 19 2003

Head K. (2000), 'Gravity for Beginners', .Material presented at Rethinking the Line: The Canada-U.S. Border Conference, Vancouver, British Columbia, October 22, 2000.

Head, K. and Ries, J. (1998), 'Immigration and trade creation: econometric evidence from Canada', Canadian Journal of Economics, pp. 47-62.

Helliwell, J. F. (1998), 'How Much Do National Borders Matter?' (Washington DC: Brookings Institution).

Helliwell, J. F. (1999), 'Language and trade'. In A. Breton (ed.), 'Exploring the Economics of Language', Canada: Canadian Heritage.

Helpman, E. and Krugman P. R., (1985) 'Market Structure and Foreign Trade', MIT Press.

Helpman, E. (1987), 'Imperfect Competition and International Trade: Evidence From Fourteen Industrial Countries', Journal of the Japanese and International Economies, Vol 1 (1).

Hutchinson, W. K. (2002), 'Does ease of communication increase trade? Commonality of language and bilateral trade', Working Paper No. 02-W17, Department Economics, Vanderbilt University, Nashville.

Krugman, P. 'Geography and Trade', Cambridge: MIT Press, 1991, 136 p.

Krugman, P. (1991) 'Increasing Returns and Economic Geography', Journal of Political Economy, Vol.3, pp. 483-499.

Leamer, E. E. and Stern, R. M. (1970), 'Quantitative International Economics', Allyn & Bacon, 1970.

Leamer, E. E. (1974), 'Nominal Tariff Averages with Estimated Weights', Southern Economic Journal, 41, 34-46.

Leamer, E. E. and Levinsohn, J., (1994), 'International Trade Theory: The Evidence', December. Revised version published in G. M. Grossman and K. Rogoff (eds.), 1995, Handbook of International Economics, Vol. III, North Holland, Amsterdam.

Linder, S. B. (1961), 'An Essay on Trade and Transformation', New York: Wiley & Sons.

Linnemann, H. (1966), 'An Econometric Study of International Trade Flows', Amsterdam: North-Holland Publishing Company.

Lowry, I.S. (1966), 'Migration and Metropolitan Growth: Two Analytical Models', San Francisco: Chandler Publishing.

Nelson . R.and Gaston, N. 'The Wage and Employment Effects of Immigration: Trade and Labour Economics Perspectives', In Greenaway, D. Upward, R. and K. Wakelin (Eds.) (2002), Trade, Investment, Migration and Labour Market Adjustment. Basingstoke: Palgrave Macmillan, pp. 201-235.

Newton, I, (1687), 'Philosophiae Naturalis Principia Mathematica', London: Printed by Joseph Streater by order of the Royal Society.

Nitsch, V. (2000), 'National Borders and international trade: evidence from the European Union'. Canadian Journal of Economics, 22, 4, pp. 1091-105.

McCallum, J. (1995), 'National borders matter; Canada-U.S. regional trading patterns', American Economic Review, 85, pp. 615-23.

Min, Pyong Gap, 'Korean Immigrants in Los Angeles', in Institute of Social Science Research, ISSR Working Papers 2 (2), University of California, Los Angeles, 1990.

Paas, T. (2002a), 'European Integration and EU Eastward Enlargement Process in International Trade: Using A Gravity Approach for Exploring Bilateral Trade Flows', The 42nd Congress of the European Regional Science Association, August 27-31, 2002,Dortmund, Germany.

Paas T. (2002b), 'Gravity Approach for Exploring Baltic Sea Regional Integration in the Field of International Trade', HWWA DISCUSSION PAPER180 Hamburgisches Weltwirtschaftliches Archive, Hamburg Institute of International Economics.

Paas, T. (2001), 'Modeling International Trade Flows of the Baltic Sea Region Countries. In: Global Business & Economics Review - Anthology ', Selected Papers of Business and Economics Society International, Worcester, USA, pp.607-618

Poyhonen, P. (1963), 'A Tentative Model for Volume in Trade Between Countries', Weltwirtschaftliches Archive, Vol. 90, pp. 91-113.

Raulch J. E. and Trindade V. (2002) 'Ethnic Chinese Networks in International Trade', The Review of Economics and Statistics, February, 84 (1), 116-130.

Ravenstein, E.G. (1889) , 'The Laws of Migration', Journal of the Royal Statistical Society, vol. 48 (June 1885) and vol. 52.

Stone, R. (1954), 'Linear expenditure systems and demand analysis: An application to the pattern of British demand', The Economic Journal, 64, 511-527.

Thursby, J. G. and Thursby, M. C., (1987), 'Bilateral Trade Flows, the Linder Hypothesis, and Exchange Risk', The Review of Economics and Statistics, August, 488-95.

Tinbergen, J. (1962), 'Shaping the World Economy', New York: Twentieth Century Fund.

Wagner, D. Head, K. and Ries, J. (2002), 'Immigration and Trade in the Provinces', Scottish Journal of Political Economy, Vol. 49, No. 5, November.

Wooldridge, M. R. (2003), 'Introductory Econometrics – A Modern Approach', Thomson South-Western.

Zipf, G. (1946), 'The PHD Hypothesis: On the Intercity Movement of Persons', American Sociological Review 11: 677-686.

Appendix 1

Table 1. Elasticity comparison

Authors	Sample and period	Additional complexities of interest	Export Elasticity*	Import Elasticity*
Gould (1994)	U.S. & 47 partners, 1970-86	Differentiated products	0.02	0.01
Head & Ries (1998)	Canada & 136 partners, 1980-92	Canadian immigration policy – immigrant heterogeneity	0.10	0.31
Dunlevy & Hutchinson (1999, 2001)	U.S. & 17 partners, 1870-1910	Differentiated products	0.08	0.29
Girma & Yu (2002)	UK & 48 partners, 1981-93	Individual vs. non-individual effects	0.16	0.10
Combes et al (2002)	95 French Departments, 1993	Intra- i.e. separate departments	0.25	0.14
Rauch & Trindade (2002)***	63 countries, 1980, 1990	Business networks, differentiated and homogenous products	0.47	0.47
Wagner, Head & Ries (2002)**	5 Canadian provinces, & 160 partners, 1992-95	Common Language and random encounter specification	0.16	0.41

Source: Wagner et al (2002)

Notes

Readers are asked to refer to Wagner et al (2002) for further discussion of how these estimates were obtained

*Trade elasticities with respect to immigration.

**Estimation without fixed effects

*** Estimation with differentiated products.

Appendix 2

Table 3. Regression results

	Imports		Exports	
	Pooled OLS	Fixed Effects	Pooled OLS	Fixed Effects
Migrant Stock	0.141 (0.126)***	0.141 (0.0138)***	0.116 (0.010)***	0.116 (0.011)***
Distance	-1.415 (0.047)***	-1.415 (0.042)***	-1.456 (0.057)***	-1.456 (0.034)***
GDP	1.110 (0.020)***	1.110 (0.020)***	0.976 (0.015)***	0.976 (0.016)***
Per capita GDP	-0.655 (0.126)***	-0.655 (0.095)***	0.282 (0.113)**	0.282 (0.076)***
Remoteness	-0.540 (0.092)***	-	-1.693 (0.075)***	-
(Overall) R2	0.8567	0.7752	0.8769	0.7680
R2 – 'within'	-	0.8005	-	0.8332
R2 – 'between'	-	0.8261	-	0.7333
No. Observations	1789	1789	1786	1786

Notes:

Estimates for all dummy variables and the constant term are not shown

Robust standard errors in parenthesis – to correct for heteroskedasticity

***, **, signify 99%, 95%, confidence intervals respectively

Appendix 3

Table 4. Elasticity comparison

Authors	Sample and period	Additional complexities of interest	Export Elasticity*	Import Elasticity*
Gould (1994)	U.S. & 47 partners, 1970-86	Differentiated products	0.02	0.01
Head & Ries (1998)	Canada & 136 partners, 1980-92	Canadian immigration policy – i.e. immigrant heterogeneity	0.10	0.31
Dunlevy & Hutchinson (1999, 2001)	U.S. & 17 partners, 1870-1910	Differentiated products	0.08	0.29
Girma & Yu (2002)	UK & 48 partners, 1981-93	Individual vs. non-individual effects	0.16	0.10
Combes et al (2002)	95 French Departments, 1993	Intra- i.e. separate departments	0.25	0.14
Rauch & Trindade (2002)***	63 countries, 1980, 1990	Business networks, differentiated and homogenous products	0.47	0.47
Wagner, Head & Ries (2002)**	5 Canadian provinces, & 160 partners, 1992-95	Common Language and random encounter specification	0.16	0.41
This paper	EU-15 & 15 EU-expansion countries,1994-2001	None	0.12	0.14

Source: Wagner et al (2002)

Notes

*Trade elasticities with respect to immigration.

**Estimation without fixed effects

*** Estimation with differentiated products.

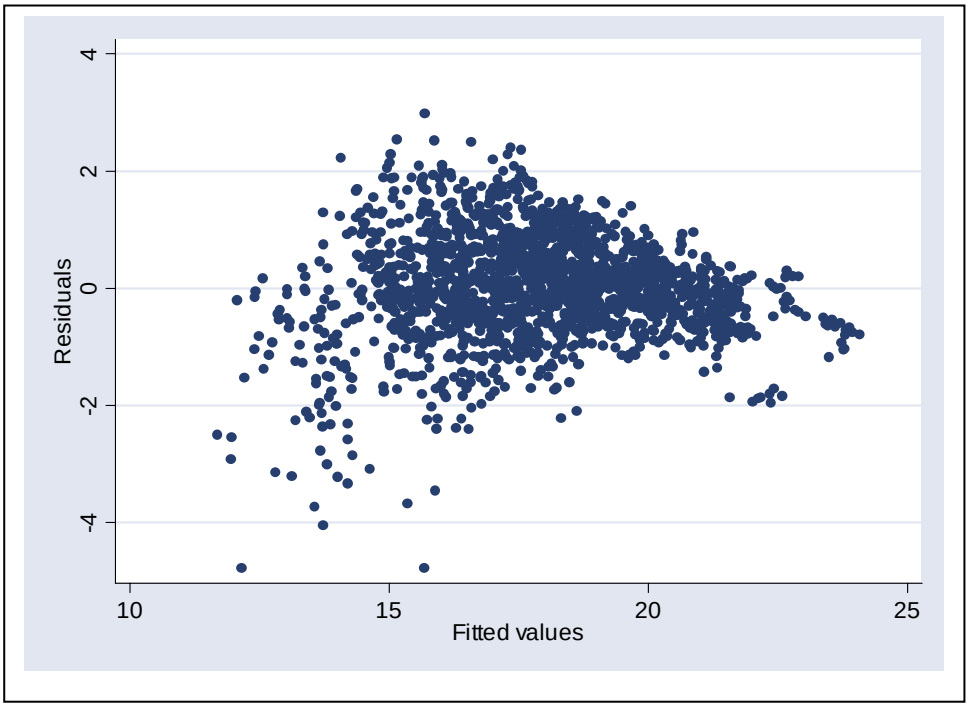
Appendix A – Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Number of Observations
Log imports					
Overall	17.90594	2.352807	7.38	23.29	N = 1789
Within		1.295605	15.83132	19.86442	n = 15
Between		1.994145	9.454621	21.70462	T-bar = 119.267
Log Exports					
Overall	18.4772	2.045353	8.29	23.34	N = 1786
Within		1.085051	16.6631	20.36908	n = 15
Between		1.756751	10.1041	22.10913	T-bar = 119.067
Log Migrant Stock					
Overall	2.949933	2.285828	0	11.45	N = 1800
Within		1.41933	1.345917	6.193167	n = 15
Between		1.828597	-3.243233	8.206766	T = 120
Log Distance					
Overall	7.221244	.6149635	4.02	8.24	N = 1800
Within		.3128473	6.788667	7.879333	n = 15
Between		.5355192	4.395244	8.157244	T = 120
Log bilateral GDP					
Overall	50.72089	1.769691	45.58924	54.97686	N = 1800
Within		1.276867	48.73689	52.88894	n = 15
Between		1.268568	47.53506	52.90756	T = 120
Log Remoteness					
Overall	-8.603115	.3958092	-9.387489	-7.913832	N = 1800
Within		.4095876	-9.387489	-7.913832	n = 15
Between		0	-8.603115	-8.603115	T = 120
Log Per Capita Incomes					
Overall	18.99909	.4590815	17.93909	20.56837	N = 1800
Within		.3651795	18.56708	19.53945	n = 15
Between		.2936342	18.24395	20.131	T = 120

Appendix B

Diagnostic Results, Part 1, results from Import Estimation

Chart 2. Plot fitted values and Residuals



F-Test for overall significance of the model

Ho: All coefficients are equal to zero. Ha: At least one coefficient is not equal to zero.	
F(25, 1763) = 392.42	Prob. > F = 0.0000

Therefore we can reject Ho that all the coefficients are simultaneously zero.

Cameron & Trivedi's decomposition of Heteroskedasticity test

Source	chi2	Degrees of freedom	P - value
Heteroskedasticity	621.53	217	0.0000
Skewness	85.54	25	0.0000
Kurtosis	9.32	1	0.0023
Total	716.39	243	0.0000

Therefore we can reject Ho that there is homoskedasticity.

Results of test for Multicollinearity

Variable	VIF	1/VIF
Log Migrant Stock	2.20	0.453523
Log Distance	1.53	0.653237
Log bilateral GDP	2.73	0.366701
Log Remoteness	3.12	0.320823
Log Per Capita Incomes	4.19	0.238787

As no VIF value is in excess of 20, or has a tolerance (1/VIF) of 0.05 we can be reasonably confident that there is no multicollinearity.

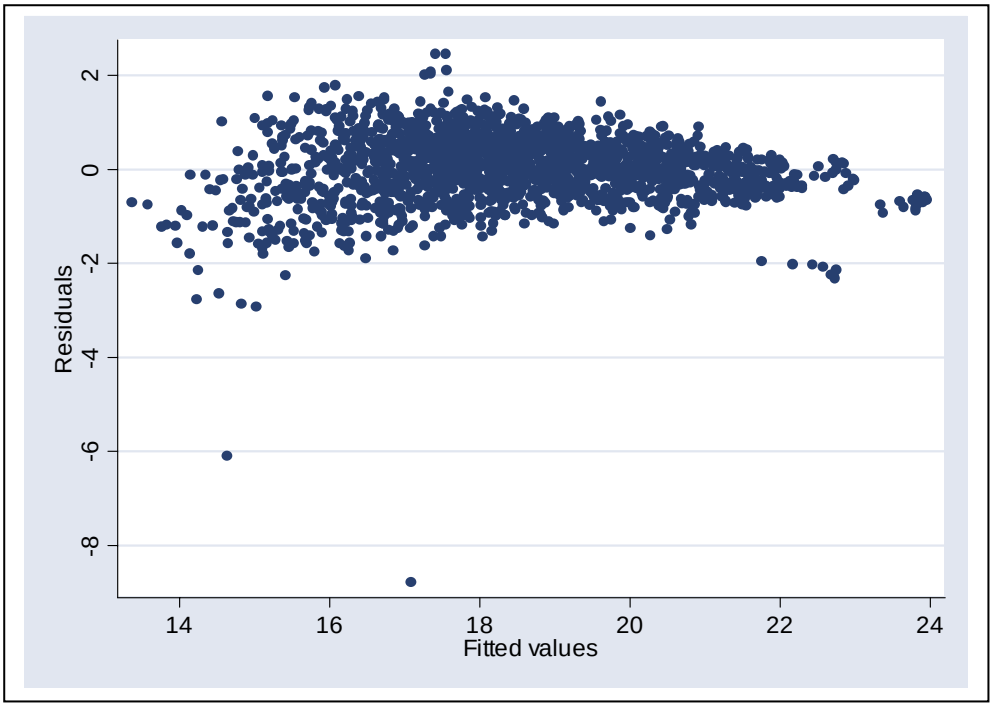
Results from Ramsey's RESET Test

Ramsey RESET test using powers of the fitted values of logimports
Ho: model has no omitted variables
F(3, 1760) = 101.55
Prob > F = 0.0000

Therefore we can reject Ho that there are no omitted variables.

Diagnostic Results, Part 2, results from Export Estimation

Chart 3. Plot fitted values and Residuals



F-Test for overall significance of the model

Ho: All coefficients are equal to zero. Ha: At least one coefficient is not equal to zero.	
F(25, 1760) = 480.09	Prob. > F = 0.0000

Therefore we can reject Ho that all the coefficients are simultaneously zero.

Cameron & Trivedi's decomposition of Heteroskedasticity – test

Source	chi2	Degrees of freedom	P - value
Heteroskedasticity	616.98	217	0.0000
Skewness	52.55	25	0.0010
Kurtosis	1.55	1	0.2125
Total	671.09	243	0.0000

Therefore we can reject Ho that there is homoskedasticity.

Results for test for Multicollinearity

Variable	VIF	1/VIF
Log Migrant Stock	2.20	0.453756
Log Distance	1.52	0.656860
Log bilateral GDP	2.75	0.363389
Log Remoteness	3.12	0.320224
Log Per Capita Incomes	4.17	0.240021

As no VIF value is in excess of 20, or has a tolerance (1/VIF) of 0.05 we can be reasonably confident that there is no multicollinearity.

Results from Ramsey's RESET Test

Ramsey RESET test using powers of the fitted values of logexports
Ho: model has no omitted variables
F(3, 1757) = 91.89
Prob > F = 0.0000

Therefore we can reject Ho that there are no omitted variables.

Appendix C

Results from Dynamic Estimation

	Static		Dynamic	
	Imports	Exports	Imports	Exports
Migrant Stock	0.141 (0.126)***	0.116 (0.010)***	0.133 (0.129)***	0.105 (0.010)***
Distance	-1.415 (0.047)***	-1.456 (0.057)***	-1.381 (.0477)***	-1.366 (0.062)***
GDP	1.110 (0.020)***	0.976 (0.015)***	0.999 (0.273)***	0.820 (0.22)***
Per capita GDP	-0.655 (0.126)***	0.282 (0.113)**	-0.659 (.0128)***	0.156 (0.111)
Remoteness	-0.540 (0.092)***	-1.693 (0.075)***	-0.257 (0.108)**	-1.281 (0.939)***
Lagged Dependent Variable	-	-	0.093 (0.019)***	0.150 (0.20)***
(Overall) R2	0.8567	0.8769	0.8589	0.8834
No. Observations	1789	1786	1777	1774

Notes:

Robust standard errors in parenthesis – to correct for heteroskedasticity

***, **, signify 99%, 95%, confidence intervals respectively