The Determinants of International Migration into the UK: A Panel Based Modelling Approach

James Mitchell and Nigel Pain

National Institute of Economic and Social Research

May 2003

Abstract

Relatively little is known about the factors behind the gross level of international migrant flows into the UK, despite the rapid growth in the number of migrants seen in recent years. This matters because assumptions about the future evolution of migration are an important component of official judgements about the potential trend rate of economic growth. In this paper we develop the first detailed econometric model of the economic and demographic determinants of annual migrant inflows into the UK from a number of different locations. Some of the factors found to be important, such as 'friends and family' effects from existing migrants, income differentials between the host and source location and the demographic structure of the source location population, have been shown to matter in other host economies as well. But we also show that it is important to allow both for developments in other potential host locations and for heterogeneity in the speed with which migration from different source locations responds to changes in economic circumstances. Neither of these factors have been given much attention to date in the applied literature on international migration. We find that the change in migration over the decade to 1998-2000 is primarily due to population growth in the source locations and the continuing pull effects from the rise in the migrant stock and per capita incomes in the UK relative to the source location.

National Institute of Economic and Social Research,
2, Dean Trench Street,
Smith Square,
London SW1P 3HE,
United Kingdom.
Tel: +44(0)20 7222 7665.
e-mail: jmitchell@niesr.ac.uk; npain@niesr.ac.uk

I. Introduction

Inward migration has once again become a widely debated political issue in the UK in recent years. This is not surprising, since the average gross annual inflow of migrants into the UK has almost doubled over the last two decades, with particularly rapid growth since the middle of the 1990s. If the higher level of (net) inward migration were to persist, it would have important economic consequences. In the 2002 Budget, the Treasury revised up their assumption of the long-term trend rate of growth of the British economy by ¼ of a percentage point (HM Treasury, 2002). A key component of this revision was the impact of an increase in the projected future growth of the working age population due to net migration, based on an assumption that the relatively high levels of (net) inward migration seen over 1997-99 would persist. Related estimates from the Government Actuaries Department (GAD) indicated that around two-thirds of the projected 5.1 million rise in the population of the UK between 2000 and 2025 would be due to migration (Population Trends, 2002).

Despite the clear importance of migration for future growth prospects, and the large body of theoretical literature on the determinants of international migration, little is known about why migration has risen so sharply into the UK, or indeed why it might be expected to continue to rise. This paper aims to help fill this gap. An understanding of the behavioural factors determining migration is necessary if the reliability of long-term forecasts of population growth and output growth are to be assessed and judgements made about the policy instruments available to affect the evolution of migrant flows.

The focus of the paper is on international migrants into the UK as recorded in the International Passenger Survey (IPS). These data are augmented by additional information on inflows from the Republic of Ireland. The data cover the majority of inward migrants, but exclude asylum seekers and visitor switchers. We estimate a panel data model for the determinants of the annual gross inflow of migrants into the UK over the period 1980-2000, distinguishing between ten different source locations. The principal reason for choosing to pursue a disaggregate approach rather than simply estimating a single relationship for total inflows, is that it allows two important questions to be tested – whether there are differences in the factors that ultimately drive migration from the different regions, and whether there are differences in the speed at which migration from each region responds to changes in the driving factors. Neither of these questions have been tested in existing empirical studies, either for the UK or elsewhere.

To date, little applied work has been undertaken on the determinants of migrant inflows into the UK. The model we estimate is much richer than those developed by van der Gaag and van Wissen (1999) and Glover e*t al.* (2001). It has some similarities with that contained in Hatton and Williamson (2002), who include the UK in their panel data model of the determinants of migration into eighty host locations since 1970, but in contrast to them we focus on gross inward migration, rather than net migration, and take specific account of heterogeneity in the determinants of migration from different locations.

Our results demonstrate clearly that biased and misleading conclusions can result from failing to allow for heterogeneity. In particular it is clear that migration from different locations responds at very different speeds to a common change in the key long-run determinants of migrant flows. But, after allowing for this, we do obtain statistical evidence in favour of the view that observed differences in migration patterns from different locations stem ultimately from differences in the evolution of the main drivers of migration across locations rather than from different long-run responses to common changes in one or more of those drivers.

An important feature of our empirical work which has been given comparatively little attention in the existing literature is a focus on economic conditions in the UK relative to those in other potential host locations, as well as relative to those in the source location from which migration takes place. We find that per capita incomes in the UK relative to those in four neighbouring Northern European economies have a significant positive impact on inward migration to the UK. Variations over time in the relative ease of entry into the UK and these other EU economies, as proxied by the past inflow of migrants as a proportion of the host population, do not however seem to have had a statistically significant effect on the aggregate level of new inflows into the UK. The other factors which are found to matter are the existing migrant stock in the UK, the level of bilateral trade between the UK and the source location, per capita incomes in the UK relative to the source location and the proportion of the source population aged 15-29.

In spite of the step-up in the number of migrants since the mid-1990s, we do not find evidence across the panel as a whole that there has been a significant, systematic rise in the level of inflows that cannot be explained by the economic and demographic forces that ultimately drive migration. This does not rule out the possibility that there has been a significant change in migration from an individual source region, but merely means that such changes are not common to migration from all regions. The aggregate level of migration over 1998-2000 is marginally higher than predicted by the estimated model, but the difference is not statistically significant. A decomposition of the changes in the average annual level of inward migration between 1988-90 and 1998-2000 reveals that higher relative per capita incomes in the UK, population growth in the source locations and a rising migrant stock account for most of the step-up in the level of inflows. All these factors suggest that the higher level of gross inflows should persist for some time, providing support for the judgement made by the UK Treasury.

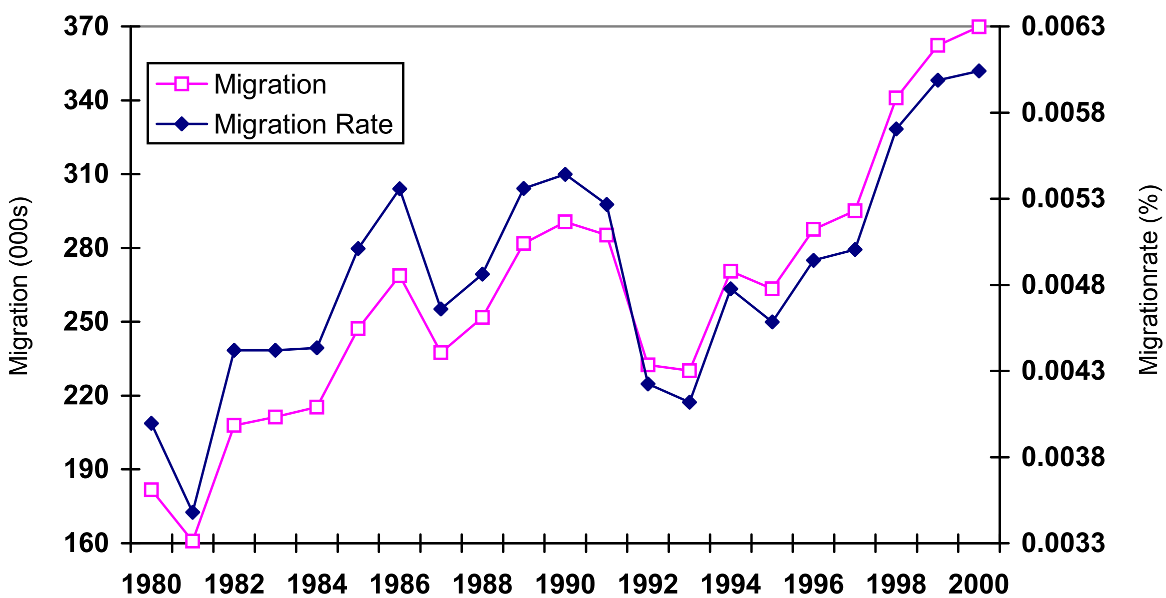
The structure of this paper is as follows. In Section II we discuss the level, and geographical origins, of inward migration into the UK over the past two decades. A simple theoretical model of the migration decision is outlined in Section III and the role and measurement of several potential key determinants of migration is discussed. Section IV reviews a number of different ways of estimating panel data models, and explains how allowance can be made for heterogeneous behaviour. The main empirical results are in Section V, with the structural stability and forecasting performance of the preferred specification being analysed in Section VI. Some concluding comments are given in Section VII.

II. International Migration into the UK

In this paper we focus on inward migration into the UK as recorded in the International Passenger Survey (IPS). The IPS is a continuous voluntary sample survey which covers about 0.2 per cent of travellers entering the UK, with people who state an intention of remaining in the UK for a year or more being classified as international migrants. Certain types of migrants are included in the official figures for Total International Migration to the UK but are not included in the IPS. In particular, the IPS excludes both asylum seekers and visitor switchers.[1] Thus the share of IPS migrants in the total annual inflow of migrants into the UK has fluctuated over time. Over the period for which consistent data are available, 1984-2000, it averaged 80.5 per cent, but in recent years it has been lower, averaging 77.1 per cent in 1999-2000.

Figure 1 shows the long-term trend in inward migration as recorded by the IPS, looking both at the flow and the migration rate. The latter is the ratio of the annual flow of IPS migrants from a particular location expressed as a proportion of the aggregate population in that location. Over the past two decades the long-term trend in the number of migrants into the UK, in fact, is very similar to that in the aggregate migration rate, with an upward trend over time, offset by cyclical downturns in the immediate aftermath of the recessions in the early 1980s and 1990s.[2]

Figure 1. Gross IPS Migrant Inflows into the UK



The IPS is a rich source of information, with migrants being classified by both their previous country of residence and their country of citizenship. In examining the economic determinants

[1] Visitor switchers are short-term visitors who subsequently stay for a year or more.

[2] The data for the total global population, as well as that for the source groups we use below, come from summing the total populations of the constituent countries using annual data from the April 2002 IMF *World Economic Outlook Database*.

of the migration decision, the last country of residence appears the most natural choice of classification, particularly as we wish to test whether migration is related to observable differences in economic conditions in the UK and the source country from which migration takes place.

Given the relatively small samples from which the country-specific information in the IPS is generated, it is not practicable to include a large number of separate source countries in an empirical exercise because of the large sampling error attached to any point estimate. Instead we have chosen to use ten separate groupings: (i) European Union (current members, including Ireland), (ii) USA, (iii) Rest of Europe (excluding Commonwealth members), (iv) Old Commonwealth (Australia, New Zealand, Canada and South Africa), (v) African Commonwealth (excluding South Africa), (vi) Asian Commonwealth, (vii) Other Commonwealth, (viii) Rest of Africa and the Middle East, (ix) Rest of Asia (including Hong Kong) and (x) Latin America.[3]

The IPS data were classified into these groups using information on country of last residence prior to migration. The membership of groups such as the Commonwealth and the EU has changed over time. Where possible we use data for the whole period based on the present day membership. Data for migration from the Irish Republic are excluded from the IPS data classified by citizenship (ONS, 2002), but are available separately and so have been added in to create a combined total.[4]

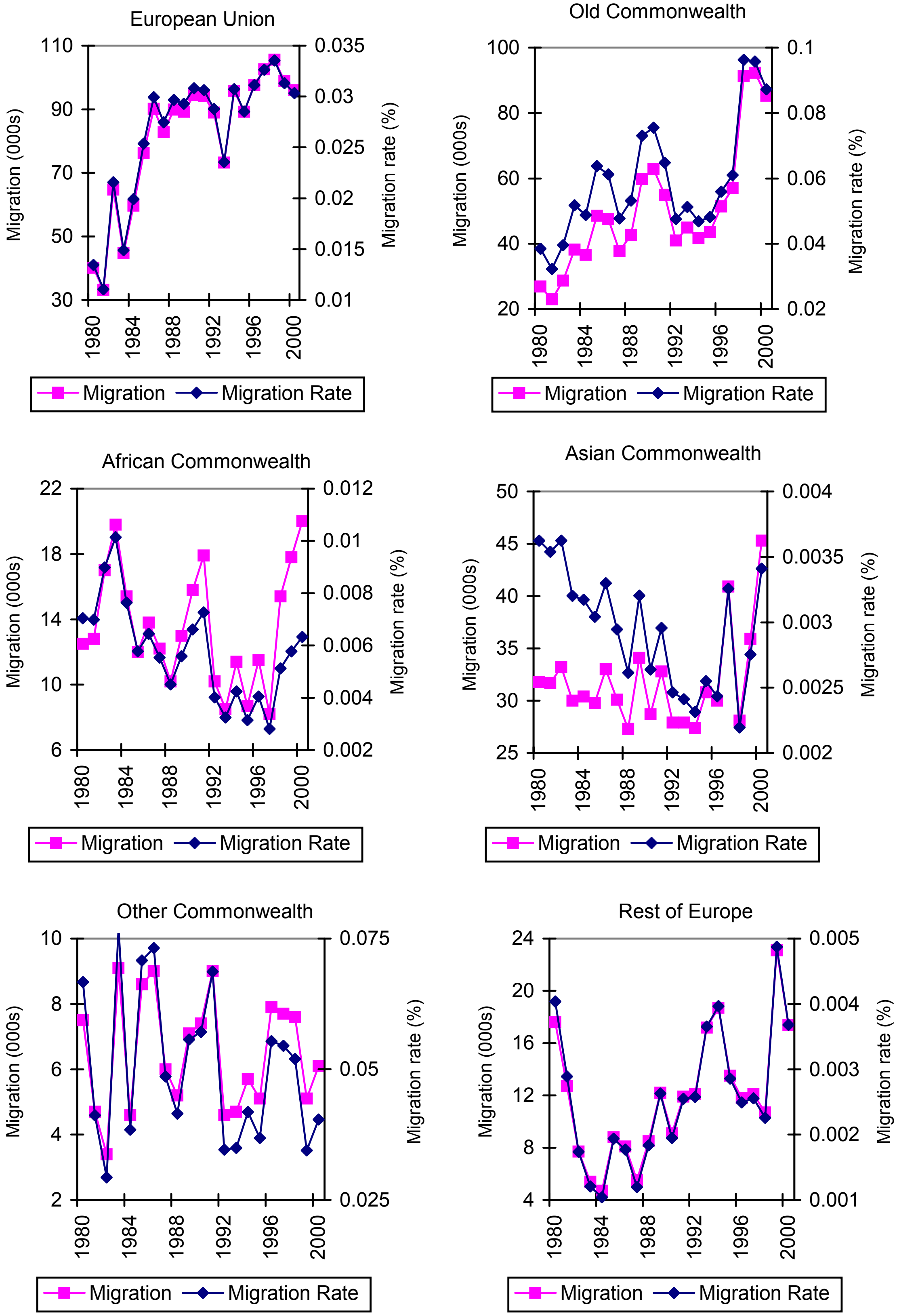
The chosen groupings cover the main regions of policy interest, and allow tests of a number of hypotheses, such as whether the determinants of migration from the Commonwealth countries and the industrialised regions differ significantly from the determinants of migration from other areas. Some of the groups, such as the Rest of Europe which includes Switzerland, Turkey and the transition economies, are relatively heterogeneous, but attempts at disaggregating them significantly further were not possible given the limitations of the available data.

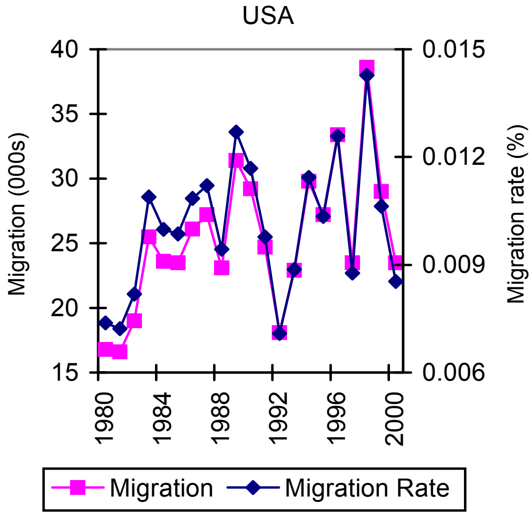
The close time series relationship between the annual migrant flow and the migration rate is also apparent in most of the ten regional groupings, as can be seen from Figure 2. However there are some exceptions, especially in the Commonwealth countries. The recent rise in the number of migrants from the African Commonwealth is not matched by a rising migration rate because of source country population growth. For the Asian Commonwealth, population growth means that the current migration rate is little different from that in the early 1980s, even though the absolute number of migrants is much higher now than at that time.

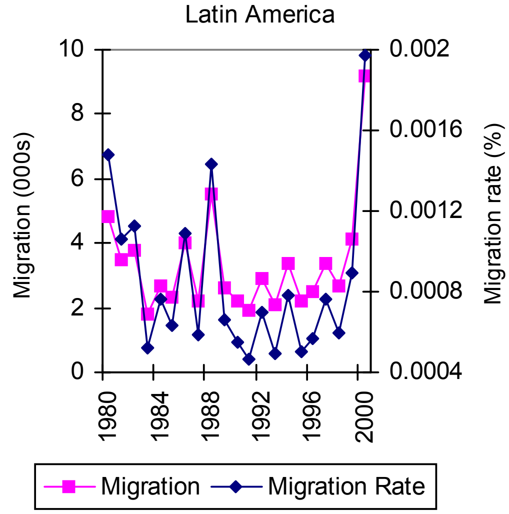
[3] Even at this level of aggregation there are potentially large sampling errors. For instance in 2000 there were an estimated 9,200 migrants from Latin America recorded in the IPS, with an estimated standard error of 2,500 (27.5%) (ONS, 2002, Table 3.21).

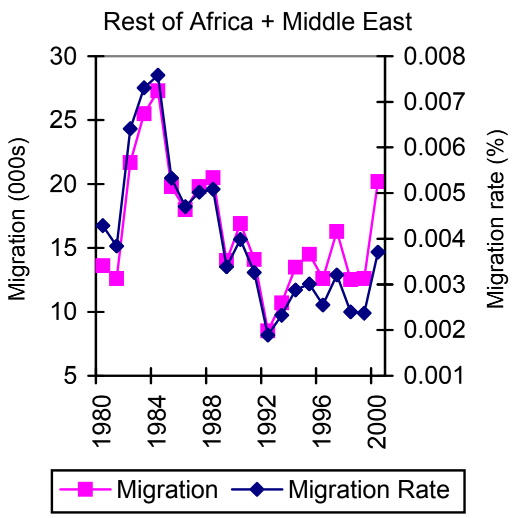
[4] Estimates for 1984-2000 were provided by the Home Office. Data for 1980-83 were obtained by interpolation using information about the number of Irish-born persons on the NHS Central Register (Eurostat, 2001, Table A1).

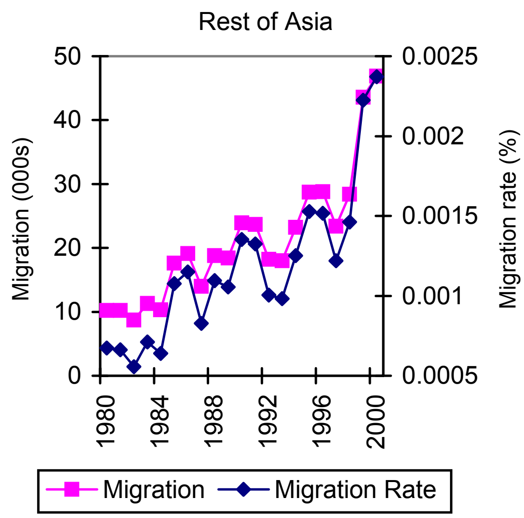
Figure 2. IPS Inward Migration by Source Location











In absolute terms, as Table 1 shows, the most important source locations are the EU and the Old Commonwealth countries, followed by the two Asian groups and the US. The numbers of migrants from the Rest of Africa group have fallen over time, as has the migration rate. The migration rate is highest in the Old Commonwealth, where annual migration to the UK is presently equivalent to just under 0.1 per cent of the total population.

Table 1. Migration Inflows By Region of Last Residence (5 year annual average, 000s)

	1981-85	1986-90	1991-95	1996-2000
Total	208.6	266.1	256.5	331.4
European Union	55.7	89.3	88.3	100.1
Old Commonwealth	35.0	50.1	45.2	75.5
African Commonwealth	15.4	13.0	11.3	14.6
Asian Commonwealth	31.0	30.6	29.4	36.0
Other Commonwealth	6.1	6.9	5.8	6.9
Rest of Europe	7.9	8.7	14.7	15.0
USA	21.6	27.4	24.5	29.6
Latin America	2.8	3.3	2.5	4.4
Other Africa and Middle East	21.4	17.8	12.3	14.8
Rest of Asia	11.6	18.8	22.4	34.2

Source: calculations from International Passenger Survey database.

III. A Theoretical Model Of International Migration

The basic framework within which most models of migration can be viewed is one in which migration decisions are based on a consideration of the relative expected future incomes from residing and working in different locations, adjusted for the costs of migrating. We illustrate this using a model of the microeconomic foundations of an individual's migration decision, adapted from the model in Hatton (1995) (see also Fertig, 2001). This can be used to derive an estimable model of the migration rate. Whilst the model does not capture all the diverse motivations for different forms of migration, it does provide a useful framework in which a wide variety of factors can be introduced.[5]

The model has two particularly important features. It incorporates uncertainty explicitly into the migration decision and acknowledges that migration decisions will be based upon expectations of future economic developments. These mean that there will not necessarily be a close connection between the observed migration rate and current differences in economic conditions in different locations. For instance, migration may take place even if current economic

[5] The so-called 'new economics of labour migration' emphasises the role of wider social factors in migration decisions, especially for migrants from developing countries, and treats the household rather than the individual as the main decision unit (Stark and Bloom, 1985; Stark and Taylor, 1991). Migration decisions in this approach are often made jointly by migrants and non-migrants. One example would be foreign migration to minimise the risk of inadequate household income in developing countries in which much of the population is dependant on agricultural incomes. Here, non-migrants benefit from the income remitted by the migrant and the risks of inadequate income are pooled if the migrant moves into an occupation where earnings are negatively correlated with those of the rest of the family. There is also empirical evidence which suggests that some individuals migrate in order to improve their comparative income position relative to other individuals or households in a relevant reference group. Two implications that stem from such models are that international migration decisions need not involve once-for-all relocation and that an increase in absolute income levels in the origin location can still be associated with higher outward migration if it also results in greater relative deprivation.

conditions are unfavourable provided that they are expected to improve in the future so that the net discounted present value of migrating is positive. At other times, when future economic conditions are uncertain, migrants may prefer to take the option of waiting for additional information. Anything which reduces uncertainty, such as improved information flows might then stimulate migration even if there is no change in current income levels.

III.1 The formal model

Suppose that the probability of migration of individual *i* from the home country (*h*) to the foreign country (*f*) depends on the difference in expected utility streams in the two locations, minus the costs of migration (z_i); we denote this difference by d_i. The utility function is assumed to be concave and dependent upon the (log of) expected income (y_j) in country *j*. Hence:

$$d_i = E\ln(y_f) - E\ln(y_h) - z_i \qquad [1]$$

Expanding $E\ln(y_j)$ around $E(y_j)$ using a second-order Taylor series expansion gives:

$$E\ln(y_j) = \ln(Ey_j) - \frac{var(y_j)}{2(Ey_j)^2}. \qquad [2a]$$

Assume that all incomes are derived from employment. Expected income will then depend on the (migrant-specific) real wage (*w*) multiplied by the probability of employment, which is a function of the employment rate (or 1 minus the unemployment rate). Assuming the uncertainty in *y* is due to *e* rather than *w*, and that the probability of employment follows a binomial distribution in which the expected value of the employment rate is *e* and the variance is *e(1-e)*, then:

$$\frac{var(y_j)}{2(Ey_j)^2} = \frac{1}{2}\frac{w_j^2 e_j(1-e_j)}{w_j^2 e_j^2} = \frac{1}{2}\frac{(1-e_j)}{e_j} \approx -\frac{1}{2}\ln(e_j). \qquad [2b]$$

Uncertainty about employment leads to it being given a greater weight than wages, so that we can express expected incomes as:

$$E\ln(y_f) = \ln(w_f) + \frac{3}{2}\ln(e_f). \qquad [2c]$$

$$E\ln(y_h) = \ln(w_h) + \frac{\gamma 3}{2}\ln(e_h), \qquad [2d]$$

where γ, γ<1, is a parameter which reflects the fact that migrating typically involves more risk than not migrating.

The difference in the expected utility streams can now be written as:

$$d_i = \ln(w_f) + (3/2)\ln(e_f) - \ln(w_h) - (\gamma 3/2)\ln(e_h) - z_i \qquad [3]$$

The decision to migrate does not depend just on current utility but also on the future values of the stream of expected utility at home and abroad. To reflect this inter-temporal dimension to the migration decision, let d_{it}^{*} denote the expectation at time *t* of the NPV of the difference in expected utility streams from *(t+1)* onwards. Then the NPV of moving today is $d_{it}^{*} + d_{it}$. Thus the individual probability of migrating at time *t* (denoted by $m_{it}=1$) is:

$$\Pr(m_{it}=1)=\Pr(d_{it}*+d_{it}>0 \cap d_{it}>0) \qquad [4]$$

The aggregate migration rate, assuming it reflects the average probability of individuals from country *h*, then can be expressed as:

$$M_t = \beta(d_t * + \alpha d_t) \qquad [5a]$$

where β measures the impact of the difference in utility streams on the aggregate migration rate and α reflects the extra weight given to current conditions, given that potential migrants could choose to wait if $d_{it}<0$.[6]

Now, assume that expectations of future utility streams are formed by a geometric series of past values of d_t such that:[7]

$$d_t* = \lambda d_t + \lambda^2 d_{t-1} + \lambda^3 d_{t-2} + \lambda^4 d_{t-3} + \ldots \qquad [5b]$$

Taking a Koyck transformation yields:

$$M_t = \beta(\alpha + \lambda)d_t - \lambda\beta\alpha d_{t-1} + \lambda M_{t-1} \qquad [6]$$

Substituting equation [3] into equation [6] gives:

$$\begin{aligned} M_t = {} & \beta(\alpha+\lambda)\left[\ln\left(w_{ft}/w_{ht}\right)+(3/2)\ln(e_{ft}/e_{ht}^{\gamma})-\bar{z}_t\right] \\ & -\lambda\beta\alpha\left[\ln\left(w_{f,t-1}/w_{h,t-1}\right)+(3/2)\ln(e_{f,t-1}/e_{h,t-1}^{\gamma})-\bar{z}_{t-1}\right]+\lambda M_{t-1} \end{aligned} \qquad [7]$$

where $\bar{z}_t$ is the mean of z_i over all *i* at time *t*.

Assume that the costs of migration from *f* to *h* are (negatively) related to the stock of previous migrant to *f* from country *h* because of network effects (Carrington *et al*, 1996), plus a time trend and a constant. The migration cost function then can be expressed as:

$$\bar{z}_t = \varepsilon_0 - \varepsilon_1 MST_{t-1} - \varepsilon_2 T \qquad [8]$$

where MST_{t-1} is the stock of migrants from *h* at the end of *t-1*, expressed as a ratio of the source country population (i.e. $MST_t = S_t / P_t$) , T is a deterministic linear trend and ε_i, $i=0,1,2$, are coefficients. The stock (*S*) diminishes at rate (1-δ) due to deaths and return immigration and

[6] Even though the net present value (NPV) of migrating this year may be positive, it may be even higher next year. In this case potential migrants have an option value to waiting.

[7] This is equivalent to rational expectations if *d* follows an AR(1) process; see Hatton (1995), p. 410.

increases due to new immigrants; $S_t = \delta S_{t-1} + M_t P_t$. Denoting the rate of population growth by g (assumed for expositional ease only to be the same across source countries), it follows that:[8]

$$MST_t = [\delta/(1+g)]MST_{t-1} + M_t \,. \qquad [9]$$

Using this expression, the costs of migration at time *t-1* can be expressed as:

$$\bar{z}_{t-1} = \varepsilon_0 - \varepsilon_1\left[\frac{(1+g)}{\partial}(MST_{t-1} - M_{t-1})\right] - \varepsilon_2(T-1) \qquad [10]$$

Substituting [9] and [10] into [7] and re-arranging, generates the following error-correction model:

$$\begin{aligned}
\Delta M_t &= \beta(\alpha+\lambda)\Delta\ln(w_f/w_h)_t + \beta(\alpha+\lambda)(3/2)[\Delta\ln(e_f/e_h^{\gamma})_t] \\
&+ \beta(\alpha+\lambda-\lambda\alpha)\ln(w_f/w_h)_{t-1} + \beta(\alpha+\lambda-\lambda\alpha)\frac{3}{2}\ln(e_f/e_h^{\gamma})_{t-1} \\
&+ \beta(\lambda\alpha-\alpha-\lambda)\varepsilon_0 + \left[\beta(\alpha+\lambda)\varepsilon_1 - \frac{\lambda\beta\alpha(1+g)}{\delta}\varepsilon_1\right]MST_{t-1} + \beta(\alpha+\lambda-\lambda\alpha)\varepsilon_2 T \qquad [11] \\
&+ \lambda\beta\alpha\varepsilon_2 + \left[\lambda + \frac{\lambda\beta\alpha(1+g)}{\delta}\varepsilon_1 - 1\right]M_{t-1}
\end{aligned}$$

where Δ denotes the first-difference operator. The migration rate is a function of relative wages and relative employment rates in the host and source country, the existing stock of migrants, the lagged migration rate and a time trend. The unemployment rate (UR) can be used in place of the employment rate if desired, by noting that e=1–UR. The lagged migration rate and the stock of migrants both enter the equation. By solving equation [11] for its implied steady-state solution we can see how each of the variables affects the migration rate in the long-run; see Section IV below. Crucially we can distinguish between dynamic and long-run determinants of migration.[9]

A wide range of additional factors can be introduced within the framework set out above, either by changing the assumptions about the factors determining expected income streams in different locations, or by changing the number of potential locations, or by changing the specification of the function determining the costs of migration.

In the context of the increasing globalisation of economic activities and the growing levels of high-skilled migrants (Regets, 2001; McLaughlan and Salt, 2002) it might be argued that the probability of employment for some migrants will depend on the level of business linkages between the source and the host countries, with business networks or multinational firms

[8] This formulation overcomes the notational inconsistency in the Hatton (1995) model discussed in Fertig (2001).

[9] We consider equation [11] a 'reduced-form' model in the sense that, like Hatton (1995) and Fertig (2001), we do not attempt to recover the structural parameters.

providing job opportunities. Indicators of bilateral trade or foreign direct investment could then be introduced within the model.[10]

III.2 The Empirical Model

Detailed surveys of the empirical literature on international migration are provided in Borjas (1989, 1994, 1999a, 1999b), Ghatak *et al.* (1996) and Mitchell and Pain (2002).[11] A wide range of economic, political and social variables have been put forward as potential determinants of migration decisions. With only 20 observations per region in our data set, it is not possible to include all the potential explanatory variables that have been used in the various empirical papers.[12] Instead we have sought to build on both the theoretical structure of the Hatton (1995) model set out above, as well as the approaches taken by Rotte and Vogler (1998) and Clark *et al* (2002), whilst focusing on six key variables:

- per capita income in the UK relative to that in the source group
- the UK unemployment rate
- per capita income in the UK relative to that in alternative host countries
- the proportion of the population in the source group aged 15-29
- the volume of bilateral merchandise trade between the UK and the source group expressed as a share of their combined GDP
- the lagged migrant stock

In this section we provide a brief explanation of the impact each of these factors is expected to have, discuss data sources and examine their cross sectional and time series variation.

Expected Incomes

Current income levels and labour market conditions are two factors which are likely to be important determinants of expected future incomes in the UK and the source country, as shown by the theoretical model set out above. In the empirical work we use both the level of relative income and the growth of incomes in the UK relative to the growth of incomes in the source location. As in many other studies we measure income in terms of GDP per capita. GDP and

---

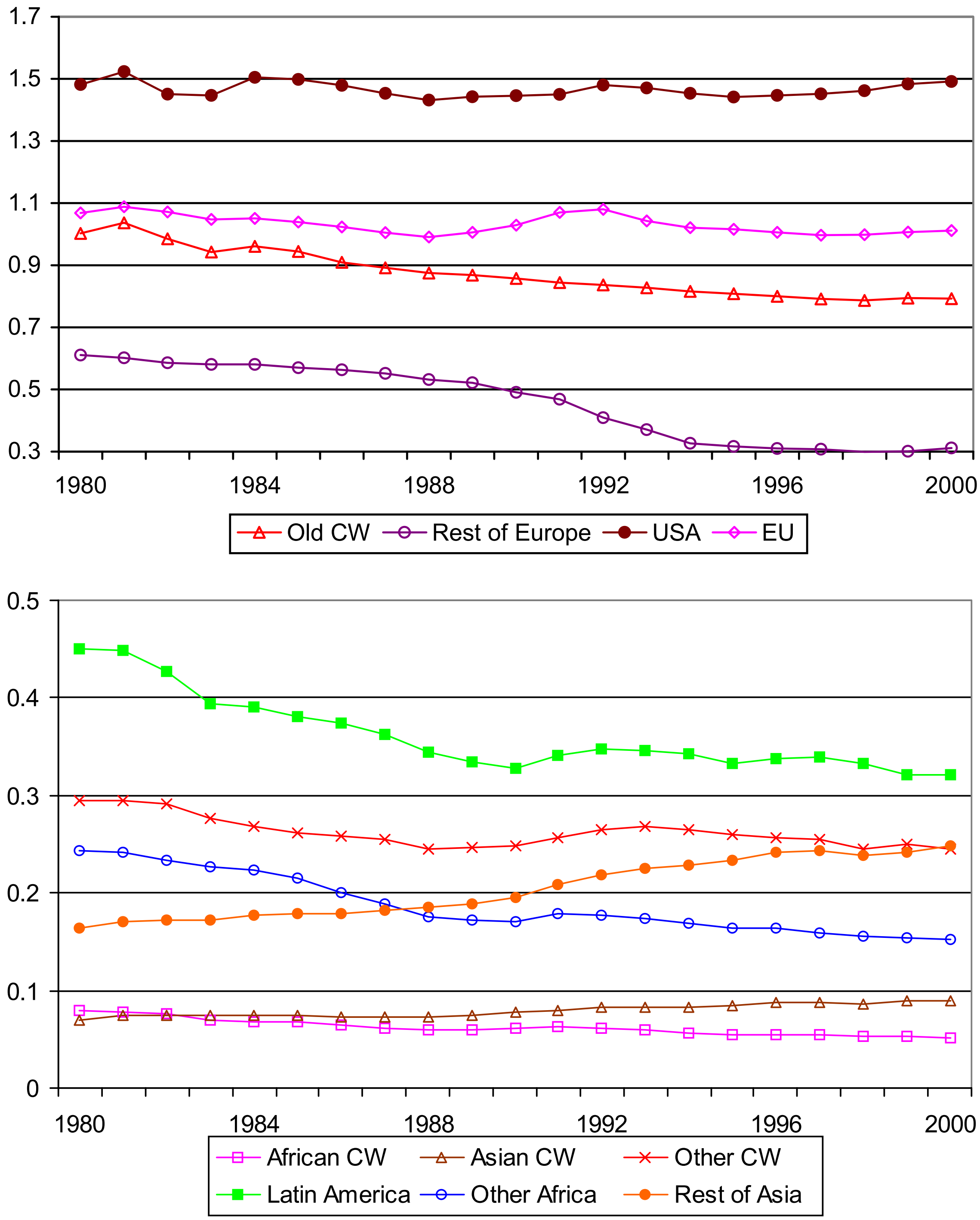
[10] A further extension would be to specify utility in terms of total, post-tax disposable incomes. In this case fiscal factors such as the provision of social security benefits and the structure of the tax system would also appear in the migration rate equation. Note however that the model also strictly requires information about these in the source country, and this might be difficult to obtain, especially for developing economies. This may be one reason why such factors do not appear in many empirical models of migration.

[11] There are additional studies of specific types of migrants such as highly-skilled workers, asylum seekers and graduates. Many of these studies utilise qualitative survey evidence. Recent examples include Robinson and Segrott (2002), Koser and Pinkerton (2002), Regets (2001), McLaughlan and Salt (2002), Winkleman (2001), Birrell et al. (2001) and Finn (2001).

[12] See Karemera et al. (2000), Hatton and Williamson (2001), Jasso et al. (1998), Clark *et al.* (2002), Rotte and Vogler (1998) and Fertig (2001) for studies that exploit a panel data set on migration.

population data are both taken from the April 2002 IMF *World Economic Outlook Database*. The GDP data are in 1995 prices, with individual country estimates converted into US dollars using 1995 Purchasing Power Parity exchange rates. GDP per capita in each of the ten source locations relative to that in the UK is shown in Figure 3.

Figure 3. Per Capita Incomes Relative to the UK ($1995 PPP and prices)



There appears to be little consistent relationship across locations between the time series pattern of relative incomes and the migration rates shown in Figure 2. In some cases the variation in the relative income measure is clearly similar to that in the migration rate. For instance, the migration rate from the Old Commonwealth and the Rest of Europe has risen over time, whilst per capita income in these locations relative to that in the UK has declined. But similar declines in incomes can be seen in Latin America and Africa and these have not been accompanied by a rise in the migration rate. Equally, the migration rate from the Rest of Asia has risen over time, even though incomes there have risen relative to the UK, which might, other things being equal, reduce the incentive to migrate for any given level of migration costs.[13]

The UK claimant unemployment rate was employed as a explanatory factor in the simple model of net migration into the UK estimated by Glover et al. (2001) and found to be significantly negatively related to net in-migration. As would be expected the unemployment rate shows marked cyclical fluctuations over time. In recent years there has also been a marked downward trend in the level, coinciding with the pick-up in the annual level of migrant inflows.[14]

Income Relative To Other Hosts

The principal theoretical innovation of this paper is to take account explicitly of the fact that some, if not all, migrants have a choice about where they choose to locate. It is quite possible that migration into the UK from source countries will depend on economic and social conditions in other potential hosts, either in the European Union or in other industrialised economies outside Europe. Very few empirical studies have attempted to model this choice.[15]

Of course an immediate difficulty lies in knowing which alternative hosts have been considered, and so an assumption has to be made. We make the arbitrary, albeit plausible, assumption that that the relevant comparator group comprises a weighted average of four Northern EU

[13] It is possible that migration is only affordable once source country incomes have reached a certain level. To allow for the possibility of an inverse U-shaped relation between economic development and migration, some authors have introduced the square of source country income into the model as well as the level; see, for instance, Rotte and Vogler (1998).

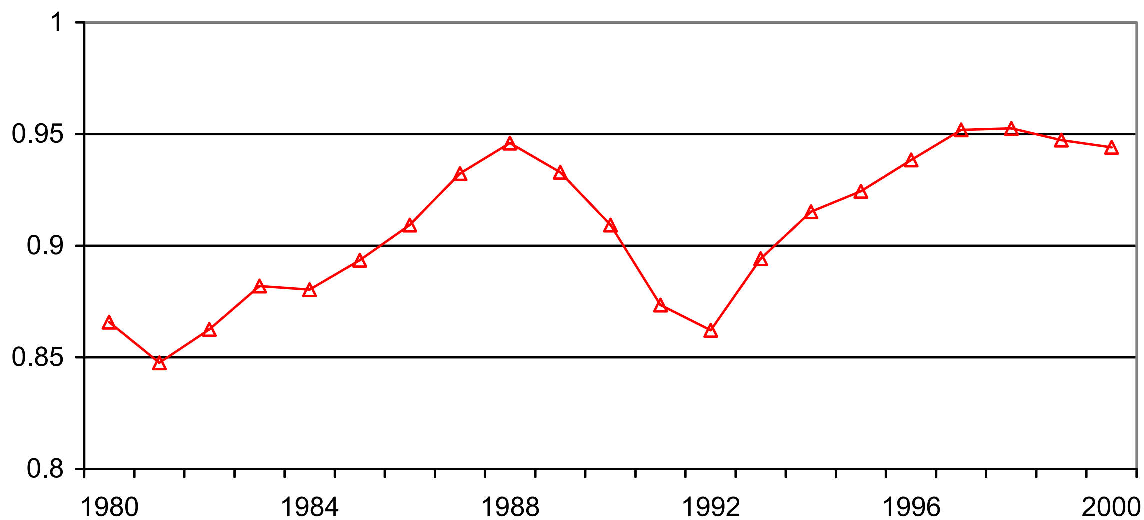
[14] In principle it might also be useful to consider labour market conditions in the other EU potential hosts and in the source region. However we did not find the former to be important in any of the empirical work, and so this is not included here, whilst we lack consistent time series data on unemployment rates in all of the source regions, especially the developing ones.

[15] In contrast the separate literature on the location choice of firms, and in particular foreign direct investment, has examined the relative characteristics of different locations; e.g. see Barrell and Pain (1999a,b). In the migration literature there has been relatively little attempt to model this choice, with the exception of van der Gaag and van Wissen (1999), who however omit a number of other potential determinants we consider. Clark *et al.* (2002) control for third country effects by including fixed effects in the panel data model. However this is not adequate if the third country effects vary over time.

economies – France, Germany, Belgium and the Netherlands, and include per capita income in the UK relative to that in these economies.[16]

This series is shown in Figure 4. It is clear that incomes in the UK have risen over the past 20 years relative to those in the EU4, although the longer-term upward rise came to a halt in the latter half of the 1990s.

Figure 4. UK Per Capita Incomes Relative to EU4 (1995 prices)



The Demographic Share

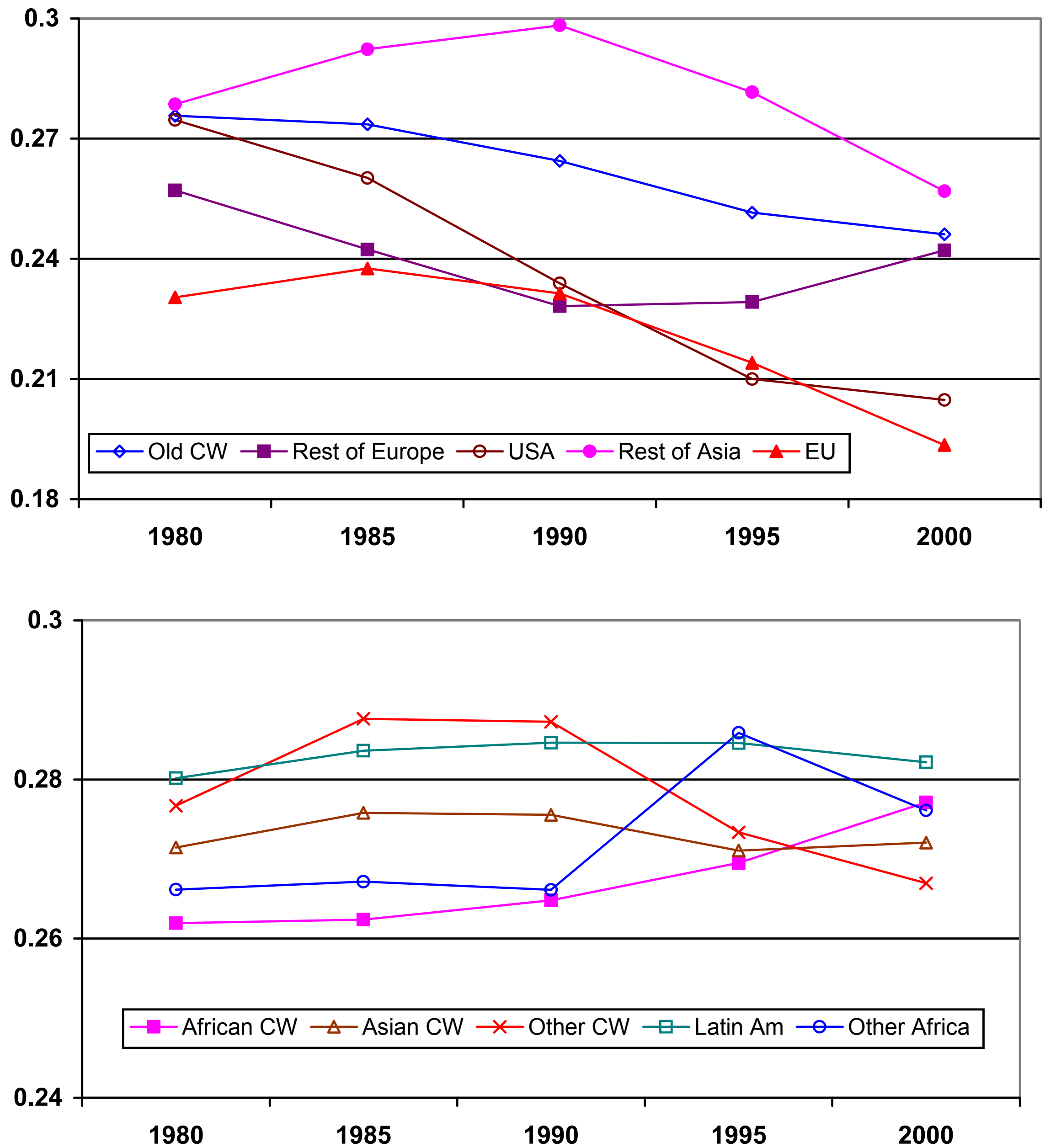
The demographic structure of the source country is included because, for any given expected income differential between the UK and the source country, the present discounted value of the gains from migration is likely to be higher for younger potential migrants. We follow other studies such as Clark *et al* (2002) and Hatton and Williamson (2002) in attempting to allow for such effects by including the proportion of the population aged 15-29 in each source location. The data are calculated from information in the *World Population Prospects: The 2000 Revision*, issued by the United Nations. Information is available at 5 year intervals, so the annual data have to be constructed by interpolation, assuming that the population proportion evolves smoothly between each fixed point. The data are plotted in Figure 5.

The share of the population aged 15-29 has declined consistently over time in the three locations with the highest per capita incomes – the EU, the USA and the Old Commonwealth, with the first two of these now having a ratio well below that in the other locations. It has also fallen in the Rest of Asia and the Other Commonwealth groups since 1990 after having risen in the 1980s. In contrast the ratio rose in the Rest of Europe in the 1990s after having declined in the

[16] We have not attempted a systematic investigation of all possible alternative groupings. However, we do comment below on some findings using a weighted average of the other economies in the G7.

1980s. A similar upward trend in the 1990s can be seen in the African Commonwealth, but in this case, unlike that of the Rest of Europe, it was not accompanied by a rise in the migration rate to the UK. This provide one illustration of the potential difficulties that may be faced in finding a common relationship for migration from different locations if attention is focused solely on a single potential determinant of migration.

Figure 5. Share of Population Aged 15-29 (%)



Bilateral Trade

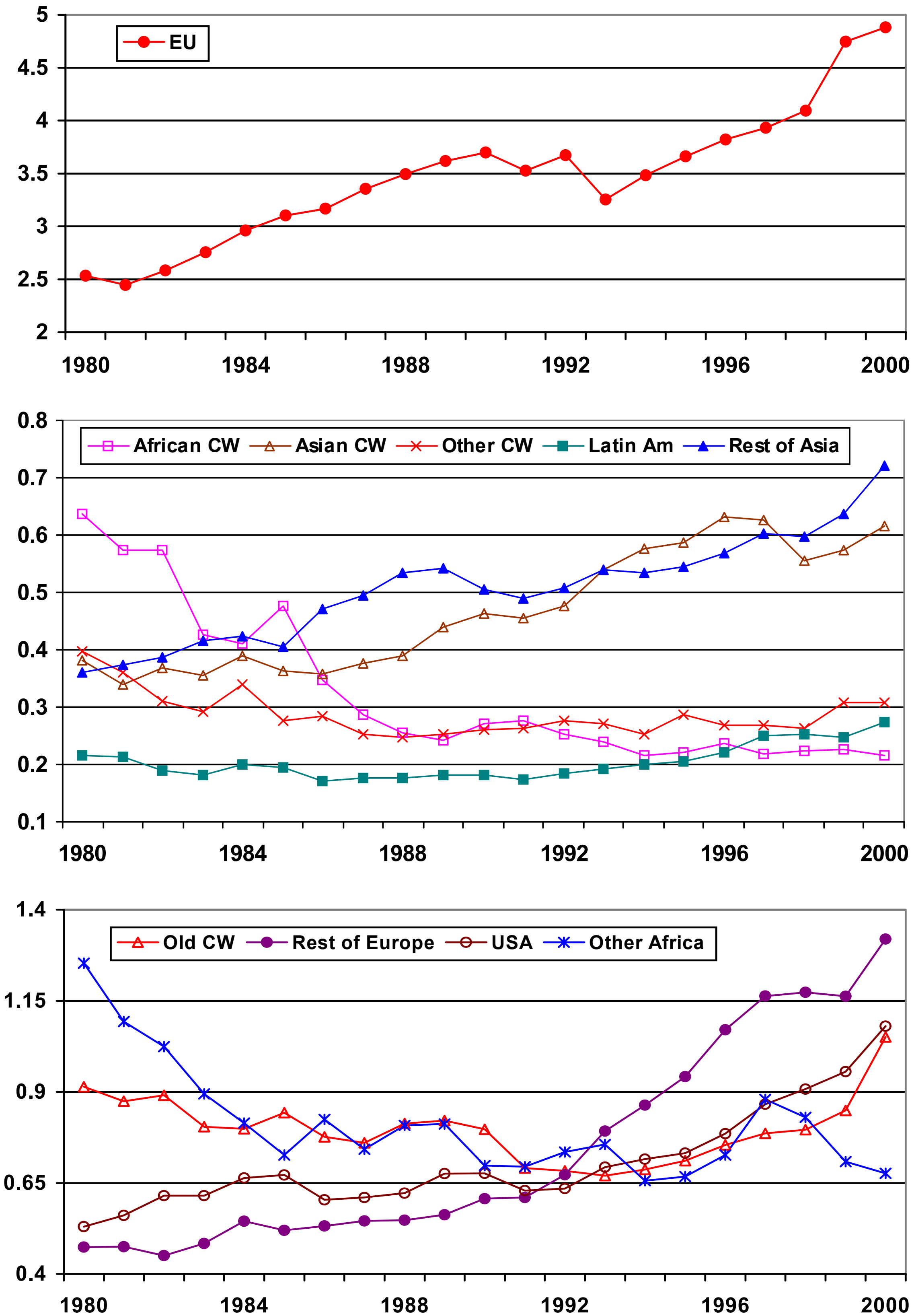
International trade may possibly capture several different factors. First, it is well known that at a global level the trade-income ratio is one indicator of globalisation, with trade volumes having grown much more rapidly than global output over the past two decades due to the greater international specialisation of many production processes and the gradual reduction of the costs of international trade, including the costs of transportation. Second, at a country-specific level bilateral trade is one indicator of the level of business linkages between economies and the economic opportunities that might be open to potential migrants, either directly through intra-company transfers or indirectly *via* networks of suppliers and customers. For both these reasons we might expect a positive linkage between trade and migration. But it is also possible that international labour movements through migration could, in some circumstances, be a substitute for trade, and so the overall effect of trade is ambiguous. Faini and Venturini (1993) and Schiff (1997) provide detailed discussions of the potential links between trade and migration.[17]

Bilateral trade data between the UK and the source locations were collected from successive issues of *Direction of Trade Statistics* published by the IMF. These data are in nominal US dollar terms. Data in constant prices are required for analysis, as it is unlikely that migration will be affected significantly by temporary fluctuations in nominal exchange rates that change the value of recorded trade but not the volume. A volume estimate was created using the UK aggregate merchandise export price deflator and aggregate non-oil merchandise import price deflator from the *Monthly Review of External Trade Statistics* converted into dollar terms. Imports into the UK from the Rest of Africa group, which includes trade with the Middle East, were deflated using a weighted average of the oil (75%) and non-oil (25%) import price deflators. The sum of the volume of exports and imports was then expressed as a ratio of the sum of UK and source location GDP in US dollars at 1995 prices. The resulting data series are shown in Figure 6.

Again there is marked heterogeneity in the patterns across different locations, although the trade ratio appears to be positively correlated with the migration rate in general. Rising migration from the EU, the Rest of Europe and the Rest of Asia has occurred at the same time as trade linkages have risen, whilst the level of trade and the migration rate from both African groups has declined over time. But there are exceptions, with little change in the level of trade with the Old Commonwealth over time, despite the rising migration rate, and a rise in the level of trade with the Asian Commonwealth which is not reflected in the migration rate.

[17] Care may also be required in dealing with cause and effect as migration may be an important determinant of trade patterns (Girma and Yu, 2002).

Figure 6. Bilateral Merchandise Trade (% of GDP, 1995 prices)



Migrant Stocks

Network effects provide another explanation of why the level of new migration might persist over time even if past migrants have changed relative factor endowments and helped to reduce expected wage differentials across locations (Carrington *et al.*, 1996). In effect, networks can be seen as both reducing the economic and social costs of migration as well as raising the initial probability of employment in the destination country. Such networks can arise in may different ways, such as family ties or through membership of the same ethnic group. The consequence of this is that the size of the existing migrant population in a host country may be an important positive determinant of the future migrant flow, as in the theoretical model outlined above.

In the empirical work we attempt to capture network effects by including the lagged migrant stock in the UK from the source location expressed as a proportion of the total current population in the source location. Again, we adopt this definition in order to ensure consistency with other studies, such as those reported by Hatton and Williamson (2002). *A priori* it is far from clear whether the relevant measure is the migrant stock alone – since all that may matter to a potential migrant is the number of friends and family in the host location – or the stock relative to the source population. The latter matters if migration decisions are also influenced in part by the level of contacts with friends and family in the source location that will be foregone if migration takes place.[18]

Unfortunately, the migrant stock data are likely to be subject to a considerable degree of measurement error.[19] This is because the data have to be interpolated over the sample period using just two data points, based on the stock of foreign-born citizens resident in the UK at the time of the 1981 and 1991 Censuses.[20] Further details are given in Appendix 1. There are three difficulties with this approach. The first is that the change in the migrant stocks between these two points differs considerably from the cumulated inflow because of deaths and outward migration. For instance, the cumulated migrant inflow into the UK from the other 14 European Union countries (including Ireland) over 1982-91 was 786,000. Yet the stock of EU-born citizens resident in the UK in 1991 was only 52,000 higher than in 1981. A second difficulty is that the Census data refer to foreign-born citizens, whilst the IPS migration data we use include returning UK citizens and are based on country of last residence. In theory at least, it is quite possible that none of the flow of migrants from a particular location were actually born in that

[18] It might also be argued that consideration should be given to the number of friends and family in the host location relative to those available in other potential host locations. Consistent time series census-based data on migrant stocks are however difficult to obtain and thus could not be used here.

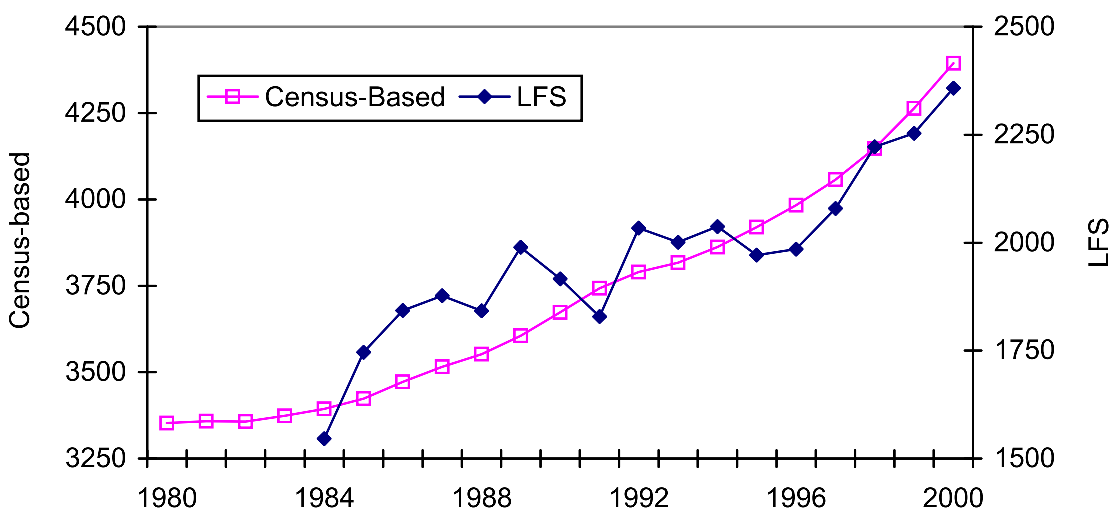
[19] The damaging statistical effects of this measurement error can be minimised in a number of ways, depending upon the particular assumption made about the form of the measurement error. Indeed, the ARDL modelling approach we suggest below can, under certain conditions (see Pesaran, 1997), obtain consistent estimates even when regressors are not strictly exogenous.

[20] At the time this work was undertaken, preliminary data from the 2001 Census had yet to be released.

location. UK citizens can account for up a third of the annual migrant inflow in some cases.[21] A third, and related difficulty is that some of the foreign-born population will have entered as asylum seekers and visitor switchers rather than as IPS migrants. There is clearly considerable potential for errors to arise because of all these differences in definition. This should be borne in mind in interpreting the results below. It is perhaps not too surprising that we find that the sign, magnitude and significance of the coefficients obtained on the migrant stock term are very sensitive to the model specification and estimation technique being used.

An alternative source of data from 1984 is the *Labour Force Survey* (Dobson *et al*, 2001, Table 11.1). However this differs from the Census in that it refers to foreign nationals by citizenship rather than by country of birth. In 1991 for instance, the stock of foreign nationals is estimated to be just over 1.8 million, whereas the Census data for that year suggest that the stock of foreign-born citizens was more than twice this number. The LFS is also a sample survey rather than a comprehensive count of the whole population and has been subject to definitional changes over time. For all these reasons we did not use it as a source of information. Nonetheless, as can be seen from Figure 7, the upward trend over time in our aggregate interpolated Census-based migrant stock is broadly similar to that in the annual LFS data, although the latter is more variable, especially in the 1980s.

Figure 7. Estimates of the UK Foreign Population (000s)



Other Factors

We also include the lagged migration rate, allowing for persistence in the level of migration from different locations. Incorporating this variable makes it possible to examine a richer array of behaviour. In particular, it is now feasible to test whether the short-run impact of a change in one of the driving variables affects migration from different locations in different ways, even if

[21] In 2000, 28½ per cent of all IPS immigrants had British citizenship (ONS, 2002, Table 3.1). Over a third of migrants from the EU and the Old Commonwealth had British citizenship.

the responses are identical in the long-run. This distinction is often ignored when modelling the determinants of migration.

In estimation we also allow for country-specific fixed effects. These will pick-up all factors that vary across source regions but do not vary over time. Possible examples include distance from the UK, language and any effects that stem from membership of the same supranational groups as the UK, such as the Commonwealth or the EU.

It is clear that there are many reasons to expect that the influence of many of the potential determinants of migration will vary over time. Variations may be due to changes in government immigration policy,[22] or spatial restructuring, such as the accession of economies into the European Union or demographic changes. The globalisation process may also matter, with the improved speed and detail of information flows, reducing the costs of migrating, and the changing location of production affecting the geographical mobility of the high-skilled. Some of these factors may be captured through other factors included in estimation, notably the trade-GDP ratio. But we also report some tests of stability to assess whether structural changes have occurred in recent years.

IV. Econometric Issues

In this section we highlight a number of important methodological issues that need to be considered when seeking to model migration using a panel data set of migration inflows into a country across both time and groups of countries. Panel data sets have been exploited before in addressing the determinants of migration, although not for the UK. Recent examples include Karemera *et al.* (2000), Hatton and Williamson (2001), Clark *et al.* (2002) and Rotte and Vogler (1998). But relatively little attention appears to have been given to testing the reported models against other less restrictive benchmark models. As we show in our empirical work, this affects the reliability of the resulting estimates and the conclusions that can be drawn.

Traditional panel data techniques, such as fixed or random effects estimators, used in most empirical studies of migration constrain all coefficients aside from the intercept to be the same across groups. This implies that migration inflows into, say, the UK from different countries are determined in the same way by the same set of explanatory variables.[23] No allowance is made, for example, for the fact that either the impact or adjustment to changes in, say, income

[22] The UK government does not operate formal quota systems that cap the annual number of potential immigrants, and so we do not include a specific indicator for the policy stance in the estimation model. Such indicators can be particularly important for understanding migrant inflows in other countries; see, for example, Clark et al (2002). It is possible that the interpretation and enforcement of regulations such as work permit schemes may vary either across time or across occupations where there are perceived labour market shortages within the UK, but there are no official data that allow this to be measured easily.

[23] Special empirical techniques may be needed if schemes are in place which impose a ceiling on the total level of migration each year, such as annual quotas on the numbers of migrants admitted. This should not be a problem for the UK.

differentials may vary across countries according to how developed they are. Furthermore, little account has previously been taken of possible dynamic effects. Studies that have sought to explain the determinants of migration using panel data sets have, in general, estimated static relationships (pooled OLS, fixed or random effects estimators) where migration today is not affected directly by migration yesterday. This is inconsistent *inter alia* with Hatton's (1995) model of migration where migration at time *t* depends on migration at time *t-1*. We adopt an alternative approach. In contrast to the traditional approach of starting with a restricted model, such as fixed or random effects, we start with a relatively unrestricted model estimated separately for each group. The restrictions implied by the more restricted estimators are then nested as special cases of our model and, rather than just being imposed, can be statistically tested.

The starting point of an econometric analysis seeking to explain the determinants of migration across both groups and time should be, in part, determined by the nature of the panel data set. In particular, the number of groups (N) and the number of time-series observations (T) should be considered. When T is reasonably large separate time-series regressions can be estimated for each group. Pesaran and Shin (1999) propose an approach based on the use of autoregressive distributed lag (*ARDL*) models that is appropriate for the examination of long-run relationships regardless of the time series properties of the individual regressors. In contrast to tests for cointegration, 'pre-testing' to establish the order of integration of the regressors is not required. This is attractive given that unit root tests are known widely to have low power.[24]

A second consideration is the size of N. When N is small relative to T then, following Zellner (1962), the Seemingly Unrelated Regression Equation (*SURE*) approach can be used. The attraction of *SURE* is that it allows the contemporaneous error covariances across groups to be estimated. However, *SURE* is inappropriate when N is quite large relative to T as in our data set. We do, however, in the empirical application below, for completeness, consider SURE estimation but gain degrees of freedom by imposing some other restrictions on the model estimated.

[24] Despite this, we did examine the properties of the regressors using two types of panel unit root test. The first type of tests have as their null hypothesis that all the time-series in the panel are I(1), that is they have a unit root and no fixed mean or deterministic linear trend to which they tend to revert over time, (Im, Pesaran and Shin, 2003; hereafter IPS). The second type of tests have as their null hypothesis that all the time-series in the panel are stationary, I(0), albeit perhaps about a deterministic linear trend (Hadri, 2000). It is important to note when using panel unit root tests that if the null hypothesis is rejected this does not imply that the null hypothesis is rejected for all groups, i.e. there may be a mixture of I(0) and I(1) variables in the panel. Unsurprisingly, given the relatively small time series dimension of our data set (1980-2000), the results were found to be inconclusive and varied according to the test undertaken. Nonetheless, there is some evidence from at least one of the IPS tests that the main variables of interest have compatible time series properties. The two variables common to all panel members – income in the UK relative to the EU4 and the UK unemployment rate - both appear to be stationary over the sample period on the basis of univariate Dickey-Fuller unit root tests.

The ARDL modelling approach

Our starting point is the estimation of separate *ARDL* models for each group in the panel. Suppose that given data on time periods *t=1,…,T* and groups, *i=1,…,N*, we wish to estimate an *ARDL(p,q,q,…,q)* model:

$$\ln(m_{it}/P_{it}) = \sum_{j=1}^{p} \lambda_{ij} \ln(m_{i,t-j}/P_{i,t-j}) + \sum_{j=0}^{q} \partial'_{ij} x_{i,t-j} + \mu_i + \varepsilon_{it} \qquad [12]$$

where $x_{it} = [\ln(S_{i,t-1}/P_{i,t-1}), D_{it}, \ln(Y_{it}), \ln(TR_{it}), \ln(YR_t), U_t]'$ and *i* denotes the location group, *m* the migration flow, P the source location population (so that in the context of Section III.1, M=*m*/P), S the migrant stock, D the demographic share, Y and YR are per capita income in the UK relative to the source location and the EU4, respectively, TR is the bilateral trade-GDP ratio and U is the UK unemployment rate. Group-specific fixed effects are denoted as μ_i, these pick up any factors that differ across groups but not across time. The coefficients on the lagged dependent variable λ_{ij} are scalars and δ_{ij} is the vector of coefficients on the explanatory variables x_{it}. This equation can be augmented further with additional dynamic terms, such as income growth, as well as outlier dummies and time dummies as required.

The disturbances, ε_{it}, are assumed to be independently and identically distributed across *i* and *t*, with zero means and variances $\sigma^2_i>0$ and finite fourth moments.[25] They are also assumed to be distributed independently of the regressors x_{it}. Importantly, however, the ARDL approach can be rendered valid even when the regressors are not strictly exogenous. This is important given the potential endogeneity of many determinants of migration.[26] On the assumption that the x_{it} follow finite order auto-regressive representations, the inclusion of additional lags of the regressors in the ARDL model in fact renders it possible to obtain consistent estimates (Pesaran, 1997).

Equation [12] can be re-written in error-correction form:

$$\Delta\ln(m_{it}/P_{it}) = \phi_i \ln(m_{i,t-1}/P_{i,t-1}) + \beta_i' x_{it} + \sum_{j=1}^{p-1} \lambda^*_{ij} \Delta\ln(m_{i,t-j}/P_{i,t-j}) + \sum_{j=0}^{q-1} \delta^{*'}_{ij} \Delta x_{i,tj} + \mu_i + \varepsilon_{it} \qquad [13]$$

[25] Attempts have been made within the ARDL approach to allow for non-zero contemporaneous covariances between groups through, for example, cross-sectional de-meaning (Pesaran, Shin and Smith, 1999). Pesaran, Smith and Akiyama (1999) provide a fuller discussion and illustration of de-meaning. We do not, however, follow such a route here; instead we consider SURE.

[26] Traditional panel estimators, such as fixed and random effects, when estimated *via* (Ordinary or Generalised) Least Squares assume that the set of explanatory variables are exogenous and not correlated with the disturbance term. If this assumption is violated then the estimators are no longer consistent. *A priori* one may expect certain explanatory variables not to be exogenous. For example, the wage ratio, or the share of young people in the population, may be endogenous because both depend upon the level of migration. To combat these potential problems, and ensure estimators are consistent, instrumental variable (IV) estimators have been considered occasionally; see for example Hatton and Williamson (2001).

where: $\phi_i = -(1 - \sum_{j=1}^{p} \lambda_{ij})$, $\beta_i = \sum_{j=0}^{q} \delta_{ij}$, $\lambda_{ij}^{*} = -\sum_{m=j+1}^{p} \lambda_{im}$, *j=1,2,…,p-1*

and: $\delta_{ij}^{*} = -\sum_{m=j+1}^{q} \delta_{im}$, *j=1,2,…,(q-1).*

Assuming the *ARDL(p,q,q,…,q)* model is stable, in the sense that all values of the scalar *z* satisfying $\sum_{j=1}^{p} \lambda_{ij} z^{j} = 1$, *i=1,2,…,N*, lie outside the unit circle, we know $\phi_i < 0$. This ensures that a long-run relationship exists between $\ln(m_{it}/P_{it})$ and x_{it}. This long-run relationship is defined as:

$$\ln(m_{it}/P_{it}) = -(\beta_i'/\phi_i)x_{it} + \eta_{it} \qquad [14]$$

for each *i=1,2,…,N*, where η_{it} is a stationary process. Let $\Theta_i = -(\beta_i'/\phi_i)$. The stability assumption also ensures that the order of integration of $\ln(m_{it}/P_{it})$ is at most equal to that of x_{it}.

The long-run relationship [14] can then be substituted back into equation [13] to yield:

$$\Delta \ln(m_{it}/P_{it}) = \phi_i\left[\ln(m_{i,t-1}/P_{i,t-1}) - \Theta_i' x_{it}\right] + \sum_{j=1}^{p-1} \lambda_{ij}^{*} \Delta \ln(m_{i,t-j}/P_{i,t-j}) + \sum_{j=0}^{q-1} \delta_{ij}^{*\prime} \Delta x_{i,t-j} + \mu_i + \varepsilon_{it} \qquad [15]$$

where $\left[\ln(m_{i,t-1}/P_{i,t-1}) - \Theta_i' x_{it}\right]$ is the error correction component and ϕ_i indicates the speed of adjustment to the long-run equilibrium relationship.[27] The interpretation of equation [15] is that the migration rate changes over time in order to eliminate any differences between the past migration rate and the current desired rate given by the linear combination of the long-run economic and demographic driving factors. When ϕ_i does not differ significantly from –1, adjustment is instantaneous. If it is insignificantly different from zero, the migration rate does not adjust towards any particular long-run level generated by a fixed linear combination of the variables included in the equilibrium-correction term.

If interest lies in the average, across groups, long-run relationship between the dependent variable and the regressors, consistent estimates can be obtained using the Mean Group (MG) estimator (Pesaran and Smith, 1995). The MG estimator is given by the mean across countries of the estimated long-run and adjustment coefficients:

$$\hat{\Theta}_{MG} = N^{-1} \sum_{i=1}^{N} \hat{\Theta}_i \qquad [16]$$

[27] Equation [15] involves a partial adjustment type error correction. It can be easily re-written to incorporate a lagged error-correction component $\left[\ln(m_{i,t-1}/P_{i,t-1}) - \Theta_i' x_{i,t-1}\right]$ (Pesaran, Smith and Akiyama, 1999).

where $\hat{\Theta}_i$ is the vector of long-run coefficients obtained by estimating the ARDL for each group by OLS. The standard error of the MG estimator can be consistently estimated (Pesaran, Smith and Im, 1996) as:

$$\hat{V}(\hat{\Theta}_{MG}) = (N[N-1])^{-1} \sum_{i=1}^{N} (\hat{\Theta}_i - \hat{\Theta}_{MG})(\hat{\Theta}_i - \hat{\Theta}_{MG})' \qquad [17]$$

Estimating ARDL models separately across groups does not impose any homogeneity restrictions across groups. However, an important distinction between alternative panel data estimators is the extent to which they impose homogeneity across groups with respect to variances, short or long-run regression slope coefficients and intercepts. If the coefficients are homogeneous across groups it is more efficient to pool the data and impose the homogeneity restriction rather than estimate unrestricted *ARDL* models for each of the *N* groups.

In fact, *a priori* (economic) reasoning often suggests that long-run relationships should be homogenous across groups, with the speed of adjustment to this long-run and the short-run coefficients being the main differences across groups. This leads us to the next least restrictive estimator, the Pooled Mean Group estimator.

The Pooled Mean Group (PMG) estimator

The PMG estimator imposes equality of one or more of the long-run coefficients explicitly, but allows dynamics and error variances to differ across groups. Pesaran, Shin and Smith (1999) propose a Maximum Likelihood estimator that imposes this equality restriction and will lead to lower standard errors of the estimates if the restriction is valid. The PMG estimator is also likely to be less sensitive to outliers than the MG estimator.[28] The restriction that $\Theta_i = \Theta$ can be tested statistically using either a Likelihood Ratio test or a Hausman (1978) test; see Pesaran, Shin and Smith (1999) and Pesaran, Smith and Im (1996) for details.

Traditional panel data estimators: fixed effects

Common panel data estimators, like the fixed effects estimator, are restricted forms of the general ARDL model considered in equation [12]. The fixed effects model, in dynamic form, is a restricted version of the ARDL model in equation [12], and is given by:

$$\ln(m_{it}/P_{it}) = \sum_{j=1}^{p} \lambda_j \ln(m_{i,t-j}/P_{i,t-j}) + \sum_{j=0}^{q} \partial'_j x_{i,t-j} + \mu_i + \varepsilon_{it} \qquad [18]$$

where $\sigma^2_i = \sigma^2$. So the fixed effect estimator imposes equality of all slope coefficients and all error variances, and only allows the intercept to differ across countries. If there is in fact

[28] The MG estimator is consistent for large N and T. But for small T there will be lagged dependent variable bias that causes downward bias. Large N will not help eradicate this problem. However, (see Pesaran, Shin and Smith, 1999), for the PMG estimators this downward bias may be offset by the upward heterogeneity biases discussed in Pesaran and Smith (1995). It is of interest that Pesaran and Zhao (1999) have proposed a bias corrected MG estimator.

heterogeneity across groups the fixed effects estimator will be biased. The restrictions implied by the fixed effects estimator are testable.

Equation [18], like equation [12], can be equivalently re-written in a non-linear form so that the dependent variable becomes the change in the migration rate:

$$\Delta \ln(m_{it}/P_{it}) = \phi_i\left[\ln(m_{i,t-1}/P_{i,t-1}) - \Theta' x_{it}\right] + \sum_{j=1}^{p-1}\lambda_j^{*}\Delta \ln(m_{i,t-j}/P_{i,t-j}) + \sum_{j=0}^{q-1}\delta_j^{*'}\Delta x_{i,t-j} + \mu_i + \varepsilon_{it} \qquad [19]$$

and the right hand side of equation [19] contains an equilibrium-correction term with long-run parameters that can be compared with equation [20] below.

Many empirical studies of migration just use a static version of equation [18]:

$$\ln(m_{it}/P_{it}) = \partial' x_{it} + \mu_i + \varepsilon_{it} \qquad [20]$$

The fixed effects estimator can provide consistent estimates of the average, across groups, effect of the exogenous variables on the dependent variable in a static model, even in the presence of heterogeneity (of the form where coefficients differ randomly across groups). But it does not do this (even as $T \rightarrow \infty$) in the case of a dynamic model (Pesaran and Smith, 1995). In this instance the fixed effects estimator of the long-run coefficients is inconsistent when coefficients differ across groups.

V. Empirical Results

This section contains the main empirical results. We begin by estimating a standard fixed effects model that does not allow for heterogeneity across groups. The implicit restrictions are shown to be rejected by the data. We then turn to estimators that are robust to the existence of heterogeneous behaviour. In all cases our model is one for the migration rate, defined as the annual flow of inward migrants from each of the 10 source locations as a ratio of the total population of that location. We have a consistent data set for the years 1980-2000, with the sample period used being determined by the number of lagged terms included in each of the different models we estimate.

V.1 Conventional Panel Data Results

In order to move from an unrestricted system of separate equations, with different parameters on each explanatory variable in each source location, to a model with a single set of parameters common to all source locations two sets of restrictions are required. These are that each of the independently estimated equations has a common error variance, and that each has common parameters. If these restrictions are not rejected statistically, then it is more efficient to estimate a single parsimonious model with the restrictions imposed. Most, if not all, of the existing applied time series papers on international migration estimate a restricted model of this kind, but without testing the implicit restrictions. If the underlying determinants of migration from each source location are heterogeneous, then it is less likely that the restrictions required to return to a single relationship common to all locations will be accepted by the data, and estimates from the restricted model may be biased.

The outcome from estimating a number of different models with location specific fixed effects but common parameters on all behavioural variables is reported in Table 2. All the equations have the form of either equation [19] or [20].[29] It is clear that the results cannot be regarded as statistically reliable, even though they appear theoretically plausible in many instances. In each specification the implicit restrictions that have been imposed on an equation with different slope parameters for each of the 10 groups (but a common variance) are jointly rejected by the data. The imposition of a common variance across all groups is also rejected in these models, and the significant first order serial correlation test statistics are indicative of an underlying mis-specification.

Column [1] contains the results from a simple panel regression with fixed effects for each group, but no dynamic terms (see equation [20] above). Three factors appear to be important determinants of migration - the trade ratio, UK income relative to that of the EU4 and the current stock of foreign-born UK residents. Adding in lagged migration and income growth terms improves the overall fit of the model, but as can be seen from the second column (estimation of equation [19]) makes little difference to the significance of the long-run parameters, with the exception of the migrant stock term, which now becomes insignificant. The coefficients on the demographic share and unemployment rate terms are now in line with prior expectations, but remain insignificant. The income growth term appears to be well determined, as is the coefficient on the equilibrium-correction term, which suggests that a statistically valid long-run relationship may exist.

In the first two columns of Table 2 all the time-varying factors common to each migrant group are being captured through the terms in relative UK-EU4 incomes and the UK unemployment rate.[30] An alternative approach is to replace these two variables with (T-1) time dummies, one for each year of the estimation period.[31] The coefficients on the remaining regressors are shown in the third column of Table 2. Again, only the trade variable is individually significant, although the equilibrium-correction term as a whole still appears to be well determined.

[29] We consider specifications with up to one lag, i.e. p=q=1 in [18]. For the conventional panel data estimators in Table 2 we include only a dynamic term in income growth in the UK relative to the source country (ΔY_{it}). A more general analysis is allowed for when we turn to ARDL models in which all variables are allowed to enter with as many lags as supported by standard statistical criteria.

[30] These are the two variables that vary over time, but not across migrant groups. Hence they can pick up time-varying effects that are common to all groups.

[31] One time dummy is excluded to avoid perfect colinearity with the group-specific fixed effects. The reported parameters are not dependent on which one is excluded.

In the fourth column of Table 2 the results from augmenting the specification in the second column with a separate deterministic time trend for each group in the panel are reported. Whilst this results in an improvement in the overall fit of the equation, the implicit restrictions required to pool the separate groups continue to be rejected by the data. It also generates marked changes in some of the coefficients. Most notably, the coefficient on the migrant stock term has become negative and significant. The coefficient on the unemployment rate has become significant whilst the size and significance of the trade variable has dropped.

Table 2. Basic Panel Results

Dependent variable: [1]: Migration Rate; [2]-[4]: Δ Migration Rate

Sample Period: 1981-2000

	[1]	[2]	[3]	[4]
Long-run parameters				
Migrant stock	0.6197 (2.6)	0.4168 (1.0)	0.4363 (1.1)	-1.8326 (3.0)
Relative income UK/Source	-0.0485 (0.2)	-0.1119 (0.4)	0.0288 (0.1)	0.3478 (0.6)
Bilateral Trade	0.5322 (4.7)	0.5952 (3.6)	0.5989 (3.8)	0.2424 (1.2)
Demographic share	-0.0056 (0.3)	0.0221 (0.8)	0.0007 (0.0)	0.0705 (1.5)
Relative income UK/EU4	1.4132 (2.0)	2.3714 (2.2)		1.9367 (2.4)
UK unemployment rate	0.0094 (0.8)	-0.0064 (0.4)		-0.0339 (2.4)
Dynamic parameters				
Income growth UK/Source		2.2401 (2.8)	2.1422 (2.4)	2.6200 (3.3)
Equilibrium-correction		-0.5899 (8.2)	-0.5883 (7.4)	-0.8110 (9.1)
Log Likelihood	-24.55	-0.77	12.21	22.16
Standard Error	0.2852	0.2546	0.2506	0.2336
Serial Correlation	LM(1)=19.4*	LM(1)=9.80*	LM(1)=13.05*	LM(1)=21.22*
Test of Pooling Parameters	LR(54)=181.2*	LR(72)=171.3*	LR(54)=141.4*	LR(72)=196.5*
Test of Pooling Variances	LR(9)=45.4*	LR(9)=47.1*	n.a.	LR(9)=59.0*

Notes: (i) all variables in logarithms apart from the demographic share and the unemployment rate.
(ii) t-statistics in parentheses.
(iii) An * denotes a diagnostic test statistic significant at the 5% level.

Although the ten unrestricted trend terms are jointly significant on the basis of a chi-squared likelihood ratio test [denoted LR(10)=45.9], this does not mean that a simple time series model of migration using only an intercept and a time trend will suffice, as the set of eight stochastic explanatory variables used in Table 2 are also jointly significant [LR(8)=49.45]. This indicates that a model with behavioural elements is to be preferred to a purely statistical one, and suggests that simple projections of future migration on the basis of past trends might well be misleading.[32]

The results in Table 2 suggest that incomes in the UK relative to the EU4 are an important determinant of inward migration. To explore the sensitivity of this finding to the choice of comparator group, we also estimated versions of the specification in column [2] of Table 1 with UK per capita incomes relative to a weighted average of the other economies in the G7. When this variable was used in place of the UK-EU4 measure, it also had a significant positive coefficient, but the fit of the equation was not as good as the original equation. This suggests that relative incomes are an important determinant of migration choice, but also suggests that the EU4 may be closer to the 'true', unobservable comparator group than the G7. This was confirmed when the UK-G7 relative income variable was added to the specification in column [2] of Table 2. The resulting coefficient was incorrectly signed and was insignificant with a t-statistic of just 0.1.

V.2 Models Allowing for Heterogeneity

The diagnostic tests reported at the foot of Table 2 suggest that there is considerable parameter heterogeneity across the respective groups, suggesting the need to use alternative estimation techniques that are robust to the presence of heterogeneity in order to obtain consistent and efficient estimates of the key long-run behavioural parameters. We have experimented with two such techniques, estimating SURE and PMG models.

The SURE approach differs from PMG by allowing for non-zero off-diagonal elements in the variance-covariance matrix of the residuals from the individual groups. This makes it potentially more powerful, but there are far fewer degrees of freedom for any given model than in the PMG approach because of the need to allow for the covariances of errors across the separate equations. This would not be problem if for panels with a large time series dimension, but in our case with T=20 observations and N=10 panel members it could easily do so. We therefore allowed dynamics to enter only through the lagged migration rate and the income growth term in each individual group equation in the SURE model.

In contrast, in the PMG approach we were able to begin with an unrestricted first-order ARDL model, using three different model selection criteria in turn to determine the specification used

[32] Although it is difficult to provide a full explanation of the changes that result in column [4], the most likely explanation is that the migration rate, which panel unit root tests showed to be stationary, contains a mixture of trend and mean stationary variables. The deterministic trend element in some groups can be captured either by the deterministic trend component of the other explanatory variables, as in column [2], or by a separate trend which is also highly collinear with many of the remaining explanatory variables as in column [4]. The latter results in a model with a better statistical fit, but is not as easy to interpret. An alternative strategy, which we investigate below, is to see whether different estimation techniques that allow for heterogeneous effects across groups generate statistically acceptable, and readily interpretable results.

for each individual group.[33] These different equations were then estimated jointly with a common long-run relationship imposed. In each case we found that the restrictions required to impose a common long-run relationship in the set of unrestricted equations could not be rejected on the basis of a Hausman test.

Table 3. SUR and PMG Long-Run Pooled Parameters

Sample Period: 1982-2000

	[1]	[2]	[3]	[4]
	SUR	**PMG: SBC**	**PMG: AIC**	**PMG: HQ**
Migrant stock	0.470 (1.8)	1.167 (5.4)	0.508 (1.9)	0.671 (2.5)
Relative income UK/Source	-0.144 (0.6)	0.514 (2.7)	0.776 (4.1)	0.738 (3.8)
Bilateral Trade	0.632 (4.6)	0.356 (3.0)	0.049 (0.4)	0.123 (1.0)
Demographic share	0.055 (2.7)	0.117 (6.8)	0.028 (2.1)	0.033 (2.5)
Relative income UK/EU4	2.441 (4.8)	1.776 (3.3)	2.488 (2.7)	2.314 (2.7)
UK unemployment rate	-0.004 (0.3)	-0.015 (1.3)	-0.011 (0.7)	-0.005 (0.3)
Hausman Test of Pooling		ChiSq(6)=12.48	ChiSq(6)=6.58	ChiSq(6)=4.71

Notes: see Table 2. SBC – Schwarz Bayesian Criterion; AIC – Akaike Information Criterion; HQ – Hannan-Quinn Information Criterion.

The SURE and PMG estimates are shown in Table 3. The results from the SURE approach are similar to those from the fixed effects panel (column [2] in Table 2), with the exception of that for the demographic share. This is now significant and has a coefficient roughly double that found in the fixed effect model.

Moving to the PMG results in columns [2]-[4] of Table 3, we found that it was not possible to reject the restrictions required to impose a common long-run relationship across all 10 source locations in any of the three different dynamic specifications obtained using the separate model selection criteria. However it is clear that the parameters of the resulting relationship are sensitive to the choice of dynamic structure resulting from the different model selection criteria. The most noticeable differences are between the findings from the dynamic ARDL model selected using the SBC (column [2]) and that obtained using the other two criteria. The former has much larger effects from the migrant stock and the trade ratio than the other two.

[33] We estimated all combinations of ARDL models with up to one lag (with only 20 years of annual data it is not feasible to include higher order lags) and used the Akaike Information Criterion (AIC), the Schwarz Bayesian Criterion (SBC) and the Hannan-Quinn (HQ) criterion to select the preferred model.

All three PMG estimates differ from the fixed effects panel and the SURE estimates by indicating significant positive effects on migration from per capita income in the UK relative to that in the source location, which is consistent with theoretical priors. However they continue to find little evidence of a significant effect from the UK unemployment rate. The demographic share is significant in all three PMG estimates, but there is considerable variation in the long-run coefficient. The one robust finding which seems to emerge from all the different estimation strategies is that migration into the UK is positively related to per capita incomes in the UK relative to those in the EU4.

Given that we have stressed the existence of considerable heterogeneity in the factors driving migration from different source groups, it might be regarded as surprising that we cannot reject the imposition of common long-run coefficients in the PMG approach. There are three main reasons why this result arises. First, the specification of the individual group equations in the PMG approach is much richer than in the basic panel models in Tables 1 and 2, with many different dynamic terms being included in the former. Given the relatively short time span of the sample period it may well be that the observed heterogeneity in the conventional panel approach arises from different speeds of adjustment to changes in the underlying factors driving migration. For example, it may be the case that, other things being equal, migration from each location responds to a change in relative incomes in an identical fashion but does so at different speeds. The PMG approach allows for this, but the conventional fixed effects panel does not, and may generate biased parameters as a result. It is worth noting that the long-run parameters from the SURE estimates, which allow for heterogeneous dynamics to a limited extent only, are much closer to those from the standard panel estimators in Table 2.

The second explanation for the common long-run parameters in the PMG approach stems from the panel approach. It is not necessarily the case that the identified long-run relationship is valid for every group in the panel, but if it is valid for 8 or 9 groups, and inappropriate for just 1 or 2 groups, then it is not unreasonable to conclude that the balance of the evidence across the panel as a whole points to it being important.

Third, it should be remembered that the Hausman test uses a variance-covariance matrix for the long-run parameters obtained by the Mean Group (MG) method, which averages the parameters obtained by estimating the model independently for each panel member. In many cases the long-run parameters in these individual equations are estimated imprecisely, so that the MG parameters have large standard errors, making it less likely that an alternative set of parameter estimates (such as the common long-run parameters) will be (jointly) statistically different from them.

V.3 Implications for Migration Behaviour

With three alternative statistically valid long-run relationships obtained from the PMG approach we are left with the choice of which to use for policy analysis and further inference. Given the obvious sensitivity of the results to the specification and modelling technique used, we chose to take a simple unweighted average of the three separate estimates. This generates the following underlying long-run relationship:

$$\ln(m_{it} / P_{it}) = 0.782 \ln(S_{i,t-1} / P_{i,t-1}) + 0.059 D_{i,t} + 0.676 \ln(Y_{it}) + 0.176 \ln(TR_{it}) + 2.193 \ln(YR_{t-1}) - 0.010 U_{t-1} \quad [21]$$

The long-run parameters imply that a 1% rise in the migrant stock will ultimately raise the level of the migration rate by 0.78%. Evaluated using the data for 2000, this implies that raising the migrant stock by 1000 will increase the annual inflow of IPS migration by 65 people. Evaluated at the sample mean, the increase would be 55 people. Both estimates are somewhat higher than the estimates obtained by Hatton and Williamson (2002), although some of their figures are for net rather than gross immigration.

The coefficient on the demographic share variable implies that a rise of 1 percentage point in the share of the population aged 15-29 in all source locations will ultimately raise the total inflow of migrants into the UK by 5.9%. At 2000 levels this implies a rise of between 21-22,000 in the annual number of immigrants. Of course it needs to be remembered that a change of this magnitude in the demographic share will occur only slowly, and has not been observed at all in some of the source locations throughout the last two decades (see Figure 5).

A sustained rise of 1% in per capita income in the UK relative to the source location raises the annual level of migration by 0.676%. Aggregating across countries and using 2000 values, this implies an increase in the total IPS inflow of approximately 2,500 per annum. A sustained rise of 1% in UK incomes relative to the EU4 will raise inflows by around 8,100 per annum. It should be noted that these estimates refer to sustained changes in relative incomes. If ideas and technologies gradually spread across international borders, there is no necessary reason why UK relative incomes should be able to change permanently in this way.

A rise of 1% in the volume of bilateral trade between the UK and the source location will raise the annual inflow of migrants from that location by 0.176%, other things being equal. Finally, a rise of 1 percentage point in the UK unemployment rate will reduce the annual flow of migrants by around 1%.

A clearer idea of the long-term importance of each of these factors in changing the level of inward migration over time can be obtained by using the estimated equations to calculate the effects of actual changes in the independent variables on the level of inward migration. From [21] the growth of migration between time *t-n* and time *t* can be expressed as:[34]

[34] Strictly, this is only an approximation as the calculations could also be made with models allowing for non-instantaneous dynamic adjustment. In this case account also needs to be taken of the presence of the lagged dependent variable which embodies past movements in the independent variables.

$$\left[\frac{m_{i,t}}{m_{i,t-n}}\right]=\left[\frac{P_{i,t}}{P_{i,t-n}}\right]\left[\frac{(S_{i,t-1}/P_{i,t-1})}{(S_{i,t-n-1}/P_{i,t-n-1})}\right]^{0.782}\left[\frac{TR_{i,t}}{TR_{i,t-n}}\right]^{0.176}\left[\frac{Y_{i,t}}{Y_{i,t-n}}\right]^{0.676}\left[\frac{YR_{t}}{YR_{t-n}}\right]^{2.193}\left[\frac{\exp(0.059D_{i,t}-0.01U_{t-1})}{\exp(0.059D_{i,t-n}-0.01U_{t-n-1})}\right]$$

[22]

This formula can be used to evaluate the contribution of movements in each of the explanatory demographic and economic variables to changes in migration from a given source region. These can then be combined in a weighted sum to calculate the overall effect of each of these variables on total migration by noting that:

$$\left[\frac{\sum_i m_{i,t}}{\sum_i m_{i,t-n}}\right]=\left[\frac{m_{1,t-n}}{\sum_i m_{i,t-n}}\right]\left[\frac{m_{1,t}}{m_{1,t-n}}\right]+\left[\frac{m_{2,t-n}}{\sum_i m_{i,t-n}}\right]\left[\frac{m_{2,t}}{m_{2,t-n}}\right]+\ldots \qquad [23]$$

In using this approach it should be remembered that the choice of years over which to carry out the calculation is arbitrary and the results may be sensitive to the choice of the base year, as the weights in [23] are time-varying. Here we calculate the contribution of each factor to the overall change in the average annual level of IPS migration (including migration from Ireland) between 1988-90 and 1998-2000. In the latter period average annual inflows were 83,400 per annum higher than in the former period, equivalent to a proportionate rise of nearly 30%. Table 4 reports the contribution of each of the main explanatory variables. For the friends and family contribution we use the change in the migrant stock as a proportion of the source region population.

Table 4. Accounting For Changes In Annual Migration Over 1998-2000 and 1988-90

Source Region Population	36,100
Friends and Family	22,500
Source Region Demographic Share	-31,300
UK –Source Relative Per Capita Incomes	36,800
UK-EU4 Relative Per Capita Incomes	8,700
UK Unemployment Rate	8,200
Bilateral Trade	10,500
Residual	-8,500
Total	83,400

The overall change in migration over the decade to 1998-2000 is more than accounted for by the positive impact of population growth in the source regions, the continuing pull effect from the rise in the migrant stock and the rise in per capita incomes in the UK relative to those in other countries. The growth in bilateral trade also made an important contribution, accounting for one-eighth of the overall change, reflecting the strong growth in trade over this period despite the relatively low estimated coefficient. In contrast, demographic changes over this period actually made an overall negative contribution to total migration. This was because the growth in the 15-29 age group share occurred primarily in regions from which there was a relatively low level of migration in the base years. These changes were more than outweighed by the effect of the fall in the 15-29 age group ratio in the US and the EU over this period, both of which are relatively important source locations. Finally there is a small negative residual factor, reflecting the fact that the actual rise in migration levels over the decade to 1998-2000 was marginally lower than suggested by the evolution of the long-run driving factors over this period. As we show in the next section, this may be because actual migration adjusts towards its long-run equilibrium level with a lag.

V.4 Modelling Dynamic Adjustment

It is also of interest to ask whether the speed with which migration from different locations responds to changes in the main driving factors differs significantly across source locations. *A priori*, the expectation must be that it does, given the evidence of heterogeneity in the basic panel results and the sensitivity of the PMG estimates to the underlying dynamic structure of the estimated model for each location.

To investigate this, we initially estimated a separate equation for each of the ten locations in which the change in the migration rate was regressed on income growth in the UK relative to that in the source, and the residuals from [21]. In effect the latter perform the same role as the equilibrium-correction term in equation [15], but the long-run parameters are now imposed from the PMG approach rather than being freely estimated. However the speed of adjustment parameter is allowed to vary across each group. In a small number of cases individual dummy variables were included to remove large one-time outliers in order to ensure that the residuals in the estimated equations were normally distributed. This provides a relatively parsimonious way of treating outliers and ensures that there is a greater number of degrees of freedom remaining, which raises the power of any hypothesis tests.[35]

Tests of various restrictions on this set of equations are reported in Table 5. In contrast to the results obtained for the basic panel model with common coefficients for all parameters (Table 2), the restriction that each group has a common variance can now be accepted. However, the joint imposition of common parameters on both the income growth and the equilibrium-correction term is rejected, confirming that there are significant differences in the dynamic behaviour of migration for different source locations. Further tests showed that this was solely due to differences in the speed of adjustment parameter on the equilibrium-correction term. It was possible to impose restrictions to yield a parsimonious model with a common income

[35] In the ARDL models underlying the PMG estimates, some of these outliers are likely to have been picked up by additional dynamic terms specific to individual source groups. This should not affect the long-run structure that emerges from the PMG approach.

growth parameter across all source locations but two separate speed of adjustment parameters. The resulting model is reported in Table 6.

Table 5. Test of Restrictions on the Dynamic Mean Group Model

Common variances	LR(9) = 15.16
Common parameters	LR(18) = 51.85*
Common income growth parameter	LR(9) = 4.52
Common equilibrium-correction parameter	LR(9) = 44.82*
Common income growth parameter plus two sub-group equilibrium-correction parameters	LR(17) = 18.46

Notes: see Table 2. Equilibrium-correction term as defined in equation [21] in main text.

Table 6. The Restricted Dynamic Mean Group Model

Common income growth parameter	2.6898 (4.3)
Equilibrium-correction parameter Group 1	-0.2746 (3.8)
Equilibrium-correction parameter Group 2	-0.8801 (10.1)
Log Likelihood	24.85
Standard Error	0.224
Serial Correlation	LM(1) = 3.69

The EU, Old Commonwealth, African Commonwealth, the Rest of Africa and the Rest of Asia were found to have a common adjustment parameter of –0.27, whilst the other five locations were found to have a common parameter of –0.88, implying a much quicker change in the migration flow to any change in the key driving forces.

It is noticeable that the five locations with a slower speed of adjustment are the ones in which there is a clear trend in the migration rate over time in Figure 2. In these locations the linear combination of the main driving variables is generating an underlying change in the 'desired' level of migration over time, but because of persistence in migration patterns it takes some time for the actual migration rate to adjust towards the evolving long-run level. In the second group of locations, the migration rate at the end of the sample period is not greatly different to that at the start, with the possible exception of the rest of Europe. All that is being observed is stochastic fluctuations around some constant underlying migration rate delivered by the linear combination in [21].

Given that the final model arrived at appears to have a considerable degree of parameter homogeneity imposed, it is worth asking whether it is a significant improvement on the basic panel models we began with. The diagnostic tests in Table 6 suggest that it is compared to the models in Table 2, with an improved equation fit, and no evidence of mis-specification through significant first-order serial correlation. Moreover, it is not possible to accept the restrictions required to yield a model in which the short-term behaviour of migration is identical across groups. This appears to be the main source of heterogeneity in the behaviour of migration from different locations. Thus the two-step modelling approach appears to have been worthwhile, even though the heterogeneity in the data, and potential measurement error in interpolated variables such as the migrant stock, means that it is still difficult to draw strong policy conclusions because the parameters of interest cannot be estimated too precisely.

VI. Forecast Performance and Structural Stability

VI.1 Forecast Performance

It is possible to compare the different models estimated by evaluating their forecast performance. Having established that there is considerable heterogeneity across groups, we examine whether allowing for this heterogeneity in fact helps to deliver more accurate forecasts. This is of particular interest if models of migration are to be used in projections of the future population.[36] Forecasts for the aggregate migration inflow are obtained by summing the forecasts from the 10 disaggregate (location-specific) forecasts, and multiplying through by the population in those locations to derive the migration inflow from the migration rate.

We consider the forecasting performance of three models – the basic fixed effects panel model reported in column [2] of Table 2, the restricted dynamic mean group model with the long-run from the PMG estimates reported in Table 6 and equation [21] of the main text, and a Mean Group model with unrestricted, heterogeneous dynamics and error variances. We report two summary statistics, the Root Mean Square Error (RMSE) of the forecasts and the corrected Diebold-Mariano [DM] statistic proposed by Harvey *et al* (1997) which tests the null hypothesis of equal predictive performance between two models, assuming a quadratic loss function.

To provide a benchmark against which to compare the forecasting performance of the different panel models, two autoregressive models for the aggregate migration inflow were also estimated. The first model is estimated in levels with an unrestricted coefficient on lagged

[36] Current official practice in many countries, including the UK, is less parametric, with official forecasts being produced through a combination of time series models and judgement (see GAD (2002) and Population Trends (2002)). Projections of total migration into the UK are derived by summing up projections for the different components of total migration, namely the International Passenger Survey (IPS) figure, plus visitor switchers, migration from the Irish Republic and asylum seekers. Forecasts of IPS migration are made by summing forecasts of flows between the UK and four different groups of countries, the Old Commonwealth, the New Commonwealth, the EEA and Rest of the World. This is done using exponential smoothing, a procedure that is *ad hoc* in the sense that it is not defined with respect to a properly defined statistical model (Harvey (1993), pp.109-113). Judgement, and knowledge, of the likely movement of potential explanatory variables, are then used to 'correct' this figure, and forecast the remaining components of migration into the UK. Hollmann *et al* (2000) provide related information on the forecasting procedures used by the US Census Bureau, although again no details of a formal model are presented. Aggregate migration forecasts are again computed by summing up disaggregate (source country specific) forecasts, rather than by simply forecasting the total.

migration a year ago, and the second is estimated in first differences. It is important to consider such benchmark models since it is well known that parsimonious non-structural models, such as models in first differences, often deliver better short-term forecasts than structural models; see Clements and Hendry (1999). This is because these models can guard against unforeseen events such as structural breaks.

We undertake two separate tests. In the first, we consider the in-sample forecasting performance of our models by focusing on their respective fitted values obtained by estimating each model once over the full sample period (1981-2000). The second set of tests seeks to evaluate the forecasting performance of the models in 'quasi-real' time. This is achieved by splitting the sample period into estimation and forecasting periods, and recursively re-estimating the model over the forecasting period. We examine one-step ahead forecasts for each year from 1996 to 2000 and two-step ahead forecasts for each year from 1997 to 2000. For example, the first of these forecasts is made by estimating the model on a sample ending in 1995 and then deriving 1 and 2 year ahead estimates using the actual values of the explanatory variables and the parameters from the model estimated to 1995. We then re-estimate the model on a sample ending in 1996 and produce new 1 and 2 step ahead forecasts for 1997 and 1998 respectively, and so on.[37]

The results for the first test using the fitted values from the whole-sample regressions are reported in Panels A-C of Table 7. We report the RMSEs of the fitted values from the various models against the outturn over three different sample periods. The RMSE assuming that the mean observation from the sample period as a whole is used as a separate forecast is also presented in the sixth column of the Table. This is, of course, equal to the standard deviation of the migration inflow over the sample period. Corrected Diebold-Mariano statistics testing the forecasts of the disaggregate models against those from the two aggregate models, are also presented.[38] The lower the p-value, presented in parentheses below the corrected Diebold-Marino statistic, the greater the indication that we can reject the null hypothesis of equal forecasting performance between the two competing models. If we can reject it completely then the model with the lower RMSE is to be preferred to that with the higher RMSE.

[37] The out-of-sample forecasts were computed in models with all dummy variables that were zero over the restricted estimation period (but non-zero over the forecast period) excluded. This is because they were collinear with the intercept.

[38] The p-values are computed assuming a one-sided alternative hypothesis, i.e. we test the null of equality of forecast variances against the alternative of better predictive ability in the model with the lowest forecast variance.

Table 7. Summary In-sample Forecast Statistics

	[1]	[2]	[3]	[4]	[5]	[6]
A. 1983-2000	Fixed effects	PMG	MG	AR(1) in levels	AR(1) in differences	Mean
RMSE	17.23	18.79	23.19	23.62	25.18	46.19
DM against AR in differences	-1.71 [0.05]	-2.07 [0.02]	-0.57 [0.28]			
DM against AR in levels	-1.16 [0.13]	-1.38 [0.09]	-0.08 [0.46]			
B. 1990-2000	Fixed effects	PMG	MG	AR(1) in levels	AR(1) in differences	Mean
RMSE	16.18	18.58	25.45	25.83	26.68	47.01
DM against AR in differences	-1.37 [0.09]	-1.73 [0.05]	-0.22 [0.41]			
DM against AR in levels	-1.17 [0.13]	-1.33 [0.10]	-0.05 [0.47]			
C. 1995-2000	Fixed effects	PMG	MG	AR(1) in levels	AR(1) in differences	Mean
RMSE	12.96	16.28	28.56	20.96	23.44	43.81
DM against AR in differences	-0.99 [0.18]	-1.38 [0.11]	0.98 [0.18]			
DM against AR in levels	-0.97 [0.18]	-1.69 [0.07]	1.54 [0.09]			
1-step ahead forecasts computed recursively						
D. 1996-2000	Fixed effects	PMG	MG	AR(1) in levels	AR(1) in differences	Mean
RMSE	18.77	20.07	28.23	34.79	25.49	38.00
DM against AR in differences	-0.52 [0.31]	-0.70 [0.26]	0.38 [0.36]			
DM against AR in levels	-1.14 [0.15]	-1.50 [0.10]	-0.86 [0.21]			
2-step ahead forecasts computed recursively						
E. 1997-2000	Fixed effects	PMG	MG	AR(1) in levels	AR(1) in differences	Mean
RMSE	32.29	32.28	51.46	64.99	47.94	33.62
DM against AR in differences	-0.90 [0.21]	-0.76 [0.25]	0.12 [0.45]			
DM against AR in levels	-1.25 [0.14]	n.a.	-1.14 [0.16]			

Note: p-values in parentheses

The results show that, particularly over the longer sample periods, the forecasts from the panel data models (fixed effects, PMG and MG) beat those from the aggregate autoregressive models, in the sense that they have a lower, and in some cases significantly lower, RMSE. This is encouraging and indicates that disaggregate forecasts of migration beat aggregate, atheoretical ones. The behavioural model of migration provides significant extra information over and above that from a purely atheoretical time series model with no economic content. Given the trending nature of total migration it is not surprising that the mean forecast performs badly.

Comparing the three panel data models, it is of note that the fixed effects model, a restricted model that imposes common short-run dynamics, despite the fact that they are rejected by the data, as well as a common long-run, not only beats the time-series models but also our preferred specification, in terms of having the lowest RMSE. This is consistent with the view that parsimonious models, that are inappropriate for structural inference, can deliver better forecasts (Clements and Hendry, 1999).

However, the corrected Diebold-Mariano p-values show that our preferred model, the PMG model, in fact does best against the benchmark models, in the sense that its p-values are lower than those of the other models. This model delivers forecasts that are statistically better, at around a 90% level of significance, than those from the two benchmark auto-regressive models.

The results for the second test are reported in Panels D and E of Table 7. The picture is similar to that from the first test, with the fixed effects panel model and the PMG model having lower RMSEs than the other models. However it is not possible to conclude that the differences are significant on the basis of the Diebold-Mariano tests.[39]

The results also confirm that, as might be expected, the uncertainty around any point forecast of migration rises with the forecast horizon.[40] This is seen by comparing the RMSE from the two-step ahead forecast with those from the one-step ahead forecast. It is also noticeable that the lower RMSE in the fixed effects panel model relative to that in the restricted PMG model in the one-step ahead forecasts, is not present in the two-step ahead forecasts. It is not possible to draw any firm conclusions from this, particularly given the relatively small sample size, but it suggests that the statistically-acceptable restricted PMG model may be a better model from which to make multi-year projections.

[39] In one case the DM statistic could not be calculated due to a singular variance-covariance matrix. This is an accepted problem with DM tests.

[40] There are a number of ways of highlighting the uncertainty associated with any given point forecast. Both GAD and the US Census Bureau produce high and low variant forecasts around their central migration forecasts, computed using alternative assumptions of future migration and population (Shaw, 1994). An alternative approach is to consider the performance of past projections; see Shaw (1994) and GAD (2002). Other options suggested by the econometric literature on forecasting when producing a model based forecast, are to derive analytical expressions for the uncertainty associated with the forecasts or, if this is not possible, to use simulation techniques. The latter two techniques would facilitate the construction of 'confidence bands', for a given significance level, around the point, forecast, or even density forecasts that fully capture forecast uncertainty.

VI.2 Structural Stability

A related question is whether the estimated models are structurally stable over the entire sample period. This question is of interest given the apparent step-up in the number of migrants since the mid-1990s. In particular, it is of interest to ask whether there is any evidence that the level of migration in any of the years from 1995 onwards has been significantly higher than the equations can otherwise explain through the economic and demographic forces driving migration. There are a number of different ways in which this question can be approached. The simplest is to augment the estimated models with a (1,0) dummy variable which is unity for all panel members in a single year and zero at all other times. This dummy will pick up all otherwise unexplained factors that are common to all panel members in that year. So, for example, if migration from all locations was systematically higher (in proportionate terms as we use a logarithmic specification) than could otherwise be explained in year *t*, a dummy variable for year *t* would have a positive, and possibly significant, coefficient.

We consider two models – the basic fixed effects panel model reported in column [2] of Table 2 and the restricted dynamic pooled mean group model with the long-run from the PMG estimates reported in Table 6 and equation [21] of the main text. For each model we tested the (joint) significance of various subsets of dummies, with the results being summarised in Table 8. There is no evidence of significant structural change across the panel as a whole in any time period.[41] For example, in Row A we report the outcome of augmenting the two models with separate dummy variables for each year from 1995 to 2000. The coefficients on these two sets of variables both proved to be jointly insignificant. The absence of any evidence of systematic under or overprediction of migration across the panel as a whole means that the rise in the number of migrants in recent years can largely be explained by changes in the main driving forces behind migration.

Table 8. Tests Of Structural Stability

Time Period	Panel Model	PMG Model
1995-2000	LR(6)=7.51	LR(6)=2.41
1996-2000	LR(5)=4.26	LR(5)=1.27
1997-2000	LR(4)=4.24	LR(4)=0.99
1998-2000	LR(3)=2.58	LR(3)=0.52
1999-2000	LR(2)=2.27	LR(2)=0.37
2000	LR(1)=1.42	LR(1)=0.25

Note: Panel Model is column [2] of Table 2, and PMG model is the one in Table 6. LR denotes a likelihood ratio test.

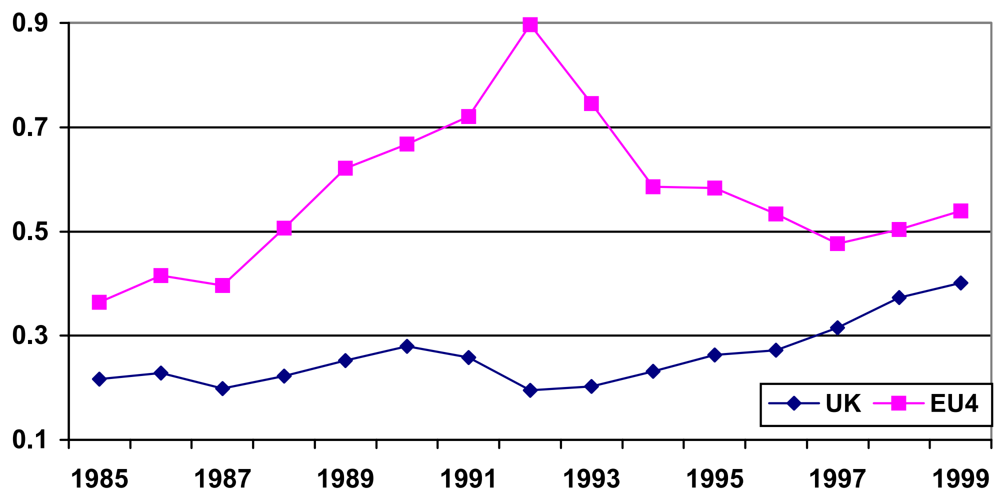
[41] This does not rule out the possibility of structural change in the pattern of migration from an individual region.

VI.3 Has the UK Become Relatively Easier to Enter?

One possible influence on migration that is not included in the estimated models in this paper is an explicit quantitative measure of migration policy (Hatton and Williamson, 2002). This absence is largely because the UK does not operate a direct system with quotas, but relies on informal variations in the implementation of entry criteria, such as the evidence required in order to obtain work permits. But as migrants face choices, it might be argued that the rising inflow of migrants into the UK in recent years is due in part to the existence of an unchanged entry regime in the UK at a time when entry criteria in other possible host countries have been tightened. If this were the case, then explicit changes in immigration policies in the UK would be one potential policy available through which to affect the level of migrant inflows.

To investigate this issue, we augment two of the estimated models with the (one-year) lagged ratio of migrant inflows into the UK, expressed as a proportion of the UK population, relative to the same ratio for the EU4 countries. We use Eurostat data for non-national immigrants over 1985-1999, as this is the only comparable data available for gross immigration. The data are taken from *Demographic Statistics 1960-99*, updated using subsequent issues of the *Demography Yearbook*.

Figure 8. Non-national Migrant Inflows (% of host population)



The data are plotted in Figure 8. Throughout the period shown, immigration as a share of the host population has been higher in the EU4 than in the UK. However it is clear that the gap has narrowed in recent years, reflecting both a rising level of migration into the UK as well as a relative weakening in the level of inflows into the EU4 compared to the late 1980s and early 1990s. To the extent that these differences reflect policy choices, it could be argued that the relative improvement in the ease with which it is possible to enter the UK may be one factor behind the rising level of inward migration. A related argument, consistent with the model in Epstein (2002), is that migration may be affected by 'herd behaviour'. If potential migrants are imperfectly informed about the characteristics of different locations, then they may augment their information set through observing, and subsequently following, others who have chosen a particular location.

We undertook various tests to see whether the past relative migrant inflow into the UK mattered for current migration into the UK by re-estimating two equations over a shorter sample period from 1986-2000 and augmenting them with the (log of the) one year lagged migrant inflow ratio. Again, the two models used were the basic fixed effects panel model reported in column [2] of Table 2 and the restricted dynamic pooled mean group model with the long-run from the PMG estimates reported in Table 6 and equation [21] of the main text.

Four separate hypotheses were tested, with the test statistics reported in Table 9. There was no evidence of statistically significant effects from the relative migration ratio on the migration rates from all destinations, from 5 developing regions or from either the 2 African regions or the 2 Asian regions, although the estimated coefficient was generally positive. The effect of the lagged relative migration ratio on migration from African and Asian regions is clearly stronger than it is on migration from other regions, but it is still not statistically significant. Thus there does not appear to be any statistical evidence to support the proposition that the relative improvement in the ease with which it is possible to enter the UK is a factor behind the recent rise in the level of inward migration.

Table 9. Tests Of Whether The Relative UK-EU4 Migration Ratio Matters

Source Regions	Panel Model	PMG Model
All	LR(1)=0.94	LR(1)=0.01
5 Developing	LR(1)=0.49	LR(1)=0.33
2 African	LR(1)=0.38	LR(1)=2.07
2 Asian	LR(1)=1.57	LR(1)=1.46

Notes: The 5 developing regions are the Rest of Europe, the African Commonwealth, the Asian Commonwealth, the Rest of Africa and Middle East and the Rest of Asia. See also Table 8 notes.

VII. Concluding Comments

In this paper we have developed the first detailed econometric model of the economic and demographic determinants of annual migrant inflows into the UK. The model provides a framework which can be utilised to obtain more soundly based projections of the future level of migrant inflows. Analysis of the factors responsible for the growth of *gross* inward migration in the late 1990s suggests that the higher level seen at that time should persist. To this extent, our results provide support for the assumptions made in HM Treasury (2002) about *net* inward migration in the calculations of the trend rate of growth of the British economy.

Some of the factors we identify as important determinants of inward migration flows have been shown to be important for understanding developments in other economies as well. Examples include 'friends and family' effects from existing migrants, income differentials between the host and source location and the demographic structure of the source location population. But we also show that it is important to allow both for developments in other potential host locations and for heterogeneity in the speed with which migration from different source locations responds to changes in economic circumstances. Neither of these factors have been given much attention to date in the applied literature on international migration.

Although we do not find any statistical evidence that international variations in migration policies have a significant direct influence on the level of inward migration in the UK, this does not mean that there are no policies available with which the government can affect migrant flows. It is just that most operate in an indirect manner *via* their impact on the main economic forces driving migration. Any policy which is aimed at improving productivity levels in the UK will, by definition, affect migrant inflows *via* the impact on per capita incomes. Equally, supply-side policies to improve the workings of the labour market and measures to improve the openness of the economy to international trade (and investment) will also have positive influences over time.

A more direct effect on migrant flows could possibly be achieved by examining the friends and family link more closely. The balance of evidence in this paper suggests that measures to change the existing migrant stock could have a long-lasting effect on the subsequent level of migrant inflows. It would be wise to explore this relationship further given the relative lack of precision with which the total migrant stock can be measured and the relative sensitivity of the empirical estimates of network effects to the choice of estimation technique.

Appendix 1. Interpolating The Migrant Stock

In this Appendix we provide a brief description of the methods used to interpolate the foreign-born migrant stock. At the time this project was undertaken there were only two observations available in our sample period, from the Censuses in 1981 and 1991. We assume these figures are measured on a mid-year basis.

Setting t=1981 and letting S and m denote the migrant stock and the gross migrant inflow respectively, we initially generated a set of estimates:

$$\hat{S}_{t+10} = S_t + \sum_{i=1}^{10} \alpha \frac{(m_{t+i} + m_{t+i-1})}{2}. \qquad \text{[A1]}$$

An average of migrant flows over two years is used as the stock; it is assumed to be measured on a mid-year basis. The parameter α (α≤1) is an adjustment factor to allow for the fact that some migrants are returning British nationals. We assume the following fixed weights over time:

EU	0.6	Other Commonwealth	0.75
USA	0.75	Rest of Europe	0.9
Old Commonwealth	0.666	Latin America	1.0
African Commonwealth	0.75	Rest of Africa and Middle East	0.9
Asian Commonwealth	1.0	Rest of Asia	1.0

Using [A1] a residual adjustment factor, R, was calculated as:

$$R = \left\lfloor S_{t+10} - \hat{S}_{t+10} \right\rfloor / 10. \qquad \text{[A2]}$$

This residual factor will incorporate trends such as the emigration and deaths of foreign-born nationals residing in the UK.

The residual adjustment factor was then added to [A1] to compute revised stock estimates for the missing years 1982-1990:

$$\hat{S}_{t+i} = \hat{S}_{t+i-1} + \alpha \frac{(m_{t+i} + m_{t+i-1})}{2} + R. \qquad \text{[A3]}$$

The estimated migrant stocks were then used to generate an implied region-specific time-varying attrition rate, δ_t using the conventional stock-flow identity that:

$$S_t = S_{t-1}(1-\delta_t) + \alpha \frac{(m_t + m_{t-1})}{2}. \qquad \text{[A4]}$$

The calculated attrition rate for 1982-1991 was projected forward for the period 1992-2000 and backwards for the period 1980-81 using the fitted values from country (*j*) specific regressions of the attrition rate on a constant and a deterministic time trend:

$$\delta_{jt} = \beta_j + \gamma_j \, TIME_t + \varepsilon_{jt}. \qquad \text{[A5]}$$

These attrition rates, along with the migrant flows, were then used to generate the stock data outside the interval 1981-91 using equation [A4].

References

Barrell, R. and Pain, N. (1999a), 'Trade restraints and Japanese direct investment flows', *European Economic Review*, 43, 1, 29-45.

Barrell, R. and Pain, N. (1999b), 'Domestic institutions, agglomerations and foreign direct investment in Europe', *European Economic Review*, vol.43, pp.925-934.

Birrell, B., Dobson, I.R., Rapson, V. and Smith, T.F. (2001), 'Skilled labour gains and losses', report for the Department of Immigration and Multicultural Affairs, Commonwealth Government of Australia.

Borjas, G.J. (1989), 'Economic theory and international migration', *International Migration Review*, 23, 457-485.

Borjas, G.J. (1994), 'The economics of immigration', *Journal of Economic Literature*, 32, 1667-1717.

Borjas, G.J. (1999a), 'Economic research on the determinants of immigration: Lessons for the European Union', World Bank Technical Paper No. 438.

Borjas, G.J. (1999b), 'The economic analysis of immigration', in *Handbook of Labor Economics Volume 3A*, (eds. Ashenfelter, O. and Card, D.), Elsevier, North Holland.

Carrington, W.J., Detragiache, E. and Vishwanath, T. (1996), 'Migration with endogenous moving costs', *American Economic Review*, 86, 909-930.

Clark, X., Hatton, T.J. and Williamson, J.G. (2002), 'Where do U.S. immigrants come from, and why?', NBER Working Paper 8998 (http://www.nber.org/papers/w8998)

Clements, M.P. and D.F. Hendry (1999), *Forecasting Non-Stationary Economic Time Series*, The MIT Press, London.

Dobson, J., Koser, K., McLaughlan, G. and Salt, J. (2001), 'International migration and the United Kingdom: recent patterns and trends', Home Office Research and Development Statistics Directorate Working Paper No.75.

Epstein, G.S. (2002), 'Informational cascades and the decision to migrate', CEPR Discussion Paper No. 3287.

Eurostat (2001), 'Indicators of migration between the Republic of Ireland and the United Kingdom', Eurostat Working Paper 3/1998/E/n°1.

Faini, R. and Venturini, A. (1993), 'Trade, aid and migrations: some basic policy issues', *European Economic Review*, 37, 435-442.

Fertig, M. (2001), 'The economic impact of EU enlargement: assessing the migration potential', *Empirical Economics*, 26, 707-720.

Finn, M.G. (2001), *Stay Rates of Foreign Doctorate Recipients from US Universities*, Oak Ridge Institute for Science and Education, Oak Ridge, TN.

GAD (2002), 'National Population Projections: 2000-based', Government Actuaries Department, National Statistics, Series PP2, No. 23.

Ghatak, S., Levine, P. and Wheatley Price, S. (1996), 'Migration theories and evidence: an assessment', *Journal of Economic Surveys*, 10, 159-198.

Girma, S. and Yu, Z. (2002), 'The link between immigration and trade: evidence from the United Kingdom', *Weltwirtschaftliches Archiv*, 138, 115-130.

Glover, S., Gott, C., Loizillon, A., Portes, J., Price, R., Spencer, S., Srinivasan, V. and Willis, C. (2001), 'Migration: an economic and social analysis', RDS Occasional Paper No. 67, The Home Office.

Hadri, K. (2000), 'Testing for stationarity in heterogeneous panel data', *Econometrics Journal*, 3, 148-161.

Harvey, A.C. (1993), *Time Series Models*, 2nd edition, Harvester-Wheatsheaf.

Harvey, D. I. and S. J. Leybourne and P. Newbold (1997), 'Testing the Equality of Prediction Mean Square Errors', *International Journal of Forecasting*, 13, 273--281.

Hatton, T.J. (1995), 'A model of UK emigration, 1870-1913', *Review of Economics and Statistics*, 77, 407-415.

Hatton, T.J. and Williamson, J.G. (2001), 'Demographic and economic pressure on emigration out of Africa', IZA Discussion Paper No. 250.

Hatton, T.J. and Williamson, J.G. (2002), 'What fundamentals drive world migration?', CEPR Discussion Paper No. 3559.

Hausman, J.A. (1978), 'Specification tests in econometrics', *Econometrica*, 46, 251-72.

HM Treasury (2002), *Trend Growth: Recent Developments and Prospects*, available at: http://www.hm-treasury.gov.uk/media//D6678/ACF521.pdf

Hollmann, F.W., Mulder, T.J. and Kallan, J.E. (2000), 'Methodology and assumptions for the population projections of the United States: 1999 to 2100', Population Division Working Paper No. 38, U.S. Census Bureau, Department of Commerce, Washington DC, USA.

Im, K-S., Pesaran, M.H. and Shin, Y. (2003), 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics*, 115, 53-74.

Jasso, G., Rosenzweig, M.R. and Smith, J.P. (1998), 'The changing skill of new immigrants to the United States: Recent trends and their determinants', NBER Working Paper 6764 (http://www.nber.org/papers/w6764).

Karemera, D., Oguledo, V.I. and Davis, B. (2000), 'A gravity model analysis of international migration to North America', *Applied Economics*, 32, 1745-1755.

Koser, K. and Pinkerton, C. (2002), 'The social networks of asylum seekers and the dissemination of information about countries of asylum', report for Home Office, available at: http://www.homeoffice.gov.uk/rds/pdfs2/socialnetwork.pdf

McLaughlan, G. and Salt, J. (2002), 'Migration policies towards highly skilled foreign workers', report for Home Office, available at: http://www.homeoffice.gov.uk/rds/pdfs2/migrationpolicies.pdf

Mitchell, J. and Pain, N. (2002), 'The determinants and methods of modelling international migration: an overview of the literature', mimeo, NIESR.

Morrell, J. and Pearson, R. (2002), 'Knowledge migrants – the motivations and experiences of professionals in the UK on work permits', report prepared for the Department of Trade and Industry.

ONS (2002), *International Migration*, Series MN no.27, Office for National Statistics, London.

Pesaran, M.H. (1997), 'The role of economic theory in modelling the long-run', *Economic Journal*, 107(440), 178-191.

Pesaran, M.H. and Shin, Y. (1999), 'An autoregressive distributed lag modelling approach to cointegration analysis', in Strom, S. (ed.) *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press.

Pesaran, M.H., Shin, Y. and Smith, R.P. (1999), 'Pooled mean group estimation of dynamic heterogeneous panels' *Journal of the American Statistical Association*, 94, 621-634.

Pesaran, M.H. and Smith, R.P. (1995), 'Estimating long-run relationships from dynamic heterogeneous panels', *Journal of Econometrics*, 68, 79-113.

Pesaran, M.H., Smith, R.P. and Im, K-S. (1996), 'Dynamic linear models for heterogenous panels', in Mátyás, L. and Sevestre, P. (eds.) *The Econometrics of Panel Data*,), Kluwer Academic Publishers, Dordrecht, The Netherlands.

Pesaran, M.H., Smith, R.P. and Akiyama, T. (1998), *Energy Demand in Asian Developing Economies*, Oxford University Press, Oxford.

Pesaran, M.H. and Zhao, Z. (1999), 'Bias reduction in estimating long-run relationships from dynamic heterogeneous panels', in Hsiao, C., Lahiri, K., Lee, L-F. and Pesaran, M.H. (eds.) *Analysis of Panels and Limited Dependent Variables: A Volume in Honour of G. S. Maddala*, Cambridge University Press, Cambridge.

Population Trends (2002), '2000-based National population projections for the United Kingdom and its constituent countries', *Population Trends*, 107, 5-13.

Regets, M.C. (2001), 'Research and policy issues in highly-skilled international migration: a perspective with data from the United States', IZA Discussion Paper No. 366.

Robinson, V. and Segrott, J. (2002), 'Understanding the decision-making of asylum seekers', Home Office Research Study 243.

Rotte, R. and Vogler, M. (1998), 'Determinants of international migration: empirical evidence for migration from developing countries to Germany', CEPR Discussion Paper No. 1920.

Schiff, M. (1997), 'South-North migration and trade: survey and policy implications', *Policy Research Working Paper No. 1696*, The World Bank.

Shaw, C. (1994), 'Accuracy and uncertainty of national population projections for the United Kingdom', *Population Trends*, 77, 24-32.

Stark, O. and Bloom, D. (1985), 'The new economics of labour migration', *American Economic Review Papers and Proceedings*, 75, 173-178.

Stark, O. and Taylor, J.E. (1991), 'Migration incentives, migration types: the role of relative deprivation', *The Economic Journal*, 101, 1163-1178.

van der Gaag, N. and van Wissen, L. (1999), 'Analysis and forecasting of international migration by major groups (Part II)', Working Paper, Eurostat, 3/1999/E/no.9.

Winkleman, R. (2001), 'Why do firms recruit internationally? Results from the *IZA International Employer Survey* 2000', IZA Discussion Paper No.331.

Zellner, A. (1962), 'An efficient method of estimating seemingly unrelated regressions, and tests for aggregation bias', *Journal of the American Statistical Association*, 57, 348-68.