Multilateral Resistance to Migration∗

Simone Bertoli\textsuperscript{a} and Jesús Fernández-Huertas Moraga\textsuperscript{b}

\textsuperscript{a}Robert Schuman Centre, European University Institute\textsuperscript{†}
\textsuperscript{b}FEDEA\textsuperscript{‡} and IAE, CSIC

March 14, 2011

Abstract

The scale of migration flows between two countries does not only depend on their relative attractiveness, but also on the one of alternative destinations. Following the trade literature, we term the influence exerted by other destinations on bilateral flows as Multilateral Resistance to Migration, and we show how it can be accounted for when estimating the determinants of bilateral migration flows in the context of a general individual random utility maximization model. We propose the use of the Common Correlated Effects estimator (Pesaran, 2006) and apply it to high-frequency data on the Spanish immigration boom between 1997 and 2009. Compared to more restrictive estimation strategies developed in the literature, the bias goes in the expected direction: we find a smaller effect of GDP per capita and a larger effect of migration policies on migration flows.

Keywords: international migration, economic determinants, migration policies, time-varying attractiveness, multiple destinations

JEL classification codes: F22, O15, J61.

∗We would like to thank Lídia Brun for her excellent research assistance in compiling the database on Spanish immigration policies; Jesús Fernández-Huertas Moraga received financial support from the ECO2008-04785 project funded by the Spanish Ministry for Science and Innovation. The usual disclaimers apply.

†Via delle Fontanelle, 19, I-50014, San Domenico di Fiesole; email: simone.bertoli@eui.eu

‡Jorge Juan, 46, E-28001, Madrid; email: jfernandezhuertas@fedea.es (corresponding author)
1 Introduction

The responsiveness of the scale of migration flows to varying economic conditions - both in sending and recipient countries - and to changing immigration policies at destination represents a central topic in the international migration literature. While some recent contributions have provided econometric analysis of aggregate data where the identification strategy is fully consistent with the proposed underlying individual-level migration decision model (Beine, Docquier, and Ozden, 2011; Grogger and Hanson, 2011; Ortega and Peri, 2009),\(^1\) others have relied on econometric specifications that have not been fully micro-founded (Clark, Hatton, and Williamson, 2007; Pedersen, Pytlikova, and Smith, 2008; Mayda, 2010; Theoharides, McKenzie, and Yang, 2010).

This methodological difference notwithstanding, these papers share a crucial feature, as Hanson (2010) observes that the literature is characterized by a long-standing tradition of “estimating bilateral migration flows as a function of characteristics in the source and destination countries only”. Still, would-be migrants sort themselves across alternative destinations, so that it is important to understand whether this econometric approach allows to control for the possible dependence of the scale of migration between two countries upon the time-varying attractiveness of other migrants’ destinations. Hanson (2010) argues that “failing to control other migration opportunities could [...] produce biased estimates”, and this issue resembles the one raised by Anderson and van Wincoop (2004) with respect to the estimation of the determinants of bilateral trade flows.

Trade between two countries does not depend on bilateral trade costs only, but rather on the relationship between these costs and the costs with the other trading partners; Anderson and van Wincoop (2004) refer the attractiveness of trading with other partners as multilateral resistance to trade.\(^2\) Similarly, the scale of migration flows between a dyad represented by an origin and a destination country does not depend on the attractiveness of the latter, but also on how this relates to the opportunities to move to other destinations. Following the terminology introduced by Anderson and van Wincoop (2004), we refer to the attractiveness of other destinations as Multilateral Resistance to Migration.

---

\(^1\)Bertoli, Fernández-Huertas Moraga, and Ortega (2010) analyze the income-sensitivity of international migration flows using individual-level data.

\(^2\)Baldwin (2006) observes that this is nothing more than a specific case of the general principle that “relative prices matter”.

---

2
This paper directly addresses the concern raised by Hanson (2010). First, it explicitly relates the stochastic properties of the underlying individual migration decision model to the need to control for the multilateral resistance to migration. Second, it shows how unbiased estimates of the determinants of bilateral migration flows can be obtained from characteristics of the source and destination countries alone even when the multilateral resistance to migration matters. Third, it applies the proposed econometric approach - which draws on Pesaran (2006) - to analyze the determinants of migration flows to Spain over 1997-2009 using high-frequency administrative data.

The paper presents a general random utility maximization (RUM) model that describes the migration decision problem that individuals face. The theoretical model - which generalizes the one presented in Ortega and Peri (2009) - shows that multilateral resistance to migration represents an issue for the analysis of aggregate data whenever the stochastic component of location-specific utility is such that the independence of irrelevant assumptions fails. The derivation of the econometric specification from the random utility maximization model reveals that multilateral resistance to migration, which is unobservable for the econometrician, gives rise to an endogeneity problem, as the regressors are correlated with the error term, and this also exhibits serial and spatial correlation.

We show that the multilateral resistance to migration term entering the error of the equation that describes the determinants of aggregate migration flows on the basis of the RUM model can be expressed as the inner product of a vector of origin-specific factor loadings and a vector of time-specific common effects. This entails that the structure of the error term coincides with the multifactor error model presented in Pesaran (2006). Pesaran (2006) proposed an estimator, the Common Correlated Effects (CCE) estimator, which allows to derive consistent estimates from panel data when the error follows this structure, i.e. it is serially and spatially correlated, and the regressors are endogenous. The CCE requires to estimate a regression where the cross-sectional average of the dependent and of all the independent variables are included as auxiliary regressors: consistency of the estimates follows

---

3The converse is also true: if the independence of irrelevant alternatives characterizes the individual migration decision problem, then the time-varying attractiveness of other destinations can be disregarded in the econometric analysis, as in Grogger and Hanson (2011) and Beine, Docquier, and Ozden (2011).

4Driscoll and Kraay (1998) allow to address the violation of the classical assumptions on the error term, but still requires exogeneity of the regressors, which does not hold when multilateral resistance to migration is an issue.
from the fact that the multilateral resistance to migration term can be approximated by an
origin-specific linear combination of the cross-sectional averages (Pesaran, 2006).

The adoption of the CCE estimator allows us to address the challenge posed by multi-
lateral resistance to migration using the same data that are traditionally employed in the
literature. This approach is more general than the one proposed in Mayda (2010), who
includes a weighted average of income per capita in the other destinations as a control for
their time-varying attractiveness, and the one in Ortega and Peri (2009), which is valid only
under more restrictive assumptions on the underlying RUM model and which does not allow
to identify the effects of origin-specific variables.

The proposed econometric approach is applied to the analysis of the determinants of
bilateral migration flows to Spain between 1997 and 2009, when this country experienced
an unprecedented boom in immigration. In fact, Spain recorded “the highest rate of growth
of the foreign-born population over a short period observed in any OECD country since
the Second World War” (OECD, 2010): the immigrant share went from 3 percent of the
population in 1998 to 14 percent in 2009 (INE, 2010b). Migration data come from the
Estadística de Variaciones Residenciales (EVR; INE, 2010a), an administrative dataset
collected on a monthly basis by the Instituto Nacional de Estadística. A key feature of
the EVR is that it provides us with high-frequency data, which give to the dataset the
longitudinal dimension that is required to be confident about the application of the CCE
estimator (Pesaran, 2006).

The data from the EVR, which have been aggregated by quarter, have been combined
with data from IMF (2010a) and World Bank (2010) on real GDP and population at origin
for 61 countries, which represent 87 percent of the total flows to Spain over our period of
analysis. Furthermore, we have compiled information about the various facets of Spanish
immigration policies - such as bilateral visa waivers and agreements on the portability of
pension rights - which have been shown to be relevant determinants of recent immigration
to Spain (Bertoli, Fernández-Huertas Moraga, and Ortega, 2011).

---

5Hanson (2010) wonders whether this is “a sufficient statistic for other migration opportunities”.
6These figures can only be compared with Israel in the 1990s, when “immigration increased Israel’s
population by 12 percent between 1990 and 1994, after emigration restrictions were lifted in an unstable
Soviet Union” (Friedberg, 2001), at a time when Israel had not yet joined the OECD.
7Data from the International Financial Statistics (IMF, 2010a) have been also combined with data from
the World Economic Outlook (IMF, 2010b), and various Central Banks, as described in the Appendix A.3.
Our results show that ignoring the multilateral resistance to migration term biases the estimation of the determinants of migration flows to Spain. In addition, the direction of the bias is the one we could expect from the existence of multilateral resistance to migration. The effect of GDP at origin on migration flows to Spain is two thirds of that found in a model without multilateral resistance to migration, although it is still negative and significant: a 1 percent drop in GDP per capita in a country increases its emigration rate to Spain by 3.1 percent. This bias is in the opposite direction of that found on the impact of migration policies. The only migration policy that has a significant effect on migration flows to Spain is the adoption of a visa waiver. This effect only turns significant when the multilateral resistance to migration is accounted for: establishing a visa waiver for a country multiplies its emigration rate to Spain by a factor of 4,\(^8\) while the estimated effect when multilateral resistance is not controlled for is not significantly different from zero.

The paper is related to four strands of economic literature. First, the papers that analyze the determinants of bilateral migration flows using panel data (Clark, Hatton, and Williamson, 2007; Lewer and den Berg, 2008; Grogger and Hanson, 2011; Mayda, 2010; Ortega and Peri, 2009; Simpson and Sparber, 2010; Pedersen, Pytlikova, and Smith, 2008; Beine, Docquier, and Ozden, 2011). In terms of the structure of the data, the paper is closely related to Clark, Hatton, and Williamson (2007) and Theoharides, McKenzie, and Yang (2010), which estimate the determinants of bilateral flows to one destination, the United States, and from one origin, the Philippines, respectively.\(^9\)

Second, we draw on the papers that have analyzed high-frequency migration data. Specifically, Hanson and Spilimbergo (1999) and Orrenius and Zavodny (2003) who analyze monthly migration flows from Mexico to the United States.

Third, the theoretical and empirical analysis presented here is related to the papers in the trade literature that discuss the relevance of multilateral resistance to trade (Anderson and van Wincoop, 2003, 2004; Baldwin, 2006).

Fourth, the paper is related to the contributions in the econometric literature that present

\(^8\)This huge effect is in line with the findings of Bertoli, Fernández-Huertas Moraga, and Ortega (2011) for the case of Ecuadorian migration to Spain.

\(^9\)The analysis is also related to the papers that estimate the influence of demographic factors (Hanson and McIntosh, 2010b,a) and migration networks (Edin, Fredriksson, and Åslund, 2003; Munshi, 2003; McKenzie and Rapoport, 2010; Bertoli, 2011) upon migration flows; these effects are controlled for but not estimated in our paper.
estimators which allow to deal with violations on the classical assumption about the variance structure of the error term (Driscoll and Kraay, 1998; Hoechle, 2007; Coakley, Fuertes, and Smith, 2002), and with the endogeneity of the regressors (Pesaran, 2006; Bai, 2009; Pesaran and Tosetti, 2011).\footnote{Endogeneity of some of the regressors, such as GDP at origin, goes beyond the effect exerted by multilateral resistance to migration: Mishra (2007) and Docquier, Ozden, and Peri (2010) show how wages at origin respond to migration whereas Borjas (2003) and Ottaviano and Peri (2010) among many others show how wages at destination respond to migration, and Bugamelli and Paternó (2009) analyze the relationship between migrants’ remittances and current account reversals, and they conclude that remittances lower the probability of such a reversal; Anderson (2011) explores the implications for the estimation strategy when GDP is endogenous to migration flows.}

The paper is structured as follows: Section 2 presents the random utility maximization model that represents the individual migration decision problem; Section 3 analyzes the relationship between the stochastic properties of the RUM model and the need to control for multilateral resistance to migration in the econometric analysis through the CCE estimator proposed by Pesaran (2006). Section 4 presents the sources of the data used in the econometric analysis and the descriptive statistics. Section 5 discusses the estimates, and the empirical relevance of multilateral resistance to migration for the case that we have analyzed. Finally, Section 6 draws the main conclusions.

2 From individual decisions to aggregate flows

We present here a random utility maximization model that describes the location choice problem that would-be migrants face, which gives us the basis for deriving the determinants of bilateral aggregate migration flows. To keep it as general as possible, we do not specify the factors that influence location-specific utility.

2.1 Random utility maximization model

Consider a set of individuals, indexed by $i$, originating from a country $j$ belonging to a set $H$, who have to chose their preferred location among countries belonging to the set $D_j = D \cup \{j\}$.\footnote{The sets $H$ and $D$ can overlap.} Let the elements in $D_j$ be indexed by $k$; the utility that the individual $i$ from country $j$ obtains from opting for destination $k$ is given by:
\[ U_{ijk} = V_{jk} + \epsilon_{ijk} = x_{jk}' \beta + \epsilon_{ijk} \tag{1} \]

where \( x_{jk} \) is a vector of factors - which can include location- or dyad-specific elements, and \( \epsilon_{ijk} \) is the stochastic term. Let \( \Sigma_j = (\epsilon_{jk} \epsilon_{jk}') \) be the covariance matrix of the error term, with \( j \in H \), with the outer product being made over the destinations \( k \in D_j \). We assume that \( \Sigma_j \) is a block-diagonal matrix, i.e. the stochastic component of utility \( U_{ijk} \) can be correlated across subsets of destinations. Let the set \( D_j \) be partitioned into \( n \geq 1 \) subsets \( d_j \) and let us denote with \( d(j,k) \) the subset to which destination \( k \) belongs to for individuals from country \( j \). We assume that the error term follows a Generalized Extreme Value distribution, so that: \(^{12}\)

\[ \forall h, k \in D \mid h \in d(j,k) : \frac{E(\epsilon_{jk}' \epsilon_{jh})}{\sqrt{E(\epsilon_{jk}' \epsilon_{jk}) E(\epsilon_{jh}' \epsilon_{jh})}} = \rho_{d_d(j,k)} > 0 \]

and

\[ \forall h, k \in D \mid h \notin d(j,k) : \frac{E(\epsilon_{jk}' \epsilon_{jh})}{\sqrt{E(\epsilon_{jk}' \epsilon_{jk}) E(\epsilon_{jh}' \epsilon_{jh})}} = 0 \]

We also assume that \( d(j,j) \) is a singleton for any \( j \in H \) (i.e. the country of origin has no close substitute among the other potential destinations). \(^{13}\)

These assumptions on the stochastic term imply that the probability \( p_{ijk} \) that individual \( i \) will opt for destination \( k \in D_j \) is given by:

\[ p_{ijk} = \frac{e^{V_{jk}/\tau_{d(j,k)}}}{e^{(1-\tau_{d(j,k)}) IV(d_j)}} \left[ \sum_{d_j} e^{\tau_{d_j} IV(d_j)} \right]^{-1} \tag{2} \]

where the dissimilarity parameter \( \tau_{d_j} \) is equal to \( \sqrt{1-\rho_{d_j}} \) (Heiss, 2002), \(^{14}\) and the inclusive value \( IV(d_j) \) is defined as:

\(^{12}\)As shown in Bertoli, Fernández-Huertas Moraga, and Ortega (2010), this specification of the stochastic term in (1) entails that we can regard the countries belonging to the same subset of destinations as being close substitutes to each other: a variation in the location-specific utility of a country belonging to a subset produces a larger impact on the distribution of individuals within the subset rather than across subsets.

\(^{13}\)This specification is more general than both Grogger and Hanson (2011), who assume that \( \Sigma_j \) is the product between a scalar and an identity matrix, and Ortega and Peri (2009), provided that \( \#D > 2 \), who assume that all the destinations in \( D \) belong to a unique subset.

\(^{14}\)If a subset \( d_j \) of destinations is a singleton, then \( \tau_{d_j} = 1. \)
\[ IV(d_j) = \ln \left( \sum_{l \in d_j} e^{V_{jl}/\tau_{dj}} \right) \]

The ratio of the probability \( p_{ijk} \) over the probability of opting for the origin country \( j \), \( p_{ijj} \), is given by:

\[
\frac{p_{ijk}}{p_{ijj}} = \frac{e^{V_{jk}/\tau_{d(j,k)}}}{e^{(1-\tau_{d(j,k)})IV[d(j,k)]} e^{V_{jj}/\tau_{d(j,j)}}} \tag{3}
\]

As we have assumed that the origin country forms a singleton in the partition of \( D_j \), we can rewrite (3) as follows:

\[
\frac{p_{ijk}}{p_{ijj}} = \frac{e^{V_{jk}/\tau_{d(j,k)}}}{e^{V_{jj}/\tau_{d(j,j)}}} \left[ e^{(1-\tau_{d(j,k)})IV[d(j,k)]} \right]^{-1} \tag{4}
\]

This ratio depends on the deterministic component of utility in location \( j \) and \( k \), and on the deterministic components of utility in all the locations \( l \in d(j,k) \). Taking the log of (4) and using the definition of the inclusive value, we get:

\[
\ln \left( \frac{p_{ijk}}{p_{ijj}} \right) = \left( \frac{x_{jk}}{\tau_{d(j,k)}} - x_{jj} \right)^\prime \beta - \left( 1 - \tau_{d(j,k)} \right) \ln \left( \sum_{l \in d(j,k)} e^{V_{jl}/\tau_{d(j,k)}} \right) \tag{5}
\]

### 2.2 Migration flows and Multilateral Resistance to Migration

Imagine that individual migration decisions are observed over a set \( T \) of periods; the log of the scale of migration flows to country \( k \) at time \( t \in T \) over the size of the population which opts for the origin country \( j \), \( y_{jkt} \), can be derived from the RUM model by averaging (5) over the set of individuals \( i \). The result is given by:

\[
y_{jkt} = \left( \frac{x_{jkt}}{\tau_{d(j,k)}} - x_{jjt} \right)^\prime \beta + r_{jkt} + \eta_{jkt} \tag{6}
\]

The error term \( \eta_{jkt} \) is orthogonal to \( x_{jkt} \) and \( x_{jjt} \), serially uncorrelated, and independently and identically distributed over the set of origin-destination pairs, and \( r_{jkt} \) is equal to:

\(^{15}\)The assumption of no serial correlation in \( \eta_{jkt} \) is unnecessary for the estimation approach proposed in Pesaran (2006).
\[ r_{jkt} = \left( \tau_{d(j,k)} - 1 \right) \ln \left( \sum_{l \in d(j,k)} e^{V_{jlt}/\tau_{d(j,k)}} \right) \] (7)

We call \( r_{jkt} \) Multilateral Resistance to Migration, as this term captures the influence upon migration from country \( j \) to country \( k \) at time \( t \) exerted by the opportunities to migrate to other destinations. This term is a non-increasing function (as \( \tau_{d(j,k)} \leq 1 \) for all \( j \in H \) and \( k \in D^j \)) of \( V_{jlt} \), the deterministic component of location-specific utility in (1). This effect is strictly related to the close substitutability of destinations belonging to the same nest discussed in Bertoli, Fernández-Huertas Moraga, and Ortega (2010): an increase in \( V_{jlt} \) redirects towards the destination country \( l \) proportionally more individuals that would have opted for destination \( k \in d(j,l) \) than individuals who would have stayed in the country of origin \( j \), thus reducing \( y_{jkt} \). Hence, the scale of bilateral migration flows \( y_{jkt} \) between country \( j \) and country \( k \) is non-increasing (respectively, non-decreasing) in the value of any element of the vector \( x_{jlt} \) which increases (decreases) the utility associated to migration to country \( l \).

The multilateral resistance to migration \( r_{jkt} \) will be, in general, (i) serially correlated, as the resistance to migration exerted by other destinations is likely to evolve slowly over time, (ii) spatially correlated, as the same alternative destination can be regarded as a close substitute for \( k \) by different origins, and (iii) correlated with the regressors. To provide an intuition of the endogeneity problem due to multilateral resistance to migration, consider a likely key macro determinant of the scale of migration flows, namely GDP per capita at origin, which enters the vector \( x_{jlt} \).

GDP per capita at origin \( j \) can correlate with GDP per capita in some of the destination countries - which are included in \( r_{jkt} \); these can occur because of the exposure to common economic shocks, or because of a partial business cycle synchronization due to trade and investment flows.

Consider that \( r_{jkt} \) is unobservable for the econometrician, as it depends (i) on the value of determinants of location-specific utility for countries other than \( j \) and \( k \), (ii) on the (unknown) composition of the set \( d(j,k) \) for any origin-destination dyad second and (iii) on a parameter, \( \tau_{d(j,k)} \), that cannot be identified with aggregate data.\(^{16}\)

\(^{16}\)Bertoli, Fernández-Huertas Moraga, and Ortega (2010) estimate the dissimilarity parameter using individual-level data on Ecuadorian migrants to the United States and Spain.
3 Estimation strategy

The properties of the stochastic term $\epsilon_{ijk}$ in (1) are closely related to the shape of the multilateral resistance to migration term $r_{jkt}$ in (6), which in turn determines which is the appropriate estimation strategy to analyze the determinants of bilateral aggregate migration flows. We consider first the implications of the assumptions about the stochastic term in the random utility maximization model which have been proposed in the literature (Grogger and Hanson, 2011; Beine, Docquier, and Ozden, 2011; Ortega and Peri, 2009), and then we move to the more general assumptions adopted in this paper.

3.1 Grogger and Hanson (2011) and Beine, Docquier, and Ozden (2011)

Grogger and Hanson (2011) and Beine, Docquier, and Ozden (2011) assume that $\epsilon_{ijk}$ follows an Extreme Value Type-1 distribution; this entails that each destination $k \in D_j$ forms a singleton, and:

$$\forall j \in H, k \in D, t \in T : r_{jkt} = 0$$

which, in turn, implies that:

$$y_{jkt} = (x_{jkt} - x_{jkt})'\beta + \eta_{jkt}$$  

as the dissimilarity parameter $\tau$ is equal to 1. The estimation of (8) does not pose specific challenges, as the multilateral resistance to migration term disappears when the stochastic component $\epsilon_{ijk}$ which enters into the individual migration decision problem is such that the independence of irrelevant alternatives holds.$^{17}$

3.2 Ortega and Peri (2009)

Ortega and Peri (2009) lift the assumption that IIA holds in the individual migration decision model, and they assume that all the destinations $D$ belong to the same subset, i.e. $\forall k \in D :$

$^{17}$Grogger and Hanson (2011) verify that the regression coefficients remain stable when destinations are removed from the choice set of prospective migrants, as a violation of the IIA assumption would entail instability of the estimated coefficients (Hausman and McFadden, 1984).
\[ d(j, k) = D. \] Under this assumption, the multilateral resistance to migration term simplifies to:

\[ r_{jkt} = \left( \tau - 1 \right) \log \left( \sum_{l \in D} e^{V_{jlt}/\tau} \right) \]  

(9)

We can observe that (9) is invariant across origin countries, and this allows Ortega and Peri (2009) to estimate (6) controlling for the multilateral resistance to migration through the inclusion of origin-time dummies in the set of regressors. This approach does not allow to identify the effects of any regressor which does not vary across destinations, such as GDP at origin; for instance, it would not allow to identify any of the regressors in the basic specification presented in Clark, Hatton, and Williamson (2007).18

### 3.3 A more general approach

Let us go back to the general specification for the multilateral resistance to migration term in (7), which is reproduced here for convenience:

\[ r_{jkt} = \left( \tau_{d(j, k)} - 1 \right) \ln \left( \sum_{l \in d(j, k)} e^{V_{jlt}/\tau_{d(j, k)}} \right) \]

The equation to be estimated is:

\[ y_{jkt} = \left( \frac{x_{jkt}}{\tau_{d(j, k)}} - x_{jkt} \right) \beta + \varepsilon_{jkt} \]

where:

\[ \varepsilon_{jkt} = r_{jkt} + \eta_{jkt} \]  

(10)

The characteristics of the multilateral resistance to migration \( r_{jkt} \) entail that the error term \( \varepsilon_{jkt} \) in (10) is not well-behaved. Specifically, \( \varepsilon_{jkt} \) is spatially correlated across both origin and destination countries, as it depends on elements that are not dyad-specific. Furthermore, it also exhibits serial correlation, as multilateral resistance to migration evolves smoothly over time. When the error term is serially and spatially correlated, OLS still provides

18The estimation strategy pursued by Ortega and Peri (2009) also entails that, if one wants to control for unobserved destination-specific shocks which influence all incoming bilateral migration flows, then only the effects of dyadic variables, i.e. variables which vary both over destination and over origin countries, can be identified.
consistent estimates of the coefficients $\beta$ (Driscoll and Kraay, 1998), but the standard errors will be incorrect. Driscoll and Kraay (1998) propose an approach to estimate the standard errors of the coefficients which is robust to non-spherical errors, which can be implemented following Hoechle (2007).

Still, the approach by Driscoll and Kraay (1998) addresses only some of the challenges posed by multilateral resistance to migration, as it requires exogeneity of the regressors. The presence of $r_{jkt}$ in the error term gives rise to endogeneity, as:

$$E\left[\left(\frac{x_{jkt}}{\tau_{d(j,k)}} - x_{jjt}\right), \varepsilon_{jkt}\right] \neq 0$$

To provide more intuition about why this occurs, consider the case where visa policy at destination enters the vector $x_{jkt}$, and GDP at origin is one of the elements of $x_{jjt}$. Visa policies - which can exert a substantial influence on the scale of bilateral migration flows (Bertoli, Fernández-Huertas Moraga, and Ortega, 2011) - can be coordinated at the supranational level. For instance, the list of third countries whose nationals need a visa to enter the European Union is determined by the European Council: when a country is included in this list, a simultaneous change in the bilateral visa policies towards this country adopted by EU member states is observed. As far as EU countries are perceived as close substitutes by would-be migrants from third countries, we have that $x_{jkt}$ correlates with $r_{jkt}$.

This entails that we need an estimator that is also able to handle the endogeneity of the regressors.\footnote{The use of external instruments is hardly an option here, as endogeneity is not confined to a regressor, but to all relevant determinants of the scale of migration flows.}

Following Pesaran (2006), we aim at controlling for the unobservable multilateral resistance to migration term with a dyad-specific linear combination of cross-sectional averages of the dependent and independent variables. Let $q_{jlt} = e^{V_{jlt}/\tau_{d(j,k)}}$, and let us define the dyad-specific average of $q_{jlt}$ as follows:

$$\tilde{q}_{jl} = \frac{\sum_{t \in T} q_{jlt}}{T}$$

Using a Taylor expansion around $\tilde{q}_{jl}$, we can approximate the multilateral resistance to migration term $r_{jkt}$ introduced in (7) as:
\[ r_{jk}t \approx \tilde{r}_{jt} + (\tau_{d(j,k)} - 1) \sum_{l \in d(j,k)} \frac{(q_{jlt} - \tilde{q}_{jl})}{\sum_{l \in d(j,k)} \tilde{q}_{jl}} \]  \hspace{1cm} (11) 

where \( \tilde{r}_{jk} \) is the value of multilateral resistance to migration in correspondence the average values \( \tilde{q}_{jl} \), for \( l \in d(j,k) \). Now, let us define:

\[ \gamma_{jk} = I(j, l, k)(\tau_{d(j,k)} - 1) \]

where \( I(j, l, k) \) is an indicator function which takes value 1 if \( l \in d(j,k) \), and 0 otherwise. Using this notation, we can rewrite (11) more compactly as follows:

\[ r_{jk}t \approx \tilde{r}_{jt} + \sum_{l \in D} \gamma_{jk} (q_{jlt} - \tilde{q}_{jl}) \sum_{l \in d(j,k)} \tilde{q}_{jl} \]  \hspace{1cm} (12)

We define the vector \( q_{jt} \) as:

\[ q_{jt} = \left( \frac{(q_{j1t} - \tilde{q}_{j1})}{\sum_{l \in d(j,1)} \tilde{q}_{jl}}, \frac{(q_{j2t} - \tilde{q}_{j2})}{\sum_{l \in d(j,2)} \tilde{q}_{jl}}, \ldots \right)' \]

and the vector \( q_t \) as \( (q_{1t}', q_{2t}', \ldots)' \).\(^{20}\) Let also the vector \( \gamma_{jk} \) be defined as:

\[ \gamma_{jk} = \left( J(j, 1)\gamma_{11}, J(j, 1)\gamma_{12}, \ldots, J(j, 2)\gamma_{21}, J(j, 2)\gamma_{22}, \ldots \right)' \]

where \( J(j, i) \) is an indicator function which takes value 1 if \( j = i \), and 0 otherwise.\(^{21}\) Using this vector notation, (12) can be rewritten as follows:

\[ r_{jk}t \approx \tilde{r}_{jt} + \gamma_{jk}' q_t \]  \hspace{1cm} (13)

Using (13), we can rewrite the equation to be estimated as:

\[ y_{jk} = \frac{\left( x_{jk}t - \tilde{x}_{jkt} \right)}{\tau_{d(j,k)}}' \beta + \tilde{r}_{jk} + \gamma_{jk}' q_t + \eta_{jk} \]  \hspace{1cm} (14)

The error term \( \gamma_{jk}' q_t + \eta_{jk} \) in (14) follows the multifactor error model described in Pesaran (2006), as the linear approximation of the multilateral resistance to migration is given by

\(^{20}\) The vector \( q_t \) has a number of elements equal to the number of origin-destination dyads.

\(^{21}\) The number of non-zero elements in the vector \( \gamma_{jk} \) is equal to the number of destinations belonging to the nest \( d(j,k) \).
the inner product of a vector of dyad-specific factor loadings, \(\gamma_{jk}\), and the a vector of time-specific common factors \(q_t\).\(^{22}\) This structure reflects the effect of common factors that exert an uneven influence on the various panels: a variation in the attractiveness of location \(l \in D\) can influence the scale of migration between any dyad of origin and destination countries, with the influence on some dyads being possibly zero.

The multifactor error model described in Pesaran (2006) does not impose limits on the (finite) number of unobserved elements in the vector \(q_t\), nor it requires this number to be known. Pesaran (2006) demonstrates that \(\gamma'_{jk} q_t\) can be expressed as a dyad-specific linear combination of the cross-sectional averages of the dependent and of the independent variables. Specifically, he demonstrates that a consistent estimate of \(\beta\), \(b_{CCE}\), can be obtained from the estimation, through OLS, of the following regression:

\[
y_{jkt} = \left( \frac{x_{jkt}}{\tau_{d(j,k)}} - x_{jjt} \right)' \beta + d_{jk} + X'_{jk} \tilde{z}_t + \eta_{jkt}
\]

where \(d_{jk}\) is a dummy for the dyad \((j, k)\), and the vector of auxiliary regressors \(\tilde{z}_t\) is equal to:

\[
\tilde{z}_t = \frac{1}{\sum_{(j,k)} \omega_{jkt}} \left( \sum_{(j,k)} \omega_{jkt} y_{jkt}, \sum_{(j,k)} \omega_{jkt} x'_{jkt}, \sum_{(j,k)} \omega_{jkt} x'_{jjt} \right)'
\]

and \(\omega_{jkt}\) is the weight assigned to each origin-destination dyad at time \(t\) in the estimation. Pesaran (2006) refers to this estimator as the common correlated effects (CCE) estimator; consistency of \(b_{CCE}\) follows from the fact that \(X'_{jk} \tilde{z}_t\) converges in probability to \(\gamma'_{jk} q_t\) as the cross-sectional dimension of the panel goes to infinity (Pesaran, 2006). Monte Carlo simulations in Pesaran (2006) also show the good finite sample properties of the CCE estimator, which produces satisfactory results already when \(N = 30\) and \(T = 20\).

4 Data and descriptive statistics

Our dataset has three main components: migration flows to Spain in the 1997-2009 period; migration policies in Spain during the same period; and quarterly real GDP series for the countries of origin of migrants to Spain. Here, we first present each of these components and we look at their main characteristics, then we provide the relevant descriptive statistics

\(^{22}\)Bai (2009) refers to the same structure of the error term as “interactive fixed effects”.

4.1 Migration flows

The migration flows data come from the *Estadística de Variaciones Residenciales* (EVR). This is an administrative dataset collected by the Spanish *Instituto Nacional de Estadística* (INE). The EVR gathers all the variations in the municipal registry (*Padrón Municipal de Habitantes*) throughout the year: each observation in the EVR corresponds either to an inscription in or to a cancelation from the *Padrón*, and it includes information on the date in which the variation occurred, and on the age, gender and country of birth of the individual to whom the variation refers to. We use the observations referring to the first inscription of foreign-born individuals coming from abroad in the *Padrón* to measure immigration flows to Spain: the EVR contains 6,166,133 of these observations between January 1997 and December 2009,\(^{23}\) related to individuals from 208 countries of origin.\(^{24}\)

By restricting our attention to inscriptions of foreign-born individuals coming to Spain from abroad, we are obtaining an almost perfect measure of gross immigration inflows. The measure would be perfect if every individual registered immediately upon arrival. Although registration is not mandatory, most immigrants eventually do register, independently of their legal status, as registration gives them access to all basic municipality services, most notably free health care and education (Bertoli, Fernández-Huertas Moraga, and Ortega, 2011). The Appendix A.1 discusses in detail the accuracy of the EVR in measuring immigration flows to Spain, comparing EVR figures with those that can be obtained from alternative data sources.

Figures 1 and 2 plot the monthly and quarterly series of immigration flows to Spain over our period of analysis according to the EVR. Despite the large apparent variability in the overall immigration series, there does not seem to be relevant seasonal patterns in the data. None is found if we regress quarterly data on year and quarter dummies: the quarterly dummies are not significant. For the monthly data, a regression on year and month dummies shows the months of August and December as those in which registrations are significantly lower (between 15 and 20 percent) than in the rest of the year, coinciding with the summer and winter holidays in Spain.

There are three noticeable spikes in the series: the first one corresponds to the January

---

\(^{23}\) As recalled in the introduction, these figures correspond to an unprecedented - even from an international perspective (OECD, 2010) - surge in immigration.

\(^{24}\) The EVR also codifies some former states, such as the USSR or Yugoslavia.
Figure 1: Monthly Immigration Inflows to Spain 1997-2009 (EVR)

![Graph showing monthly immigration inflows to Spain 1997-2009 (EVR)]

Figure 2: Quarterly Immigration Inflows to Spain 1997-2009 (EVR)

![Graph showing quarterly immigration inflows to Spain 1997-2009 (EVR)]
Figure 3: Total flows and excluding immigrants from Bulgaria and Romania (1997-2009)

2000 law that ensured access to basic services for those registered; the second one can be associated to the 2005 massive amnesty and happened in November 2004; finally, the third one has to do with the accession of Romania and Bulgaria to the EU in January 2007, taking into account that Romanians have created the largest immigrant community in Spain (see Figure 3 for the evolution of total flows excluding the two most recent EU member states).

Our analysis aggregates the EVR data at the quarterly level, as this is the finest period of time for which we can gather information on the economic conditions at origin. We restrict our sample to the origin countries with a positive total number of immigrants in all the 52 quarters included in our period of analysis: 98.6 percent of total migration flows to Spain between January 1997 and December 2009 originated from these countries, whose population represents 86 percent of the world total.

In our empirical analysis below, our dependent variable will be the log of the emigration rate to Spain from a given origin country over a quarter, consistently with the model presented in Section 2. This is calculated as the total number of immigrants to Spain from

---

25 The share of the observations where the recorded migration flow is equal to zero is much lower than in the dataset employed by Beine, Docquier, and Ozden (2011), where it stands at 36 percent; Beine, Docquier, and Ozden (2011) assess the sensitivity of their estimates to the inclusion of these zero observations, and they validate the estimates obtained from the specifications where these observations are dropped from the sample as “results are highly robust to various econometric techniques accounting for the large proportion of zeros”. A similar conclusion is reached also by Grogger and Hanson (2011).

26 The EVR would allow us to build gender- or age-specific measures of immigration flows, but this choice turns out to be immaterial: the correlation between total and male flows stands at 0.989, while the correlation
origin country $j$ who registered during a given quarter divided by the population of that country of origin $j$ in that year.\footnote{Our population figures are taken from the World Development Indicators (World Bank, 2010), and vary only at the yearly level.}

### 4.2 Spanish migration policies

We gather data on Spanish migration policies between 1997 and 2009; specifically, we codify the following policies which are likely to influence bilateral migration flows in the EVR: \((i)\) general policies - the 2000 Amnesty, the 2005 Amnesty; \((ii)\) bilateral policies - visa agreements, double nationality agreements, social security agreements, agreements on the signature of labor contracts at origin; and \((iii)\) multilateral treaties - membership to the EU-15, membership to the Schengen area, 2004 EU enlargement, 2007 EU enlargement. The Appendix A.2 describes the definition and sources of these variables.

Our database comprises 8 EU-wide agreements transposed into Spanish Law through Decrees,\footnote{The EU enlargement to 25 members that applied from May 1, 2004 is one such entry in our database} 48 national Laws, Resolutions and Orders dealing with migration issues,\footnote{These include, for example, the 2005 amnesty that applied from February 7, 2005 to May 7, 2005.} and 94 bilateral agreements between Spain and origin countries regarding matters such as the need of a visa to enter Spain, portability of social security benefits, legal recognition of educational degrees, etc. We have taken the data from the web pages of the Ministry for Labor and Immigration and the Boletín Oficial del Estado, a daily official bulletin where all Spanish legislation is published.

We model these migration policies as dummy variables that change from 0 to 1 from the month the policy is applied. For instance, the 2000 Amnesty is modeled as a 0 before January 2000 and as a 1 afterwards. Another example, already studied by Bertoli, Fernández-Huertas Moraga, and Ortega (2011) is the bilateral agreement between Ecuador and Spain regarding the need of a visa for Ecuadorians to enter Spain. We model this as a dummy taking value 1 when a visa is needed to enter Spain and value 0 otherwise. In the Ecuadorian case, this means the value of the visa dummy is 0 before August 2003 and 1 after that date. We present a more detailed description of the construction of the dataset in the Appendix.
This set of ten variables is able to explain, in a simple OLS regression, up to 54 percent of the total variation on the log of the monthly or quarterly emigration rates to Spain by country of origin. This shows that our migration policy specification has a good deal of variability and potential explanatory power.

### 4.3 Economic conditions at origin

Our estimation strategy requires the use of high-frequency data, and we were able to gather quarterly real GDP data for 61 origin countries, representing 87 percent of total migration flows to Spain over the 1997-2009 period. As detailed in the Appendix A.3, our data sources are the International Financial Statistics (IMF, 2010a), the April 2010 issue of the World Economic Outlook (IMF, 2010b) and the data published by some Central Banks.

We divide our quarterly real GDP series by the yearly population figures from the World Development Indicators (World Bank, 2010) to obtain real GDP per capita series that we use as a proxy for the time-varying economic conditions at origin. Since the series vary widely in terms of base year, adjustments on seasonality, base currency and other aspects, we construct a country-specific seasonally-adjusted real GDP per capita index (setting the index equal to 100 in the first quarter of 2000). The raw correlation between the log of the GDP per capita index by quarter and country of origin and the log of the emigration rate to Spain is 0.05. In a simple regression of the two variables, the coefficient on the GDP per capita index is 0.8 and is only able to explain 0.3 percent of the variation in quarterly emigration rates.

### 4.4 Summary statistics

When combining our migration flows, migration policies and real GDP per capita datasets, we are left with 3,020 observations. Out of the 6,166,133 immigrants who, according to the EVR, entered Spain between January 1997 and December 2009 coming from 208 countries, we keep in our sample 5,341,586 immigrants coming from 61 countries, which host 51 percent of the world population. Figure 4 shows that these 61 countries keep the basic time series structure of the overall number of immigrants.

We present in Table 1 some summary statistics of this emigration rate (expressed in
In order to allow a straightforward comparison, we also construct a country-specific index for emigration rates. We weight observations by the population of the country of origin since we are interested in exploring determinants of emigration rates over the whole population.

Table 1 shows that the variability is much more substantial in the emigration rate than in the GDP per capita during the period. The mean emigration rate per quarter to Spain was 32.88 emigrants per 1,000,000 inhabitants with a maximum in the sample of 3,099 emigrants in the first quarter of 2007 from Romania and a minimum of 0.01 in the first quarter of 1997 from Indonesia. For the country-specific index, the average of 268 reflects the growth in migration rates from 2000. The relative maximum (15,740) corresponds to Paraguay in the first quarter of 2007 whereas the minimum (0.30) is Ecuador in the first quarter of 1997. For the GDP per capita index, the average value (weighted by population) in the sample is 115 with a minimum of 70 for Venezuela in the first quarter of 2003 and a maximum of 223 for Georgia in the second quarter of 2008. We can observe the scatter-plot of the log of both indexes in Figure 5.

It is perhaps more informative to look directly at the time series evolution of the variables the way they will be used in the empirical analysis below. The following series of figures present this representation for the log of the emigration rate and the log of real GDP per capita for the four top emigrant sending countries to Spain during the period: Romania (809,857 emigrants), Morocco (666,798 emigrants), Ecuador (490,580 emigrants) and
Table 1: Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>mean</th>
<th>s.d.</th>
<th>min</th>
<th>max</th>
<th>obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emigrants to Spain per 1,000,000 inhabitants</td>
<td>32.88</td>
<td>136.75</td>
<td>0.01</td>
<td>3,098.78</td>
<td>3,020</td>
</tr>
<tr>
<td>Emigration rate index (2000q1=100)</td>
<td>267.58</td>
<td>381.83</td>
<td>0.30</td>
<td>15,470.04</td>
<td>3,020</td>
</tr>
<tr>
<td>Real GDP per capita index (2000q1=100)</td>
<td>115.17</td>
<td>19.91</td>
<td>69.61</td>
<td>223.34</td>
<td>3,020</td>
</tr>
<tr>
<td>January 2000 Amnesty</td>
<td>0.83</td>
<td>0.37</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>November 2004 Amnesty</td>
<td>0.44</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>EU-15</td>
<td>0.11</td>
<td>0.31</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>Schengen Area</td>
<td>0.09</td>
<td>0.28</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>EU May 2004 Eastern Enlargement</td>
<td>0.01</td>
<td>0.10</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>EU May 2007 Romania and Bulgaria Enlargement</td>
<td>0.002</td>
<td>0.05</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>Visa requirement</td>
<td>0.57</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>Bilateral Agreement on Nationality</td>
<td>0.05</td>
<td>0.23</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>Bilateral Agreement on Social Security</td>
<td>0.13</td>
<td>0.33</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
<tr>
<td>Bilateral Agreement on Contracts at Origin</td>
<td>0.02</td>
<td>0.13</td>
<td>0</td>
<td>1</td>
<td>3,020</td>
</tr>
</tbody>
</table>

Note: quarterly series on 61 countries (1997-2009), all descriptive statistics are weighted by population at origin; see the Appendix A.2 for a description of the immigration policy variables.

Figure 5: Emigration and GDP at origin, selected sample (1997-2009)
Figures 6 to 9 show that, despite a general upward time trend in most of the series that the empirical analysis will have to account for, there is substantial time and cross-sectional variation to be exploited in the dataset.

5 Econometric analysis

The econometric analysis of the determinant of bilateral migration flows to Spain over 1997-2009 follows the steps entailed by the estimation strategy outlined in Section 3. We report
Figure 8: Emigration and GDP at origin, Ecuador

Figure 9: Emigration and GDP at origin, Colombia
here the equation to be estimated, derived on the basis of the RUM model presented in Section 2:

\[ y_{jkt} = \left( \frac{x_{jkt}}{\tau_{d(j,k)}} - x_{jjt} \right) \beta' + r_{jkt} + \eta_{jkt} \]

Consistently with the model, the dependent variable \( y_{jkt} \) is represented by the log of the quarterly migration rate to Spain for each of the 61 origin countries included in our sample. The vector \( x_{jkt} \) contains a number of dyad-specific elements, represented by the bilateral immigration policies and multilateral treaties described in Section 4.2, while we control for all origin-invariant factors - such as the level of GDP or unemployment at destination - through the inclusion of quarter fixed effects, and for all time-invariant factors - such as cultural or linguistic proximity - through origin fixed effects. The vector \( x_{jjt} \) includes (various lags of) the log of real GDP per capita at origin, and origin-year fixed effects to control for all unobserved origin- and dyad-specific time-varying determinants of bilateral migration flows.\(^{30}\)

With respect to our measure of GDP at origin, we include lagged values given that we have high-frequency migration data, and one can reasonably assume that would-be migrants do not react instantaneously to changes in economic conditions at origin. We relied on the Akaike and Bayesian Information Criteria and on Likelihood Ratio tests in order to select the optimal lag structure for each specification as suggested in Canova (2007), thus avoiding \textit{ad hoc} choices. The optimal number of lags selected was four with all methods.

As a first step, we assume, as in Grogger and Hanson (2011) and Beine, Docquier, and Ozden (2011) that the stochastic term in the individual location-specific utility follows an Extreme Value Type-I distributions, so that multilateral resistance to migration disappears, and (14) simplifies to:\(^{31}\)

\[^{30}\text{The origin-year fixed effects also render our GDP per capita and emigration rate series stationary although the CCE estimator can accommodate unit roots.}\]

\[^{31}\text{Observe that, as multilateral resistance to migration } r_{jkt} \text{ does not enter the equation to be estimated, endogeneity should not be a pressing concern here: some of the crucial facets of the Spanish policy stance towards immigration are determined at the EU level and bilateral migration flows to Spain can be expected to exert only a very limited - if any - impact on economic conditions at origin. Remember that the largest emigration rate in our sample is 0.3 percent of the Romanian population in the first quarter of 2007. The median emigration rate in the sample is just 0.0002 percent.}\]
$$y_{jkt} = (x_{jkt} - x_{jkt})' \beta + \eta_{jkt}$$ (16)

This equation is estimated with a two-way error component model, and the results are presented in the first data column in Table 2. The model controls for origin-year fixed effects. The inclusion of this very rich structure of fixed effects allows us to control for those determinants of migration, such as demographic factors (Hanson and McIntosh, 2010a,b) or migrant networks (Munshi, 2003; Edin, Fredriksson, and Äslund, 2003; McKenzie and Rapoport, 2010; Beine, Docquier, and Ozden, 2011; Bertoli, 2011), which evolve at a pace that is slower than the frequency of our panel data. This substantially reduces the variability in the data that we are exploiting to identify the coefficient vector $\beta$ but we are still able to precisely estimate the effect of GDP variations on migration decisions.

According to the first data column in Table 2, a 1.0 percent increase in real GDP per capita leads, after four quarters, to a 4.7 percent reduction in the migration rate to Spain.32

The estimates from this specification are consistent as long as multilateral resistance to migration does not influence bilateral migration flows to Spain. From Section 3, we know that this would induce spatial and serial correlation in the error term, and we follow Frees (1995) and Wooldridge (2002) to test for the presence of cross-sectional dependence and an autoregressive structure in the residuals.33

Table 2 shows that the null hypotheses of both tests are strongly rejected,34 and this suggests that bilateral migration flows to Spain could be influenced by multilateral resistance to migration. This entails that the standard errors provide an incorrect basis for inference, and we re-estimated the same specification resorting to the method proposed by Driscoll and Kraay (1998) to obtain standard errors which are robust to serial and cross-sectional dependence in the error term.35 The estimates in the second data column in Table 2 show

32 Note that this effect is notably larger than that found by Clark, Hatton, and Williamson (2007) for US immigration, which stands at 0.44; differently from them, we consider both legal and illegal immigration and exploit within-year variability in GDP. Our country-year fixed effects allow us to control for a much wider set of possible confounding factors that evolve slowly over time.

33 We opted for the test for cross-sectional dependence proposed by Frees (1995) over the alternative test proposed by Pesaran (2004) as the latter could lack power and “miss out cases of cross-sectional dependence where the sign of the correlations is alternating” (De Hoyos and Sarafidis, 2006), as the multilateral resistance to migration needs not to be positively correlated across different countries of origin.

34 The two tests are implemented following De Hoyos and Sarafidis (2006) and Drukker (2003) respectively.

35 The method by Driscoll and Kraay (1998) is implemented following Hoechle (2007).
Table 2: Determinants of migration

Dependent variable: log of quarterly emigration rate

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation methods</td>
<td>FE</td>
<td>FE</td>
<td>CCE</td>
</tr>
<tr>
<td>Regressors</td>
<td>Lags</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log real GDP per capita</td>
<td>1</td>
<td>-1.57</td>
<td>-1.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.24]***</td>
<td>[0.52]***</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1.05</td>
<td>-1.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.26]***</td>
<td>[0.40]**</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-0.93</td>
<td>-0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.27]***</td>
<td>[0.52]*</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-1.18</td>
<td>-1.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.26]***</td>
<td>[0.62]*</td>
</tr>
<tr>
<td>Visa requirement</td>
<td>0</td>
<td>-0.15</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.13]</td>
<td>[0.23]</td>
</tr>
<tr>
<td>Other migration policy controls</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Quarter fixed effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Origin-year fixed effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2,776</td>
<td>2,776</td>
<td>2,776</td>
</tr>
<tr>
<td>Countries of origin</td>
<td>61</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Frees’ test (p-value)</td>
<td>13.19 (0.00)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wooldridge’s test (p-value)</td>
<td>24.17 (0.00)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cross-sectional averages (p-value)</td>
<td>-</td>
<td>-</td>
<td>2.11 (0.00)</td>
</tr>
<tr>
<td>GDP per capita, cumulated effect</td>
<td>1</td>
<td>-1.57</td>
<td>-1.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.24]***</td>
<td>[0.52]***</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-2.63</td>
<td>-2.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.28]***</td>
<td>[0.51]***</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-3.56</td>
<td>-3.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.33]***</td>
<td>[0.74]***</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-4.74</td>
<td>-4.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.37]***</td>
<td>[1.03]***</td>
</tr>
</tbody>
</table>

Notes: standard errors in brackets; *** p<0.01, ** p<0.05, * p<0.1; observations are weighted by population at origin; the number of lags of log real GDP per capita has been determined according to AIC, BIC and LR tests to identify the optimal lag structure following Canova (2007); specification (2) includes standard errors computed following Driscoll and Kraay (1998); for specification (3), we present an F-test that the coefficients on cross-sectional averages in the CCE estimator are jointly zero, calculated on F(659,1,378).
that income at origin remains a significant determinant of bilateral migration flows - though the correction by Driscoll and Kraay (1998) substantially inflates its standard error, while the effect of the visa policy is still not significant. In Grogger and Hanson (2011), the effect of the visa waiver was marginally significant.

Still, these estimates are biased and inconsistent as multilateral resistance to migration is likely to make the regressors endogenous, as discussed in Section 3. Before resorting to the CCE estimator proposed by Pesaran (2006), which we have shown to be well-suited to address this specific form of endogeneity, it is interesting to consider the expected direction of the bias induced by multilateral resistance to migration with respect to the estimated coefficients of the GDP at origin and of the visa policy.

If real GDP per capita at origin correlates positively with real GDP per capita in some destinations that would-be migrants perceive as close substitutes to Spain, then the coefficient estimated in specifications (1)-(2) in Table 2 is downward biased. This occurs because an increase in GDP at origin is associated with an improvement in the attractiveness of other alternative destinations: if this is not controlled for, then the estimated effect of GDP at origin also captures the reduction in migration flows to Spain due to the increased attractiveness of other destination countries. This might be the relevant case with our dataset: Bertoli, Fernández-Huertas Moraga, and Ortega (2010) provided evidence that prospective migrants from the third largest origin country, Ecuador, regard Spain and the US as close substitutes, and the correlation between real GDP per capita in Ecuador and in the US stands at 0.57 once origin-year and quarter fixed effects are controlled for.  

A similar line of reasoning suggests that the coefficient of the visa requirement estimated in specifications (1)-(2) in Table 2 is upward biased. A change in the Spanish visa policy towards one origin country occurs when also the other EU member states are adopting an identical change, when this decision follows a regulation by the European Council. An instance of such a change occurred in March 2001, when the citizens of the countries which were candidate to accession at that time were granted visa-free access to the EU by the EC Regulation No. 539/2001.  

---

36 The correlation between real per capita GDP in Colombia, the fourth largest origin country, and the US stands at 0.76, strengthening the expectation about the direction of the bias from neglecting multilateral resistance to migration. The correlations between other top destinations and their main alternatives also go in the same direction: 0.40 of Romania with the US and 0.48 of Morocco with France.

37 The sample countries to which the EC Regulation No. 539/2001 applied to are Bulgaria, Czech Republic,
migrate to other EU destinations: if this effect is not controlled for, then the estimated effect of the Spanish visa requirement also captures the increase in migration flows to Spain due to changes in the visa policy in other member states, biasing the negative coefficient upwards.\footnote{Similar arguments can be applied to other variables that we control for, such as those referring to the 2004 and 2007 EU enlargements.}

The third data column in Table 2 presents the estimates obtained from the Common Correlated Effects estimator proposed by Pesaran (2006). As shown in the bottom panel of Table 2, the cross-sectional averages of the dependent and independent variables which are introduced as auxiliary regressors are jointly significant, which is required for the estimator to be valid.\footnote{We do not test for the serial correlation of the residuals as the CCE estimator is consistent also when \( \eta_{jkt} \) in (16) is serially correlated.}

With the CCE estimator, we find that a 1.0 percent increase in real GDP per capita leads, after four quarters, to a 3.1 percent reduction in the migration rate to Spain. This effect is only 66 percent of the one estimated in specifications (1)-(2), confirming in this case the expectation that neglecting the influence of the multilateral resistance to migration biases the coefficient of GDP downwards. Similarly, the estimated negative coefficient of the visa requirement is now highly significant, and much larger than the one obtained in the previous specifications. The introduction of a visa requirement for non-immigrant admission to Spain reduces the size of migration flows by 74 percent. This can be compared with the effects from specifications (1)-(2), which pointed to a much smaller (14 percent) and non significant reduction of migration flows. The CCE large estimated effect is in line with the findings on Ecuadorian migration to Spain in Bertoli, Fernández-Huertas Moraga, and Ortega (2011).

6 Concluding remarks

The possible dependence of bilateral migration flows upon the time-varying attractiveness of other destinations represents a source of concern for the econometric analysis of the determinants of migration (Hanson, 2010), as it can introduce an omitted variable bias. This paper has explored the relationship between the stochastic properties of the individual migration decision problem, and the presence of such a bias: when the independence of irrelevant alternatives does not characterize individual migration choices, then bilateral aggregate flows

\[ \text{migration to other EU destinations: if this effect is not controlled for, then the estimated effect of the Spanish visa requirement also captures the increase in migration flows to Spain due to changes in the visa policy in other member states, biasing the negative coefficient upwards.} \]

The third data column in Table 2 presents the estimates obtained from the Common Correlated Effects estimator proposed by Pesaran (2006). As shown in the bottom panel of Table 2, the cross-sectional averages of the dependent and independent variables which are introduced as auxiliary regressors are jointly significant, which is required for the estimator to be valid.

With the CCE estimator, we find that a 1.0 percent increase in real GDP per capita leads, after four quarters, to a 3.1 percent reduction in the migration rate to Spain. This effect is only 66 percent of the one estimated in specifications (1)-(2), confirming in this case the expectation that neglecting the influence of the multilateral resistance to migration biases the coefficient of GDP downwards. Similarly, the estimated negative coefficient of the visa requirement is now highly significant, and much larger than the one obtained in the previous specifications. The introduction of a visa requirement for non-immigrant admission to Spain reduces the size of migration flows by 74 percent. This can be compared with the effects from specifications (1)-(2), which pointed to a much smaller (14 percent) and non significant reduction of migration flows. The CCE large estimated effect is in line with the findings on Ecuadorian migration to Spain in Bertoli, Fernández-Huertas Moraga, and Ortega (2011).

6 Concluding remarks

The possible dependence of bilateral migration flows upon the time-varying attractiveness of other destinations represents a source of concern for the econometric analysis of the determinants of migration (Hanson, 2010), as it can introduce an omitted variable bias. This paper has explored the relationship between the stochastic properties of the individual migration decision problem, and the presence of such a bias: when the independence of irrelevant alternatives does not characterize individual migration choices, then bilateral aggregate flows...
depend on the opportunities to migrate to other countries, and we labeled this effect Multi-
ilateral Resistance to Migration.

Consistent estimates of the determinants of bilateral migration flows can be obtained in
the presence of multilateral resistance to migration adopting the Common Correlated Effects
estimator proposed by Pesaran (2006). This approach is more general than others proposed
in the literature, which either rely on an *ad hoc* controls for the time-varying opportunities to
migrate to other destinations (Mayda, 2010), or require more restrictive assumptions on the
stochastic properties of the model and do not allow to identify the effects of origin-specific
variables (Ortega and Peri, 2009).

This approach is applied to the analysis of high-frequency Spanish administrative data
on bilateral migration flows between 1997 and 2009, which are found to respond quickly
and significantly to variations in economic conditions at origin, and to changes in the le-
gal provisions for non-immigrant admission. The econometric analysis shows the empirical
relevance of the concern expressed by Hanson (2010) in our data: if not accounted for, mul-
tilateral resistance to migration would bias upwards the estimated effect of GDP at origin
and downwards the effect of visa policies upon bilateral migration flows to Spain.

References


1229–1279.

Series No. 594, Frankfurt.


Economics*, 95(1), 30–41.


A Data sources

A.1 Migration flows

A.1.1 The Estadística de Variaciones Residenciales

The EVR is an administrative dataset: municipalities are responsible for keeping the municipal registry up to date and the INE just compiles the information received from the municipalities about variations in the Padrón. The EVR registers changes of status in the Padrón, both inscriptions and cancelations, with each observation corresponding to one variation. We use the observations referring to the first inscription of foreign-born individuals coming from abroad in the Padrón to measure immigration flows to Spain: the EVR contains 6,166,133 of these observations between January 1997 and December 2009.

We can assess the accuracy of the EVR in measuring immigration flows to Spain by comparing it with other possible sources. These alternative sources are represented by the 2001 Population Census, the 2007 Encuesta Nacional de Inmigrantes (ENI), a special survey for foreign-born individuals, and various rounds of the quarterly Spanish labor survey, Encuesta de Población Activa (EPA). The ENI was a special immigrant survey which was only ran once between the last months of 2006 and the first months of 2007, with a sample of approximately 15,000 immigrants. The ENI and the EPA provide information about the year of arrival to Spain of all immigrants, although the EPA does not contain this information for the foreign-born who obtained Spanish citizenship.

A.1.2 Total yearly flows

Our EVR data span the 1997-2009 period. Figure A.1 compares gross immigration flows into Spain according to the EVR with the 2001 Population Census and the ENI, and it shows that the EVR underestimates migration flows before 2000. In January 2000, the Spanish government enacted a new immigration law which included both an amnesty and a provision guaranteeing that immigrants would have access to basic public services such as health and

\[40\] The methodology to locate immigrants was based on past Padrón data and it is exactly the same methodology used by the EPA.

\[41\] This entails that we have to take the EPA numbers as a lower bound, as one out of six immigrants residing in Spain in 2010 was naturalized.

\[42\] Recall that, the 2007 ENI being a survey, it might fail to enumerate recently arrived immigrants.
education for their children as long as they register in the Padrón (see Bertoli, Fernández-Huertas Moraga, and Ortega (2011)). This shows as a spike in the 2000 EVR data that can be attributed to the registration of both newly arrived immigrants and of those who had come to Spain in the earlier years but had not registered yet. The EVR and ENI series then pretty much coincide for the years 2002 and 2003 but they diverge again for 2004 and 2005 (we do not include the 2006 ENI arrivals because the survey was administered partly in the last months of 2006). There are three possible explanations for the 2004 and 2005 divergence: (i) the ENI might be underestimating the number of newly arrived immigrants because of a sampling problem, and of the bias due to the emigration of the foreign-born; (ii) the 2005 amnesty may have induced more and more illegal immigrants to register; and (iii) the 2004 EU enlargement may have made immigrants from Eastern Europe register massively, even though they may have arrived much earlier.

First, let us consider the likely magnitude of sampling problems in the 2007 ENI. As the methodology to locate immigrants for the ENI is exactly the same methodology used for the EPA, a comparable dataset would be the EPA for the first quarter of 2007, which interviewed around 10,500 immigrants. Figure A.2 shows the implied immigration flows to Spain for different rounds of the first quarter of the EPA between 2006 and 2010. The

---

35

---

43 This may have been also helped by the 2000 and 2001 amnesties, even though being legally in the country is irrelevant for the registration.

44 Incidentally, the 2007 EU enlargement to Romania and Bulgaria is behind the great surge in immigration flows in 2007, as shown in Section 4.1.
standard errors for the ENI and EPA numbers are between 10,000 and 25,000 immigrants. Hence, what Figure A.2 shows is that sampling problems in the ENI could go on average around half of the way in explaining the difference between the ENI and the EVR. Taking the year 2005, where the discrepancy between the EVR and the ENI is greatest (with 700,000 and 300,000 immigrants respectively), the most recent rounds (2009 and 2010) of the EPA report that 500,000 immigrants arrived.

What about the bias due to the emigration of the foreign-born? The unique data source is represented by the EVR itself, as Figure A.2 shows that there is little hope of gauging the size of return migration or re-migration to third countries from the comparison of different rounds of the EPA. A problem is represented by the fact that it is not mandatory to cancel from the Padrón before leaving the country, and the law, while making inscription attractive, does not provide incentives to cancel registrations, and this entails that many episodes of emigration of foreign-born individuals are likely to remain unreported. Fortunately, the law changed in November 2003 (Ley Orgánica 14/2003) so that non-EU immigrants (which represented around 60 percent of the total immigrant population in 2005) must renew their inscription every two years, otherwise they are removed from the Padrón, with the corresponding variation being recorded in the EVR.\footnote{The EVR does not provide information on the country of destination of the foreign-born who do cancel from the Padrón, and this is why we do not refer to these variations as instances of return migration.}

Thus, reliable EVR estimates of the emigration of the foreign-born should be available since 2006. It must be noted though that these figures may be reliable in terms of magnitude
(i.e. every observation corresponds to an instance of emigration of a foreign-born), but not necessarily in terms of timing since there could be, at most, a two-year lag between the actual departure and the variation recorded in the EVR. The analysis is further complicated by the fact that the EVR does not provide the information about the date of the first inscription of the foreign-born who cancel from the Padrón. Taking all of this into account, Figure A.3 shows the yearly figures of the emigration of the foreign-born according to the EVR.

If we assume, because of the two-year delay, that outflows recorded in 2006 (the first year to which the new law applies) correspond to actual departures in 2004, they would represent 18 percent of the 2004 gross inflow, whereas 2007 outflows correspond to 30 percent of the 2005 gross inflow. Given the uncertainty about the timing of the flows, all that can be said is that emigration could potentially go a large part of the way in explaining the discrepancies between the EVR figures for 2004 and 2005, and the corresponding figures from the ENI. Together with the sampling design problem, emigration of the foreign-born could even go all the way in explaining the observed difference.

With respect to point (iii) above, we can safely disregard the role of the 2004 EU enlargement for the 2004 difference in flows. There are two reasons for this: first, none of the enlargement countries accounts for a relevant share of immigration to Spain, with less than 14,000 immigrants in total (2 percent of the 2004 inflow); second, Spain - unlike Sweden, Ireland or the United Kingdom - imposed restrictions on mobility for two years after the enlargement. Thus, the immigrants’ situation did not really change until 2006.

In our empirical analysis, we exploit cross-country high frequency variations in migration
flows so that all of the discussed measurement problems in the EVR end up being absorbed by our quarter fixed effects. Measurement problems related to particular countries of origin would only be an issue as long as they may not be absorbed as well by our origin-year fixed effects.

### A.1.3 Total quarterly flows

The EPA represents the only other data source for which variation on migration flows at the quarterly level can be obtained. By subtracting the stock of migrants in a given quarter from the stock of migrants in the following quarter we can obtain a measure of net migration flows. The quarterly migration series that we produce with this methodology can be compared to the net migration flows obtained from the EVR (recall that figures for the emigration of the foreign-born can only be considered reliable after 2006). This is what is done in Figure A.4.

The comparison of both time series indicates that the general trend and magnitude of the flows is highly comparable in the two sources, especially taking into account the standard errors associated with the EPA net flow. The raw correlation between the two net flows series is 0.78, which is extremely high considering the uncertainty involving the timing of emigration flows from the EVR.
A.1.4 Immigration flows by origin

The EVR and the ENI look much more alike when we move to a country-level analysis. The correlation coefficient between origin-year observations from the EVR and from the ENI between 1997 and 2005 is 0.88 (calculated over 604 country-year observations). If we restrict the ENI sample to those country-year pairs for which there were at least 10 observations, we are left with 177 country-year observations, for which the correlation with the EVR is still 0.84.\(^{46}\)

We run basic regressions to check to what extent ENI origin-year observations can explain EVR origin-observations: when we did so for the 177 common observations for which there were at least 10 individuals in the ENI sample, the result is a coefficient reassuringly equal to 1.00.\(^{47}\) This could hide differences on a country by country basis but, when we run origin-specific regressions (with the caveat that the highest number of observations is 9 in these regressions), we could not reject the coefficient on the ENI numbers being 1 for any country but Colombia at a 95 percent confidence level (with a p-value of 0.0497).

When we run a regression with origin fixed effects, the resulting coefficient was 0.97 (not statistically different from 1 at a 99 percent confidence level). We also run a regression with time fixed effects exploiting the cross-sectional variation in the data, and the estimated coefficient was again 1.00. However, the last specification shows that the 2004 and 2005 year fixed effects are significant, which indirectly suggests that point (iii) above about the 2005 amnesty (announced in the last months of 2004) did play a relevant role.

A.2 Spanish immigration policies

We detail below how each of the ten variables that describe how Spanish immigration policies changed over our period of analysis (1997-2009) were built, and the corresponding legal sources.

**January 2000 Amnesty** - the dummy variable takes value 1 for all countries from January 2000 (source: *Ley Orgánica 4/2000*).

\(^{46}\)The share of total immigrants covered by this restriction is 85 percent both in the ENI and in the EVR.

\(^{47}\)Of course, this result hides differences on a year by year basis: running yearly regressions, we obtain a coefficient around 0.3 for the years 1997-1999, around 1 for 2000-2003, 1.3 in 2004 and 2.3 in 2005.
November 2004 Amnesty - the dummy variable takes value 1 for all countries from November 2004 (source: Real Decreto 2393/2004).

EU-15 - the dummy variable takes value 1 if the country of origin belongs to the European Union as of 1997. Thus, it is 1 for Austria, Belgium, Denmark, Finland, France, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, United Kingdom, Germany and Sweden (source: www.europa.eu.int).

Schengen Area - the dummy variable takes value 1 from the inclusion of a country in the Schengen Area. It is 1 in the whole sample period for Belgium, France, Luxembourg, Monaco, Netherlands, Portugal and Germany; 1 from November 1997 for Italy, San Marino and the Holy See; 1 from December 1997 for Austria; 1 from April 2000 for Greece; 1 from April 2001 for Denmark, Finland, Iceland, Norway and Sweden; 1 from April 2008 for Hungary, Malta, Poland, Latvia, Estonia, Lithuania, the Czech Republic, Slovakia and Slovenia; and 1 from January 2009 for Switzerland (source: www.europa.eu.int).

EU May 2004 Eastern Enlargement - the dummy variable takes value 1 from May 2004 for Cyprus, Hungary, Malta, Poland, Latvia, Estonia, Lithuania, Czech Republic, Slovakia and Slovenia (source: www.europa.eu.int).

EU January 2007 Romania and Bulgaria Enlargement - the dummy variable takes value 1 from January 2007 for Romania and Bulgaria (source: www.europa.eu.int).

Visa requirement for non-immigrant admission - the dummy variable takes value 1 for those countries and periods for which a visa was required to enter Spain. It is 1 for all values with the exception of the following: members of the EU-15 group; Andorra; Iceland; Norway; Liechtenstein; Croatia; country-month pairs for which the Schengen area dummy is 1; Eastern Enlargement (2004 and 2007) countries plus Switzerland from April 2001; Chile; Peru; Argentina; Bolivia until March 2007; Colombia until December 2001; Ecuador until July 2003; Venezuela; Paraguay; Brazil; Uruguay; Mexico; Costa Rica; El Salvador; Guatemala; Honduras; Panama; Nicaragua; Australia; New Zealand; Canada; United States; South Korea; Brunei; Israel; Japan; Malaysia; Singapore; Antigua and Barbuda, Bahamas, Barbados, Saint Kitts and Nevis, Seychelles and Mauritius from June 2009. The sources are Schengen Area regulations (www.europa.eu.int; www.maec.es and www.boe.es) and bilateral agreements of Spain with Latin American countries (www.mtin.es and www.boe.es).

Bilateral agreement on double nationality - the dummy takes value 1 if a bilateral agreement on double nationality with Spain exists: Costa Rica, Guatemala, Honduras,
Nicaragua, Dominican Republic, Argentina, Bolivia, Colombia, Chile, Ecuador, Paraguay and Peru (sources: www.mtin.es and www.boe.es).

**Bilateral agreement on social security** - the dummy takes value 1 if a bilateral agreement on Social Security with Spain exists: Mexico, Argentina, Brazil, Chile, Ecuador, Paraguay, Peru, Uruguay and Venezuela for the whole period; Panama until May 2000; Dominican Republic from July 2006; and Colombia from March 2008 (sources: www.mtin.es and www.boe.es).

**Bilateral agreement on contracts at origin** - the dummy takes value 1 from the moment when a bilateral agreement that allows to sign in the country of origin a labor contract with a Spanish employer is applied: Colombia from June 2001, Ecuador from July 2001, Dominican Republic from February 2002 and Peru from August 2004 (sources: www.mtin.es and www.boe.es).

**A.3 GDP data**

We gathered real GDP quarterly data for all countries of origin with a positive total number of immigrants in all the quarters, and we were able to find these data for 61 origin countries, representing 87 percent of total immigration flows to Spain between 1997 and 2009.\(^{48}\) The main data source was represented by IMF (2010a), which we combined with data from IMF (2010b) and from various Central Banks.\(^{49}\) When the original series of real quarterly GDP data were not seasonally adjusted, we implemented the adjustment regressing the log of real GDP on a linear time trend and quarterly dummies, as suggested by Baum (2006).

---

\(^{48}\)The population residing in these countries amount to 51 percent of the world total.

\(^{49}\)IMF (2010b) provides information on the rate of growth of real quarterly GDP for several countries and regional aggregates; for each origin country, Table A.1 reports whether the figures from IMF (2010b) are country- or region-specific.
<table>
<thead>
<tr>
<th>Country</th>
<th>Source</th>
<th>from</th>
<th>to</th>
<th>SA</th>
<th>obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Australia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Austria</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Belgium</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Bolivia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>WEO, LAC</td>
<td>2009q4</td>
<td>2009q4</td>
<td>no</td>
<td>1</td>
</tr>
<tr>
<td>Brasil</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>WEO, EE</td>
<td>1999q1</td>
<td>2001q4</td>
<td>no</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2002q1</td>
<td>2009q4</td>
<td>no</td>
<td>32</td>
</tr>
<tr>
<td>Canada</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Chile</td>
<td>IFS</td>
<td>1999q1</td>
<td>2002q4</td>
<td>no</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>WEO, LAC</td>
<td>2003q1</td>
<td>2009q4</td>
<td>no</td>
<td>28</td>
</tr>
<tr>
<td>Colombia</td>
<td>IFS</td>
<td>1999q1</td>
<td>1999q4</td>
<td>yes</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2000q1</td>
<td>2009q4</td>
<td>yes</td>
<td>40</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>WEO, LAC</td>
<td>1999q1</td>
<td>1999q4</td>
<td>yes</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2000q1</td>
<td>2009q4</td>
<td>no</td>
<td>40</td>
</tr>
<tr>
<td>Croatia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Denmark</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Dom. Republic</td>
<td>Central Bank</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Ecuador</td>
<td>IFS</td>
<td>1997q1</td>
<td>2007q3</td>
<td>yes</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Central Bank</td>
<td>2007q4</td>
<td>2009q4</td>
<td>yes</td>
<td>9</td>
</tr>
<tr>
<td>Egypt</td>
<td>WEO, MENA</td>
<td>1999q1</td>
<td>2001q4</td>
<td>yes</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2001q1</td>
<td>2009q4</td>
<td>no</td>
<td>32</td>
</tr>
<tr>
<td>El Salvador</td>
<td>WEO, LAC</td>
<td>1997q1</td>
<td>2005q4</td>
<td>yes</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2006q1</td>
<td>2008q1</td>
<td>yes</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>WEO, LAC</td>
<td>2008q2q1</td>
<td>2009q4</td>
<td>yes</td>
<td>7</td>
</tr>
<tr>
<td>Finland</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>France</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
</tbody>
</table>

(continued)
Table A.1: Data sources for quarterly real GDP (continued)

<table>
<thead>
<tr>
<th>Country</th>
<th>Source</th>
<th>from</th>
<th>to</th>
<th>SA</th>
<th>obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgia</td>
<td>WEO, CIS</td>
<td>1999q1</td>
<td>2002q4</td>
<td>yes</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2003q1</td>
<td>2009q4</td>
<td>no</td>
<td>28</td>
</tr>
<tr>
<td>Germany</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Greece</td>
<td>WEO, Euro</td>
<td>1999q1</td>
<td>2000q4</td>
<td>yes</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2001q1</td>
<td>2009q4</td>
<td>no</td>
<td>36</td>
</tr>
<tr>
<td>Guatemala</td>
<td>WEO, LAC</td>
<td>1999q1</td>
<td>2000q4</td>
<td>no</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Central Bank</td>
<td>2001q1</td>
<td>2009q4</td>
<td>yes</td>
<td>36</td>
</tr>
<tr>
<td>Hungary</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Iceland</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>India</td>
<td>WEO, country</td>
<td>1999q1</td>
<td>2006q4</td>
<td>no</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2007q1</td>
<td>2009q4</td>
<td>no</td>
<td>12</td>
</tr>
<tr>
<td>Indonesia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Iran</td>
<td>IFS</td>
<td>1997q1</td>
<td>2007q4</td>
<td>no</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>WEO, Emerg.</td>
<td>2008q1</td>
<td>2009q4</td>
<td>no</td>
<td>8</td>
</tr>
<tr>
<td>Ireland</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Israel</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Italy</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Japan</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Jordan</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Mexico</td>
<td>WEO, LAC</td>
<td>1999q1</td>
<td>2002q4</td>
<td>no</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>1997q1</td>
<td>2003q1</td>
<td>yes</td>
<td>28</td>
</tr>
<tr>
<td>Morocco</td>
<td>WEO, MENA</td>
<td>1999q1</td>
<td>2004q4</td>
<td>yes</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2005q1</td>
<td>2009q4</td>
<td>yes</td>
<td>20</td>
</tr>
<tr>
<td>Netherlands</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>WEO, LAC</td>
<td>1999q1</td>
<td>2002q4</td>
<td>no</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Central Bank</td>
<td>2003q1</td>
<td>2009q4</td>
<td>no</td>
<td>28</td>
</tr>
<tr>
<td>Norway</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Panama</td>
<td>WEO, LAC</td>
<td>1999q1</td>
<td>2003q2</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2003q3</td>
<td>2006q1</td>
<td>no</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>WEO, LAC</td>
<td>2006q2</td>
<td>2009q4</td>
<td>yes</td>
<td>15</td>
</tr>
<tr>
<td>Paraguay</td>
<td>WEO, LAC</td>
<td>1999q1</td>
<td>2005q4</td>
<td>yes</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2006q1</td>
<td>2008q3</td>
<td>no</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>WEO, LAC</td>
<td>2008q4</td>
<td>2009q4</td>
<td>yes</td>
<td>5</td>
</tr>
</tbody>
</table>

(continued)
### Table A.1: Data sources for quarterly real GDP (continued)

<table>
<thead>
<tr>
<th>Country</th>
<th>Source</th>
<th>from</th>
<th>to</th>
<th>SA</th>
<th>obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peru</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Philippines</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Poland</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Portugal</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Romania</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Russia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q3</td>
<td>no</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>WEO, country</td>
<td>2009q4</td>
<td>2009q4</td>
<td>no</td>
<td>1</td>
</tr>
<tr>
<td>Slovakia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Slovenia</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>South Africa</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>South Korea</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Sweden</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Switzerland</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Turkey</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>no</td>
<td>52</td>
</tr>
<tr>
<td>Tunisia</td>
<td>WEO, MENA</td>
<td>1999q1</td>
<td>2000q4</td>
<td>yes</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2001q1</td>
<td>2007q4</td>
<td>no</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>WEO, MENA</td>
<td>2008q1</td>
<td>2009q4</td>
<td>yes</td>
<td>8</td>
</tr>
<tr>
<td>Ukraine</td>
<td>WEO, EE</td>
<td>1999q1</td>
<td>2000q4</td>
<td>no</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>IFS</td>
<td>2001q1</td>
<td>2009q4</td>
<td>yes</td>
<td>36</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Uruguay</td>
<td>Central Bank</td>
<td>1997q1</td>
<td>2008q4</td>
<td>no</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>WEO, LAC</td>
<td>2009q1</td>
<td>2009q4</td>
<td>no</td>
<td>4</td>
</tr>
<tr>
<td>USA</td>
<td>IFS</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
<tr>
<td>Venezuela</td>
<td>Central Bank</td>
<td>1997q1</td>
<td>2009q4</td>
<td>yes</td>
<td>52</td>
</tr>
</tbody>
</table>

Notes: SA describes whether the original series was seasonally adjusted.