Forcasting Swiss Immigration: a Spatial Dynamic Panel Data Model

Alice Milivinti

December 14, 2015

When the absolute number of migrants started to become significant policy-makers of immigration countries started to ask for robust forecasts of the flows of new comers. A rudimentary answer was provided by National Statistical Offices using deterministic models which are still popular even though the importance of stochastic projections has been strongly underlined (Lutz, Sanderson, and Scherbov (1999) and Keilman, Pham, and Hetland (2002)).

However, in contrast to this simplistic methods, migration theory has a long interdisciplinary tradition attempting to explain potential migration drivers using complex interactions among socio-economic variables (Karemera, Oguledo, and Davis (2000), Hatton and Williamson (2002), Mayda (2005), Clark, Hatton, and Williamson (2007), Pedersen, Pytlikova, and Smith (2008), Kim and Cohen (2010), Mayda (2010)).

This paper links the optimal quantity of migration emerging from the neoclassical utility theory, that looks at migration as a result of a cost-benefit analysis in a two-country system (Lewis (1954), Ranis and Fei (1961) Harris and Todaro (1970)), and the empirical literature that tries to precisely estimate and forecast the number of migrants (Alvarez-Plata, Brücker, and Siliverstovs (2003), Brücker and Siliverstovs (2006), Cappelen, Skjerpen, and Tønnessen (2014)). The resulting migration equilibrium has, as empirical counterpart, an econometric model which includes a number of established migration drivers as well as unexplored ones.

The empirical analysis is done exploiting the longitudinal structure of a panel data model. This allows to control for the impact of omitted variables as well to uncover dynamic relations while identifying unit-specific characteristics. Earlier studies have already stressed both the importance of path-dependency and countries heterogeneity in the prediction of future migratory flows (Brücker and Siliverstovs (2006)). Nonetheless, it is crucial to study not only how specific country characteristics impact human mobility, but also the interplay of different states features in a globalized and co-integrated migration regime. To address this the present study employs a dynamic panel model with a spatial dimension, which gives a potential explanation to cross-country linkages. These cross-sectional linkages should identify the economic, demographic and political network that causes and affects peoples' movements. The choice to simultaneously analyse dynamic and spatial components is not only the result of the micro-foundations of the model but also of the aggregate nature of the data used. In macro panels, contrary to micro ones, unites are represented by countries, therefore, cross-sectional units are generally limited (N is small) and the time series are quite long (T is large) whereas the opposite if often true for micro panels. Moreover, static interdependency is an appealing feature of macro panels (Canova and Ciccarelli (2013)) which might be useful for modelling the dynamics of international migration. This attribute, on the one hand, implies that cross-unit endogenous variables interdependencies are likely to be important in explaining the dynamics of multi-country data and, on the other hand, results in correlated errors among different countries. Translated in the current contest it means that migrants with different nationalities may impact one another according to some unobserved characteristics. This aspect raises the issue of cross-sectional dependence which is treated here with a spatial model since it is believed that the presence of a spatial autocorrelation plays a role in shaping migration trajectories.

Spatial dependence might be allowed in two different ways: directly and indirectly. The former indicates that migration behaviours of neighbour countries influence one another while the latter implies that migration decisions are affected by some unobservable variables which are spatially correlated. Hence, the spatial term enters the model directly in the estimation equation through spatial lag and in the error equation. In order to measure the spatial dependence between countries a $n \times n$ Euclidean distance matrix (W) is constructed. The spatial model takes the following form:

$$
\begin{aligned}
y_{it} &= \gamma' \sum_{k=1}^{K} y_{it-k} + \beta' X_{it} + \delta' U_t + \lambda \sum_{j=1}^{N} w_{ij} y_{jt} + \varrho' \sum_{j=1}^{N} w_{ij} X_{jt} + \pi' \sum_{j=1}^{N} w_{ij} y_{jt-k} + \varepsilon_{it} \\
\varepsilon_{it} &= \alpha_i + \rho \sum_{j=1}^{N} w_{ij} \varepsilon_{jt} + \epsilon_{it}, \quad \epsilon_{it} \sim iid(0, \sigma^2), \quad K = 7
\end{aligned} \tag{1}
$$

The dependent variable y_{it} is the value of immigration from country i to Switzerland, y_{it-k} is the dynamic part of the model, X_{it} is a vector of country and time specific variables, Z_i is a vector of country specific time invariant variables and K_t is a vector of time specific country invariant variables. The spatial elements are $\lambda \sum_{j=1}^{N} w_{ij} y_{jt} + \pi' \sum_{j=1}^{N} w_{ij} y_{jt-k}$, the spatial lags, $\rho \sum_{j=1}^{N} w_{ij} \varepsilon_{jt}$, the spatial error term and $\varrho' \sum_{j=1}^{N} w_{ij} X_{jt}$, the spatial weighted exogenous variables, where w_{ij} is the k–element of a matrix of spatial weights (W).

Four different version of the spatial model in Equation 1 are performed: one with only a spatial lag (SAR model), one with a spatial error component (SEM model), one with both (SAC model) and the spatial dynamic Durbin model

(SDM).

Once the nature of the equilibrium and the characteristics of the empirical estimates have been clarified, it is possible to understand why the model can don better forecasts than its competitors/alternatives

The result is a set of estimated parameters that allow to improve the predictions accuracy.

The aim of the current paper is twofold: the construction of a macro panel dataset which includes potential drivers of Swiss immigration and the finding of a suitable estimation methods allowing for the most accurate forecasts. Switzerland gives an interesting case study with its significant immigration since the 1980s, its wide range of different immigrant nationalities, approximately 23% of foreign permanent population and the availability of administrative data. The last facet have consented to merge different data sources in the construction of a dataset which includes economic, demographic, institutional, cultural, historical and geographical variables. The final balanced version of the panel includes 153 countries for 31 years (1981-2011). The potentiality of the data in the estimation of migration drivers will be explored using dynamic and spatial-dynamic panel data methods. First, the model is computed using common estimation methods, such as, pooled OLS, fixed effect, Generalized Method of Moments (GMM) regressions. Second, in order to take into account potential dependencies across countries, dynamic spatial regressions are implemented using a Quasi Maximum Likelihood estimator for spatial-dynamic panel data (Yu, Jong, and Lee (2008)). Testing different estimators is a crucial point since they could lead to very different results as shown, for example, by Alecke, Huber, and Untiedt (2001). Moreover, the comparison between different models will give an insight about which would exhibit the best prediction power and will then be the most suitable to implement forecasts. Third, the prediction power of each model is tested and the performances are ranked according to the Mean Square Root Forecast Error. As expected, dynamic models under-perform with respect to the spatial-dynamic ones which are employed in the forecasting exercise of Swiss immigration.

Keywords: *International Migration; Forecasting; Dynamic Spatial Panel Model; Spatial Autocorrelation*

JEL Classification: F22; C53; C23

References

Alecke, Björn, Peter Huber, and Gerhard Untiedt (2001). “What a difference a constant makes: how predictable are international migration flows?” In: *Migration Policies and EU Enlargement The Case of Central and Eastern Europe: The Case of Central and Eastern Europe*, p. 200163.

Alvarez-Plata, Patricia, Herbert Brücker, and Boriss Siliverstovs (2003). *Potential migration from Central and Eastern Europe into the EU-15: An update.* European Commission, Directorate-General for Employment and Social Affairs, Unit A. 1.

Brücker, Herbert and Boriss Siliverstovs (2006). “On the estimation and forecasting of international migration: How relevant is heterogeneity across countries?” In: *Empirical Economics* 31(3), pp. 735–754. ISSN: 03777332. DOI: 10.1007/s00181-005-0049-y.

Canova, Fabio and Matteo Ciccarelli (2013). “Panel Vector Autoregressive Models A Survey”. In: *ECB Working Paper Series*(15), pp. 1–53. ISSN: 07319053. DOI: 10.1108/S0731-9053(2013)0000031006.

Cappelen, Ådne, Terje Skjerpen, and Marianne Tønnessen (2014). “Forecasting Immigration in Official Population Projections Using an Econometric Model”. In: *International Migration Review.*

Clark, Ximena, Timothy J Hatton, and Jeffrey G Williamson (2007). “Explaining US immigration, 1971-1998”. In: *The Review of Economics and Statistics* 89(2), pp. 359–373.

Harris, John R and Michael P Todaro (1970). “Migration, unemployment and development: a two-sector analysis”. In: *The American economic review*, pp. 126–142.

Hatton, Timothy J and Jeffrey G Williamson (2002). *What fundamentals drive world migration?* Tech. rep. National Bureau of Economic Research.

Karemera, David, Victor Iwuagwu Oguledo, and Bobby Davis (2000). “A gravity model analysis of international migration to North America”. In: *Applied Economics* 32(13), pp. 1745–1755.

Keilman, Nico, Dinh Quang Pham, and Arve Hetland (2002). “Why population forecasts should be probabilistic - Illustrated by the case of Norway”. In: *Demographic Research* 6, pp. 409–453. ISSN: 14359871. DOI: 10.4054/DemRes.2002.6.15.

Kim, Keuntae and Joel E Cohen (2010). “Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach Beyond Gravity1”. In: *International Migration Review* 44(4), pp. 899–932.

Lewis, W Arthur (1954). “Unlimited Supplies of Labour”. In: *Manchester school.*

Lutz, Wolfgang, Warren Sanderson, and Sergei Scherbov (1999). “Expert-Based Probabilistic Population Projections”. In: *Population and Development Review* 24(Supplementary Issue: Frontiers of Population Forecasting), pp. 139–155. ISSN: 00987921. DOI: 10.2307/2808054.

Mayda, Anna Maria (2005). “International migration: a panel data analysis of economic and non-economic determinants”. In:

Mayda, Anna Maria (2010). “International migration: A panel data analysis of the determinants of bilateral flows”. In: *Journal of Population Economics* 23(4), pp. 1249–1274.

Pedersen, Peder J, Mariola Pytlikova, and Nina Smith (2008). “Selection and network effects—Migration flows into OECD countries 1990–2000”. In: *European Economic Review* 52(7), pp. 1160–1186.

Ranis, Gustav and John CH Fei (1961). “A theory of economic development”. In: *The American Economic Review*, pp. 533–565.

Yu, Jihai, Robert de Jong, and Lung-fei Lee (2008). “Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large”. In: *Journal of Econometrics* 146(1), pp. 118–134.