

WZB

Wissenschaftszentrum Berlin
für Sozialforschung



Samuel D. Schmid and Marc Helbling

Validating the Immigration Policies in Comparison (IMPIC) Dataset

Discussion Paper

SP VI 2016–202

December 2016

WZB Berlin Social Science Center

Research Area

Migration and Diversity

Emmy Noether Junior Research Group

Immigration Policies in Comparison

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
www.wzb.eu

Copyright remains with the author(s).

Discussion papers of the WZB serve to disseminate the research results of work in progress prior to publication to encourage the exchange of ideas and academic debate. Inclusion of a paper in the discussion paper series does not constitute publication and should not limit publication in any other venue. The discussion papers published by the WZB represent the views of the respective author(s) and not of the institute as a whole.

Samuel D. Schmid and Marc Helbling
Validating the Immigration Policies in Comparison (IMPIC) Dataset

Discussion Paper SP VI 2016–202
WZB Berlin Social Science Center (2016)

Affiliation of the authors

Samuel D. Schmid
European University Institute
E-mail: Samuel.Schmid@EUI.eu

Marc Helbling
WZB Berlin Social Science Center and University of Bamberg
E-mail: marc.helbling@wzb.eu

Abstract

*by Samuel D. Schmid and Marc Helbling**

The aim of this paper is to discuss the external and internal validity of the newly created Immigration Policies in Comparison (IMPIC) dataset. After presenting its theoretical conceptualization, we compare the IMPIC to other datasets in this policy field. Next, using a variant of principal component analysis, we empirically analyze its sub-dimensions. Among other things, and contrary to some expectations in the extant literature, we find that there appears to be a comprehensive and consistent dimension comprising immigration policies for the fields of labor migration, family reunification, and asylum seekers. We also offer two typologies, which can be used to map the most important dimensions of variation. These validity tests allow us to better understand what the IMPIC dataset measures, what its main dimensions are, and how it can be compared to other indices that measure immigration policies

Keywords: Immigration policy, open borders, internal validity, external validity, principal component analysis, index-building

* We would like to thank Rainer Bauböck and Maarten Vink for their support and detailed feedback on earlier versions of this manuscript. Also, we are indebted to Maggie Peters for providing useful information on her immigration policy index. The two authors of this paper divided the work as follows. Marc Helbling conducted the external validity tests, while Samuel D. Schmid analyzed the internal validity. The other sections were mainly written by Samuel D. Schmid.

Table of Contents

1.	Introduction and overview	2
2.	Context	3
2.1	Strategies of index-building	3
2.2	Measuring the openness of borders	4
2.3	The theoretical conceptualization of the IMPIC	5
3.	External validity	7
4.	Internal Validity	12
4.1	Data and methods	15
4.2	Empirical analysis	16
4.3	Discussion	20
5.	Conclusion	26
6.	References	28

Appendix

- I. CATPCA with diffuse loadings
- II. Correlation matrix: Policy fields and control mechanisms
- III. Robustness tests: Some examples

1. Introduction and overview

The first attempts to measure migration-related policies in the social sciences relied on a “national model approach” derived from typologies and their resulting ideal types (Koopmans 2013: 696). Examples include “models of citizenship” (Safran 1997), “citizenship regimes”¹ (Koopmans et al. 2005), or “idioms of nationhood” (Brubaker 1992). The most salient distinction these works make is between citizenship policies based on (allegedly “ethnic” and German) *jus sanguinis* versus (allegedly “civic” and French) *jus soli*. But more recently, these models and their usefulness for empirical analysis have been questioned. They are considered too static, too simplistic, or too normative (Helbling and Vink 2013: 552; see also Finotelli and Michalowski 2012: 233-5; Bertossi and Duyvendak 2012). Instead, it has become clear that contemporary citizenship (as well as broader integration) policies constitute rather complex regimes that do not necessarily correspond to categorical distinctions, reflecting differences in degree much more than differences in kind (Bauböck 1996: 67). Accordingly, recent research on citizenship (and integration) has begun to move beyond single and small-N comparative case studies and has developed numerous composite indices that measure such policies comparatively in a quantitative way (for an overview see Helbling 2013, 2016 and Goodman 2015). And among other things, recent research has demonstrated that, when viewed comprehensively, citizenship policies defy easy “ethnic-civic” distinctions. Instead, with respect to the birthright dimension of citizenship it has been shown that *jus soli* and *jus sanguinis* provisions are not opposite and contradictory constructs, but are in fact part of two statistically independent dimensions, at least across Europe (Vink and Bauböck 2013).

A similar story can be told for the measurement of immigration policies (understood as policies that regulate immigrant entry, as opposed to integration and citizenship), where typological approaches (see Boucher and Gest 2014: 5) have been superseded by composite comparative policy indices (for overviews see Bjerre et al. 2015, Helbling 2016, and Goodman 2015). The latest addition to these indices is the IMPIC (Immigration Policies in Comparison; Helbling et al. 2016). Based on the framework developed by Munck and Verkuilen (2002), which can be seen as the hallmark of methodological maturity and reflection in index-building in the realm of democracy measurement, it aims to offer a more rigorous conceptualization, measurement, and aggregation than extant indices. However, though the IMPIC is indeed derived against the background of a comprehensive theoretical conceptualization, its underlying empirical dimensionality remains untested. Hence, it is an open question whether the theoretically posited dimensions also represent coherent constructs in empirical terms—

¹ The concept of two-dimensional citizenship regimes crafted by Koopmans and colleagues is an example of a typology that has been refined over time so as to classify countries not only according to types, but also degrees (see Koopmans et al. 2012, Ruedin 2015).

that is, whether the items that go into each dimension are correlated highly enough to warrant their usage as measures depicting consistent sources of variation in empirical analyses. In addition to such a test of internal validity, the IMPIC dataset also needs to be validated externally by comparing it to extant indices.

This paper attempts to do this. We assess both the empirical dimensionality—that is, the internal validity—as well as the external validity of the IMPIC. To do so, we first situate our endeavor in the field of index-building and in the context of some of the debates on and empirical endeavors in measuring the openness of borders. Next, we introduce and discuss the theoretical conceptualization of the IMPIC. Against this background, we evaluate the IMPIC dataset’s external validity. Subsequently, we formulate some possible expectations with regard to the internal validity and empirical dimensionality of the IMPIC. We argue that multiple views are possible, and that the conceptualization employed by the IMPIC, which understands and maps immigration policies as clustering around a differential treatment of various target groups as well as the additional element of “immigration control,” is only one of them. For this analysis, we employ methods of optimal scaling (categorical principal component analysis; CATPCA) as well as standard principal component analysis (PCA). After a discussion of the results, the paper concludes by summarizing the main insights.

2. Context

Before we discuss the conceptualization of the IMPIC, let us briefly address some contextual issues.

2.1 Strategies of index-building

Two basic strategies of index-building can be contrasted. The first is based on theoretical deduction, conceptual relevance, and a “constitutive” view of concept formation (see Goertz 2006). Within this approach, what belongs to which dimension of a concept is derived from theoretical considerations—that is, from a theory about what, in an ontological sense, “constitutes” the dimension of a concept. The second strategy is based on empirical induction, correlations between indicators, and a “latent variable” view of concept formation (see Bollen 1989). Within this approach, what belongs to which dimension of a concept is derived from empirical considerations—that is, from an empirical analysis of what dimensions can be statistically described as consistent in terms of the correlational patterns all the indicators exhibit. In contrast to the first approach, this correlational view ultimately sees indicators as

the empirical manifestations of the concepts or “latent variables” that are assumed to cause them.

Like many others, we believe that, with regard to the indices used for causal² analysis, the best strategies for index-building rest on solid theoretical foundations, while at the same time attending to empirical plausibility. This is why in this paper, besides testing the IMPIC’s external validity, we confront a theoretically well-grounded index with its empirical dimensionality—in the hope of building even more ideal constructs for empirical analysis.

2.2 Measuring the openness of borders

The concept of “open borders” (Carens 2013) or “porous borders” (Benhabib 2004) has attracted much attention in normative political theory. However, it has been disputed whether such one-dimensional notions are useful to analyze contemporary immigration policies empirically. For instance, Shachar and Hirschl (2014) contend that

[d]ebates about migration and globalization can no longer exclusively revolve around the dichotomy between open versus closed borders. Countries simultaneously engage in *both* opening *and* closing their borders, but they do so selectively—by indicating quite sharply who they desire to bring in (namely, those with specialized skills and talent, or, as we shall later see, deep pockets) and erecting higher and higher legal walls to block out those deemed “unwanted” or “too different.”

This view is confirmed by the initiators of another project that aims to comprehensively measure immigration policies (the IMPALA project; Beine et al. 2015). Based on a selection of data for nine countries between 1999 and 2008, they “challenge the idea that any one country is systematically the most or least restrictive toward admissions” (ibid. 1-2). Rather, their data reveal more complex regimes that treat various groups differentially. For example, as we have already seen in the above quote, low-skilled labor immigrants usually face much higher barriers to entry than high-skilled ones (and the rights they are granted also differ in their extent; see Ruhs 2013). This highlights the importance of distinguishing different target groups and entry tracks. However, it must be noted that none of these studies have tested the dimensionality of their data using principal component or factor analysis. Other projects have focused on the admission of low-skilled immigrants specifically (Peters 2015; Shin 2016), and they have used standard principal component analysis (PCA) to construct a composite one-dimensional score.

² For normative evaluations (or descriptive purposes), a purely deductive strategy of index-building seems more adequate (see Blatter, Schmid, and Blättler 2015, 2016).

2.3 The theoretical conceptualization of the IMPIC

The IMPIC defines immigration policy as “government’s statements of what it intends to do or not do (incl. laws, regulations, decisions, or orders) in regards to the selection, admission, settlement and deportation of foreign citizens residing in the country” (Helbling et al. 2016: 4). To measure these policy outputs,³ a focus on different target groups is also important for the conceptualization of the IMPIC. Accordingly, it distinguishes between four so-called “policy fields,” which at the same time “reflect the main reasons why states may accept immigrants” (ibid. 5): for economic reasons (labor immigration), social reasons (family reunification), humanitarian reasons (refugees and asylum seekers), and for cultural and historical reasons (co-ethnics).⁴

As a second step, the IMPIC distinguishes different regulations as policy dimensions that span across and beyond these four fields. These dimensions are grouped into two levels. The first is called “modus operandi,” which separates *regulations* from *control mechanisms*. The former refers to “binding legal provisions that create or constrain rights” (Dreher 2002, cited in Helbling et al. 2016: 7), while the latter are “mechanisms that monitor whether the regulations are adhered to” (ibid.). Thus, the “modus operandi” level captures *how* laws operate (for examples, see ibid. 7). Moreover, while each policy field has its own regulations, control mechanisms cover all policy fields, and they also include elements that refer to irregular immigrants.

The second level are the “locus operandi”. It accounts for the fact that “states regulate and control immigration not only at their borders, but also within their territories” (Helbling et al. 2016: 7). Accordingly, for both regulations and control mechanisms, it differentiates between laws that operate *externally* and laws that operate *internally*.

Moreover, there are several sub-dimensions of regulations. External regulations consist of *eligibility requirements*, which stipulate the criteria for immigrants to qualify for a certain entry track, and *conditions*, which define additional requirements that need to be fulfilled (Helbling et al. 2016: 7). Internal regulations consist of *security of status*, which comprises all policies that regulate the duration of permits and access to long-term settlement, and *rights associated*, which define both the rights immigrants receive with regard to access to the labor market as well as how immigrants are monitored within the territory. All these elements are summarized in Table 1.

³ The IMPIC focuses solely on policy outputs rather than implementation and outcomes, and it tries to keep these policies as distinct as possible from immigrant integration and citizenship policies.

⁴ The aspect of educational or student immigration is left aside.

Table 1: The IMPIC conceptualisation of immigration policy

Policy dimensions			Policy fields			
Modus operandi	Locus operandi	Sub-dimension	Labour migration	Asylum / refugees	Family reunification	Co-ethnics
Regulation	External	Eligibility				
		Conditions				
	Internal	Security of status				
		Rights associated				
Control	External					
	Internal					

Source: Helbling et al. (2016: 6)

The IMPIC further differentiates between different groups of immigrants within each policy field. For labor immigrants, it distinguishes between low-skilled, high-skilled, self-employed, and unspecified migrants. For family reunification, the IMPIC differentiates between sponsors that are citizens (including Second Country Nationals in the EU⁵) versus sponsors that are Third Country Nationals. For the category of asylum, it distinguishes recognized refugees, asylum seekers, and people with humanitarian protection. Finally, in the field of co-ethnics there are up to four different entry tracks. However, in a third of all countries in the sample, no special entry track for co-ethnics exists (Helbling et al. 2016: 5).⁶

The entry tracks of these different groups can be disaggregated into single observations of country-years for each comparable track. This raises the following question: On which level of aggregation is the dimensionality of the IMPIC most adequately tested? We choose to analyze the IMPIC on the level of regulations of each of the four policy fields (eligibility, conditions, security of status, and rights associated; 12 items), plus the two variables concerning control (internal and external; 2 items). We opt for this medium level of aggregation mainly because it yields a number of variables adequate for an informative PCA.

⁵ For national comparisons, the IMPIC only includes the aspects of law that are still regulated by the member states themselves. EU legislation is separately recorded in the dataset, and we do not analyze it here.

⁶ This also has to do with the sample. It seems safe to assume that in non-Western/non-OECD contexts co-ethnic immigration policies may be more prevalent.

3. External validity

One way to test the external validity of an index would be to compare its content with the content of other indices in order to see whether it covers the same or similar dimensions. But there is no agreed-upon authoritative definition or conceptualization of immigration policies. We are thus not in the position to validate our index in comparison to a generally accepted conceptualization of immigration policies.

We can, however, still analyze to what extent our conceptualization is similar to others. A certain degree of similarity can be seen as a sign of validity, as it shows us that our understanding of immigration policies is similar to that of other researchers. Dissimilarity does not, however, necessarily mean that an index is not valid if a different conceptualization has been chosen. The very reason for building a new index often stems from dissatisfaction with existing indices. The aim of an index is thus to cover new aspects. In any case, the comparison of different indices allows us to better understand our new index.

Comparing the content of the existing indices is not a straightforward task, as the indices have been conceptualized in different ways (precisely because there is not yet a common understanding of this concept). We therefore take our conceptualization as a point of reference to see to what extent other indices cover the same dimensions. For the following comparison we first differentiate between the different policy fields. We add citizenship as an additional policy field because there are several indices that also include information on this policy field in their databases.⁷

In Table 2 we list most of the existing and especially the more recent databases on immigration policies (see also Bjerre et al. 2015). It appears that there are certain indices that measure one specific policy field, such as labor migration or asylum, whereas others take a more comprehensive view. Of the latter group, all the indices, with one exception, cover the policy fields that target the four main migrant groups. Timmer and Williams (1998) do not cover asylum seekers and refugees as this category is not relevant for the period under investigation in their study (1860–1930).

There is clear disagreement about whether or not citizenship policies should be included in such a database. Three of the comprehensive indices do include this item whereas the others decided against it. While a clear-cut separation between the policy fields of migration, integration, and citizenship is sometimes difficult, in our opinion citizenship issues constitute a policy area that is different from immigration policies, at least in conceptual terms (Helbling

⁷ Some of them also include information on more general integration policies or immigrant rights (e.g., Ruhs 2013; Peters 2015). We exclude this aspect in our discussion, apart from the instances where we mention the IMPIC subdimension of “Rights associated,” which is related to this dimension.

2013: 559-560; see also Hammar 1990).⁸ This does not, of course, pose a problem as long as a dataset allows for differentiation between these policy areas.

Table 2: Policy fields covered by different indices

	Labor migration	Asylum seekers/ refugees	Family reunification	Control/ irregular migration	Citizenship
Timmer and Williams (1998)	X		X	X	
Thielemann (2003)		X			
Hatton (2004)		X			
Mayda (2005)	X	X	X	X	
Lowell (2005)	X				
Givens and Luedtke (2005)	X	X	X	X	X
Cerna (2008)	X				
Klugman and Pereira (2009)	X	X	X		
Ruhs (2013)	X		X		X
Peters (2015)	X	X	X	X	X
IMPALA (Beine et al. 2015)	X	X	X	X	X
IMPIC (Helbling et al. 2016)	X	X	X	X	
DEMIG (De Haas et al. 2016)	X	X	X	X	X

The policy fields or the migrant groups targeted by regulations and control mechanisms are only one crucial aspect of how to conceptualize immigration policies. Within these fields, further differentiations can be made. We analyze these on the basis of the four most recent databases, which provide more detailed conceptualizations than the older indices and which are also at least partly accessible to the public (Peters 2015; De Haas et al. 2015; Beine et al. 2015; Helbling et al. 2016).

Some databases differentiate between different track types within each policy field. Entry tracks, for example, are the basic unit of the IMPALA database. The number of tracks can vary enormously between countries. For the year 2008, Beine et al. (2015: 9) indicate that the number of tracks varied between 15 and 64 for their nine pilot countries. The IMPALA project does not make any further differentiations, as its aim is to code immigration policies in as much detail as possible, thereby covering all possible regulations. Peters (2015) is much more focused. In her project, tracks are irrelevant, as her dataset focuses on low-skilled

⁸ Whether citizenship policies (and integration policies) are correlated with immigration policies is an empirical question. Related research either finds no correlation or, at least, divergent political logics (Givens and Luedtke 2005) or negative correlations (Ruhs 2013).

migrants only. Moreover, she codes only 12 different items that correspond to some degree to the policy fields we identified in Table 1.

The DEMIG and IMPIC projects are different from the other two as they code a large but limited number of aspects. De Haas et al. (2015: 10) differentiate between 28 different policy tools that range from surveillance technology to access to social rights. These policy tools cover many of the 69 items that are included in the IMPIC database (Bjerre et al 2016: 45). Moreover, the DEMIG project differentiates between four different policy areas that are to some extent comparable to the three policy dimensions differentiated in the IMPIC project.

As we see in Table 3 the first DEMIG category is “Border and land control,” which regulates the external and internal border controls that aim to secure the national territory (e.g., surveillance, detention, and sanctions of fraudulent acts). This covers similar aspects as the internal and external control mechanisms in the IMPIC project, including, for example, information-sharing systems between countries, carrier liability rules for transporting undocumented migrants, the surveillance of admitted refugees, and sanctions against employers that have hired irregular migrants. However, unlike the DEMIG project, the IMPIC database makes it possible to differentiate between internal and external regulations. This allows us to account for the fact that states regulate and control immigration not only at their borders, but also within their territories.

The second DEMIG category is “Legal entry and stay,” which covers issues related to entry and stay permits as well as regularizations. This strongly corresponds to the external regulations of the IMPIC project, which measure how difficult it is to cross national borders, but it also corresponds to a sub-dimension of the internal regulations, “Security of status,” that concerns the duration of residence and the possibilities to renew permits.

The third DEMIG category measures post-entry rights and other aspects of integration of a target group, including the access to citizenship. In the IMPIC project this corresponds to the category “Rights associated,” which includes aspects that go beyond the rights of a special status, for example, vocational training rights for labor migrants or labor rights for refugees. The IMPIC category is, however, defined in a narrower way than the DEMIG category, and it does not include access to citizenship. “Exit” is the fourth DEMIG category, and it covers regulations concerning the forced or voluntary exit or return from a territory of a target group. In addition, it includes policy measures that aim to regulate the state’s relations with its citizens living abroad as well as with their descendants. These aspects are not included in the IMPIC database.

Overall we see that the two projects cover very similar aspects. Still, there are also some crucial differences. While the DEMIG project covers aspects concerning exit that are not

covered by the IMPIC, the latter has a more fine-grained categorization that makes it possible to study more specific aspects of immigration policies.

Table 3: Policy dimensions of the IMPIC and DEMIG projects

IMPIC			DEMIG
<i>Modus operandi</i>	<i>Locus operandi</i>	<i>Sub-dimension</i>	<i>Policy areas</i>
Regulation	External	Eligibility	Legal entry and stay
		Conditions	
	Internal	Security of status	Integration
		Rights associated	
Control	External	Border and land control	
	Internal		
			Exit

Given the fact that the DEMIG and IMPIC databases are conceptualized in a similar way, we should expect them to also correlate highly. A direct comparison is, however, not possible as the DEMIG project only codes policy changes and not the absolute levels of restrictiveness of specific policies. Unlike the IMIPC database, which includes information on the absolute levels of restrictiveness, the DEMIG project can therefore not compare policies across countries. Still, it is possible to compare policy trends across the two datasets. We can test to what extent the data of the two datasets provide a similar picture of how policies evolved over time between 1980 and 2010—the time period that is covered by both datasets.

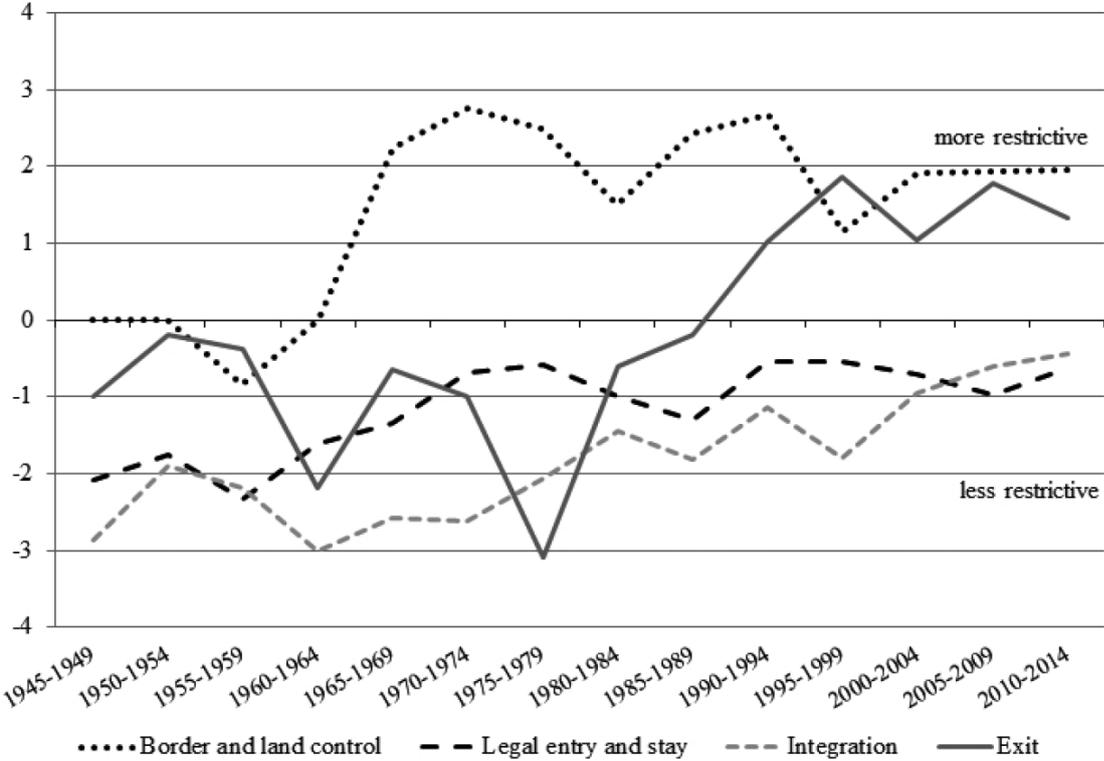
Figure 1 indicates the evolution of the four DEMIG policy areas. Since only policy changes have been coded the graph indicates whether a policy area has become more restrictive or more liberal. Values higher than zero indicate a trend toward more restrictiveness and values lower than zero indicate a trend toward less restrictiveness. So even if a line remains positive with the same value, this indicates that regulations have become more restrictive each year. For the three policy areas that can also be measured with the IMPIC data, we see that since the 1980s border and land controls have continuously become more restrictive. By contrast, the areas “Legal entry and stay” as well as “Integration” have become consistently less restrictive since 1980.

Figure 2 shows the evolution of the DEMIG policy areas when measured with the IMPIC data for the 22 liberal democracies included in the DEMIG dataset. These countries vary between 0 (liberal) and 1 (restrictive). We have created new indices that correspond to the

DEMIG areas as shown in Table 2. “Border and land control” corresponds to the external and internal control mechanisms, “Integration” to rights associated, and “Legal entry and stay” to the external regulations plus security of status. While the scales are not comparable in absolute terms, we observe the same trends: Border controls became continuously more restrictive between 1980 and 2010 whereas the other two policy areas became more open.

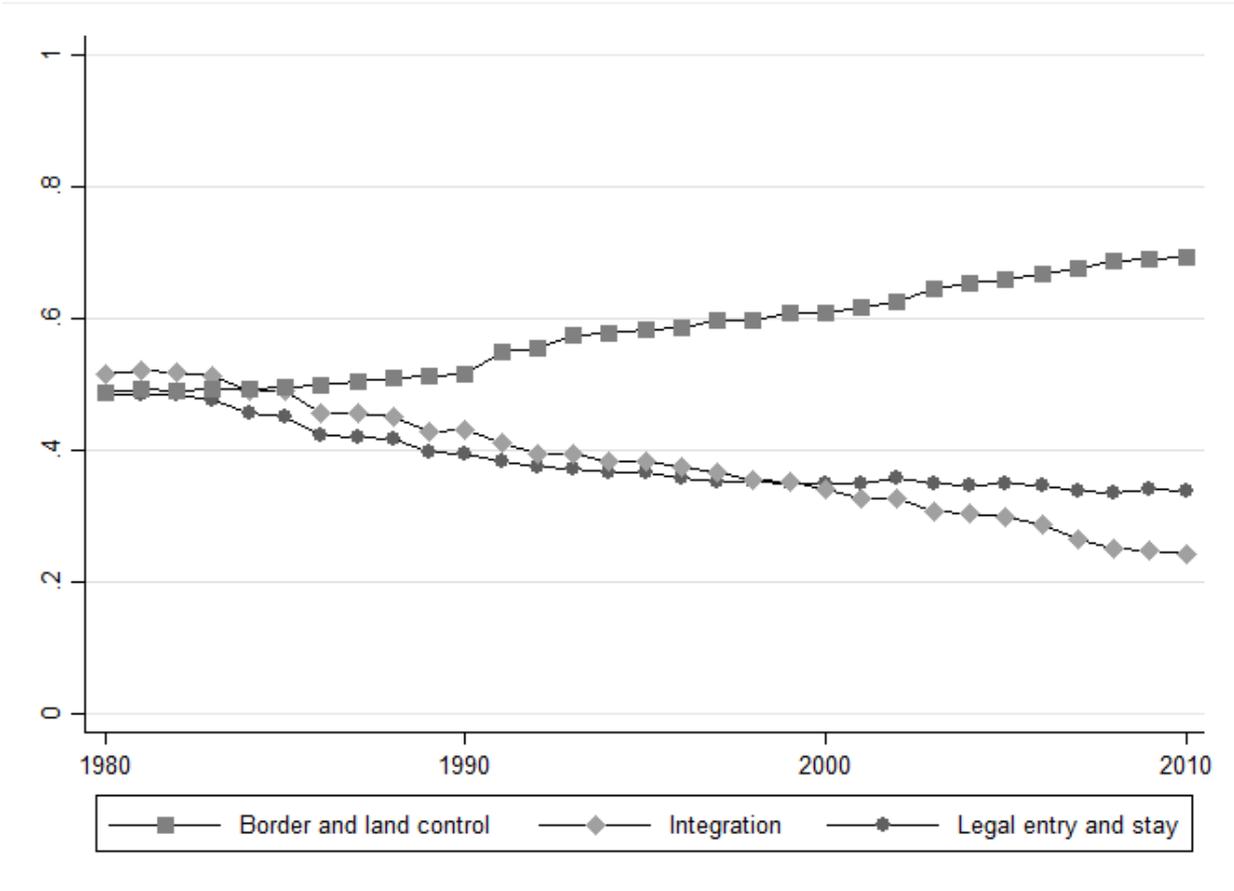
Comparisons with other existing datasets are difficult as they are not only conceptualized differently but also cover only specific aspects for individual years. In more detailed analyses we correlated the indices by Ruhs (2013) and Peters (2015) and integration indices from the MIPEX (Huddleston and Niessen 2011) and ICRI projects (Koopmans et al. 2012; other datasets were not accessible) with sub-indices we developed that come close to their conceptualizations. In most instances, we found positive correlations that varied from rather weak to relatively moderate with Pearson’s *r* values around 0.5 and 0.6. It thus appears that the general trends are the same as in other projects but that there are still some large differences. This might be due to different conceptualizations. Final conclusions are, however, difficult to draw given the relatively small overlap of country/year data points.

Figure 1: Evolution of DEMIG policy areas



Source: De Haas et al. (2016: 20)

Figure 2: Evolution of DEMIG policy areas with IMPIC data



4. Internal validity

Having compared the IMPIC dataset with other datasets, we now turn to its internal validity. Can we expect the regulations to empirically cluster along the lines demarcating the policy fields, and the control mechanisms to form a separate dimension? Against the background of the conceptualization of the IMPIC, we argue that we can. Immigration policy is mainly structured according to different target groups, each of which enters a host country on specific grounds (see above). Labor immigrants enter for economic reasons (e.g., to increase their earnings), and when deciding how to admit them, government actors in host countries (as well as business interest groups and domestic labor who seek to influence policy) apply economic considerations such as efficiency (e.g., maximizing the benefits of immigration for economic growth) or distribution (e.g., making sure immigration does not harm the lowest-paid workers in the economy; Ruhs 2013: 5). Refugees and asylum seekers enter for

humanitarian reasons⁹ (e.g., because they are persecuted in their origin country or are fleeing from war), and when deciding how to admit them, government actors in host countries (as well as non-state actors such as human rights NGOs seeking influence) apply moral and humanitarian considerations (e.g., how generous the state wants to be in offering protection, within the bounds of international law; cf. Gibney 2004). Co-ethnics enter for cultural and historical reasons (e.g., to return to the original “homeland” or to a former colonizer) and when deciding how to admit them, government actors in host countries apply considerations that relate to moral obligations toward historically relevant groups (e.g., offering kin-minorities preferential access to its territory) as well as considerations relating to nation-building (cf. Joppke 1998, 2005). Finally, family immigrants enter for social reasons (e.g., to reunite with their partner and children); and, when deciding how to admit them, state actors in host countries apply considerations about the extent of the right of individuals to family life (which is legally constrained in liberal democracies; see Joppke 1998) and societal cohesion. The latter aspect cross-cuts the former three, as claims to immigration rights of family members can be made by immigrants admitted under any category. However, it can still be expected to form a consistent dimension that indicates an underlying degree of liberalism with regard to the interpretation of this individual right across all target groups.¹⁰

Internal and external control mechanisms transcend these four policy fields to an extent that they can be considered a separate dimension. This dimension is, essentially, about “immigration control,” and it is structured mainly according to considerations about enforcement, public order, and national security. This expectation is reinforced by the fact that control mechanisms also refer to the treatment of irregular immigrants, who can potentially belong to any type of group listed above, but for whom—by definition—no legal entry track exists (or they use a given entry track and then overstay).

In light of these arguments, we expect the empirical dimensionality to closely mirror the theoretical conceptualization of the IMPIC—at least on the level of aggregation on which we test it. Immigration policies should configure along the four dimensions demarcated as regulations in the various policy fields, and along the dimension of control, leading to a five-dimensional outcome.

Hence, we do not expect that the internal and external aspects of the regulations in the different fields cluster on different dimensions, or that they are even negatively correlated.

⁹ Of course, immigrants trying to enter a country on this basis may be economic migrants. However, states generally do not admit economic migrants through the channel of asylum, and we can assume that most of their asylum requests are rejected. Whether these migrants then leave the country or stay there as irregular migrants is another question, which is covered by certain items on control mechanisms in the IMPIC.

¹⁰ Sometimes, this aspect is even considered an immigrant right in itself rather than an aspect of immigration policy openness (Ruhs 2013).

This goes against certain claims in the literature that posit a trade-off between “openness” and “rights,” at least in the realm of labor migration (Ruhs 2013; see also Peters 2015, who finds that immigration policy openness and immigrant rights form different dimensions for low-skilled migrants). Rather, we hold that openness (i.e. the external regulations stipulating conditions and eligibility) and rights (i.e. the internal regulations covering rights associated and the security of status) are part of the same dimensions for each policy field; they are, in fact, positively correlated. Therefore, we argue that it is the policy fields as a whole rather than the sub-dimensions that act as the main structuring factors.

But does that mean that the variations in the different policy fields are completely independent? That is, are countries that are open on one dimension not necessarily open on other dimensions as well? Let us consider an alternative argument that goes against the assumption of policy differentiation amounting to multi-dimensionality. As broader and more fundamental concerns about security and public order, national identity and social cohesion, or the sustainability of the welfare state pervade all immigration policy fields, admitting immigrants to a country may also be seen as an issue structuring political conflict along a single dimension. Comprehensive stances toward immigration can be conceptualized as manifestations of the positioning of a national community toward “the outside” in the context of globalization; and the highly politicized area of immigration policy may be a battleground for political competition that divides politics along the lines of a new one-dimensional “integration-exclusion” or “globalization cleavage” (Kriesi et al. 2012).

Nevertheless, there is another group that is distinct from all other immigrants: co-ethnics. In the public discourse, this group is often not even perceived as immigrants. “They belong to us, not to them,” a fellow co-ethnic might say. It is also possible that co-ethnics form one dimension, while—in line with the above argument—all others form another, very comprehensive dimension covering asylum, family reunification, and labor immigration. Meanwhile, since control mechanisms in a way transcend the policy fields, it is unclear how they align with these two dimensions. Taking the argument very seriously, though, we could expect them to line up with the broad dimension covering “ordinary” (non-co-ethnic) immigrants.

As a consequence, we can expect different numbers of dimensions. The first argument suggests five dimensions that reflect the four policy fields plus control mechanisms. The second argument suggests two dimensions. On the one hand, there may be one comprehensive dimension covering three policy fields (family, labor, and asylum) as well as control mechanisms. And on the other hand, there are policies targeting co-ethnics. However,

there might even be another dimension if we expect control mechanisms to be separate, leading to a total of three dimensions.

4.1 Data and methods

We use the IMPIC dataset, which covers 33 OECD countries from 1980 to 2010, resulting in 1,023 country-year observations (Helbling et al. 2016). Since all country-years are pooled, this set-up is likely to produce a pattern of auto-correlation for each indicator with itself over time. Whether this poses a problem for principal component analysis (PCA; which is the method we use, see below) has been contested. For a long time, many studies using PCA with non-independent data appear to have cited Jolliffe (2002: 299; see e.g., Leong and Goswami 2015), which states that as long as the main goal of the PCA is descriptive and not inferential (which applies to us), the non-independence problem poses no serious threat. In addition, this pooled approach has already been used to construct immigration policy indices by applying standard PCA to ordinal data (Peters 2015; Shin 2016). However, recent research in the field of engineering has questioned the feasibility of what has been labeled the “naïve approach” of pooling, arguing that time dependency produces a tendency to find more components to adequately describe the data (Vanhatalo and Kulahci 2015). Nevertheless, due to a lack of alternatives, we follow the majority of studies, which do not regard this as a serious problem. To ensure the validity of our results, however, we conduct robustness tests using certain sub-sets of our data to validate our results (these results are shown and discussed in appendix III). All but one robustness test confirm that our results also hold for different sub-sets of the data.

As we already indicated, the IMPIC indicators are coded with a restrictiveness scale from 0 to 1 (and several points in between), with 1 indicating maximum *restrictiveness* (theoretical minima and maxima are used; see Helbling et al. 2016: 11). All observations with missing values (e.g., many countries do not have policies for co-ethnics) are given the most restrictive score, since this implies that a country has not opened a specific legal pathway for people to enter, for instance, on the grounds of co-ethnicity. However, before the analysis all the variables are recoded from 0–1 to 1–2 in order to use the CATPCA package in SPSS, which treats values of 0 as missing.

The underlying concept of *restrictiveness* can be seen as continuous. However, any attempt to code laws and regulations means that, in the real world, such data are ordinal—or that they can be transformed into ordinal scales (e.g., fees, which are numerical; as done in Helbling et al. 2016). Hence, though the IMPIC presents a “*linear restrictiveness scale*” (Helbling et al. 2016: 13; emphasis added), for the purposes of this analysis, we treat the scored IMPIC data

as ordinal. We therefore use categorical instead of standard principal component analysis (CATPCA; Meulman et al. 2005). This is a method of optimal scaling that allows variables to be scaled at various levels and that aims to model non-linear relationships. We select a spline ordinal scaling for all variables. Hence, the information in the observed variable on the ordinal rank of objects is preserved in the optimally scaled variable, but without assuming that the intervals between consecutive categories are equal. Since CATPCA is very similar to standard PCA, the output of such an analysis can be interpreted in the same way (Linting et al. 2007: 27-8).

However, to facilitate the extraction of clearly interpretable components, we apply an extension. If the outcome of a dimensional analysis with CATPCA leads to diffuse dimensions, negative and positive loadings on different dimensions or cross-loadings in general, this can be seen as evidence for the need to rotate the solution. This is why, whenever necessary, we run a standard PCA using the transformed variable quantifications obtained by CATPCA (which can be treated as linear constructs) and rotate this solution using a common orthogonal VARIMAX rotation. This strategy is recommended by proponents of CATPCA (Linting et al. 2007).

Finally, it must be noted that we evaluate the different solutions and components with regard to their theoretical plausibility, their explained variance, their Eigenvalues, and Cronbach's alpha (to assess internal consistency of a component). We do not apply further, more sophisticated model fit statistics such as the so-called RMSEA or screeplots. We also fall between using PCA as a confirmatory tool and using it as an exploratory tool; we have some theoretical expectations, yet we want to remain open to other possibilities. Last but not least, to keep matters simple, note that we assume the different components to be orthogonal instead of testing the statistical adequacy of various oblique rotations. Instead, the correlations of the averaged items of each policy field are shown in a correlation matrix in appendix I.

4.2 Empirical analysis

Our first argument theoretically posits five dimensions: four policy fields plus control. As a first step, we thus extract a five-dimensional solution. Since the initial result displays a diffuse pattern involving many cross-loadings that cannot be given a clear theoretical interpretation (see appendix II), we conduct a standard PCA using the variable quantifications obtained by the CATPCA and rotate the solution to obtain a clearer pattern of loadings. Note that, however, for the standard PCA we do not specify the extraction of five dimensions *ex ante*—

rather, we use the common method of a minimum Eigenvalue of 1 as the extraction criterion. This provides an additional test of the validity of a five-dimensional solution: If this procedure produces five dimensions, this can be considered strong evidence for the validity of this way of seeing the data.

The result is simple (Table 4). There indeed appear to be five dimensions, which configure along the policy fields and control mechanisms. The first dimension describes policies toward co-ethnics and is extremely consistent. However, note that the extremely high loadings and the spectacular value for alpha are somewhat artificial. They stem from the fact that there were many missing values for countries that had no policies toward co-ethnics. And since we replace all these missing values with the value for maximum restrictiveness (which is substantially correct according to the IMPIC framework) a high correlation of these items was to be expected. Besides, since co-ethnics are a rather special and less common group of immigrants, the importance of this result in terms of explained variance and Eigenvalue should not be overstated.

Table 4: Component loadings of a standard PCA as a follow-up to CATPCA

IMPIC sub-dimensions	Dimension/Component				
	1	2	3	4	5
Family eligibility quantification		.953			
Family conditions quant.		.925			
Family security quant.		.911			
Family rights quant.		.832			
Labor eligibility quant.			.938		
Labor conditions quant.			.942		
Labor security quant.			.938		
Labor rights quant.			.725		
Asylum eligibility quant.				.897	
Asylum conditions quant.				.723	
Asylum security quant.				.805	
Asylum rights quant.				.857	
Co-ethnics eligibility quant.	.991				
Co-ethnics conditions quant.	.989				
Co-ethnics security quant.	.968				
Co-ethnics rights quant.	.992				
Internal control quant.					.821
External control quant.					.856
Cronbach's alpha	.994	.968	.940	.896	.679
Explained variance (%)	36.988	24.037	10.374	9.254	6.959
Eigenvalues	6.658	4.327	1.867	1.666 ¹¹	1.253

PCA based on the transformed variable quantifications of the CATPCA in appendix II; N = 1023; extraction criterion: Eigenvalue > 1; VARIMAX rotation with Kaiser normalization; Cronbach's alpha calculated only for the items that belong to the dimension; loadings < 0.4 not shown

¹¹ The devil is in the details.

The other dimensions are consistent as well, and more realistically and importantly so. Family, labor and asylum immigration policies appear to form coherent and distinct empirical constructs, though their importance in terms of explained variance diminishes from one to the next. In addition, internal and external control mechanisms constitute a separate dimension. However, it is the least consistent according to alpha (though it is still sufficiently reliable; a value of above 0.6 is considered sufficient), and it adds only little to the proportion of explained variance. Finally, when taken together, the five dimensions explain almost 88 percent of the variance in all items.

Our alternative argument has suggested that family, labor, and asylum regulations may be part of a single comprehensive dimension. And indeed, a one-dimensional CATPCA solution provides evidence for this claim (Table 5). The Eigenvalue of this dimension is high, and it explains almost half of all variation in the items. With regard to the most common immigrant groups, immigration policies can therefore also be seen as a single broad dimension that, so we would interpret, signifies the general openness of the territorial boundaries of a national political community. The items for co-ethnics and controls form something separate, albeit diffusely so. This suggests that more dimensions are needed to depict the variation in all the items. This is why we now turn to a two-dimensional solution.

Table 5: Component loadings of a CATPCA: One-dimensional solution

IMPIC sub-dimensions	Dimension 1
Family eligibility	.733
Family conditions	.802
Family security	.819
Family rights	.826
Labor eligibility	.781
Labor conditions	.781
Labor security	.782
Labor rights	.576
Asylum eligibility	.602
Asylum conditions	.652
Asylum security	.679
Asylum rights	.702
Co-ethnics eligibility	-0.167
Co-ethnics conditions	-0.200
Co-ethnics security	.018
Co-ethnics rights	-0.060
Internal control	-0.345
External control	-0.327
Cronbach's alpha	.915
Explained variance (%)	47.587
Eigenvalue	6.662

Principal component analysis for categorical data (CATPCA);
N = 1023; variable principal normalization; loadings > 0.4
bold

In a two-dimensional specification, the first dimension again shows that policies for family, labor and asylum immigration form a comprehensive and consistent one-dimensional continuum (Table 6). However, the proportion of explained variance is markedly lower than in the one-dimensional solution (about 37 percent). The items for co-ethnics all load extremely highly on the second dimension. Co-ethnics can thus indeed be seen as a special group. The only flaw is that two asylum items now also moderately load on this dimension. However, their loadings are so much smaller than those of co-ethnics that it seems reasonable to ignore this real-world complexity. An additional finding is the diffuse character of control mechanisms. This suggests that they may lie on a third dimension. However, as we show in appendix II, for this level of aggregation this hypothesis does not hold.

Table 6: Component loadings of a CATPCA: Two-dimensional solution

IMPIC Sub-dimensions	Dimension/Component	
	1	2
Co-ethnics eligibility		.957
Co-ethnics conditions		.954
Co-ethnics security		.946
Co-ethnics rights		.958
Family eligibility	.737	
Family conditions	.803	
Family security	.820	
Family rights	.827	
Labor eligibility	.791	
Labor conditions	.792	
Labor security	.792	
Labor rights	.589	
Asylum eligibility	.581	
Asylum conditions	.641	
Asylum security	.641	.478
Asylum rights	.662	.433
Internal control	-0.316	-0.252
External control	-0.331	.106
Cronbach's alpha	.900	.814
Explained variance (%)	36.988	24.037
Eigenvalues	6.658	4.327

Principal component analysis for categorical data (CATPCA); N = 1023; variable principal normalization; loadings > 0.4 bold

In the last step of our analysis, we climb up the ladder of abstraction. Each policy field and the control mechanisms are now covered by a single item, which reflects their average value. Against the background of the above analyses, we expect three dimensions. Family, labor, and asylum should cluster along one dimension, while co-ethnics and control mechanisms should be separate, also from each other. As a CATPCA with a three-dimensional solution leads to fuzzy results with many cross-loadings (not shown), we are going to concentrate here on the

rotated solution of a follow-up standard PCA based on the variable quantifications obtained by CATPCA. Will we finally get the neat arrangement that our second argument could not fully pin down empirically? The answer is yes (Table 7).

Table 7: Component loadings of a standard PCA based on a previous CATPCA

IMPIC sub-dimensions	Dimension/Component		
	1	2	3
Family quantification	.806		
Labor quantification	.666		
Asylum quantification	.857		
Co-ethnics quantification		.959	
Control quantification			.984
Cronbach's alpha	.649	-	-
Explained variance (%)	40.291	21.883	18.213
Eigenvalues	2.015	1.904	0.991

PCA based on the transformed variable quantifications of a CATPCA; N = 1023; extraction criterion: 3 dimensions; VARIMAX rotation with Kaiser normalization; Cronbach's alpha calculated only for the items that belong to the dimension; loadings < 0.4 not shown

Also from this angle it seems that the most fundamental (and consistent) dimension in immigration policy concerns regulations targeting family, labor, and asylum immigrants. Co-ethnics are special, and so are control mechanisms. This is what our second argument ultimately suggested, but what we could only now show—from almost the top of the ladder of abstraction.

Most of these results are robust over different cross-sections and time periods, that is, they also hold for sub-samples of the IMPIC dataset. We show some of these robustness tests in appendix III. Hence, we can be sufficiently confident that serial auto-correlation does not pose too serious a threat to the validity of our analyses.

4.3 Discussion

Against the background of the conceptualization of the IMPIC, our theoretical discussion made two arguments. The first was that we can expect the IMPIC conceptualization to be valid in the sense that there are four empirically separate policy fields of internal and external regulations targeting the various potential immigrant groups, and an additional dimension comprising internal and external control mechanisms. The second was that it may also be possible that regulations targeting the three most common immigrant groups—family, labor, and asylum immigrants—exhibit a common pattern of variation, leading to a broad dimension of immigration policy, while regulations targeting co-ethnics are special. How control

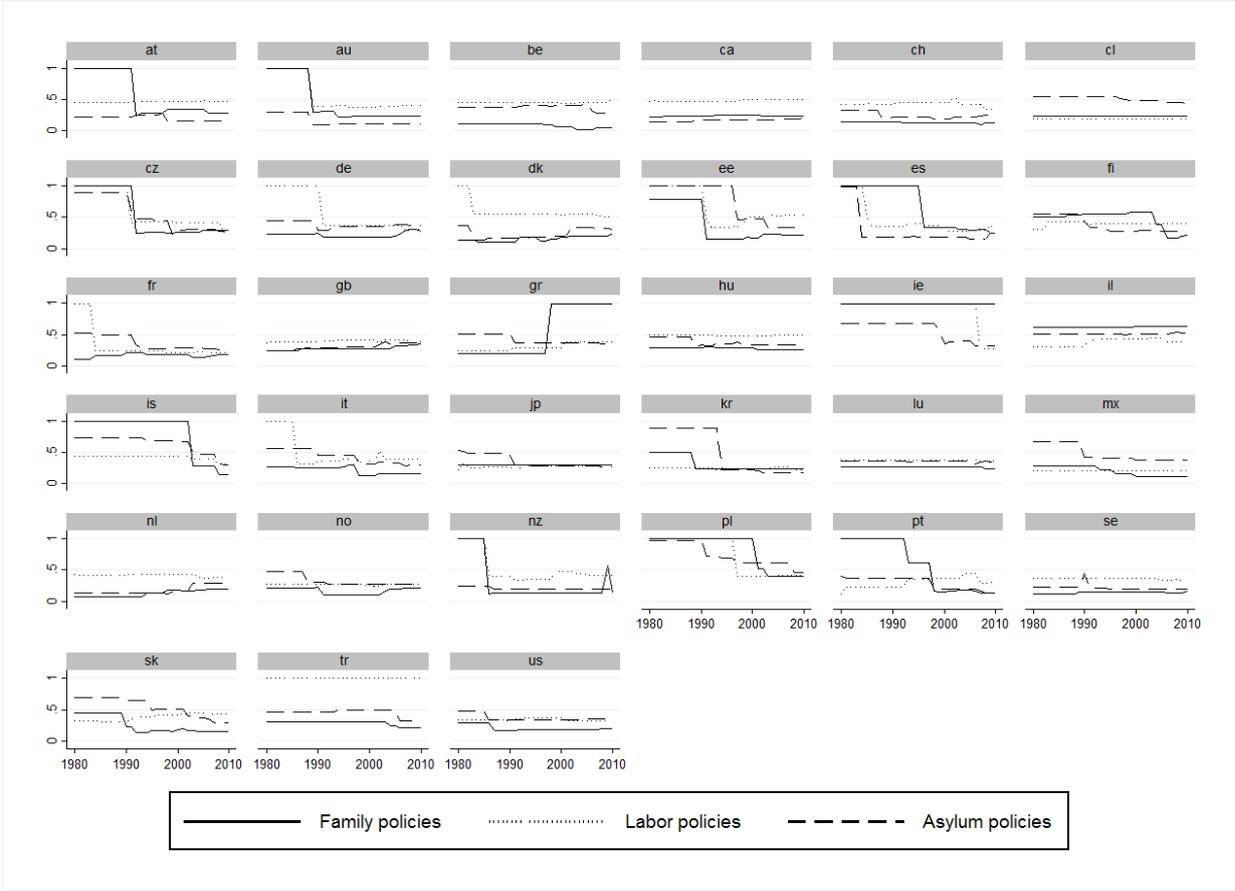
mechanisms fit into this second picture remained somewhat unclear, though we ultimately speculated that they should form a different dimension as well.

The empirical analysis provides evidence for both perspectives. When we specify five dimensions, the results show that a five-dimensional solution neatly describes the distinct policy fields and control mechanisms. However, the one- and two-dimensional solutions indicate that family, labor, and asylum immigration capture a broad underlying tendency of the openness of a country's borders. Finally, a three-dimensional solution on a higher level of aggregation shows that next to this comprehensive first dimension, co-ethnic immigration and control mechanisms form separate constructs.

To illustrate the complementarity of these findings, we have plotted the aggregated IMPIC scores of the regulations of the three most important policy fields—family, labor, and asylum—as a time series for all countries (Figure 4). Note that, because the IMPIC has a theoretically clear interpretation, we use the IMPIC scores rather than the optimally scaled variables from our CATPCAs.

The evolution of these three policy fields over time suggests that, in many countries, they differ to a considerable degree. However, no country exhibits a clear pattern of completely opposite directions (recall that for principal component analyses only common variance patterns, not common levels of values, matter). Instead, the policies often roughly approximate each other, and in some cases they appear tightly linked and even similar in terms of levels (e.g., in the Czech Republic or Hungary). Hence, while a multi-dimensional view naturally allows us to map more of the complexities in policy differentiation, a one-dimensional simplification does not seem to come at the cost of an unacceptably high loss in information (which, in principle, is inescapable when aggregating multiple indicators in any case).

Figure 4: IMPIC family, labor, and asylum policy developments



Notes: The graph displays the development of policy regulations between 1980 and 2010 for 33 OECD countries. The index varies between 0 (liberal) and 1 (restrictive).

If it fits the purpose of a specific research project interested in causal analysis, we therefore recommend two options for aggregating the IMPIC scores. The first option treats the different policy fields and control mechanisms as separate. The second combines the internal and external regulations of family, labor, and asylum immigration policies as one dimension, while treating regulations for co-ethnics as a separate dimension. Also, because there is a family resemblance in the relationship of the different dimensions (cf. Helbling et al. 2016: 13; Goertz 2006), we recommend using unweighted arithmetic means to allow for symmetric substitutability. However, of course this does not mean that this is the only valid aggregation procedure. Depending on the research question, and on a substantive theory about the relationships between the policy fields, one may end up with different aggregations, also in terms of mathematical operations and weightings. And when seen from a “constitutive” view of concept formation and index building—based on the deductive conceptualization of the IMPIC—the complete aggregation of all items can still be seen as valid. However, this is especially true for descriptive or normative-evaluative purposes. We are more skeptical when it comes to causal analysis. A high level of aggregation may not only disguise important

disaggregate differences, but could also yield a one-dimensional empirical construct that is statistically inconsistent.

Another way to use these results is by constructing two-dimensional typologies. A first possible typology differentiates between a dimension capturing labor, family, and asylum immigrants and a dimension covering policies toward co-ethnics. Since the first three groups are the most common immigrants, their dimension could simply be called, in somewhat abstract terms, “openness toward immigration.” But the second dimension comprising policies toward co-ethnics also constitutes immigration, albeit a special kind. In theoretical terms, therefore, the crucial difference between the two is that policies toward co-ethnics can be interpreted as *particularistic*, while policies toward all other groups could be labeled as *universalistic*. The former is *particularistic* because it opens a channel only for a particular ethno-cultural or historically important group. And the latter is *universalistic* because it is based on liberal rights that open channels for all common regular immigrants. This leads to the following typology of immigration regimes (Table 8).

Table 8: A two-dimensional IMPIC typology of immigration regimes I

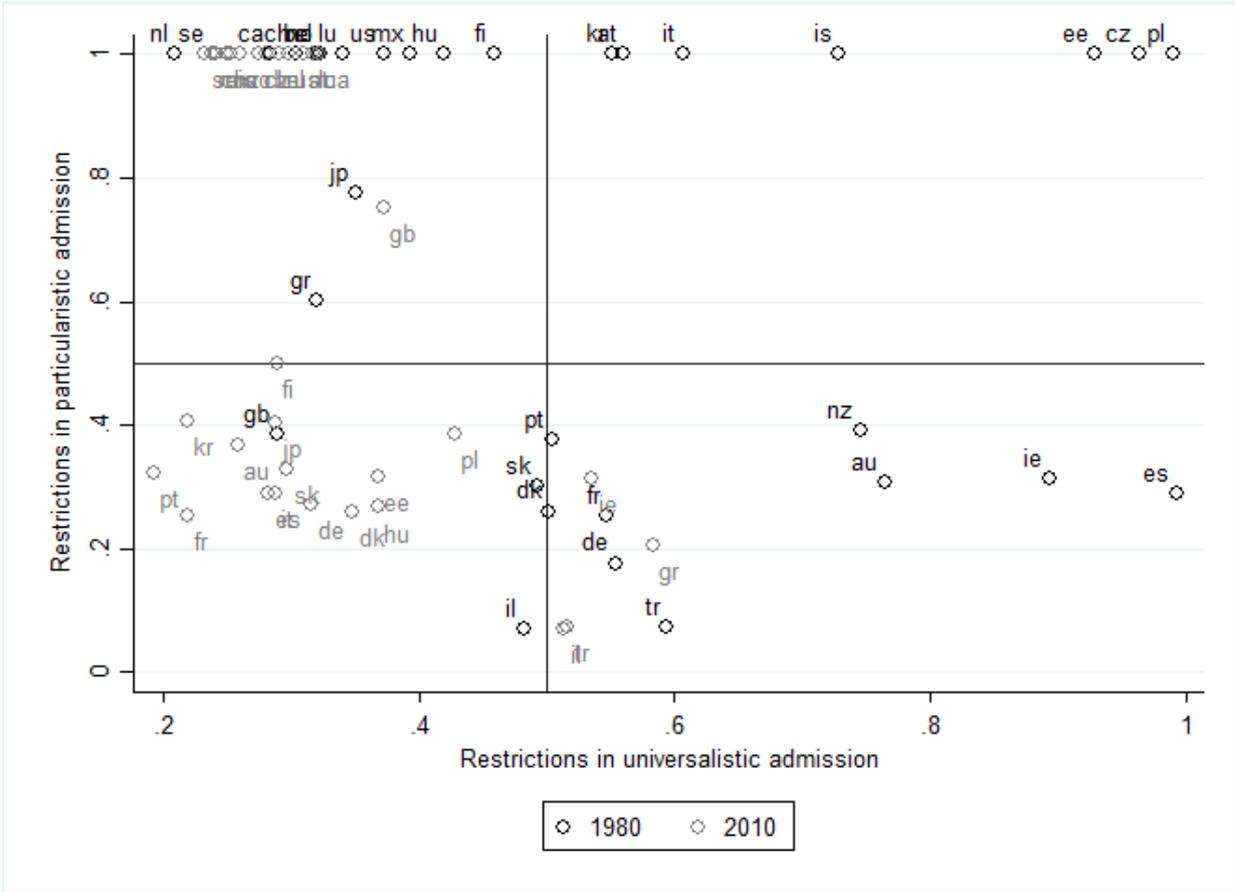
	Universalistic admission	
Particularistic admission	<i>Liberal</i>	<i>Restrictive</i>
<i>Restrictive</i>	universalistic openness	closed
<i>Liberal</i>	open	particularistic openness

Note that we use the term “open” here in a non-idealistic and empirically relative sense, without implying that all border controls would be abolished in a perfectly open setting (as would be true only for the reality of open borders as they exist in the EU Schengen Area or in arguments for open borders such as in Carens 2013). Additionally, also note that this typology is somewhat “unbalanced.” Policies toward co-ethnics are not nearly as important as all the other policies. Co-ethnic immigrants indeed constitute a special and usually rather small group, which is also not particularly salient in the public discourse, at least across the OECD countries. Hence, this typology may be only of limited value. Besides, the universalistic admission dimension seems far more important, also because it goes against previous arguments and findings of a highly differential and potentially contradictory treatment of different immigrant groups (although these arguments mostly pertain to different skill levels in labor immigration, which we did not test in this paper).

Still, it is interesting to see how the countries configure within this property space (Figure 5). After all, this typology does not lead to a rough characterization of countries by simple

categorical ideal types, but allows us to map each country precisely. One notices immediately the artifact produced by the non-existence of co-ethnic immigration policies, with a row of observations clustering at the upper end. Other than that, this scatterplot reveals significant variation (also over time, from 1980 to 2010, indicated by the different signs). Moreover, the universalistic dimension seems heavily skewed toward more open borders (0.5 is the medium value on the IMPIC scale; however, it is also a special threshold because it indicates the presence of a legal provision; see Helbling et al. 2016: 11–2).

Figure 5: Universalistic versus particularistic admission



The second typology also starts from the finding that the universalistic dimension has emerged as the most important. Like its second dimension, however, it uses the control mechanisms, which were shown to measure a different concept, even though they did not load consistently on one dimension other than in the five-dimensional rotated solution and in the final three-dimensional solution with aggregated policy fields and control mechanisms. Hence, the first dimension comprises the most important immigration regulations, and the second dimension comprises the internal and external control mechanisms. As we have seen,

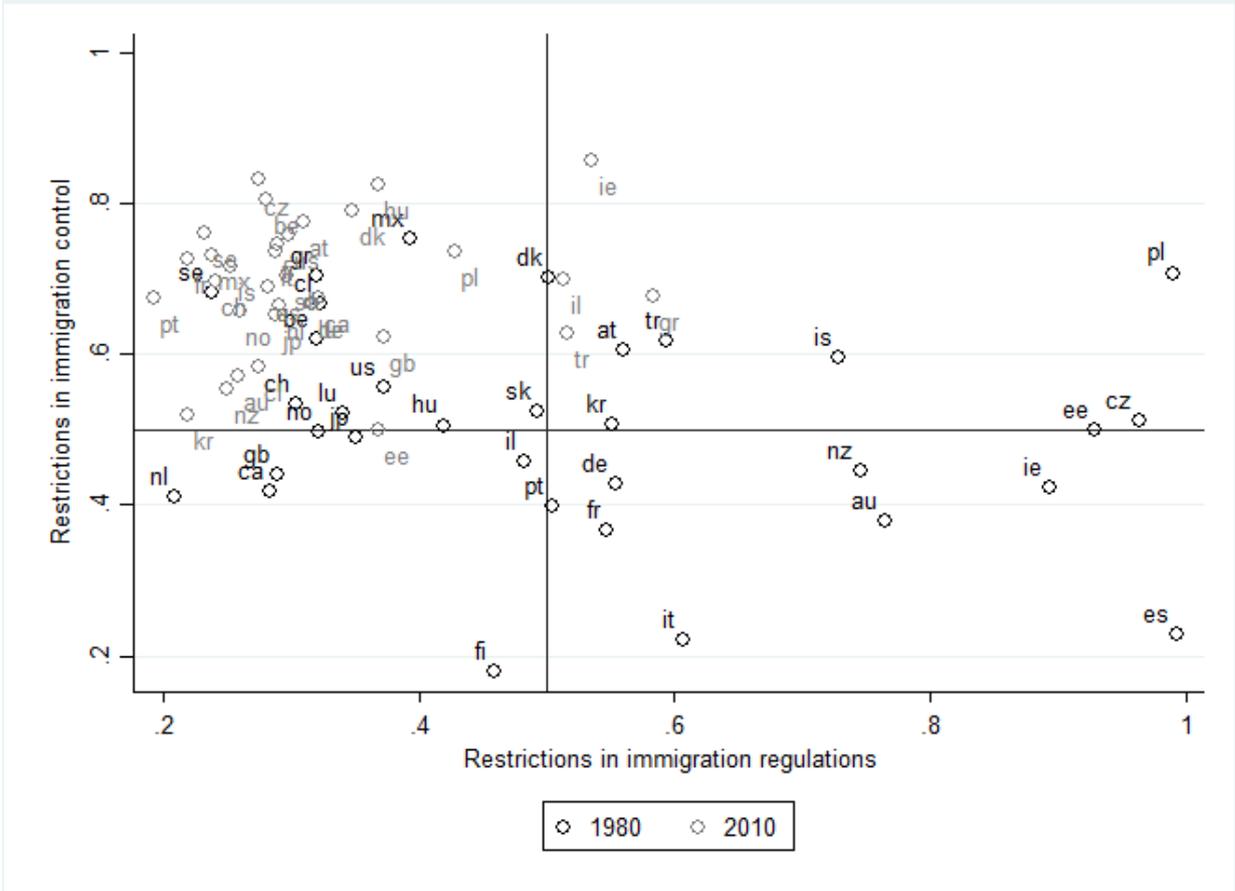
immigration control indicates how strongly the regulations are enforced and how irregular migrants are treated. This leads to the following, alternative typology (Table 11).

Table 11: A two-dimensional IMPIC typology of immigration regimes II

Immigration controls	Immigration regulations	
	<i>Liberal</i>	<i>Restrictive</i>
<i>Strong</i>	strongly enforced open borders	strongly enforced closed borders
<i>Weak</i>	weakly enforced open borders	weakly enforced closed borders

The countries' scores in this property space also reveal significant and interesting variation across space and time (Figure 6). However, not only immigration regulations (which are equivalent to what we called universalistic admission above) but also control mechanisms appeared to be skewed. This results in the finding that most observations show a certain degree of strongly enforced open borders.

Figure 6: Immigration regulations versus immigration control



5. Conclusion

The IMPIC offers a new and theoretically as well as methodologically rigorous way of comprehensively measuring immigration policies in the sense of the external and internal dimensions of territorial admission. The external validity tests have shown that, because of this comprehensive approach, the index varies greatly in content in comparison to other indices. Combined with the highly divergent coverage of countries and years (with the IMPIC normally having a much higher coverage), the correlations with other indices may be of limited informational value. Still, though only weak to moderate in magnitude, the correlations we calculated with comparable sub-dimensions of various indices were all positive. More importantly, however, the external validity of the IMPIC database is especially bolstered by a comparison to the similar DEMIG database, which detects the same long-term patterns of policy evolution for different sub-dimensions of immigration policy.

With respect to its internal validity, this paper has shown that the policy fields and sub-dimensions of regulations and control mechanisms that make up the conceptualization of the IMPIC are not only theoretically consistent, but also empirically valid—at least when we apply a two-step strategy combining CATPCA with standard PCA using a subsequent rotation on the level of aggregation. This suggests that the logic of immigration policy-making can be seen as structured according to diverging purposes relating to different groups of immigrants and to immigration control. We have argued that for labor immigration the purposes are economic, while for family immigration they are social and also tap into broader liberal considerations. For asylum policies, humanitarian and moral considerations are most relevant, and for co-ethnics communitarian aspects come to the fore. Finally, control mechanisms pervade and transcend all policy fields, as they describe how rigorously all the regulations are enforced and how irregular migrants are treated. They form a separate dimension.

In addition, we have presented an alternative argument that can also be supported by empirical evidence. Our analyses show that there is a comprehensive and—in terms of Eigenvalues, explained variance, and Cronbach's alpha—statistically important and consistent first dimension lumping together policies targeting the most common categories of labor, family, and asylum immigration (but excluding the special category of co-ethnics and control mechanisms). Against this background, one may also argue that policy-makers do not act in such a differentiated way after all. As broader and more fundamental concerns about security and public order, national identity and social cohesion, or the sustainability of the welfare state pervade all policy fields, admitting immigrants to a country may also be seen as a rather one-dimensional issue. This would suggest that the logic of immigration policy-making is not

so much about rational purposes relating to different groups of immigrants, but instead constitutes a more basic dimension of political contestation about how the national community and its identity are construed in its relation to the “outside world” in the context of globalization. This battleground divides politics along the lines of a new one-dimensional “integration-exclusion” or “globalization cleavage” (Kriesi et al. 2012). When seen in this light, the dichotomy and continuum between open and closed borders would appear more useful to describe empirical patterns of immigration policies than commonly assumed. However, we should not forget that this result may hinge upon the fact that the level of aggregation on which we tested the IMPIC’s dimensionality was rather high. Future analyses could go deeper and uncover the more complex architectures of immigration policies aggregated along different entry tracks. This is particularly useful to distinguish policies toward labor immigrants with different skill levels and to assess their dimensionality. At any rate, these final reflections highlight the nature of index-building and scale construction as a constant dialogue between deductive specifications about conceptual formation and conceptual relevance, inductive testing of empirical consistency, and careful theoretical considerations that are open to multiple interpretations.

6. References

- Bauböck, Rainer (1996): 'Social and Cultural Integration in Civil Society', in: Rainer Bauböck, Agnes Heller, Aristide Zolberg (eds.): *The Challenge of Diversity. Integration and Pluralism in Societies of Immigration*, Avebury, Aldershot, UK, 67-132.
- Beine, Michel; et al. (2015): "Comparing Immigration Policies: An Overview from the IMPALA Database". *International Migration Review*, DOI: 10.1111/imre.12169.
- Benhabib, Seyla (2004): *The rights of others. Aliens, residents, and citizens*. Cambridge: Cambridge University Press.
- Bertossi, Christophe; Duyvendak, Jan Willem (2012): "National models of immigrant integration: The costs for comparative research." *Comparative European Politics* 10 (3), 237–247.
- Bjerre, Liv; Helbling, Marc; Römer, Friederike; Zobel, Malisa (2015): "Conceptualizing and Measuring Immigration Policies: A Comparative Perspective", *International Migration Review* 49(3), 555–600.
- Blatter, Joachim; Schmid, Samuel D.; Blättler, Andrea (2015): *The Immigrant Inclusion Index (IMIX): A Tool for Assessing the Electoral Inclusiveness of Democracies with Respect to Immigrants*. Working Paper No. 08, Second Edition (Working Paper Series „Glocal Governance and Democracy“, Political Science Department of the University of Lucerne; ISSN 1662-923X).
- Blatter, Joachim; Schmid, Samuel D.; Blättler, Andrea C. (2016): "Democratic Deficits in Europe: The Overlooked Exclusiveness of Nation-States and the Positive Role of the European Union", *Journal of Common Market Studies*. DOI: 10.1111/jcms.12491.
- Bollen, Kenneth A. (1989): *Structural Equations with Latent Variables*. New York: Wiley.
- Boucher, Anna; Gest, Justin (2015): "Migration studies at a crossroads: A critique of immigration regime typologies". *Migration Studies* 3(2), 182–198.
- Brubaker, Rogers (1992): *Citizenship and nationhood in France and Germany*. Cambridge, MA: Harvard University Press.
- Carens, Joseph H. (1987): "Aliens and Citizens: The Case for Open Borders". *The Review of Politics* 49(2), 251–273.
- Carens, Joseph H. (2013): *The ethics of immigration*. New York: Oxford University Press.
- Cerna, L. (2008): "Towards an EU Blue Card? The Delegation of National High Skilled Immigration Policies to the EU Level." COMPAS Working Paper No. 65. Oxford: ESRC Centre on Migration, Policy and Society.
- Determinants." IZA Discussion Papers No. 1590. Bonn: Institute for the Study of Labor.
- Finotelli, Claudia; Michalowski, Ines (2012): "The Heuristic Potential of Models of Citizenship and Immigrant Integration Reviewed." *Journal of Immigrant & Refugee Studies* 10(3), 231–240.
- Gibney, Matthew J. (2004). *The Ethics and Politics of Asylum: Liberal Democracy and the Response to Refugees*. Cambridge: Cambridge University Press.
- Givens, Terri, and Adam Luedtke (2005). "European Immigration Policies in Comparative Perspective: Issue Salience, Partisanship and Immigrant Rights", *Comparative European Politics*, 3(1), 1–22.
- Goertz, Gary (2006). *Social Science Concepts. A User's Guide*. Princeton: Princeton University Press.
- Goodman, S. W. (2015): "Conceptualizing and Measuring Citizenship and Integration Policy: Past Lessons and New Approaches", *Comparative Political Studies* 48(14), 1905–1941.
- Haas, Hein de; Natter, Katharina; Vezzoli, Simona (2015): "Conceptualizing and measuring migration policy change". *Comparative Migration Studies*. DOI: 10.1186/s40878-015-0016-5.

- Haas, Hein de; Natter, Katharina; Vezzoli, Simona (2016): “Growing Restrictiveness or Changing Selection? The Nature and Evolution of Migration Policies”. *International Migration Review*. DOI: 10.1111/imre.12288.
- Hammar, Tomas (1990): *Democracy and the nation state. Aliens, denizens and citizens in a world of international migration*. Aldershot: Ashgate.
- Hatton, T. J. (2004): “Seeking asylum in Europe”. *Economic Policy* 19(38), 6–62.
- Helbling, Marc (2013): “Validating integration and citizenship policy indices”, *Comparative European Politics* 11(5), 555–576.
- Helbling, Marc (2016): ‘Immigration, Integration and Citizenship Policies: Indices, Concepts and Analyses’ In Gary P. Freeman and Nikola Mirilovic (eds.): *Handbook on Migration and Social Policy*. Cheltenham: Edward Elgar, 28–41.
- Helbling, Marc, and Vink, Maarten P. (2013): “The use and misuse of policy indices in the domain of citizenship and integration”, *Comparative European Politics* 11(5), 551–554.
- Helbling, Marc; Bjerre, Liv; Römer, Friederike; Zobel, Malisa (2016): “Measuring Immigration Policies: The IMPIC Database”. *European Political Science*, DOI: 10.1057/eps.2016.4.
- Huddleston, T., and J. Niessen (2011). ‘Migrant Integration Policy Index’. Brussels: British Council and Migration Policy Group.
- Jolliffe, I. T. (2002): *Principal component analysis*. 2nd edition. New York: Springer .
- Joppke, Christian (1998): “Why Liberal States Accept Unwanted Immigration,” *World Politics* 50(2), 266–93.
- Joppke, Christian (2003). ‘Citizenship between De- and Re-Ethnicization’, *European Journal of Sociology*, 44(3), 429–58.
- Joppke, Christian (2005): *Selecting by Origin: Ethnic Migration in the Liberal State*. Cambridge: Harvard University Press.
- Klugman, J., and I. M. Pereira (2009): *Assessment of National Migration Policies: An Emerging Picture on Admissions, Treatment and Enforcement in Developing and Developed Countries*. HDRP-2009-48. New York: Human Development Report Office (HDRO), United Nations Development Programme (UNDP).
- Koopmans, Ruud (2013). “Indices of immigrant rights: What have we learned, where should we go?” *Comparative European Politics* 11(5), 696–703.
- Koopmans, Ruud; Michalowski, Ines; Waibel, Stine (2012): “Citizenship Rights for Immigrants: National Political Processes and Cross-National Convergence in Western Europe, 1980–2008”, *American Journal of Sociology* 117(4), 1202–1245.
- Koopmans, Ruud; Statham, Paul; Giugni, Marco; Passy, Florence (2005): *Contested citizenship. Immigration and cultural diversity in Europe*. Minneapolis: University of Minnesota Press.
- Kriesi, Hanspeter; et al. (2012): *Political conflict in Western Europe*. Cambridge, New York: Cambridge University Press.
- Leong, Victoria; Goswami, Usha; Dick, Frederic (2015): “Acoustic-Emergent Phonology in the Amplitude Envelope of Child-Directed Speech.” *PLoS ONE* 10(12). DOI: 10.1371/journal.pone.0144411.
- Linting, M., Meulman, J.J., Groenen, P.J.F. and Van der Kooij, J.J. (2007) “Nonlinear principal components analysis: Introduction and application“, *Psychological Methods* 12: 336–358.

- Lowell, L. B. (2005): *Policies and Regulations for Managing Skilled International Migration for Work*. New York: United Nations Expert Group Meeting on International Migration and Development.
- Mayda, A. M. (2005) "International Migration: A Panel Data Analysis of Economic and Non-Economic
- Meulman, J.J. and Heiser, W.J. (2005) *Categories 14.0*. Chicago, IL: SPSS.
- Munck, Gerardo L., and Jay Verkuilen (2002). "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices", *Comparative Political Studies* 35(1), 5–34.
- Peters, Margaret E. (2015): "Open Trade, Closed Borders Immigration in the Era of Globalization". *World Politics* 67 (1), 114–154.
- Ruedin, Didier (2014). "Increasing validity by recombining existing indices: MIPEX as a measure of citizenship models", *Social Science Quarterly* 96(2), 629-638.
- Ruhs, Martin (2013): *The Price of Rights. Regulating International Labor Migration*. Princeton: Princeton University Press.
- Safran, William (1997): "Citizenship and Nationality in Democratic Systems: Approaches to Defining and Acquiring Membership in the Political Community", *International Political Science Review* 18(3), 313-335.
- Shachar, Ayelet and Ran Hirschl (2014): 'On Citizenship, States, and Markets', *Journal of Political Philosophy* 22, 231-257.
- Shin, Adrian J. (2016): "Tyrants and Migrants: Authoritarian Immigration Policy". *Comparative Political Studies*. DOI: 10.1177/0010414015621076.
- Thielemann, E. R. (2003): "Does Policy Matter? On Governments' Attempts to Control Unwanted Migration." IIS Discussion Papers No. 9, Dublin: Institute for International Integration Studies, Trinity College.
- Timmer, Ashley S.; Williams, Jeffrey G. (1998): "Immigration Policy Prior to the 1930s: Labor Markets, Policy Interactions, and Globalization Backlash". *Population and Development Review* DOI: 10.2307/2808023.
- Vanhatalo, Erik; Kulahci, Murat (2016): "Impact of Autocorrelation on Principal Components and Their Use in Statistical Process Control". *Quality and Reliability Engineering International* 32(4), 1483–1500.
- Vink, Maarten P.; Bauböck, Rainer (2013): "Citizenship configurations: Analysing the multiple purposes of citizenship regimes in Europe". In *Comparative European Politics* 11 (5), 621–648.

Appendix

I. Correlation matrix: Policy fields and control mechanisms

	Family	Labor	Asylum	Co-ethnic	Control
Family	1.00				
Labor	0.41***	1.00			
Asylum	0.44***	0.32***	1.00		
Co-ethnic	-0.14**	-0.20**	0.05*	1.00	
Control	-0.20***	-0.20***	-0.14**	0.02	1.00

*** p < 0.001; ** p < 0.01; * p < 0.05

Entries are Pearson correlation coefficients (Spearman rank correlation coefficients look very similar)

II. CATPCA with diffuse loadings

The following table shows the CATPCA solution on which we based the follow-up standard PCA for the five-dimensional solution. It displays diffuse loadings, though it is striking that the first dimension also turns out to be both important and consistent here. This is further evidence for our finding of a broad universalistic dimension of immigration policy.

Table A1: Component loadings of a CATPCA: Five-dimensional solution

IMPIC sub-dimensions	Dimension/Component				
	1	2	3	4	5
Family eligibility	.737		.575		
Family conditions	.809		.492		
Family security	.812		.497		
Family rights	.821		.444		
Labor eligibility	.783		-.464		
Labor conditions	.779		-.467		
Labor security	.783		-.465		
Labor rights	.574		-.410		
Asylum eligibility	.578			.594	
Asylum conditions	.644			.410	
Asylum security	.658	.442		.463	
Asylum rights	.678			.437	
Co-ethnics eligibility		.963			
Co-ethnics conditions		.962			
Co-ethnics security		.952			
Co-ethnics rights		.965			
Internal control					.706
External control					.831
Cronbach's alpha	.898	.813	.523	.443	.309
Explained variance (%)	36.557	23.894	10.974	9.551	7.844
Eigenvalues	6.580	4.301	1.975	1.719	1.412

Principal component analysis for categorical data (CATPCA); N = 1023; variable principal normalization; loadings < 0.4 not shown

III. Robustness tests: Some examples

The logic of our robustness tests is simple. We present two tests with distinct examples. First, we choose country-years from 1980, 1995, and 2010 to ensure maximum distance between data points in time (therefore minimizing auto-correlation), but still enough country-years for a PCA having to digest so many variables ($N = 99$). The results are in Table A2 and A3 below. They appear to be robust. The same result is found if we apply the “decade test” to the five-dimensional solution (separate analyses for each decade from 1980 to 2010; not shown here).

Table A2: Component loadings of a CATPCA: Five-dimensional solution

IMPIC sub-dimensions	Dimension/Component				
	1	2	3	4	5
Family eligibility	.733		.533		
Family conditions	.797		.471		
Family security	.802		.452		
Family rights	.810		.440		
Labor eligibility	.790		-.479		
Labor conditions	.786		-.483		
Labor security	.791		-.480		
Labor rights	.593		-.461		
Asylum eligibility	.615			.630	
Asylum conditions	.681			.372	
Asylum security	.681	.433		.445	
Asylum rights	.676			.467	
Co-ethnics eligibility		.956			
Co-ethnics conditions		.955			
Co-ethnics security		.940			
Co-ethnics rights		.959			
Internal control					.607
External control					.717
Cronbach's alpha	.905	.809	.541	.444	.144
Explained variance (%)	38.199	23.558	11.354	9.566	6.433
Eigenvalues	6.876	4.240	2.044	1.722	1.158

Principal component analysis for categorical data (CATPCA); $N = 99$; variable principal normalization; loadings < 0.4 not shown

Table A3: Component loadings of a standard PCA based on the previous CATPCA

IMPIC sub-dimensions	Dimension/Component				
	1	2	3	4	5
Family eligibility quantification		.947			
Family conditions quant.		.925			
Family security quant.		.918			
Family rights quant.		.837			
Labor eligibility quant.			.934		
Labor conditions quant.			.938		
Labor security quant.			.934		
Labor rights quant.			.758		
Asylum eligibility quant.				.917	
Asylum conditions quant.				.704	
Asylum security quant.				.808	
Asylum rights quant.				.880	
Co-ethnics eligibility quant.	.991				
Co-ethnics conditions quant.	.989				
Co-ethnics security quant.	.964				
Co-ethnics rights quant.	.991				
Internal control quant.					.807
External control quant.					.828
Cronbach's alpha	.993	.961	.935	.879	.670
Explained variance (%)	38.199	23.558	11.354	9.566	6.433
Eigenvalues	6.876	4.240	2.044	1.722	1.158

PCA based on the transformed variable quantifications of the CATPCA in Table 2; N = 99; extraction criterion: Eigenvalue > 1; VARIMAX rotation with Kaiser normalization; Cronbach's alpha calculated only for the items that belong to the dimension; loadings < 0.4 not shown

Second, we take another one of our main findings, the one-dimensional solution, to see whether this dimension is present in the same way across all three decades covered by the IMPIC. We argued that the resulting comprehensive universalistic dimension of immigration policy can be seen in the context of re-structuring political conflict in the age of globalization. However, as we all know, globalization in 2010 was not what it was in 1980; though the term was already used back then. Hence, the robustness check assesses whether the empirical implication of our argument extends back to the 1980 world of immigration policy-making.

Based on the first test during the 1980s, this hypothesis is corroborated (Table A5). The same holds for the 1990s (Table A6). However, quite ironically, it does not hold where we are most certain about the argument. From 2000 to 2010 the situation is radically different (Table A7). Family and labor immigration policies still load highly, albeit negatively on the first dimension. Asylum has much lower loadings, though, and does not seem to fully belong to this dimension any more. Instead, co-ethnic policies come to the fore—and they are in juxtaposition to the other policy fields, as only their items load positively on the dimension.

What do we make of this? First, it seems that during the first decade of the new millennium something changed (indeed, as we all know, at the dawn of the new millennium, there was an event that changed the world fundamentally and in ways that are very relevant

to immigration). Among other things, states strengthened efforts to select immigrants by origin (Joppke 2005). This would not be big news for our dimensionality tests so far because we could expect that this to play out in the particularistic dimension of an immigration regime. Instead, however, co-ethnic policies now appear to creep into the universalistic dimension—that is, particularistic policies and universalistic policies no longer seem independent. Instead, the evidence suggests a trade-off between openness toward co-ethnic and openness toward other immigrants. This may be read as two countertrends meeting face to face: the re-ethnicization of citizenship and immigration policies directed toward diasporas, and the de-ethnicization of citizenship and immigration policies upholding universalist and non-discriminatory ideals (cf. Joppke 2003). This taints our analysis with incomplete robustness. However, this finding is so interesting that it seems worth investigating more deeply in order to connect it to broader trends over these decades.

**Table A4: Component loadings of a CATPCA:
One-dimensional solution for 1980s**

IMPIC sub-dimensions	Dimension 1
Family eligibility	.660
Family conditions	.772
Family security	.801
Family rights	.761
Labor eligibility	.859
Labor conditions	.859
Labor security	.859
Labor rights	.848
Asylum eligibility	.679
Asylum conditions	.546
Asylum security	.570
Asylum rights	.683
Co-ethnics eligibility	-.358
Co-ethnics conditions	-.300
Co-ethnics security	-.350
Co-ethnics rights	-.304
Internal control	-.344
External control	-.216
Cronbach's alpha	.914
Explained variance (%)	40.728
Eigenvalue	7.331

Principal component analysis for categorical data (CATPCA); N = 330; variable principal normalization; loadings > 0.4 bold

**Table A5: Component loadings of a CATPCA:
One-dimensional solution for 1990s**

IMPIC sub-dimensions	Dimension 1
Family eligibility	.760
Family conditions	.777
Family security	.774
Family rights	.804
Labor eligibility	.791
Labor conditions	.796
Labor security	.802
Labor rights	.566
Asylum eligibility	.591
Asylum conditions	.679
Asylum security	.621
Asylum rights	.628
Co-ethnics eligibility	-.367
Co-ethnics conditions	-.383
Co-ethnics security	-.262
Co-ethnics rights	-.243
Internal control	-.120
External control	-.188
Cronbach's alpha	.901
Explained variance (%)	37.214
Eigenvalue	6.698

Principal component analysis for categorical data (CATPCA); N = 330; variable principal normalization; loadings > 0.4 bold

**Table A6: Component loadings of a CATPCA:
One-dimensional solution for 2000s**

IMPIC sub-dimensions	Dimension 1
Family eligibility	-.602
Family conditions	-.574
Family security	-.628
Family rights	-.751
Labor eligibility	-.685
Labor conditions	-.722
Labor security	-.743
Labor rights	-.666
Asylum eligibility	-.480
Asylum conditions	-.309
Asylum security	-.378
Asylum rights	-.281
Co-ethnics eligibility	.710
Co-ethnics conditions	.789
Co-ethnics security	.713
Co-ethnics rights	.727
Internal control	-.277
External control	.218
Cronbach's alpha	.895
Explained variance (%)	35.952
Eigenvalue	6.471

Principal component analysis for categorical data (CATPCA); N = 363; variable principal normalization; loadings > 0.4 bold