Conceptualizing and Measuring Citizenship and Integration Policy: Past Lessons and New Approaches

Sara Wallace Goodman[1]

Abstract

Research on immigration and citizenship has become one of the fastest growing areas in political science, with one trend being the boon of comparative citizenship and immigration policy indices. This article discusses methodological concerns with this enterprise. The first half addresses issues with policy indices, including (a) concept validity and boundary maintenance and (b) measurement, compensability, and aggregate index use. The second half examines why these problems matter in hypothesis testing and for inference by replicating three policy index–using studies, rerunning analyses with different indices to test consistency of findings. These tests underscore a central finding: What scholars know about the effects of immigration and citizenship policy is subject to data and sample selection. The article concludes with a number of recommendations and strategies for moving forward. These approaches will not only strengthen this growing research agenda but also mainstream migrant-related policy studies into larger literatures in comparative politics.

Keywords

citizenship, immigration, integration, methodology

[1]University of California, Irvine, USA

Corresponding Author:
Sara Wallace Goodman, Department of Political Science, University of California, Irvine, 3151 Social Science Plaza B, Irvine, CA 92697-5100, USA.
Email: s.goodman@uci.edu

Introduction

Research on immigration and citizenship has become one of the fastest growing areas in political science, with an increasing number of articles in major disciplinary journals (Hollifield & Wong, 2013), the formation of a new American Political Science Association (APSA) section on Migration and Citizenship, and a proliferation of subfield journals accompanied by rising impact factors. This surge is not unwarranted; immigration has brought about profound change to conceptions of national membership and belonging in nation-states. The impact of immigration is particularly pronounced in Western Europe, where demographic shifts, economic transformation, and party politics make immigration a central issue.

Although much of the first wave of research has been qualitative in nature (e.g., Brubaker, 1992; Favell, 1998; Schain, 2008), a prolific new research trend seeks to capture and explain state policy and practice through the development and analysis of policy indices. Examples range from systematic and comparative examinations of citizenship (European Union Democracy Observatory [EUDO] on Citizenship, 2011; Howard, 2009; Janoski, 2010) and immigrant integration policy (Banting & Kymlicka, 2013; Hutcheson & Jeffers, 2012; Koopmans, Michalowski, & Waibel, 2012; Koopmans, Statham, Giugni, & Passy, 2005; Migration Policy Group [MPG], 2011) to newer schema for scoring and classifying immigration policy more generally. These endeavors are welcome as they allow researchers to examine more cases (both across space and time) and, thus, generate more generalizable inferences through rigorous hypothesis testing (more than even the most adroit qualitative research could manage). Moreover, indices enable researchers to build generalizable understandings across political science subfields, offering new perspectives and data on questions such as what impact far-right parties have on immigrant-related policies, whether citizenship policies and status affect an immigrant's socioeconomic integration, as well as myriad questions on citizen attitudes and behavior. Although research into these questions is made possible through this index proliferation, we can observe important differences in how conceptually similar policies are scored (Helbling, 2013). Furthermore, despite index proliferation, we have also seen only limited examples of index implementation (e.g., Bloemraad & Wright, 2014; Dinesen & Hooghe, 2010; Dronkers & Vink, 2012; Fitzgerald, Leblang, & Teets, 2014; Koopmans et al., 2012; Wright & Bloemraad, 2012). That makes this an opportune time to take stock of existing integration and citizenship indices—inspired by ongoing methodological discussions that have positively shaped the field of democracy studies[1]—and apply some practical notes that might aid in ongoing development and overall use in political and sociological analysis.

This article identifies two central concerns with migration and citizenship policy indices. The first is concept validity. Integration, citizenship and immigration policies all make members or "insiders" out of immigrants, but through very different procedures and policies. And while there are policy overlaps in existing indices, there are also significant differences in policy dimension coverage. In the absence of shared definitions and vigorous boundary maintenance, this variation will only perseverate to create substantive problems of interpretation and inference. Careful and informed index selection can avoid conflation of concepts and yield more accurate interpretations of the explanatory power of certain institutional structures, particularly where an effect of policy is paired with a corresponding theoretical justification.

The second issue is measurement. This is not so much a concern that policies are measured accurately, but rather that measures validly reflect the concepts they purport to represent. One reason to question the validity of indices is that they only convey what policies are intended to do. Actual outcomes, whether intended or accidental, can be quite different and this has direct consequences on analysis. Second, they operate from a functionalist assumption, that policies are designed to have an effect as opposed to more symbolic objectives, for example, signaling. A third reason is that all citizenship and integration indices assume an additive relationship between policies, aggregating static policy dimensions—from labor market practices to access to citizenship to family reunification provisions—into one index or several subindices. This neglects the possibility that these subpolicies differently affect outcomes, where some policies are more determinist than others. This can be addressed through explicit theorization about the relationship between policy dimensions preceding not only analysis but also operationalization and measurement.

These problems of index construction directly affect hypothesis testing and inference, whereby what scholars "know" about the causes and consequences of immigration and citizenship policy are vulnerable to these measurements and their flaws. For example, there is a lack of consensus in the literature over the question of whether integration policies help (Hoehne & Michalowski, 2015; Wright & Bloemraad, 2012), hinder (Koopmans, 2010), or having a limited effect (Ersanilli & Koopmans, 2010; Goodman & Wright, 2015) on immigrant integration outcomes. These conflicting understandings are sustained by problems of concept validity and measurement. Therefore, the final section of this article illustrates how problems of concept validity and measurement affect hypothesis testing and knowledge accumulation through a series of replication and replacement tests (where I rerun popular analyses with conceptually similar indices). The three replacement studies

conducted here each highlight a different consequence of the aforementioned index issues: Koopmans et al. (2012) on issues of boundary maintenance, Dinesen and Hooghe (2010) on index holism and theoretical specification, and Wright (2011) on sample selection. Each in their own way, these examples show that when we change the index—even when it is conceptually similar, covers a broadly similar sample of cases, and strongly correlates— we observe divergent results, thus altering what we cumulatively know about integration and citizenship policy. This does not mean rejecting a null hypothesis about policy effects or identifying the "true" effect of policy, but rather increasing our understanding of appropriate scope conditions and conceptual contexts of found results.

The article concludes by offering a series of prescriptive recommendations for index users. These suggestions do not fix the problems of indices but rather help the user navigate through them, and range from increased clarity over scope conditions to greater use of robustness checks and guidelines for theory-informed index selection. The conclusion also advocates complementary qualitative, case-based research to verify findings when possible to discern mechanisms and map policy impact. These exercises will contribute to an increase in theory-building and testing, a strengthening of this growing, interdisciplinary research agenda, and the "mainstreaming" of citizenship and integration studies into larger literatures in comparative politics.

A Survey of Existing Indices

The comparative study of citizenship and integration policy has come a long way. What started out as a vestige of the nationalism literature—in which understandings of and policies for conferring citizenship were defined as mirror images of either ethnic or civic concepts of nationhood (Brubaker, 1992)—has expanded considerably to include dynamic measures and analyses of policy. This is visible in the array of increasingly sophisticated indicators and multimethod studies. Many would point out that the near-frenzied interest in identifying and analyzing citizenship policy by scholars is matched only by the activity of policy makers themselves, where the past 20 years have seen more changes to nationality laws than the previous 50 combined.

Indeed, the 2000s has been a boon decade for the development of policy indicators. Today, there are (at a minimum count) five major independent, comprehensive, multicountry indices that measure citizenship and/or immigrant integration policy: the Migrant Integration Policy Index (MIPEX) produced by the MPG (2011); Marc Morjé Howard's (2009) Citizenship Policy Index (CPI); Ruud Koopmans, Statham, et al.'s (2005) and Ruud Koopmans, Michalowski, et al.'s (2012) Indicators for Citizenship Rights of Immigrants

(ICRI); Thomas Janoski's (2010) Barriers to Naturalization Index (BNI); and EUDO Citizenship's (2013) Citizenship Law Indicators (CITLAW).[2] These expansive projects build upon the achievements of earlier comparative indices, including Waldrauch and Hofinger's (1997) Legal Obstacles to Integration (LOI) index and Kymlicka and Banting's (2006, 2013) Multicultural Policy (MCP) index. There are also a number of recent as well as ongoing efforts to do similar classificatory work with immigration policy, including the ongoing International Migration Policy and Law Analysis (IMPALA; Beine et al., 2015; Gest et al., 2014) and Immigration Policies in Comparison (IMPIC; Helbling, Bjerre, Römer, & Zobel, 2015) database projects, as well as Martin Ruhs's (2013) Openness and Migrant Rights Indicators. Where the migration indices show more conceptual discipline and scope than citizenship and integration indices, much can be gleaned from these earlier experiences, in which, I argue, the entrepreneurial spirit of developing and applying unique indicators has generally outpaced a concurrent conversation on methodology.[3]

The commonalities and differences of measurement in citizenship and integration policy indices have been explored elsewhere. In particular, Marc Helbling (2013) summarizes a number of recent citizenship and integration indices and, after running a series of correlational tests, concludes that "our expectations are borne out in practice: Indices that cover the same policy fields are highly correlated, whereas those that cover different items are not empirically related to each other" (p. 11). Helbling (2013) continues by noting "correlations between integration and citizenship indices are lower than correlations within the two groups themselves," (p. 11) suggesting not only index overlap within specific policy domains but also important conceptual differences between them. It is, therefore, imperative that indices are an accurate reflection of concepts appropriate to the research question at hand. For the purposes of review here, Table 1 summarizes coverage and content of citizenship and integration policy indices.

In terms of coverage, this table conveys the proliferation of indices, but it does not sufficiently portray their limited geographic treatment. The study of citizenship and integration has largely been confined to Europe. CITLAW provides the most expansive coverage of indicators across Europe (including under-examined countries like Serbia, Macedonia, and Turkey) while MIPEX expands outside of Europe to include countries like Canada, the United States, Japan, and South Korea. Plainly apparent in the table, however, are the significant overlaps in time. Given the late 1990s to mid-2000s is widely acknowledged to be a period of major citizenship change in Europe, we see policy measures taken in similar years: Policy is collected in a preceding "snapshot" year (e.g., 1970s, 1980s, 1990s) as well as after (2005, 2008, 2006) and, sometimes, also during (in the case of ICRI's 2002 score).

Table 1. Citizenship and Integration Policy Index Coverage and Content.

	LOI (1997)	CPI (2009)	MIPEX (2010)	BNI (2010)	MCP (2013)	ICRI (2012)	CITLAW (2013)	Fitzgerald, Leblang, and Teets (2014)
N	8	27	33	18	21	8	36	18
Years	1995	1980s 2008[a]	2010[b]	1970-2005	1980 2000 2010[c]	1980 1990 2002 2008[d]	2011	1980-2006
Citizenship policy								
Access to citizenship (first generation)	X	X	X	X	X	X	X	By birth
Access to citizenship (second generation)	X	X	X	X		X	X	By birth
Material conditions	X	X	X	X		X	X	X
Conditions for settlement	X		X					
Loss/expulsion				Women after marriage		X	X	
Integration policy								
Antidiscrimination			X		X	X		
Access to labor market	X		X			Public service		
Family reunification	X		X			X		

(continued)

Table 1. (continued)

	LOI (1997)	CPI (2009)	MIPEX (2010)	BNI (2010)	MCP (2013)	ICRI (2012)	CITLAW (2013)	Fitzgerald, Leblang, and Teets (2014)
N	8	27	33	18	21	8	36	18
Years	1995	1980s 2008[a]	2010[b]	1970–2005	1980 2000 2010[c]	1980 1990 2002 2008[d]	2011	1980–2006
Political participation						Voting rights		
Education rights			X		X	X		
Cultural and religious rights			X		X	X		

I have excluded other indices from the table, namely, Goodman's (2014) CIVIX, as it is sufficiently analytically distinct, covering only cultural conditions (e.g., language and country knowledge requirements) for entry, settlement, and citizenship. Although there is some overlap, CIVIX is unique in concept and measure (for correlations that illustrate this, see Helbling, 2013). X denotes fulfillment of the cell; LOI = legal obstacles to inclusion; CPI = citizenship policy index; MIPEX = Migrant Integration Policy Index; BNI = barriers to naturalization index; MCP = multicultural policy index; ICRI = Indicators for Citizenship Rights of Immigrants; CITLAW = Citizenship Law Indicators; CIVIX = civic integration policy index.

a. 2014 updates are available in (Howard & Goodman, in press).

b. MIPEX was also collected in 2004 and 2007. However, together, these cannot be treated as time series as different coding rules are used in each data set, making comparability a significant issue. A new wave of comparable data will be released in 2015.

c. The multiculturalism policy index measures in comparable policy areas, but only to the extent that states include or exclude policies of multiculturalism (e.g., measuring parliamentary/constitutional affirmation of multiculturalism, school curriculum, media representation/sensitivity, dress code exemptions, the funding of ethnic groups, funding of bilingual education, and affirmative action for disadvantaged immigrant groups).

d. An updating and expansion of cases is currently underway.

7

In terms of thematic content, each of these indices covers a variety of policy dimensions. The most basic aspects of citizenship are the rules for acquisition. With the exception of CITLAW (which builds over 27 modes of nationality acquisition into four, averaged measures of acquisition as well as 11 modes of loss into two, averaged measures), most indices code only for birth-based acquisition (i.e., whether a state permits citizenship acquisition by being born in the territory—*jus soli*—or not for the first and second generations, or *double jus soli*) and residence-based acquisition (e.g., naturalization). Fitzgerald et al.'s index (problematically labeled a "citizenship policy index" 8 years after Howard, 2006, 2009, published one by the same name) makes no generational distinction in birth-based acquisition, using instead a dichotomous measure of whether a provision exists or not. All indices include a variety of coverage on material conditions. Material conditions for citizenship acquisition in each database apply to conditions of naturalization, including residency duration, language and country knowledge requirements, and whether renunciation of a prior citizenship is warranted. However, more comprehensive databases go beyond this basic structure to include conditions ranging from health, welfare requirements, administrative fees, and wait time. Only the Legal Obstacles to Integration—the first "citizenship" index—and MIPEX include conditions for obtaining permanent residence/settlement as part of its index, even though qualitative work has established this to be an important dimension of obtaining citizenship.[4] Finally, only CITLAW (as previously mentioned) and ICRI cover the loss of citizenship as a significant dimension of policy.

It is where citizenship policy content and integration policy content coverage diverge that concept differentiation becomes highly visible. While all indices include citizenship policy, it is easy to discern exclusive citizenship measures (CPI, BNI, CITLAW, and Fitzgerald et al., 2014) from broader integration measures (legal obstacles to inclusion [LOI], MIPEX, ICRI, multicultural policy index [MCP]), in which citizenship is but one component. In terms of integration policy coverage, MIPEX and ICRI include a wide variety of policies that range from those listed in Table 1 but include such specificity as whether there are state-funded imams in prison (ICRI), Islamic religious classes in state schools (ICRI, MIPEX), whether there are immigrant consultative bodies (ICRI, MIPEX) or equality agencies (MIPEX), as well as whether third-country nationals may take up self-employment under conditions equal to EU nationals (MIPEX), to name but a few. These policies are important, but clearly lie outside the scope of citizenship.

As previously mentioned, more recent endeavors to code and categorize immigration policy take on slightly more differentiation, where differences

between criteria for labor (Ruhs, 2013) and low-skilled (Peters, 2015) migrants render indices more unique to one another. Table 2 summarizes index coverage and content of migration policy indices, including ongoing projects (IMPIC and IMPALA).[5]

We can observe greater coverage by indices in terms of scope and space through the inclusion of a more diverse set of countries across the globe and for longer stretches of time. However, there is still a lot of overlap. Policy coverage between IMPIC and IMPALA, in particular, appears almost indistinguishable: Each covers all categories of immigrants (IMPALA including student regulations; IMPIC including coethnics) and deportation rules. Moreover, where the demarcation between immigration and citizenship is clear—Hammar (1985) distinguishes between *immigration policies* as rules that deal with conditions for entry, while *immigrant policies* deal with conditions of domestic membership after entry, for example, rights and benefits—All these indices include dimensions of both. IMPIC includes immigrant rights and assistance as well as other internal control mechanisms, such as how states regulate or keep tabs on immigrants already within the territory, while IMPALA does a comprehensive coding of citizenship regulation. Databases by Peters (2015) and Fitzgerald et al. also include citizenship rules. Although citizenship indices have been in production for over a decade, none of these later studies expand on existing scholarship, choosing instead to consistently reinvent the wheel.

As it stands, there are at least 10 indices that overlap to some degree on immigration, citizenship, and integration policy, and none build on previous or contemporaneous work. This considerably muddles conceptual distinction. It also raises a question of how much these competing enterprises are progressive in establishing Lakatosian novel facts, where the incentive for index creators is to differentiate and establish a niche rather than build common understanding.

In of itself, index proliferation is not a problem. In fact, many arguments can be made that it strengthens the literature as a whole: The "marketplace of ideas" can accommodate many indices as it increases choice, encourages greater conceptual differentiation, and, in theory, adds more data points in terms of years, countries, policies, and so on. However, with more choice comes an increased obligation for the user to select correctly as the inferences we draw not only shape our understanding of policy but, of greater to concern to comparative politics as a whole, what we know about citizens, government, policy making, and democracies as a whole. Moreover, as immigration is among the central political issues in Europe today, these inferences touch nearly every aspect of political life more generally.

Table 2. Immigration Policy Indices Coverage and Content.

	Ruhs (2013)	Fitzgerald, Leblang, and Teets (2014)	Peters (2015)	IMPALA (Beine et al., 2015; Gest et al., 2014)	IMPIC (Helbling, Bjerre, Römer, & Zobel, 2015)
N	46	18	19	25	33
Years	2009	1980-2006	Various years-2010	1960-2010	1980-2010
Immigration policy					
Border regulation (intake rules)	X	"Entry restriction" (restrictive or liberalized)	X	X	X
By category					
Labor	X		Low skill	X	X
Humanitarian			X	X	X
Family			X	X	X
Other				Students	co-ethnics
Demand restrictions	X		X	X	X
Enforcement (deportation)			X	X	X
Immigrant policy					
Immigrant rights and assistance	X		X	X	X
Citizenship		X	X		
Immigration enforcement					Immigrants within country

There is much more variation in countries covered for migration policy than citizenship. For example, IMPIC and Fitzgerald et al. cover OECD countries, IMPALA and Ruhs's migration and openness indices include a wider swath of cases, while Peters covers a smaller sample of states (from the United States to Saudi Arabia) over a wider period of time (her earliest coding year is Canada in 1783). Unshaded with X denotes fulfillment of the cell; shaded cells denote that certain policy areas (defined in the first leftmost column) are not present in indexes.

Concept Validity and Differentiation

Given the empirical rate of index construction has proceeded without a concurrent discussion of concept validity, and having examined extant measures to identify differentiation in coverage, Robert Adcock and David Collier (2001) advise to work backward to properly identify and maintain boundaries for systematized concepts. Why do concept validity and differentiation matter? Irene Bloemraad and Matthew Wright (2014) offer a simple explanation: "Conceptual clarity matters in translating normative philosophy and political rhetoric into empirical analysis: precision in terms helps identify evidence-based measures" (p. S302). Bear Braumoeller (2014) also hints at the consequences of concept validity on measurement: "In a perfect world of realized observations measured on a tidy, cardinal scale, measurement requires little thought. In practice, we don't often get to ignore all of these issues" (p. 42). But while immigration can easily be distinguished from integration and citizenship as rules that regulate the border, that is, external dimensions of movement, how should we demarcate between integration and citizenship policies?

Conceptual ambiguity partly results from the incentive to differentiate, as stated above, and partly from the interdisciplinary nature of the citizenship and migration subfield itself. Political scientists define citizenship as a legal category and examine policies that confer or block the acquisition of status to produce CPIs (Howard 2009; Fitzgerald et al., 2014) whereas sociologists (ICRI) or publicly oriented research institutes (MIPEX) define citizenship as a series of rights (drawing more directly on the tradition of T. H. Marshall) to produce citizenship *rights* indices (or, to establish differentiation, integration indices). But, in both approaches, citizenship acquisition is one type of integration (legal integration), and sits laterally to other kinds of integration-promoting policies, including access to the labor market (economic integration), opportunities for political participation (political integration), and so on. Therefore, we can interpret citizenship policy as fitting into integration policy indices as a subset category, that is, as one of many indicators. However, integration policies are not a subset of citizenship.[6] The problem is not inherently one of parallel enterprise—each index group measures different concepts—but of concept stretching and incongruence between concept and measurement.

One way to distinguish the member *making* of citizenship policy from the member *shaping* of integration policy is by examining process. Both sets of policy can produce inclusive or exclusive membership, but take distinct paths in doing so. Membership is a vital element of the democratic polity; it both confers legitimacy in democratic political systems where rule

makers are elected and it maintains a sense of common purpose by defining and reaffirming shared traits. Drawing on Charles Tilly's (1996) relational definition of citizenship, we can understand membership as a type of association between an individual and a state as "a continuing series of transactions between persons and agents of a given state in which each has enforceable rights and obligations" (p. 8). Although in contemporary immigrant-receiving societies, citizenship is no longer the exclusive membership category, where new conditions for belonging are attached to the acquisition of other status categories like temporary and permanent residence, this definition establishes the quintessential membership status as one with *enforceable* ties.

As such, there is something consequential about *establishing* membership that differentiates it from *promoting* membership. The former confers obligations to members by conveying status while the latter enables membership by acknowledging minority rights and needs, among other approaches. On a fundamental level, policies that encourage inclusion are not the same as policies that structure formal member creation by establishing conditionality. In other words, one group of policies may seek to enable inclusion whereas another group of policies requires it in some form. Civic integration policies serve as a good illustration. These policies promote "citizen-like" skills, like language proficiency and country knowledge, but through mandatory naturalization and settlement tests, integration courses, contracts, interviews, and so on. Participation seeks to yield overall improvements in immigrant integration, but at their core, they are requirements without which status cannot be obtained (Goodman, 2014).

Although policies of promotion and policies of requirement have similar audiences and outcomes, that is, improving a migrant's formal status or position in the labor market, they are conceptually set apart by distinct processes. As such, we can distinguish citizenship policy (and, where appropriate, residence-establishing policies) as *member conditioning*, while integration policy is defined as *member enabling*. Both are consequential in affecting life chances, but the former proceeds through formal requirements while the latter through group- or individual-based enhancement. One could also think of the difference in terms of agency and hierarchy; in member-conditioning scenarios, the immigrant rises to meet the demands of the state, while in member-enabling scenarios, the state lowers itself to accommodate, promote, and alter the life changes of the immigrant.[7]

This important and fine-grained distinction between citizenship and integration policy is not lost on index designers; conflation is not a problem as differentiated aspects of indices are kept apart (Munck & Verkuilen, 2002). Yet, when aggregate indices are used in analysis and where interpretation

comes into play, one faces the potential pitfall of conceptual slipping. Ariel Ahram (2013) defines this problem as arising "when different taxonomical schemas are used to organize what purports to be the same concepts" (p. 1), typically as qualitative methods array cases in nominal and ordinal categories while quantitative methods array them along basic interval or ratio scales. Ahram (2013) continues by discussing how this can produce analytical ambiguity in terms of categorizing cases, as well as cause difficulty in generating valid generalizations as, in citing John Gerring, "it violates the core assumption of unit homogeneity and equivalence." (p. 8)

It is therefore incumbent upon the user to carefully select and consistently use the measure that reflects the appropriate concept (e.g., member-enabling vs. member-conditioning policies) as defined by the research design. This may even involve index consumers taking pieces of an index and assembling their own (e.g., Dronkers & Vink, 2012). That said, selection is particularly challenging when indices are not already conceptually delineated. For example, Koopmans et al. (2012) seek to analyze the "evolution of immigrant rights" (interchangeably, and confusingly, also referred to as both "citizenship rights for immigrants" and "immigrant citizenship rights"), categorizing indicators as disparate as citizenship access to rights for religious minorities under one masthead (pp. 1211, 1215). This operationalization of "immigrant rights" is clearly distinct from the similarly titled "migrant rights" index by Ruhs, which overlaps only by including citizenship, focusing predominately on territorial access (particularly that of high-skilled works) and the bevy of economic benefits that might attract them.[8] It is also difficult when serial correlational tests presume to establish validity,[9] as Koopmans et al. (2012) claim to establish by comparing ICRI with, for example, Kymlicka and Banting's (2006) more focused look at multicultural policies. Helbling (2013) illustrates this point by way of Goodman's (2014) civic integration policy index (CIVIX), noting that through unique policy coverage of only civic integration requirements across three policy gates (immigration, settlement, and citizenship) and weak correlational test results, CIVIX is not invalid but rather "almost unrelated to any of the other indices and thus seems to represent a different dimension" (p. 12). Correlations are only a test of convergent validity if they cover similar policy dimensions, as Howard (2009) conducts by testing CPI validity against only the "access to nationality" dimensions of both MIPEX and ICRI (pp. 34-35).

Given the post hoc construction of concept validity, and that the index enterprise is now almost a decade gone, the onus rests on users to conduct a process of careful index selection, apply case knowledge to identify appropriate indicators of interest—or informed justification for why an index can be used wholesale—and take a "pragmatic" (Adcock & Collier, 2001; Collier & Adcock, 1999) approach to analysis, recognizing that a single approach may

not always be appropriate. Beginning with the research question, conceptual precision about what is to be examined, whether it is intake rules, the institutional, domestic policy context generally, or specifically requirements of legal status acquisition, should be explicit and precede database selection. Immigration, integration, and citizenship policies each carry different characteristics and different theoretical expectations. Distinguishing member enabling versus member conditioning can shape more conceptually distinct hypotheses and predictions, as policies that *compel* outcomes—from immigrant behavior to levels of trust—proceed down a different causal path than policies that *coerce* outcomes. As such, boundary maintenance of concepts can move apace with empirical analysis so long as careful and informed index selection is conducted, which will yield more accurate interpretations of the explanatory power of certain institutional structures.

Capturing the Realities of Policy in Measurement

The problem of concept validity leads to a second major concern: There are distinct differences between policy design and practice. These differences also manifest in how policy is designed and how it is measured. Echoing problems of conceptual slippage and boundary maintenance, these issues often parallel how policies are categorized qualitatively versus how policies are scored quantitatively. As such, inferences may not be accurate if correspondence is not maintained between concept and measurement.

Regarding the first gap between policy and practice, there is no consistent incorporation of policy practice among citizenship, integration, and immigration indices. Peters and Ruhs both exclude practice in migration policy indices—a difference Ruhs (2013) describes as a concentration on laws and regulations ("rights on paper") instead of rights in practice (p. 12). However, among citizenship indices, both CPI and BNI include naturalization rates to try and correct for the potential gap. Howard (2009) uses naturalization rate as a weighted "corrective" to citizenship policy scores, wherein otherwise inclusive citizenship policies "on the books" might be negatively adjusted if empirical evidence shows a low take-up rate (pp. 209-217). Janoski (2010) places naturalization rates at the center of analysis as his study is primarily and uniquely concerned with a "political-institutional explanation of naturalization and nationality" (p. 15), that is, the effects of nationality law on nationality practice. This issue can only be ameliorated with increased specification and explicit theorization about policy expectations. Interpretation of results must then flow from more-refined expectations, accompanied by rationales for why we may see or empirically observe gaps. For example, many scholars have noted that Germany's 2000 citizenship liberalization "on

the books" did not bring about the anticipated rise in naturalizations (see Green, 2013, p. 341; Schönwälder & Triadafilopoulos, 2012; Street, 2014). This sort of qualification would be appropriate to include in index usage and analysis. Or, consider that policy may not always be designed to produce its purported effect; integration policy may, instead, be a political maneuver to pacify the anxieties of the public (e.g., Permoser, 2012) or pursue an ulterior motive altogether (Goodman & Wright, 2015).

The second gap—between what is designed and how it is measured—is a problem that precedes interpretation and applies to index construction, where policy domains carry distinct values but indices do not provide clear guidelines of a theorized relationship for potential interactions. Composite indices assume a priori that no subindex interactions (e.g., labor market access, citizenship access) are meaningful to alter an outcome of interest beyond their contemporaneity, even though they are potentially theoretically distinct (Why should labor market accessibility rules "behave" similarly to family reunification rules?). As such, indices are constructed by simply adding up policy scores.

Furthermore, there are nonadditive interactions that could be relevant, such as multiplicative relationships or those based on a combination of variables that comprise thresholds (e.g., Boolean, fuzzy-set Qualitative Comparative Analysis). As such, theories not only need to predict the anticipated effect of policy on outcomes of interest but also specify the nature of that effect, be it through conjunctural causation ($X1$ and $X2$ produce Y), disjunctural causation ($X1$ or $X2$ produce Y), or "INUS" ("insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result") causation ([$X1$ and $X2$] or [$X3$ and $X4$] produce Y).[10] If $X1+...+XN$ are all collapsed into one variable—X—we forfeit increased analytical precision. This is not a problem of indices per se, where transparent coding and disaggregate scores are reported, but a particular concern where aggregate scores are used in modeling without a priori considerations of subpolicy effects or interactions. Through wholesale employment of indices, users assume an additive relationship between policies and may miss the true relationship at work, or misrepresent a policy's effect altogether.

Finally, aggregate scores—or even subscale averages of indicators—can produce problems of compensability. A brief example to illustrate this point should suffice. MIPEX (MPG, 2011) includes 148 policy indicators across seven different policy areas on migrant integration (labor market mobility, family reunion, education, political participation, long-term residence, access to nationality, and antidiscrimination). Each indicator within these seven fields is scored between 0 and 100, where 100 represents the "most inclusive" or "most favorable to immigrant integration." Total, aggregated

scores for Germany and the United Kingdom are both 57/100. Subscale scores differ widely in that some exhibit large margins (e.g., antidiscrimination scores are 48 and 86, respectively) while others are scored identically, namely, "access to nationality" (both scored at 59/100). These subscores would be surprising to any country expert. Although both countries have undergone significant reform in the past 15 years—from Germany's citizenship policy liberalization in 1999 to Britain's "quiet citizenship revolution" (Kelly & Byrne, 2007)—few would categorize those changes as resulting in convergence.[11]

Unpacking these nationality scores, Germany scores high in eligibility (90/100), whereas Britain's score is more modest (65). Looking closely at coding, this difference is largely attributed to the allowance of periods of absence preceding citizenship acquisition in Germany. Moving to dual citizenship, we see Germany with a midrange score on dual citizenship (50) where Britain's full allowance earns a maximal score of 100. These important policy differences are then lost through averaging and aggregation. It is not a traditional problem of compensability, where a concept is not complete if certain components are absent (e.g., democracy without elections),[12] but it distorts policy contours, problematically suggesting that all policies are equally important and, ultimately, complicates inference.

The Effects of Index Selection on Inference: Three Examples

How do these problems of concept validity and measurement ultimately affect analysis? These concerns are not merely academic; they influence the substantive inferences scholars make about popular hypotheses in the immigration, integration, and citizenship literatures. The bevy of (mainly qualitative) subfield literature already examines the causes of policy, focusing on the role of politics (Howard, 2009; Messina, 2007), institutional legacies (Favell, 1998; Hansen, 2000; Janoski, 2010), or both (Goodman, 2014), while the indexing enterprise has opened the door to analyzing policy consequences. Regarding the latter, scholars have found evidence supporting functional hypotheses, where the institutional environment of citizenship policy—along with individual-level predictors[13]—strongly and significantly accounts for citizenship acquisition (Dronkers & Vink, 2012; Janoski, 2010; Vink, Prokic-Breuer, & Dronkers, 2013).[14] In terms of attitudes, Ariely (2012) shows how citizenship mitigates nationalist attitudes, in which the more inclusive a citizenship policy is, the weaker the tie between national identity and xenophobia. Likewise, those who understand national identity in inclusive or civic terms tend to hold more favorable views of immigration

(Citrin, Sears, Muste, & Wong, 2001). Finally, Fitzgerald et al. (2014) show how an inclusive citizenship policy environment is associated with larger bilateral migration flow.

Like citizenship, immigrant integration policy also exhibits some functional effects. Wright and Bloemraad (2012) interrogate the effects of both policies on an immigrant's sociopolitical integration, finding that while an "independent effect of closed or open citizenship regimes is harder to discern" (p. 89), inclusive citizenship policy amplifies the effect of multiculturalism.[15] Not only integration policies such as multiculturalism but also labor market preparation and family reunification have also shown to be significant in affecting individual attitudes by positively influencing a migrant's subjective well-being (Hadjar & Backes, 2013) but not level of trust (Dinesen & Hooghe, 2010) or national pride (Reeskens & Wright, 2014). And, for the general public, multiculturalism increases social trust (Kesler & Bloemraad, 2010) and decreases prejudice against minorities (Weldon, 2006).

Given the sheer volume of substantive policy inferences derived from these and other studies, which in turn affect how we understand not only migration politics but also democratic participation and political identity, how much are these insights shaped by index selection and, thus, problems of conceptual boundary maintenance and measurement? To investigate this question, I conduct a series of replication tests, wherein I then replace variables using policy indices with conceptually similar alternatives. The first replication/replacement study—Ruud Koopmans et al.'s (2012) cross-national analysis of determinants of "citizenship rights" policies—examines policy indices as a dependent variable, where $Y = f(X, \text{controls})$, replacing the dependent variable Y for conceptually similar alternatives. The second two studies examine policy indices as independent variables: (i) Peter Dinesen and Marc Hooghe's (2010) analysis of factors shaping immigrant levels of trust in Europe and (ii) Matthew Wright's (2011) study of the effects of multicultural policy, citizenship policy, and social spending on normative boundaries of national belonging. These three studies are representative of the variation in comparative studies in that they vary in the policy index used and subject matter. As such, replication/replacement tests also convey different lessons about the consequences of index selection: Koopmans et al.'s study underlines the importance of boundary maintenance; Dinesen and Hooghe's study raises the issue of compensability and index holism; and Wright's highlights the effects of sample selection, which—while heretofore unmentioned—intimately relates to index selection as a natural consequence. As such, each distinctly showcases how the indices we use determine the kinds of answers we get and, thus, the nature of the inferences we make. Therefore,

acting with an awareness of concept and measurement problems is paramount to increasing analytical validity and inferential utility.

First, a word on the replication and replacement procedure: Data sets were either made available by the initial researcher or constructed based on published guidelines. The first step for each study was to replicate initial findings of relevant models of interest. Second, to move to the replacement stage and examine consistency of findings with those produced by rerunning analyses with alternative indices, I synchronized the *N* of the original model and replacement index to match sample countries and (closest) year. As different indices cover different countries (e.g., MIPEX covering 33 vs. ICRI covering 10), I compare coefficients against identical sample sizes. In most cases, this results in reducing the number of observations.[16] Thus, significance may change from that of the original article. With this maximally cautious modeling strategy, I then test the consistency of arguments across indices. As such, these are not tests to formally reject a null about policy effects or identify the "true effect" of policy. Finally, for the sake of brevity, tables presented here only report coefficients and standard errors of study variables in the replicated and replacement models, though I make reference to the original model's findings where possible and are available (along with full models) in the Supplementary Appendix.

Concept Validity Across Indices

In the first study, Koopmans et al. (2012) "investigate to what extent [historical differences] have endured [in citizenship rights for immigrants] since 1980 and which factors can explain expansions and restrictions" (p. 1203). They construct an original data set—the ICRI—to capture rights regimes as their dependent variable, measuring and aggregating 41 indicators of immigrant-related policy across eight policy arenas, in 10 countries, and at four different points in time (1980, 1990, 2002, and 2008).[17] The authors find evidence that variation in level of immigrant rights is strongly determined by path dependence, in which prior level of rights has a powerful effect on future levels. Keeping states on inclusive paths includes a high share of immigrant-origin voters, while "high vote share of right-wing populist parties reduces subsequent levels of immigrant rights" (Koopmans et al., 2012, p. 1229).

A replication test of the control model yields similar coefficients,[18] in which 1980 level of rights, share of immigrant-origin voters, and vote share of right-wing populist parties are all statistically significant on an *N* of 30 (10 countries in 1990, 2002, and 2008).[19] However, in replacement tests, reported in Table 3, we begin to see some differences. The first replacement tests were run with conceptually most similar indices, in that they cover both integration

Table 3. Comparing Koopmans Et al. With Alternative Multicultural and Citizenship Indices.

	Dependent variable							
	Citizenship rights for immigrants							
	Replicated control model (ICRI 2008)	MIPEX (2007)	Replicated control model (ICRI 2008)	MCP (2002, 2008)	Replicated control model (ICRI 2008)	CPI (2009)	Replicated control model (ICRI 2008)	BNI (1990, 2002)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
European Union membership	ns	ns	ns	ns	ns	ns	ns	ns
Strength of judicial review	ns	ns	−0.073* (0.035)	ns	ns	ns	ns	ns
1980 level of rights (ICRI 1980)	0.890** (0.197)	37.128***a (9.196)	0.550** (0.145)		0.866* (0.246)		0.763*** (0.147)	
1980 level of rights (MCP 1980)				1.435**b (0.255)				
1980 level of rights (CPI 1980)						ns		
1980 level of rights (BNI 1980)								0.377* (0.133)
Share of immigrant-origin voters	ns	ns	ns	ns	ns	0.515* (0.197)	0.039* (0.014)	ns

(continued)

19

Table 3. (continued)

	Dependent variable							
	Citizenship rights for immigrants							
	Replicated control model (ICRI 2008)	MIPEX (2007)	Replicated control model (ICRI 2008)	MCP (2002, 2008)	Replicated control model (ICRI 2008)	CPI (2009)	Replicated control model (ICRI 2008)	BNI (1990, 2002)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote share of right-wing populist parties	ns	ns	−0.018** (0.005)	ns	ns	ns	ns	0.020** (0.005)
Left-party government incumbency	ns	ns	ns	ns	ns	ns	ns	ns
Economic growth	ns	ns	ns	ns	ns	ns	ns	ns
Constant	0.294** (0.073)	68.737*** (3.43)	0.052 (0.093)	1.365 (0.391)	0.306*** (0.088)	−1.766 (2.063)	−0.063 (0.08)	0.176* (0.082)
Adjusted R^2	.72	.67	.80	.62	.62	.54	.75	.74
N	10	10	20	20	8	8	20	20

MIPEX does not produce scores for 1980s policy, so ICRI 1980 was used. ICRI = Indicators for Citizenship Rights of Immigrants; MIPEX = Migrant Integration Policy Index; MCP = multicultural policy index; CPI = citizenship policy index; BNI = barriers to naturalization index; ns = not significant.

a. z = 4.13, p < .001.

b. z = 6.878, p < .001

*p < .05. **p < .01. ***p < .001.

20

and citizenship. In comparing the replicated control model with the MCP for 1980, 2002, and 2008 (Models 3 and 4), path dependence remains strongly significant, but vote share of right-wing populist parties completely loses significance with the MCP data set. With MIPEX II (Models 1 and 2),[20] we observe path dependence coefficients as significant, though the effect is much stronger with MIPEX data. In this test ($N = 10$), both right-wing vote share and share of immigrant voters lose significance.

Moving to citizenship indices—which offer more conceptual precision as a subset of the aforementioned, generalized integration indices—Janoski's (2010) BNI maintains significance for vote share of right-wing populist parties but the replicated sample control model loses significance (Models 8 and 7, respectively). Share of immigrant-origin voters loses significance with BNI data and is only weakly significant in the ICRI data set ($p < .05$). Inversely, using CPI data (Model 6), share of immigrant-origin voters approaches significance ($p < .05$) while 1980 level of rights (measured with CPI 1980 scores) and share of right-wing populist parties both report as not significant. In the replicated sample control model (Model 5), where $N = 8$ to match the N of CPI 2009, only 1980 level of rights (i.e., the path dependent argument) is significant. As significance is so low across the board in these models, this certainly underscores a recommendation to increase sample size where possible. At the same time, it also calls into question the appropriateness of this type of test for studying policy causes. Although both the CPI and synchronized replication model suggest that far-right parties are not significant in policy outcomes, Howard (2009) does make a cogent argument for the effectiveness of far-right parties in blocking citizenship policy liberalization using qualitative data in comparative case studies (also see Messina, 2007).

In sum, cross-index comparisons provide external validity to shore up Koopmans et al.'s argument about path dependence. However, both electoral variables lose significance against integration policy indices (i.e., predominately defined by member-enabling policies) and are significant only in citizenship indices (i.e., exclusively defined by member-conditioning policies). These different findings are particularly notable given Koopmans et al.'s (2012, Table 2) strong reliance on correlation tests to establish reliability: ICRI correlates with BNI at $r = .74$ ($p < .001$), CPI at $r = .74$ ($p < .001$), MCP (cultural rights dimension only) at $r = .81$ ($p < .001$), and MIPEX correlations range from $r = .74$ to $r = .94$ (p at least $< .05$). Yet, as the replacement tests show, conceptually similar and reliable indices do not always yield similar effects. In other words, measurement validity does not always translate to concept validity.

These insights about electoral variables may provide a more focused answer to the question of *under what conditions* voter constituency and

political representation are significant, namely, particularly configurations of citizenship policy, not integration policy. Cross-index findings suggest greater conceptual precision and theoretical specification in the model can lead to stronger findings and more precise inference. Moreover, the low number of countries included in the analysis raises questions on the appropriateness of tests and whether these hypotheses are testable or relevant outside of the European context.

Compensability and Index Holism

Dinesen and Hooghe (2010) examine "how trust develops among non-western immigrants in their new country of residence in Western Europe" (p. 698) by looking at respondents in the second and third waves of the European Social Survey. In addition to a series of individual-level characteristics, they include integration policy as a country level variable to discern "whether migration policy matters for trust." To measure this policy, they use the MIPEX II data set, using aggregate country scores. In the initial study, Dinesen and Hooghe (2010) conclude that "integration policy does not influence trust of immigrants" (p. 715) in either a general sense (as a main effect in a sample size including immigrants and natives) or for either first- or second-generation migrants (estimated by creating interaction terms between each generation and integration policy).

In the replication study,[21] similar results were found with the exception that a general integration policy effect was observed to be approaching significance (at the $p < .10$ level) in addition to the specific effect of policy on first-generation levels (original model reports significance at $p < .05$). This finding is consistent when replacing aggregate MIPEX scores with other generalized indices of integration policy (see Table 4), where both MCP 2000 (Model 2) and ICRI 2002 (Model 4) show similar significance with fewer observations. However, we see some important differences. In the original study, an interaction between first-generation migrants and integration policy approaches significance in affecting trust while the interaction effect for second-generation migrants does not. In replacement studies with fewer observations, the first-generation interaction loses significance whereas the second-generation interaction with policy becomes robustly significant (MCP 2000 at $p < .001$, ICRI 2002 at $p < .05$).[22] The emergence of significance in the second-generation interaction not only reflects how inference is a function of index selection but also presents new opportunities for theorization. Wright and Bloemraad (2012, 2014) have already begun to consider integration differences between first- and second-generation immigrants as unique.

Table 4. Comparing Dinesen and Hooghe With Alternative Multicultural Indices.

	Dependent variable			
	Measures of generalized trust			
	Control revised	MCP (2000)	Control revised	ICRI (2002)
	(1)	(2)	(3)	(4)
Natives reference				
Second generation	0.32 (0.24)	−0.06 (0.20)	0.28 (0.38)	−0.21 (0.49)
First generation	1.62** (0.48)	1.54** (0.39)	1.54 (0.97)	1.42 (0.88)
Country-level variables				
Integration policy	0.07† (0.03)	0.04* (0.01)	0.00 (0.00)	0.27*a (0.11)
Interactions				
Second generation × Policy	−0.03 (0.03)	−0.03** (0.01)	0.00 (0.00)	−0.22*b (0.09)
First generation × Policy	−0.03 (0.05)	0.00 (0.02)	0.00 (0.01)	−0.09 (0.19)
Constant	−1.57** (0.41)	−1.13* (0.46)	−1.19* (0.37)	−0.66† (0.34)
R^2	.25	.25	.25	.25
N	44,346	44,346	32,361	32,361

Study variables only. Full models available in the Supplementary Appendix. MCP = multicultural policy index; ICRI = Indicators for Citizenship Rights of Immigrants.
a. $z = 2.42$, $p < .01$.
b. $z = 2.40$, $p < .01$.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

As integration policy is significant, but only just, it raises a more probing question that moves analysis from *whether* integration policy matters to *why*. To probe whether the conceptual distinctions between member conditioning and member enabling makes a difference in analysis, I replace the aggregate MIPEX variable,[23] which consists of over 100 policy indicators across six different policy areas on migrant integration (the same policy areas as the 2010 wave minus education), with disaggregated variables, simultaneously running each policy area subscore in the control model. Given the diversity of policy categories that otherwise comprise the single, aggregated MIPEX variable, and in light of aforementioned compensability issues, wholesale

usage reflects a type of under-theorization about which dimensions of integration policy might actually have an effect on immigrant trust.

   As we see in Table 5, labor market, family reunion, political participation, and long-term residence lose significance but remain within the standard error. By contrast, access to nationality and antidiscrimination not only remain significant but also become even more so (at $p < .001$). This suggests that the integration policy effect is strongly driven by these two subscales and, thus, that the original inclusion of MIPEX wholesale waters down the effect of integration policy on immigrant trust. Sufficient theorization a priori about why specific dimensions of integration policy might affect immigrant trust would not only increase model precision but also shape new, testable hypotheses about the effects of citizenship and other significant policy areas. Complementing this type of analysis with case studies to probe at causal mechanisms would further strengthen claims about policy effects on immigrant trust.

Sample Selection

Finally, Wright (2011) examines the effect of policy regimes (multiculturalism, citizenship, and social welfare) on shaping the normative boundaries of the national political community according to public opinion. He probes questions on national community membership made available in the National Identity Module of the 1995 and 2003 International Social Survey Program. To measure multicultural policy regimes, Wright uses an early version of the MCP, which tap policies in the domain of "polyethnic rights" (see Kymlicka & Banting, 2006), from bilingual education to school curriculum. To measure citizenship, Wright creates a simple dichotomous variable to score jus sanguinis (citizenship through birth–familial ties) versus jus soli (citizenship through birth in a territory) provisions. The overall findings of the study suggest "multiculturalism does not encourage ethnocentrism" but "it does not appear to have a *positive* effect either" (Wright, 2011, p. 617). Regarding citizenship, Wright finds that citizens in liberal citizenship regimes prioritize "achievable" (as opposed to "ascriptive") markers of belonging more than citizens in restrictive regimes.

   Table 6 replicates the effect of multicultural and citizenship policy regimes on Wright's "two-item ascriptive" index, which identifies the relative priority respondents place on ascriptive boundaries, including ancestry and nativity, to those of achievable attributes, like respect for laws and institutions and "feeling" like a national.[24] In the original study, Wright (2011) observes, "the relationship between multiculturalism policy and these outcomes remains weak" and by contrast jus soli remains "negatively

Table 5. Comparing Dinesen and Hooghe With Disaggregated MIPEX Indices.

	Dependent variable						
				Measures of generalized trust			
	Control revised	Labor	Family reunion	Political participation	Long-term residence	Access to nationality	Antidiscrimination
Natives reference							
Second generation	0.32 (0.23)	0.09 (0.25)	0.15 (0.25)	0.16 (0.23)	0.22 (0.18)	0.22 (0.19)	0.17 (0.20)
First generation	1.62** (0.48)	1.26* (0.57)	1.68* (0.62)	1.50** (0.49)	1.56** (0.63)	1.58** (0.41)	1.62** (0.41)
Country-level variables							
Integration policy	0.01†a (0.00)	0.00 (0.02)	0.02 (0.03)	0.02 (0.03)	0.00 (0.03)	0.06** (0.02)	0.05**a (0.01)
Interactions							
Second generation × Policy	−0.00 (0.00)	0.02 (0.02)	0.01 (0.04)	−0.01 (0.03)	0.00 (0.00)	−0.03*b (0.01)	−0.02 (0.02)
First generation × Policy	−0.00b (0.00)	0.04 (0.04)	−0.02 (0.04)	0.07 (0.05)	0.00 (0.01)	−0.03 (0.03)	−0.06† (0.03)
Constant	−1.57** (0.41)	−1.29* (0.47)	−1.51** (0.44)	−1.27* (0.45)	−1.23* (0.47)	−1.45*** (0.30)	−1.51** (0.36)
R^2	.25	.25	.25	.25	.25	.25	.25
N	45,144	45,144	45,144	45,144	45,144	45,144	45,144

Study variables only. Full models available in the Supplementary Appendix. MIPEX = Migrant Integration Policy Index.
a. z = 5.75, p < .001.
b. z = 2.5, p < .01.
†p < .10. *p < .05. **p < .01. ***p < .001.

25

Table 6. Comparing Wright With Alternative Multicultural and Citizenship Indices for 2003 Data.

	Dependent variable					
			Two-item ascriptive index			
	Replicated model	MCP	Replicated model adjusted	ICRI	Replicated model adjusted	CPI
	(1)	(2)	(3)	(4)	(5)	(6)
Multiculturalism	−0.08 (0.06)		−0.06** (0.02)		−0.25** (0.06)	−0.13* (0.05)
Average MCP		−0.01 (0.01)				
ICRI (2002)				−0.06* (0.03)		
Jus soli	−0.06 (0.04)	−0.07 (0.04)	−0.04* (0.02)	−0.05* (0.02)	−0.12* (0.05)	
CPI (1980)						−0.03** (0.01)
Social spending (2003)	−0.02** (0.004)	−0.02** (0.004)	−0.03** (0.002)	−0.03** (0.003)	−0.02** (0.004)	−0.03** (0.003)
Constant	0.11 (0.12)	0.1 (0.12)	−0.21** (0.05)	−0.18** (0.06)	0.18† (0.1)	0.51** (0.11)
N	12,172	12,172	5,395	5,395	6,903	6,903
Bayesian information criterion	2,521.03	2,525.25	1,474.70	1,475.94	1,130.86	1,132.05

Study variables only. Full models are available in the Supplementary Appendix. MCP = multicultural policy index; ICRI = Indicators for Citizenship Rights of Immigrants; CPI = citizenship policy index.
†$p < .10$. *$p < .05$. **$p < .01$.

associated with ascriptive national identity in a statistically significant way" (p. 611). In replicating Wright (Model 1), multiculturalism[25] remains insignificant but citizenship loses significance (though social spending remains strongly significant).[26]

While the previous two studies highlight a direct problem of index selection—concept validity and compensability in aggregate indices, respectively—this third study showcases a corollary consequence of index selection: sample selection. Policy indices handle nonrandom data, and exclusion based on particular attributes is often the case. None of these indices capture the entire universe of potential cases, that is, every country in the world (or, for that matter, regional or substate policies of integration). Sample selection—typically driven by practical concerns outside of a research design, like language proficiency, funding, or data availability—takes place in which countries and policies are included or excluded from the processes of data collection and coding. As a result, and to recall patterns identified in Table 1, most indices overlap in studying Western Europe, although many also expand beyond in some capacity. For example, MCP includes the Anglo-settler states of Australia, Canada, the United States, and New Zealand, and MIPEX, as previously mentioned, includes Japan and South Korea.

In examining the consistency of Wright's findings with other indices—ICRI to replace Wright's multiculturalism score (Model 4) and CPI in place of his jus soli score (Model 6)—results change drastically. In both models, significance grows as *N* decreases. In the synchronized replication to ICRI 2001 (Model 3)—where the number of observations is cut by over half—multiculturalism and citizenship both become strong in statistical significance (Wright's multiculturalism variable at $p < .01$). In the synchronized replication to CPI 1980 (Model 5), both the control model and the CPI-replaced model (Model 6) boast statistical significance for these policies.

In sum, a smaller *N* amplifies the effect of policy regimes, most likely as it kicks Canada, Australia, and the United States out of the sample. As looking at different countries brings into play different substantive considerations, it is a call to researchers to deliberately articulate potential scope—but also limitations—of inference. It may be that multiculturalism and citizenship policy regimes are increasingly salient in Europe or, on the contrary, particularly nonsalient in Anglo-receiving states. Dronkers and Vink (2012), for example, argue in their study on the effects of citizenship policy on naturalization rates that immigrant naturalization in settler states "has a different dynamic" than the European context (p. 391). Moreover, just as different sample sizes can aid in circumscribing appropriate scope conditions for inference, it can also direct our attention to the potential of omitted variables.

Strategies for Moving Forward

The bumper crop of citizenship and migration indices has propelled the field of study to the precipice of cutting edge analysis. It has also enabled an increasing number of scholars across diverse regions and areas comparative study to take more seriously the role of the immigrant-related policy environment. As such, the take-away from the cautionary tone presented here is not that the indexing endeavor should be abandoned. Quite the contrary, the efficient representation of complex and evolving policy into parsimonious, comparable, and exportable indicators has significant value. Yet, as the issues raised here and examples demonstrate, what we know about these policy environments very much depends on what kind of data we use. Conceptual imprecision—where studies lack a clear differentiation of member-enabling versus member-conditioning policies—can produce misleading conclusions. Theoretically underspecified uses of indices may miss the greater picture of *which* policies are actually having an impact on the outcome of interest and *how*, which in turn affects the growing and cumulative knowledge about migrant-related policies and effects.

There is any number of strategies to ameliorate these problems and move forward. First and foremost, progressive research agendas begin with thoughtful, conceptually consistent, theory-informed, and—above all—explicit index and case selection criteria. Deliberate selection must assess the appropriateness of fit of a potential index in terms of both policy and countries covered. Crucially, it would require an upfront consideration of the selected index and whether it is appropriate to use whole cloth or mine for subscales of interest. Selecting definitional characteristics to a concept is a balancing act, where "maximalist definitions tend to be so overburdened as to be of little analytical use" and minimalist definitions "run the risk of omitting relevant attributes in the definition of a concept" (Munck & Verkuilen, 2002, p. 9). And policy variables need be as orthogonal as possible, particularly where other—and potentially overlapping—policies are under consideration. Aggregate policy measures reduce empirical complexity for the sake of parsimonious comparison. And while this is oftentimes necessary for painting descriptive pictures of policy trends and patterns, as well as increase the likelihood of index utilization by experts in other subfields and research areas, this reduced image may not convey what policies actually are, what policies do, or how policies relate to one another. Finally, it would go a long way to strengthen the credence of selection if users would engage directly with competing indices, articulating the trade-offs of using one index versus another or explicating why the selected index is most appropriate given the abundance of alternatives.

Second, there are not—as of yet—that many testable hypotheses for pre-dicting effects of immigrant-related policy. It is not that models fail to meet correct assumptions; it is that we have very few theoretical expectations to inform our model specifications. Therefore, researchers might consider using indexes not only for hypothesis testing but, as a first step, in hypothesis build-ing. These projects would help to develop broad theoretical arguments by defining parameters and scope conditions as well as refining operationaliza-tion of concepts. They could range from large-*N* data analysis to plausibility probe case studies (Eckstein, 1975), where preliminary studies on untested theory determine whether further and more robust testing needs to take place. These are not theory-confirming but developing exercises. With these efforts, researchers could proceed to examine these findings through empirical test-ing, under new conditions, through comparative case study, or with a larger pool of countries.

Relatedly, a third strategy is to—where possible and space permitting—pair quantitative results with qualitative analysis. An investigation of causal mechanisms through process tracing can not only strengthen arguments and triangulate findings but also identify which elements of policy are driving results. Increased analytical rigor for example, would not only identify which policies exert an impact on an outcome of interest but also identify *how*, that is, specification of causal mechanisms, recognizing that not all citizenship and migration policy dimensions are defined by the same normative expecta-tions nor "behave" similarly in practice. Pairing large-*N* analysis with case studies not only aids at unpicking the causal story, it may be specifically use-ful in addressing the principle-implementation gap, where policy and practice diverge. Moreover, where there is always the possibility that policies arise precisely because they are intended to be symbolic or ineffectual, case studies can play a role in interpreting null findings. At their best, mixed or multi-method approaches validate measures and strengthen inferences. Falling short of these goals, even poorly done complementary methods signal to the reader that they have thought through the issue at hand, which is valuable in of itself.[27]

Finally, there are a few steps in the area of hypothesis testing that can strengthen index-driven studies as a whole. The existing practice to show strength in indices is to run correlations with others indices, as shown in Koopmans et al. But where this only captures measure validity, robustness tests with other indices would help to establish concept validity (see, for example, Citrin, Levy, & Wright, 2014). It would also strengthen confidence in statistical significance which does not necessarily—in of itself—indicate robust results. As such, using different indices help avoid "Type I errors" (false positives). To wit, it bears reminding that replication procedures should

be explicit. Finally, where Europe-only and other medium-*N* studies offer a comparative advantage in reaching results by reducing noise with a narrow set of parameters, increasing the number of cases (i.e., countries) enables researchers to test the extent to which inferences hold across a variety of contexts, including temporal or geographic limits to theories. Case extension should not be haphazard, adding more countries should be a theoretically driven process to test hypotheses across identified variation of interest, such as colonial experience, regime type, GDP, migration flow, and so on.

Knowing which attributes of migration policy matter and how is difficult, and can often require iterated theorization and systematic testing. Yet, it is imperative, given the number of indices, the upward trend of studies on migration and citizenship in the past decade, the interdisciplinary popularity of the field of study, as well as the central role immigration and belonging continue to play in national politics in Europe and elsewhere. Any work that scholars can do on our end to increase analytical rigor, theoretical precision, and empirical insight seems well worth the time.

Notes

1. This literature spans over a decade not only from Munck and Verkuilen (2002) to Lindberg, Coppedge, Gerring, and Teorell (2014) but also include Coppedge et al. (2011) and others.

2. Of note, Citizenship Law Indicators (CITLAW) provides citizenship indicators but does not report them in an aggregate index per se, appropriately leaving it to the researcher/modeler to theorize the relationship between, for example, *jus soli* and naturalization. In terms of other indices, the first citizenship policy index was the legal obstacles to inclusion (LOI) by Waldrauch and Hofinger

(1997). Also see Keith Banting and Will Kymlicka's (2013) multiculturalism policy index (MCP), Sara Wallace Goodman's (2014) civic integration policy index (CIVIX), and Edward Koning's (2012) unnamed index.

3. The only exceptions are Waldrauch and Hofinger (1997), Goodman (2012), and Bjerre, Helbling, Römer, and Zobel (2014).

4. For example, in some cases (e.g., Austria, Denmark, France, Luxembourg, The Netherlands, and Spain), an immigrant must hold permanent residence status prior to applying for citizenship (see Goodman, 2010, p. 8).

5. I have excluded immigration studies that only code policy change, for example, de Haas, Natter, and Vezzoli (2014), as they do not convey levels at t_0 for purposes of comparability.

6. The one exception is *civic* integration requirements that make status acquisition conditional upon completion or language or country knowledge.

7. This, of course, does not imply that the immigrant is passive in the member-enabling approach. Proactive behavior is required to take advantage of state policy (e.g., allowance for mother-tongue education, carving out political space for Islamic consultative bodies, etc.), but the point is that nonparticipation does not yield a negative decision on the formal acquisition of status.

8. Of course, the inclusion of citizenship in both of these indices itself can be debated as to whether it is an "immigrant right" at all, where states require some degree of closure for effective governance and resource allocation (Aristotle, 1984). One can also debate what length of residence and level of material requirements is normatively "acceptable" as a barrier delineating insiders and outsiders.

9. Index validity is crucial as there is no authoritative, comparative source for national citizenship or integration policy. As such, researchers rely on interpretation of a variety of documents, from national law to amendments to procedural guidelines to circulars. Moreover, some studies rely on immigration scholars to interpret these documents (Migrant Integration Policy Index [MIPEX]) while others rely on legal scholars (CITLAW) or individual interpretation of secondary material (Indicators for Citizenship Rights of Immigrants [ICRI], BNI). Some interpretations are understandably more nuanced than others and some national law clearer than others.

10. These examples are drawn from Braumoeller (2014, p. 43).

11. As such, their divergence is visible in other scores. Citizenship policy index [CPI] scores Germany's citizenship policy in 2008 at 2.04/6, whereas the United Kingdom scores much more liberal at 4.97. In presenting disaggregated *jus soli* and naturalization scores, CITLAW also reveals some differences: Germany's *jus soli* liberalization only scores 0.48/1 (compared with the United Kingdom's 0.6) with naturalization at 0.57 (compared with 0.72).

12. Compensability would be even more problematic were we to expect certain thresholds, for example, a cutoff point between "inclusive" and "exclusive" practices, which MIPEX does not state. Howard's CPI, for example, establishes these. Distributed between a scale of 0 to 6, scores between 0 to 1.5 are

"restrictive," 1.5 to 4 are "medium," and 4 to 6 are "liberal." However, because each of the three citizenship policy dimensions (jus soli, naturalization, and dual citizenship) is theoretically driven and capped at a maximum score of 2, the absence of values in one area cannot be hidden by high values in another, as is possible in MIPEX.

13. Examples include speaking host country language, married or not, education level, skill level, and so on. For more, see Yang (1994) and Chiswick and Miller (2008).

14. These insights are complemented by research that shows it is not only policy but procedures—that is, the implementation and practice of policy—that matter (Bloemraad, 2002; Huddleston, 2014).

15. Although evidence shows that inclusive citizenship *policy* facilitates socioeconomic inclusion, research on actual citizenship *status* is more mixed. Political scientists have determined that status does not affect socioeconomic incorporation (Ersanilli & Koopmans, 2010; Maxwell, 2012) whereas economists have shown citizenship status as significant in employability and income (Bevelander & DeVoretz, 2008).

16. Functionally, this also limits the comparison with observations in Western Europe, speaking to the issue of sample selection which I discuss later.

17. Authors are currently undergoing an expansion of this index, see icri.wzb.eu

18. Note that other variables are recorded as "*ns*" (not significant). Due to the process of stepwise regression, in which models are built by successively adding variables that meet the selection criterion, variables that do not meet this threshold are excluded from analysis without generating coefficients.

19. For this and other replication tests, see Supplementary Appendix.

20. MIPEX does not cover the same amount of time as ICRI, as it only has comparable measures for 2007 and 2010. Therefore, I pick only 1 snapshot year for comparison—MIPEX 2007 and ICRI 2008, thus lowering the sample size of both to $N = 10$.

21. Replication is not exact to the reported original model as discrepancies were found in MIPEX II calculations. My results import MIPEX II scores directly from the MIPEX.eu website.

22. This effect is also visible in the "Access to Nationality" subscale of the disaggregated MIPEX model (at $p < .05$).

23. Each subindicator is averaged within its category to produce a single score for each policy dimension, which is then added together to produce an aggregate score for each country.

24. Wright's study runs 24 different models for the year 2003 alone, conducting tests where policy variables are entered separately as well as simultaneously. I replicate the latter, as shown in Table 4, Panel B, column 3 of the original article (Wright, 2011, p. 613).

25. Wright uses an early version of the multicultural policy index (MCP), averaging 1980 and 2000 scores. In the first replication (Model 2), I use the formal scores from MCP averaging the same years and results are consistent. These two averages correlate at $r = .95$.

26. Wright's reported coefficients differ from the replication model. It could be argued that the missing minority variable invalidates this replication, but this is

unsubstantiated on both empirical and theoretical grounds, not to mention misses the point of the replication here. Empirically, in the original analysis, coefficients associated with the ethnic minority variable were zero and statistically insignificant. Therefore, in terms of actual effect, being an ethnic minority had no statistically determinable effect on attitudes. Theoretically, this replication tests the consistency of Wright's conclusions using different citizenship and multiculturalism policy specifications. Minority status is a variable that, aside from its name, appears to otherwise have no effect on the dependent variable. Its inclusion or exclusion should not therefore affect robust correlations (social spending, for example, retains its sign and significance throughout). The replication provided here—kindly guided through personal communication with Wright—uses the same method and includes the same number of observations, supporting my argument that citizenship and multiculturalism indicators are problematic. As a check, differences were not due to minute differences in how software packages compute mixed effects models, as replications in R and STATA 13 corroborated one another.

27. I wish to thank an anonymous reviewer for this point.

Supplemental Material

The supplementary appendices are available at http://cps.sagepub.com/suplemental

References

Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, *95*, 529-546.

Ahram, A. I. (2013). Concepts and measurement in multimethod research. *Political Research Quarterly*, *66*, 280-291.

Ariely, G. (2012). Do those who identify with their nation always dislike immigrants? An examination of citizenship policy effects. *Nationalism and Ethnic Politics*, *18*, 242-261.

Aristotle. (1984). *The politics*. Chicago, IL: University of Chicago Press.

Banting, K., & Kymlicka, W. (2013). Is there really a retreat from multiculturalism policies? New evidence from the multiculturalism policy index. *Comparative European Politics*, *11*, 577-598.

Beine, M., Boucher, A., Burgoon, B., Crock, M., Gest, J., Hiscox, M., . . . Thielemann, E. (2015). Comparing immigration policies: An overview from the IMPALA database. *International Migration Review*. Advance online publication. doi:10.1111/imre.12169

Bevelander, P., & DeVoretz, D. J. (2008). *The economics of citizenship*. Malmö, Sweden: Malmö University Press (MIM).

Bjerre, L., Helbling, M., Römer, F., & Zobel, M. (2014). Conceptualizing and measuring immigration policies: A comparative perspective. *International Migration Review*. Advance online publication. doi:10.1111/imre.12100

Bloemraad, I. (2006). *Becoming a citizen: Incorporating immigrants and refugees in the United States and Canada*. Berkeley: University of California Press.

Bloemraad, I., & Wright, M. (2014). "Utter failure" or unity out of diversity? Debating and evaluating policies of multiculturalism. *International Migration Review, 48*(Suppl. s1), S292-S334.

Braumoeller, B. (2014). Analyzing interactions: Four alternative models. *Newsletter of American Political Science Association Organized Section on Qualitative and Multi-Method Research*, *12*(1), 41-46.

Brubaker, R. (1992). *Citizenship and nationhood in France and Germany*. Cambridge, MA: Harvard University Press.

Chiswick, B. R., & Miller, P. W. (2008). *Citizenship in the United States: The roles of immigrant characteristics and country of origin* (IZA Discussion Paper, Vol. 3596). Retrieved from http://repec.iza.org/dp3596.pdf

Citrin, J., Levy, M., & Wright, M. (2014). Multicultural policy and political support in European democracies. *Comparative Political Studies*, *47*, 1531-1557.

Citrin, J., Sears, D. O., Muste, C., & Wong, C. (2001). Multiculturalism in American public opinion. *British Journal of Political Science*, *31*, 247-275.

Collier, D., & Adcock, R. (1999). Democracy and dichotomies: A pragmatic approach to choices about concepts. *Annual Review of Political Science*, *2*, 537-565.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., . . . Teorell, J. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*, 247-267.

de Haas, H., Natter, K., & Vezzoli, S. (2014, March 23). *Compiling and coding migration policies: Insights from the DEMIG POLICY database* (IMI Working Paper 87/DEMIG Paper 16). Oxford, UK: International Migration Institute, University of Oxford.

Dinesen, P. T., & Hooghe, M. (2010). When in Rome, do as the Romans do: The acculturation of generalized trust among immigrants in Western Europe. *International Migration Review*, *44*, 697-727.

Dronkers, J., & Vink, M. P. (2012). Explaining access to citizenship in Europe: How citizenship policies affect naturalization rates. *European Union Politics*, *13*, 390-412.

Eckstein, H. (1975). Case studies and theory in political science. In F. I. Greenstein & N. W. Polsby (Eds.), *Handbook of political science: Political science: Scope and theory* (Vol. 7, pp. 94-137). Reading, MA: Addison-Wesley.

Ersanilli, E., & Koopmans, R. (2010). Rewarding integration? Citizenship regulations and the socio-cultural integration of immigrants in the Netherlands, France and Germany. *Journal of Ethnic and Migration Studies*, *36*, 773-791.

European Union Democracy Observatory on Citizenship. (2013). *Citizenship Law Indicators (CITLAW)*. Retrieved from http://eudo-citizenship.eu/indicators/eudo-citizenship-law-indicators

Favell, A. (1998). *Philosophies of integration: Immigration and the idea of citizenship in France and Britain*. New York, NY: St. Martin's Press.

Fitzgerald, J., Leblang, D., & Teets, J. C. (2014). Defying the law of gravity: The political economy of international migration. *World Politics*, *66*, 406-445.

Gest, J., Boucher, A., Challen, S., Burgoon, B., Thielemann, E., Beine, M., . . . Hiscox, M. (2014). Measuring and comparing immigration, asylum and naturalization

policies across countries: Challenges and solutions. *Global Policy*, *5*(3), 261-274. doi:10.1111/1758-5899.12132

Goodman, S. W. (2010). *Naturalisation policies in Europe: Exploring patterns of inclusion and exclusion* [EUDO Citizenship Comparative Reports (RSCAS/EUDO-CIT-Comp. 2010/7)]. Florence, Italy: EUDO Citizenship, Robert Schuman Centre for Advanced Studies, EUI.

Goodman, S. W. (2012). Measurement and interpretation issues in civic integration studies: A rejoinder. *Journal of Ethnic and Migration Studies*, *38*, 173-186.

Goodman, S. W. (2014). *Immigration and membership politics in Western European*. New York, NY: Cambridge University Press.

Goodman, S. W., & Wright, M. (2015). Does mandatory integration matter? Effects of civic requirements on immigrant socioeconomic and political outcomes. *Journal of Ethnic and Migration Studies*. Advance online publication. doi:10.1080/1369183X.2015.1042434

Green, S. (2013). Germany: A changing country of immigration. *German Politics*, *22*, 333-351.

Hadjar, A., & Backes, S. (2013). Migration background and subjective well-being. A multilevel analysis based on the European social survey. *Comparative Sociology*, *12*, 645-676.

Hammar, T. (1985). *European immigration policy: A comparative study*. New York, NY: Cambridge University Press.

Hansen, R. (2000). *Citizenship and immigration in post-war Britain: The institutional origins of a multicultural nation*. New York, NY: Oxford University Press.

Helbling, M. (2013). Validating integration and citizenship policy indices. *Comparative European Politics*, *11*, 555-576.

Helbling, M., Bjerre, L., Römer, F., & Zobel, M. (2015). Measuring immigration policies: The IMPIC database. *European Political Science*.

Hoehne, J., & Michalowski, I. (2015). Long-term effects of language course timing on language acquisition and social contacts: Turkish and Moroccan immigrants in Western Europe. *International Migration Review*. Advance online publication. doi:10.1111/imre.12130

Hollifield, J. F., & Wong, T. K. (2013). International migration: Cause or consequence of political change? *Newsletter of American Political Science Association Organized Section on Migration and Citizenship*, *1*(1), 3-9.

Howard, M. M. (2006). Comparative citizenship: An agenda for cross-national research. *Perspectives on Politics*, *4*, 443-455.

Howard, M. M. (2009). *The politics of citizenship in Europe*. New York, NY: Cambridge University Press.

Howard, M. M., & Goodman, S. W. (in press). The politics of citizenship and belonging in Europe. In C. M. Swain (Ed.), *Debating immigration* (2nd ed.). New York, NY: Cambridge University Press.

Huddleston, T. (2014). *Between policies and outcomes: Surveying immigrants' perceptions of the obstacles in naturalisation procedures*. Paper presented at the 21st International Conference of Europeanists, Washington, DC.

Hutcheson, D., & Jeffers, K. (2012). *Citizenship status and the integration of immigrants: CITINT indicators*. Florence, Italy: EUDO Citizenship, Robert Schuman Centre for Advanced Studies.

Janoski, T. (2010). *The ironies of citizenship: Naturalization and integration in industrialized countries*. New York, NY: Cambridge University Press.

Kelly, R., & Byrne, L. (2007). *A common place*. London, England: Fabian Society.

Kesler, C., & Bloemraad, I. (2010). Does immigration erode social capital? The conditional effects of immigration-generated diversity on trust, membership, and participation across 19 countries, 1981-2000. *Canadian Journal of Political Science*, *43*, 319-347.

Koning, E. A. (2011). Ethnic and civic dealings with newcomers: naturalization policies and practices in twenty-six immigration countries. *Ethnic and Racial Studies*, *34*, 1974-1994.

Koopmans, R. (2010). Trade-offs between equality and difference: Immigrant integration, multiculturalism and the welfare state in cross-national perspective. *Journal of Ethnic and Migration Studies*, *36*, 1-26.

Koopmans, R., Michalowski, I., & Waibel, S. (2012). Citizenship rights for immigrants: National political processes and cross-national convergence in Western Europe, 1980-2008. *American Journal of Sociology*, *117*, 1202-1245.

Koopmans, R., Statham, P., Giugni, M., & Passy, F. (2005). *Contested citizenship: Immigration and cultural diversity in Europe*. Minneapolis: University of Minnesota Press.

Kymlicka, W., & Banting, K. G. (2006). *Multiculturalism and the welfare state: Recognition and redistribution in contemporary democracies*. Oxford, UK: Oxford University Press.

Lindberg, S. I., Coppedge, M., Gerring, J., & Teorell, J. (2014). V-Dem: A new way to measure democracy. *Journal of Democracy*, *25*, 159-169.

Maxwell, R. (2012). *Ethnic minority migrants in Britain and France: Integration trade-offs*. New York, NY: Cambridge University Press.

Messina, A. M. (2007). *The logics and politics of post-WWII migration to Western Europe*. New York, NY: Cambridge University Press.

Migration Policy Group. (2011). *The Migrant Integration Policy Index (MIPEX)*. Available from http://mipex.eu/

Munck, G. L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative ideas. *Comparative Political Studies*, *35*, 5-34.

Permoser, J. M. (2012). Civic integration as symbolic politics: Insights from Austria. *European Journal of Migration and Law*, *14*, 173-198.

Peters, M. (2015). Open trade, closed borders: Immigration policy in the era of globalization. *World Politics*, *67*, 114-154.

Reeskens, T., & Wright, M. (2014). Host-country patriotism among European immigrants: A comparative study of its individual and societal roots. *Ethnic and Racial Studies*, *37*, 2493-2511.

Ruhs, M. (2013). *The price of rights: Regulating international labor migration*. Princeton, NJ: Princeton University Press.

Schain, M. (2008). *The politics of immigration in France, Britain, and the United States: A comparative study*. New York, NY: Palgrave Macmillan.

Schönwälder, K., & Triadafilopoulos, T. (2012). A bridge or barrier to incorporation? Germany's 1999 citizenship reform in critical perspective. *German Politics & Society*, *30*, 52-70.

Street, A. (2014). My child will be a citizen: Intergenerational motives for naturalization. *World Politics*, *66*, 264-292.

Tilly, C. (1996). *Citizenship, identity and social history*. New York, NY: Cambridge University Press.

Vink, M. P., Prokic-Breuer, T., & Dronkers, J. (2013). Immigrant naturalization in the context of institutional diversity: Policy matters, but to whom? *International Migration*, *51*(5), 1-20.

Waldrauch, H., & Hofinger, C. (1997). An index to measure the legal obstacles to the integration of migrants. *New Community*, *23*, 271-285.

Weldon, S. A. (2006). The institutional context of tolerance for ethnic minorities: A comparative, multilevel analysis of Western Europe. *American Journal of Political Science*, *50*, 331-349.

Wright, M. (2011). Policy regimes and normative conceptions of nationalism in mass public opinion. *Comparative Political Studies*, *44*, 598-624.

Wright, M., & Bloemraad, I. (2012). Is there a trade-off between multiculturalism and socio-political integration? Policy regimes and immigrant incorporation in comparative perspective. *Perspectives on Politics*, *10*, 77-95.

Yang, P. Q. (1994). Explaining immigrant naturalization. *International Migration Review*, 449-477.

Author Biography

Sara Wallace Goodman is associate professor of political science at the University of California, Irvine. Her research examines institutions of democratic inclusion, including citizenship, integration, and related member-making policies. She is the author of *Immigration and Membership Politics in Western Europe* (Cambridge University Press, 2014).