Article

Drivers of Human Migration: A Review of Scientific Evidence

Dino Pitoski * [ID], Thomas J. Lampoltshammer [ID] and Peter Parycek [ID]

Department for E-Governance and Administration, Danube University Krems, Dr.-Karl-Dorrek-Straße 30, 3500 Krems an der Donau, Austria; thomas.lampoltshammer@donau-uni.ac.at (T.J.L.); peter.parycek@donau-uni.ac.at (P.P.)
*   Correspondence: dino.pitoski@donau-uni.ac.at (D.P.)

Abstract: While migration research is at the peak of its productivity, a substantial gap persists between scientific evidence and policy action. As societal complexity increases, migration theory loses track on the numerous factors of human migration; the information on the most relevant factors affecting human migration (i.e., migration drivers), essential for policy decision-making, are hidden and dispersed across the ever-growing literature. Introducing a novel approach to conducting a literature review, emphasizing an unbiased selection of literature and the approach to analysing literature by coding, we collect evidence on the most pertinent migration factors. The study establishes a methodology for a quick but rigorous, collaborative gathering of evidence, as well as an initial inventory and an interactive map of nearly 200 factors working at different migration corridors.

Keywords: migration drivers; migration factors; review methodology; evidence collection; relevance ranking

1. Introduction

The ongoing events of mass migration towards Europe, also referred to as a "migration crisis", clearly demonstrated the necessity for policy interaction on a bilateral, regional, and global scale (Reslow 2019). Adding the temporal dimension into the mix, as well as the conflict of interest between the short-term and long-term benefits in an international, collective environment, the phenomenon "migration" can be described—based on the definition of Levin et al. (2012)—as a "super-wicked problem".

Along with the societies' rising complexity—to which the intricate phenomenon of migration is undoubtedly one of the most significant contributors—it has become increasingly difficult to understand both tangible and intangible infrastructures, and to predict the unintended consequences and side effects of actions. Within this complexity, Sanderson (2002) describes policies as "conjectures" that are only as good as the evidence they are based on.

Although Migration Studies produced a surplus of geography-specific evidence, a clear gap has been noticed between these signs and actual policies. According to Baldwin-Edwards et al. (2018), this gap can be attributed to three main factors. One is the ongoing "paradigm war" between positivists, interpretivists and critical approaches among migration researchers, in which the nature of what counts as evidence is itself contested; see, in relation, the discourses on "structure" versus "agency" by Bakewell (2010), and "migration optimism" and "migration pessimism" by de Haas (2010). Another factor is the way evidence gets interpreted, in terms of capacity and knowledge at both individual and institutional level policy, and in terms of the overall narrative that each policy advises (see also Boswell et al. 2011). The third factor is associated to the political environment in which the policymaking takes place. This setting, which is built upon theories, and manifests elaborate social and political realities, often lacks a neutral and objective stance due to policymaking becoming entangled in political processes (see also Sutcliffe and Court 2005).

To sharpen migration matters, policy has recently been putting a lot of focus to quantitative data, which is tempting, as it often seems easier to define and measure success

of actions in this way (see e.g., migration-related indicators as part of the Indicators for Sustainable Development discussed by Piper 2017). Yet, the diversity in data collection methods and measurement principles often leads to their incorrect use and interpretation (Singleton 2016; Wiśniowski 2017). At the same time, the actual indicators of performance in which the collected data become employed are very simplistic, and the comprehensive data analyses and reporting methods are disregarded (Pitoski et al. 2021). What is systematically missing from this quantitative perspective, is the information on the causal mechanisms underlying migration; mechanisms essential for explaining, as well as predicting migration (see Willekens 2018). At the same time, all attempts to elaborate on these mechanisms remain "articulated only verbally, and not in formal, mathematical terms" (see Bijak 2011, p. 42). This leaves much room for subjective assertions on specific factor-migration relationships.

Clear and accurate insights on migration factors must (have) come from systematically traced relationships, which normally constitute a migration theory. Yet, over the years, Migration Studies have produced more than twenty theories, which interchange, overlap and contrast hundreds of factors considered important for human migration (for a thorough review of theories that is still up to date, see, e.g., Bodvarsson and Van den Berg 2013). As the complexity of human life increased over an exponential scale (consult Scholten 2020), together with the awareness about the vastness of factors propelling human migration, it became more difficult, and perhaps less viable, to establish theories. The fact is that for several decades, migration theory has remained at an impasse (de Haas 2014).

This situation is inconsistent with the fact that Migration Studies is at the peak of its productivity. Pisarevskaya et al. (2020) show that the number of articles, as well as the number of journals focusing on migration, has tripled in the past three decades. Migration Studies "came of age" in terms of the diversity of topics present in research. The field experienced a "transition from geographies to mobilities, and from the governance of migration to the governance of migration-related diversity, race and racism, discrimination, and social– psychological issues", which further "indicates a shifting attention [...] from questions of 'who' and 'what' towards 'how' and 'why'" (Pisarevskaya et al. 2020, p. 24). However, works focused on these latter questions—particularly regarding the aspect of "why"—are difficult to identify due to the sheer amount of research produced, especially since migration theory stopped gathering answers into coherent volumes.

What is currently needed is a systematic means of extracting evidence from this enormous body of research. There is a need for establishing "when and why some [migration] drivers are more important than others, which combinations are more potent than others, and which are more susceptible to change through external intervention" (Van Hear et al. 2018, p. 1). In a way, there is a need to recollect the evidence on migration drivers[1]. This challenge is also acknowledged by the European Commission, which, recently, has made substantial investments in systematising migration research, with investigation into migration drivers as one of the essential research domains (European Commission 2018).

To our knowledge, thus far, there have been no attempts to establish a comprehensive list, or, better yet, a relevance ranking of migration drivers, which would be based on a systematic and objective collection of evidence from migration research. For policymakers and scientific scholars alike, such ranking holds significant value. For policymakers, this identification would allow for understanding the causes that have most of the effect to migration, calibrating the indicators they currently use for monitoring and forecasting migration, as well as executing targeted actions on the most relevant causes and seeing whether these produce the desired outcomes in the future. For scholars, the identification practically exposes gaps in research; namely, the under-researched migration drivers and

---

[1] In this work we disregard the semantic nuances as to when and whether it is appropriate to use the terms drivers, causes, reasons, determinants, or any other word symbolising explanatory factors of migration (see Carling and Collins 2018). Although we are aware of the sensitivity surrounding specific terms being translated into analytical concepts in migration theory, we use "most relevant factors" and "drivers" (of human migration) interchangeably.

geographies in which these operate. The new research on under-investigated drivers and driver geographies, then, provides an updated and ever more credible information used to formulate migration policies. Hence, our key goal in this paper is to develop that relevance ranking mechanism, which would be useful as an overview of relevant migration factors, but would also offer a transparent, rigorous method that facilitates sustained monitoring, and updating, of the provided relevance ranking over time. By making the collection of scientific evidence on drivers of human migration transparent and systematic, we aim to narrow the gap between evidence and policymaking.

2. Systematic Literature Reviews for Evidence Based Policymaking

The notion of what is regarded under the term "evidence" is critical to designing policies. Solesbury (2001), argumenting on the fundamental linkages between evidence and scholarly research, refers to the Oxford English Dictionary and its definition of evidence, namely as "the available body of facts or information indicating whether a belief or proposition is true or valid" (see also Pearsall and Hanks 1998).

The methods by which these facts or information are brought to light, however, are limitless, and continuously contested. The contestations come about both when these are mutually compared and individually assessed, and both in terms of their adequacy to collect and validate facts in general and in particular case. Perhaps the most typical "methods debate" in migration research is on the adequacy of qualitative versus quantitative methods, which corresponds to the general disciplinary fragmentation of migration researchers (see Castagnone 2011).

Limitless are also the knowledge channels and study materials that comprise all the facts and information, delivered by the diverse and contested methods, on a certain topic. Especially in Social Sciences, where Migration Studies are inherently situated, there is an overwhelming explosion of information, which information is retrievable through different bibliographic databases, as well as through different searching and relevance ranking algorithms.

Thus, any analyst, especially a politician that needs to incorporate multiple perspectives (i.e., evidence on a multitude of affairs) into decision-making, has a problem with reliability of evidence, overload of potential evidence-bearing material and evidence dispersion.

What comes about as the most appropriate means of collecting evidence, is systematic reviewing of literature. Systematic reviews are the key tool in developing the evidence base (Tranfield et al. 2003). In medicine, as the precursor of the evidence-based policy "movement", systematic reviews have been recognized as a remedy for the science-policy gap (consult Young et al. 2002). On the other hand, in the same field, the investigation on the usability of systematic reviews amongst the policymakers suggests what is needed is: (i) more pragmatism; (ii) increased combination of scientific evidence with governance principles; and (iii) further persuasion to translate complex evidence into simple stories (see Cairney and Oliver 2017). Meanwhile, science still struggles to offer rigorous methods to synthesize evidence of diverse types, generated by diverse methodologies (see Dixon-Woods et al. 2005).

In this work, we try to satisfy some of these governance needs, building a pragmatic method for the systematic review of literature that produces a straightforward overview of migration drivers with their current relevance. We do that while being aware of the offered guidelines for systematic reviewing (s.a. SWiM, see Campbell et al. 2020) which are implicitly incorporated in our work. We follow closely the guidelines outlined by Boaz et al. (2002), for systematic reviews in the domain of Social Sciences. We establish the quick but rigorous review method which (i) is protocol-guided; (ii) focuses on answering specific questions; (iii) identifies as much of the relevant research as possible; (iv) appraises the quality of research included in the review; (v) synthesizes the research findings in the reviewed studies; (vi) aims to be objective as possible about research to remove bias; and (vii) prompts updating to remain relevant. The method is, moreover, dedicated towards

eliminating the bias in terms of selection of evidence materials, as well as towards disabling the debates on the strength of methods used in individually reviewed works.

In what follows, we describe the methodology as it was being developed to pursue our concrete research aim, splitting the presentation into two conceptual parts: on the selection of review materials, and on the coding and analysis of the selected materials.

3. Evidence Materials on Migration Drivers—Methodology of Selection

The main concept pursued by this article are the direct[2] factor-migration relationships, evidenced by scientists, across various geographies. The underlying idea was that the "stockpiling" of evidences on these factor-migration relationships will produce a scaled distribution of relevance of migration factors, per specific country or country pairs, as well as aggregately when relaxing the geographical constraints. This expectation, on the factor-relevance scaling, is in accordance with the ubiquitous Pareto principle, by which a relatively small number of causes bears most of the effect in terms of the phenomenon observed.

To ensure that we got to an extensive amount of evidence-bearing materials, while making the amount of these materials manageable, we have set some boundaries before the literature selection process. One boundary was related to the research methodologies deployed across the works we eventually reviewed, that may have been considered as maximally reliable in delivering evidence. In that regard, we considered the evidence-bearing studies to be only those based on: (1) inferential statistics, where measured migration flows are matched against estimated migration factors; and (2) surveys—both based on interviews and questionnaires-where people declare their reasons for migrating. Other review papers and alike (such as works based on theoretical argumentations supported exclusively by citing other authors) we did not regard as studies directly evidencing factor-migration relationships, while studies based on the observation of human (migrant) behaviour were, in our case, considered not applicable.

Furthermore, we have set the boundary on the knowledge channels that might be considered as credible in terms of delivering these evidence-bearing studies to the public. Apart from the academic-publication outlets, considerable knowledge on human migration factors has been brought forward by both governmental and non-governmental organizations (NGOs), as well as by consultancy agencies. Although their literature does not necessarily undergo a classical peer review (unlike that of their academic counterparts), it is demonstrably grounded in observations made by these institutions' research staff, which observations are, presumably, directly employed as evidence in policymaking. Hence, we considered the policy/NGO/consultancy knowledge channels to be equally reliable in bringing forward the evidence, and incorporated this source of literature into the collection process.

Moreover, the decision on the reliable knowledge channels intimately relates to the decision on the library, or the set of libraries that should be browsed in the data-gathering process. This refers to the selection of appropriate online bibliographic databases that will ensure the inclusiveness of all above admitted channels, and point to the richest assortment of materials on the particular topic. In that regard, we opted for the three most well established bibliographic databases: Google Scholar (Scholar), which includes policy and consultancy literature; Scopus; and Web of Science (WoS). These three libraries have been confirmed to, jointly, guarantee the greatest coverage on just about any matter, after having been extensively analysed for differences in search results in terms of offered titles, the amount of results, and the citation information (Harzing and Alakangas 2015).

Within the limits of the three libraries, time restrictions were imposed on the selection as well, as to get to the manageable, while historically valid, set of review materials. The

---

2　Direct (factor-migration), as a complementary to indirect (factor-factor-migration) influences, have been the only concept that we could have feasibly reviewed, as the evidences on the indirect influences are not commonly delivered by any of the admitted scientific methods, knowledge channels or bibliographic databases (defined in the subsequent text).

restrictions were based on the following logic: if one wants to know which factors currently influence migration, one has to investigate the materials dating from as close as possible (or, rather, available) to the present day. This assumption is justifiable by the notion that science continuously builds upon historic knowledge, filling in the gaps spotted in prior research, while maintaining the facts that are still valid; thus, pertinent migration factors analysed in the past should reappear in the more recent research, and in this way the theory gets updated. We, therefore, decided to extract the evidence from migration-factors studies published from 2000 onwards, while within the oldest publications from this collection, we selected those that rely on data dated year 1990 onwards.

Upon setting the above boundaries, we commenced the literature refinement and review process. At this point, the readers are asked to consult the proces flowchart in Figure 1. The initial search phrase we chose for browsing through the selected libraries (steps S1–S3), was "determinants of migration" (texts containing that exact phrase anywhere in text or title). Another phrase deployed for the search was "drivers of migration", which we considered, in line with the observations made by Carling and Collins (2018), to be increasingly employed in the vocabulary of migration scholars. An examination of the quality of results from both of the queries showed that "determinants of migration" returned more evidence-bearing materials, as they are defined above, on the factors that affect migration[3].

On the list of initial search results (over 2000 titles), in order to get to the desired evidence-bearing materials, we performed three rounds of refinement. In the first round (step S4), studies that did not observe human migration (e.g., that focused on the migration of animal species, human cells, digital data, etc.) were eliminated. Studies that were immediately recognized as covering the topics other than factors of human migration were removed in this round as well. An example would be a study on the spread of diseases among various migrant groups. Furthermore, we took out a minor set of results for which the software returned unreadable characters, or studies written in languages other than English.

In the second round of refinement (steps S5–S6), for the rest of more than 1000 results connected to human migration, every publication was examined by reading the text thoroughly, to the point where we could clearly distinguish those works centred on the factors influencing human migration, and those only indirectly linked to the topic (including literature reviews by other authors). What remained was the share of works directly dealing with the matter of interest: evidence on the factor-migration influences.

The third round of refinement was performed during a later phase in the content analysis (unfolding in next section), where additional studies were removed, as follows:

- studies that concentrated on population-geographical segments too specific to be able to confidently claim wider, at least country-level, validity (e.g., migration incentives of Polish construction workers, Italian university graduates, farmworkers, physicians, and the like)
- studies analysing migration intentions, not the already pursued migration decisions, or studies analysing post-accession migration
- studies for which duplicates have been discovered, most often working papers that preceded a later-published journal article, by the same authors and on the same topic (where only the later publications were maintained).

Ultimately, 163 studies were selected for further analysis (step S7); 123 of them were published across 92 scientific journals, while 40 were issued as conference, discussion or

---

[3] Several queries were performed in addition: "reasons for migration", "migration factors", "causes migration", "influences migration", "why people migrate", where we have analysed the match of titles. From over thousands of results returned in each query, the co-occurrence of titles has shown to be extremely small; maximum of 141 repeating titles was found when comparing queries "migration factors" and "causes migration". Due to enormous amount of results, and since the results of the two queries we considered enough to satisfy the purpose of this study that focuses on the methodology for migration drivers evidence collection and ranking, the titles obtained from these additional queries have not been considered as part of this article. Results for the two queries covered in the article are available in the Supplementary Materials. The list of additional titles and an analysis of the co-occurrence of results is available from authors upon request.

working papers (e.g., by The World Bank, IZA, MPRA, and other institutes). The list of titles, from the original libraries' output to after each step of reduction, can be found in the Supplementary Materials.
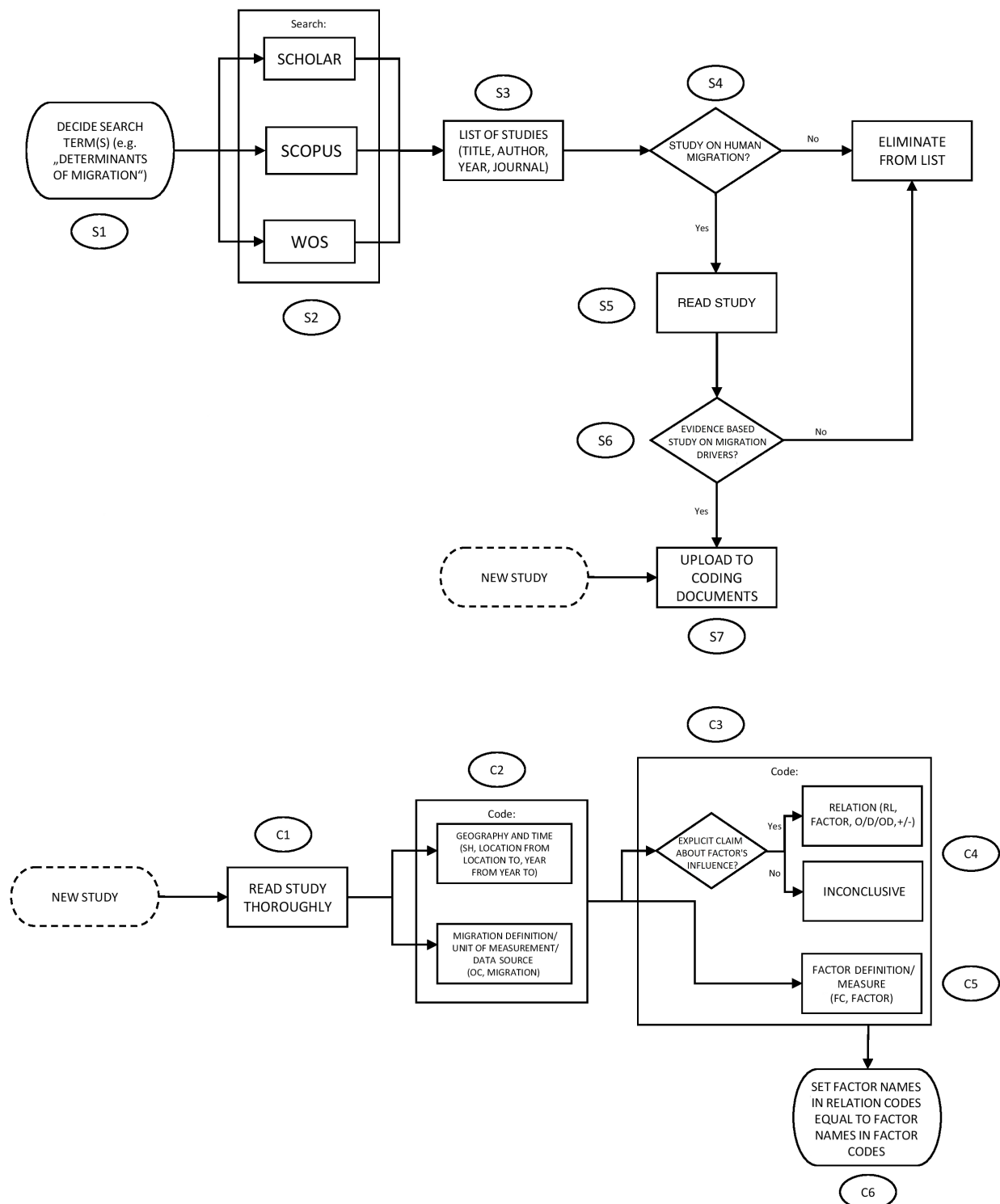

Figure 1. Literature selection (S) and coding (C) process.

4. Content Analysis of the Evidence Materials

We assessed the selected 163 studies using the method of content analysis (Krippendorff 2004; Mayring 2014), for which we deployed MAXQDA software (VERBI Software 2017). The coding scheme comprised the four main concepts: relations, factors, spheres, and outcomes.

Relation codes were used to mark parts of text where factor-migration influence was explicitly reported. These included the name of the factor influencing migration, location of its influence (origin, destination, or link), and the direction of influence (positive or negative). Factor codes designated the text segments in which the exact definitions for factors analysed for their influence on migration were provided, including the way in which these have actually been measured. Sphere codes pertained to "time-geographies": the specific countries and specific year-to-year periods where the above relations were being noted. With outcomes we were marking the text segments containing precise definitions of migration, and its means of measurement. In this regard, it is important to note that tracing the definition of migration and the way in which it is gauged represents one of the long-standing issues in migration research (see Willekens 2018); therefore, to note down these definitions means to open up an inventory of migration definitions that may, at some point in the future, serve scientists and politicians on the way to standardize the way migration is defined and measured. The full list of migration definitions, the ways in which it has been measured, and migration data sources, is retrievable from the Supplementary Materials.

In the sequel, we describe in detail the process and the distinct sequence of coding. At this point the reader is invited to consult again Figure 1. In the first step (C1), the complete text was read to get acquainted with its structure and derive insight into how the concepts were distributed across the study. Usually, factors, spheres and outcomes were found in the sections describing data and methodologies, while relations were traced in the results sections of the studies reviewed.

In the second step (C2), two relatively easily discernible concepts were coded: spheres, and outcomes. Spheres referred to the exact geographic and temporal segment in which the factor-migration relationships were being validated. Generally, a sphere code would read: SH, COUNTRY FROM-COUNTRY TO, YEAR FROM-YEAR TO. In cases where studies were focused on internal migration, the same country was marked down as both origin and destination. Simultaneously, we coded the outcomes, whereby we identified each portion of a text specifying: (i) the definition of migration; (ii) the exact way by which migration was measured; and (iii) the information on the dataset from which migration measurement followed in the reviewed study. The outcome, coded everywhere uniquely as OC, MIGRATION, included different directions of migration: immigration to a country/region, emigration from a country/region, the bilateral exchange between countries or the multilateral exchange across a region, as well as migration within a country (internal migration).

In the third step (C3), we traced the essential concepts demonstrating the factors' influence on the outcome: factor-migration relations. These were coded as RL, FACTOR, O/D/OD, +/−, whereby O/D/OD marked the factors' effect on migration at the origin (O), at the destination (D), or along a link (OD), while +/− pertained to whether the effect was found to be positive or negative (C4). Only statements where authors explicitly claimed that a factor impacts migration were coded as relations; ambiguous relations (e.g., would, could, might) were coded as INCONCLUSIVE (C4). Through such conservative approach, we ensured to have applied a high amount of rigour in the analysis, while reducing complexity in covering a very large, and rapidly growing, body of literature. With the review materials based on regression analysis, it was relatively straightforward to trace and code the clearly established relations, as these would normally follow numeric results shown in regression tables, which (estimates, standard errors and other statistics) could be compared with written statements following these results. Studies that did not involve inferential statistics were somewhat harder to evaluate; for instance, from a summary

chart showing relative proportions of each factor's relevance derived from aggregated survey responses, one could not have clearly defined a threshold for which factors are more important than others. In that sense, too, strictly explicit verbal statements were sought for in order to corroborate the exact relations. In addition, claims about a factor's effect that would have occurred in the same study but prior to administering a more sophisticated method of analysis (e.g., conclusions based on descriptive statistics preceding conclusions based on regression analyses), were omitted. Likewise, any claims of existence on an influential relation that later in the same piece would be contradicted (for example, through a robustness check in a regression analysis), were recoded as INCONCLUSIVE.

The last steps in the coding sequence (C5–C6) involved coding for the factors whose names were provisionally created when coding for relations, based on authors' wording in relation-coded text segments. These were looked up throughout the rest of the text, to find the precise definitions of these factors. All factor codes and all relation codes in the coding system were then updated to carry unique factor names, and reflect precisely the actually investigated factors' definitions. Assigning "proper" code names presented a challenging task due to subtle (and less subtle) variations in terminology used to define both one and the same factor. We overcame this problem wherever possible by tracing the exact way a factor was measured in the study, including the source data for this measurement. For example, although the authors may have discussed the influence of "income", if GDP per capita was the factor actually investigated for its influence on migration, the code names for both relations and factors were updated to GDP per capita as a sole factor name. Where turning to exact measurement principle would be inapplicable (e.g., the statement "I migrated for work" within a survey about migration decisions), we aimed for a pragmatic code name, keeping up to a standard (in this example, "EMPLOYMENT").

The completed coding system comprised over 2800 coded instances from which the following parameters (column headings) could have been immediately derived: (1) document (study) name; (2) code names; (3) coded segment of text; (4) period from; (5) period to; (6) geography from; and (7) geography to. On top of these parameters, for each study, we looked into the details on its journal issue and noted an additional quality parameter, the matching 2017 Journal Impact Factor (Clarivate Analytics 2018), if the journal was listed. As the impact factor values of the reviewed publications ranged from 0.5 to 9.5, for studies issued under journals unlisted, as well as for conference, discussion and working papers, we chose an arbitrary (lower) value of 0.1. These values were appended as additional columns (weights) next to each coded factor and factor-migration relation to compare weighted frequencies with the unweighted frequencies of their occurrences. The complete coding system, including journal information and journal-ranking-based weights, is available in the Supplementary Materials.

5. Results

The focus of this study is the design of the very methodology for the systematic extraction of evidence on human migration drivers. The designed methodology, essentially one main result of this study, is elaborated through the description of its application in the two previous sections (also supported by the process flowchart), so its replication is straightforward for any potential user. In this section, we concentrate further on our additional aim: the results coming from our application and our coded review sample.

The initial data feed, of close to 200 migration factors that we gathered using our method, is directly useful to commence an interactive visualization of human migration drivers—the Migration Drivers Map—as a comprehensive tool for monitoring migration factors working around the globe. This map we initiated at https://public.tableau.com/pro file/dinopitoski#!/vizhome/DriversofHumanMigrationReviewofScientificEvidenceDriver sMap/Story1. The map covers all geographies that have been examined by authors in the coded studies, with all factors that have been factually proven as influencing human migration across the examined works. By clicking on specific countries or the set of countries on the map, one is directly informed about the factor-migration relationships working

in and between these countries, where the origin-destination positioning is shown next to the list of factors and the specific study evidencing their influence). Also, the regions which constitute of more countries where there has been no clear distinction in the original studies, of which specific factors actually relate to which specific country in the region, are shown separately, each region in its own map interface. The Map is set out to be continuously updated, by adding new studies, new geographies, and new factors proven to be working in these geographies, after gathering and coding the scientific studies in the same consistent way.

From an overlapped image of all geographies occurring on the map, we can obtain a total "ranking of attention" that the particular world geographies received in terms of being studied for migration drivers. This information is contained in Figure 2, demonstrating the spatial coverage of this study. The map highlights are based on the number of reviewed studies that analysed migration drivers within a specific country or region. This information is particularly interesting concerning the new investigations towards under-represented geographical research areas.

In Figure 3, we present the overall relevance ranking of migration drivers based on the factor-migration codes derived from the 163 analysed publications. The chart contains the weighted and unweighted frequencies of evidenced spurring (+) or deterring (−) effects, of factors of migration working specifically at the origins (O), destinations (D), or the origin-destination links (OD). Each factor-migration relation is essentially a (weighted) count of studies that have proven its existence using any of the recognized observation-based scientific methods. The reader should note that the ranking in the figure is reduced to top 30 relations, which stem from an aggregate observation (all covered geographies), and that we do not claim their historic or geographic validity across countries. The ranking is envisaged as to be continually updated with old works selected for review by additional rounds of our literature-selection process, as well as with any freshly published work on migration drivers. However, for the purpose of this review, we judged the body of documents carefully sifted and reviewed is suitable to vouch for some "asymptotically valid" conclusions. In order to obtain the insight into the current relevance ranking for factors valid on any precise geography, the reader should consult the Supplementary Materials, or visit the above website for the interactive map of migration drivers.

A general way to interpret the relations charted in Figure 3 is as follows, starting from those at the top. The higher the education level in the country, the higher the outmigration from that country/region. The second most frequently proven relation can be interpreted as: the greater the physical distance between countries/regions, the lower the migration between these countries/regions. The third relation is the so-called network effect: the greater the community of origin country/region migrants at the destination, the greater the inflow of new migrants from the same country/region to that destination. We leave the reader to analogously interpret the rest of the factor relations, whose complete ranking is retrievable from the Supplementary Materials.

When ranking relations at an aggregated level, one might obtain, as we have, reciprocal factor-migration relations in the same ranking list. For instance, in the overall ranking in Figure 3, the unemployment rate, shown very frequently to be positively linked to outmigration from the origin, is also very often demonstrated to be negatively related to in-migration to the destination. Thus, the relevance of the unemployment rate as an influential migration factor, in a general observation, jumps up by several ranks relative to other factors. Similar upscaling effect may be established when complementing the factors of GDP per capita with income, as well as few other relations with factor name variations (e.g., the simultaneously negative effect of old-aged and positive effect of young-aged population).
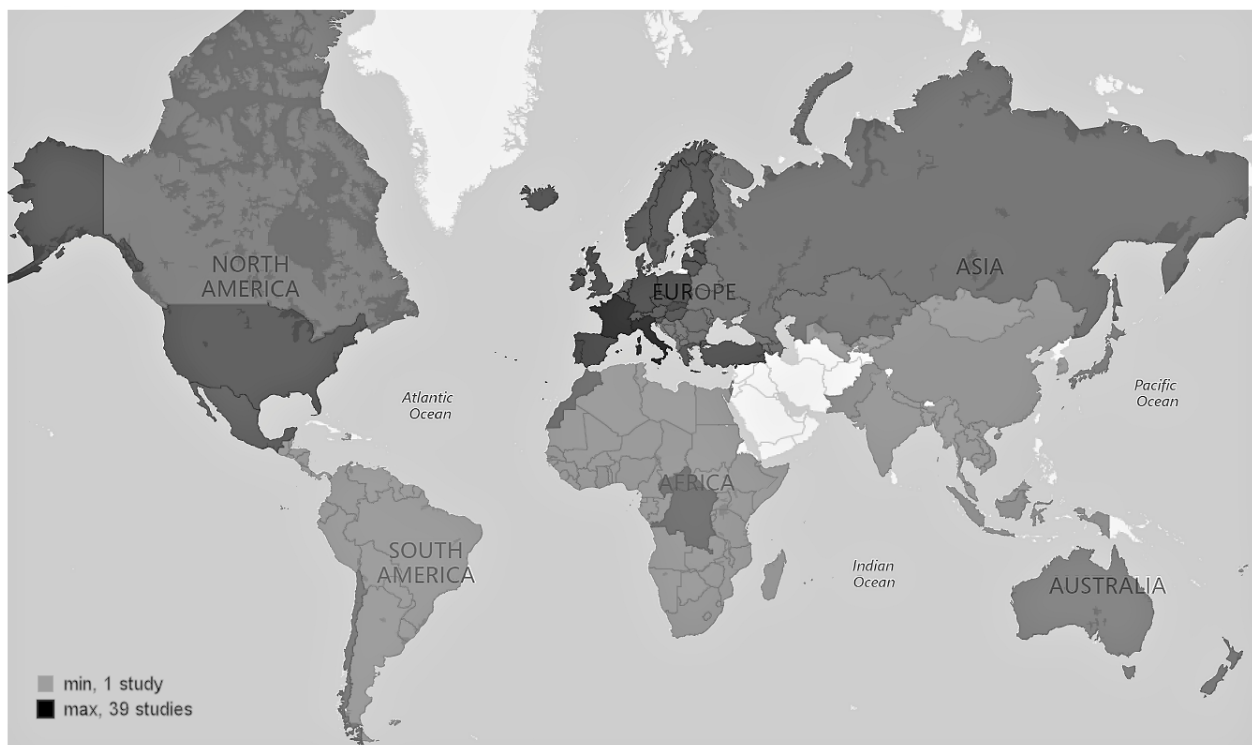
Figure 2. Geographical regions analysed for migration factors by frequency of dedicated studies.

Besides the reciprocal, ambivalent relations can be traced in our rankings as well, where some factors appear to manifest both deterring and spurring features. Examples for such ambivalent relations include relations containing the factor of unemployment rate, but also of age and marital status. The proof of some factors' effect to migration, moreover, becomes amplified when one jointly observes the factors that are inherently related, such as the population size and population density, or the income differential at a link and income's deterring effect at origins or attracting effect at destinations. Notwithstanding these ambivalent relations and "duplications", the representation of results in which both the direction and location of influence are incorporated into the ranking is preferred, as being consistent with exact denotations established in the materials reviewed. Moreover, this consistent representation enables the identification of like contradictions, which in turn can assist researchers in investigating why these occur, and aid eventual discussions on the causal mechanisms composed of multiple factors being at work.

A reader interested in a general overview of factors, relaxed from the information on direction and location in the relation codes, may turn to Figure 4, which displays an aggregated view based on the frequency of proven factors, and their suggested grouping. The frequency of each factor/group being evidenced in literature sample is proportional to the size of the field in the rings constituting the diagram. Provided in the figure are about fifty of 194 factors in total, which fifty have been proven as influential to migration the most number of times. The complete ranking list of factors is available in the Supplementary Materials.

The logic behind the grouping of factors in Figure 4, starting with the larger groups (income and trade, (un)employment), is as follows. Income, and income differential, can fairly be considered as a single factor, and the same applies for employment and employment differential. Furthermore, GDP per capita (often synonymized with income), per capita income, relative inequality (expressed by the Gini coefficient that is derived through income), as well as wages and wage differentials, could all be regarded as one factor of income. The same goes for the (un)employment rates, (un)employment differentials and employment opportunities, even the dependency ratio, which by definition incorporates the share of the working population; all these can essentially be classified under a single

category, of employment. The rest of the factors were grouped using the similar rationale, while the grouping can also be done in a more classical approach, with headings such as economic factors, demographic factors, cultural factors, political factors, etc.
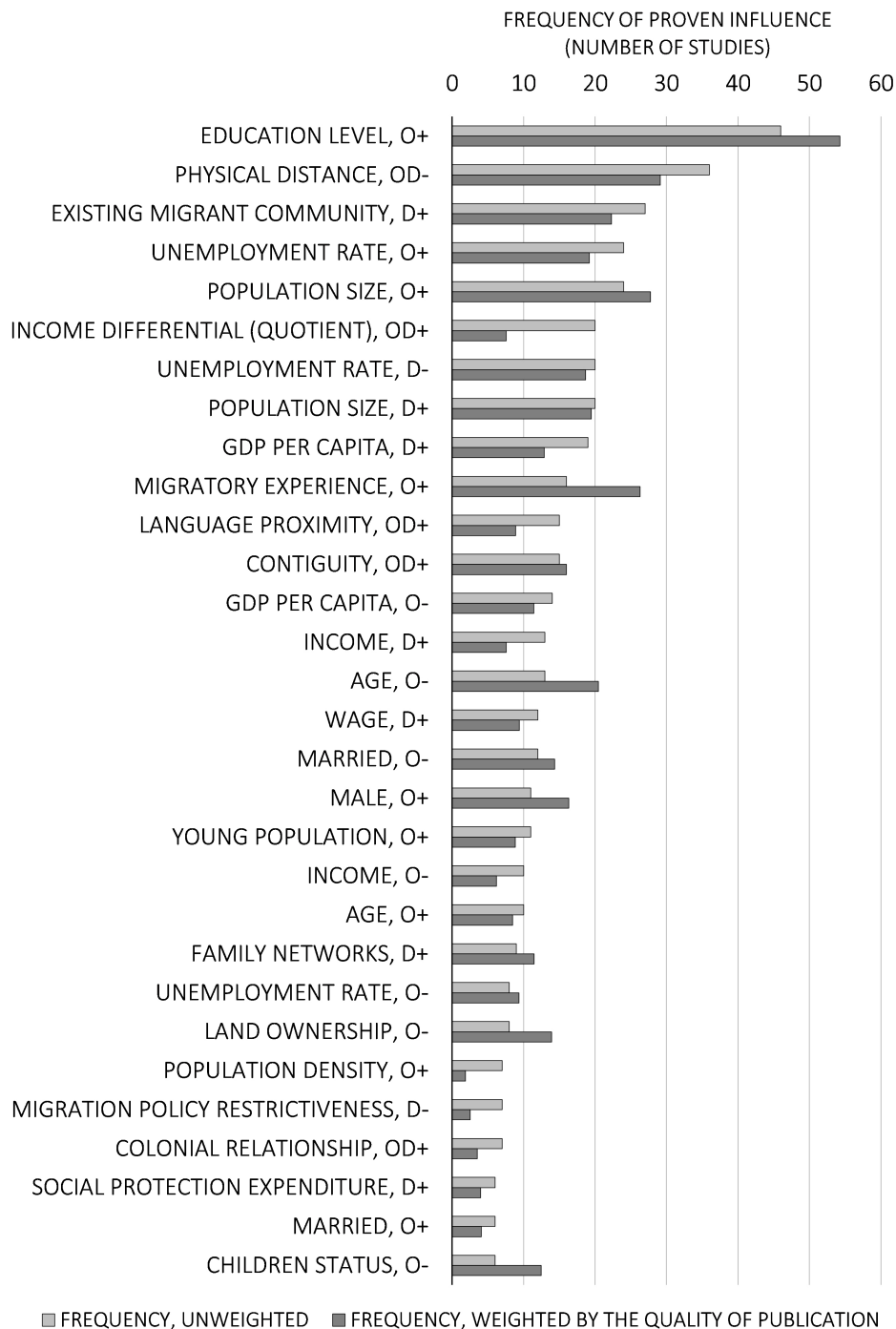

Figure 3. Top factor-migration relationships (aggregated for the studied sample).
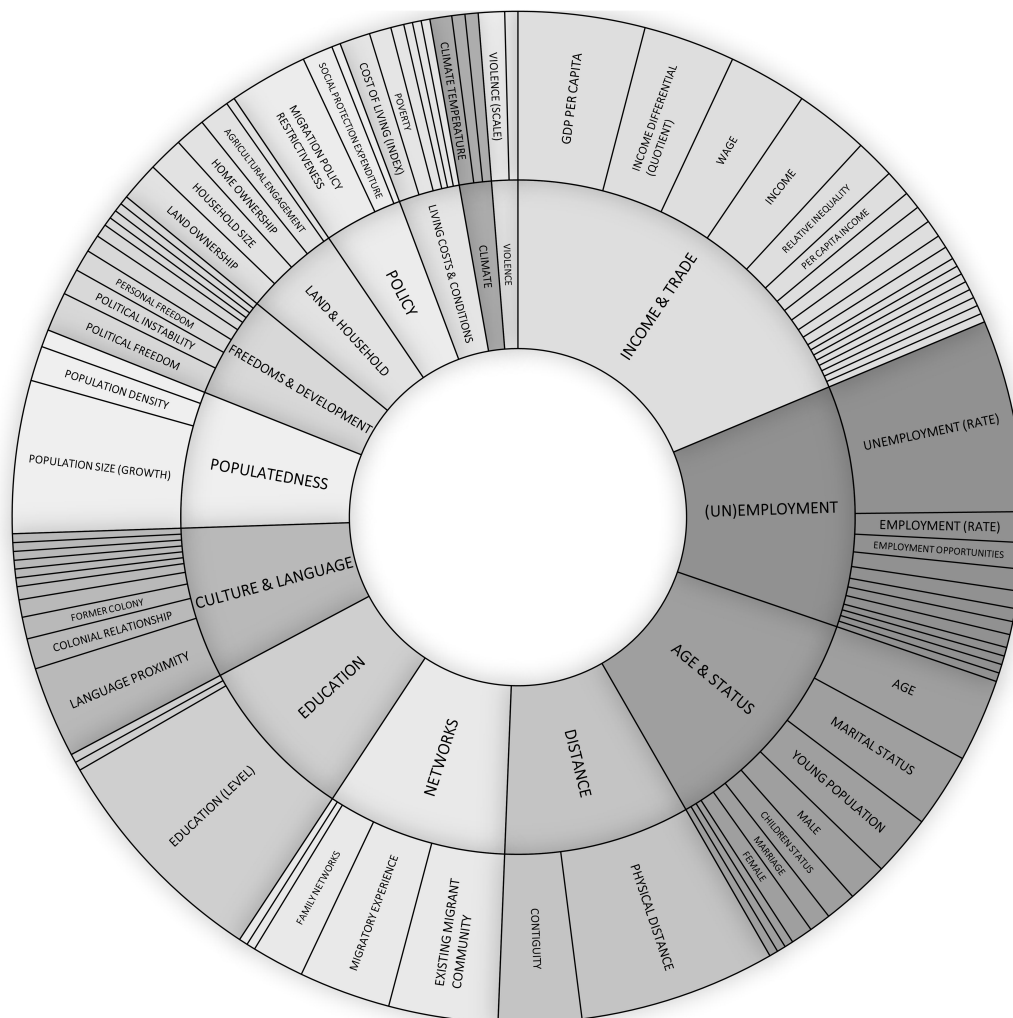
Figure 4. Top factors of migration, grouped (aggregated for the study sample).

The ranking of relations and factors obtained through frequency of proof in the initial literature review sample resonates well with some established facts in migration theory. That, e.g., more educated people are more prone to migrate out of their resident regions, or, that physical distance between areas impedes migration, is nothing new. Every relation from those at the top on the ranking list—the network effect, the population effect, migratory experience, etc.—is a well and long known migration-theoretical proposition (see the aforementioned theoretical overview by Bodvarsson and Van den Berg 2013). However, the relativity of the intensities, directions and locations of influence of these factors has, at least to our knowledge, never been established precisely, and presented comprehensively. Moreover, the top relations and factors that are predominantly discussed by both theorists and politicians, often divert all the attention from the very long list of under-researched, and even missing drivers. We refer to some examples of these under-investigated factors as we conclude this article.

6. Conclusions

While migration research is at the peak of its productivity, a substantial gap persists between scientific evidence and policy action. A highly politicized topic, migration has not been managed adequately by scholars in terms of systematising evidence, especially in regard to migration drivers as being the building blocks of both migration theorizing and policymaking. Systematic reviews on migration drivers practically do not exist, although systematic reviewing is the most reliable way to gather evidence from the vast body of

literature. At the same time, systematic reviews are often not straightforward enough to be used in policy practice.

In this study, we presented a systematic methodology that enables the ongoing observation of developments in migration research. Consequently, the scientific evidence is fully exposed and allows both researchers and practitioners to understand how the data was gathered, which conclusions were made, as well as if and how far these have been reflected in the arguments on certain policy options.

We focused on research that contains evidence of the observed, direct effects of migration factors. An unbiased sample, covering as many as possible geographies, was methodically extracted from thousands of studies, and systematically reviewed as to establish some asymptotically valid evaluations of these effects. The methodology for the unbiased selection and coding of evidence-bearing publications is prescribed in detail, which should alleviate future data feed. This is expected to enable continuous updates to the factors' influence rankings, especially by incorporating evidence from new studies that cover the under-researched factors and geographies, recursively identified in the ranking list. We have also initiated an interactive tool—Migration Drivers Map—for the representation of the traced factor-migration relations working at different geographies, a tool that is useful for policymakers to maintain a comprehensive overview of these relations, and for scientists to recognize which factors and geographies around the globe are being under-investigated. Besides establishing the overall methodology, the interactive tool and migration factor rankings based on the review sample, we collected the various definitions, measurement principles and data sources for these factors, and for human migration. All factor-migration relations are noted per specific country and time period occurring and are sortable per any desired geographic and time span, as well as aggregable for a universal look into the "drivers of human migration".

6.1. Limitations

The presented research has some limitations that readers should be aware of when interpreting and applying the results. First, the selection of relevant literature and materials for review is by nature a difficult process, especially in terms of selecting key phrases for the actual search. In addition, the ranking and search algorithms of the bibliographic databases that were used are usually black boxes; thus, researchers have to have trust the validity and comprehensibility of the returned outcomes. To reduce further bias, we chose to code only the "absolutist" claims on factor-migration relationships, thereby striving for the highest level of rigor and fairness possible regarding the subsequently derived ranks. However, more refined results on the strength of each factor-migration relationships could have been achieved by involving fuzzy logic, where strength of verbal expressions would also be ranked. Alternatively, a meta-analysis could have been performed to identify average effects, but this would limit the review only to works based on statistical inference. These possibilities have been considered, but the diversity of data and measurement principles for different factors and migration, the diversity of geographies involved, and the diversity of models alone did not suggest a reachable outcome.

Furthermore, as the relevance ranking is determined by the built-up (weighted) frequency of the proven effect for particular factors, the crucial point comes in the form of how factors themselves came about in the studies reviewed. This selection is often made following previous works and factors of proven importance, which are then combined with newly introduced factors of interest in the next step. Upon re-identifying significance, this increases the overall frequency count, yet does not guarantee a complete picture in terms of all existing migration factors.

Our method is entailed also by spatio-temporal limitations. For example, regarding population (geographic) validity, it is impossible to prove that the identified drivers are ultimately applicable for all segments of population. Likewise, from a historical angle, the significance of drivers during the time of a particular study is evident (or also regarding the timestamp of the used dataset), yet it is unknown to which degree the importance and

validity have degraded—or even vanished—over time. This fact once more emphasizes the notability of the presented methodology, which facilitates continuous monitoring and consequently a rigorous, transparent evaluation of the evolution of migration factors.

6.2. Future Research

As the methodology of this study was devised by analysing specifically migration literature, we see direct benefits in replicating the methodology for other sub-domains of migration research. One example where this is straightforwardly replicable is the sub-domain of migration infrastructures; replacing the question "why people migrate" with the question "how people migrate", or by which means they migrate, and across which time-geographies.

Based on the results of the identified migration factors, at several points we made general observations about the under-investigated factors. One specific example is the factor of physical distance and the specifics of its measurement. Given the frequency of proof of strong effect of physical distance-whose simplistic approximation by means of the great circle distance between country capitals, has, at numerous instances, been put into question—we suggest complementing the investigations on the effects of distance with factors that indicate transportation connectivity between countries or regions. For that aim, vast amounts of data are available, in terms of free schedules of shipping lines, airlines and online road maps, from which it would be possible to not only infer real distances (thus, more precisely estimate transportation costs), but also infer transportation opportunities. Another overlooked factor that becomes ever more relevant is the information connectivity, with broadband access and broadband usage being its major components. For these components, data are also readily available online, e.g., from the World Bank Open Data. This list of under-investigated aspects continues (e.g., consider the climate change "group" of factors such as droughts, floods, extreme temperatures, etc.), which holds immense potential for future research.

In that sense, instead of serving as the finished inventory of relevant facts, we suggest the established methodology to serve as a basis for policy analysts and fellow researchers, to jointly work on a growing and refining a migration observatory. This observatory would ideally be grounded in fresh studies on under-investigated factors in under-investigated geographies, and be established as an online platform like the one we initiated, with geographical mapping of factor-migration relations.

References

Bakewell, Oliver. 2010. Some reflections on structure and agency in migration theory. *Journal of Ethnic and Migration Studies* 36: 1689–708. [CrossRef]

Baldwin-Edwards, Martin, Brad K. Blitz, and Heaven Crawley. 2018. The politics of evidence-based policy in Europe's "migration crisis." *Journal of Ethnic and Migration Studies* 45: 1–17. [CrossRef]

Bijak, Jakub. 2011. *Forecasting International Migration in Europe: A Bayesian View.* London: Science+Business Media.

Boaz, Annette, Deborah Ashby, and Ken Young. 2002. Systematic reviews: what have they got to offer evidence based policy and practice? ESRC working paper 2 *ESRC UK Centre for Evidence Based Policy and Practice London*. Available online: https://www.kcl.ac.uk/sspp/departments/politicaleconomy/research/cep/pubs/papers/assets/wp2.pdf (accessed on 17 November 2020).

Bodvarsson, Örn B., and Hendrik Van den Berg. 2013. *The Economics of Immigration: Theory and Policy*, 2nd ed. New York: Springer Science+Business Media.

Boswell, Christina, Andrew Geddes, and Peter Scholten. 2011. The Role of Narratives in Migration Policy-Making: A Research Framework. *The British Journal of Politics and International Relations* 13: 1–11. [CrossRef]

Cairney, Paul, and Kathryn Oliver. 2017. Evidence-based policymaking is not like evidence-based medicine, so how far should you go to bridge the divide between evidence and policy? *Health Research Policy and Systems* 15. [CrossRef]

Campbell, Mhairi, Joanne E. McKenzie, Amanda Sowden, Srinivasa Vittal Katikireddi, Sue E. Brennan, Simon Ellis, Jamie Hartmann-Boyce, Rebecca Ryan, Sasha Shepperd, James Thomas, and et al. 2020. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* l6890. [CrossRef]

Carling, Jørgen, and Francis Collins. 2018. Aspiration, desire and drivers of migration. *Journal of Ethnic and Migration Studies* 44: 909–26. [CrossRef]

Castagnone, Eleonora. 2011. Bridging the gap between qualitative and quantitative research methods. The case of biographical approach. ESF Strategic Workshop: New approaches for researching the determinants of migration processes. *International Migration Institute Oxford*. Available online: https://www.fieri.it/wp-content/uploads/2013/04/ESF-oxford-_note-Castagnone_145732-12102012_ita.pdf (accessed on 17 November 2020).

Clarivate Analytics. 2018. 2017 Journal Impact Factor, Journal Citation Reports. *Clarivate Analytics*. Available online: https://clarivate.com/webofsciencegroup/solutions/journal-citation-reports/ (accessed on 17 November 2020).

de Haas, Hein. 2010. Migration and development: A theoretical perspective. *International Migration Review* 44: 227–64. [CrossRef]

de Haas, Hein. 2014. Migration theory: Quo vadis? IMI Working Paper No. 100 / DEMIG Project Paper No. 24. *International Migration Institute Oxford*. Available online: https://www.migrationinstitute.org/publications/wp-100-14 (accessed on 17 November 2020).

Dixon-Woods, Mary, Shona Agarwal, David Jones, Bridget Young, and Alex Sutton. 2005. Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research and Policy* 10: 45–53. [CrossRef]

European Commission. 2018. CROSS-MIGRATION: Current European and cross-national comparative research and research actions on migration. *European Commission*. Available online: https://cordis.europa.eu/project/rcn/212877_en.html (accessed on 17 November 2020).

Harzing, Anne-Wil, and Satu Alakangas. 2015. Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics* 106: 787–804. [CrossRef]

Krippendorff, Klaus. 2004. *Content Analysis: An Introduction to its Methodology*, 2nd ed. Thousand Oaks: Sage.

Levin, Kelly, Benjamin Cashore, Steven Bernstein, and Graeme Auld. 2012. Overcoming the tragedy of super wicked problems: Constraining our future selves to ameliorate global climate change. *Policy Sciences* 45: 123–52. [CrossRef]

Mayring, Philipp. 2014. Qualitative content analysis—Theoretical foundation, basic procedures and software solution. *Social Science Open Access Repository Klagenfurt*. Available online: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173 (accessed on 17 November 2020).

Pearsall, Judy, and Patrick Hanks. 1998. *The new Oxford dictionary of English*. Oxford: Clarendon Press Oxford, vol. 16.

Piper, Nicola. 2017. Migration and the SDGs. *Global Social Policy: An Interdisciplinary Journal of Public Policy and Social Development* 17: 231–38. [CrossRef]

Pisarevskaya, Asya, Nathan Levy, Peter Scholten, and Joost Jansen. 2020. Mapping migration studies: An empirical analysis of the coming of age of a research field. *Migration Studies* 8: 455–81. [CrossRef]

Pitoski, Dino, Thomas J. Lampoltshammer, and Peter Parycek. 2021. Human migration as a complex network: appropriate abstraction, and the feasibility of Network Science tools. In *Data Science—Analytics and Applications.* Edited by Peter Haber, Thomas J. Lampoltshammer, Manfred Mayr and Kathrin Plankensteiner. Wiesbaden: Springer Vieweg, pp. 113–120.

Reslow, Natasja. 2019. EU External migration policy: Taking stock and looking forward. *Global Affairs* 5: 273–78. [CrossRef]

Sanderson, Ian. 2002. Evaluation, policy learning and evidence-based policy making. *Public Administration* 80: 1–22. [CrossRef]

Scholten, Peter. 2020. Mainstreaming versus alienation: conceptualising the role of complexity in migration and diversity policymaking. *Journal of Ethnic and Migration Studies* 46: 108–26. [CrossRef]

Singleton, Ann. 2016. Migration and asylum data for policy-making in the European Union-The problem with numbers. *Liberty and Security in Europe* 89. Available online: https://ssrn.com/abstract=2749508 (accessed on 17 November 2020).

Solesbury, William. 2001. Evidence based policy: Whence it came and where it's going. *ESRC UK Centre for Evidence Based Policy and Practice London*. Available online: https://www.kcl.ac.uk/sspp/departments/politicaleconomy/research/cep/pubs/papers (accessed on 17 November 2020).

Sutcliffe, Sophie, and Julius Court. 2005. Evidence-based policymaking: What is it? How does it work? What relevance for developing countries? *Overseas Development Institute London*. Available online: https://www.odi.org/publications/2804-evidence-based-policymaking-what-it-how-does-it-work-what-relevance-developing-countries (accessed on 17 November 2020).

Van Hear, Nicholas, Oliver Bakewell, and Katy Long. 2018. Push-pull plus: Reconsidering the drivers of migration. *Journal of Ethnic and Migration Studies* 44: 927–44. [CrossRef]

Tranfield, David, David Denyer, and Palminder Smart. 2003. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management* 14: 207–22. [CrossRef]

VERBI Software. 2017. MAXQDA 2018 [computer program]. *VERBI Software Berlin*. Available online: https://www.maxqda.com (accessed on 17 November 2020).

Willekens, Frans. 2018. Towards causal forecasting of international migration. *Vienna Yearbook of Population Research* 16: 199–218. Available online: http://austriaca.at:8080/?arp=0x003a1295 (accessed on 17 November 2020). [CrossRef]

Wiśniowski, Arkadiusz. 2017. Combining Labour Force Survey data to estimate migration flows: The case of migration from Poland to the UK. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 180: 185–202. [CrossRef]

Young, Ken, Deborah Ashby, Annette Boaz, and Lesley Grayson. 2002. Social Science and the Evidence-based Policy Movement. *Social Policy and Society* 1: 215–24. [CrossRef]