International Public Administrations and the Perception of Crisis: 
The Case of UNRWA and Palestine Refugees

Ronny Patz*, Svanhildur Thorvaldsdottir** and Klaus H. Goetz***
Department of Political Science, LMU Munich

Paper presented at the Panel
“Regional and Global Crisis Management“ (P317)
ECPR General Conference, Oslo, 6-9 September 2017

The draft and the data presented are work in progress – Comments are welcome
- Please contact the authors before quoting (via ronny.patz@gsi.lmu.de) -

Acknowledgments: This paper presents insights and empirical data gathered in the context of the project “Resource Mobilization in International Public Administration: Strategies for the Financing of International Public Policy”, funded by the German Research Foundation (DFG) as part of the Research Unit “International Public Administration” (http://ipa-research.com). The project follows up on our DFG-funded research on budgeting procedures and budgeting administration in international organizations. A key finding – discussed extensively in the Special Issue “Resourcing International Organizations” in Global Policy—was the increasing importance of the resource mobilization function in international bureaucracies. In the current project, we therefore study this function more closely, with particular focus on the domain of Global Refugee Policy.

We thank Salma Nosseir and Maximilian Haag for their research assistance.

* Ronny Patz: ronny.patz@gsi.lmu.de
** Svanhildur Thorvaldsdottir: svanhildur.thorvaldsdottir@gsi.lmu.de
*** Klaus H. Goetz: Goetz.LMU@gmail.com
Abstract

In this paper, we analyze the perception of policy crisis and budgetary pressure by the bureaucracy of the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA), the UN agency responsible for supporting several million Palestine refugees. Conducting a computer-assisted content analysis of UNRWA annual reports from 1951 to 2016, we show that it is possible to measure variation in the degree of policy crisis perception over the full history of the organization on the basis of administrative documents. We also discuss the challenges of constructing a dictionary that allows for a valid measurement of budgetary pressures over time. Our empirical findings indicate that the degree of perceived policy crisis and budget pressure of UNRWA vary strongly over time, against the common notion that the agency is an organization in permanent crisis. Against theoretical expectations, policy crisis and budgetary pressures do not evolve in parallel throughout UNRWA’s past 66 years, suggesting that in certain periods resource mobilization as one means of crisis management could be more successful than in other periods.
1. Introduction

With the outbreak of the Syrian war in 2011 and the resulting so-called “refugee crisis”, those international organizations that are responsible for catering the needs of millions of refugees have received renewed global attention. Whereas most attention has been on UNHCR, on IOM, on the role of the European Union (see Attina (2017) and Rossi (2017), in this panel) or on those UN agencies, such as the World Food Programme (WFP), that are supporting the collective humanitarian efforts, one key UN refugee agency in the region is often overlooked: the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA).

Building on the initial relief work of the United Nations Relief for Palestine Refugees (UNRPR) from 1948 to 1950 (Richardson 1950), UNRWA became the agency with the main responsibility for Palestine refugees (UN 1949). This large population of refugees is, and has been from the beginning, outside the realm of UNHCR’s mandate (for a discussion on the origins and merits of this distinction, see for example Kagan (2009) and for resulting difficulties see Goddard 2009). In a region where (geo)political crises are frequent and consequential for humanitarian and social work related to refugees (Rosenfeld 2009), Palestine refugees have been affected by the ongoing regional conflict as much as populations of refugees that fall under the responsibility of UNHCR. This makes UNRWA a pertinent case to study for crisis perception and crisis management. This is the more pertinent as the present situation of UNRWA has also been termed an “unprecedented” financial crisis by the organization itself (UNRWA 2015), underlining that crisis management can imply both policy crisis management as well as budgetary or financial crisis management. Looking into the history of UNRWA, this double crisis situation is not at all unprecedented since its creation in 1950. Repeated regional crises have required repeated efforts, for example to ensure additional resource mobilization when the financial needs of the organizations were higher than the resources available (Al-Omari 2015).

The key question for this paper is how to measure whether an international public administration (IPA) such as UNRWA is facing budgetary stress as a result of increased resource needs when a policy crisis, such as a regional or geopolitical conflict, affects its population of concern. We argue that, both, policy crisis and budgetary stress are not objective. Instead, they are first and foremost a matter of perception of the international bureaucrats faced with situations that they consider crises that result in budgetary stress. This perception may be shared by or it may be brought into the international bureaucracy by actors.
in their environment, but the main methodological challenge is how to measure these perceptions over extended time periods and across organization.

We present a computer-assisted (i.e. dictionary-based) content analysis of all UNRWA annual reports from 1951 to 2016 – a total of 66 reports – supplemented with a qualitative analysis of additional documents as a solution to the first part of this methodological challenge, i.e. measurement over extended time periods. The content analysis relies on a pre-existing general sentiment dictionary in combination with a custom-built dictionary for budget- and resource-related matters in international organizations constructed to measure crisis and budget pressure over time. We validate these measurements with a third dictionary and provide other materials to discuss in how far the dictionaries allow for the identification of those years that can be considered years with significant policy crisis (such as conflicts in the region affecting Palestine refugees) as well as of those years in which UNRWA perceived increased budgetary pressures or life-threatening financial situations.

In the next section (Section 2), we will introduce UNRWA as a very specific international public administration before developing a theoretical framework in Section 3 to explain the link between policy crisis and budgetary pressures, and why IPAs would signal both through annual reports. In Section 4, we discuss the methodological options and challenges of content analysis, before applying these methods to UNRWA’s annual reports in Section 5. Section 6 concludes the paper.

2. UNRWA: an international bureaucracy with state-like functions but voluntary funding

UNRWA is a remarkable organization in the population of international public administrations, i.e. in the population of bureaucracies that are at the core of many international organizations (for a review of literature see Liese and Weinlich 2006; Busch 2014; Eckhard and Ege 2016). Since its creation in 1950, the agency has been providing not just crisis relief but permanent state-like services – including housing, schools, and hospitals – to millions of Palestine refugees in Gaza, the West Bank, Jordan, Lebanon, and Syria (Rosenfeld 2009). This makes UNRWA, first, an international organization with 193 UN member states1 as its key principal(s), with the delicacy of acting in a highly salient and difficult geopolitical environment. Second, it is also a sizeable “street-level bureaucracy” (see

---

1 The intergovernmental Advisory Commission of UNWRA comprises only 27 member states and 3 observers.
Smith 2012 on this type of public administration) with millions of “clients” in various degrees of precarious situations. The most visible sign of UNRWA’s “uniqueness” is the fact that around 80% of its staff are teachers (Rosenfeld 2010, 288-9).

This special situation is even more remarkable looking at the resource situation of the agency: It is the number ten UN entity with regards to annual expenditures (2015 figures, http://www.unsceb.org/content/FS-F00-03), which are higher than those of all UN Specialized Agencies except for WHO. During the 2016-17 biennium, it only had baseline appropriations from the UN Regular Budget of about US-$ 55 million, financing mainly the ~150 regular budget posts. However, the largest part of UNRWA’s budget is based on voluntary (“extrabudgetary”) contributions, reaching a budgeted US-$ 2.4 billion in 2016-17. These funds are paying for “UNRWA local staffing numbering 30,905” (UN 2017). The central role that voluntary contributions play in UNRWA since its creation have become a general trend in international organizations, in particular in the UN system (Goetz and Patz 2017; Graham 2017; Michaelowa 2017), but the permanency of the tasks of UNRWA outside crisis relief make its situation special. The dominance of voluntary funding means none of the permanent work of the agency is actually permanently assured (Brynen 2003). Thus, for its small share of the UN regular budget, the agency has to deal with the budgetary procedures of a 193-member state strong international organization while having to fundraise a massive volume of voluntary resources, just like a privately run humanitarian organization, but with additional (geo)political constraints and without the right to raise taxes or raise funding on global capital markets. In UNRWA’s own words (UNRWA 1977, §13):

“UNRWA has been assigned quasi-governmental responsibilities by the international community, but it does not have the authority Governments have to acquire resources to meet these responsibilities.”

Whereas in the field of refugee studies, UNRWA has received some detailed attention in the course of its history (see for example the special issue in Refugee Survey Quarterly N°28/2-3 or the recent volume edited by Hanafi (2014)), scholars of IOs and IPAs have not paid any (special) attention to this organization since the 1970s (Buehrig 1971; Forsythe 1971a, 1971b; Perlmutter 1971). In recent comparative research where it was part of a wider population of IOs studied, it was found that despite its high amount of voluntary contributions, just a small share comes from private sources (Graham 2017) and that most of its funding is accounted for as unspecified voluntary contributions (Bergmann and Fuchs 2017).
The budget process in UNRWA has been described as “unorthodox” (Buehrig 1971, 167), even though some changes have happened since the 1970s (Barney 2003). In its earliest years, until 1955, the budget procedure was driven by the Commissioner-General (or Director) together with the – then very small – intergovernmental Advisory Commission of UNRWA, who proposed a budget but gave only very little detail to the General Assembly (Buehrig 1971, 175). From 1955, the Commissioner-General took the lead on proposing a budget though his annual reports, which were not even discussed in UN’s Advisory Committee on Administrative and Budgetary Questions (ACABQ) and the General Assembly’s Fifth Committee (Buehrig 1971, 176), the key bodies responsible for UN budgeting (see Patz and Goetz 2017b). Over time, this has changed and in recent decades UNRWA’s budget proposal is an integral part of the UN’s regular budget process (see for example the respective budget proposal for UNRWA’s in 2018-19 budget in UN (2017b)). And even though only about 150 of UNRWA’s (international) posts are financed from the UN’s regular budget, the large part of the budget financed from voluntary resources is also approved, albeit “implicitly” (Bartholomeusz 2010, 455). A detailed description of the current budget procedure (see UNRWA 2012) underlines the clear separation between the “internal” administrative process of drafting and presenting the budget at the UN’s New York headquarters, where only the small amount of international posts is formally considered.

The functioning of the agency based on mainly voluntary contributions has been a source of problems throughout its history (Brynen 2003). In his 1981 annual report (UNRWA 1981, §26), the Commissioner-General highlighted, again, that UNRWA was in a permanent state of budgetary stress and recalled that there are different ways of financing the operations of the organization:

“during the thirty-fifth session of the General Assembly several delegations suggested that the General Assembly should abandon the voluntary system for financing UNRWA. Suggestions were made that UNRWA's expenditure should be put on the regular budget of the United Nations or that a separate budget should be established and Member States pay according to an agreed assessment […]. It is unthinkable to continue as at present, with the Agency lurching from crisis to crisis.”

This situation in the 1980s was by far not exceptional, and so when the agency faced another of its repeated funding shortfalls in the second half of the 1990s, this was not “a new episode in the history of UNRWA but rather a recurrent one”, meaning that “the Agency has had to develop communication skills to market its mission and to secure funding” (Bocco 2010,
233). Seeing the financial difficulties UNRWA faced in the wake of the Syrian war, a former UNRWA Comptroller noted that this still applied today (Al-Omari 2015)

3. Theoretical expectations: policy crises, budgetary stress and resource management as crisis management

The overall situation of UNRWA described in the previous section may be specific to the agency, but the different aspects of crisis are not. However, the notion of “crisis” in international organizations and the implications of international bureaucracies in crisis management are not well-researched yet:

“The role IOs may play in defining and tabling a crisis is … overlooked. This is remarkable, since the agenda-setting capabilities of IO secretariats […] gave long been recognized.” (Olsson and Verbeek 2013, 326)

One exception, including in recent contributions, are crisis-like situations when international organizations face repeated financial challenges (Patz and Goetz 2017a, 125-7). This can include situations of ‘objective’ organizational crisis, in which the survival of agencies such as UNESCO is questioned (Hüfner 2017; Eckhard, Patz and Schmidt 2016). There are multiple factors that can induce these situations of extreme budgetary pressure in an IO, and a recent review of literature on budgetary stress in IOs has highlighted at least three types of budgetary stress: (1) permanent or structural budget stress due to underfunding; (2) sudden budgetary stress due to loss of revenue; and (3) sudden budgetary stress due to an unexpected excess demand for action (Patz and Goetz 2017a, 126). It has also been shown that financial crises in IOs can be more extreme (and thus more consequential) when multiple forms of budgetary stress come on top of each other (see Eckhard, Patz and Schmidt 2016 on UNESCO financial crises in the 1980s and after 2011). Thus, we expect that structural or permanent budgetary shortfall, when combined with a sudden income loss or sudden expenditure needs in situations where income cannot be generated rapidly, may result in situations that organizations (or their environment) perceive as extreme financial stress. Whereas others have focused on sudden stress caused by extreme income losses in IOs, i.e. more narrowly defined as “financial crisis” inside IOs, for example as a result of the withdrawal of financing from the US (Hüfner 2017; Eckhard, Patz and Schmidt 2016; Beigbeder 1979), this paper focuses on financial stress induced by sudden excess demand in the wake of policy crisis as perceived by an international secretariat.
Whereas there are multiple qualitative or quantitative measures of *policy crises* that can reasonably be expected to cause direct excess demands on international organizations and their bureaucracies, in many cases the question of what constitutes a crisis is a matter of perception and, where this perception is made public, a matter of communication. Whether a policy challenge is a crisis or a matter of routine may differ from one organization to the next, or from one (historical) context to another (Olsson and Verbeek 2013, 325-6). We consider that operational international organizations and their bureaucracies dealing with conflict-response and humanitarian aid, peacekeeping and peacebuilding, or other forms of service-provision to populations in need are most at risk for sudden excess demands as a result of regional and/or geopolitical crises (intra- or inter-state conflicts, political instability, natural catastrophes etc.). For organizations with core missions in refugee or humanitarian policy, a refugee or humanitarian “crisis” can be considered a *policy crisis* affecting those populations that they already serve or that fall under their mandates, comparable to how a pandemic event is a (health) policy crisis for an organization like the World Health Organization (WHO).

Measuring budgetary stress caused by excess expenditure demands as well as the resulting additional income needs in the wake of perceived policy crises is also difficult in objective terms: Whereas income losses should be clearly represented in a drop in budgetary figures, excess demand may even exist when budget figures increase (yet not sufficiently). Thus, measuring budgetary crises requires a sense of how the financial situation of the organization is perceived by actors within the organization (or by close observers outside of the organization). In the case of international bureaucracies, especially those involved in operational activities that are usually funded through voluntary contributions, we expect that these IPAs want to signal their perception of policy crisis to their member state principals, donors and, eventually, to the public at large as an indication for increased income needs. Since this may not result in a positive response, at least not immediately, we expect signaling of budgetary stress to come directly with, or shortly after, the time at which IPAs signal the perception of policy crisis. Thus, whereas the underlying hypothesis is that policy crises can cause extreme budgetary stress (H1) in international bureaucracies, the underlying assumption (H0) is that IPAs actually use documents (such as regular or special reports) or speech (such as interventions from Directors-General) to signal this stress to outside specific audiences, i.e. to provide argumentation for additional funding requests, to explain resource shifts or to warn about potential funding cuts. Communication of the perception of policy crisis and budgetary stress thus can be seen as one means of crisis management. In combination, this results in the following hypothesis:
Hypothesis. Under conditions of (perceived) policy crisis, international bureaucracies are expected to signal increased (perceived) budgetary stress to principals, donors and/or the public through documents or speech produced by the bureaucracy.

Where policy crisis and budgetary stress come together, crisis management thus does not just involve addressing policy issues and supporting populations in need, but it also requires resource management as one of the means of overall crisis management. While this paper is primarily concerned with the relationship formulated in the hypothesis above, it is expected that, under conditions of financial stress, international public administrations have the option to either centralize budgeting to control expenditure (see also Goetz and Patz 2016) or to centralize resource mobilization to focus the political attention of the bureaucracy towards mobilizing additional funding (Patz and Goetz 2017a). The higher the share of voluntary contributions, the more likely it is that we should observe a focus on resource mobilization efforts. This overall framework and theoretical model is represented in Figure 1 below.

In the next section, we discuss how to measure the perception of policy crisis and budgetary stress over extended time periods and across organizations, considering the sources available from international public administrations in which such signaling is expected commonly take place.
4. The case for computer-assisted content analysis to measure international public administrations’ perception of policy crises and budgetary stress

Computer-assisted content analysis and natural language processing (NLP) are becoming increasingly common methods for data collection and data analysis in the social sciences. Building on seminal contributions on content analysis by Krippendorff (2004), Neuendorf (2002) and others (see review by Duriau 2007), recent advances can be found both in international relations (e.g. King and Lowe 2003) and political science, but also in organizational studies and public management (e.g Pandey, Pandey and Miller 2017), i.e. key fields that provide the conceptual bases for the study of resource mobilization by international public administrations (cf. Patz and Goetz 2017c). However, traditional news (Dalton, Beck and Huckfeldt 1998; De Vreese and Semetko 2001; Young and Soroka 2012b) and social media have been the most frequently used sources for analyses in these fields (Hegelich 2017; Ceron 2017). In addition, party political documents (Slapin and Proksch 2008; Gemenis 2013; Grimmer and Stewart 2013), speeches (Eggers and Spirling 2016; Greene and Haber 2016) or press releases (Grimmer 2009) that can be represented in clear-cut political dimensions and for which sufficiently large validation data sets exist from hand-coded material have been another source in political science studies using automated content analyses. The advantage of these sources is that they are mostly public, and meant to be public, representing public discourses. Because of their publicness, these sources are (increasingly) available in large quantities, and can be compared with previous research using these sources qualitatively or measuring similar phenomena through other types of data.

Studying international organizations, and in particular the work of their administrative bodies, has often been made difficult by an absence of news media coverage of their work; debates taking place behind closed doors; restrictive archival policies due to the diplomatic nature; and inconsistent data that makes it hard to compare changes both over time and across international organizations. However, with increasingly better online documentation of meeting documents and speeches; with standardized financial and budgetary reporting such as those fostered by the introduction of International Public Sector Accounting Standards (IPSAS) in recent years (Bergmann and Fuchs 2017); and with systematic document registers making the output of IOs public, there exists now a sufficiently large and accessible data source for systematic content analysis. For organizations that have been difficult to study, especially with regards to substantive longitudinal studies and to medium-N comparative analyses, quantitative and automated content analysis of newly available text corpora promises substantial and new insights. From unsupervised topic modeling and state
preference measurement based on speeches in the UN General Assembly covering over 45 years (Baturo, Dasandi and Mikhaylov 2017; Gurciullo and Mikhaylov 2017a, 2017b) to the policy output of IOs over 35 years (Lundgren et al. 2017), automated content analysis clearly has entered the realm of IO studies in the narrow sense, using text produced by IOs and in IOs as data to understand over-time and cross-sectional dynamics.

Applying these methods to the study of international public administrations to measure the perception of the international bureaucracies requires identifying types of documents that represent bureaucratic perspectives – i.e. produced by the bureaucracy without formal intervention from other actors amending those documents – and that are comparable in nature, at least over time but ideally also across organizations. One particularly valid and reliable source to analyze stability and change in IPAs’ perception of their external and internal environment should be routinely produced documents, such as quarterly or annual reports or key speeches held regularly by administrative leaders. Directed at a broader audience, but in particular to member state principals and to major donors, they are meant to reflect an objective analysis of the activities of each organization from the perspective of the international bureaucracy. At the same time, they are meant to signal to all principals, financial supporters and potentially to the public at large which challenges the organization has faced or is expected to face in the past or upcoming period. Changes in topics, sentiments or terminology in those reports from one year to the next (one decade to the next, etc.) can be expected to reflect actual changes in perception, reflecting both positive and negative developments that can be expected to echo phenomena such as financial and geopolitical crises that IPAs face, and to identify at what times these crises are most acute and the measures needed to address those crises are most substantive.

5. The history of UNRWA as a series of external crisis and budgetary stress: a computer-assisted content analysis of UNRWA annual reports 1951-2016

Finding a valid and reliable measurement of how international public administrations perceive their external environment over long periods of time is challenging. Measuring the perceived presence or absence of external crisis directly affecting the work of IPAs, and thus the need to put in place appropriate crisis management, is only one of the possible ways to address the perception of their environment by IPAs. In this section, we present a computer-assisted content analysis approach to identifying those years over the course of its 66-year history in which UNRWA’s administrative leadership signaled through its annual reports (a) the
perception of policy crisis as one for of external crisis, and (b) the perception of budgetary pressure as a form of organizational crisis.

UNRWA is a pertinent case for developing a way to measure how policy crises resulting from regional and geopolitical conflicts may result in extreme budget-related organizational crises and how, potentially, resource mobilization efforts by international bureaucracies are adapted as a result to these crises. It is pertinent because prior research considered that UNRWA is an organization in a “perpetual budget crisis” (Misselwitz and Hanafi 2010, 385). This crisis is created by an almost permanent discrepancy between budgeted income (i.e. resource needs) and actual income (Bartholomeusz 2010, 474), but also by events such as UNRWA’s extreme financial crisis in the 1990s, which is said to be “not a new episode … but rather a recurrent one” (Bocco 2010, 233).

Data collection and preparation of data analysis

We selected all 66 annual reports published by UNRWA from 1951 to 2016 as the population of cases. The full text of most of these cases was available via UNISPAL, the UN page dedicated to “The Question of Palestine”. Where full texts were not available (1963-66, 1972-76, 1981, 1992, 1994-96, 1999-2001), we used the scanned versions available via the UN’s general document system, applied an automated text recognition function and corrected manually where the text in these documents was not correctly recognized. We then removed all formatting, tables of contents, headers and sub-headers, foot- and endnotes, tables and performed a few smaller cleaning operations, retaining only the introductory letters (usually by the head of UNRWA and the by the Chairman of UNRWA’s intergovernmental Advisory Commission) and the textual paragraphs of the main reports.

To measure the level of policy crises that affected the organization, we apply a general sentiment analysis based on the Lexicoder Sentiment Dictionary (Young and Soroka 2012a; for a discussion see Young and Soroka 2012b). This dictionary allows the classification of text, employing a 1-gram bag-of-words approach, using a list of negative words and a list of positive words, allowing to count how often words from each of these two categories appear in a text. Terms like “crisi*”, “war”, “emergency*”, “conflict*” or “distruct*” are part of the dictionary’s negative term list (with a total of 2858 terms), suggesting that policy crisis and

---

2 The report for 2017 was not yet available at the time of writing.
5 The texts as manually cleaned until here are available for replication at [LINK].
regional conflicts should be detected by the dictionary. We calculate the negativity of each annual report as the share of negative words among all words (negative plus positive) identified through the dictionary as a proxy for the occurrence of geopolitical or regional conflicts affecting the organization, assuming that conflicts and their effects should increase the use of negative words in relation to the positive words in a text.

To measure the level of financial crisis, we construct a custom dictionary of budget-, finance- and resource-mobilization-related words that allows us to identify different aspects of financial reporting within an annual report. Based on the knowledge of budgeting and resource mobilization terminology in the UN system (see for example Goetz and Patz 2017; Graham 2016, 2017; Patz and Goetz 2017a, 2017b, 2017c; JIU 2007, 2014) and the selective reading of a few of the annual reports, we constructed a custom dictionary with keywords (1-grams) that reflect four key topics related to the financial situation of IOs, in particular in the UN system: (1) budgetary pressure terms; (2) resource mobilization-related terms; (3) general financial matters, in particular on the income side; and (4) expenditure-related terms (see Annex 1 for this custom dictionary). We use this dictionary to measure the share (%) of budgetary pressure-related terms among all terms included in (1)-(4) at the level of documents (i.e. (1)/(1-4)%).

The findings

Following the description of the data selection and clearing as well as the choice and construction of the two dictionaries, we applied both the Lexicoder sentiment dictionary and our custom dictionary to measure the two key variables in our hypothesis – perception of policy crisis (with overall negativity as proxy) and budgetary pressure in the 66 annual reports of UNRWA. Figure 2 below shows the results, together with selective key events affecting Palestine refugees during the period analyzed.
The first observation is that there is substantive variation in the negativity of UNRWA annual reports from 1951 to 2016, ranging from 24% negative terms in 1960 to 46% negative terms (as shares of all terms identified by the Lexicoder dictionary). This suggests that this measure can indeed detect changing perceptions of negativity and policy crisis over time in administrative documents. Comparing the increases of negativity with selected real-life events – especially the Six-Day-War in 1967, the beginning of the 1st Intifada in 1987 and the beginning of the 2nd Intifada in 2000, and the start of the Syria War onwards – underlines that there seems to be some correlation between outside events and the perception of increasing policy crisis inside the organization. The peak in negativity in 1991 comes at a time when the combined crisis of the 1st Intifada and the Gulf War came together.

Falsifying the assumption of our hypothesis would require finding no link between the negativity in the annual reports and share of budget-pressure related terms among all financial and budget-related terms in our custom dictionary. Over the complete 66-year period, the correlation between both measures is very weak (0.05), but a closer observation of the two curves shows that until the mid-1980s the curves seem to behave relatively similar, but that especially in the late 1980s and during the 1990s the relationship goes into opposite directions. Thus, whereas the perception of an extreme policy crisis in the late 1980s/early 1990s does not seem to be associated with the perception of extreme budgetary pressure, the
reduction of perceived policy crisis in the early 1990s came with a steep increase of perceived budgetary pressures. One interpretation could be that the financial effects of policy crisis in the late 1980s only had negative budgetary effects in the early 1990s, but the question would be why there was this change to a relatively long time-lag. An alternative explanation (Bocco 2010, 233) could be that UNRWA’s financial crisis in the 1990s has resulted from a shift in donor support to other crises that resulted from the end of the Cold War. Said differently, synchronous policy crises without direct effect to the population of concern of UNRWA can still have a negative budgetary effect if UNRWA donors divert funding to other organizations. Instead of effecting the expenditure side, synchronous crises thus can have effects on the income side.

This interpretation would still require an understanding of why the policy crisis in the late 1980s is associated with reduced perceived budgetary pressure. One interpretation could be that resource mobilization in the 1980s was particularly successful to react to the challenging policy environment. Looking at the evolution of resource mobilization terminology over time (see Figure 3 below), there are first indications that this could be the case, with a first small peak of this terminology around that time, but more detailed research is necessary to better understand these dynamics.
Figure 3. Evolution of resource mobilization terminology (see custom dictionary in Annex I) in comparison to the budgetary pressure terminology.

These observations made on the basis of the custom dictionary also have to be taken with a grain of salt, as this dictionary (unlike the Lexicoder dictionary) has not yet been tested and validated. Whereas we undertake a first validation of the findings below, more detailed validity tests are necessary to confirm that we are actually measuring perceived budgetary pressure with this measurement instrument.

Validation of findings

Following the traditional means of validation (Zeller and Carmines, 1980), we assess first the content validity of our findings to show that we are actually measuring what we intend to measure; then we look at criterion validity to see whether our findings match alternative sources or ways of measuring; and, finally, we take a first look at construct validity, asking whether our measure behaves as theoretically expected.

As far as content validity of the variable “Perception of Policy Crisis” of the main hypothesis is concerned, it is challenging to assess in how far the negativity captured by Lexicoder Sentiment Dictionary captures particular policy crises that affected UNRWA. However, as we have discussed above, at least some of the terms included in the Lexicoder dictionary actually measure terms that are clearly linked to policy crisis and conflicts. With regard to criterion validity, Figure 4 below shows a comparison with another sentiment dictionary developed to
capture sentiments in financial reporting (first version presented by Loughran and McDonald 2011). The measurement of overall negativity from both dictionaries (as share of all terms captured by both dictionaries) is strongly correlated (.89), indicating that both measure similar concepts. The fact that the dictionary developed for the analysis of sentiments in financial reporting finds a higher share of negative terms for most of the history of UNRWA could represent the constantly bad financial situation of the UNRWA. The only time that the two curves behave significantly different is in the early to mid-1990s, in which we find significantly less negativity in the general sentiment analysis (LSD) but a much lower decrease and even small increase in the negativity of the dictionary that measures financial reporting. This difference in trend matches our observations for the differences between the trends in the 1990s between the overall negativity (Lexicoder) and the budgetary pressure (custom dictionary) measurements. One could argue that the Loughran and McDonald dictionary captures the combined crisis, but makes it difficult to separate different trends between perceived policy crises and perceived budgetary pressures.

![Comparison of dictionary measurements of negativity in UNRWA annual reports](image)

Figure 4. Comparison of the measures of two different dictionaries (Lexicoder and Loughran & McDonald) applied at the document level.

Looking more specifically at the validity of the measurement of our second variable “Perceived budgetary pressure” through our custom dictionary (see Annex I), there are a number of observations to be made. At present, the main shortcoming of the dictionary is that the bag-of-words approach in our analysis is a 1-gram based dictionary, i.e. that we do not
measure phrases like “financial crisis”, which would represent perception of budgetary pressure, but instead only measure “financial”, which is a general income term in our dictionary, thus reducing the overall negativity in our measure (which measure share of budget pressure terms to all terms). At the same time, terms that are meant to measure perception of budget pressure like “financial” or “cut*” may refer to situations that are not related to the financial situation of UNRWA itself (i.e. the UNRWA report talks about third actors) or may not even refer to financial aspects at all (e.g. “cut*” also captures the physical act of “cutting”). Thus, to assess the content validity of our custom dictionary, we need additional validation.

For criterion validity, we first looked at qualitative measurements from other researchers using the same data. Our custom dictionary’s measurement of financial problems indicating the move towards a financial crisis from 1967 onwards to a peak in 1969 do indeed match Buehrig’s (1971) qualitative analysis, which was also mainly based on the annual reports of UNRWA. Our data suggests 11.48% financial problems-related terms in 1966, compared to 7.45% in 1967, 12.66% in 1968 and 15.86% in 1969. Here’s the description by Buehrig (1971, 173) for these four years:

“At the end of 1967 UNRWA’s reserves were higher than the year before [1966, the authors]. However, they dropped sharply in 1968 and were reduced to a bare minimum in 1969.”

Given that 1967 is the year in which Israeli-Arab relations deteriorated after the Six-Day-War, this also speaks for a certain degree of construct validity, because we expect one cause of financial crises in IOs working in the field of global refugee policy to be the increase in resource needs as a consequence of violent conflicts that lead to significant increases of refugee populations.

Nevertheless, not all qualitative accounts do match our observations: Whereas Rempel (2010) notes “a series of financial crises in the 1970s” as a result for lack of funding and additional expenses following the 1967 war and the need to relocate the agency headquarters from Beirut to Vienna due to the civil war in Lebanon in 1978, our data indicate a rather pronounced crisis right after 1967 (see above) and then a major crisis in the early 1970s and an increasingly difficult situation towards the beginning of the 1980s. There are two ways of interpreting this lack of criterion validity: either our own data lacks precision, or the qualitative description by Rempel is less precise. The overall direction is clear in both accounts, but the details of timing don’t match. Looking into the content of the annual reports
from the 1970s, we find that the construct validity that underlines both Rempel’s and our assessment still match, meaning that we both find that the financial crises are the result of regional conflicts (the Six Day War, the Lebanese civil war).

An alternative way to validate the document-level, bag-of-words findings – and potentially an alternative way of measuring the concepts we are interested in – is to proceed to a sentence-level analysis. For this, we tokenized the annual reports into natural sentences, and perform the measurement with the three dictionaries at the sentence level. Of the 66,802 sentences identified, only 13,445 (20.1%) actually contain at least one word from our custom dictionary, suggesting that a sentence-level analysis could find different results than a document level-analysis on the measurement of the perception of budgetary pressure. As Figure 5 below indicates, the general trend of the document-level approach to the sentence-level approach is virtually the same.

Assessing the content validity of our measurement, we selected the 2,292 sentences that contained at least one term from the budget pressure category of our custom dictionary. Of these, only 1,743 are identified as having a negative sentiment in the Lexicoder dictionary,

---

Note that this figure is artificially high as the sentence-level analysis will require further cleaning of the source documents to ensure that not too many false sentences are extracted. The most common example in our dataset are the numbers plus dots at the beginning of paragraphs that are currently counted as sentences.
and only 1,655 are identified as negative in the Loughran & McDonald dictionary. This could indicate that not all budgetary pressure terms actually represent a negative context. On the other hand, 2,744 sentences containing at least one term from the “income” category of our custom dictionary (each of these reduces the share of perceived budget pressure terms) contained at least one negative term from each of the two sentiment dictionaries, showing that many neutral or positive terms in the custom dictionary may also appear in negative contexts (such as “financial” and “crisis”). This clearly indicates the need for additional sentence-level validation through human coders to better understand which terms in the custom dictionary contribute to the most valid measurement of perceived budgetary crisis, and which combination of custom term dictionary and sentiment dictionary provides the most valid identification of sentences representing the perception of budget pressure.

4. Conclusions

We have argued in this paper that understanding crisis management in and by international organizations and their administrations requires a better understanding of when and how policy crises are actually perceived as such by international bureaucracies. Our theoretical assumption was that the perception of policy crises, in the context of UNRWA in particular regional conflicts affecting Palestine refugees, should be related to increased perception of budgetary pressure. It has then been demonstrated that a computer-assisted content analysis of UNRWA annual reports based on pre-existing sentiment dictionaries is able to detect shifts in negativity over time, which can reasonably be interpreted in shifts of perception of policy crisis by UNRWA’s bureaucracy that are related to actual outside events, thus appear relatively valid. This supports our assumption that international bureaucracies use documents and speech to indicate policy crisis to outside audiences, and that it is possible to make a valid measurement over extended time periods.

The patterns showing the evolution of the perception of budgetary pressure by UNRWA’s bureaucracy as measured through a custom dictionary developed by the authors indicate that the latest “unprecedented” budgetary crisis of UNRWA does not appear to be unprecedented, for example compared to the situation in the late 1960s or in the mid-1990s. However, there are still quite a number of questions regarding the validity of the measurement of the perception of budgetary pressure, and further refinement of the dictionary itself and a better understanding of the best level of measurement (document, paragraph, sentence) is necessary to increase the precision of this measure. Nevertheless, the measure presented here already
captures important shifts in resource availability over time, and, in combination with the sentiment analysis, provides for interesting findings such as the inverse trends between the perception of policy crisis and budgetary stress in the late 1980s (policy crisis increases, budgetary stress decreases) and in the 1990s (policy crisis decreases, budgetary stress increases). A first analysis suggests that the former may indicate successful resource mobilization as part of the policy crisis management strategy, with donors being ready to support the solution of the policy crisis. The latter indicates that budgetary stress may result from crises unrelated to the population of concern of UNRWA, simply because these other crises shift donor attention away from the organization. In particular, these latter findings speak to the emerging debate on resourcing international organizations (Goetz and Patz 2017; Michaelowa 2017) and the conclusion that wider organizational fields need to be studied to understand dynamics that exceed the realm of single international organizations (Goetz and Patz 2017, 10).

With regard to crisis management, this means that we require a better understanding of what constitutes a policy crisis affecting only one organization versus policy crises that affect a whole field of organizations. And, where (perceived) policy crises come with (perceived) budgetary crises (as one form of organizational crisis), overall crisis management will have to address both challenges at the same time.
References


Documents cited


### Annex I – Custom Dictionary (version 1)

<table>
<thead>
<tr>
<th>Budgetary pressure</th>
<th>Resource mobilization</th>
<th>Income</th>
<th>Expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td>austerity</td>
<td>appeal*</td>
<td>budget*</td>
<td>cost*</td>
</tr>
<tr>
<td>bankrup*</td>
<td>donor*</td>
<td>contribut*</td>
<td>expenditure*</td>
</tr>
<tr>
<td>borrow*</td>
<td>extra-budg*</td>
<td>financi*</td>
<td>spend*</td>
</tr>
<tr>
<td>burden*</td>
<td>extrabudg*</td>
<td>fund</td>
<td>spent</td>
</tr>
<tr>
<td>claims</td>
<td>fund-rais*</td>
<td>funded</td>
<td></td>
</tr>
<tr>
<td>constrain*</td>
<td>fundrais*</td>
<td>funding*</td>
<td></td>
</tr>
<tr>
<td>cut*</td>
<td>genero*</td>
<td>funds</td>
<td></td>
</tr>
<tr>
<td>deficit*</td>
<td>lobby*</td>
<td>resource*</td>
<td></td>
</tr>
<tr>
<td>depleted</td>
<td>mobiliz*</td>
<td>source*</td>
<td></td>
</tr>
<tr>
<td>Inflation</td>
<td>partner*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-reimbursement</td>
<td>pledg*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>outstanding</td>
<td>pric*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>shortage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shortfall*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>underfund*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unfund*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zero-growth</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ANNEX II – R Code for analyzing the text at document level

#Preparation: COPY all UNRWA annual reports into the working directory
#Preparation: COPY all three dictionaries (Lexicoder Sentiment Dictionary, the Custom Dictionary prepared in Yoshikoder, and the Loughran/McDonald dictionary for financial reporting into the working directory

#load quanteda and readtext packages
library("quanteda",
  lib.loc="/Library/Frameworks/R.framework/Versions/3.4/Resources/library")
library("readtext",
  lib.loc="/Library/Frameworks/R.framework/Versions/3.4/Resources/library")

#Read all UNRWA annual reports
UNRWA.annualreports <- readtext("*.txt")

#Create a text corpus based on all texts
UNRWA.corpus <- corpus(UNRWA.annualreports)

# Create a data-frame-matrix based on the corpus
UNRWA.dfm <- dfm(UNRWA.corpus)

#Tokenize the dataframe and remove punctuation, hyphens and numbers
UNRWA.dfm <- dfm(tokens(UNRWA.corpus, remove_punct=TRUE, remove_hyphens=TRUE, remove_numbers=TRUE) )

#Read in the three dictionaries
dictionary.LSD<-dictionary(file="LSD2011.CAT", format="wordstat", concatenator = ",")
dictionary.Custom<-dictionary(file="UNRWA_Dictionary.ykd", format="yoshikoder", concatenator = ",")
dictionary.LM<-dictionary(file="LoughranMcDonald2014.CAT", format="wordstat", concatenator = ",")

#Compare dictionaries with UNRWA data frame matrix, counting number of occurrences of words in each category of the dictionaries with the respective annual reports. The output is one line per document (i.e. year) and the columns containing the count of word occurrence per dictionary category.
results.LSD<-dfm_lookup(UNRWA.dfm, dictionary.LSD)
results.Custom<-dfm_lookup(UNRWA.dfm, dictionary.Custom)
results.LM<-dfm_lookup(UNRWA.dfm, dictionary.LM)

#Procedure output documents
write.csv(results.LSD, file="resultsLSD.csv")
write.csv(results.Custom, file="resultsCustom.csv")
write.csv(results.LM, file="resultsLM.csv")