The Foerder Institute for Economic Research
at Tel Aviv University

מכון למחקר כלכלי על שם
ד"ר ישעיהו פורדר
על יד אוניברסיטת תל אביב

The Eitan Berglas School of Economics

בית-הספר לכלכלה ע"ש איתן ברגלס

עמותה רשומה

Bureaucracy in quest of feasibility

Herve Cres a, Itzhak Gilboa, Nicolas Vieille

Working paper 5-2025

The Foerder Institute for Economic Research
and
The Sackler Institute of Economic Studies

Journal of Mathematical Economics 114 (2024) 103047

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Bureaucracy in quest of feasibility☆

Hervé Crès[a], Itzhak Gilboa[b,c,*], Nicolas Vieille[b]

[a] NYU Abu Dhabi, United Arab Emirates
[b] HEC, Paris, France
[c] Reichman University, Israel

ARTICLE INFO

Manuscript handled by Editor C Thomas

Keywords:
Decision theory
Organizations
Condorcet
Arrow's theorem
Bureaucracy

ABSTRACT

A bureaucracy has to determine the values of many decision variables while satisfying a set of constraints. The bureaucracy is not assumed to have any objective function beyond achieving a feasible solution, which can be viewed as "satisficing" à la Simon (1955). We assume that the variables are integer-valued and the constraints are linear. We show that simple and (arguably) natural versions of the problem are already NP-Hard. We therefore look at decentralized decisions, where each office controls but one decision variable and can determine its value as a function of its past values. However, an attempt to consult more than a single past case can lead to Condorcet-style consistency problems. We prove an Arrovian result, showing that, under certain conditions, feasibility is guaranteed only if all offices mimic their decisions in the same past case. This result can be viewed as explaining a status quo bias.

1. Introduction

1.1. Motivation

The framework of decision theory suggests an image of an organization as an entity that maximizes an objective function, subject to certain constraints. The most prominent example is the textbook model of a firm as a profit-maximizing agent. Many not-for-profit organizations can also be viewed as coherent decision makers with objective functions that measure their performance. When it comes to polities and other large organizations whose goals are less crisply defined, economists often view the organization as a game played among several utility-maximizing agents, diverging in their goals, their information, and so forth. (See, for instance, Milgrom and Roberts, 1992.) However, there are situations in which organizations are run by bureaucracies that seem to care about feasibility without having a very well-defined objective function to distinguish among feasible solutions.

Consider a university that plans its course offering for the coming year. The problem involves a complex "teaching matrix" that has to match students, course sections, teachers, classrooms, and time slots. There are some obvious constraints to satisfy: each class should be assigned a teacher, a classroom, and weekly time slots. Each teacher should be assigned to classes adding up to her teaching load, while respecting constraints on the topics a person can teach. The number of students in a class cannot exceed the classroom size. And on it goes. Once all these constraints are satisfied, one might start thinking about optimization, perhaps taking into account teachers' preferences for time slots, students' preferences for teachers, etc. Yet, it seems that the major part of the bureaucracy's problem is finding a feasible solution, rather than finding the best one among several such solutions. Consider, for concreteness, two mistakes a bureaucracy can make in this context:

Mistake 1: Assigning two classes to the same classroom at the same time slot;

Mistake 2: Assigning two teachers to two sections in a way that could be improved upon, according to the two teachers' preference, by a swap.

There are at least three reasons for which Mistake 1 draws more organizational attention than Mistake 2. First, the sheer magnitude of the cost involved: Mistake 2 would make the two teachers less content that they could have been, whereas Mistake 1 generates considerable mess. Second, measurability: the degree to which teachers are content with their time slots is a bit vague; by contrast, conflicting schedules are easily observable. Lastly, information also distinguishes between

☆ We thank Antoine Billot, Eddie Dekel, David Schmeidler, and Yoram Weiss for conversations and references, as well as the editors and two anonymous referees for comments that greatly improved the paper. We gratefully acknowledge support from the Investissements d'Avenir ANR -11- IDEX-0003/ Labex ECODEC No. ANR - 11-LABX-0047 as well as from Tamkeen under the NYUAD Research Institute award for project CG005 and from the Center for Institutional Behavioral Design (Crès) and from ISF Grant 1443/20, the AXA Chair for Decision Sciences at HEC, the Foerder Institute at Tel-Aviv University and the Sapir Center for Economic Developmemt (Gilboa).

* Corresponding author at: HEC, Paris, France.
E-mail addresses: herve.cres@nyu.edu (H. Crès), tzachigilboa@gmail.com (I. Gilboa), vieille@hec.fr (N. Vieille).

https://doi.org/10.1016/j.jmateco.2024.103047
Received 22 September 2023; Received in revised form 21 July 2024; Accepted 8 August 2024
Available online 15 August 2024

the mistakes: avoiding Mistake 2 requires information about teachers' preferences, which may not be shared by all; by comparison, avoiding Mistake 1 does not require more than knowledge of regulations (or just common sense in obvious cases).

Given these considerations, it makes sense to assume that various officers within a bureaucracy would be driven mostly by satisfying constraints, avoiding obvious calamities, and may not attempt to achieve optimality based on less measurable goals and potentially unavailable information. This leads us to model bureaucracies as seeking feasible, rather than optimal solutions.

Clearly, a bureaucracy that only seeks a feasible solution can be viewed as utility-maximizing, where the utility function is binary, treating all feasible solutions as equally good (and all infeasible ones as equally bad). The focus on such special utility functions can provide insights that may be harder to observe in a more general model. By analogy, Simon (1947) satisficing behavior can be viewed as a result of discounted expected utility maximization (Gittins, 1979), but the generality of the latter might becloud insights provided by the former. In fact, this analogy is not coincidental: in our model a bureaucracy attempts to reach a feasible solution, and it may be viewed as "satisficed" as long as it manages to do so.

1.2. Outline

We start, in Section 2, with the general presentation of the model, and a brief comment on computational complexity. We note that even highly simplified versions of problems bureaucracies deal with turn out to be computationally intractable. We then suggest that, for that reason, bureaucracies are likely to attempt to repeat decisions that worked well in the past.

In Section 3 we impose the additional constraint that each office within a bureaucracy can only rely on its own past decisions. This assumption attempts to capture limitations on information flow within an organization. Under this assumption, we show that setting a variable to its most commonly observed value may result in a problem: a Condorcet-style example shows that, following this rule, a sequence of decisions that satisfy a constraint may result in a decision that does not. Past cases are analogous to voters, the constraint — to the transitivity condition. Along similar lines, we prove an Arrovian theorem, stating that, under mild conditions, the only way past feasible choices can be aggregated into a new one, while respecting feasibility, is by mimicking a single case. Thus, focusing on one, "most similar" case in the past is equivalent to declaring one voter a dictator.

As opposed to Arrow's theorem (and related results that followed), our assumptions and result have a descriptive rather than a normative flavor. Arrow's assumption of IIA (Independence of Irrelevant Alternatives) is replaced in our context by the assumption that the bureaucracy is constrained to make decisions in a decentralized way. This assumption is based on organizational constraints, and has no normative flavor. Similarly, Arrow's result conveys a negative message, because a dictatorship is considered to be a bad outcome. By contrast, in our model the result that only a single case will be mimicked is morally neutral. It is, however, offered as a descriptive model of the way bureaucracies might function.[1]

Further discussion, and a survey of the related literature are deferred to Section 4.

2. Model

The bureaucracy is viewed as solving an integer programming problem at each period t. It controls a vector of $n_t \geq 1$ non-negative integer-valued decision variables $\mathbf{x}_t = \left(x_{1,t}, \dots, x_{n_t,t}\right)'$ and it has to assign values to them so that

$$A_t \mathbf{x}_t \leq \mathbf{b}_t$$

where A_t is an $m_t \times n_t$ matrix of real numbers, and $\mathbf{b}_t$ is a vector of m_t extended real numbers, $b_{j,t} \in \mathbb{R} \cup \{\infty\}$ for $j \leq m_t$. We denote the set of possible right-hand-side (RHS) vectors by $B_t \equiv (\mathbb{R} \cup \{\infty\})^{m_t}$. As $A_t, \mathbf{b}_t$ are not constrained by sign, there is no loss of generality in assuming only constraints of the type $\leq$.

The integer-valued variables $\left(x_{1,t}, \dots, x_{n_t,t}\right)$ denote decisions such as whether a given teacher is assigned to a given class, or whether the teacher-class combination is assigned a certain classroom. Thus, the n_t decision variables in period t may result from higher-dimensional arrays.

To simplify notation, we will assume that n_t and A_t are independent of t. This involves no loss of generality because, at each period t, we can take the union of all decision variables that appeared in periods $\tau \leq t$, which is a finite set of variables. That is, if the ith variable was not actually present in period τ, it can be formally included in the problem with the constraint $x_{i,\tau} \leq 0$. Similarly, we can consider the matrix A that consists of the union of all rows of the matrices $\left(A_\tau\right)_{\tau \leq t}$ and, when a certain constraint does not appear in period τ, set $b_{j,\tau} = \infty$. Thus, we assume that at each period t the bureaucracy faces the problem

$$A \mathbf{x}_t \leq \mathbf{b}_t$$

with period-independent number of variables n, an $m \times n$ matrix A, and period-dependent RHS vector $\mathbf{b}_t \in B_t = (\mathbb{R} \cup \{\infty\})^m \equiv B$.

Before completing the description of the model, we comment, in Section 2.1, on the complexity of the bureaucracy's problem. Because the problem is computationally complex, it seems natural to seek guidance in history. Thus, in Section 2.2 we formally introduce past problems, and define the special case in which the same problem has been repeated throughout history. In Section 3 we formally impose the additional assumption, namely, that each variable is a function only of its own past values, and prove the main result.

2.1. A comment on complexity

It is well-known that integer programming (IP) is, in general, NP-Hard. Similarly, the nonemptiness of the feasible set of an IP problem is NP-Complete. (See Gary and Johnson, 1979.) However, organizations need not solve general IP problems. We therefore highlight the fact that the type of problems bureaucracies face suffices to generate high computational complexity.

Specifically, consider the following problem. There are two employees and n tasks that need to be performed. Each task $i \leq n$ can be performed by each employee, and it takes a_i units of time (irrespective of the employee chosen to perform it). The bureaucracy has a budget of $b_1, b_2 \geq 0$ hours of each employee's time, and the question is: Is there an allocation of tasks to employees such that no employee is asked to work more hours than she is supposed to? Formally, define

Problem 1 (*TASK-ASSIGNMENT*)**.** Given $n \geq 1$, integers $\left(a_i\right)_{i \leq n}$, and $b_1, b_2 \geq 0$, is there a subset $S \subset \{1, \dots, n\}$ such that $\sum_{i \in S} a_i \leq b_1$ and $\sum_{i \notin S} a_i \leq b_2$?

Clearly, this problem can be stated as an IP problem. Suppressing the time index t, let there be n decision variables, where $x_i \in \{0, 1\}$ indicates whether task $i \leq n$ is assigned to employee 1 $\left(x_i = 1\right)$ or to employee 2 $\left(x_i = 0\right)$. A vector $\mathbf{x} = \left(x_i\right)_{i \leq n}$ is a feasible solution iff

$$\begin{aligned} x_i &\leq 1 \qquad \forall i \leq n \\ \sum_{i \leq n} a_i x_i &\leq b_1 \\ \sum_{i \leq n} -a_i x_i &\leq b_2 - \sum_{i \leq n} a_i \end{aligned} \tag{1}$$

(where the last inequality is a rearrangement of $\sum_{i \leq n} a_i \left(1 - x_i\right) \leq b_2$).

We note that

[1] A similar contrast exists between our result and the impossibility result of List and Pettit (2002), who use an Independence assumption and obtain a result by which only dictatorial aggregation of opinions is guaranteed to retain consistency.

Remark 2. TASK-ASSIGNMENT is NP-Complete.[2]

Thus, even a highly restrictive version of the course scheduling problem discussed above is too complex to be solved correctly (in general). In particular, we cannot expect a polynomial-time algorithm to find a feasible solution for this problem whenever one exists.

2.2. Consulting history

In the absence of efficient algorithms for finding feasible solutions, a bureaucracy may use heuristics, and, specifically, consult its past decisions. Indeed, problems that organizations face are typically not encountered without a context. A person who is new to a management position would usually start by learning what has been done by others holding the position before her. If the problem she is facing has been encountered in the past, she may mimic the solutions found by her predecessors.

Importantly, history may serve as a useful guide only if the current period is not dramatically different from the past. Indeed, it is possible that the current problem has nothing to do with past ones. For example, it might be the case that for any j with $b_{j,t+1} < \infty$ we have $b_{j,\tau} = \infty$ for all $\tau \leq t$, so that past decisions solved different problems. In this case the feasibility of past problems does not imply that the present problem is feasible. Further, the bureaucracy is facing an NP-Complete problem for which history does not provide any guidance. The tendency to consult history is therefore justified by the implicit assumption that history does not change too much across periods. We will focus on the extreme, and presumably simplest cases, in which history is constant, and a feasible solution has been found in each period in the past.

Formally, consider a history $h_t = \left(\left(\mathbf{b}_1, \mathbf{x}_1\right), \ldots, \left(\mathbf{b}_t, \mathbf{x}_t\right)\right)$ and a RHS $\mathbf{b}_{t+1}$. The pair $\left(h_t, \mathbf{b}_{t+1}\right)$ is *regular* if

(i) $\mathbf{b}_\tau = \mathbf{b}$ for all $\tau \leq t+1$;

(ii) $\mathbf{x}_\tau$ is feasible for all $\tau \leq t$ (that is, $A\mathbf{x}_\tau \leq \mathbf{b}$).

The case of regular $\left(h_t, \mathbf{b}_{t+1}\right)$ should be viewed as a benchmark, or a minimal requirement: any reasonable decision procedure should be able to find a feasible solution in such cases.

Bureaucracy is decentralized, to a degree. For reasons having to do with time and expertise, different offices control different decision variables. We consider here the simplified (extreme) case in which each variable $x_{i,t}$ is controlled by a separate office, and, while setting the value of $x_{i,t}$, the i'th office can only consult its own past decisions, $\left(x_{i,\tau}\right)_{\tau<t}$. Importantly, this last assumption is not suggested as universally accurate. We study its implications, which can sometimes serve as a descriptive model of the way bureaucracies operate. Importantly, our result can also be used for normative purposes: by indicating when decentralization may lead to inconsistencies, the result implicitly suggests that in these situations it may be advisable to bear the cost of a centralized decision.

How should $x_{i,t}$ depend on $\left(x_{i,\tau}\right)_{\tau<t}$? One possibility is to consider the choice that has been made most frequently in the past, along the lines of "this is how we normally do it". Another would be to pick the choice that was made in the single most-similar problem in the past. Naturally, many other possibilities exist. For example, one can aggregate over past cases but give higher weight to more similar cases, resulting in a procedure that is akin to kernel classification in statistics (see Akaike, 1954; Parzen, 1962; Silverman, 1986). Alternatively, one may "upgrade" the most-similar problem approach and allow the decision maker to consult several similar problems, along the lines of k -nearest neighbor classification (Fix and Hodges, 1951, 1952).

In the next section we discuss the consistency problems that may arise in using multiple past cases, and prove the main result, showing that these problems can be avoided only if a single case is consulted.

3. Feasibility and the prominence of a single case

3.1. A condorcet-style problem

Consider the following example, which is another highly simplified version of the course scheduling problem discussed in the introduction. Mary can teach any one of three courses, but her teaching load is only two courses. Let $x_{i,t} \in \{0,1\}$ indicate whether she is assigned to course $i = 1, 2, 3$ at period t. Thus, the three variables have to satisfy a single constraint,

$$x_{1,t} + x_{2,t} + x_{3,t} = 2$$

(which, to fit the mold of our model, can be written as two inequalities, $x_{1,t} + x_{2,t} + x_{3,t} \leq 2$ and $-x_{1,t} - x_{2,t} - x_{3,t} \leq -2$). Assume that these courses belong to different programs and are run by different branches of the administration, so that each variable $x_{i,t}$ is governed by a different office.

Next assume that we are at period $t = 4$, and that the three past problems ($\tau = 1, 2, 3$) had the same constraint. However, different feasible solutions were found in these past periods.[3] Specifically, suppose that the observed past decisions are

$x_{i,\tau}$	$\tau = 1$	$\tau = 2$	$\tau = 3$
$i = 1$	1	1	0
$i = 2$	1	0	1
$i = 3$	0	1	1

If each office i selects for $x_{i,t}$ the value that has been most commonly chosen in the past, we will obtain the assignment $x_1 = x_2 = x_3 = 1$, which violates the constraint of Mary's teaching load.

We therefore find that doing "what has been most commonly done in the past" may lead to infeasibility: even if history consists only of solutions that satisfy a certain set of constraints, choosing the most common value *for each* x_i *separately* may lead to an assignment of values which is infeasible. The analogy to Condorcet's paradox is inevitable: in this paradox, a majority vote among agents is taken, for each pair of alternatives separately. Agents who have transitive preferences may yield a majority vote that is not transitive. In our case, the role of voters is played by past cases, and transitivity is replaced by linear constraints. Indeed, the matrix above can be interpreted as an algebraic representation of Condorcet's example: there are three candidates, $\{a, b, c\}$, and three voters, $\{1, 2, 3\}$. Preferences are assumed to be strict[4] and are given by the matrix above, where each row i reflects preferences between a given pair of alternatives: $i = 1$ – the pair (a, b), $i = 2$ – (b, c), and $i = 3$ – (c, a). The assignment $x_{i,\tau} = 1$ means "Voter τ prefers the first element in the pair i to the second". (For example, $x_{1\tau} = 1$ stands for the strict preference $a \succ_\tau b$ for voter τ, while $x_{1,\tau} = 0$ – for $b \succ_\tau a$.) The constraint $x_{1,\tau} + x_{2,\tau} + x_{3,\tau} \leq 2$ rules out the preference cycle $a \succ_\tau b \succ_\tau c \succ_\tau a$ (while the converse cycle would be ruled out by a similar inequality, namely, $x_{1,\tau} + x_{2,\tau} + x_{3,\tau} \geq 1$). The matrix corresponds to a profile of strict, transitive preferences ($a \succ_1 b \succ_1 c$; $c \succ_2 a \succ_2 b$; $b \succ_3 c \succ_3 a$) and, in particular, each voter τ satisfies $x_{1,\tau} + x_{2,\tau} + x_{3,\tau} \leq 2$. Yet, the majority preference has

[2] TASK-ASSIGNMENT is clearly in NP. To see that it is NP-Complete, consider the following version of KNAPSACK (which is known to be NP-Complete, see (Gary and Johnson, 1979)):

Problem KNAPSACK: Given integers $a_1, \ldots, a_n$ and b, is there a subset $K \subset \{1, \ldots, n\}$ such that $\sum_{i \in K} a_i = b$?

To reduce KNAPSACK to TASK-ASSIGNMENT, given a KNAPSACK instance $\left(a_1, \ldots, a_n, b\right)$ one may define a TASK-ASSIGNMENT instance by the same integers $\left(a_i\right)$ and the RHS values $b_1 = b; b_2 = \sum_i a_i - b$.

[3] The reasons for which different choices have been made in the same problem are beyond the scope of our model. We assume that such histories may occur. We return to this point in the Discussion section.

[4] One would need six (rather than three) variables to describe not-necessarily-strict preferences.

$x_1 + x_2 + x_3 = 3$. Correspondingly, in our example each past case is compatible with a given constraint, but the "majority" case is not.

Similarly, the problem we face here is also analogous to the "doctrinal paradox" (or the "discursive dilemma") of List and Pettit (2002). In their example, three judges have to vote on the validity of three propositions, p, q, and r. All judges agree with the "doctrine" that $((p \wedge q) \leftrightarrow r)$ that is, that r can be established if and only if both p and q can. When the judges vote on each proposition separately, majority vote may be inconsistent with the doctrine.[5]

3.2. An Arrow-style theorem

Consider a given period $t+1 \geq 2$. We will consider different histories of length t, $h_t = \left(\left(b_1, x_1\right), \ldots, \left(b_t, x_t\right)\right)$. Focusing on the decisions made in the past, such a history defines an $n \times t$ matrix, so that $x_{i,\tau}$ is the value chosen for the variable x_i at period $\tau \leq t$. The decision at time t, $x_{i,t}$, is made for each i separately, as a function of past decisions made for the same variable. Assume, then, that there is a function

$$f : \mathbb{Z}_+^t \rightarrow \mathbb{Z}_+$$

(with $\mathbb{Z}_+$ denoting the nonnegative integers) such that, for each i,

$$x_{i,t+1} = f\left(x_{i,1}, x_{i,2}, \ldots, x_{i,t}\right)$$

A function f *retains the status quo* if $f(c, \ldots, c) = c$ for all $c \in \mathbb{Z}_+$. It is natural to assume this property in our context: if history offers but a single past decision, it makes sense that this decision be repeated. Clearly, this assumption corresponds to a Pareto or a Unanimity assumption in the impossibility theorems of Arrow (1950) or List and Pettit (2002): in these results, each entry is the vote of an individual in society, and, in case of unanimity, the requirement that society agree with it has a strong normative flavor. By contrast, in our model an entry is a past problem, and the assumption that identical decisions in the past would yield the same decision in the current problem is interpreted descriptively, attempting to capture regularities in the behavior of bureaucracies. The independence assumptions in these models may also have a normative flavor. In our model, the corresponding assumption is that decisions are decentralized, and it is justified based on practical constraints.[6]

A function f will be called a *most-similar-case function* if there exists $\tau \leq t$ such that

$$f\left(x_{i,1}, x_{i,2}, \ldots, x_{i,t}\right) = x_{i,\tau}$$

for all $\left(x_{i,1}, x_{i,2}, \ldots, x_{i,t}\right) \in \mathbb{Z}_+^t$.

As mentioned above, regular pairs $\left(h_t, \mathbf{b}_{t+1}\right)$ can be considered as minimal tests for a suggested function f. Formally, a function $f : \mathbb{Z}_+^t \rightarrow \mathbb{Z}_+$ is *consistent* if it generates a feasible decision for each regular $\left(h_t, \mathbf{b}_{t+1}\right)$.

If the constraint matrix A does not generate much interaction among the decisions, decentralization may not raise any difficulties. Assume, for example, that each constraint involves only one variable. In that case the bureaucracy's decisions may well be decentralized without inconsistencies being a concern. We are interested, however, in problems where decisions do interact. Specifically, we assume that the matrix A satisfies the following condition.

A $(m \times n)$ matrix A contains *two potentially conflicting rows* if there exist two rows $i_1, i_2 \leq m$ and three columns $j_1, j_2, j_3 \leq n$ such that $a_{i_1 j} > 0$ and $a_{i_2 j} < 0$ for $j = j_1, j_2, j_3$. In our example there was one constraint, namely,

$$x_1 + x_2 + x_3 = 2.$$

Translating this constraint to $\leq$ inequalities, it would take the form

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 2 \\ -x_1 - x_2 - x_3 &\leq -2. \end{aligned}$$

These two constraints would appear in the matrix A as

$$\begin{matrix} +1 & +1 & +1 \\ -1 & -1 & -1 \end{matrix}$$

which clearly define A as containing two potentially conflicting rows. Differently put, the condition of "containing two potentially conflicting rows" generalizes the example we started out with, by allowing any three positive values (not necessarily all 1) associated with any three negative values (not necessarily all -1). It turns out that this condition is sufficient to establish our result.

Theorem 3. *Assume that A contains two potentially conflicting rows. Then f is consistent and retains the status quo if and only if it is a most-similar-case function.*

This result can be interpreted as suggesting that a decentralized bureaucracy would be more likely to come up with feasible solutions, relying on past choices, if it picks a single past case to mimic than if it attempts to aggregate over cases. Or, to take it literally, the only way to guarantee consistency is for all branches of the bureaucracy to mimic the solution to one past problem rather than to aggregate solutions to several of them. Thus the single past case plays the role of the dictator in Arrow's theorem: in the latter, choosing one individual to call the shots is the only way to guarantee consistency; in our case – a single past case determines the values of variables, and, again, this is the only way to guarantee consistency.

Interpreting impossibility results such as Arrow's, the identity of the dictator is not an issue. One need not ask, "OK, so who is the dictator, then?" because the analysis is normative and its bleak conclusion is a dead end, at least of a given research path. By contrast, in our case the analysis is descriptive, and it makes sense to ask the next question, "Which case, then, is going to be mimicked by all branches of the bureaucracy?". Luckily, our model has additional structure that suggests a natural candidate for the "dictator case": past cases are ordered by time, and it makes sense to choose the most recent case. Recency may be a "focal point" for the coordination of various branches of the bureaucracy, and it is likely to be more robust than other alternatives to various perturbations of the model. For example, people are more likely to recall the most recent than the least recent case. More importantly, people are more likely to agree on what the most recent case was — while they may disagree on the point at which history started.

To conclude, when different branches consider their past decisions in order to make a current one, it appears that the strategy, "Let us do what we usually do", is more likely to cause problems than "Let us do what we did last year". With this interpretation, our analysis may be viewed as a justification for a status-quo bias.

4. Discussion

4.1. Related literature

Organizations may be thought of as monolithic, rational decision makers, maximizing expected utility under constraints. In the context of the decisions of a firm, the rational decision maker model goes back to Smith (1776), Marx (1867), and Durkheim (1893), with an emphasis on efficiency of production in the early 20th century (Taylor, 1911; Follett, 1918; Fayol, 1919). More generally, any organization

[5] In fact, our example can be precisely mapped to the doctrinal paradox, if we define the doctrine to be that "Mary is *not* be assigned to course 3 (r) iff she is assigned to course 1 (p) *and* she is assigned to course 2 (q)".

[6] Arrow's IIA can be viewed as a normative statement about preference aggregations, and it can also be viewed as a robustness requirement, making the model's recommendation insensitive to manipulations.

that satisfies the axioms of rational choice can be viewed as an expected utility maximizer.

However, organizations do not always seem to be sufficiently coherent to be ascribed a utility function and a subjective probability that would describe their choices via the expected utility maximization paradigm. One may espouse a different view, according to which organizations are games played by different agents, who have different utility functions and perhaps also different beliefs. Viewing organizations in this way originates with Weber (1921, 1924). He viewed bureaucracy as a way of establishing legitimate authority, and of achieving maximal efficiency. Buchanan and Tullock (1962) viewed the state as comprising rational agents with different goals. Niskanen (1971, 1975) analyzed bureaucracy as a production entity, and questioned its efficiency.

Our model is akin to the second strand in the literature, as it does not view the organization as a monolithic agent. However, it differs from the models mentioned above in that it does not ascribe an explicit "payoff" function to the bureaucracy.[7]

Ours is by no means the only model that goes beyond the rational choice paradigm, whether applied to the organization as a whole or to components thereof. Simon (1947) and March and Simon (1958) pointed out the bounded rationality that characterizes organizational decision making. Burns and Stalker (1961) suggested that mechanistic bureaucracies are ill-adapted to deal with changing environments. Kanter (1977) argued that power inside an organization may not be easy to define, and suggested that the seemingly powerful are often powerless.[8] Bendor and Moe (1985) analyzed bureaucracies using bounded rationality models.

More generally, there are many other images that have been used to describe organizations. Morgan (2006) mentions metaphors such as machines (Taylor, 1911; Fayol, 1919; Weber, 1924), organisms (Parsons, 1951; Burns and Stalker, 1961), brains (Sandelands and Stablein, 1987; Walsh and Ungson, 1991; March, 1999), cultures (Ouchi and Wilkins, 1985), and political systems (Burns, 1961; March, 1962). For the most part, these images have not been formally modeled.

The present paper shares much of its motivation with Gilboa and Schmeidler (2011). In particular, that paper formally models organizations as entities that make decisions without a clear utility function, and with a tendency to be consistent with past decisions.

The main result of the paper, presented in Section 3, is related to Arrow's impossibility theorem (Arrow, 1950) and its generalizations by Wilson (1972) and Rubinstein and Fishburn (1986), as well as to the recent literature on judgment aggregation, starting with List and Pettit (2002), and followed by, among others, Vieille (2006), Dietrich (2010), Dietrich and List (2010), Dietrich and Mongin (2010), Dokow and Holzman (2010a,b) and Nehring and Puppe (2010a,b) (see List and Polak, 2010, for an introduction and a survey). Most of the analysis in this literature focuses on binary variables, such as whether alternative a is preferred to b, whether proposition p is true or not, etc. Our result can be viewed as extending the algebraic approach to integer variables that are not necessarily binary. Another distinction between our model and existing ones is that cases, as opposed to voters, are naturally ordered by time. Therefore, when a most-similar-case needs to be picked, history offers a natural candidate, whereas there is no natural candidate for a dictator in the social choice context.

[7] The bureaucracy in our model can be viewed as following the path of least resistance. Mathematically, this can clearly be described as maximizing some objective function. But we do not interpret this function as a "payoff" or "utility" function that can also be interpreted as an explicit goal (as can "profit" be in the theory of the firm).

[8] For extensive introductions to organization theory, see Handel (2003), Hatch and Cunliffe (2006), and Scott and Davis (2007).

4.2. General discussion

Consider a dynamic process of case arrivals, and assume that, at each period t, a function f that satisfies our assumptions is applied. If the set of constraints is constant (as assumed for a regular history), then the decisions at each period will be identical. Indeed, the assumption of retaining the status quo suffices to dictate choice at each period. In this case we will not observe the values of f for other, non-constant histories, and will not be able to distinguish between a function that mimics the most-similar case and other functions that retain the status quo.

Our theorem therefore depends on the richness of possible histories in the domain of f, including histories in which decisions were made in ways that did not satisfy the corresponding assumptions for shorter histories. There are various unmodeled phenomena that may justify this richness assumption. For example, past decisions might have been constrained in additional ways; or may have been made by different officers; or may have resulted from "trembling hand" deviations from the corresponding f's values.[9] In any event, as is often the case with Arrovian-style impossibility results, the message of the theorem depends on the richness of the domain of f.

There are many examples of organizational decisions that require coordination among different offices within a bureaucracy. For example, different expenses, controlled by different offices, should be assigned to various budgets, where not all expense-budget pairs are allowed. Local police is asked to approve events and demonstrations, which need to be coordinated. Our result is relevant to these problems and to many others that can be modeled by IP.

Real bureaucracies face problems that are more complicated than those modeled in our Theorem 3. Typically, the RHS vectors will vary from period to period, and the new vector, $\mathbf{b}_{t+1}$, may not have been encountered in the past at all. Additional difficulties that arise when problems do not appear in precisely the same way is that the most recent case need not be the most similar one. As the judgment of similarity may be subjective, such situations are prone to generate infeasible solutions. Thus, when facing novel problems (or problems that are sufficiently different from their immediate predecessors), the organization may have to engage in coordinated planning, where different offices' decisions are taken together. In this sense, our result can be viewed as identifying conditions under which different offices of a bureaucracy can use recent history as a coordination device, and tell these apart from conditions in which coordinated planning is necessary.

CRediT authorship contribution statement

Hervé Crès: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Itzhak Gilboa:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Nicolas Vieille:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization.

Declaration of competing interest

We hereby declare that we have no financial or personal relationships with other people or organizations that could have influenced our work.

Data availability

No data was used for the research described in the article.

[9] This is analogous to the definition of a strategy in an extensive form game, which is defined also at nodes that are inconsistent with itself.

Appendix. Proof of Theorem 3

It is immediate that a most-similar-case function is consistent and retains the status quo. We wish to prove the converse.

Let there be given an $m\times n$ matrix A, which contains two potentially conflicting rows. That is, there exist two rows $i_1, i_2 \in \{1,\ldots,n\}$ and three columns $j_1, j_2, j_3 \in \{1,\ldots,m\}$ such that $a_{i_1 j} > 0$ and $a_{i_2 j} < 0$ for $j = j_1, j_2, j_3$. W.l.o.g., we take these rows and columns to be, respectively, the first two rows and the first three columns of A.

For notational simplicity, we denote the first three entries of the first two rows by a, b, and c, and $-a'$, $-b'$, and $-c'$ respectively, so that all six numbers a, b, c, a', b', and c' are positive. We will be interested in the sub-matrix generated by these six numbers. Specifically, for $d, d' \in \mathbb{R}\cup\{-\infty,+\infty\}$ we define the system $S(d,d')$ in the variables $y = (y_1, y_2, y_3)$ as

$$\begin{cases} ay_1 + by_2 + cy_3 & \leq \quad d \\ a'y_1 + b'y_2 + c'y_3 & \geq \quad d'. \end{cases}$$

Thus, for $y \in \mathbb{Z}^3_+$ we say that y is a solution to $S(d,d')$ if the above inequalities hold.

We let $t \geq 1$, and $f : \mathbb{Z}^t_+ \to \mathbb{Z}_+$ be a consistent map that retains the status quo. We will prove that f is a most-similar-case function.

We will use the consistency property only through Lemma 1 below.

Lemma 1. *The map f satisfies the following property. Let $d, d' \in \mathbb{R}\cup\{-\infty,+\infty\}$, and, for $\tau = 1,\ldots,t$, let $y(\tau) \in \mathbb{Z}^3_+$ be a solution to the system $S(d,d')$. Then the (three-dimensional) vector $(f(\mathbf{y}_1), f(\mathbf{y}_2), f(\mathbf{y}_3))$ is also a solution to $S(d,d')$, where $\mathbf{y}_i$ stands for the t-dimensional vector $(y_i(\tau))_{\tau=1,\ldots,t}$.*

Proof. For each $\tau = 1,\ldots,t$, let $x(\tau)$ be the n-dimensional vector obtained by appending $n-3$ zeroes to $y(\tau)$. Set $b_1 = d$, $b_2 = -d'$, and $b_i = +\infty$ for $i > 2$, so that $x(\tau)$ solves $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ for each τ.

Since f retains the status quo, $f(\mathbf{x}_j) = 0$ for $j = 4,\ldots,n$. Since f is consistent, the n-dimensional vector $(f(\mathbf{y}_1), f(\mathbf{y}_2), f(\mathbf{y}_3), 0,\ldots,0)$ solves $A\mathbf{x} \leq \mathbf{b}$ as well. ∎

Given $d, d' \in \mathbb{R}\cup\{-\infty,+\infty\}$, and three t-dimensional vectors α, β, and γ, we will slightly abuse terminology and say that (α,β,γ) *solves* $S(d,d')$ when $(\alpha(\tau),\beta(\tau),\gamma(\tau))$ solves $S(d,d')$ for each τ. Lemma 1 thus says that $(f(\alpha), f(\beta), f(\gamma))$ is a solution to $S(d,d')$ whenever (α,β,γ) solves $S(d,d')$.

The proof proceeds in two steps. In **Step 1**, we prove that f coincides with a most-similar-case function on the set of inputs $\{0,1\}^t$. In **Step 2**, we remove the latter restriction.

Step 1: The restriction of f to $\{0,1\}^t$ is a most-similar-case function.

Throughout **Step 1**, we restrict inputs in $\{0,1\}^t$, and use the following piece of notation. Given a set $B \subset [1;t]$ (where $[1;t] := \{1,\ldots,t\}$), we denote by $\vec{1}_B \in \{0,1\}^t$ the indicator function of B. That is, $\vec{1}_B(\tau) = 1$ iff $\tau \in B$, and we similarly denote by $\vec{0}_B$ the vector defined by $\vec{0}_B(\tau) = 0$ iff $\tau \in B$. We will abbreviate $\vec{0}_{[1;t]}$ and $\vec{1}_{[1;t]}$ to $\vec{0}$ and $\vec{1}$, respectively. Note that $\vec{1}_B = \vec{0}_{\bar{B}}$ where $\bar{B}$ is the complement of B in $[1;t]$. Any vector $\alpha \in \{0,1\}^t$ can be written as $\alpha = \vec{1}_B$, for $B = B_\alpha := \{\,\tau \mid \alpha(\tau) = 1\,\}$. We abuse notation and write $\bar{\alpha}$ to denote $\vec{0}_{B_\alpha} = \vec{1} - \alpha$.

Observe that $f(\vec{0}) = 0$ and $f(\vec{1}) = 1$ because f retains the status quo.

Lemma 2. *For every $\alpha \in \{0,1\}^t$, one has $f(\alpha) \in \{0,1\}$.*

Proof. Assume to the contrary that $f(\alpha) \geq 2$ for some $\alpha \in \{0,1\}^t$. Choose $\mathbf{y}_1 = \alpha$, $\mathbf{y}_2 = \mathbf{y}_3 = \vec{0}$. Note that $(\alpha,0,0)$ solves $S(a,0)$ but $(f(\alpha), f(\vec{0}), f(\vec{0})) = (f(\alpha),0,0)$ does not solve $S(a,0)$, contrary to Lemma 1. ∎

In the rest of Step 1 of the proof, we derive consequences of Lemma 1, with $d := \max(a+b, a+c, b+c)$ and $d' := \min(a',b',c')$. Note that the set of integer-valued solutions of $S(d,d')$ in $\{0,1\}^3$ consists of all vectors in this set, apart from $(0,0,0)$ and $(1,1,1)$.

Lemma 3. *For every $\alpha \in \{0,1\}^t$, one has $f(\bar{\alpha}) = 1 - f(\alpha)$.*

Proof. Assume to the contrary that $f(\alpha) = f(\bar{\alpha}) = \delta \in \{0,1\}$ for some α. Observe that $(\alpha, \bar{\alpha}, \vec{\delta})$ solves $S(d,d')$ (since $(\alpha(t), \bar{\alpha}(t), \vec{\delta}(t))$ is either $(1,0,\delta)$ or $(0,1,\delta)$). Yet, $(f(\alpha), f(\bar{\alpha}), f(\vec{\delta})) = (\delta,\delta,\delta)$ does not solve $S(d,d')$ – a contradiction. ∎

Lemma 4. *f is non-decreasing on $\{0,1\}^t$ (w.r.t. the product order).*

Proof. Assume to the contrary that $f(\alpha) = 1$ and $f(\beta) = 0$ for some $\alpha \leq \beta$. Since $\alpha \leq \beta$, we know that, for every τ, if $\alpha(\tau) = 1$, then $\beta(\tau) = 1$, that is, $\bar{\beta}(\tau) = 0$. This implies that $(\alpha(\tau), \bar{\beta}(\tau), \vec{1})$ solves $S(d,d')$ for each τ. Thus, $(f(\alpha), f(\bar{\beta}), f(\vec{1}))$ solves $S(d,d')$ as well. Yet $f(\bar{\beta}) = 1$ by Lemma 3, so that $(f(\alpha), f(\bar{\beta}), f(\vec{1})) = (1,1,1)$ – a contradiction. ∎

Lemma 5. *Let $B, C \subset [1;t]$ be given. If $f(\vec{1}_B) = 1$ and $f(\vec{1}_C) = 1$, then $B\cap C \neq \emptyset$ and $f(\vec{1}_{B\cap C}) = 1$.*

Proof. Assume to the contrary that $f(\vec{1}_{B\cap C}) = 0$, so that $f(\vec{1}_{\bar{B}\cup\bar{C}}) = 1$. Plainly, $1_{\bar{B}\cup\bar{C}}(\tau) = 0$ as soon as $1_B(\tau) = 1_C(\tau) = 1$ and $1_{\bar{B}\cup\bar{C}}(\tau) = 1$ as soon as $1_B(\tau) = 1_C(\tau) = 0$. Therefore, $(\vec{1}_B, \vec{1}_C, \vec{1}_{\bar{B}\cup\bar{C}})$ solves $S(d,d')$. Yet, $(f(\vec{1}_B), f(\vec{1}_C), f(\vec{1}_{\bar{B}\cup\bar{C}})) = (1,1,1)$ – a contradiction. Hence we obtain $f(\vec{1}_{B\cap C}) = 1$. This implies $B\cap C \neq \emptyset$, since $f(\vec{1}_\emptyset) = f(\vec{0}) = 0$. ∎

By exchanging the roles of zeroes and ones, one gets the following version of Lemma 5.

Lemma 6. *Let $B, C \subset [1;t]$ be given. If $f(\vec{0}_B) = 0$ and $f(\vec{0}_C) = 0$, then $B\cap C \neq \emptyset$ and $f(\vec{0}_{B\cap C}) = 0$.*

Lemma 7. *The restriction of f to $\{0,1\}^t$ is a most-similar-case function.*

Proof. Denote by S_1 the intersection of all sets $B \subseteq [1;t]$ such that $f(\vec{1}_B) = 1$, and by S_0 the intersection of all sets C such that $f(\vec{0}_C) = 0$. Thanks to Lemmas 5 and 6, S_0 and S_1 are non-empty. By Lemma 4, $f(\alpha) = 1$ if and only if $\alpha \geq \vec{1}_{S_1}$ and similarly, $f(\alpha) = 0$ if and only if $\alpha \leq \vec{0}_{S_0}$.

We now prove that the sets S_0 and S_1 coincide and that this common set is a singleton. Pick any element $\tau \in S_0$. Since the inequality $\vec{1}_{\{\tau\}} \leq \vec{0}_{S_0}$ does not hold, one must have $f(\vec{1}_{\{\tau\}}) = 1$ and therefore $\vec{1}_{\{\tau\}} \geq \vec{1}_{S_1}$ and it follows that $S_1 = \{\tau\}$. Since τ was an arbitrary element of S_0, for every $\tau, \tau' \in S_0$ we have $S_1 = \{\tau\} = \{\tau'\}$ and thus $\tau = \tau'$. Hence S_0 is a singleton, $S_0 = \{\tau\}$ and, as we concluded that $S_1 = \{\tau\}$, we also have $S_0 = S_1$.

Note now that, for $\delta \in \{0,1\}$, $f(\alpha) = \delta$ as soon as $\alpha(\tau) = \delta$. That is, $f(\alpha) = \alpha(\tau)$ for every α, as desired. ∎

For clarity, we will henceforth assume that the unique element of $S_0 = S_1$ is $\tau_* = 1$.

Step 2. We now remove the restriction of inputs to $\{0,1\}^t$, and we prove that $f(\alpha) = \alpha(1)$ for every $\alpha \in \mathbb{Z}^t_+$.

We proceed by induction and prove that, for each $k \geq 1$, one has $f(\alpha) = \alpha(1)$ for every $\alpha \in [0;k]^t$. Assume thus that the latter property holds for some $k \geq 1$ (with $k = 1$ proven in Step 1), and let $\alpha \in [0;k+1]^t$ be given. The proof that $f(\alpha) = \alpha(1)$ goes by contradiction.

Assume first that $f(\alpha) > \alpha(1)$. Fix $d' = 0$ so that the second constraint in $S(d,d')$ is satisfied by every $\alpha \in \mathbb{Z}^t_+$. By possibly permuting the first three columns of A, we may assume that $a \leq b, c$. Set $d = a\alpha(1) + (b+c)k$, and observe that

$$d \geq (b+c)k \geq a(k+1) \tag{2}$$

where the last inequality follows from the facts that $b, c \geq a$ and that $k \geq 1$.

Let $\beta, \gamma \in [0;k]^t$ be given by $\beta = \gamma = (k,0,\ldots,0)$. By the choice of d, $(\alpha(1), \beta(1), \gamma(1))$ is a solution to $S(d,d')$. Moreover, for $\tau > 1$, one has

$$a\alpha(\tau) + b\beta(\tau) + c\gamma(\tau) = a\alpha(\tau) \leq a(k+1) \leq d,$$

where the first inequality follows from the fact that $\alpha \in [0; k+1]^t$ (and that $\beta(\tau) = \gamma(\tau) = 0$) and the second – from inequality (2). Thus, $(\alpha(\tau), \beta(\tau), \gamma(\tau))$ is a solution to $\mathcal{S}(d, d')$ for all $\tau \geq 1$. Hence, $(f(\alpha), f(\beta), f(\gamma))$ is a solution to $\mathcal{S}(d, d')$. Yet $f(\beta) = f(\gamma) = k$ by the induction hypothesis, and therefore, $af(\alpha) + bf(\beta) + cf(\gamma) > a\alpha(1) + (b + c)k = d$ – a contradiction.

Assume now that $f(\alpha) < \alpha(1)$. Fix $d = +\infty$ so that the first constraint in $\mathcal{S}(d, d')$ is satisfied by every $\alpha \in \mathbb{Z}_+^t$. By possibly permuting the first three columns of A, we may assume that $a' \leq b', c'$. Set $d' = a'\alpha(1)$. Let $\beta, \gamma \in [0; k]^t$ be given by $\beta = \gamma = (0, k, \ldots, k)$. By the choice of d', $(\alpha(1), \beta(1), \gamma(1))$ is a solution to $\mathcal{S}(d, d')$. As for $\tau > 1$, one has

$$a'\alpha(\tau) + b'\beta(\tau) + c'\gamma(\tau) \geq (b' + c')k \geq 2a'k \geq a'\alpha(1) = d'$$

where the last inequality follows from the fact that $\alpha(1) \leq k + 1 \leq 2k$.

Hence, $(f(\alpha), f(\beta), f(\gamma))$ is a solution to $\mathcal{S}(d, d')$. Yet $f(\beta) = f(\gamma) = 0$ by the induction hypothesis, and therefore, $a'f(\alpha) + b'f(\beta) + c'f(\gamma) = a'f(\alpha) < a'\alpha(1) = d'$ – a contradiction.

References

Akaike, H., 1954. An approximation to the density function. Ann. Inst. Statist. Math. 6, 127–132.

Arrow, K.J., 1950. A difficulty in the concept of social welfare. J. Polit. Econ. 58, 328–346.

Bendor, J., Moe, T.M., 1985. An adaptive model of bureaucratic politics. Am. Polit. Sci. Rev. 79, 755–774.

Buchanan, J.M., Tullock, G., 1962. The Calculus of Consent: Logical Foundations of Constitutional Democracy. The University of Michigan Press, Ann Arbor.

Burns, T., 1961. Micropolitics: Mechanisms of organizational change. Admin. Sci. Q. 6, 257–281.

Burns, T., Stalker, G.M., 1961. The Management of Innovation. Tavistock, London.

Dietrich, F., 2010. The possibility of judgment aggregation on agendas with subjunctive implications. J. Econom. Theory 145, 603–638.

Dietrich, F., List, C., 2010. Majority voting on restricted domains. J. Econom. Theory 145, 512–543.

Dietrich, F., Mongin, Ph., 2010. The premiss-based approach to judgment aggregation. J. Econom. Theory 145, 562–582.

Dokow, E., Holzman, R., 2010a. Aggregation of binary evaluations. J. Econom. Theory 145, 495–511.

Dokow, E., Holzman, R., 2010b. Aggregation of binary evaluations with abstentions. J. Econom. Theory 145, 544–561.

Durkheim, E., 1893. The Division of Labor in Society. Free Press, Glencoe, Il, English translation, 1949.

Fayol, H., 1919. General and Industrial Management. Pitman, London, English translation, 1949.

Fix, E., Hodges, J., 1951. Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties. Technical Report 4, Project Number 21-49-004, USAF School of Aviation Medicine, Randolph Field, TX.

Fix, E., Hodges, J., 1952. Discriminatory Analysis: Small Sample Performance. Technical Report 21-49-004, USAF School of Aviation Medicine, Randolph Field, TX.

Follett, M.P., 1918. The New State: Group Organization and the Solution of Popular Government, 1923 ed. Longman, Green, and Co, New York.

Gary, M., Johnson, D.S., 1979. Computers and Intractability: A Guide to the Theory of NP-Completeness. W. Freeman and Co, San-Francisco, CA.

Gilboa, I., Schmeidler, D., 2011. Authorization Decisions. Working Paper.

Gittins, J.C., 1979. Bandit processes and dynamic allocation indices. J. R. Stat. Soc. Ser. B 41, 148–177.

Handel, M.J., 2003. The Sociology of Organizations. Sage Publications, Thousand Oaks, CA.

Hatch, M.J., Cunliffe, A.L., 2006. Organization Theory, second ed. Oxford University Press, Oxford.

Kanter, R.M., 1977. Men and Women of the Corporation. Basic Books, New York.

List, C., Pettit, P., 2002. Aggregating sets of judgments: An impossibility result. Econom. Philos. 18, 89–110.

List, C., Polak, B., 2010. Introduction to judgment aggregation. J. Econom. Theory 145, 441–466.

March, J.G., 1962. The business firm as a political coalition. J. Politics 24, 662–678.

March, J.G., 1999. The Pursuit of Organizational Intelligence. Blackwell, Malden, MA.

March, J.G., Simon, H.A., 1958. Organizations. John Wiley and Sons, New York.

Marx, K., 1867. Capital. Foreign Languages Publishing House, Moscow, English Translation, 1954.

Milgrom, P., Roberts, J., 1992. Economics, Organization, and Management. Prentice Hall, New Jersey.

Morgan, G., 2006. Images of Organization. Sage Publications, Thousand Oaks, CA.

Nehring, K., Puppe, C., 2010a. Abstract arrowian aggregation. J. Econom. Theory 145, 467–494.

Nehring, K., Puppe, C., 2010b. Justifiable group choice. J. Econom. Theory 145, 583–602.

Niskanen, W., 1971. Bureaucracy and Representative Government. Aldine Transaction, New Brunswick.

Niskanen, W., 1975. Bureaucrats and politicians. J. Law Econom. 18, 617–644.

Ouchi, W., Wilkins, A., 1985. Organizational culture. Annu. Rev. Sociol. 11, 457–483.

Parsons, T., 1951. The Social System. The Free Press, New York.

Parzen, E., 1962. On the estimation of a probability density function and the mode. Ann. Math. Stat. 33, 1065–1076.

Rubinstein, A., Fishburn, P.C., 1986. Algebraic aggregation theory. J. Econom. Theory 38, 63–77.

Sandelands, L.E., Stablein, R.E., 1987. The Concept of Organization Mind, Research in the Sociology of Organizations, Vol. 5. JAI Press, Greenwich, CT.

Scott, W.R., Davis, G.F., 2007. Organizations and Organizing: Rational, Natural, and Open System Perspectives. Pearson Education, Upper Saddle River, NJ.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London and New York.

Simon, H.A., 1947. Administrative Behavior. McMillan, New York.

Smith, A., 1776. The Wealth of Nations, 1957 ed. Appleton Century Crofts, New York.

Taylor, F.W., 1911. The Principles of Scientific Management. Harper, New York.

Vieille, N., 2006. A Generalized Aggregation Impossibility Result. Mimeo.

Walsh, J.P., Ungson, G.R., 1991. Organizational memory. Acad. Manag. Rev. 16, 57–91.

Weber, M., 1921. Economy and Society: An Outline of Interpretive Sociology. University of California Press, Berkeley, CA, English translation, 1968.

Weber, M., 1924. The Theory of Social and Economic Organization. Free Press, Glencoe, Il, English translation, 1947.

Wilson, R., 1972. Social choice theory without the Pareto principle. J. Econom. Theory 5, 478–486.