Process tracing change management: the reform of the Italian judiciary

Simone Busetti
Department of Management Engineering,
Politecnico di Milano Dipartimento di Ingegneria Gestionale, Milano, Italy, and
Giancarlo Vecchi
Department of Management,
Economics and Industrial Engineering – Policy Analysis Team,
Politecnico di Milano Dipartimento di Ingegneria Gestionale, Milano, Italy

Abstract
Purpose – In 2009, the Italian Government initiated a national programme to improve the management of judicial offices. Programme implementation has been patchy and unsatisfactory in all but a few cases. Against this background, the Law Court of Milan has achieved exceptional results and is now recognised as a good practice benchmark for Italy. The purpose of this paper is to investigate this case in order to reconstruct the local conditions for successful implementation of the national programme.

Design/methodology/approach – To test a theory of the programme based on leaders’ engagement, their access to managerial knowledge, and the transfer and consolidation of that knowledge, the present study applies process tracing, a qualitative method that uses Bayesian reasoning to improve the accuracy of within-case inferences.

Findings – The analysis shows how programme and context features interacted to support change. In particular, while the national programme succeeded in providing resources for leader engagement and knowledge access, the transfer and consolidation of managerial knowledge depended largely on a brokerage function performed locally between consultants and magistrates.

Originality/value – The paper sheds light on the local conditions for change management and does so by employing an innovative qualitative method that improves the reliability of within-case inferences.

Keywords Change management, Judiciary, Process tracing, Public sector modernization

1. Introduction
By international standards, the Italian judicial system is performing poorly (OECD, 2013a; Esposito et al., 2014; Ministero della Giustizia (Italian Ministry of Justice), 2015). Scholars and practitioners have advanced a number of possible reasons for this critical condition, such as the high level of litigiousness in the Italian society (Buonanno and Galizzi, 2014; CEPEJ, 2014); the limited availability of alternative ways of resolving disputes (Castelli, 2013); and tensions between political and judicial powers (Guarnieri, 2015). In addition, however, deficiencies in the management of Italian judicial offices are also blamed; despite long-running debate, efforts to modernise and reform managerial practices began only in the late 2000s (Steelman and Fabri, 2008; Vecchi, 2013a).

The present paper investigates the implementation of a national programme for improving the management of judicial offices and describes an in-depth case study of the Law Court of the Milan District (LCM). Amid patchy implementation of this programme, Milan stands out as a success story and therefore as a possible source of insight into how modern management tools can be introduced to judicial offices.

For present purposes, we adopt the perspective of programme designers and policy makers at the national level. Confronted with the poor implementation results across the country, a prospective designer might look to the LCM in envisaging a more
effective intervention. Learning from the LCM success requires an understanding of whether and how the programme helped to initiate change in the LCM and – given the likely influence of local factors – how these can be integrated into a new national programme and reproduced in different contexts.

To begin, the analysis develops a three-step theory of how the programme may have contributed to the LCM success by engaging leaders, providing previously inaccessible knowledge, and securing knowledge transfer. This theory is tested using process tracing (PT) (Beach and Pedersen, 2013; Bennett and Checkel, 2015), which is an innovative method for making within-case inferences from a single case study. PT is one of the most promising developments to emerge from recent debate concerning qualitative methodology (see Brady and Collier, 2010), and the public management context represents fertile ground for the deployment of this method.

The paper proceeds as follows. Section 2 provides an overview of both the national programme and the LCM case. Section 3 introduces the method of PT, and Section 4 elaborates the three-step programme theory referred to above. Based on PT prescriptions, Section 5 discusses the prior likelihood of the hypothesis and sets out the relevant evidence. Finally, Sections 6-8 describe the case material and discuss the probative value of the evidence collected.

2. Changing the Italian judiciary

Between 2007 and 2008, the European Commission allocated EU structural funds to finance a reform programme for the Italian judicial system. During this period, Italy’s Ministry of Justice, Ministry of Public Administration, and regional governments designed a reform programme called “Diffusion of Best Practices in the Italian Judicial Offices” (BPJO).

BPJO provided funding for technical assistance to enhance the quality of civil and penal justice attained in selected Prosecutors’ Offices, Ordinary Law Courts, and Courts of Appeal by improving efficiency, external accountability, and communication skills under the innovative frame of “justice as public service”. Regional governments would receive funding to hire consultants to collaborate with internal personnel in the design of new management tools, mainly to optimise procedures and improve relations with users by means of ICT solutions. With total funding of almost 40 million euros, 191 judicial offices (about 30 per cent of the total) were invited to participate in the programme, and activities commenced in 2009.

As the judicial sector had always been exempt from previous reforms targeting public administration and therefore lagged far behind most of the Italian public sector, BPJO represented a major opportunity. However, the first evaluation reports were discouraging, as about 70 per cent of the participating offices delivered insufficient or deficient outputs (see www.risorseperlagiustizia.it; Vecchi, 2013a, b; Xilo, 2014). The heterogeneity of results owed in part to weak national coordination, as a political crisis severely undermined the national steering bodies just one year after the programme commenced, leaving regional governments and local judicial offices with no overarching guidance. In addition, because BPJO was structured as a national framework that provided resources and set general goals (such as improved use of ICT or enhanced communication with the public), it was left to individual offices to define and implement projects congruent with that frame. For a sector with no experience of driving managerial innovation, weaknesses of coordination at national level may have proved particularly harmful.

Against a background of widespread failure, the LCM case is a clear outlier, even when compared to good performers. In fact, Milan achieved almost double the average number of interventions and led the way in experimenting with new tools and procedures. Unlike most other offices, Milan implemented projects across a wide range of organisational areas, entailing a profound reorganisation of internal working arrangements such as the distinction between back-office and front-office processes, along with the adoption of a
A strong “service” approach (e.g. implementation of a dedicated front-office for users), and the introduction of some near-revolutionary IT solutions (e.g. digitisation of civil trials procedures, algorithmic allocation of case files to magistrates). Finally, and uniquely, the LCM agreed several projects in conjunction with connected judicial offices, embodying a comprehensive vision of workflows and placing a great emphasis on the user service experience. Table I lists all projects implemented by the LCM.

Overall, the LCM evaluations returned positive results, especially regarding the performance of revised managerial tools. Between 2011 and 2015, despite staff reductions (−10 per cent of magistrates; −30 per cent of administrative staff), the LCM managed to significantly reduce its backlog. Proceedings older than three years were reduced from 22 to 17.5 per cent of the total backlog, which was cut by 12 per cent. Finally, the average duration of civil proceedings was reduced by 10 per cent (see Law Court of Milan, 2013; Tribunale Ordinario di Milano, 2016; OECD, 2013b).

The transfer to other offices of some LCM project experiments (in particular, the software algorithm for allocating case files) is further evidence of the relevance and quality of these interventions. As the LCM is now seen as a national reference for modernisation of the judiciary (OECD, 2013b), the sections that follow investigate that success in an attempt to specify the respective contributions of BPJO and local factors.

3. The method: introducing PT

In the management sciences, case studies are widely used for theory development and have produced some inspiring works (e.g. Eisenhardt, 1989; Eisenhardt and Graebner, 2007).

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Relevancea</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Improve the efficiency of administrative services</td>
<td>Medium</td>
</tr>
<tr>
<td>2</td>
<td>Re-engineer budget management processes (justice expenses)</td>
<td>High</td>
</tr>
<tr>
<td>3</td>
<td>Re-engineer human resource management/administrative staff</td>
<td>High</td>
</tr>
<tr>
<td>4</td>
<td>Re-engineer judicial staff offices (civil and penal sector) and design and implement new front-office procedures</td>
<td>High</td>
</tr>
<tr>
<td>5</td>
<td>Re-engineer notification procedures</td>
<td>Medium</td>
</tr>
<tr>
<td>6</td>
<td>Improve quality of services and new back-office procedures for witnesses</td>
<td>High</td>
</tr>
<tr>
<td>7</td>
<td>Re-engineer management of judicial consultants</td>
<td>Medium</td>
</tr>
<tr>
<td>8</td>
<td>Re-engineer management of judicial court files</td>
<td>Medium</td>
</tr>
<tr>
<td>9</td>
<td>Re-engineer management of human resources/judges</td>
<td>High</td>
</tr>
<tr>
<td>10</td>
<td>Use CAF model for organisational self-assessment</td>
<td>Low</td>
</tr>
<tr>
<td>11</td>
<td>Re-engineer voluntary jurisdiction administrative offices</td>
<td>High</td>
</tr>
<tr>
<td>12</td>
<td>Design and implement new Office for Relations with the Public (in collaboration with the Court of Appeal and the Office of the Public Prosecutor)</td>
<td>High</td>
</tr>
<tr>
<td>13</td>
<td>Conduct feasibility study to transfer parts of the offices in a new building (in collaboration with the Court of Appeal and the Office of the Public Prosecutor)</td>
<td>Medium</td>
</tr>
<tr>
<td>14</td>
<td>Implement design and organisational development of the Innovation office</td>
<td>High</td>
</tr>
<tr>
<td>15</td>
<td>Implement new performance management system</td>
<td>Medium</td>
</tr>
<tr>
<td>16</td>
<td>Conduct audit of hardware, software, and platform for ICT continuous monitoring</td>
<td>Medium</td>
</tr>
<tr>
<td>17</td>
<td>Install new software and procedures to support hearing planning</td>
<td>High</td>
</tr>
<tr>
<td>18</td>
<td>Install new Intranet system (in collaboration with the Court of Appeal and the Office of the Public Prosecutor)</td>
<td>Medium</td>
</tr>
<tr>
<td>19</td>
<td>Deploy new ICT systems and restyle the website in conjunction with the special plan of the Ministry of Justice</td>
<td>High</td>
</tr>
<tr>
<td>20</td>
<td>Conduct feasibility study of a new database for penal sector sentences</td>
<td>Medium</td>
</tr>
<tr>
<td>21</td>
<td>Enact guidelines for Voluntary Jurisdiction Sector services</td>
<td>High</td>
</tr>
</tbody>
</table>

Table I.

Projects implemented in Milan

Note: aBased on interviews with LCM leaders
However, the lack of any clear rule (such as “results are significant at $p < 0.05$”) leads to ambiguity in appraising case study evidence and associated inferences (Siggelkow, 2007). In the change management literature, however if the data is rich on the factors underpinning successful change, scholars often fail “in specifically demonstrating if and how these factors actually affected the change processes” (Kuipers et al., 2014, p. 15; see also Fernandez and Pitts, 2007). As one promising means of improving rigour and transparency in collecting qualitative evidence, PT may prove a useful addition to the methodological toolkit of management scholars. PT is a qualitative method for making within-case inferences in single cases (Beach and Pedersen, 2013; Bennett and Checkel, 2015). Although there is no standard way of conducting a PT study, the method can be condensed into two basic components relating to theory building and theory testing. With regard to the former, PT owes its name to a focus on unpacking causal processes – that is, explicating why and how a cause and an outcome are related. In this respect, PT resonates with theory-based approaches to evaluation (Astbury and Leeuw, 2010; Pawson, 2013), as well as the study of mechanisms for the analysis and transfer of good practices (Bardach, 2004; Barzelay, 2007; Busetti and Dente, 2018) and the broader methodological debate on opening the black box of causality (Brady and Collier, 2010).

As an approach to theory testing, PT combines Bayesian-inspired data collection, evidence assessment, and theory updating, providing novel tools for small-n or unique interventions. To begin, researchers must assess prior confidence in their hypothesis ($H$), based on preliminary evidence supporting $H$ (e.g. results of previous studies), critically discussing their expectations to define an initial direction for evidence collection and assessment. The second step is to evaluate the impact of the evidence on posterior confidence in $H$. After mapping the evidence in support of the hypothesis, the probative value of that evidence must be assessed by considering true and false positive rates. The true positive rate (also called certainty or necessity) represents the probability of finding the evidence conditional to $H$ being true. A high true positive rate means that the evidence should be there if $H$ is true; a low rate means that finding the evidence does not necessarily confirm $H$. Conversely, the false positive rate is the probability of finding the evidence conditional to the probability of $H$ being false. A low false positive rate indicates that the evidence is highly specific (or unique, or sufficient) to $H$. A high false positive rate means that many alternative hypotheses might explain that same piece of evidence.

As one typical example from PT textbooks, being in the vicinity on the day of a murder has both a high true positive value and a high false positive value. This kind of evidence (the so-called hoop test) has a limited confirmatory power if found and a telling disconfirmatory effect when not found. Conversely, being caught with a smoking gun close to the crime scene is an example of evidence with a negligible false positive rate but a low true positive rate. Finding such evidence (the smoking gun test) greatly increases confidence in $H$ (i.e. it has great confirmatory power) while not finding it has only limited disconfirmatory value (given the low true positive, $H$ can be confirmed even in the absence of such evidence – in other words, one can be charged even when not found with a smoking gun). Table II summarises the different types of test and their impacts on the likelihood of $H$[1].

<table>
<thead>
<tr>
<th>Type of test</th>
<th>True+</th>
<th>False+</th>
<th>Empirical results</th>
<th>Impact on $H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking gun</td>
<td>Low</td>
<td>Low</td>
<td>Passed</td>
<td>+++</td>
</tr>
<tr>
<td>Hoop test</td>
<td>High</td>
<td>High</td>
<td>Passed</td>
<td>=</td>
</tr>
<tr>
<td>Doubly decisive</td>
<td>High</td>
<td>Low</td>
<td>Passed</td>
<td>+++</td>
</tr>
</tbody>
</table>

Table II. Classification of evidentiary tests
Combining priors with true and false positive rates provides a powerful and transparent tool for guiding the search for evidence and assessing the probative value of the collected evidence. Note that although this process increases the reliability of within-case inference (and therefore internal validity), the generalisability of the results remains limited (i.e. external validity is low). In this respect, the method is especially useful for studying critical or special cases (such as success) to understand the factors and processes underlying those outcomes.

4. A theory of programme success

Granted that implementing reforms is no easy task, the management literature offers extensive advice about how to conduct a process of change, addressing such issues as how to win over dissenters, how to optimise the speed and scale of change, and how to consolidate outcomes (Lewin, 1947; Kotter, 1996, 2014; Armenakis and Bedeian, 1999; Tdhem By, 2005). Within this, a significant body of literature addresses the characteristics of managerial change in public organisations (see, for instance, Fernandez and Rainey, 2006; Rainey and Bozeman, 2000). Most adopt the perspective of change agents within organisations and specify strategies and implementation tactics to support innovative change (Cels et al., 2012; Rainey, 2009; Behn, 1991). From this perspective, committed leaders are the heroes of all stories of successful change management (Hargrove, 1994; Boin and Christensen, 2008) and are commonly seen to mobilise the factors that determine successful change (see for instance, Kotter, 1996). However, when dealing with professionally dominated administrations in the public sector, the literature shows how the technical knowledge and organisational power of the professionals who dominate the public sector can play a significant role in resisting managerial reforms (Ferlie and Geraghty, 2005). These organisations commonly lag behind in terms of managerial practices and capacity to implement change. They are further hindered by leaders who, rather than driving the reform process, are often reluctant, disinterested, or even openly resistant to change.

As outlined above, the national BPJO programme provided resources for technical assistance, framed by a set of reform objectives that included the improved use of ICT technologies, reengineering of workflows, and a service approach based on enhanced communication and attention to users. However, the national programme was conceived as a trigger, leaving responsibility for defining and implementing specific actions to local level, with regions and individual offices hiring consultants to develop their individual paths to change. Given this approach, it may be hypothesised that the programme’s contribution depended on the engagement of local leaders when offered an opportunity to initiate the process, the access it provided to requisite managerial knowledge by targeting actual deficiencies in judicial offices, and whether and how it contributed to consolidating and transferring that knowledge by adjusting it to local realities.

These three potential causal contributions are explained in depth below as components of a programme theory of how BPJO may have worked. While the first deals mainly with how change is initiated, the second and third concern the effective design of the programme as a technical assistance-based intervention. Following PT standards, the three-step programme theory is represented diagrammatically in Figure 1.

![Figure 1. Hypothesised programme theory of BPJO](image-url)
A – Leaders’ engagement
The first step relates to how leaders engage in the process of change (Stoker, 1991). As noted above, there is an assumption that change is costly and risky (Brown and Osborne, 2013), and that leaders are not necessarily benevolent innovators who are equipped and ready for change (Kelman, 2005; Pawson, 2013). From the perspective of national policy makers, the central question concerns how the programme might provide an opportunity to trigger change, transforming leaders into “programme entrepreneurs” (Kingdon, 2011; McAdam et al., 2003). In line with suggestions present in the literature, BPJO may have supported engagement in several ways: by creating a sense of urgency (Kotter, 1996), working as a focusing event (Abramson and Lawrence, 2001; Birkland, 2006), providing a vision (Kotter, 2014), supplying resources (Boyne, 2003), overcoming resistance (Judson, 1991), or helping to remove other barriers to change (Kotter, 2014).

B – Knowledge access
This step investigates how leaders access managerial knowledge that was previously unavailable to the organisation. To this end, organisational leaders must connect with providers of managerial knowledge, creating a “weak tie” with an outside network (Granovetter, 1983; Considine et al., 2009). The goal of BPJO was to provide support through consultants trained in both public management and IT technologies for administrative modernisation. Here, it becomes necessary to verify whether the consultants actually provided that knowledge, whether sources other than the BPJO network provided important knowledge, and whether organisational leaders might have had other opportunities to access the requisite knowledge.

C – Knowledge transfer
Knowledge access does not inevitably imply successful knowledge transfer. Scholars working on innovations and knowledge diffusion claim that knowledge translation often remains a challenging task once new knowledge has been accessed (Osborne and Brown, 2005; Greenhalgh et al., 2013), all the more so when such knowledge is complex or alien to the target population (Hansen, 1999). In this respect, not only must access be provided to new knowledge but there must also be a stable and collaborative connection with the knowledge providers (Klijn and Koppenjan, 2015). Frequent contact can help knowledge transfer by triggering positive feedback (Schön, 1983; Kolb et al., 2011; Argote, 2013); improving organisational learning (Gilson et al., 2009) and knowledge absorption (Cohen and Levinthal, 1990); and providing the requisite resources for translation between consultants and organisational leaders (Weick and Quinn, 1999). In the case of BPJO, technical assistance was conceived as an extensive collaboration with local implementation agents (a one-year service on average) in order to develop both tailored projects and local capacity for future unassisted project development. The relatively long duration of this service and the need to draft tailored projects prompted ongoing work at local level and may have facilitated successful knowledge transfer by providing for repeated contact opportunities and collaboration, monitoring of experimental pilot projects, and general development of trust.

5. Assessing prior confidence and mapping the evidence
Before proceeding to any empirical work, PT requires an assessment of prior confidence in the hypothesis. The three-step programme theory described above is congruent with the programme characteristics and the existing literature. In assessing prior likelihood, it is also useful to take account of any preliminary information about the programme, results of comparable cases, and features of the local context.
As noted above, the BPJO structure was relatively weak, as it delegated most interventions to local level and provided only general guidelines for implementation. Additionally, the BPJO central taskforce was dismantled following the fall of the national government in 2011. These circumstances may have had a major impact, especially in a sector (the judiciary) with no previous experience of managerial reforms. In addition, the high level of failures across the country makes it more likely that local factors played a major role in the success of the change process at the LCM. This proposition is reinforced by Milan’s suitability for a programme of administrative modernisation. As one of the main judicial offices in Italy, the LCM is at the centre of important economic, political, and judicial processes. In this respect, it is reasonable to suppose that this office would be equipped with special resources, competences, and absorptive capacities.

In sum, given that programme success is generally rare across the population, that the programme design was weak, and that the case may be atypical, prior confidence in any strong causal contribution from BPJO seems likely to be quite low. In this respect, although the programme may have facilitated leaders’ engagement, knowledge access, and knowledge transfer as hypothesised above, the LCM results may well have been heavily dependent on local factors. Low prior confidence means that it is important to remain alert to causal factors other than the programme, or at least to consider how BPJO may have combined with local causal factors. Low confidence also means that the empirical material must provide a great deal of confirming evidence if the hypothesis is to find support.

Having formulated a hypothesis and assessed prior confidence, PT recommends that possible evidentiary material should be mapped, and its true and false positive rates should be discussed. This latter discussion was completed prior to the fieldwork, but for the sake of presentation, it is discussed in the empirical sections along with the findings. Figure 2 maps the relevant empirical tests for each step in the hypothesis.

The case study benefited from diverse sources of evidence that included semi-structured interviews, direct observation of meetings, and analysis of documents and meeting minutes. Details of the interviews are reported in Table AI. The informants included both programme designers and coordinators at the national level, as well as all top implementation actors in the LCM. The interview guidelines were based on the empirical tests mapped in Figure 2.

Figure 2.
Mapping the evidentiary material
6. The programme engages organisational leaders

This step sought to verify whether and how the programme triggered leaders’ motivation, supported existing motivation, and helped reformers to commit to organisational change.

A first test considers whether LCM leaders had commenced other reform efforts before applying to BPJO (a1). Note that, to prove the programme’s success, BPJO need not be the sole motivator in the LCM’s history (i.e. the test has a low true positive rate). Conversely, it is highly unlikely that the programme would have had no influence on actors’ engagement if the first reforms coincided with BPJO (i.e. low false positive).

In this case, the evidence is mixed. In fact, the president of the LCM had commenced a number of initiatives prior to BPJO implementation. One major innovation was the experimental introduction of the so-called “Processo Civile Telematico” (PCT) (the “digital trial”, involving an online connection between the Court, the barristers/lawyers, and the accused), which was subsequently expanded and diffused across all of Italy’s judicial offices. Another (failed) initiative was the attempt to establish the so-called “Ufficio per il Processo” (Trial Office), which was to include trainers or temporary staff to support individual magistrates in their tasks. However, these projects were not only less comprehensive than BPJO but commenced when the programme was already being drafted at the central level. Some interviewees reported that the national commitment to BPJO created a favourable perception of the environment that helped reformers at local level in anticipating change.

Another relevant test would be whether, when implementing past reforms, leaders failed to maintain engagement or encountered implementation difficulties that were resolved once BPJO was underway (a2). As compared to relying only on local resources, accessing BPJO needs to have made a difference for the change agents (i.e. the evidence is essential or has a high true positive). Interestingly, the test has a low false positive, as it is hard to imagine other causal factors parallel to the programme (making it a doubly decisive test). The evidence confirms the test, so providing strong support for Step A.

The interviewees reported that projects that commenced prior to BPJO suffered a number of implementation difficulties. On the one hand, funds were insufficient for anything more than pilot projects. The PCT (the major project implemented before BPJO) gained strong momentum from the programme, with additional funds provided to fine-tune the system by means of an improved IT infrastructure, several extensions, and a dedicated taskforce. More generally, interviewees noted that previous interventions were fragmented and had no clear horizon or consistency of guidance. In sum, regardless of previous efforts, BPJO provided a “stable framework for change” as a platform that supported reformers in implementing both brand new and previously stalled interventions.

A third test examined LCM leaders’ assertions about the importance of BPJO in triggering and supporting their engagement (a3). No testimony is unbiased, and in this case, one would expect LCM leaders to overstate their independent motivation, so downplaying the effect of the programme. Conversely, declarations of the programme’s causal power may be especially significant. In PT terms, this means that such testimonies have a relatively low true positive (i.e. they are not certain) and a low false positive (i.e. they are highly specific).

The case evidence meets the test requirements. Leaders testified that the programme was highly relevant for the organisation and the management of all initiatives. All interviewees agreed that a full-blown strategy for change could have not been planned without BPJO; the programme opened a window of opportunity for willing reformers and served as the backbone of the modernisation process by providing both funding and knowledge. In this respect, interviewees emphasised that existing connections and reform efforts could not have produced results without the momentum generated by the programme. BPJO helped to overcome resistances and to engage the whole administration.

A further test was to explore whether the programme may have generated individual gains that triggered instrumental engagement (a4). Such evidence is not needed to support
the hypothesis and is unlikely to be found (i.e. it has a low true positive value, as it is hard to imagine leaders admitting that they pursued individual gains). However, evidence of that kind would provide important confirmation of the programme’s ability to trigger engagement (i.e. low false positive). In that respect, the case analysis yielded mixed results. As expected, managerial leaders did not explicitly support any suggestion of individual gains from the programme. However, interviews and meeting minutes indicate that reforming the administration and promoting the LCM as a leading national player in judicial modernisation was considered important both for internal promotion and for national visibility. The use of the Social Responsibility Budget (part of BPJO projects) for external promotion of the LCM is a good case in point. In sum, although the programme had no explicit in-built mechanisms for individual gain, it may have served the instrumental values of some programme leaders.

If no factor other than the programme emerged as triggering engagement, this would represent one final support for Step A (a5). No programme is a magic wand, and programme effects are typically complemented by other factors. In other words, one should not necessarily expect to find that evidence in the fieldwork (i.e. low true positive), but that would be significant for the hypothesis (i.e. low false positive).

Remarkably, the test was partly confirmed, as external factors were not considered particularly relevant. At the same time, the arrival in 2007 – two years before BPJO implementation – of the new president and the former director of the Ministry of Justice prompted an increased propensity to innovate. These factors serve to reiterate that the BPJO opportunity was seized by capable leaders who were willing to exploit it.

Table III summarises the empirical results and their impact on confidence in Step A. In sum, the effect of BPJO is confirmed, although qualified by the importance of a favourable organisational context, in which leaders were already striving to reform the office.

As expected, previous attempts to modernise the administration were plagued by several limitations, and LCM leaders lacked key resources for progressing the necessary reforms. Congruent with the hypothesis, as well as resources, BPJO provided a vision – a stable framework for change and a sense of urgency that established a platform for envisaging wider reform beyond the fragmented efforts of the past. This framing facilitated an expansion of previous efforts, implementation of already planned interventions, enlargement of the reform coalition within the organisation, and victory against resistance.

### 7. Knowledge access

Step B tested how leaders accessed new knowledge by verifying whether BPJO prompted leaders to engage with a formerly unconnected network, and whether this led to the provision of knowledge that was new to the organisation.

One possible test would be to verify that the organisation had no previous connection with potential knowledge holders and, in particular, no connection of an intensity comparable to that provided by BPJO (b1). If the LCM had access to equivalent knowledge

<table>
<thead>
<tr>
<th>A – The programme engages organisational leaders</th>
<th>True +</th>
<th>False +</th>
<th>Empirical results</th>
<th>Impact on A</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1 – No previous efforts at reform</td>
<td>Low</td>
<td>Low</td>
<td>Failed</td>
<td>=</td>
</tr>
<tr>
<td>a2 – Deficient engagement or problematic implementation of previous reforms</td>
<td>High</td>
<td>Low</td>
<td>Passed</td>
<td>+++</td>
</tr>
<tr>
<td>a3 – BPJO considered relevant by LCM leaders</td>
<td>Low</td>
<td>Low</td>
<td>Passed</td>
<td>+++</td>
</tr>
<tr>
<td>a4 – BPJO provided individual incentives to LCM leaders</td>
<td>Low</td>
<td>Low</td>
<td>Partly passed</td>
<td>+</td>
</tr>
<tr>
<td>a5 – No factor other than BPJO mentioned by managerial leaders</td>
<td>Low</td>
<td>Low</td>
<td>Partly passed</td>
<td>+</td>
</tr>
</tbody>
</table>

Table III. Assessing the evidence for Step A
holders in the absence of the programme, BPJO may not have added much to the existing capacity; if the programme enhanced knowledge access, one would expect to find evidence supporting the test (i.e. high true positive). At the same time, if the Court had no such connection before the programme, it would be hard to think of alternative explanations for the newly established connection (i.e. low false positive). Low false positive and high true positive means that the test is doubly decisive for Step B. Interestingly, the evidence provides support for the test.

It also seems relevant here that judges possess independent power, with self-government and formal membership. Their isolation, both from other parts of the administration and from external actors, is long-standing and further ensures their independence. In this respect, their routine network is typically confined to their members, other courts, and other judicial institutions at national level. The LCM has limited relations with local barristers/lawyers’ associations and other local organisations. Other public administrations in Milan, such as the Municipality and the Chamber of Commerce, although actively innovating within, lacked the capacity and opportunity to transfer their knowledge to the LCM. Other courts with links to the LCM were less developed and could provide no help in the modernisation process.

The LCM had some connections with managerial and IT experts; in particular, one consultant to the court had some managerial background, and some magistrates had taken summer courses on management run by a group of scholars from the University of Bologna. However, these were mostly individual contacts that would not be capable of triggering large-scale organisational change.

A follow-up test for b1 might ask whether managerial leaders supported the idea that, without BPJO, the consultants’ knowledge would have not been accessed (b2). This test risks the same bias as in test a3 above, as LCM leaders might be expected to resist any portrayal of change as unduly dependent on the programme (i.e. low true positive). However, supporting statements would be particularly significant for increasing confidence in B (low false positive).

All interviewees (with no exceptions among LCM leaders) agreed that there would have been no opportunity to access that knowledge without BPJO. While some contacts with other experts existed before BPJO, these were individual connections that lacked the necessary strength and consistency and did not encompass the range of professional profiles brought together by BPJO.

In sum, the evidence collected is confirmatory. In fact, no link existed before BPJO, nor could any have been established by other means. The programme is therefore acknowledged to be a valuable (and otherwise unavailable) source of knowledge. Table IV summarises the probative value of the two empirical tests, their results, and the impact on the likelihood of Step B.

### 8. Knowledge transfer

As mentioned earlier, knowledge transfer is a key element for an administration like the judicial sector, which is particularly unfamiliar with managerial practice. This final

<table>
<thead>
<tr>
<th>B – Knowledge access</th>
<th>True</th>
<th>False</th>
<th>Empirical results</th>
<th>Impact on B</th>
</tr>
</thead>
<tbody>
<tr>
<td>b1 – Little or no connection with knowledge holders before the programme</td>
<td>High</td>
<td>Low</td>
<td>Passed</td>
<td>+++</td>
</tr>
<tr>
<td>b2 – Organisational leaders agree that access to BPJO knowledge would have been</td>
<td>Low</td>
<td>Low</td>
<td>Passed</td>
<td>+++</td>
</tr>
<tr>
<td>difficult without the programme</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Assessing the evidence for Step B
step serves to illuminate how knowledge was transferred to the administration. The evidence should support the existence of specific in-built programme mechanisms for knowledge transfer – for instance, by providing ongoing opportunities for exchanging knowledge. The alternative hypothesis is that the transfer depended on other resources extraneous to the programme.

For Step C, a first test was to verify as a background condition that a learning process was needed for effective use of the knowledge accessed through the programme (c1). This test is important as a means of proving that the knowledge to be transferred was complex and that dedicated mechanisms for knowledge transfer were needed (i.e. high true positive). Nonetheless, evidence of a learning process – even if assisted by the programme – is not in itself sufficient to conclude that BPJO helped knowledge transfer (i.e. high false positive).

The evidence confirms the test. Managerial innovation for judicial offices represented a new departure – not only for an administration unfamiliar with such processes but also for the consultants, who mostly had no experience of judicial offices. Interestingly, the consultants testified that the first phase was intense in terms of learning how a court worked and how managerial and IT tools could be adapted. This meant that mutual learning was in part required because of an initial lack of knowledge on both sides. However, the programme assisted this learning by setting up dedicated task forces, where consultants and officials at different levels met periodically.

For a deeper understanding of the programme’s effect, it was useful to examine the contact opportunities offered by BPJO after the opening phase (c2). Mechanisms for knowledge adjustment and transfer can range from simple fine-tuning of contracted interventions to the co-design of new projects. To support the hypothesis, the evidence should prove that the transfer was complex and was tackled explicitly by the programme (i.e. high true positive). However, the relevance of the programme does not exclude the contributing (or indeed predominant) impact of other factors (i.e. high false positive).

The test finds supporting evidence, as meeting minutes and direct observations confirm that there were continuous contacts between the consultants and the target organisation. As the programme extended over several years, there were many opportunities to fine-tune and adjust interventions. The BPJO programme design as initially specified at the central level did not include explicit mechanisms for knowledge transfer. In fact, the design was light and allowed for great flexibility at the local level. After winning the bid, the consultants began an intense dialogue with LCM leaders to redesign and adjust interventions. To this end, a Strategic Committee was established, enabling programme leaders within the LCM to meet periodically with the consultants (sometimes more than once a week). In addition, the LCM Innovation Office, which had been established some months before the programme commenced, focussed entirely on BPJO implementation and became the operational office, under the jurisdiction of the Strategic Committee.

In considering factors contributing to knowledge transfer, it is interesting to explore the existence and development of trust between LCM leaders and consultants. For a relatively closed organisation with unique special powers and professional knowledge, trusting the consultants might in fact be of relevance in ensuring substantial and lasting organisational change. In this respect, it is useful to examine whether the programme helped to develop such trust (c3). Congruence of the evidence with the test would increase the importance of BPJO in assisting knowledge transfer (i.e. low false positive). However, it should be noted that evidence of the programme’s effect on trust is not necessary in claiming that BPJO supported transfer (i.e. low true positive).

The test again finds confirming evidence. Although the university presence among the consultants was seen to assure reliable knowledge, there was initial mutual distrust, and the consultants took a number of steps to build trust. In particular, they began work before the contract was finally formalised, and they offered additional services free of charge.
One of these – a two-day meeting at Lake Como – was designed to address one trust problem by socializing magistrates and consultants. To signal the importance of the service, the first post-contract meetings were attended by the Rector of the Politecnico di Milano. The interviewed magistrates reported their initial fear of being evaluated by the consultants during the first period in particular. They saw the service as an intrusion into their job and knew that the consultants had no previous experience of judicial offices, and they were therefore keen to check what the consultants were doing. Despite this initial mistrust, the relationship developed into a collaborative effort after the first interventions, yielding several spin-off joint projects beyond the main service.

Finally, it is important to verify whether the fieldwork reveals other factors unrelated to the programme as essential elements in knowledge transfer (c4). While this would not wholly devalue the contribution of BPJO, it would indicate possible gaps between the consultants and the LCM. In terms of the programme’s relevance, the absence of any such factors would represent a smoking gun rather than necessary evidence (i.e. low true positive) but would constitute sufficient evidence if found (i.e. low false positive).

The evidence points to the importance of factors extraneous to the programme. The novelty of the knowledge to be transferred, the need to adapt that knowledge, and the relative closure of the target organisation made transfer difficult. Although magistrates wanted change, they were unsure of its direction and feared the consequences of that transformation. Additionally, the consultants were unsure (at least initially) how to apply their knowledge to a judicial office. These difficulties were overcome by means of a brokerage function performed by the LCM president’s personal consultant; direct observations and interviews confirm that the consultant played a fundamental role in bridging the gap. His professional profile enabled him to move easily between the two groups, and he was trusted by all of the court’s top leaders. Through ongoing negotiation, he was able to smooth the groups’ interactions and to manage information flows in both directions. Notably, his position as a broker was acknowledged by court leaders and consultants alike.

Table V summarises the evidentiary material for Step C. The generally low level of definition of the BPJO framework meant that significant adjustment was necessary at local level once the programme had started. Although there was some initial distrust between the LCM leaders and the consultants, BPJO implementation helped to increase contacts and build trust.

However, a key factor for the success of knowledge transfer was the action of the personal consultant to the president of the court, who was able to direct consultants towards the main critical points in the administration, translating LCM leaders’ needs and facilitating information flows between the two groups.

9. Conclusions
The Italian courts are typical of organisations that resist administrative reform. Magistrates constitute a closed group with a homogenous culture, self-government, a weak
organisational hierarchy, and few managerial competences. In addition, BPJO was weak and vaguely designed, with no governance at national level. This made the LCM’s success particularly interesting. To test whether and how the programme was responsible for success, and whether case-specific factors interacted with the programme, the present study builds and tests a theory of programme success based on leaders’ engagement, knowledge access, and knowledge transfer.

The LCM case points to an interplay of local and programme features that together account for the success of implementation. While knowledge access is confirmed with no notable qualifications, the evidence in respect of both leaders’ engagement and knowledge transfer warrants further comment. In relation to leaders’ engagement, BPJO represented a ready-made opportunity for would-be innovators. In fact, LCM leaders were active in reforming their office even before the launch of BPJO, but they needed the programme to support their engagement and to provide the necessary momentum. This tends to confirm the ability of public sector leaders to make use of national programmes and political opportunities as focal events in support of their reform projects (Abramson and Lawrence, 2001). It should also be noted that BPJO was not only an enabler (i.e. a provider of resources), but also served as a framework for envisaging a more comprehensive reform strategy, which is congruent with the importance of framing change management efforts within a wider vision (Kotter, 1996, 2014).

In relation to knowledge transfer, the causal effect of the programme, though weaker, is nevertheless confirmed. Certainly, BPJO facilitated repeated interactions between the two groups through dedicated meetings that helped to build mutual trust. However, the brokerage action performed by the consultant to the LCM president must be taken into account as an additional causal factor external to programme design. The visible importance of this brokerage function directs designers to increase their attention to this “translation process” between consultants and organisational leaders. This translation by the broker was important not only as hypothesised (i.e. to adjust managerial knowledge to a novel context) but also to increase magistrates’ trust by ensuring that the new management tools respected their professional principles.

The method of PT warrants some final comments. The development of empirical tests before examining the case material not only serves to guide the collection of evidence but also ensures transparency in data collection and in updating confidence in a given hypothesis. Granted that reasoning in terms of true and false positive rates lacks the clean edges of statistical significance testing, this is nonetheless an important tool for reviewing the significance of case study data. In addition, preliminary assessment of a test’s probative value is an incentive to formulate demanding tests and clearly stated hypotheses for ease of empirical scrutiny. Finally, while enhancing the rigour of case analysis, the method remains open to alternative explanations that may emerge during fieldwork. By selecting several case-specific factors, the approach outlined here both enables effective testing of an initial hypothesis and enriches the explanation. In short, PT seems likely to increase accuracy and transparency while maintaining an openness to emergent causal factors.

Note

1. The reasoning is formalised in the Bayes theorem as a formula combining these three measures (prior likelihood, false positive rate, and true positive rate) to update confidence estimates (i.e. measuring the updated probability of H conditional to finding a piece of evidence with certain true and false positive rates). For examples of formal treatments, see Befani and Stedman-Bryce (2017). For a first attempt to apply the Bayesian formula to a full case study, see Fairfield and Charman (2017). Note, however, that the informal Bayesian updating used throughout the paper is the dominant approach; both methodologists (Beach and Pedersen, 2013) and substantive scholars (Busetti and Dente, 2017) refrain from using the formal Bayes formula.
References


Osborne, S.P. and Brown, K. (2005), Managing Change and Innovation in Public Service Organizations, Routledge, Abington.


Appendix

<table>
<thead>
<tr>
<th>Interview</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magistrate responsible for national implementation of BPJO</td>
<td>9 July 2013</td>
</tr>
<tr>
<td>Former president of LCM</td>
<td>8 July 2013, 22 November 2013</td>
</tr>
<tr>
<td>Magistrate responsible for implementation of BPJO in the LCM</td>
<td>6 February 2014, 8 May 2015</td>
</tr>
<tr>
<td>Magistrate responsible for individual interventions in the LCM (1)</td>
<td>8 July 2013</td>
</tr>
<tr>
<td>Magistrate responsible for individual interventions in the LCM (2)</td>
<td>9 July 2013</td>
</tr>
<tr>
<td>Manager of the Innovation Office in the LCM</td>
<td>8 July 2013</td>
</tr>
<tr>
<td>Leader of Consultants</td>
<td>9 July 2013, 26 September 2014</td>
</tr>
</tbody>
</table>

Table AI.

Interviews

Corresponding author
Simone Busetti can be contacted at: simone.busetti@unimib.it

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com