



HEC
PARIS

Insper

The Assessment Paradox: Practices Dissemination and Impact Evaluation

Fernando Deodato Domingos
Bertrand Quélin
Sandro Cabral
Sergio Lazzarini

June 19th , 2020 – SIOE

CSC literature recent developments

Scholars have focused on:

- **Arrangements generating spillovers** (e.g., Agarwal, Audretsch and Sarkar, 2010; Agarwal, Campbell, Franco and Ganco, 2016; Cabral, Lazzarini and Azevedo, 2013; Yang, Phelps and Steensma, 2010).
- **Creation of social externalities** (e.g., Rangan, Samii and Wassenhove, 2006; Kivleniece and Quelin, 2012).
- **Performance measurement implications over the public domain** (e.g., Behn, 2003; Campbell, 2002; Gerrish, 2016; Newcomer, 1997; Wholey and Hatry, 1992)

...however, there has been **scant attention to inherent tensions** that may emerge when partners would like to **promote positive externalities and at the same time measure the performance of collaborative efforts** (e.g., Cabral et al., 2013; Cabral, 2017; Ethiraj and Levinthal, 2009; Kroeger and Weber, 2014).

The Assessment Paradox

Inherent tensions may emerge when CSC partners would like to promote **dissemination of practices** and at the same time **measure the performance** of collaborative efforts.

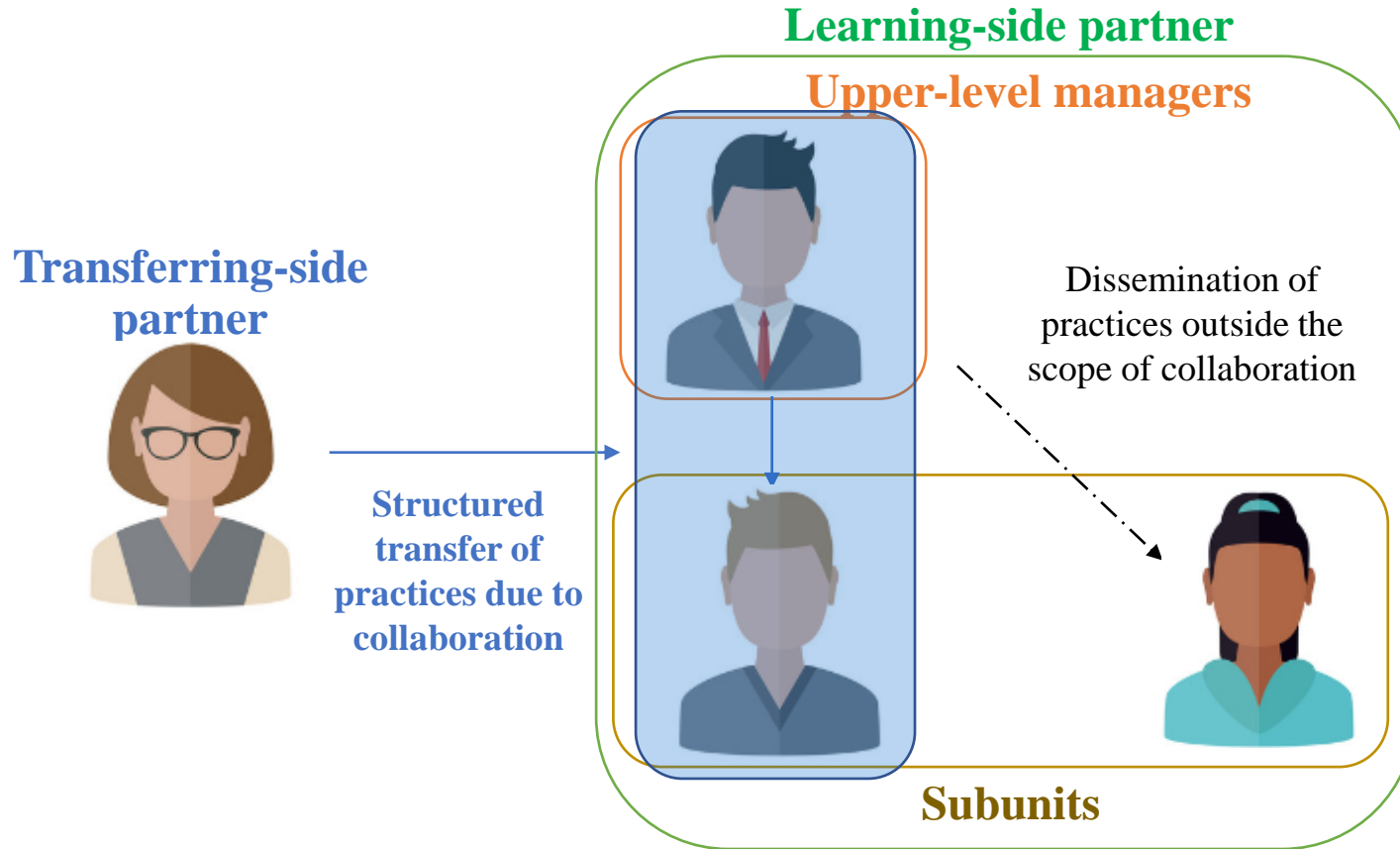
If managers should disseminate the practices known to positively impact performance of their subordinate subunits...

...but the process of properly measuring impact of a collaborative agenda prevent control contamination...

...The **assessment paradox** emerges.

(in other words, when collaborations aim to **promote generalized learning** and at the same time to **measure its impact**).

In a nutshell...



[H0] Baseline Hypothesis

Performance-enhancing managerial practices

Collaborations focused on transferring **proper managerial practices** (Bromiley and Rau, 2014) **were proven efficient** in several settings, as:

- the management of **health care systems** (e.g., Bloom, Propper, Seiler and Reenen, 2015; Banerjee, Duflo and Glennerster, 2008)
- **educational outcomes** (e.g., Bloom, Lemos, Sadun and Rennen; 2015; Duflo, Hanna and Ryan, 2012)
- **prisons** (e.g., Cabral et al., 2013).

For instance, **private or nonprofit actors** may bring **new technologies** and **improved managerial practices** that can be used by **public units** (e.g., Bruce, Figueiredo and Silverman, 2019; Cabral, Mahoney, McGahan and Potoski, 2019; Quélin, Cabral, Lazzarini and Kivleniece, 2019).

[H0] Baseline Hypothesis: The higher the level of structured transfer of performance-enhancing managerial practices, the higher the performance of targeted subunits.

Impact measurement and CSCs inherent tensions

Pursuing **both** dissemination of practices and accurate assessment of impact can be challenging, particularly in the context of CSCs.

CSC partners' efforts generally seek **cross-sector learning that often extrapolates the domain of the focal collaboration** (e.g., Cabral, Lazzarini and Azevedo, 2013; Luo and Kaul, 2019; Yang, Phelps and Steensma, 2010).

- Yet, allowing for practices dissemination beyond the target units may severely **distort the assessment of causal impact** if they also increase the performance of untargeted units serving as control groups.
- Assessment is a paradox as partners simultaneously aim at two goals that are nearly incompatible.
- Overall, such tensions arise when the social meaning of assessment diverges from its empirical merits.

[H1] Main Hypothesis

The inherent tensions between dissemination of practices and impact evaluation

- We propose that learning-side managers face inherent tensions once they get access to performance-enhancing practices and should also concern about impact evaluation purposes.
- Managers who better internalize practices, properly following guidelines and implementing managerial routines, are the ones who will mostly promote dissemination of practices and consequently hinder impact assessments. Formally:

[H1] Hypothesis 1: The higher the level of learning-side's internalization of received practices, the more impact assessments understate the transferring-side's impact on the targeted subunits.

[H2] Moderator Hypothesis

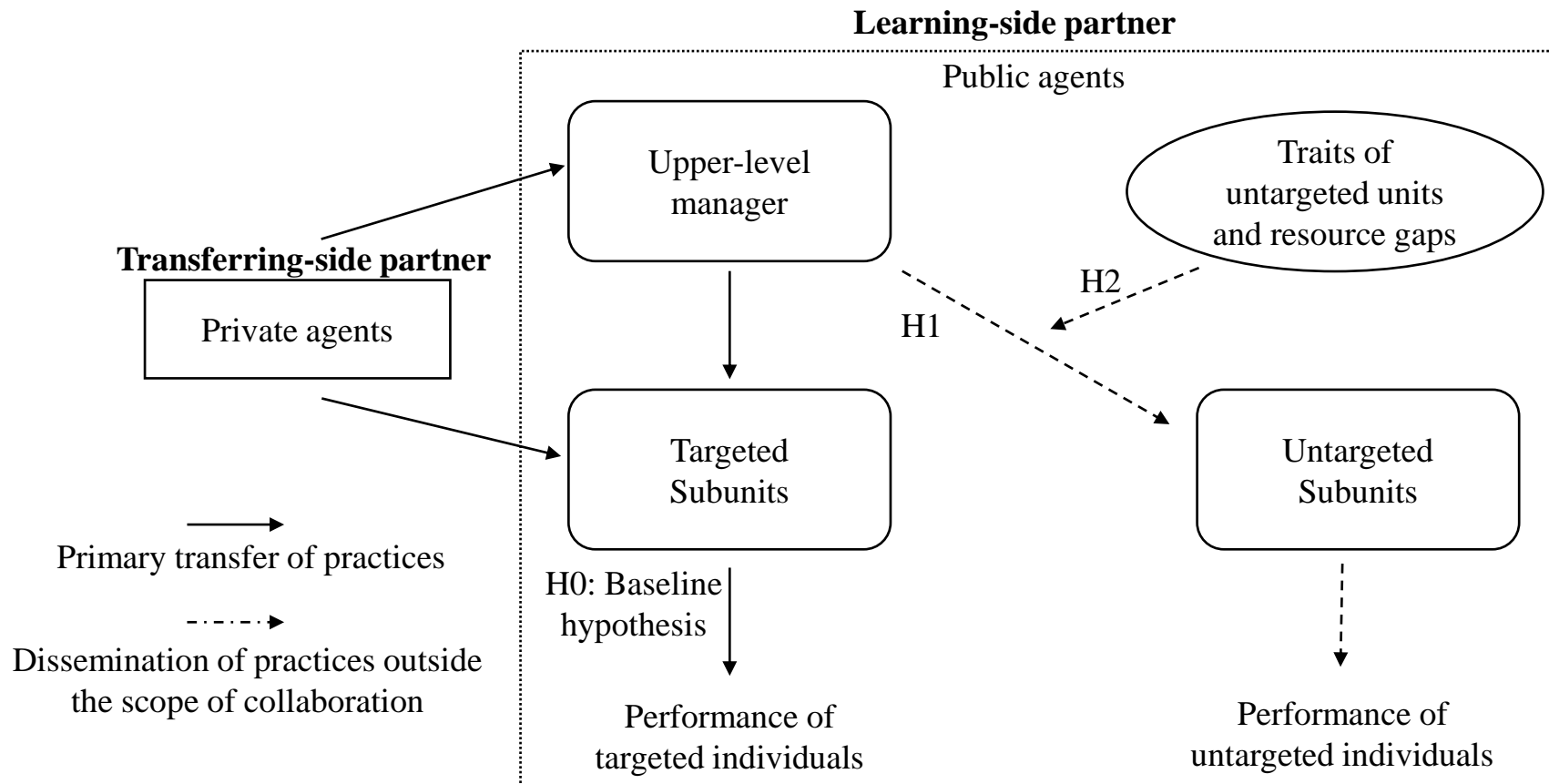
Organizational traits and resource gaps

- This sort of unstructured dissemination of practices does not rely on external constituencies securing successful adoption.
- The untargeted subunits presenting the least resources would have no mechanism to secure the success of the performance-enhancing practices. Specifically, those resource gaps could concern staff human capital or available physical resources.

[H2.a] Hypothesis 2.a: The effect of performance-enhancing practices dissemination on untargeted subunits' is positively moderated by the untargeted subunits' human capital.

[H2.b] Hypothesis 2.b: The effect of performance-enhancing practices dissemination on untargeted subunits' is positively moderated by the untargeted subunits' physical capacity.

Conceptual framework



Context and Empirical Design

- **Unibanco Institute:** nonprofit focused on fostering public education in Brazil.
- **Youth of the Future Program:** through partnerships with the State Secretariats of Education, aimed at improving management in high school, already adopted by 3,000 schools in 11 Brazilian states.
 - **Motivation:** poor performance of Brazilian students according to worldwide measures (e.g., PISA)
- **Knowledge-based collaboration:** Although the public schools are randomly assigned to receive the program, the most recent phase of the collaboration included managerial training for the managers of all school districts as well, in addition to the trainings aiming the targeted school principals.

Data and variables

Primarily, the intervention represented an exogenous shock influencing the transfer of practices directly to schools, but currently all schools are under the influence of school district managers who received managerial training.

- **Dependent variables:** Measures of school performance (Math and Languages).
- **Independent variables:** For H0, targeted schools and untargeted schools ($T=1$); for H1, we rely on an index developed by the private institute to observe the upper-level managers' internalization of managerial practices. Besides, we code two dummy variable indicating the high-capacity schools: one based on their human capital (H2.a) and other their physical capacity (H2.b).
- **Control variables:** We count on several important features at the schools, teachers and students' level.

Methodology

All models reported are **clustered OLS regressions** at the school's districts level. We thus specify the following general model explaining the educational outcomes following a **difference-in-differences methodology**:

$$Y_{i,t} = TargetedSchool_i + PracticesInternalization_i +$$

$$+ Post * TargetedSchool_i + Post * PracticesInternalization_i + \quad (1)$$

$$+ Post * TargetedSchool_i * PracticesInternalization_i + Controls_i + \epsilon_{i,t},$$

The triple interaction between post treatment, targeted schools and internalization of practices aims to disentangle the effect of schools' district internalization of practices between targeted and untargeted schools (H1).

Second, we investigate the organizational resource gaps as a potential moderator for successful dissemination of practices (H2).

Math performance considering schools' treatment and upper-level managers' internalization of practices

OLS clustered regressions	(1)	(2)	(3)	(4)
DV: Schools' Math average score	Only school's treatment	W/ controls	W/ internalization of practices	W/ triple interaction
Targeted school	3.779 (2.296)	-0.938 (1.611)	-0.943 (1.590)	-0.740 (1.644)
Post treatment	1.237 (1.409)	2.552 (1.499)	2.379 (1.377)	2.108 (1.304)
(H0) Targeted # Post	4.392*** (0.966)	3.366*** (0.996)	3.377*** (0.975)	3.803*** (0.943)
Internalization of practices			5.460*** (1.019)	6.141*** (1.791)
Internalization # Post			0.876 (1.113)	2.397 (1.523)
Targeted # Internalization				-1.020 (1.871)
(H1) Internalization # Targeted # Post				-2.396** (1.103)
Distance to closest targeted school		-0.0742 (0.122)	-0.0740 (0.122)	-0.0796 (0.116)
Controls and FE	N	Y	Y	Y
Constant	259.8*** (4.305)	231.6*** (15.91)	232.4*** (15.80)	235.0*** (15.97)
Observations	1,496	1,496	1,496	1,496
R-squared	0.022	0.669	0.669	0.670

Notes: The table present OLS regressions estimates explaining schools' performance regarding their students' average math score. The difference-in-difference estimates refer to the average treatment effect. The control variables include all schools', teachers', and students' features, as described in our methodology section. Robust standard errors in parentheses (clustered at the schools' district level); *** p<0.01, ** p<0.05, * p<0.1

Language performance considering schools' treatment and upper-level managers' internalization of practices

OLS clustered regressions	(1)	(2)	(3)	(4)
DV: Schools' Language average score	Only school's treatment	W/ controls	W/ internalization of practices	W/ triple interaction
Targeted school	3.953* (2.150)	0.0578 (1.273)	-1.071 (1.377)	-1.067 (1.401)
Post treatment	4.766** (1.785)	5.363*** (1.431)	5.716*** (1.589)	5.516*** (1.607)
(H0) Targeted # Post	4.312*** (1.228)	3.187*** (0.890)	3.285*** (1.108)	3.595*** (1.182)
Internalization of practices			3.786*** (1.170)	3.776** (1.628)
Internalization # Post			1.485 (1.423)	2.590 (1.999)
Targeted # Internalization				0.0443 (1.389)
(H1) Internalization # Targeted # Post				-1.743 (1.287)
Distance to closest targeted school		-0.0191 (0.134)	-0.152 (0.0888)	-0.154* (0.0876)
Controls and FE	N	Y	Y	Y
Constant	252.6*** (4.076)	234.5*** (17.00)	232.8*** (18.20)	233.9*** (17.68)
Observations	1,496	1,496	1,496	1,496
R-squared	0.037	0.619	0.641	0.641

Notes: The table present OLS regressions estimates explaining schools' performance regarding their students' average Language score. The difference-in-difference estimates refer to the average treatment effect. The control variables are the same as in Table 3. Robust standard errors in parentheses (clustered at the schools' district level); *** p<0.01, ** p<0.05, * p<0.1.

Untargeted schools' performance and the potential moderator effect of resources

OLS clustered regressions	(1)	(2)
Subsample of untargeted schools	Math performance	Languages performance
Internalization of practices	3.169*	-0.317
	(1.697)	(2.519)
Post treatment	1.824	4.071*
	(1.651)	(2.087)
Internalization # Post	3.873*	5.185*
	(1.938)	(2.627)
Human capital	-9.027**	-10.08**
	(4.028)	(4.235)
Internalization # Human Capital	4.311	-0.243
	(5.386)	(6.025)
Human Capital # Post	2.947	0.821
	(4.651)	(5.127)
(H2.a) Internalization # Human Capital # Post	-1.413	4.841
	(7.031)	(7.691)
Physical Capacity	-1.366	-2.296
	(3.026)	(4.224)
Internalization # Physical capacity	5.142	8.265
	(4.701)	(5.619)
Physical Capacity # Post	0.736	3.710
	(3.842)	(4.182)
(H2.b) Internalization # Physical Capacity # Post	-3.279	-10.46
	(6.378)	(6.268)
Distance to closest targeted school	0.615**	0.492
	(0.263)	(0.339)
Controls and FE	Y	Y
Constant	245.7***	245.5***
	(18.02)	(24.50)
Observations	631	631
R-squared	0.717	0.664

Notes: The table present OLS regressions estimates explaining the subsample of untargeted schools' performance regarding their students' average score (Specification 1 for Math and Specification 2 for Languages). The difference-in-difference estimates refer to the average treatment effect. The control variables are the same as in Table 3. Robust standard errors in parentheses (clustered at the schools' district level); *** p<0.01, ** p<0.05, * p<0.1.

Main findings

[H0] Baseline Hypothesis: The higher the level of structured transfer of performance-enhancing managerial practices, the higher the performance of targeted subunits.

- **Supported:** according to prior literature.

[H1] Hypothesis 1: The higher the level of learning-side's internalization of received practices, the more impact assessments understate the transferring-side's impact on the targeted subunits.

- **Supported:** consequently, it emerges the assessment paradox.

[H2] Hypothesis 2: The effect of performance-enhancing practices dissemination on untargeted subunits' is positively moderated by the untargeted subunits' [2.a human capital / 2.b physical capacity]

- **Not supported:** we acknowledge these common practices may be imitable and easily implemented (Bromiley and Rau, 2014), what even attenuates the assessment paradox as the managerial practices may not require resourceful schools for proper dissemination.

Conclusions

- **Strategic management literature implications**
 - Assessment paradox may represent an important tradeoff for managers with both goals of willing to measure impact while disseminating practices;
 - Cross-sector collaborations potentially attenuate such assessment paradox;
 - Insights to practice transfer literature concerned to the environment and actors involved in such practice: how do marginal costs to dissemination influence dissemination?
- **Managerial implications to public policy and managerial decision-making related to knowledge-based collaborations**
 - Practices dissemination potentially understate the private efforts within the targeted subunits.
 - Collaborations including consulting-related activities could not be long-term profitable for consulting firms if their partners rapidly develop expertise to transfer best practices to their own units without external monitoring.



Appendix

Limitations and future research

- Framework requires **imitable and easily replicable practices**.
- Research design: **we face the assessment paradox ourselves**.
 - Future research could deeply explore causal linkages.
- Does **boundary conditions restrict the analysis to cross-sector?**
 - Desirable/enforced dissemination of practices;
 - Desirable/enforced impact evaluation.

Robustness checks

- **Other DV specifications:** we use different specifications for our educational outcomes (math and language scores) and findings are maintained.
- **Using different specifications for “internalization of practices”:** we check whether results are the same using different factor analysis extraction methodologies for our index of internalization of practices, as the maximum-likelihood instead of the principal axis factoring (Costello and Osborne, 2005). The findings are maintained, concerning the full sample and original index.
- **Dropping “internalization of practices” outliers:** findings are similar to the full sample analysis (dropping 5%; 2,5% highest and 2,5% lowest). Robust to other thresholds.
- **Differences between higher x lower “internalization of practices” districts:** there is no substantial difference between schools from highest or lowest compliance districts.

Internalization of practices measure

Unibanco Institute

The Private Initiative Index for Internalization of practices

Index composition and score categories	Learning of the Managers' Role	Internalization of Routines	Interactions with Regional Supervisors	Interactions with Schools	Interactions with Secretary
Lowest score: 1 out of 4	Executes only documentation	No proper planning, mainly bureaucratic	None or few meetings	None or few meetings	No or few meetings
Score: 2 out of 4	Documents and presents some concerns regarding general pedagogical issues	Basic planning with no focus on students' learning	Meetings only when demanded or less than monthly	At least a meeting per school every 3 months	At least a meeting every 3 months
Score: 3 out of 4	All previous concerns, and mentions the importance of connections with schools	Planning focused on students' learning, but with no routines regarding data analyses	At least a monthly meeting is required	At least a meeting per school, between 1 and 3 months	At least a meeting, between 1 and 3 months
Highest score: 4 out of 4	All previous concerns, and emphasize the importance of close monitoring and feedback with schools	Planning focused on students' learning, implementing routines with data analyses	At least a meeting every two weeks is required	At least a monthly meeting with every school	At least a monthly meeting