Sources of performance improvement in global pollution elimination fund: A learning progressions hypothesis

David Grover
Grenoble École de Management
12 Rue Pierre Semard
Grenoble 38000
FRANCE
david.grover@grenoble-em.com

Claudio R. Lucinda
University of São Paulo
908 Av. Prof. Luciano Gualberto
São Paulo, SP, 05508-010
BRAZIL
claudiolucinda@usp.br

March 2023

ABSTRACT

We study the performance of a multi-lateral financing organization responsible for funding pollution elimination projects in low- and middle-income countries. We find that the organization's performance improved by around 20 percent over two decades and that a large part of this improvement arose from experience gained internally by the organization itself. We identify a paradox in the way that the task iterations performed by this and other organizations can be measured to accumulate into organizational experience. We find that differently-accumulated task iterations produce experience variables with coefficients that vary by a factor of 4. We propose that learning progressions – a theory that we transpose from the pedagogical sciences to the organizational learning literature - may explain this paradox. The findings of this research enrich the organizational performance and learning literature with insights from the pedagogical sciences. They also provide insights for improving the proficiency of multi-lateral financing organizations at funding global pollution elimination activities today.

KEY WORDS

Organizational performance, learning curve, global pollution, learning progressions

ACKNOWLEDGEMENTS

The authors are grateful to the Secretariat of the Montreal Fund for the Protection of the Ozone Layer for facilitating data access.

DATA DECLARATION

The data and code used in this research are available upon request and within the bounds of approved use by the Secretariat.

1. Introduction

We study the performance of a multi-lateral public financing organization that facilitated the elimination of a global pollutant over two decades. We test explanations for why the organization's performance improved over time and advance a new hypothesis for organizational performance improvement generally that is based in the idea of learning progressions.

The multi-lateral financing organization, like the one that we study here, is the organizational vehicle that the community of nations has chosen for channelling large amounts of financial capital into climate change mitigation and adaptation projects globally. One of the largest of these organizations today is the Green Climate Fund (GCF), based in Seoul, which received around US $10 billion from donor countries between 2015 and 2021 and directed it into projects in low- and middle-income countries (GCF, 2021).[1] Donor countries have contributed more than US $50 billion to these organizations to date (Lee et al., 2023). The aspiration is for the GCF alone to be receiving and distributing more than US $100 billion *per year* by the year 2030 (Lattanzio, 2017; Markandya et al., 2015).

These organizations are at the public administrative forefront of addressing contemporary global environmental issues, yet little is known about how their performance evolves over time and/or what drives this change. In fact, most related research looks at how donor countries can or should channel more money to these organizations (Bowman and Minas, 2019; Cui et al., 2020; Cui and Huang, 2018) rather than at how these organizations can generate more impact for the same amount of money. The impact of this type of organization matters because political and business leaders consistently rank public sector failure on 'climate action' among the top risks to global economic stability (World Economic Forum, 2022).

We find that the performance of the multi-lateral financing organization that we study improved over time. Probably the most widely regarded explanation for performance improvement in any organization is the organizational learning or experience curve (Argote et al., 2021; Jaber, 2011;

[1] Because low- and middle-income countries are expected to account for more than 90 percent of the *growth* in GHG emissions by 2050 (Edenhofer et al., 2014; OECD, 2012).

Thompson, 2012). We find strong evidence that the organization in question benefitted in performance terms from its own internal experience.

However, we also identify a paradox related to the relationship between organizational performance (improvement) and experience. The paradox is that the magnitude of the estimated 'experience effect' varies by a factor of four depending on the way that the task iterations performed by the organization are measured to accumulate into experience. Why would the exact same set of tasks iterations, measured differently, produce such different experience coefficients?

We propose that learning progressions[2] bear on the way that the task iterations performed by an organization accumulate into experience and that learning progressions are what produce different varieties of experience, which in turn have different 'saliences' for organizational performance and improvement. The idea of the learning *progression* emerged in pedagogical science and not in economic or management science. Learning progression theory has not been applied to or tested in a 'workplace organization' to our knowledge.[3]

Learning progressions theory holds that learning outcomes in a group of learners can be improved, relative to some baseline, by exposing the group to the elements of a topic domain in an order that facilitates an accumulating, self-reinforcing, pathway to understanding (Duschl, 2019; Jin et al., 2019; Liu and Jackson, 2019; National Research Council, 2007; Taguma et al., 2019). The idea of learning progression falls somewhere between the organizational learning that occurs through unstructured, on-the-job experience and formal workplace training interventions designed to build specific skills or knowledge (Hammer et al., 2019; Krueger and Rouse, 1998; Nadler et al., 2003; Salas et al., 2012).

In hypothesizing a role for learning progressions in the performance-experience relationship our research builds on several prior studies. Prior related work has tested how organizational

---

[2] Also 'experience progressions', 'conceptual trajectories', 'learning trajectories', 'developmental corridors', 'developmental progressions', and 'learning continua' (Jin et al., 2019). A common definition of a learning progression is 'a sequence of successively more complex ways of thinking about a practice or content that develop over time' (Ford, 2008).

[3] By which we mean an organization that is populated by workers and managers and where while no 'teaching' formally occurs, as in a school organization, learning nonetheless occurs.

performance responds to heterogeneity in the task iterations that organizations perform. One study of cardiothoracic surgeries performed by surgeons defined a surgery (a task iteration) as 'successful' if the patient survived beyond a time threshold and 'unsuccessful' if the patient did not (KC et al., 2013). That study found that successful surgeries associated with faster surgeon performance improvement than unsuccessful ones. Another related study considered how airline safety performance responded to prior safety incidents (task iterations) that had either homogeneous or heterogeneous causes (Haunschild and Sullivan, 2002). That study found that certain airlines learned more from accidents with heterogeneous causes than with homogeneous causes, because heterogeneous causes demanded deeper search and learning by the organization for understanding. Another related study estimated the effect of different 'varieties' of experience on the time it took to perform joint replacement procedures in a teaching hospital (Reagans et al., 2005). That study found that experience that accumulates in individual workers, in teams, and in the organization as a whole all contributed in different ways to performance improvement.

This paper makes several contributions to the organizational performance literature and to the empirical problem of financing global pollution elimination. First, it demonstrates the paradox in the performance-experience relationship[4] discussed above and proposes that organizational learning progressions are a plausible explanation for it. Second, it documents how the performance of a single multi-lateral financing organization changed over 22 years while working to eliminate a single global pollutant, chlorofluorocarbons (CFC). During the period in question CFC use fell by over 97 percent in the low- and middle-income countries where the organization's financing activities were targeted. This means that we study the performance of the relevant organization during the entirety of a virtually complete 'transition' away from a single global pollutant. Third, we find that the organization's performance improved by around 20 percent over the period and that the organization's own internal experience accounts for a much larger share of that improvement than all time-based forces originating outside the organization. We compare the learning effect of geography-, industry- and project scale-based varieties of experience and find that scale-based experience consistently associates with the fastest performance improvement. We also introduce a novel methods approach to the literature on the organizational performance-experience relationship by using creating counterfactual experience variables based on semi-random sequences of task

[4] By which we mean learning / experience curve theory and/or the cost-quantity relationship (Thompson, 2012).

iterations and using them to elucidate the learning progression that we hypothesize exist in the observed data.

2. Learning progressions and organizational performance

Organizational experience is widely regarded as contributing to organizational performance improvement, including in financial services organizations (Ashworth et al., 2004; Barnett et al., 1994; Drexler and Schoar, 2014). At the heart of the organizational performance-experience relationship are the iterations of a task that the organization performs. A task is a discrete action that the organization performs repeatedly in producing a good or service (Harris, 2017; William and Harter, 1899; Yelle, 1979). The organization's ability to perform the task is not directly observed, but held to improve with the number of iterations of the task that it performs ('experience') (Argote, 2013; Fiol and Lyles, 1985; Jaber, 2011; Levitt and March, 1988; Thompson, 2012). Experience is typically represented as the number and/or intensity of prior iterations performed. Experience can arise directly when the organization performs the task itself or indirectly as other organizations perform the task (Baum and Ingram, 1998; Epple et al., 1991; Thornton and Thompson, 2001). It is supposed that experience improves performance by helping the organization recognize and avoid errors, delays, defects and mistakes that can occur in performing the task (Cohen and Bacdayan, 1994; Jaber, 2011; Levitt and March, 1988; Wright, 1936) and / or by revealing ways for it to improve final product or service quality (Dosi et al., 2017).

The hypothesis that we develop is that the strength of the performance-experience relationship depends on how the task iterations that the organization performs accumulate into experience in reality, and how the analyst represents experience via measurement empirically. The hypothesis focuses on the term E_t in Equation 1, which shows the standard performance-experience relationship (Argote et al., 2021; Levitt et al., 2013; Sinclair et al., 2000). Conceptually, E_t represents the level of organizational ability or experience in the period prior to performing a given iteration of a task. C_t represents the level of performance of a task iteration at time t, A the level of performance of a task iteration in some initial period, and β the relationship between performance and experience.

Equation 1: Standard organizational performance-experience relationship

$$C_t = AE_t^{\beta}.$$

Equation 2 shows how the term E_t is typically conceptualized and measured in variable form. Q_t is some indication of the number of iterations of the task that the organization performed in each period. $t = 1$ is the first period and $T - 1$ is the period prior to the current period. A typical way of measuring the experience variable is to arrange the period-specific quantities in chronological order, then take their running sum. The resulting values are sometimes depreciated by a constant to represent how the salience of experience that is generated in one period might decay before it is drawn on in a subsequent period (Darr et al., 1995; Egelman et al., 2017).

Equation 2: Experience over time only

$$E_t = \sum_{t=1}^{T-1} Q_t$$

We hypothesize that learning progressions enter the picture when the experience variable is measured to accumulate over time but within a set of categories that characterise the task iterations. The categories might be geographic regions in a country, product lines in a firm, or production technologies in an industry. Equation 3 illustrates this. For illustration we can suppose that the categories are geographic regions. Subscript *i* indexes regions 1, . . . , *I*. This means that $Q_{i,t}$ is the sum of the task iterations performed in each region-period and $E_{i,t}$ is the running sum of those quantities within each region over time.

Equation 3: Experience over time and task characteristics

$$E_{i,t} = \sum_{t=1}^{T-1} Q_{i,t}$$

The following examples illustrate the range of category types that task iterations and so experience have been measured to accumulate over. A study situated at a mail delivery company measured organizational performance as the mean cost per mail item delivered in each of 27 service regions, in each month, and experience as the cumulative quantity of mail items delivered in each region-month (Wiersma, 2007). The categories in this study are geographic regions. A wind farms in

China measured performance as the cost of electricity production at the level of the wind farm, and experience as the cumulative quantity of wind farm capacity built by the project developer (Tang and Popp, 2016). The type of category in this study is project developers (firms). A study situated at a single chemicals manufacturing firm measured the performance as the average cost of producing 224 chemical products, in each month, and experience as cumulative quantity of chemical production in each month. The type of category in this study is chemical products. A study of goods manufactured by Indian firms measured performance as the price of the goods produced by each firm in each year, and experience as the cumulative quantity of each good manufactured by each firm in each year (Dosi et al., 2017). The type of category in this study is 'firm-goods'.

All four of those studies find that performance improves with greater 'experience', except the strength of that relationships is essentially incomparable across studies, not least because the different category types represent fundamentally different types of experience. The empirical element of our research shows that the categories that task iterations are measured to accumulate over has a large impact on a) the strength of the relationship of interest and b) the extent to which experience explains total variation in organizational performance.

We show this by measuring the exact same set of task iterations to accumulate over different category types, as in Equation 4. In Equation 4, subscript (*x*) represents the different category types. These are categories defined by geography (countries, denoted *g*), categories defined by industry (industries, denoted *j*), and categories defined by the scale of the task iteration (scale bands, denoted *s*). Regardless of category type and experience variable, the task iterations are always measured to accumulate in the chronological order in which they occur.

Equation 4: Experience over different category types

$$E_{(x)t} = \sum_{t=1}^{T-1} Q_{(x)t}$$

Our hypothesis is that category types reflect something like learning progressions and that learning progressions explain this variation in the strength of the performance-experience relationship to an extent. Learning progression research emerged around the early 2000s as education scientists, practitioners and policymakers were looking for ways to improve learning outcomes in formal instructional environments (schools) (National Research Council, 2007, 2004). Learning progressions research arose partly in response to the critique that the curricula in many schools were too fragmented, discontinuous, and non-accumulating to achieve the learning outcomes that experts and policymakers saw as desirable (Duschl et al., 2011; National Research Council, 2007, 2004).

Learning progressions research proposes that groups of learners pass through distinct, identifiable cognitive stages as they acquire progressively greater competence in a topic domain. The cognitive stages that groups of learners pass through have been described as being akin to 'rungs on a ladder'. If the main rungs in a learning process can be identified and characterised through research, it is supposed, then learning outcomes might be ensured or improved by emphasizing the rungs in curriculum design and instructional practice choices (Duschl et al., 2011; Duschl, 2019; National Research Council, 2007). Learning outcomes of interest typically include: speed of apprehension, depth of understanding, duration of knowledge retention, and facility in applying knowledge in different circumstances (Reed and Wolfson, 2021; Stevenson et al., 1994).

The idea that learning progressions exist is based on theories of how cognitive change occurs in groups of learners aged 0 to 12 years old but also how it occurs in adults. In one view of how cognitive change occurs, learners transition away from intuitive ways of thinking about a topic domain (pseudo theories) towards more reason-based ways of thinking (scientific theories based on evidence and logic) (Jin et al., 2019). In this transition, learners update their systems of understanding by allowing the weaker ideas they once held to be swept away by stronger ones (Chen et al., 2019; Duschl et al., 2011; Perry, 1997), sometimes self-reflectively so, when the learners are adolescents or adults.[5]

The following are concrete examples of learning progressions that education scientists and practitioners have proposed and found at least observational evidence to support. Several learning progressions have been proposed for the way that preschool learners come to comprehend why

[5] In another view of how cognitive change occurs, groups of learners progressively accumulate 'facts' in a topic domain, but rather than those facts eventually converging to some fixed conceptual structure that the instructor considers 'correct', learners acquire the ability to re-configure the facts to the variable demands of whatever context they are operating in (di Sessa and Minstrell, 1998). Thus what is 'learned' is a kind of capacity-for-plasticity of the application of facts (Hatano and Inagaki, 2016). This means that evidence that learning has occurred is partly gauged by the range of contexts in which learners are able to apply the facts for the generation of insight (Duschl, 2019; National Research Council, 2004).

objects sink or float. The progression conjectures that learners first comprehend the concept of ‘density’, then ‘relative density’, and then ‘force and submersion’, and finally ‘buoyancy’ (Gao et al., 2020; Paik et al., 2017). Another learning progression has been proposed for the way that middle school students come to comprehend a model of aquatic ecology (Lehrer et al., 2008; Lehrer and Schauble, 2012). In this progression, students first comprehend the concept of ‘quantity’, then ‘measurement’, then ‘relations among quantities’, then ‘representation’. Another progression proposes how high school students learn to justify a claim to knowledge through the process of scientific argumentation (Berland and McNeill, 2010). In this progression, students first learn to state and defend claims, then to question others’ claims and defences, then to evaluate others’ claims and defences, then to revise their own and others’ claims. Another progression proposes how third-year university students come to understand the origins and evolution of the flora and fauna on island ecosystems through exposure through fieldwork (Meyers and Nulty, 2009).

The purpose of these examples is to illustrate the learning progression concept and not to imply that the types of substantive knowledge therein would necessarily be useful to a workplace organization. The only instance we are aware of where a learning progression has been proposed and applied in a ‘workplace organization’ is in fact in a training program for pre-service teachers (Álvarez Valdivia and Lafuente Martínez, 2019). Beyond this we are aware of no empirical research in the learning curve field or elsewhere that has proposed or tested for learning progressions in non-school workplace organizations.

Our basic hypothesis is that the category types that task iterations within an organization can be measured to accumulate over represent something akin to learning progressions. The first and weakest hypothesis that we test is that the coefficients on the experience variables that we measure for the organization in question are each significantly different from zero. The second is that the coefficients on those variables are different from one another. The third is that each is different from the coefficient on a counterfactual experience variable where task iterations are measured to accumulate over randomly-defined category types.

3. Empirical setting

We test our hypotheses using data on CFC elimination projects that were funded by the Multilateral Fund for the Implementation of the Montreal Protocol (the Fund).[6] The Fund was created around 1990 for the specific purpose of implementing a provision in the Montreal Protocol on the Protection of the Ozone Layer[7] (the Protocol). The provision involved providing financial assistance to low- and middle-income countries to meeting their commitments under the Protocol.[8] The Fund is loosely attached to the United Nations Development Program, it is based in Montreal, and it continues to operate at the time of writing.

The Fund is comprised of a Secretariat (13 professional staff plus 10 general service staff) and a 14-member Executive Committee. The Fund receives money from donor countries and channels it to recipient countries in the form of projects that eliminate CFC and other ozone-depleting substances (ODS). The Fund disbursed around US $2.7 billion to such projects during the period that we observe (June 1991 to December 2012), which is essentially its first 22 years of operation.

Our analysis focuses only on those projects that involved the elimination of CFC. The level of CFC consumption in the countries where the Fund funded projects fell by more than 97 percent during the period that we observe (see Table 1, appendix). This means that we study the Fund's performance during a virtually complete 'transition' away from a single global pollution.

The Fund's day-to-day work involves reviewing, revising and approving proposals for projects. It also regularly liaises with the international organizations (such as UNIDO) and bilateral development agencies (such USAID) that implement the projects in the field. The CFC elimination projects themselves were typically located at industrial facilities that manufacture refrigeration equipment, foam products, or cosmetics. Examples of specific projects are given below.

---

[6] This section draws on extensive documentation produced by the Fund itself about its internal policies and procedures (The Multilateral Fund Secretariat, 2022a, 2022b) as well as interviews carried out with Fund employees over several years.

[7] A treaty signed in 1989 by 197 countries agreeing to reduce the use of industrial chemicals that deplete the global ozone layer.

[8] Figure 3 (appendix) shows the countries that were eligible for financial assistance from the Fund when it began operating in 1991 ('Article 5' countries in the language of the Protocol, of which there were around 140).

The Fund's project approval work serves its legal remit under the Protocol, which is to meet all of the 'agreed incremental costs' (UNEP, 2000) of all eligible countries in eliminating CFC (and ODS). 'Incremental' means all capital and operating costs that are directly connected to converting equipment and processes to a technical standard that eliminates CFC pollution. The term 'incremental' is carefully defined in the Fund's operating protocol (The Multilateral Fund Secretariat, 2022a, 2022b).

We believe that the project data that we use are credible because they are the same data the Fund uses for its own internal financial accounting, reporting and control. The Fund never 'rejects' project proposals. When a project proposal is received that contains non-allowable costs, the Secretariate revises the proposal with the proposer until all costs are approvable. Once the Secretariate deems a project approval, it transfers it to the Executive Committee for final approval. The Executive Committee met approximately three times per year during the period that we observe, for approximately one week at each meeting. The Executive Committee approved proposals by consensus wherever possible.

Our conversations with the Fund employees as well as the internal documents that we reviewed suggest that the Fund got better at approving projects through its own internal experience with that process. Fund employees told us that both the Secretariate and Executive Committee got better at defining what 'incremental costs' are and at recognizing and eliminating non-allowable costs in project proposals. The documents show that the Secretariate and Executive Committee developed procedures for resolving disagreement over incremental / non-allowable costs. In the early days of the Fund's operation and prior to the creation of those procedures, disagreement was a time-consuming aspect of project approval that fell almost entirely on the Executive Committee (The Multilateral Fund Secretariat, 2022a). Fund employees told us that the Fund maintained a list of CFC elimination technologies that it considered reliable, and regularly updated the list as new technologies emerged. Fund employees perceived this practice to have improved the quality of match between project proposals and technologies over time. The Fund's internal documents show that it retained independent consultants to evaluate the strengths and limitations of the projects that it approved. These consultants helped the Fund identify 'major causes of delays and action taken to overcome difficulties' in project implementation (The Multilateral Fund Secretariat, 2022a).

4. Research design and data

We observe 2,440 CFC elimination projects. These are all of the CFC elimination projects that Fund ever funded and that the Executive Committee ever approved.[9] The Executive Committee approved the first CFC project in its fifth meeting, in November 1992, and the last in its 68th meeting, in December 2012. It is difficult to say what share of the 97 percent reduction in CFC use cited above is accounted for by the Fund's activity. Our exchanges with Fund employees imply more than half.

For each project we observe the total project cost in US dollars, the meeting in which it was approved, the quantity of CFC it eliminated in tons,[10] the country where the project was implemented, and the industry sector of the facility where the project occurred.

We fit the data to a standard performance-experience model as in Equation 5 below, where the unit of observation is the individual project. The dependent variable measures the performance of the Fund as the project unit cost, in US dollars per ton of CFC eliminated. The independent variables measure experience in terms of the number of tons of CFC that the projects eliminated, as discussed below.

The empirical model is derived from Equation 1 by taking logs of both sides. The term $E_{(x),t}$ is as in Equation 4. c_i denotes project unit cost, α is an intercept, and $e_{(x)t}$ is one of the three experience variables of interest (geography, industry, scale). The βs are interpretable as elasticities (or by transformation, learning rates). ε_i is a project-specific error. The project-level errors are assumed to be independent of one another but not necessarily identically distributed. Under this assumption, the estimates of β should be consistent, but the errors need to be adjusted to account for the departure from the i.i.d assumption, which is done in the estimation procedure. Estimation is by OLS.

Equation 5: Empirical model

$$\ln c_i = \alpha + \beta \ln(e_{(x)t}) + \varepsilon_i$$

[9] We verified this point with Fund employees: it did not approve a single CFC project after its 68th meeting in December 2012.

[10] 'Quantity of CFC eliminated' means the number of tons of CFC that the industrial facility was consuming in each year prior to the project being implemented (Decanio and Norman, 2005).

4.1. Dependent variable

The dependent variable is the cost in current US dollars per ton of CFC eliminated by the project. This is the *expected* cost as approved by the Executive Committee at the time of its meeting, before the project was implemented, and not the final, realized project cost. We used the expected cost because this seemed like a better measure of the Fund's own performance than realized cost. Realized cost, we were concerned, would reflect to some degree the performance or influence of implementing agencies, contractors, country authorities, and/or facility owners that were involved in project implementation.[11]

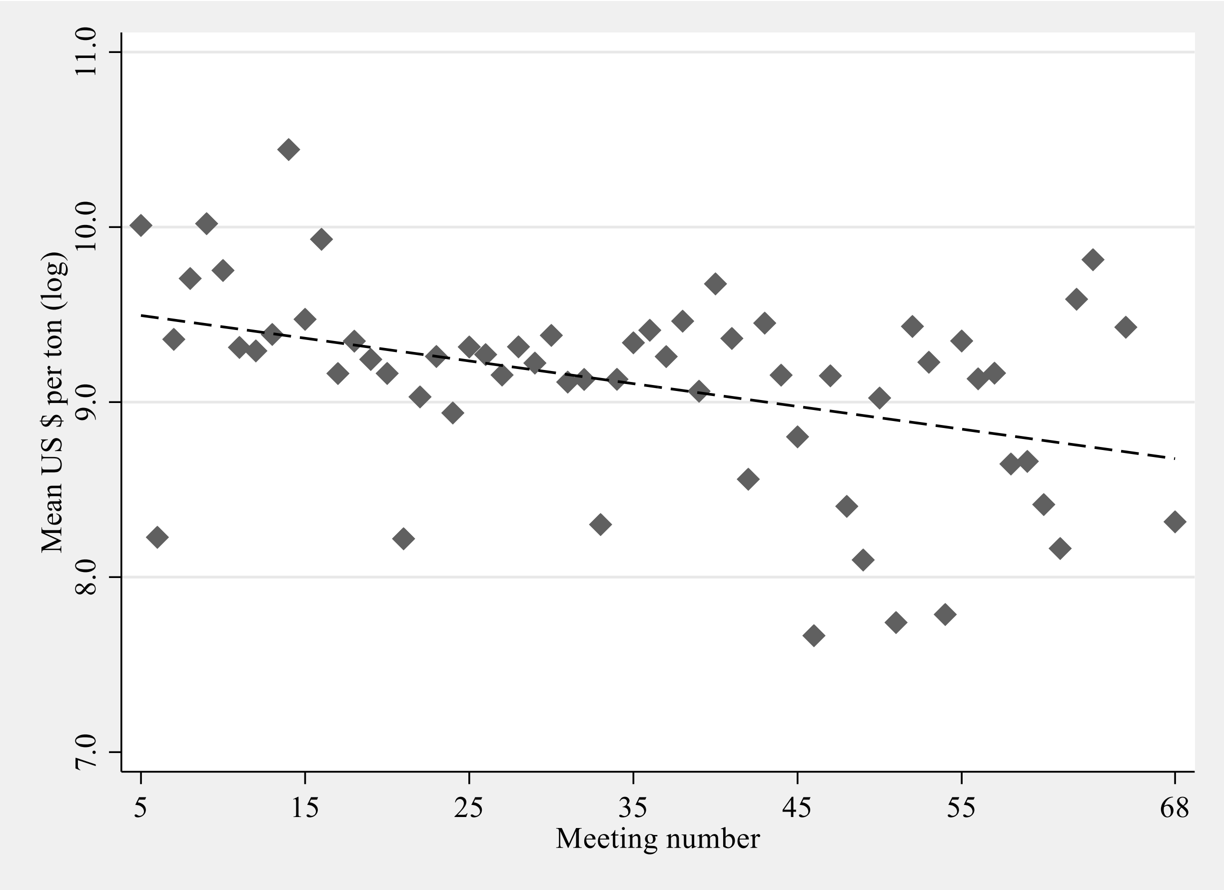
In the raw data, the mean cost per ton across all projects is US $12,316 and the median is US $7,472. The distribution is right-skewed, and for 30 projects, the cost is greater than US $83,333 (p99) per ton.[12] This is skewness is another reason that we log the dependent variable.

Figure 1 shows the mean cost per ton (log) for the CFC projects approved by the Executive Committee in each meeting. In that figure only, the project cost values are adjusted for US dollar inflation before calculating the mean cost per ton. The bivariate correlation coefficient between that version of the dependent variable and meeting number is -.41. When project cost is unadjusted for inflation it is -.21.

[11] We also tested realised cost as the dependent variable. Realized cost was on average 1.7 percent greater than approved cost. Using the realized cost did not substantially alter our results.

[12] The least expensive project (US $24 per ton) was approved in 2005 for implementation at a cosmetics manufacturing facility in Cote d'Ivoire (project number IVC/ARS/46/INV/23), industry sector classification 'Aerosol'. The median project (US $7,472 per ton) was approved in 1998 for implementation at a facility in Brazil that manufactured rigid foam for packaging electronics and dentistry products (BRA/FOA/25/INV/110), industry sector 'Foam'. The most expensive project (US $750,000 per ton) was approved in 1996 for implementation at a facility in China that manufactured compressor units for household refrigerators (CPR/REF/20/INV/185), industry sector 'Refrigeration'.

Figure 1: Mean cost per ton of CFC elimination projects by Executive Committee meeting



4.2. Independent variable

The three experience variables of interest are calculated as shown in Equation 3. For the 'geography' experience variable, g indexes 138 countries. For the 'industry' experience variable, j indexes 11 industry sectors.[13] For the 'scale' experience variable, s indexes five scale bands.[14] All three experience variables have a minimum value of zero. All three experience variables are calculated from the full set of CFC elimination projects and exclude none of them. All three experience variables include all task iterations up to an including the meeting before the current meeting.

[13] For example Aerosol, Desiccant, Foam, Refrigerant, and Solvent.

[14] These are calculated as quintiles from the quantity of CFC eliminated by each project (range .1 to 6,850 tons).

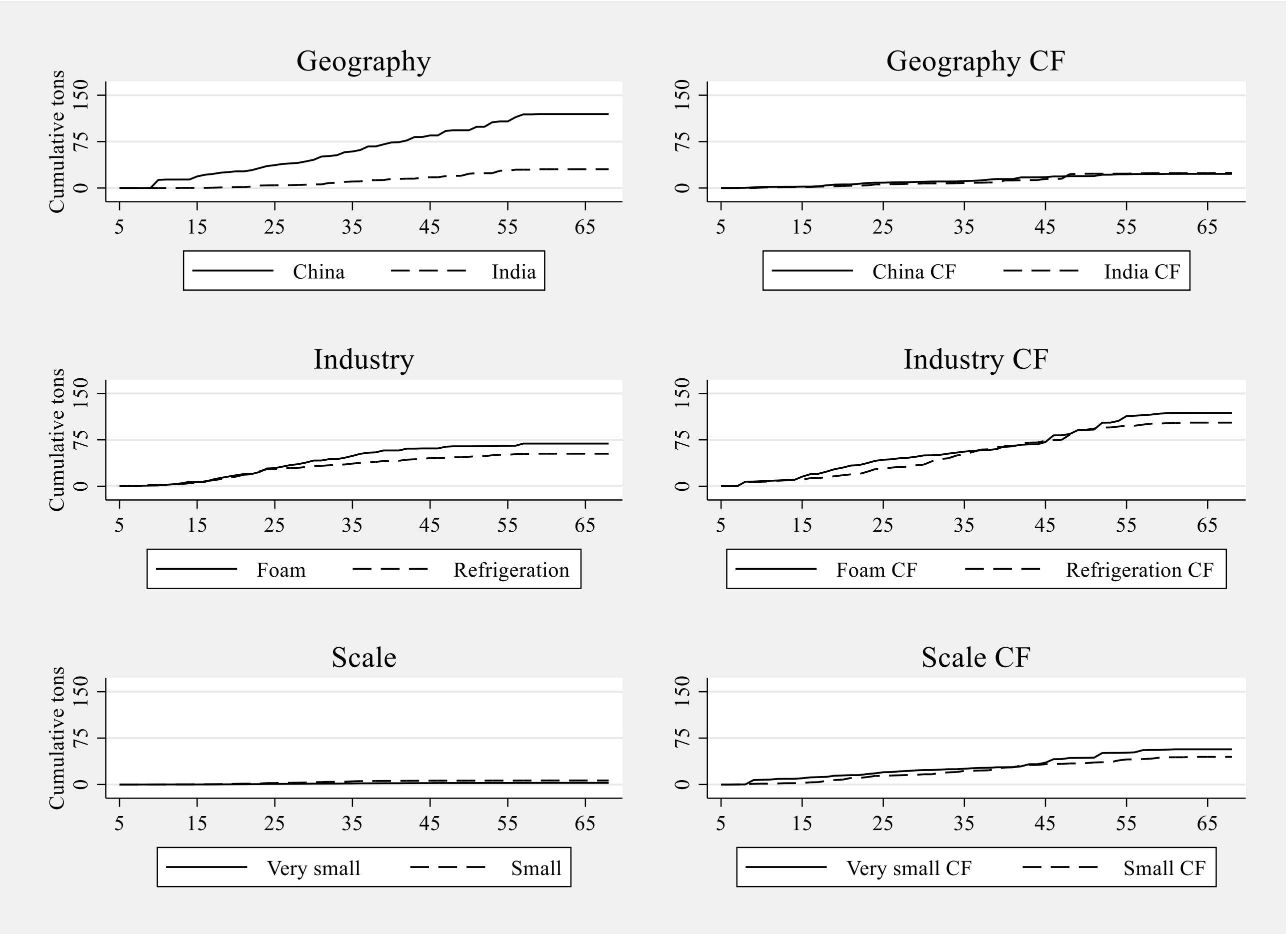
Most empirical research on the performance-experience relationship evaluates whether one or more experience variable coefficients are significantly different from zero. We are also interested in that, but we are also interested in whether the coefficients are significantly different from one another. This tests whether underlying variables are measuring something different, what we hypothesize to be learning progressions.

In comparing the experience coefficients to one another, we were concerned that that the number of categories in the underlying experience variables could confound the comparison of interest, that of learning progressions. To address this we created three counterfactual experience variables, each designed to replicate the category structure of the 'real' experience variable. This approach is inspired by the synthetic control group literature (Abadie, 2021). The real and counterfactual geography experience variables are an example. In the real geography variable, there is an 'Albania' category, and it contains five projects. We constructed the counterfactual geography variable to also contains a category with five projects, except that the projects that comprise it are drawn from the full pool of projects at random. The real geography variable also contains an 'Algeria' category with 25 projects, and the counterfactual geography variable also contains a category of the same size, but with projects drawn randomly from the remaining ones. This is how all 138 categories of the counterfactual geography variable were populated. The projects in the counterfactual experience variables accumulate as in Equation 4, just like the real experience variables.

Figure 2 shows how the level of the geography, industry, and scale band experience variables evolves over time. Only the two categories of each variable with the largest number of projects are shown. The left column shows that at the time of almost any given meeting, the Fund had different levels of experience available to it depending on the category type used to measure experience.

For example, at the time of the Executive Committee's 45th meeting, in April 2005, the Fund had about 60,000 more tons of experience with China projects than with India projects. At the time of that same meeting, it has about 15,000 more tons of experience with foam than with refrigeration projects. The difference also extends to the categories of scale experience although this is difficult to read because although the very small and small categories of that variable had many projects, the cumulative tons of experience in each was relatively small.

Figure 2: Experience measured by channel and category



Note: Cumulative tons in thousands, CF denotes counterfactual version.

4.3. Control variables

The most important control variables are 60 indicator variables for each Executive Committee meeting (the fifth meeting is the omitted category). These variables are intended to control for all time-related influences on the Fund's performance other than the Fund's own internal experience.

These influences are potentially many. One is technological change in CFC elimination technology that occurred exogenously to the Fund, but which may have pushed down the cost of pollution elimination projects by expanding the range and quality of technologies available (Andersen and Sarma, 2012; Decanio and Norman, 2005; Nordhaus, 2014). Another is the adoption of general purpose technologies by the Fund itself, such as digital information and communication technologies, which may have improved operating efficiency and reduced overhead costs (Ashworth et al., 2004). Another is the effect of experience gained by the Fund in approving projects that eliminated substances other than CFC (experience arising from 'non-focal' activities (Argote et al., 2021; Lapré and Tsikriktsis, 2006; Pisano et al., 2001)). Others include personnel changes at the Fund that may have influenced organizational performance (Huckman and Pisano, 2006; Taylor and Greve, 2006), temporal mismatches between the level of personnel and the volume of projects proposals needing approval (Adler, 1990), the level of donor country contributions available to be disbursed, US dollar inflation, and movements in the exchange rate between the US dollar and the currency of country where the project was implemented.

Another control variable is included to account for the size of the project. This is measured as the number of months that the Executive Committee estimated that the project would take to complete. This variable accounts for the extraneous effect on the cost of eliminating CFC of project economies of scale (Arce, 2014; Gruber, 1992).

Another control variable is included to account for whether the project also eliminated an ODS other than CFC. Projects eliminating multiple chemicals might have been more expensive because they were more complex or because the non-CFC ODS was more expensive to eliminate. This variable takes a 1 if the project eliminated one or more ODS additional to CFC (52 projects).

Several dozen additional indicator variables are also included in the regressions as robustness checks. These correspond to the categories that the projects are measured to accumulate over. The rationale for including these is discussed in the results section.

Table 3 (appendix) shows bivariate correlations for all variables except the meeting indicator variables. It shows the expected negative relationship between project cost per ton of eliminating CFC and each 'real' experience variable. It also shows a positive relationship between each pair of real and counterfactual experience variables.

5. Results

Specifications 2, 3 and 4 of Table 1 show that the magnitudes of the coefficients on the geography, industry and experience variables are different. The coefficient on scale experience is four times larger than that on geography experience. The only thing that changes across specifications is the type of category that the projects are measured to accumulate over. Nothing else changes: not the set of projects themselves, the control variables in the model, or the estimation procedure.

These differences are not due to gross omitted variable bias because all specifications in Table 1 include all of the control variables. Specification 1 of Table 1, which includes the control variables only, shows that they explain about 8 percent of overall variation in the dependent variable. The coefficients on the 60 meeting indicator variables are omitted for presentation but these account for almost the entire 8 percent. Of the 60 meeting indicators, 56 have negative coefficients and 35 have negative coefficients that are significant at the 5 percent level (see Figure 4, appendix). This means that relative to the time of the Executive Committee's fifth meeting (November 1992 - the base category) the Fund's performance was better in most subsequent meetings. We interpret this to mean that the Fund's performance benefitted from time-based forces other than the Fund's own internal experience.

Table 1: Fund performance and experience

	Dependent variable: US $ per ton of CFC eliminated				
	(1)	(2)	(3)	(4)	(5)
Experience, accumulating by:					
Geography categories (gt)		$-.07^{***}$			$-.01$
		(.01)			(.01)
Industry categories (jt)			$-.12^{***}$		$-.11^{***}$
			(.02)		(.01)
Scale categories (st)				$-.30^{***}$	$-.29^{***}$
				(.01)	(.01)
Project length - months	$.19^{***}$	$.18^{***}$	$.23^{***}$	$.21^{***}$	$.24^{***}$
	(.05)	(.05)	(.05)	(.04)	(.04)
Multi-pollutant project	$.65^{***}$	$.73^{***}$	$.49^{***}$	$.64^{***}$	$.50^{***}$
	(.13)	(.13)	(.13)	(.11)	(.11)
N projects	2,440	2,440	2,440	2,440	2,440
Meeting controls	Y	Y	Y	Y	Y
Adj R^2	.08	.12	.10	.32	.33
Adj R^2 change from (1)	-	.04	.02	.23	.25
Learning rate	-	5	8	19	26

Note: Unit of observation is the CFC elimination project. Coefficients on meeting indicator variables omitted for presentation (see Figure 4, appendix). Dependent variable and experience variables in logs; experience variable coefficients interpretable as elasticities. Heteroscedasticity-corrected standard errors in parentheses. Learning rate interpretation: unit cost roughly halves for an X-fold increase in experience.

Evidence of the difference between the experience variables is also apparent in the share of overall variation in the dependent variable that each one explains. The adjusted R^2 values in Table 1 show that geography experience variable explains 12 percent of overall variation, industry experience 10 percent, and scale experience 32 percent. Here the explanatory power of scale experience is three times that of industry experience.

Similar differences emerge in the *difference* in the adjusted R^2 value for each of those specifications (2, 3 and 4) relative to specification 1. Specification 1 contains no experience variable at all. The 'Adj R^2 change from 1' line in Table 1 shows that adding geography experience and changing nothing else in the model explains an additional 4 percent of overall variation, adding industry experience explains an additional 2 percent, and adding scale experience explains an additional 23 percent. Here the explanatory power of scale experience is five times greater than that of geography experience.

Our first (and most conservative) hypothesis is that internal experience gained by the Fund through the project approval process contributed to the Fund's performance improvement over time. We evaluate the evidence for this through the significance level of the coefficients on three experience variables in specifications 2, 3, and 4 in Table 1. When the coefficients on these variables are estimated separately, as in those specifications, each is different from zero at the 1 percent level. When the coefficients on these variables are estimated jointly (specification 5), the geography experience coefficient is not significantly different from zero at conventional levels but the industry and scale experience variables are. [15] We interpret this as evidence that internal experience facilitated performance improvement at the Fund.

One interesting result is that the three experience variables do not show signs of collinearity when estimated jointly (in specification 5). One might expect collinearity because each experience variable is measured from the exact same set of projects. In fact, the highest variance inflation factor (VIF) score on any of the three experience variables in specification 5 s 2.37. This is well below the threshold of concern of 10 (Studenmund, 2021). We interpret this to mean that the experience variables are capturing varieties of experience that are to some extent distinct from one another, and cumulative, in their relationship to Fund performance. Further, the adjusted R^2 value in specification 5 shows that including all three variables increases, by 2 percent, the share of

[15] Since we evaluate the hypothesis in terms of coefficient significance level, it matters how the errors are handled in the estimation procedure. In all specifications in Table 1 the errors are robust to violation of the assumption of being identically distributed. We also tested how significance level changes when the errors are clustered on the categories that define the experience variable. For example, when testing geography experience, we clustered the errors on countries. This hardly affects the significance of any of coefficients. These tests are shown in Table 6 (appendix).

overall variation that is explained by experience. On this basis, we calculated the learning rate in specification 5 as the sum of the learning rates implied by each individual experience variable.

The magnitude of the experience variable coefficients in specifications 3, 4 and 5 in Table 1 are not out of line with those reported in prior research (DUTTON; SCHRTAZ). They imply that a 1 percent increase in experience associates with a fall in the cost of eliminating CFC (improvement in Fund performance) of between -.07 and -.30 percent, approximately. Transformed into learning rates, those estimates imply that unit cost fell by about 50 percent with each increase in the Fund's own internal experience of between 5 and 19 percent.

Our second hypothesis is that if learning progressions are relevant to the performance-experience relationship, then the coefficients on the three experience variables should be *different from one another* (rather than different from zero as in hypothesis 1).

We were concerned that a formal comparison of the magnitude of any two experience coefficients might be invalid if the coefficients themselves were strongly mis-estimated. To address this concern, we added to the model indicator variables that correspond to all of the category types that we measure experience to accumulate over, except for those for the focal category type. For example, when testing the geography experience, we added full sets of indicator variables for the industry and scale categories.

These tests produced slightly smaller experience variable coefficients (relative to Table 1) when the experience variables were tested separately (specifications 1, 2 and 3 in Table 7, appendix). The *extent* of the difference among them persists. When the experience variables are tested jointly (specification 4), the coefficients are generally not significant. That specification includes indicator variables that correspond to *all* of the category types. We choose to proceed with the results in Table 1 and not those in Table 7. This is because the variation in project unit cost (Fund performance) that the category indicator variables capture is precisely the basis for expecting learning progressions to be relevant. This is to say that if learning progressions exist, this is because 'more' learning happens across similar project types than dissimilar project types. The similarities between project types already enter the model in the way that projects accumulate into experience within categories.

We evaluated the evidence for hypothesis two by comparing the experience coefficients in specification 5 of Table 1 to one another. We did this using F-tests where the null hypothesis is that the coefficients for a given pair of experience variables are equal. We ran the F-test on the three possible pairs of coefficients in that specification: geography and industry, geography and scale, and industry and scale. All three tests reject the null that the coefficients in the pair are equal. The test and p-values for the pairs are, respectively: 45.32 (.000), 357.86 (.000) and 88.08 (.000). We interpret this as evidence that the experience variable coefficients are different to one another. We infer that this difference is due to the category types that the projects are measured to accumulate over, in line with our learning progressions hypothesis, in part because this the only thing that differs in the underlying variables.

Our third hypothesis is that the coefficients on the ‘real’ experience variables are different from those on the counterfactual experience variables. We first examined the evidence for this hypothesis by evaluating whether the coefficients on the counterfactual experience variables in specifications 1, 3 and 5 of Table 2 are significantly different from zero. They are not. This is notable because, if any of the counterfactual variable coefficients *had* been significantly different from zero, this would have meant that projects that were randomly selected into categories *also accumulated into experience that improved Fund performance*. This would have invalidated the logical basis for our belief that the real category types that we observe in the data capture similarities among projects that facilitated cross-project learning.

We also evaluated hypothesis three by comparing the adjusted R^2 value in the specification with the real experience variable to that in the specification with the counterfactual version. In all three pairs of specifications in Table 2 (1 and 2 for geography, 3 and 4 for industry, 5 and 6 for scale) the adjusted R^2 is greater for the ‘real’ specification.

We also tested for systematic differences between the coefficients in the counterfactual and real specifications, using a Chow test. We did this for each pair of specifications (geography, industry, scale as before). The null hypothesis in this test was that the pair of coefficients is equal. The tests rejected the null of equality for all three pairs of coefficients. All three tests rejected the null of equality at the 1 percent level ($p = .000$). The ‘Chow’ line of Table 1 shows the test statistic.

Strictly speaking, a Chow test is appropriate when the aim is to compare coefficients for the same variable across models fitted to different subsets of the data (Clogg et al., 1995). (It would be appropriate for example if the aim was to estimate the effect of worker age on wage by gender; the age-wage relationship would be estimated for the men in the data, then for the women, then the coefficients would be compared across models). We were concerned that our application of the test might not be appropriate because we are comparing coefficients for *different variables* across models fitted to the *same data*.

To address this, we compared the coefficients another way, using a Durbin-Wu-Hausman test. The Durbin-Wu-Hausman test evaluates the null hypothesis that every pair of coefficients between two given specifications is the same (not just that the coefficients for a single variable are the same). In our situation, we consider that the specification with counterfactual experience is the partial, incomplete specification, and that the specification with real experience is the full, complete specification. This test rejected the null of equality for all three pairs of specifications. All three test values were significant at the 1 percent level of significance. The 'Hausman' line of Table 2 shows the test value.

Table 2: Comparison of counterfactual and real experience

	Dependent variable: US $ per ton of CFC eliminated					
	(1)	(2)	(3)	(4)	(5)	(6)
Experience accumulates by:						
Geography categories CF	-.01					
	(.01)					
Geography categories (gt)		-.07***				
		(.01)				
Industry categories CF			.01			
			(.01)			
Industry categories (jt)				-.12***		
				(.02)		
Scale categories CF					-.04	
					(.04)	
Scale categories (st)						-.30***
						(.01)
N projects	2,440	2,440	2,440	2,440	2,440	2,440
Meeting controls	Y	Y	Y	Y	Y	Y
Adj R^2	.08	.12	.08	.10	.08	.32
Adj R^2 change from CF	-	.04		.02	-	.23
Learning rate	1	5	-1	8	3	19
Chow	-	98.9	-	46.5	-	608.5
Hausman	-	28.2	-	31.3	-	31.8

Note: Basic model set up is same as Table 1 (see that note). These regressions compare the coefficients on the real and counterfactual versions of each experience variable. All control variables omitted for presentation. All experience variable coefficients interpretable as elasticities. Learning rate interpretation: unit cost roughly halves for an X-fold increase in experience. Chow: test of no difference between the coefficients on the real and counterfactual experience variable pair. Test values shown; all significant at the 1 percent level. Hausman: test of no difference between all coefficients in the specification with the real and counterfactual experience variables. Test values shown; all significant at 1 percent level.

We interpret the results of the Durbin-Wu-Hausman tests as evidence that each specification containing the real version of the experience variable is different to that containing the counterfactual version. We also see substantially greater overall explanatory power in the real experience specifications relative to the counterfactual ones.

6. Discussion and conclusions

Learning progressions research emerged in the pedagogical and education sciences and have never been considered or tested in an organizational workplace context as we have here. The research reported here is the first to our knowledge to propose that learning progressions are present in the way that organizations accumulate experience in practice, as well as in the way that researchers represent that experience empirically. Our findings support the idea that multiple varieties of experience are relevant for organizational performance and performance improvement (Reagans et al., 2005) and that different varieties of experience have different 'saliences' for those outcomes.

The idea that learning progressions play a role in organizational performance improvement is the most original contribution of this research. It is also the most speculative. The least speculative contribution of this research is the wide differences it has shown across experience coefficients that is due purely to the category type over which task iterations are measured to accumulate. These differences suggests that a substantial part of the variation in learning rate estimates across organizations, reported in a large body of empirical research (Argote et al., 2021; Dutton and Thomas, 1984; McDonald and Schrattenholzer, 2001), may be due to analysts' measurement decisions.

This research demonstrated those differences partly by comparing 'real' experience variable coefficients to zero, partly by experience variable coefficients to one another, and partly by comparing the 'real' experience variable coefficients to counterfactual versions. To our knowledge, this third approach has not been used in used in organizational learning research. That novel approach showed that 'experience' arising from task iterations that accumulate over randomly-defined category types does not associate significantly with organizational performance. This indirectly supports our learning progressions hypothesis.

One limitation of this research lies in the limits that observational data placed on the hypotheses that we were able to test. We were able to manipulate the category type that the task iterations that the organization performed accumulated over. We were not able to manipulate either a) the intensity of the underlying task iterations themselves (tons of CFC eliminated in this case) or b) the chronological order in which different project types occurred in reality.

More fully experimental research designs could accommodate both of these factors in testing whether and how learning progressions play a role in the organizational performance-experience relationship. Here is a hypothetical field experiment in this vein. Two groups of new employees are hired, at separate offices of the same organization. One group is administered a version of a training program where the training task iterations are of uniform intensity (say duration) and administered in random order with respect to some characteristic of the task (say intrinsic difficulty). The other group is exposed to a version of the training program where the task iterations are also uniform in duration but ordered from easiest to hardest. The groups are then evaluated in terms of their ability to perform a task or tasks and their performance is compared. A design like this would overcome the limits we recognize in our own research design.

The international community has chosen multi-lateral financing organizations like the one we studied here as the vehicle for channelling billions of US dollars in public money into climate mitigation and adaptation projects. Our research demonstrates how the performance of such an organization evolved over time, shows that performance improvement is possible, and identifies various sources of performance improvement. The Fund's own internal experience explained up to three times more of the variation in its performance than all external, time-based forces. These findings inform the management and expectations of organizations like the GCF. Younger organizations like the GCF stand to benefit more from the experience of an organization like the Fund than from their own internal experience because their own internal experience is relatively shallow (Aranda et al., 2017).

The strongest gains in organizational performance in our study associated with scale-related experience. Why would scale-related experience associate with faster improvement than geography- or industry-related experience? One possibility is that two given 'small' projects (task iterations) are more similar to one another than two given ‘Argentina’ projects (task iterations). This implies that when scale-related task iterations accumulate into experience, the resulting experience stock is more homogeneous, and more specialized, than geography-related experience. Specialized experience improves performance faster than less specialized experience, particularly in an organization whose mandate is to implement a narrow legal mandate. Specialized experience might improve performance slower in a organization whose success was connected to creating new products or to competing with other organizations through creativity (Taylor and Greve, 2006).

References

Abadie, A., 2021. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. Journal of Economic Literature 59, 391–425. https://doi.org/10.1257/jel.20191450

Adler, P.S., 1990. Shared Learning. Management Science 36, 938–957. https://doi.org/10.1287/mnsc.36.8.938

Álvarez Valdivia, I.M., Lafuente Martínez, M., 2019. Improving preservice teachers' scientific argumentative writing through epistemic practices: a learning progression approach. Journal of Education for Teaching 45, 169–185. https://doi.org/10.1080/02607476.2018.1548172

Andersen, S.O., Sarma, K.M., 2012. Protecting the Ozone Layer, 0 ed. Routledge. https://doi.org/10.4324/9781849772266

Aranda, C., Arellano, J., Davila, A., 2017. Organizational Learning in Target Setting. AMJ 60, 1189–1211. https://doi.org/10.5465/amj.2014.0897

Arce, D.G., 2014. Experience, Learning, and Returns to Scale. Southern Economic Journal 80, 938–947. https://doi.org/10.4284/0038-4038-2013.088

Argote, L., 2013. Organizational Learning. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4614-5251-5

Argote, L., Lee, S., Park, J., 2021. Organizational learning processes and outcomes: Major findings and future research directions. Management Science 67, 5399–5429. https://doi.org/10.1287/mnsc.2020.3693

Ashworth, M., Mukhopadhyay, T., Argote, L., 2004. Information technology and organizational learning: An empirical analysis. ICIS 2004 Proceedings 38.

Barnett, W.P., Greve, H.R., Park, D.Y., 1994. An Evolutionary Model of Organizational Performance. Strat. Mgmt. J. 15, 11–28. https://doi.org/10.1002/smj.4250150903

Baum, J.A.C., Ingram, P., 1998. Survival-Enhancing Learning in the Manhattan Hotel Industry, 1898–1980. Management Science 44, 996–1016. https://doi.org/10.1287/mnsc.44.7.996

Berland, L.K., McNeill, K.L., 2010. A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. Science Education 94, 765–793. https://doi.org/10.1002/sce.20402

Bowman, M., Minas, S., 2019. Resilience through interlinkage: the green climate fund and climate finance governance. Climate Policy 19, 342–353. https://doi.org/10.1080/14693062.2018.1513358

Chen, L., Tsai, J., Kao, Y., Wu, Y., 2019. Investigating the learning performances between sequence- and context-based teaching designs for virtual reality (VR)-based machine tool operation training. Computer Applications in Engineering Education 27, 1043–1063. https://doi.org/10.1002/cae.22133

Clogg, C.C., Petkova, E., Haritou, A., 1995. Statistical Methods for Comparing Regression Coefficients Between Models. American Journal of Sociology 100, 1261–1293. https://doi.org/10.1086/230638

Cohen, M.D., Bacdayan, P., 1994. Organizational Routines Are Stored as Procedural Memory: Evidence from a Laboratory Study. Organization Science 5, 554–568. https://doi.org/10.1287/orsc.5.4.554

Cui, L., Huang, Y., 2018. Exploring the Schemes for Green Climate Fund Financing: International Lessons. World Development 101, 173–187. https://doi.org/10.1016/j.worlddev.2017.08.009

Cui, L., Sun, Y., Song, M., Zhu, L., 2020. Co-financing in the green climate fund: lessons from the global environment facility. Climate Policy 20, 95–108. https://doi.org/10.1080/14693062.2019.1690968

Darr, E.D., Argote, L., Epple, D., 1995. The Acquisition, Transfer, and Depreciation of Knowledge in Service Organizations: Productivity in Franchises. Management Science 41, 1750–1762. https://doi.org/10.1287/mnsc.41.11.1750

Decanio, S.J., Norman, C.S., 2005. Economics of the “critical use” of methyl bromide under the Montreal Protocol. Contemporary Economic Policy 23, 376–393. https://doi.org/10.1093/cep/byi028

Dosi, G., Grazzi, M., Mathew, N., 2017. The cost-quantity relations and the diverse patterns of “learning by doing”: Evidence from India. Research Policy 46, 1873–1886. https://doi.org/10.1016/j.respol.2017.09.005

Drexler, A., Schoar, A., 2014. Do Relationships Matter? Evidence from Loan Officer Turnover. Management Science 60, 2722–2736. https://doi.org/10.1287/mnsc.2014.1957

Duschl, R., Maeng, S., Sezen, A., 2011. Learning progressions and teaching sequences: a review and analysis. Studies in Science Education 47, 123–182. https://doi.org/10.1080/03057267.2011.604476

Duschl, R.A., 2019. Learning progressions: framing and designing coherent sequences for STEM education. Disciplinary and Interdisciplinary Science Education Research 1, 4. https://doi.org/10.1186/s43031-019-0005-x

Dutton, J.M., Thomas, A., 1984. Treating Progress Functions as a Managerial Opportunity. Academy of Management Review 9, 235–247. https://doi.org/10.5465/amr.1984.4277639

Edenhofer, O., Pichs-Madruga, R., Sokona, Y., 2014. Mitigation of Climate Change: Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change, New York.

Egelman, C.D., Epple, D., Argote, L., Fuchs, E.R.H., 2017. Learning by Doing in Multiproduct Manufacturing: Variety, Customizations, and Overlapping Product Generations. Management Science 63, 405–423. https://doi.org/10.1287/mnsc.2015.2352

Epple, D., Argote, L., Devadas, R., 1991. Organizational Learning Curves: A Method for Investigating Intra-Plant Transfer of Knowledge Acquired Through Learning by Doing. Organization Science 2, 58–70. https://doi.org/10.1287/orsc.2.1.58

Fiol, C.M., Lyles, M.A., 1985. Organizational Learning. Academy of Management Review 10, 803–813. https://doi.org/10.5465/amr.1985.4279103

Ford, M., 2008. Disciplinary authority and accountability in scientific practice and learning. Sci. Ed. 92, 404–423. https://doi.org/10.1002/sce.20263

Gruber, H., 1992. The learning curve in the production of semiconductor memory chips. Applied Economics 24, 885–894. https://doi.org/10.1080/00036849200000056

Hammer, L.B., Wan, W.H., Brockwood, K.J., Bodner, T., Mohr, C.D., 2019. Supervisor support training effects on veteran health and work outcomes in the civilian workplace. Journal of Applied Psychology 104, 52–69. https://doi.org/10.1037/apl0000354

Harris, S.F., 2017. Learning curves: An alternative analysis. Air Force Institute of Technology, Air University, Maxwell, Alabama.

Haunschild, P.R., Sullivan, B.N., 2002. Learning from Complexity: Effects of Prior Accidents and Incidents on Airlines’ Learning. Administrative Science Quarterly 47, 609–643. https://doi.org/10.2307/3094911

Huckman, R.S., Pisano, G.P., 2006. The firm specificity of individual performance: Evidence from cardiac surgery. Management Science 52, 473–488.

Jaber, M.Y., 2011. Learning Curves: Theory, Models, and Applications. Taylor & Francis Group, Boca Raton.

Jin, H., Mikeska, J.N., Hokayem, H., Mavronikolas, E., 2019. Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature. Science Education 103, 1206–1234. https://doi.org/10.1002/sce.21525

KC, D., Staats, B.R., Gino, F., 2013. Learning from My Success and from Others' Failure: Evidence from Minimally Invasive Cardiac Surgery. Management Science 59, 2435–2449. https://doi.org/10.1287/mnsc.2013.1720
Krueger, A., Rouse, C., 1998. The Effect of Workplace Education on Earnings, Turnover, and Job Performance. Journal of Labor Economics 16, 61–94.
Lapré, M.A., Tsikriktsis, N., 2006. Organizational Learning Curves for Customer Dissatisfaction: Heterogeneity Across Airlines. Management Science 52, 352–366. https://doi.org/10.1287/mnsc.1050.0462
Lattanzio, R.K., 2017. International Climate Change Financing: The Green Climate Fund (GCF). Congressional Research Service, Washington, DC.
Lee, N., Landers, C., Matthews, S., 2023. Concessional Climate Finance: Is the MDB Architecture Working? (Policy Paper 287). Center for Global Development, Washington, DC.
Levitt, B., March, J.G., 1988. Organizational Learning. Annual Review of Sociology 14, 319–338. https://doi.org/10.1146/annurev.so.14.080188.001535
Liu, L., Jackson, T., 2019. A Recent Review of Learning Progressions in Science: Gaps and Shifts. The Educational Review, USA 3. https://doi.org/10.26855/er.2019.09.001
Markandya, A., Antimiani, A., Costantini, V., Martini, C., Palma, A., Tommasino, M.C., 2015. Analyzing Trade-offs in International Climate Policy Options: The Case of the Green Climate Fund. World Development 74, 93–107. https://doi.org/10.1016/j.worlddev.2015.04.013
McDonald, A., Schrattenholzer, L., 2001. Learning rates for energy technologies. Energy Policy 29, 255–261. https://doi.org/10.1016/S0301-4215(00)00122-1
Meyers, N.M., Nulty, D.D., 2009. How to use (five) curriculum design principles to align authentic learning environments, assessment, students' approaches to thinking and learning outcomes. Assessment & Evaluation in Higher Education 34, 565–577. https://doi.org/10.1080/02602930802226502
Nadler, J., Thompson, L., Boven, L.V., 2003. Learning Negotiation Skills: Four Models of Knowledge Creation and Transfer. Management Science 49, 529–540. https://doi.org/10.1287/mnsc.49.4.529.14431
National Research Council, 2007. Taking Science to School: Learning and Teaching Science in Grades K-8. National Academies Press, Washington, D.C. https://doi.org/10.17226/11625
National Research Council, 2004. How People Learn: Brain, Mind, Experience and School. National Academy Press, Wahsington, DC.
Nordhaus, W.D., 2014. The Perils of the Learning Model for Modeling Endogenous Technological Change. The Energy Journal 35. https://doi.org/10.5547/01956574.35.1.1
OECD, 2012. OECD Environmental Outlook to 2050. Organization for Economic Cooperation and Development. https://doi.org/10.1787/9789264122246-en
Perry, W.G., 1997. Cognitive and ethical growth: The making of meaning, in: Arnold, K., Carreiro-King, I. (Eds.), College Student Development and Academic Life: Psychological, Intellectual, Social and Moral Issues. Routledge, New York, pp. 48–88.
Pisano, G.P., Bohmer, R.M.J., Edmondson, A.C., 2001. Organizational Differences in Rates of Learning: Evidence from the Adoption of Minimally Invasive Cardiac Surgery. Management Science 47, 752–768. https://doi.org/10.1287/mnsc.47.6.752.9811
Reagans, R., Argote, L., Brooks, D., 2005. Individual Experience and Experience Working Together: Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together. Management Science 51, 869–881. https://doi.org/10.1287/mnsc.1050.0366
Reed, C., Wolfson, A., 2021. Are Learning Progressions a Useful Pedagogical Tool for Instructors? Journal of College Science Teaching 51, 32–40.

Salas, E., Tannenbaum, S.I., Kraiger, K., Smith-Jentsch, K.A., 2012. The Science of Training and Development in Organizations. Psychological Science in the Public Interest 13, 74–101. https://doi.org/10.1177/1529100612436661
Stevenson, D.L., Schiller, K.S., Schneider, B., 1994. Sequences of Opportunity for Learning. Sociology of Education 67, 184–198.
Studenmund, A., 2021. Using econometrics: A practical guide. Pearson, New York.
Taguma, M., Gabriel, F., Lim, M., 2019. Future of Educational and Skills 2030: Curriculum analysis - A Synthesis of Research on Learning Trajectories / Progressions in Mathematics. OECD, Paris.
Tang, T., Popp, D., 2016. The Learning Process and Technological Change in Wind Power: Evidence from China’s CDM Wind Projects. Journal of Policy Analysis and Management 35, 195–222. https://doi.org/10.1002/pam.21879
Taylor, A., Greve, H.R., 2006. Superman or the Fantastic Four? knowledge combination And experience in Innovative Teams. AMJ 49, 723–740. https://doi.org/10.5465/amj.2006.22083029
The Multilateral Fund Secretariat, 2022a. Executive Committee Primer - 2022. Multilateral Fund for the Implementation of the Montreal Protocol, Montreal.
The Multilateral Fund Secretariat, 2022b. Policies, Procedures, Guidelines and Criteria (As at July 2022). Multilateral Fund for the Implementation of the Montreal Protocol:, Montreal.
Thompson, P., 2012. The Relationship between Unit Cost and Cumulative Quantity and the Evidence for Organizational Learning-by-Doing. Journal of Economic Perspectives 26, 203–224. https://doi.org/10.1257/jep.26.3.203
Thornton, R.A., Thompson, P., 2001. Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding. American Economic Review 91, 1350–1368. https://doi.org/10.1257/aer.91.5.1350
UNEP, 2000. The Montreal Protocol on Substances that Deplete the Ozone Layer.
Wiersma, E., 2007. Conditions That Shape the Learning Curve: Factors That Increase the Ability and Opportunity to Learn. Management Science 53, 1903–1915. https://doi.org/10.1287/mnsc.1070.0733
William, L.B., Harter, N., 1899. Studies on the telegraphic language: The acquisition of a hierarchy of habits. Psychological Review 6, 345–375. https://doi.org/10.1037/h0073117
World Economic Forum, 2022. The Global Risks Report 2022, 17th Edition - Insight Report.
Wright, T.P., 1936. Factors Affecting the Cost of Airplanes. Journal of the Aeronautical Sciences 3, 122–128. https://doi.org/10.2514/8.155
Yelle, L.E., 1979. THE LEARNING CURVE: HISTORICAL REVIEW AND COMPREHENSIVE SURVEY. Decision Sciences 10, 302–328. https://doi.org/10.1111/j.1540-5915.1979.tb00026.x

Appendix

Figure 3: Countries where CFC elimination projects were implemented

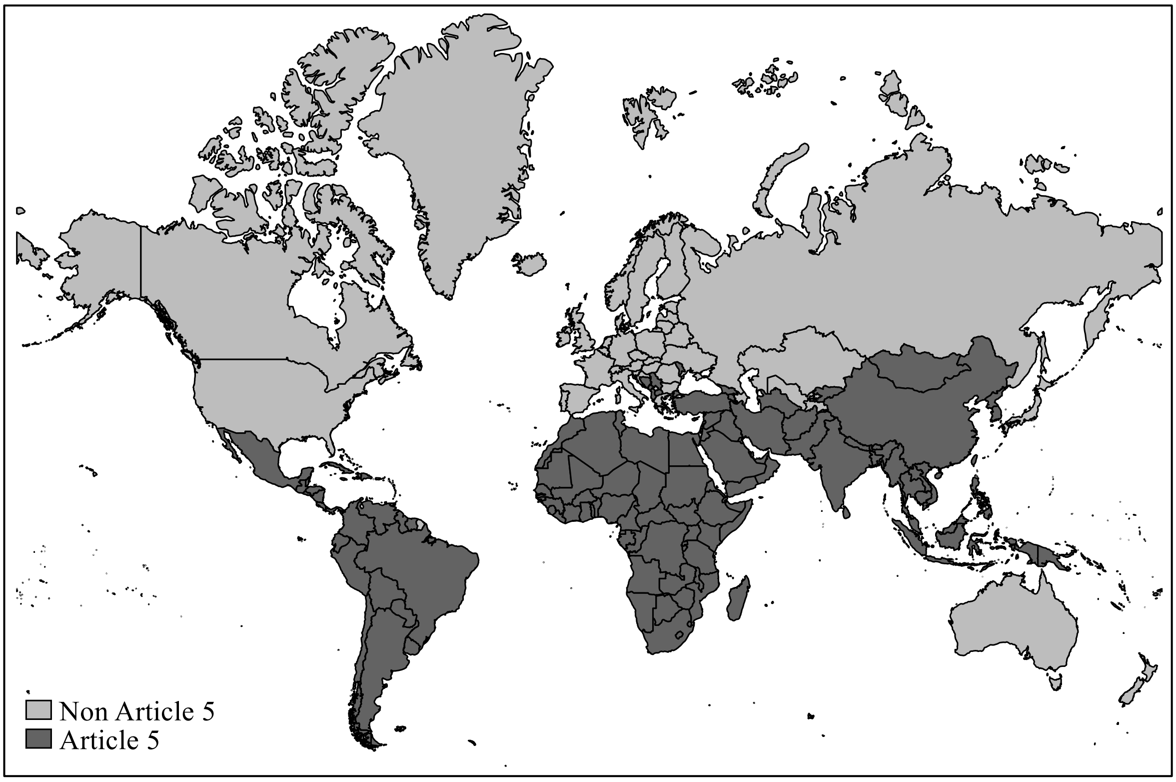


Table 3: CFC consumption level and change by country group

	Tons, 1991	Tons, 2012	Change, tons, 1991-2012	Change, %, 1991- 2012
CFC consumption by				
Low- and middle-income countries	117,462	32	-117,430	-97.3
High-income countries	498,416	-1,446	-499,862	-100.0

Note: Data as reported by individual countries to the UNEP Ozone Secretariat.

Table 4: Bivariate correlations between project unit cost and experience variables

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	Project unit cost, US $ / ton	1.00								
(2)	Geography experience – real	-.20***	1.00							
(3)	Industry experience – real	-.15***	.15***	1.00						
(4)	Scale experience – real	-.46***	.39***	.35***	1.00					
(5)	Geography experience – CF	-.05**	.09***	.34***	.28***	1.00				
(6)	Industry experience – CF	-.04*	.14***	.43***	.36***	.35***	1.00			
(7)	Scale experience – CF	-.08***	.19***	.61***	.53***	.50***	.63***	1.00		
(8)	Project length, months	.10***	.03	.18***	.01	.11***	.04*	.08***	1.00	
(9)	Multi-pollutant project	.10***	.04	-.15***	.01	.03	-.00	-.01	-.08***	1.00

Note: Bivariate correlations, N = 2,440. All variables in logs. *, **, and *** denote $p < .05$, .01, and .001.

Figure 4: Coefficients on Executive Committee meeting indicator variables

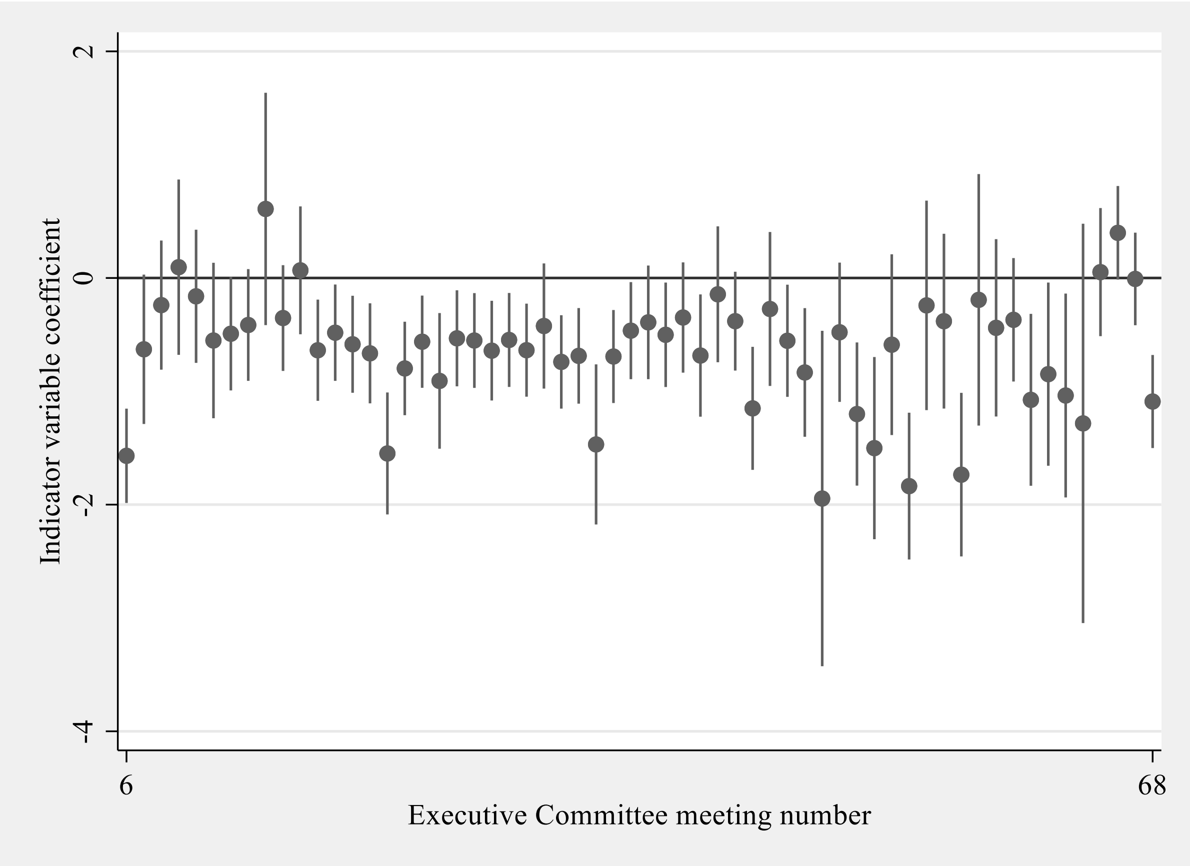


Table 5: Errors clustered on the categories of each experience variable

	1	2	3
Experience accumulates by:			
Geography categories (gt)	-.07***		
	(.01)		
Industry categories (jt)		-.12**	
		(.03)	
Scale categories (st)			-.30**
			(.05)
N projects	2,440	2,440	2,440
Adj R^2	.12	.10	.32
Learning rate	5	8	19

Note: These specifications are almost exactly the same as in specifications 2, 3, and 4 of Table 1 (see that note). These specifications test how clustered errors change the experience variable significance levels with respect to hypothesis 1. Errors are clustered on the categories of the corresponding experience variable. This increases the size of the errors relative to Table 1, but all experience coefficients remain significant.

Table 6: Tests for omitted variable bias

	(1)	(2)	(3)	(4)
Experience accumulates by:				
Geography categories (gt)	.02**			-.00
	(.01)			(.02)
Industry categories (jt)		-.11***		-.03
		(.02)		(.04)
Scale categories (st)			-.27***	-.16*
			(.01)	(.08)
N projects	2,440	2,440	2,440	2,440
Meeting controls	Y	Y	Y	Y
Geography category controls	N	Y	Y	Y
Industry category controls	Y	N	Y	Y
Scale category controls	Y	Y	N	Y
Variables in model	72	195	199	205
Adj R^2	.41	.45	.51	.51
Learning rate	-1	7	17	13

Note: Model set up and data similar to Table 1 (see that note). These specifications test how controlling for different project characteristics bears on the magnitude of the experience coefficients and the difference between them.