



**Evidence in
Practice**

Contents

Stakeholder Map
Timeline
Process Diagram
Appendices
Print

Full Report Summary + Findings

Case Studies

View Alphabetically
View Geographically
View by Discipline

MEXICO | EMPLOYMENT

Programa Primer Empleo

A national employment generation program with good intentions and a desire for rapid impact was implemented too quickly to integrate existing and emerging evidence, creating insurmountable barriers.

By María del Mar Gutiérrez





**Programa Primer Empleo
Table of Contents**

Evidence in Practice Introduction	03
Part I: The PPE Story	04
The Design Process	04
Context: 2006 -2007	04
Evidence Used	05
Targeting + Objective	05
Determining the Rules of Operation	06
Trust	06
Including the Private Sector in the Conversation	08
Program Implementation + Monitoring	08
Design Evaluation	10
Civil Service Middle Management, the Federal Government, and the Private Sector	12
Results	15
Part II: Evidence in Practice: Key Themes + Insights	17
The Role of Timing in the Incorporation of Evidence into Practice	17
Evidence Definition, Creation + Use	18
Need to Devote Exclusive Time and Resources to Learn About and Operationalize Evidence	19
Importance of Building Trust and Forging Relationships Among Stakeholders	20
Failure to Learn from Failure	20
Conclusion	21
Part III: Resources	
Stakeholder Map	22
Timeline	23
Process Diagram	24
Appendices	25

Image Credits

Cover: María del Mar Gutiérrez

Page 9: [Bud Ellison](#)

Page 14: [Geraint Rowland](#)

Page 18: [Bud Ellison](#)



The **Evidence In Practice research project at the Yale School of Management**, funded by the William and Flora Hewlett Foundation, was conducted from January 2016 to January 2018 in order to better understand the conditions under which rigorous evidence can be effectively integrated into public policies and non-governmental organization (NGO) practices in the field of international development.

The Evidence in Practice project followed a rigorous methodology comprised of three broad elements: First we conducted an initial round of expert interviews with individuals who have spent a significant portion of their professional lives attempting, researching, or promoting the integration of evidence into development practice, including academics, government officials, foundation program officers, NGO practitioners, and think-tank

directors. Second, we conducted a matched comparison of eight cases of development programs or interventions where rigorous evidence was integrated with varying degrees of effectiveness. This case study is one of the eight produced by the project. The third component, conducted in parallel to the eight case studies, consisted of interviews with prototypical representatives of each of the stakeholder groups, or individuals who could clearly describe the typical experience of enacting a particular stakeholder role. Our synthesis analysis is presented in the accompanying report.¹

Stakeholder Characterization

Based on our research, we have found it useful to think of the flow of evidence into policy and practice as an “ecosystem” in which a set of archetypical stakeholder groups interact. This set of stakeholder categories was described and reinforced by our interviewees throughout the project. While this is not a perfect description (e.g., some organizations fall within more than one stakeholder group and individuals often shift across stakeholder groups or play roles that effectively span categories), it can help frame the conversation to identify the critical roles, incentives, and relationships that animate the

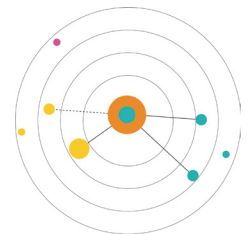
complex relationship between “evidence” and “practice.” These representative stakeholder groups are: **Researchers, Funders, Influencers², Intermediaries, Policymakers, Implementers, and Beneficiaries³**

Each of the cases thus contains a map of the specific organizations (and individuals) that defined its evolution, their structural affiliation to a stakeholder category (in some cases, organizations played more than one formal, structural role), the informal roles that certain individual actors played, as well as the key relationships between these individuals and organizations.

¹ Please see the appendix for a detailed description on Data and Methods.

² While some of our interviewees identified “Influencers” (such as the media, the general public, lobbyists, and influential individuals) as playing an important role in the evidence-to-practice eco-system, this group did not play an explicit role in the narrative of any of the case studies. So we have included the category here, though it does not appear in the stakeholder maps of the individual case studies.

³ We use the term “beneficiaries” to indicate those whom a specific policy or program is intended to help. Different analytic frameworks use various terms to describe this group, including clients, users, recipients, etc.



See the relationship between the stakeholders in this project in the stakeholder map.



Part I: **The PPE Story**

Programa Primer Empleo (PPE) launched during the first months of 2007 and was intended to create incentives for employers to generate new permanent jobs. It consisted of a government subsidy to reduce the costs of hiring, provided through a retroactive, partial reimbursement in the mandatory social security fees paid by employers who hire new workers.⁴ The program was designed by the incoming Presidential administration of Felipe Calderón, funded through the federal budget, and implemented by the Mexican Social Security Institute (the Instituto Mexicano del Seguro Social, or IMSS). The program was voluntary, but contained guidelines and restrictions that defined the eligibility of employers and employees. The decree established that the program would start operations in March of 2007 and end on November 30, 2012. In the 2007 federal budget, the program was assigned \$3 billion Mexican pesos (\$274.5 million USD 2007 equivalent).⁵

The story of Programa Primer Empleo illustrates how political and social pressures can lead policymakers to seek program results and visibility at a much faster speed and much larger scale than the available data—and the uncertainty surrounding initial assumptions—would otherwise dictate. The design, development, and management of Programa Primer Empleo also shows the importance of defining the problem to be addressed by a program in consultation with all the relevant stakeholders, which can only occur in a framework of trust among all the

actors involved. Also, the case highlights the barriers to evidence uptake. Finally, the program's implementation involved challenging choices and presents learning opportunities regarding all these topics.

The Design Process | Context: 2006 -2007

After a bitterly contested campaign season and election, Mr. Felipe Calderón was elected president by a narrow margin of 0.56 percent.⁶ The election was followed by a long, controversial post-electoral process that created an environment of enormous

political pressure. Mr. Calderón, who ran on a platform of job creation (he was the “candidate of employment”), needed quick and visible results, so the transition team started working on an ambitious plan to encourage job creation, which later became Programa Primer Empleo. This process began during the transition period that followed the elections in July 2006 and ramped up after President Calderon's inauguration on December 1, 2006. The program's creation was officially announced in January 2007, less than two months after the new administration began. Programa Primer Empleo became a flagship program of the administration, addressing the “more and better employment” campaign pledge.

⁴ Objective published in the official decree creating the program. Federal Official Gazette. January 23, 2007

⁵ Average 2007 exchange rate \$1.00 USD = \$10.9282 MXN. [Available here.](#)

⁶ “Official Results, Mexico 2006 Presidential Elections,” Journal of Democracy. [Access here.](#)



The story of Programa Primer Empleo illustrates how political and social pressures can lead policymakers to seek program results and visibility at a much faster speed and much larger scale than the available data would otherwise dictate.

Evidence Used

In 2003, a theoretical economics paper titled “A Model of the Mexican Labor Market with and without Social Security, 2003”, modeled the elasticity of employer fees paid to IMSS and their effect on the uninsured (informal) labor market. The study argued that a 6% reduction in the fees would translate to as many as 300,000 jobs in the formal, private, non-agricultural sector, mostly through the formalization of informal jobs.⁷ The authors sought to generate knowledge, “especially because of the importance of the non-affiliated labor sector in Mexico and in many other countries and the lack of

macroeconomic models to finance social security around the world that include labor market specifications in which a large group of workers are not covered”.⁸ The paper specifically focused on the transition of existing jobs between the formal and informal labor markets (not on the creation of new jobs), but the authors nonetheless believed it provided a useful point of reference for labor market dynamics and elasticities in Mexico. The direct implications of the paper for the specific, job creation objectives of PPE were much less clear.

Targeting + Objective

In line with the theoretical paper, the President’s team believed that IMSS fees were high enough to constitute a real constraint on the creation of new, formal jobs.⁹ Programa Primer Empleo thus sought to encourage formal job creation by lowering this cost to employers. The official objective of the program, stated in its creation decree was: “to support employers to generate new permanent jobs, through granting a subsidy that would be applied to the fees paid by the employers for hiring additional new workers and registering them at IMSS.”¹⁰ The program’s founding document does not explicitly address a target population, but the

⁷ The model takes a redistribution of work without coverage to have coverage in the supply, but without more hours of work being offered or more people who were not working being incorporated into the labor market. Since that is not considered, the 300,000 would not be new jobs, the jobs would be going from the informal sector to the formal one.

⁸ Nora Garro Bordonaro, Jorge Meléndez Barrón, Eduardo Rodríguez-Oreggia, *Un modelo del mercado laboral mexicano con trabajo con o sin seguro social (IMSS)* (Mexico: Universidad Iberoamericana, A.C, 2005), 5-7

⁹ Design Evaluation. CONEVAL, page 9. October, 2007.

¹⁰ Objective published in the official decree creating the program. Federal Official Gazette. January 23, 2007



name of the program and its eligibility requirements indicate that it was targeted specifically for recent graduates and people entering the formal job market for the first time. Also, the National Development Plan for 2007-2012 mentioned the program as part of the strategy to create new jobs for youth entering the formal work force.

Determining the Rules of Operation

The program design began as a simple and straightforward initiative, but policymakers soon worried about two potential secondary consequences. One was the fiscal impact, as any subsidy implies the loss of income or an expense for the government. The second had to do with the duration of the subsidy and its potential to generate long-term market distortions. These two factors, among several others, disproportionally influenced the design process. As described below, each program design decision was driven by commendable intentions and solid reasoning, but some of them nevertheless—and unintentionally—turned into barriers and restrictions that negatively affected the performance of the program.

To be eligible for the program, a candidate had to be entered into the IMSS registry as a permanent employee and had to meet the following criteria:

- *The employee had never been previously registered at the Mexican Social Security Institute (IMSS).* Programa Primer Empleo conditioned eligibility to target young people who were joining the workforce for the first time or people who had worked in the informal sector but could transition into the formal sector. This rule also sought to avoid a replacement effect.

- *An employee's enrollment in IMSS had to be in addition to the maximum permanent number of workers already registered by the hiring company.* The program aimed to create new employment, so this rule granted eligibility only to companies that increased their net workforce.
- *The employer could request the payment of the subsidy only after the tenth month of continuous full employment of the newly registered employee.* Through this rule, the program aimed to encourage full-time, long-term jobs and avoid promoting temporary jobs.
- *To participate, the company was required not to have fiscal debts with IMSS or the Mexican government.* Since it was a public program, beneficiary companies had to be up to date with their fiscal obligations.
- *The maximum duration of the subsidy would be twelve months.* The subsidy would vary with a worker's base salary, ranging from 10% to 100% of the IMSS fee. The duration was chosen in order to provide a large enough incentive for companies to hire new workers, while also limiting any harmful, long-term effects on public finances and avoiding the creation of perverse incentives, such as market distortions through the gradual incorporation of the expectation of subsidies into an organization's business model. The intention was that, a) once organizations trained and then observed the quality of a new employee for a year, they would choose to retain him or her; b) once workers were brought to the formal labor market, they would prefer to remain; and c) the subsidies would be short-term but would trigger a long-term impact.



With these eligibility requirements and historical data, IMSS officials worked with Mexican academics to estimate, through simulations based on the projected behavior of the labor market, various scenarios that had different outcomes. Another consideration was the 800,000 employers registered at IMSS, so up to that many employers could have been eligible for the program. It remained to be seen how many of them would actually seek the subsidies in practice. For a number of reasons explored below, however, the program never lived up to its expected results.

Trust

As mentioned above, what began as a straightforward initiative with good intentions soon concerned well-intentioned policymakers with unwanted or unanticipated impacts. Each identified concern was addressed through adding complexity and rigidity to the eligibility requirements, seeking to anticipate and avoid “gaming” of the system. As a result, policymakers unwittingly developed a program built, at its core, on distrust of the same employers that had been planned as vehicles to reach the intended end-beneficiaries.

This mistrust by government officials the private sector was, not surprisingly, reciprocated. The private sector often

What began as a straightforward initiative with good intentions soon concerned well-intentioned policymakers with unwanted or unanticipated impacts.

does not trust government programs because it believes their scope is rarely clear or transparently explained, much less communicated adequately. Accordingly, Programa Primer Empleo was received with skepticism. First, the program was fraught with complicated requirements and commitments. Second, the monetary incentive had strings attached, with added costs and risks, like the need to be up-to-date on all taxes and government fees. Employers had three main concerns: (1) the risk of being subject to additional audits, (2) the costs of artificially long contracts with workers, and (3) the uncertainty and administrative burden of claiming the subsidies, ex-post, from IMSS.

By law, IMSS is one of three Mexican institutions with legal auditing and sanctioning capacities. Companies thus feared that the program would provide an opportunity for IMSS to audit them. In addition, Mexican labor law was known for its rigidity. The program granted a partial, temporary subsidy, while companies absorbed the long-term financial and legal risks of recruiting individuals with no experience in the formal labor market. According to the program's rules, employers were only allowed to claim the subsidy once the new employee had completed ten months of employment. As the president of Durango's chapter of the Mexican Employers' Association (COPARMEX)



mentioned in a newspaper article, “when a new employee without experience enters the workforce of any business, he or she requires training time; however, in this process there is always desertion, approximately in the fifth or sixth month”¹¹ and under this program, the employer did not reap any benefit unless the new employee stayed at least ten months.

Including the Private Sector in the Conversation

Government officials involved during the design phase of the program had felt no need to consult with the private sector, because they believed PPE was a simple, straightforward idea. Time was also limited, since they wanted the program to start very early in the Calderón administration to make a strong political statement. The design process entailed robust discussions among a group of highly skilled and experienced policymakers, so designers were thoroughly convinced by the logic behind the program. At the time, policymakers were convinced that they had listened to outside voices, but the other stakeholders experienced the program as a unilateral, executive decision by the Federal Government.

When asked, key representatives from the private sector said they were not invited to be part of the design process and only learned about the program after it was

Government officials involved during the design phase of the program had felt no need to consult with the private sector, because they believed PPE was a simple, straightforward idea.

launched. Many mentioned that the Federal Government had been the only one involved in the design phase, and not seeking the business perspective had been one of the reasons why the program did not succeed. Private sector representatives believed that PPE was far too complex and would have required substantial modifications to become attractive. One of them explained that COPARMEX had attempted to intervene and influence the program design, but since the organization played no official role within the IMSS Technical Council,¹² its engagement had been limited to indirect political action and marketing efforts.

Program Implementation + Monitoring

IMSS was responsible for operating the program, even though it lacked the legal mandate to promote employment or provide fiscal incentives. The rationale was that, as the institution tasked with keeping track of the formal labor market and running the country’s employment database, it held the most timely and accurate employment information, so it was best positioned to oversee PPE. It also made administrative sense, as PPE operated as a government subsidy to employers who would be reimbursed for fees already paid to IMSS. At the same time, IMSS could not finance the program, as it is a tripartite body, financed by employer fees, worker contributions, and transfers from the Federal Government. These resources are used to cover workers’ pensions, provide them with medical services, and expand the medical infrastructure, and by law cannot be used for other types of programs or activities. As a result, the Ministry of Finance (SHCP) provided the necessary resources and became a central actor in the program.

¹¹ Citlalli Zoé Sánchez, *Necesario, modificar Primer Empleo: IP* Published: Wednesday August 8, 2007. *El Siglo de Durango*. Accessed August 31, 2017. [Access here](#).

¹² IMSS Technical Council is integrated by representatives of the employer, worker and government sector. In terms of IMSS Internal Regulation, the Technical Council is the legal representative and manager.



Within IMSS, two units were responsible for the program. Formally, the Inspection and Collection Unit was designated as responsible for the program, but the Planning and Evaluation Coordination department¹³ was in charge of day-to-day operations. At the same time, PPE had no formal organizational structure, only individuals assigned to it from the two aforementioned units.¹⁴ The program had national coverage, with no emphasis on specific geographical areas, so it was available throughout Mexico to companies that met the eligibility requirements. The program wanted to avoid generating additional administrative burdens for employers registering for PPE, so it was designed to run on an online platform. The system was developed so that all the registration and verification was done through the online system, avoiding extra administrative costs for the employers and reducing the program's operating costs, which would hopefully generate greater incentives for employers to register for the program, and make the benefits available to as many workers as possible in an efficient and effective way.

Given its dependence on the voluntary registration of employers, the program required effective communication to all eligible companies. Communication efforts in 2007 included printed materials, a phone

campaign, a media campaign, direct mailings, and dissemination throughout the IMSS network (see appendix 1 for details). While there was no systematic effort to measure the effectiveness of these communication campaigns, at the time PPE's implementers believed the efforts had been sufficient. They acknowledged later that it might have been good to have maintained or intensified the level of awareness-building of the program.

PPE was funded by federal resources and, by law, was subject to a series of rules and controls, including audits by the Auditor General of the Federation and a formal Design Evaluation by a Federal entity. In addition, given its tradition of careful data collection and analysis, IMSS systematically collected and analyzed program results. Despite the data and the evaluations of the program's operations, there was no systematic analysis of its intended impacts or of the reasons behind PPE's lack of acceptance among eligible companies. There were several structural reasons why this was the case. First, IMSS does not have the legal mandate to carry out impact evaluations. Second, federal money is usually earmarked with very specific restrictions on its use, which was the case for PPE. Finally, the team for PPE was clearly focused on ensuring that implementation happened ethically and with strict observance of the program's rules.



¹³ The Planning and Evaluation Coordination is a separate administrative unit which is not part of the Inspection and Collection Unit

¹⁴ Design Evaluation. CONEVAL, page 58. October, 2007.



Design Evaluation

A few years earlier, in 2004, new legislation created the National Monitoring and Evaluation System and the National Council for Social Development Policy Evaluation (CONEVAL)—the latter with the mandate to measure poverty and assess social development policy.¹⁵ Following this mandate, CONEVAL commissioned a design evaluation of Programa Primer Empleo, focusing on archival data from the Inspection and Collection Unit at IMSS. The evaluation also included a review of existing literature and of other relevant international experiences. It identified several positive features of the program, as well as a series of structural design flaws. On the positive side, the evaluation found that “the program had minimal operating costs, due to the highly-automated process which resulted in few operational field activities.”¹⁶

The evaluation’s final report, however, also identified critical shortcomings:

“Programa Primer Empleo performed well below expected results due to an inadequate program design, resting on the assumption that a subsidy to the social security fee was a determining factor to stimulate demand of formal and permanent jobs for new entry-level workers who traditionally did not access permanent hiring schemes despite the inertial dynamics of the labor market.”¹⁷

“The weakness showed by the operation of Programa Primer Empleo was in line with international experience, which conclusively points out that strategies that use a salary subsidy scheme as the only component are not effective in order to integrate a vulnerable population into permanent and formal jobs. The programs that promote a combined strategy that include subsidies to the employer, job training for vulnerable groups, and support and information services while searching for employment are more likely to succeed. Also, for combined strategies to be more likely to succeed, the training received by workers must be linked to productive needs and the benefits of the program should be targeted to the vulnerable population. Even if the subsidy provided by Programa Primer Empleo had been higher, the Program would have not worked properly because the international evidence is conclusive that such programs, operating solely through a subsidy to the employer, have not worked.”¹⁸

“There are other factors that explain the low level of permanent employment in the formal sector, such as: 1) the rigidity of labor legislation; 2) the low valuation of workers regarding social security, especially young people; 3) the high information costs faced by a vulnerable segment of the population; 4) incentives for circumvention and avoidance of employer contributions; and 5) the low educational levels of the vulnerable segments of the population.”¹⁹

¹⁵ Created by the Social Development Law, see: *¿Quiénes somos? – Creación del CONEVAL*. CONEVAL [Access here](#).

¹⁶ Design Evaluation. CONEVAL, pages 77-79. October, 2007.

¹⁷ *Ibid*, 6-7.

¹⁸ *Ibid*, 20.

¹⁹ *Ibid*, 37.



The evaluation delves into the structural impediments to youth employment because of their apparent absence in the design of PPE, which contributed to the program's lackluster results.²⁰ Because of the PPE's structural shortcomings, the evaluation issued the recommendation to terminate the program due to:

- 1) "The unviability of the Program under its current design, and 2) the relevance of modifying the PPE towards a comprehensive strategy when these efforts already occur at the Ministry of Labor and Social Welfare."²¹

Should the administration retain the program, the evaluation proposed several modifications to address some of the program's flaws, as seen in Table 1.

²⁰ Ibid.

²¹ Ibid, 86.

²² Ibid, 77-79, 86.

Table 1. Proposed Modifications to PPE (2007) ²²	
ISSUE	PROPOSED RECOMMENDATION
Potential and target populations are not well identified or adequately quantified by the Program.	The potential population should be defined as vulnerable groups that, in the absence of the Program, would not obtain permanent employment in the formal sector. The target population should be workers with low educational levels, of productive age, with little or no work experience, from low socio-economic strata and with a higher propensity to work in the informal sector, who would not otherwise join the formal job market.
The purpose of Programa Primer Empleo is to encourage the demand for vulnerable workers beyond the normal inertia of the labor market. This is not fulfilled, so the Program is subsidizing workers who were already likely to enter the formal labor force.	
The Program does not allow portability. An employee who changes jobs after registration in the program but within the eligibility period automatically loses eligibility.	Allow portability. The program should not generate additional rigidities in the labor market. Workers, especially those with lower incomes and educational levels, have greater mobility and turnover, so the program should allow them to change jobs while retaining eligibility for the Program.
The subsidy component of the worker-employer fee is not adequate, because it is based on a theoretical assumption about the dynamics of the informal labor market, not about new job creation. Furthermore, unidimensional programs, such as the provision of economic incentives to employers, have proven insufficient in previous international experiences and in the evidence of PPE itself.	This issue forces us to rethink the pertinence of keeping Programa Primer Empleo in operation due to the infeasibility of the Program under its current design, and the relevance of modifying PPE towards a multi-dimensional, comprehensive strategy that includes training and job placement (the Ministry of Labor and Social Welfare already has programs that seek to follow these principles).
OTHER ISSUES AND RECOMMENDATIONS	
The information about employers is not enough to know their needs thoroughly.	It would be important to carry out a survey to understand the employers' perspective, opinions, and suggestions. This would provide knowledge on the perceptions companies have on youth employment, training, and employment demand.
A survey of workers is also required to know their perceptions on employment, education, and job offers.	This study would serve to understand workers' disposition on types of employment and training programs.



Regarding the paper “A Model of the Mexican Labor Market with and without Social Security, 2003,” which served as the primary evidence and inspiration for the program, the evaluation states that the study was a useful theoretical analysis of labor market dynamics that, nevertheless, had no empirical validation. In fact, the early results of PPE demonstrated that the reduction of employer’s fees was not enough to increase the number of new, permanent jobs in the formal sector above the normal market inertia.²³

Finally, the evaluation concluded that previous international experiences had not been adequately incorporated. In particular, the evaluation mentions two types of international experiences focused on the employment of vulnerable workers that seemed relevant to PPE:

- 1) Programs that provide a subsidy to the employer - providing economic support to hire the worker. The two programs mentioned in the evaluation were: The New Jobs Tax Credit²⁴ and the Targeted Jobs Tax Credit.²⁵ Research on both programs concluded that neither had been effective.
- 2) Programs that promote a multidimensional strategy providing support for employers (subsidy) and employees (training and job search support). The two programs mentioned in the evaluation were: The Youth Incentive Entitlement Pilot Project²⁶ and the AFDC Homemaker-Home Health Aide Demonstrations,²⁷ with less conclusive evidence.

The CONEVAL evaluation thus concluded that the evidence on unidimensional programs, like PPE, was unequivocal: they are not sufficient to generate new jobs or insert vulnerable groups into formal and permanent jobs.

Civil Service Middle Management, the Federal Government, and the Private Sector

The team charged with the execution of PPE had been appointed to IMSS during the presidential transition, but many had had long careers as civil servants. Because of their experience in the public sector, they were skeptical of the program and the assumptions it was built upon. Regardless, it was their task, and not that of program designers, to ensure that PPE was implemented in strict accordance with its operational rules. As program results trickled in (IMSS is a rigorous collector of labor market data), the fears of the implementation team were confirmed: numbers were far below expectations and with no signs of promise.

The CONEVAL evaluation was a turning point, because it not only corroborated the internal analysis that IMSS had performed on the operation and evolution of the program, but also was an objective report from an *external* actor, indicating that the program was not having the desired results because of structural flaws in its design (and not because of ineffective implementation). Implementers agreed with the CONEVAL assessment, but also felt that employers probably feared IMSS as a fiscal control entity: The program stated that companies had to be up to date with all their fiscal obligations with the Mexican government and open to random, unannounced audits from IMSS.

²³ *Ibid*, 13.

²⁴ *New Jobs Tax Credit: USA – 1977.*

²⁵ *Targeted Jobs Tax Credit: USA – 1978.*

²⁶ *Youth Incentive Entitlement Pilot Project: USA – 1978 to 1980.*

²⁷ *AFDC Homemaker-Home Health Aide Demonstrations: USA—1983 to 1986*



As much as implementers agreed with them, most of the critical recommendations provided by CONEVAL fell beyond the institutional mandate and jurisdiction of IMSS. Implementers were legally constrained by their official roles and by program rules, so even as they attempted to integrate the new evidence, they were only able to tweak the program at the margins. This was not only insufficient, but also made it appear as if IMSS had resisted or ignored what the evidence was suggesting was needed. IMSS implementers had to keep the program running according to the federal budget and the program rules. Addressing the main recommendations, in contrast, would require involvement from higher levels in the federal government and a fundamental shift in strategy, far above the mandate or purview of IMSS.

In time, CONEVAL understood that those who operated the program were not the ones who made decisions on whether to terminate it or even on its structural design. To the evaluator's surprise, those responsible for implementing the program were transparent, open, and receptive. It was clear that there were good intentions behind all the decisions, even those that kept the core of the program design intact. As one of the evaluators noted: "I was surprised by that, but at the same time I was not. It is difficult for someone within a

government to have an evaluation and say, 'Ah, I will totally change it,' especially when it's a flagship program."

After CONEVAL's evaluation and in line with the implementers' intuition, the program underwent two major modifications to ease the requirements for companies and beneficiaries. In general terms, starting in December 2007, the rules were slightly simplified:

2007 | 2008²⁸

- Targeted population expanded: individuals who had not been employed for more than nine consecutive months registered with one employer as a permanent worker remained became eligible.
- The waiting period to claim the subsidy was shortened from ten to four months after the employee's registration.
- The registration period for employers and workers was extended to August 31, 2011 from February 28, 2011.
- The condition of being up to date with all fiscal obligations was made more flexible: the requirement for timely completion of fiscal obligations was kept, but only at the moment of claiming the subsidy and not at the time of registration.

After the first round of modifications was put into place, a document developed by the Mexican Employers' Association (COPARMEX) noted several program shortcomings, including: so many requirements that most companies would not consider it as an option; restricting the subsidy to hiring of people who had never had a formal job, and so precluding companies from hiring staff with previous experience; not taking into account the reasons behind any fiscal liabilities companies might have; and having to maintain a steady workforce, with no margin for industry dynamics or natural turnover.

The program continued to perform below its expected results, so in September 2011 the Federal Government consulted the private sector on how to restructure and relaunch PPE²⁹. COPARMEX proposed seven modifications: (1) extend the duration of the program, (2) eliminate the requirement for full fiscal compliance, (3) through a decree, confirm that there is no need to be up to date

²⁸ ASF. *Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2012. Programa Primer Empleo performance Audit (12-1-00GYR-07-0379) DS-095 page, 2. 2012 and Comisión Representativa ante Organismos de Seguridad Social: Programa Primer Empleo. December, 2011.*

²⁹ *El Empresario*. [Access here](#).



with Federal Government liabilities, (4) reduce the period to begin collecting the subsidy to two months, (5) IMSS should deposit 100% of the subsidy in the month following an employer's request, (6) the option to choose between receiving the subsidy in a bank account or a 10-day credit note, and (7) design a new communications campaign, as the first set of modifications were not well known by target companies and beneficiaries.

On November 14, 2011, the Federal Government issued a new set of guidelines for Programa Primer Empleo, which incorporated some of these recommendations:

2011³⁰

- The employer registration period was extended from August 2011 to September 2012.
- Worker eligibility was made more flexible, allowing people with limited work experience (and not only individuals with no experience) to register for the program.
- The waiting period to receive the subsidy payment was shortened from four to three months of full employment of the registered employee.



As noted before, since the beginning of the Calderón administration, the federal government had intended Programa Primer Empleo to make a big splash. When the program was launched, the signing of the creation decree was hosted by Mr. Calderón himself with the presence of union leaders, business leaders, legislators, and the media.

During his speech, he highlighted that with this program, his government was fulfilling one of the most important commitments

made to the Mexican people during his campaign. That same day, Mr. Calderón inaugurated the National Chamber of Industry's (CANACINTRA) 2007 National Industrial Convention where he announced the creation of Programa Primer Empleo.

30 ASF. Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2012. Programa Primer Empleo performance Audit (12-1-00GYR-07-0379) DS-095 page, 2. 2012 and Comisión Representativa ante Organismos de Seguridad Social: Programa Primer Empleo. December, 2011.



By 2009, the President had stopped talking about the program and during the beginning of 2010 the director at IMSS said during an interview for a newspaper that “Programa Primer Empleo was not in the institute’s agenda as a priority item.”³¹ After its first year, President Calderón’s administration became absorbed by the security agenda and PPE lost prominence. PPE was maintained because canceling it would have given the government negative attention, especially as people were critical about the focus on security at the expense of economic issues. By 2008 the program’s resources were redirected and it only kept enough funds to meet the commitments with the registered companies.

Results

Programa Primer Empleo officially came to an end with the Calderon administration, in November of 2012. At this point, the program had met 13%³² of the employee registration target and 3.2%³³ of the employers’ enrollment target. The results were far below the expectations which targeted a goal of 738,133 employees registered between 2007 and 2012³⁴ and, in the eyes of an evidence producer, “the results were not aligned with the theory of change,” the magnitude of the expected results was vastly lower than projected.

Figure 1. Programa Primer Empleo Results

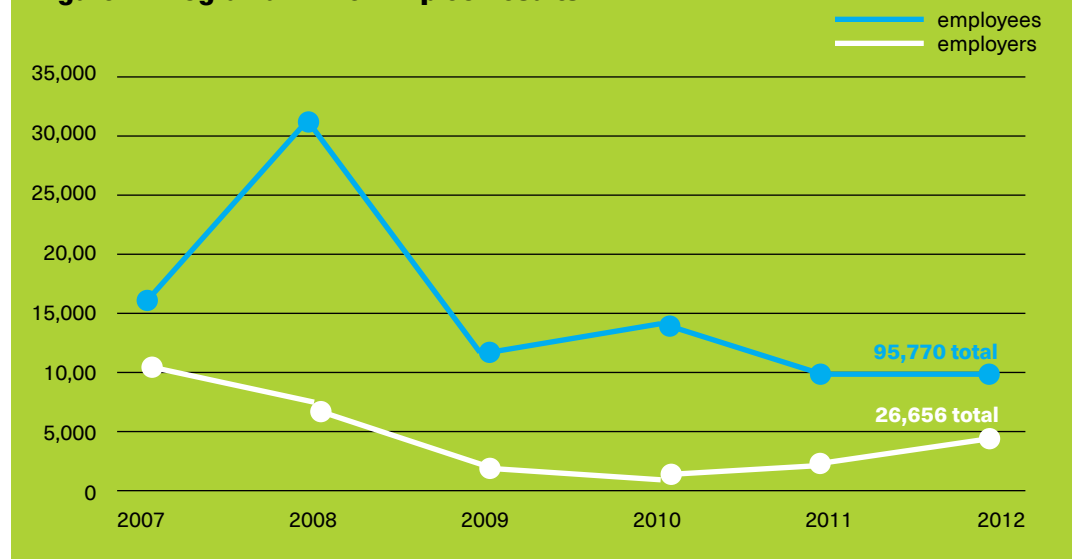


Figure 1 shows the evolution of the program throughout its history. As shown, there was an interesting spike in the number of registered employees in 2008, the same year the first major modifications were implemented to make the program more flexible. According to one of the implementers, the spike and the evolution of the statistics suggests that the first set of modifications had a positive impact, but it was not

³¹ Martínez, María del Pilar. “Programa Primer Empleo, al archivo muerto del IMSS” *El Economista*. Published March 3, 2010. [Available here](#).

³² ASF. Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2012. Programa Primer Empleo performance Audit (12-1-00GYR-07-0379) DS-095 page, 5. 2012

³³ *Ibid.*

³⁴ *Ibid.*



sustained for all the reasons described above. In addition—and in reaction to the 2008 financial crisis—the Mexican economy and labor market contracted in 2008, further affecting PPE results. The incentives provided by Programa Primer Empleo could have helped soften the impact of the recession, but once again certain design features interacted with emerging realities in unexpected ways. The subsidy was contingent on increasing the historic maximum workforce registered by a given employer, but few companies were able to increase absolute employment in a context of low growth. So even if companies hired new employees during those difficult times, they were not eligible for PPE if their entire workforce remained below their historic maximum. While the logic behind that rule was to encourage the creation of new jobs and prevent employee replacement, it actually played against the program's potential to mitigate the economic crisis. It was yet another way in which the program's design process and rigid structure hindered its own potential.

While the logic behind that rule was to encourage the creation of new jobs and prevent employee replacement, it actually played against the program's potential to mitigate the economic crisis.



Part II: **Evidence in Practice : Key Themes + Insights**

This section discusses the Evidence in Practice themes as they pertain to Programa Primer Empleo and summarizes key insights and implications for thinking about the translation of evidence to policy and practice more generally.

The Role of Timing in the Incorporation of Evidence into Practice

The different and often discordant timeframes within which researchers, policymakers, and implementers operate often hobble efforts to coordinate, let alone collaborate, on evidence-informed approaches. Electoral cycles and political windows differ from NGO funding cycles and from academic publishing rhythms. Yet each actor is bound by the timeframes of her formal stakeholder group.

Programa Primer Empleo illuminates how constraints, especially campaign promises and post-electoral pressures, play a major role in facilitating—or constraining—the incorporation of evidence into policy and practice. When designing and implementing a major government program, policy makers have incentives to start new programs and fulfill campaign promises at the beginning of their political cycle.

Within the five-month transition period before the Calderon administration began, Programa Primer Empleo was designed in order to be launched at the beginning of the administration, as a fulfillment of one of President Calderon's most important campaign promises, in an environment of social fragility and political pressure. This situation illustrates the struggle between following a truly evidence-informed design process (with sufficient time to generate and/or consider evidence, implement pilot programs, and modify the design accordingly) and the exigencies of a public policy implemented with a fast-paced timeline in order to generate a quick and big “splash”; in the case of PPE, this approach affected not only the design and launch of the program but also its ongoing implementation.

Another key lesson that can be drawn from PPE is the importance of designing public policies that retain flexibility for learning. This naturally has to be balanced with the need

for strict operating and accountability rules in the use of public funds. But as other cases have shown, operating rules can include a mandate for adjustments based on learning. They can also include a structure that allows all relevant stakeholders to remain engaged throughout the design and implementation process, to promote the integration of all relevant evidence. In the case of PPE, the sequential (and disaggregated) role played by different stakeholders meant that the program missed critical perspectives in its design and was not adaptive to useful insights about its performance.

For Programa Primer Empleo, program designers did not have much evidence to support the assumptions behind the core of the program, but the pressures to start at a major scale and have a big splash pressed the program forward. Alternatively, the launch of the program could have been seen as an opportunity to pilot the intervention and test all critical hypotheses, in order to gain insight on what was needed to achieve a larger impact.

As time progressed, and especially towards the second set of modifications, the time constraint of the program's duration and with Calderon's tenure coming to an end, any major modifications were difficult to implement, due to: 1) the structure and rules of the program



that gave little flexibility and 2) the incentive politicians have to garner credit from their initiatives. Although the program did not have the expected results, politically it was very difficult to cut the program. Policymakers were thus subject to a series of pressures and restrictions that limited their ability to integrate the results of a potentially helpful evaluation. PPE's implementers literally could not change any structural features of the program because of its rules, which lacked the mandate to collect or adapt to evidence. The political environment made it difficult for those higher in the administration to intervene. Primer Empleo helps us understand the many factors leading to the realization that evaluations do not necessarily change public policy or program design, especially when the programs are important to the political strategy.

Policymakers were thus subject to a series of pressures and restrictions that limited their ability to integrate the results of a potentially helpful evaluation.

Evidence Definition, Creation + Use

There are varying definitions and understandings of what constitutes "evidence," dependent especially on the perspectives of each stakeholder group. For example, the framing, language, and limited accessibility of academic evidence can render it less useful to other stakeholders. These diverging views of evidence create barriers across stakeholder groups, as what constitutes valid evidence for each exists in different realms and in different forms that are challenging to reconcile.

Programa Primer Empleo's design was based on a theoretical economic model developed a few years earlier. The model identified potentially critical leverage points in stimulating employment, but did not consider many underlying assumptions or identify

conditions that might occur in the field that would be conducive or resistant to its proving successful. Further, the program's rules and structures of operation, while designed with good intentions, proved to be unappealing to the target audience of employers. While the rationale behind them was correct in conceptual terms, the complexities of the field proved them to be inaccurate.

PPE illustrates many of the risks that arise in well-intentioned public policies when a comprehensive, evidence-informed approach is not followed in its design and roll out. Rather, PPE reflects the political imperative to launch full-scale, apparently "watertight" programs. PPE shows the importance of clearly identifying (a) the assumptions underlying an initiative, (b) the types of



evidence that would best shed light on those assumptions, (c) the actors who could best provide such evidence, and (d) how the initiative could adapt to such new evidence. This is especially true with programs that seek to modify behavior at scale, where inaccurate assumptions about actors' incentives can lead to unintended effects and a drain on resources.

While PPE at its core had the ambition and the potential to be a flagship program for the Calderon administration, the way it was designed and executed essentially turned it into a single (and expensive) test of a larger, unexamined hypothesis—which proved to be false. There were lost opportunities to engage key stakeholders as genuine partners, including potential employers, academic institutions, and think tanks in the design process, and implementation staff in using the evaluations' recommendations to strengthen the program. Its rigid design left committed, well-informed implementers unable to act upon emerging evidence. The narrative of the case highlights how government officials often lack the mandate, resources, and time to make use of the evidence generated by rigorous evaluations that can lead to program improvements.

Need to Devote Exclusive Time and Resources to Learn About and Operationalize Evidence

Few organizations provide incentives or carve out explicit time for managers to devote to learning about emerging evidence in their field of endeavor, and even fewer have staff explicitly devoted to learning about evidence and translating it into forms relevant for the organization. Even organizations with strong monitoring and evaluation departments often do not transform the operational data into formats that could be widely used within the organization, or beyond, to expand actors' understanding about what has been learned from past or existing programs. Data is thus used to evaluate retrospective operations, but not to improve the prospective design of new initiatives. Discovering, incorporating and translating evidence requires time, energy and funding.

As described in the case study narrative, the evaluation found that Programa Primer Empleo did not use all of the international evidence available at the time, such as the New Jobs Tax Credit and the Targeted Jobs Tax Credit from the U.S. This emphasizes the importance of not only explicitly considering different types of evidence upfront, but also factoring in the resources, bandwidth, research capacity, and time needed to ensure that relevant, timely and useful information is





made readily available to program designers and implementers. The critical analyses from the periodic evaluations of the program were also not fully taken into account. For the Primer Empleo story, considering all these factors could have helped develop a good intention to its fullest potential without investing valuable resources in a weakly conceived intervention—and ultimately to the discrediting of a flagship program.

Importance of Building Trust and Forging Relationships Among Stakeholders

The cross-stakeholder collaborations required for evidence-informed policies and practices are often difficult to initiate, develop, and sustain. Particularly when institutional incentives are lacking, personal trust, respect, and buy-in between individuals across stakeholder groups become critical to fostering the effective flow of evidence into practice.

Programa Primer Empleo illustrates the importance of stakeholder engagement to make sure policymakers, intended partners, and beneficiaries have the same understanding of the problem and potential solutions. The program's objective focused mainly on supporting employers to create permanent formal jobs, yet employers did not participate in the design of the program. Not having agreement on what would be

Another key lesson that can be drawn from PPE is the importance of designing public policies that retain flexibility for learning.

appropriate incentives or targets created confusion among the relevant stakeholders, each having a different understanding of what the program aimed to address and how to achieve its goals. PPE illustrates the importance of effective communication between key stakeholders, both upfront and ongoing, as fundamental for the successful design and implementation of a program. The case study research suggests that lack of trust blinded both sides, where policymakers complicated the program's operational rules (e.g., by requiring a company not to have fiscal debts with IMSS and by creating tight restrictions on employee eligibility) out of concern that employers could bend the rules and abuse the program. In turn, employers doubted the appropriateness of, and felt

threatened by, the rigidity of the rules—a natural reaction to rules that clearly were built on mistrust. Because of the program's apparent simplicity, together with the rushed timing and the fear of being taken advantage of by employers, policymakers did not include the private sector in the design process. As noted earlier, this process prevented the main stakeholders from being mutually invested in the program, or developing strong, productive relationships.

Even after the two rounds of program modifications had addressed some of their concerns, private sector employers were still not attracted to enroll. This illustrates that both sides did not understand the problem and each other's perspective accurately and could have benefited from more in-depth studies and collaborative participation in identifying how best to generate new employment opportunities.

Failure to Learn from Failure

Potential consequences for risk-taking and experimentation with innovative approaches are generally seen as negative and dissuade the exploration of novel, evidence-informed interventions. Fear of failure can further hinder the incorporation of novel evidence into practice, even when stakeholders recognize the value and applicability of the evidence.



Programa Primer Empleo had been operating for three years and was known not to be achieving its expected results. The evaluation identified international evidence that corroborated the lack of effectiveness of unidimensional tax incentives to promote employment. Nevertheless in October 2010 a new effort to incentivize employment was carried out. The Senate³⁵ approved amendments to the income tax law, and the Law to Promote First Employment (*Ley del Fomento al Primer Empleo*). The main difference between Programa Primer Empleo and the new Law to Promote First Employment was that the incentive in the new law consisted of an additional deduction in the employer's income tax from each employee hired and registered at IMSS for the first time.

The two programs at their core were very similar, and the new legislation did nothing to address the underlying flaws that international evidence illustrated and the lessons learned from the design and operation of Programa Primer Empleo, even though these were widely known and had been identified in PPE's evaluation.

Conclusion

The case of Program Primer Empleo illustrates several of our research project's broader themes, especially the role of timing

Although the lessons learned by the design and implementation of PPE were difficult ones, some of our informants recognized the value that PPE had on how these lessons contributed to a richer knowledge about how to support the design of public policy in the future.

in the incorporation of evidence into practice; the role of evidence definition, creation and use; the need to devote time and resources to learn about available evidence; the importance of building trust among stakeholders; and the failure to learn from failure.

Although the lessons learned by the design and implementation of Programa Primer Empleo were difficult ones, some of our informants recognized the value that PPE had on how these lessons contributed to a richer knowledge about how to support the design of public policy in the future. A high ranking official recognized Programa Primer Empleo as a failure from the public policy standpoint, but pointed out how PPE has

also served as an experimental exercise to understand the complexities of policy design and how the government could better design and implement social programs, so that they would be addressing the issues in a rigorous and productive way.³⁶

35 México, D.F., a 26 de octubre de 2010.- Dip. Jorge Carlos Ramírez Marín, presidente; Sen. Manlio Fabio Beltrones Rivera, presidente; Dip. Cora Cecilia Pinedo Alonso, secretaria; and Sen. Adrián Rivera Pérez, secretario

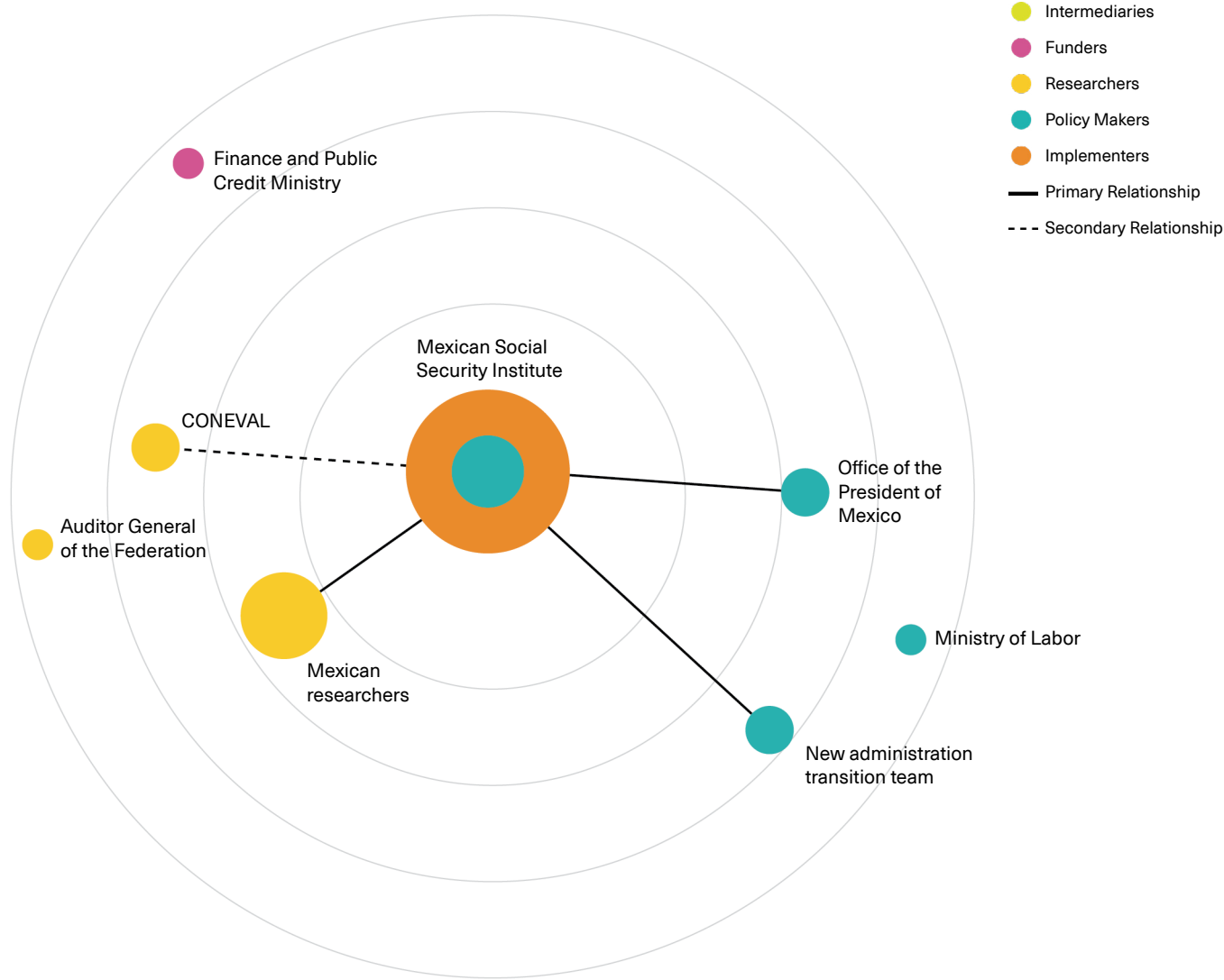
36 Miranda González, Sergio and Salgado Vega, María del Carmen. La nueva Ley Federal del Trabajo en México formaliza la flexibilidad laboral. Revista Trimestral de Análisis de Coyuntura Económica. April-June 2013. Vol. VI Number 2. [Available here.](#)



Stakeholder Map

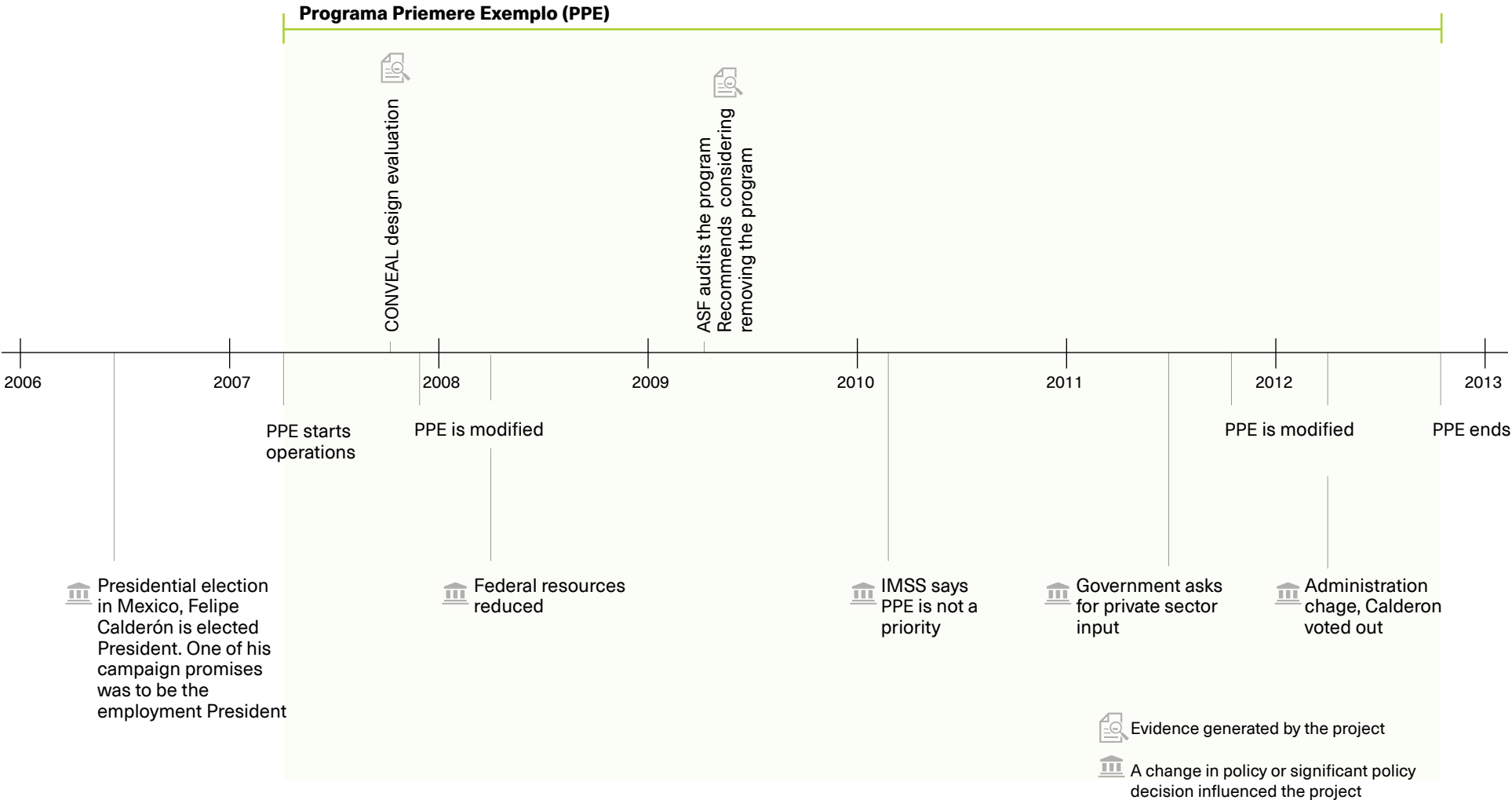
Programa Primer Empleo was mostly a government program, so the stakeholder map describes the relationships between government agencies.

This stakeholder map is a visual representation of the major stakeholders involved with this project. The importance of each of the actors is defined by their relative size, and their proximity to the center of the project. Their role is defined by the color; multiple colors indicate multiple roles. Primary relationships, denoted by solid lines, indicate the most directly significant relationships while secondary relationships, denoted by dashed lines, indicate indirect, but influential relationships. Actors not connected by lines are still involved with the project, but less directly.





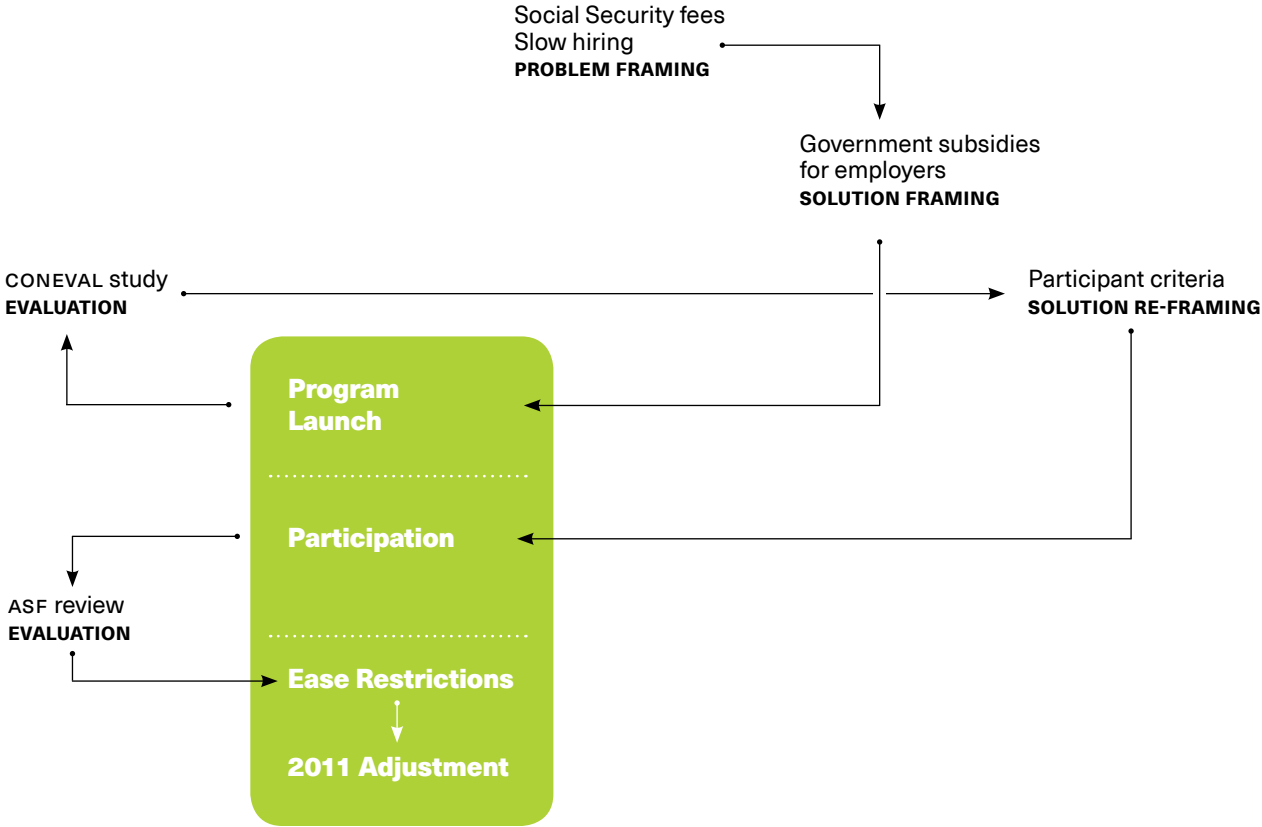
Timeline





Process Diagram

Programa Primer Empleo was launched at a national **scale**.





Contents

Stakeholder Map
Timeline
Process Diagram
Appendices
Print

**Full Report
Summary + Findings**

Case Studies

View Alphabetically
View Geographically
View by Discipline

Appendix 1

2007 Communication Actions for Programa Primer Empleo

Printed	1,001,000 brochures were distributed through IMSS sub-delegations, the President's residence when the program was launched and several unions.
Phone Campaign	From May 7 to June 22, 2007, through IMSS contact center, a telephone campaign was carried out to inform HR staff members at companies affiliated to IMSS about the benefits and requirements of Programa Primer Empleo. By the end of the campaign 78,344 companies had been contacted.
Media Campaign	From June 11 to July 10, 2007, a marketing campaign was carried out in the following media outlets: 27 radio groups, 2 television networks, 11 newspapers in Mexico City and 77 local newspapers, 24 magazines and Postcards and advertising banners.
Post Service	To promote the registration of new companies to the program, letters were sent during 2007. An invitation to all 822,978 employers was sent via the Mexican post service (SEPOMEX) (July and August) and 234,840 by email (May to July).
IMSS Webpage	A direct access to information about Programa Primer Empleo was enabled in the main IMSS webpage.
IMSS Contact Center	Through the IMSS contact center more than 100 thousand queries about the requirements and benefits of the program were answered.



Appendix: Data + Methods

The research design for the Evidence in Practice project consisted of three broad components. First, we conducted expert interviews (31) with individuals who had spent a significant portion of their professional lives attempting, researching, or promoting the integration of evidence into development practice.^{A1} This included academics, government officials, foundation program officers, NGO practitioners, and think-tank directors. To identify these experts, we first contacted individuals who had either published extensively and prominently on the topic or who had actively funded research or programs with the explicit goal of integrating evidence into practice. From this first set of experts we conducted snowball sampling until we reached a saturation point.^{A2} This initial set of interviews informed and directed the next two components, as they resulted in an initial map of the relevant stakeholders in the “evidence-to-practice ecosystem” and the hypothesized and actual paths that seemed to link them together.

Second, we conducted a matched comparison of eight cases of development programs or interventions where rigorous evidence was integrated with varying degrees of effectiveness. These cases were matched on structural, geographic, and programmatic characteristics—as well as on the extent to which evidence had informed practices—to better identify the critical factors that allowed actors in certain cases, and not others, to integrate rigorous evidence into practice.^{A3} This matching process led us to identify pairs of cases across four different countries, leveraging temporal and cross-sectional variation between them as seen in table A2.

A1 By development practice, we mean the work of government actors, NGOs, and others who are responsible for designing and executing development projects and programs.

A2 Data saturation is difficult to define and is dependent on the field of study. In this case, we defined saturation as the moment when, in a sequence of several expert interviews, no interviewee gave us information that we had not encountered before.

A3 George, A. L., & Bennett, A. (2005). Case studies and theory development in the social sciences. MIT Press, Cambridge, MA. Chapter 5.



Table A1. Expert Interviews						
	Researchers	Funders	Intermediaries	Policymakers	Implementers	Total
Phase 1: February 2015 – May 2016	10	8	6	2	5	31
Phase 2: September 2016 – June 2017	10	7	6	3	8	34

For each case, we first identified, through existing literature and interviews with subject experts, a series of key informants who had detailed knowledge of the case’s history and protagonists. These initial interviews with case experts led to the creation of a detailed actor/stakeholder map for each case, where we identified the key stakeholder groups that either participated in or were affected by the program, as well as the specific individuals who played an active role in the program’s evolution.^{A4} These stakeholder maps were validated with several informants for each of the cases. We then conducted interviews with each of the key individuals across stakeholder groups. Interviewees were asked to relate chronologies of objective events, behaviors, choices at critical junctures, and facts of the processes described.^{A5} In every instance, the goal was to identify the individuals responsible for the particular evolution of a case, as well as the specific

tactics they employed throughout the process, to better understand the rationale behind their decisions as well as the factors that led them to succeed or fail. In total, we conducted 161 interviews across the eight cases. Interviews were complemented with a wealth of archival information including media articles, private documents (donor reports, internal presentations and communications, etc.), and public documents (announcements, academic articles, editorial pieces). These data were used to trace the chronological list of events for the overall development of each case. Each storyline was developed in an extensive document that established the causal links described by the subjects and ensuring a balanced consideration of different stakeholders.^{A6}

The third component, conducted in parallel to the eight case studies, consisted of interviews with prototypical representatives of each of

the stakeholder groups, or individuals who would clearly describe the typical experience of enacting a particular stakeholder role. Using the stakeholder map and initial hypotheses as starting points, this stage focused on the dynamics that shape the interactions between stakeholder categories. The work consisted of 34 in-depth interviews with representative actors from each stakeholder group. The interviews focused on each individual’s needs, assumptions, operational constraints, main concerns, professional and ideological backgrounds, timelines, and aspirations—especially concerning the development, dissemination, and use of novel evidence in development practice. This in-depth analysis resulted in a more nuanced and detailed stakeholder and system map that more clearly identified both breakdown points and paths of connection that hinder and facilitate the exchange of knowledge and information across stakeholder groups, as well as a refined

^{A4} See Canales, R. (2016). *From ideals to institutions: Institutional entrepreneurship and the growth of Mexican small business finance*. *Organization Science*, 27(6), 1548-1573.

^{A5} Davis, J. P., & Eisenhardt, K. M. (2011). *Rotating leadership and collaborative innovation: Recombination processes in symbiotic relationships*. *Administrative Science Quarterly*, 56(2), 159-201.

^{A6} *Ibid.*



Table A2: Case Studies

Country/Program	Description	Dates of Intervention	Number of Interviews	Primary Stakeholders
South Africa				
Collaborative Analysis of Labor Intervention Effectiveness	Employment program introducing new elements to vocational training	2011 – 2016	42	Government, Researchers
FUEL: Feed, Uplift, Educate, Love	School nutrition program	2007 – present		NGO
Ghana				
Teacher Community Assistant Initiative	Remedial education program for primary school children in reading and math through teaching assistants from local communities	2010 – 2013	30	Researchers, Government
Graduating the Ultra Poor	Poverty alleviation program integrating elements of social protection, livelihoods development, and financial services	2010 – 2013		Researchers, NGO
India				
Teaching at the Right Level	Remedial education program for primary school children in reading and math	2001 – present	51	NGO, Researchers, Government
AQUA+	Water purification drops for retail sale	2010 – present		NGO
Mexico				
Progresa Oportunidades	Poverty alleviation program using conditional cash transfers	1997 – present	38	Government
Programa Primer Empleo	Employment program using government incentives for the private sector	2007 – 2012		Government



set of hypotheses about the breakdown of communication and about possible interventions to solve it.

Across the three components, we conducted a total of 226 interviews. All interviews were in-depth and semi-structured, with an average length of around 90 minutes (minimum of 60, maximum of over 120). Around two-thirds of them were done in person and the rest were conducted remotely. All interviews were recorded and transcribed verbatim.

Data analysis was conducted in several stages. Each of the 226 interview transcripts was coded extensively to identify first-order concepts related to the integration of evidence into development practice. First-order concepts include “concerns about reputation” or “short-term decision-making”. This required multiple readings of interview transcripts, field notes, and archival data to associate nearly every passage of text with one or more codes. These codes were then grouped into second-order themes,^{A7} always contrasting them with current research on the integration of evidence into practice. Second order themes included “incentive structures” or “timing misalignments”, each of which was developed extensively in a memo that explored the characteristics, tensions, and contradictions of each theme. In stage

three, we mapped the codes to each of our case narratives to detect patterns of activities, constraints, and decisions that defined the evolution of each case at critical junctures. This allowed us to identify similarities and discrepancies across cases, as well as to create comparable counterfactuals that could account for differing outcomes.^{A8}

In stage four, we created process maps, concept maps, data tables, and detailed case synopses that linked key challenges, events, and decisions to the specific alternative tactics employed by actors and then to their subsequent consequences for the development program or intervention in question. This final set of analyses revealed a somewhat consistent set of factors faced at comparable stages by actors across our different settings. Throughout our analysis, we iterated between emerging insights, existing theory, and matched comparisons across cases to identify the mechanisms that operated at critical junctures.

It is worth mentioning that, at two moments of the project (the first after our first set of expert interviews was over and the second after the completion of our initial case narratives) we hosted a workshop with two different groups of highly experienced representatives from each of the stakeholder groups. During these

workshops, we discussed our emerging findings and we gathered additional, essential insights from participants. The workshops served to validate and deepen our understanding of emerging insights.

A7 Glaser BG, Strauss AL (1980) The Discovery of Grounded Theory: Strategies for Qualitative Research (Aldine Publishers, Hawthorne, NY).

A8 We ensured consistency in coding across the different cases and authors through several mechanisms, including: a) a selection of interviews was coded by two or more coders, after which they reviewed discrepancies and agreed on their resolution, b) a common project book where all the codes were collectively kept, aggregated, and analyzed, c) a weekly meeting to review coding process and to develop a joint coding standard, d) memos were developed jointly, with contribution from and verification by the different team members, among others. [Access here.](#)