

# ارزیابی اثربخشی سیاست‌ها (۱): مفاهیم و روش‌ها





بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تاریخ انتشار:  
۱۴۰۴/۲/۱۷

شماره مسلسل: ۲۰۷۰۲  
کد موضوعی: ۲۳۰



مرکز پژوهش‌های  
مجلس شورای اسلامی

### عنوان گزارش:

ارزیابی اثربخشی سیاست‌ها (۱): مفاهیم و روش‌ها

نوع گزارش: طرح/ لایحه □ راهبردی ■ نظارتی □ پیش نویس قانونی □

### نام دفتر:

مطالعات بخش عمومی (گروه سیاست‌های حمایتی و مقابله با فقر)

### تهیه و تدوین کنندگان:

هانیه حسن‌شاهی، سیدمحمد مجتبی میرپنجی

### ناظران علمی:

سیدمحمد حسین فاطمی، سیدهادی موسوی نیک

### گرافیک و صفحه آرایی:

سمیرا صادقی عسکری

### ویراستار ادبی:

اکرم وحدانی‌فر

### واژه‌های کلیدی:

۱. سیاست‌گذاری شواهدمحور

۲. ارزیابی سیاست

۳. اثربخشی

۴. سیاست‌های حمایتی

### تاریخ شروع مطالعه:

۱۴۰۳/۰۲/۰۱



## فهرست مطالب

چکیده	۶
خلاصه مدیریتی	۷
۱. مقدمه	۸
۲. ارزیابی سیاستی چیست؟	۱۰
۳. چه زمانی باید ارزیابی صورت گیرد؟	۱۱
۳-۱. قبل از اجرای سیاست	۱۱
۳-۲. حین اجرای سیاست	۱۱
۳-۳. پس از اجرای سیاست	۱۲
۴. انواع ارزیابی و روش‌های متداول در هر نوع	۱۳
۴-۱. انتخاب نوع ارزیابی	۱۳
۴-۲. روش‌های ارزیابی اثربخشی	۱۵
۴-۳. روش‌های ارزیابی ارزش در برابر پول پرداختی	۱۸
۵. انواع داده‌ها و شیوه بهینه انتخاب آن	۱۹
۵-۱. داده‌های ثبتي و نظارتی	۲۰
۵-۲. داده‌های پیمایشی در مقیاس بزرگ	۲۰
۵-۳. داده‌های جدید که به‌طور اختصاصی برای ارزیابی ایجاد شده‌اند	۲۱
۵-۴. داده‌های رسانه‌های اجتماعی	۲۱
۶. جمع‌بندی و نتیجه‌گیری	۲۱
منابع و مآخذ	۲۲

## فهرست شکل‌ها

شکل ۱. نمودار روند تغییرات فقر شدید در جهان و ایران	۹
شکل ۲. خلاصه‌ای از ارزیابی‌های لازم در هر مرحله از سیاست‌گذاری	۱۳
شکل ۳. پرسش اصلی انواع ارزیابی	۱۵
شکل ۴. مراحل انجام ارزیابی موفق	۱۹

## فهرست جدول

جدول ۱. پیشنهاد توصیه سیاستی در رابطه با ارزیابی اثربخشی سیاست‌های حمایتی	۲۲
---------------------------------------------------------------------------	----



## ارزیابی اثربخشی سیاست‌ها (۱): مفاهیم و روش‌ها

### چکیده



این گزارش به بررسی انواع و روش‌های ارزیابی سیاستی پرداخته است تا زمینه فنی لازم را برای پیگیری و ارزیابی میزان تحقق عدالت در حوزه‌های مورد هدف سیاست‌ها فراهم کند. در اهمیت این بحث، گفتنی است که اجرای مستمر برنامه‌های متنوع حمایتی و اختصاص سالیانه اعتباراتی به آنها در کنار باقی ماندن سطوح شدید فقر در مناطق مختلف همچنان به‌عنوان چالشی جدی برای برخی از کشورها مطرح است. از همین رو چندین دهه است که اثربخشی برنامه‌های حمایتی به یکی از دغدغه‌های اصلی فقر پژوهان تبدیل شده است و بسیاری از کشورها برنامه‌های حمایتی را فقط به شرط اثرگذاری آن بر کاهش فقر اجرا می‌کنند. به‌صورت کلی ارزیابی به سه دسته تقسیم می‌شود: ارزیابی فرایند، ارزیابی اثربخشی و ارزیابی ارزش در برابر پول پرداختی. هر سه انواع ارزیابی کاربردها و مزایایی دارند؛ اما به‌طور مشخص ارزیابی اثربخشی در چند دهه گذشته بیشتر مورد توجه قرار گرفته و پژوهش‌های بسیاری را در اقتصاد و سایر رشته‌ها به خود اختصاص داده است. روش‌های ارزیابی اثربخشی نظریه-محور (ارزیابی واقع‌بینانه، تحلیل مقایسه‌ای کیفی، تحلیل مشارکت و دنبال کردن فرایند)، ارزیابی اثربخشی آزمایشی (روش‌های آزمایش تصادفی کنترل شده) و ارزیابی اثربخشی شبه‌آزمایشی (زمان‌بندی رویدادها، سری زمانی مختل شده و تفاضل در تفاضل)، مهم‌ترین روش‌های ارزیابی اثربخشی است که مورد استفاده قرار می‌گیرد.



## بیان/شرح مسئله

هرساله اعتبارات قابل توجهی برای پیشبرد طرح‌های حمایتی و محرومیت‌زدایی هزینه می‌شود. این در حالی است که متناسب با این هزینه‌ها میزان فقر (به‌ویژه در یک دهه گذشته) تغییر نیافته یا به‌کندی کاهش پیدا کرده است. در دهه‌های گذشته، بسیاری از کشورها با این چالش مواجه شده‌اند که اجرای برنامه‌های فقرزدایی لزوماً منجر به کاهش فقر نمی‌شود و از این‌رو موضوع ارزیابی اثربخشی سیاست‌های حمایتی از سوی فقر پژوهان و حتی سیاست‌گذاران بسیار مورد توجه قرار گرفته است.

## نقطه‌نظرات / یافته‌های کلیدی

در دسته‌بندی کلی، ارزیابی سیاست شامل ارزیابی فرایند، ارزیابی اثربخشی و ارزیابی ارزش در برابر پول پرداختی می‌شود. ارزیابی فرایند تلاش می‌کند کلیه مراحل اجرای سیاست (از ورودی‌ها گرفته تا آخرین مرحله اجرا) را عارضه‌یابی و گلوگاه‌ها و موانع نیل به هدف اولیه سیاست (به‌عنوان مثال ارائه خدمات به متقاضیان یا پرداخت نقدی به آنها) را بررسی و معرفی کند.

ارزیابی اثربخشی نوع خاصی از ارزیابی‌هایی است که در پی پاسخ به پرسش رابطه علی است و آن اینکه اثر یک برنامه بر پیامد مورد نظر چیست؟ این نوع ارزیابی از ارزیابی فرایند گامی فراتر نهاده و در پی بررسی این پرسش است که آیا با اجرای سیاست مورد نظر، هدف اصلی سیاست (به‌عنوان مثال کاهش فقر) محقق شده است یا خیر.

ارزیابی ارزش در برابر پول پرداختی، هزینه‌ها و منافع مداخله‌ها را با یکدیگر مقایسه می‌کند. به بیان دیگر، این نوع ارزیابی در پی پاسخ به این پرسش است که عواید به دست آمده از اجرای سیاست مورد نظر چه تناسبی با هزینه‌های انجام شده دارد و آیا بهترین مداخله برای تحقق هدف مورد نظر بوده است؟ هر سه نوع ارزیابی کاربردها و مزایایی دارند؛ اما به‌طور مشخص ارزیابی اثربخشی در چند دهه گذشته بیشتر مورد توجه قرار گرفته و پژوهش‌های بسیاری را در اقتصاد و سایر رشته‌ها به خود اختصاص داده است. روش‌های ارزیابی اثربخشی نظریه-محور (ارزیابی واقع‌بینانه، تحلیل مقایسه‌ای کیفی، تحلیل مشارکت و دنبال کردن فرایند)، ارزیابی اثربخشی آزمایشی (روش‌های آزمایش تصادفی کنترل شده) و ارزیابی اثربخشی شبه‌آزمایشی (زمان‌بندی رویدادها، سری زمانی منقطع و تفاضل در تفاضل)، مهم‌ترین روش‌های ارزیابی اثربخشی است که مورد استفاده قرار می‌گیرد. همچنین نوع داده‌های موجود یا قابل دسترس در انتخاب نوع و روش ارزیابی نیز تأثیرگذار است.

## پیشنهاد راهکارهای تقنینی، نظارتی یا سیاستی

تجربه برنامه‌های حمایتی موفق در کشورهای مختلف نشان می‌دهد نظام پایش و ارزیابی مستمر (پیش از اجرا، حین اجرا و پس از اجرا) یکی از ارکان کارآمدسازی نظام حمایتی است. ارزیابی سیاست نیازمند انتخاب نوع روش و تمهید داده‌های مورد نیاز متناسب با روش مورد نظر است. بسیاری از سیاست‌ها و برنامه‌های اجرایی فاقد داده‌های کافی و لازم برای انجام دادن انواع ارزیابی (به‌ویژه ارزیابی اثربخشی) است؛ لذا پیشنهاد می‌شود برای افزایش نظارت پیشینی و پسینی بر اجرای سیاست‌ها و در اجرای بند «الف» ماده (۵۷) قانون احکام دائمی برنامه‌های توسعه کشور (براساس آن وزارت تعاون، کار و رفاه اجتماعی مکلف به اثرسنجی برنامه‌ها و اقدامات اقتصادی بر وضعیت و کیفیت زندگی اقشار و گروه‌های هدف است)، وزارت تعاون، کار و رفاه اجتماعی، نظام ارزیابی سیاستی را در قالب دستورالعمل طراحی و ابلاغ کند که به‌موجب آن موارد لازم برای امکان ارزیابی پیشینی و پسینی سیاست‌ها (شامل نوع ارزیابی، روش و داده‌هایی که باید در طول اجرای طرح تولید شوند) از زمان طراحی سیاست توسط مجریان آن تمهید شوند.



## ۱. مقدمه

این گزارش بنا دارد به بررسی انواع و روش‌های ارزیابی سیاستی بپردازد تا زمینه فنی لازم برای پیگیری و ارزیابی میزان تحقق عدالت در حوزه‌های مورد هدف سیاست‌ها را فراهم کند. سیاست‌ها و برنامه‌های حمایتی عمدتاً با هدف تغییر ملموس در پیامدها، همچون افزایش درآمد، بهبود سطح یادگیری یا کاهش میزان بیماری، طراحی می‌شوند. اینکه آیا سیاست‌های اجرا شده واقعاً این اهداف را محقق می‌کنند یا خیر، پرسش اساسی در سیاستگذاری عمومی است که اغلب مورد بررسی قرار نمی‌گیرد. مدیران برنامه‌ها و مجریان غالباً بر اندازه‌گیری و گزارش دهی درباره منابع طرح‌ها و خروجی‌های اولیه (و نه پیامدها و دستاوردها) متمرکز می‌شوند و به موضوعاتی نظیر اینکه چه مقدار اعتبار برای برنامه یا طرح مورد نظر صرف شده یا چه ظرفیت اجرایی برای این طرح تمهید شده است، بسنده می‌کنند. غافل از آنکه آنچه برای سیاستگذار اهمیت دارد، اهداف طرح است و پرسش اصلی آن است که آیا اجرای طرح توانسته است اهداف مورد نظر را محقق کند یا خیر؟ به عنوان مثال، طرح‌های مقابله با فقر بسیاری در دنیا در حال اجراست و عمدتاً هدف از اجرای این طرح‌ها کمک به فقرا و توانمندسازی آنها برای خروج از فقر است. گزارش‌های عمده این طرح‌ها حاکی از صرف منابع قابل توجه و پرداخت به خانوارهای فقیر است که طبیعی است این حمایت‌ها کمکی به وضعیت نامناسب آنها محسوب می‌شود. اما در عمده این طرح‌ها فراموش می‌شود که هدف خروج از فقر بوده است. آیا بعد از دوره‌ای مشخص تعدادی از خانوارهای مشمول توانسته‌اند از فقر خارج شوند و دیگر نیازمند آن طرح حمایتی نباشند؟ این پرسش اصلی سیاستگذار است که باید به آن پاسخ داده شود. سیاستگذاری شواهدمحور<sup>۱</sup> مفهومی است که در فضای سیاستگذاری عمومی برای پاسخ به این دغدغه سیاستگذاران شکل گرفته است و ارزیابی اثربخشی سیاست‌ها به عنوان بخشی از دایره وسیع سیاستگذاری شواهدمحور محسوب می‌شود.

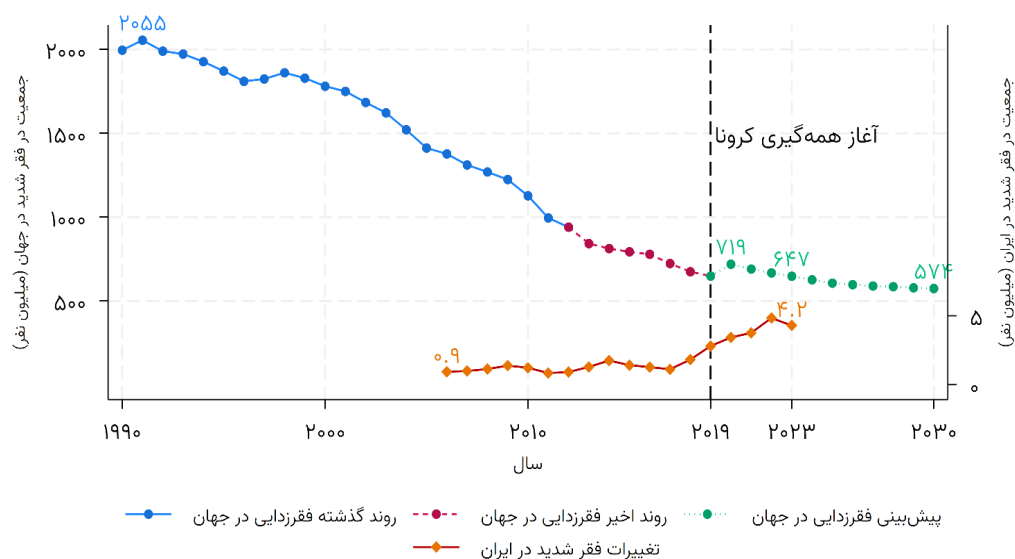
سیاستگذاری شواهدمحور در چند دهه گذشته از سوی سیاستگذاران بسیار مورد استقبال قرار گرفته است. همین تغییر رویکرد را می‌توان عامل اصلی دستاوردهای مهم در این بازه زمانی دانست. به عنوان مثال، بررسی وضعیت فقر در جهان نشان می‌دهد که جمعیت فقیر در چند دهه گذشته کاهش قابل توجهی را تجربه کرده است. بر اساس برآورد بانک جهانی، نرخ فقر در جهان در سال ۱۹۹۰ حدود ۳۸ درصد بوده که این رقم در سال ۲۰۱۹ به ۸٫۴ درصد کاهش یافته است [۶]. شکل ۱ تغییرات جمعیت فقیر (فقر شدید)<sup>۲</sup> در جهان از سال ۱۹۹۰ و همچنین پیش‌بینی تا سال ۲۰۳۰ را نشان می‌دهد. همان‌طور که در شکل ذیل قابل مشاهده است، کاهش فقر در چند دهه گذشته به میزان قابل قبولی محقق شده است.<sup>۳</sup> اگرچه چنین دستاوردهایی فقط متأثر از سیاست‌های حمایتی نیست و توسعه و رشد اقتصادی کشورها نیز بر آن بسیار تأثیرگذار است، پژوهش‌های متعدد در حوزه اقدامات فقرزدایی و چابکی نظام‌های حمایتی در تغییر رویکردها برای تحقق هدف اصلی (به عنوان مثال کاهش فقر) نشان از توجه بی‌سابقه به مفهوم سیاستگذاری شواهدمحور و ذیل آن ضرورت ارزیابی اثربخشی برنامه‌ها و طرح‌های اجرایی دارد.

1. Evidence-based Policy

۲. درآمد روزانه کمتر از ۲٫۱۵ دلار.

۳. شیوع ویروس کرونا در جهان باعث شد این روند دچار اختلال شود؛ لذا هدف‌گذاری کاهش فقر جهانی به ۲۵۰ میلیون نفر در سال ۲۰۳۰ دور از دسترس به نظر می‌رسد و پیش‌بینی می‌شود تا سال ۲۰۳۰ بیش از ۵۰۰ میلیون نفر همچنان در وضعیت فقر شدید قرار داشته باشند.

شکل ۱. نمودار روند تغییرات فقر شدید در جهان و ایران [۱]، [۲]



این رویکرد، یعنی حساسیت به روند فقرزدایی در سه دهه گذشته، موجب شده است تا پژوهشگران درباره کندی این روند در چند سال گذشته هشدار دهند و با توجه به تحولاتی همچون شیوع بیماری کرونا پیش‌بینی‌های خود را مبنی بر کاهش شدید سرعت فقرزدایی ارائه کنند. اکنون براساس پیش‌بینی بانک جهانی، همچنان حدود ۶۵۰ میلیون نفر در جهان در فقر شدید به سر می‌برند.

در شکل ۱، روند تغییرات جمعیت در فقر شدید در ایران نیز مشاهده می‌شود. همان‌طور که پیش‌تر نیز مطرح شد، افزایش نرخ فقر در کشور در دهه گذشته فقط متأثر از نظام حمایتی نیست و تحولات کلان اقتصادی را در چند سال گذشته شاید بتوان مهم‌ترین عامل افزایش فقر در کشور بیان کرد. با این حال، نمی‌توان انکار کرد که نبود کنترل بر رشد فقر شدید، ناشی از ناکارآمدی سیاست‌های حمایتی است؛ زیرا کنترل و کاهش فقر شدید به‌نوعی اولین مأموریت نظام حمایتی محسوب می‌شود. رشد جمعیت در فقر شدید در کشور در شرایطی رخ داده است که در چند سال گذشته ریشه‌کنی فقر شدید به‌عنوان مطالبه جدی از سوی سیاستگذاران مطرح شده<sup>۱</sup> و علاوه بر حمایت‌های نقدی، طرح‌های حمایتی مکملی، همچون طرح ملی کالا بزرگ الکترونیکی، نیز به اجرا درآمده است. در سال ۱۴۰۲، بیش از چهارصد هزار میلیارد تومان در قالب طرح‌هایی همچون یارانه نقدی و مستمری نهادهای حمایتی به‌صورت مستقیم به خانوارها پرداخت شده است. صرف نظر از کافی یا ناکافی بودن اعتبارات این طرح‌ها در قوانین بودجه سنواتی، پرسشی که باید در پی پاسخ آن بود، میزان موفقیت و اثربخشی سیاست‌های حمایتی است.

- آیا نحوه اجرای برنامه‌های حمایتی به‌صورت کارا انجام گرفته است؟
- چقدر از منابع در نظر گرفته شده برای هر برنامه حمایتی به فقرا اصابت کرده است؟
- آیا دولت با پرداخت این هزینه، توانسته است به هدف خود، یعنی کاهش فقر و محرومیت، دست یابد؟

۱. یکی از مهم‌ترین مطالبه مسئولان نظام، دستور شهید رئیسی مبنی بر اقدام فوری برای «ریشه‌کنی فقر مطلق» در اسفندماه سال ۱۴۰۰ است. اگرچه پس از این دستور، اراده مناسبی میان دستگاه‌های اجرایی در سراسر کشور مشاهده شد؛ با این حال، همراهی مناسبی با طرح وزارت تعاون، کار و رفاه اجتماعی (موسوم به طرح حامی) از سوی سایر دستگاه‌های فعال در این حوزه صورت نگرفت و پس از مدتی این مطالبه رئیس‌جمهور شهید به فراموشی سپرده شد.



البته بی توجهی به ارزیابی سیاست‌های حمایتی در کشور به واسطه نبود پشتوانه قانونی نیست. به‌عنوان مثال در بند «الف» ماده (۵۷) **قانون احکام دائمی برنامه‌های توسعه کشور** آمده است: «به‌منظور بسط و گسترش هر چه بیشتر عدالت اجتماعی در جامعه، کاهش فقر و نابرابری، تحقق اقتصاد مقاومتی و عدالت بنیان کردن توسعه اقتصادی، وزارت تعاون، کار و رفاه اجتماعی مکلف است نسبت به رصد شاخص‌های رفاه و تأمین اجتماعی و تقویم و اثرسنجی برنامه‌ها و اقدامات اقتصادی بر وضعیت و کیفیت زندگی اقشار و گروه‌های هدف اقدام کند». لذا این بی توجهی را باید در فقدان دانش کافی در این حوزه و همچنین ساختار نهادی ناکارآمد نظام حمایتی جست‌وجو کرد. این گزارش (و گزارش‌های آتی که با جزئیات بیشتر بحث را بسط خواهند داد) با هدف تقویت جایگاه ارزیابی‌های پسینی و پیشینی در اجرای برنامه‌های حوزه حمایتی (و بلکه در سایر سیاست‌ها)، تلاش دارد روش‌های ارزیابی سیاست‌ها و برنامه‌ها را به‌صورت کاربردی معرفی کند. پیش از تمرکز بر روش‌های ارزیابی اثربخشی در این مجموعه، گزارش حاضر به‌عنوان اولین گزارش از این مجموعه، مروری مختصر بر انواع روش‌های پایش و ارزیابی خواهد داشت و با بیان اینکه هر یک از این روش‌ها قصد دارد به چه پرسش‌هایی پاسخ دهد، تفاوت آن روش‌ها را تبیین خواهد کرد. بدین منظور، ابتدا مفهوم ارزیابی سیاستی بررسی و سپس، روش‌های مختلف ارزیابی سیاستی به‌طور جزئی‌تر تحلیل می‌شود؛ پس از آن، روش انتخاب داده مناسب بیان خواهد شد.

## ۲. ارزیابی سیاستی چیست؟



وزارت خزانهداری انگلستان در دستورالعملی برای ارزیابی سیاست، این تعریف را از ارزیابی سیاستی ارائه کرده است: «ارزیابی سیاستی، برآوردی سیستماتیک از طراحی، اجرا و نتایج سیاست دولت است که شامل درک چگونگی مداخله دولت و تأثیرات آن، افراد ذی‌نفع و چرایی آن است. همچنین، ارزیابی سیاستی شامل مشخص کردن آنچه می‌تواند بهبود یابد و چگونگی آن در کنار تخمین اثربخشی کل و مقرون‌به‌صرفه بودن است» [۵].

ارزیابی سیاستی مسیر دستیابی به هدف را برای سیاستگذار روشن می‌کند. برای مثال، سیاستی ممکن است پرهزینه و با اثرگذاری پایین باشد یا به افرادی منفعت برساند که هدف اصلی سیاستگذار نبوده است. در این حالت، ارزیابی به کمک وی می‌شتابد تا برای هدف خود (برای مثال کاهش فقر) بهینه‌ترین سیاست را انتخاب کند و ریسک‌ها و هزینه‌ها را به حداقل ممکن برساند. بنابراین، ارزیابی قبل، حین و پس از اجرای سیاست کنار سیاستگذار خواهد بود تا تصویر روشن‌تری را از نتیجه‌بخشی و انتخاب روش بهینه ارائه کند.

### ۳. چه زمانی باید ارزیابی صورت گیرد؟



به همان اندازه که ارزیابی یک سیاست ارزشمند است، به موقع انجام دادن آن نیز ضروری است. برای مثال، یک ارزیابی اصولی قبل از اجرای سیاست، می‌تواند مانع از خسارت‌های مالی و تعهدات بلندمدت (بعضاً جبران‌ناپذیر) به دولت و جامعه شود. برخلاف تصور عموم، که ارزیابی یک سیاست را فقط پس از اجرای سیاست در نظر می‌گیرند، بازه‌های ارزیابی به سه دسته کلی قبل، حین و پس از اجرای سیاست حمایتی تقسیم می‌شود. هر بازه نیازمند بررسی‌های متناسب با خود است که در ادامه ذکر می‌شود:

#### ۳-۱. قبل از اجرای سیاست

قبل از اجرای سراسری یک سیاست، ارزیاب از طریق روش‌هایی نظیر اجرای آزمایشی (اجرای طرح برای نمونه کوچک برای بررسی امکان‌پذیری اجرا در ابعاد بزرگ‌تر) و شواهد موجود به این پرسش‌ها پاسخ می‌دهد:

- انتظار می‌رود اجرای سیاست چگونه کار کند و چه شواهدی این ایده را تأیید می‌کند؟
- در چه بخش‌هایی از کار، نااطمینانی و ریسک وجود دارد؟
- این سیاست در ابعاد کوچک و محیط کنترل‌شده چگونه عمل می‌کند؟
- چه شواهدی را به‌عنوان نماینده‌ای از وضع موجود استفاده کنیم تا بتوانیم در آینده تغییر را با آنها اندازه‌گیری کنیم؟

نتیجه این ارزیابی به سیاستگذار کمک می‌کند تا طراحی نسبتاً دقیقی از مداخله و شیوه اجرای آن داشته باشد. به عبارت دیگر، از آنجایی که منابع عمومی محدود است، فرصتی برای آزمون و خطا در ابعاد بزرگ وجود ندارد. از همین رو، ارزیابی قبل از اجرای سیاست به سیاستگذار کمک می‌کند تا بهترین روش را برای حداکثرسازی رفاه جامعه و ریشه‌کن کردن فقر برگزیند. از نمونه‌های اجرای پایلوت در ایران، طرح پزشک خانواده است که در برخی از استان‌ها به صورت آزمایشی به اجرا گذارده شد تا اشکالات و میزان اثرگذاری آن پیش از اجرا در سطح کلان شناسایی شود. همچنین طرح کالابریگ الکترونیکی پیش از اجرا در سراسر کشور، در استان هرمزگان به صورت آزمایشی اجرا شد.

#### ۳-۲. حین اجرای سیاست

اجرای هر سیاست اقتصادی با هزینه‌های آشکار و پنهان همراه است. سیاست‌های رفاهی نیز از این قاعده مستثنا نیست. نمی‌توان سیاستی رفاهی را بدون تحمیل هزینه به بخشی از جامعه اجرا کرد. سیاستگذار باید هزینه‌ها و منافع هر سیاست را کنار یکدیگر ارزیابی کند تا به تصمیم بهتری دست یابد. تعیین ردیف جدید در بودجه برای پوشش نیازهای ضروری خانواده‌های فقیر، به کارگیری نیروی انسانی برای اجرای سیاست، نظارت بر درست اجرا شدن فرایند و در نهایت بازبینی نتیجه نیاز به هزینه‌های اجرایی قابل توجهی دارد. ارزیابی دقیق و پیوسته در تمامی مراحل، می‌تواند سیاستگذار را مطمئن کند که تحمیل این هزینه‌ها، در نهایت وی را به هدف مطلوب (کاهش فقر) می‌رساند. در حین اجرای سیاست، ارزیاب به دنبال بررسی مشاهدات درباره اثربخشی طرح، اجرا و نتایج در حال ظهور آن است. در این مرحله به این پرسش‌ها پاسخ داده می‌شود:

- آیا مزایای سیاست به گروه هدف اصابت می‌کند؟
- آیا سیاست طبق برنامه اجرا می‌شود؟
- آیا سیاست در مناطق مختلف و برای گروه‌های هدف متفاوت، عملکرد یکسانی دارد؟
- نشانه‌های اولیه از میزان اثرگذاری سیاست مشاهده می‌شود؟



● آیا نتایج ناخواسته هم اتفاق افتاده است؟

● آیا طراحی یا شیوه رسیدن کمک‌ها به جامعه هدف می‌تواند بهبود یابد؟

در ایران می‌توان به طرح هدفمندسازی یارانه‌ها به عنوان نمونه‌ای در حال اجرا از سیاست حمایتی اشاره کرد. چند نمونه از پرسش‌هایی که سیاستگذار در ابتدای اجرای سیاست باید به آن پاسخ می‌داد عبارت‌اند از: آیا **قانون هدفمندی یارانه‌ها** توانسته است شرایط جامعه هدف - که ضعیف‌ترین قشر جامعه‌اند - را بهبود بخشد؟ آیا افراد فقیر ساکن شهر همانند روستاییان تحت تأثیر این سیاست قرار گرفته‌اند؟ آیا مسیری که در عمل اتفاق افتاد، همان پیش‌بینی سیاستگذار بود؟ آیا می‌توان همین منابع را به شیوه صحیح‌تری تقسیم کرد تا عواید بیشتری برای قشر محروم داشته باشد؟ این پرسش‌ها کمک می‌کند تا هزینه‌هایی که جامعه از مداخله می‌پردازد، به حداقل ممکن برسد و در مقابل، منافعی که کسب می‌شود، حداکثر افراد را پوشش دهد. در واقع هدف اصلی از ارزیابی حین اجرای سیاست، اطمینان بخشی از صحت عملکرد برنامه، رفع نواقص اجرایی و گرفتن بازخورد اولیه از اثر بخشی سیاست است.

### ۳-۳. پس از اجرای سیاست

وقتی یک سیاست حمایتی به‌طور کامل اجرا شد، کل سیاست را می‌توان با بررسی داده‌های نهایی طراحی، اجرا و نتایج، بررسی کرد و در این بخش نیز تلاش می‌شود به این پرسش‌ها پاسخ داده شود:

● آیا این سیاست حمایتی به اهداف طراحی شده رسید؟

● میزان اثربخشی و هزینه سیاست چقدر بود؟

● آیا بین گروه‌های مختلف فرق داشت؟

● سیاست اعمال شده چقدر توانست بر نتیجه اثرگذار باشد؟

● آیا نتایج با آنچه در ارزیابی پیش‌بینی شده بود، مطابقت داشت؟

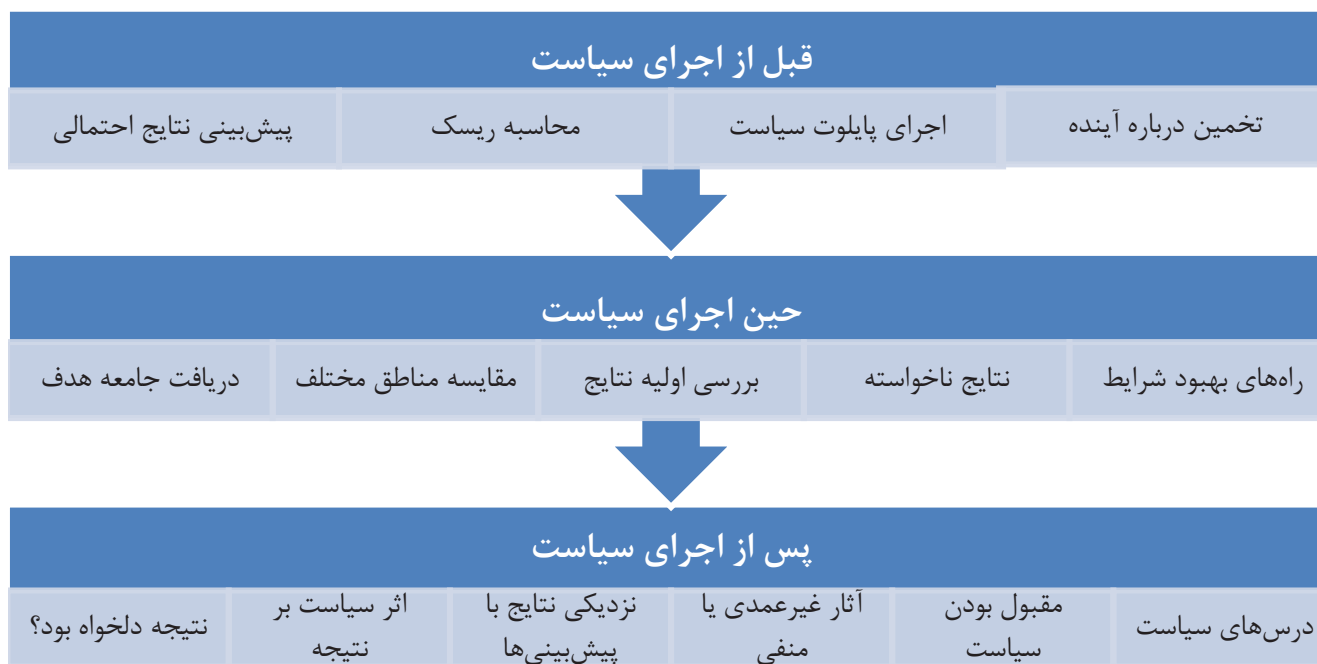
● آیا اثرهای نامطلوب غیر عمدی یا تصادفی هم وجود داشت؟

● آیا منافع این سیاست بر هزینه‌هایش برتری داشت؟

● آموخته‌های قابل انتقال به دیگران چیست؟

بنابراین در این مرحله، علاوه بر آنکه مشخص می‌شود تخمین‌ها چقدر به واقعیت نزدیک بوده است، مسیر طی شده به‌عنوان راهگشایی برای اتخاذ تصمیمات بهتر در ارزیابی آتی سیاستگذاران خواهد بود. خلاصه‌ای از ارزیابی‌های لازم در مراحل سه‌گانه قبل، حین و پس از اجرای یک سیاست حمایتی در شکل ۲ ارائه شده است.

شکل ۲. خلاصه‌ای از ارزیابی‌های لازم در هر مرحله از سیاستگذاری [۲]



## ۴. انواع ارزیابی و روش‌های متداول در هر نوع



اکنون که مفهوم اولیه ارزیابی مورد بررسی قرار گرفت و اهمیت آن بیان شد، اینجا شیوه صحیح طراحی یک ارزیابی بررسی می‌شود. برای این کار، ارزیاب ابتدا به تعیین محدوده ارزیابی می‌پردازد و سپس از میان انواع مختلف روش‌های مداخله، بهترین گزینه را متناسب با شرایط هر مداخله انتخاب می‌کند. در ادامه، این دو بخش به‌طور دقیق‌تر بررسی خواهد شد [۲].

### ۴-۱. انتخاب نوع ارزیابی

ابتدا سیاستگذار باید چارچوب درستی را از ارزیابی سیاستی خود تعیین کند. برای مثال، هدف اصلی سیاستگذار رفع فقر در جامعه است. برای دستیابی به این هدف، باید بدانیم چه نهادهایی درگیر این سیاست خواهند شد و میزان اثرگذاری‌شان چقدر است؟ رفع فقر در کدام حوزه، اولویت اول سیاستگذار است و کدام حوزه اهمیت کمتری دارد؟ تراکم فقر در چه استان‌هایی بیشتر از دیگر استان‌هاست؟ با توجه به تمامی قیود، کدام استان اولویت بالاتری برای رفع فقر دارد؟ در هر استان، چه میزان بودجه‌ای برای این هدف قابل تأمین است؟ در کدام استان‌ها به‌علت محدودیت دسترسی به داده‌های خانوار، مشکلات بیشتری در شناسایی خانوار فقیر وجود دارد؟ آیا نتایج اجرای طرح به‌صورت کمی قابل بررسی است؟ (این مورد بسیار مهم است؛ زیرا اگر نتوان نتایج را ارزیابی کرد، نمی‌توان درباره کیفیت سیاست نظری داد). بنابراین، برای تحقق هدف سیاستگذار لازم است که مرحله تعیین محدوده با دقت بالایی انجام شود و در این راستا درک نوع مداخله، پرسش‌هایی که سیاستگذار در پی پاسخ دادن به آن است و انتخاب نوع ارزیابی اهمیتی دوچندان می‌یابد. بدین منظور در نخستین قدم، از نظریه تغییر<sup>۱</sup> استفاده می‌شود. نظریه تغییر دیاگرام یا توصیفی مکتوب از استراتژی‌ها، اقدامات، شرایط و منابعی است که براساس این

1. Theory of Change



موارد چهارگانه، تغییری روی می‌دهد یا خروجی جدیدی تولید می‌شود. به بیان ساده، نظریه تغییر توضیحی برای این پرسش است که چرا فعالیت‌ها یا اقدامات خاص منجر به نتایج خاصی می‌شوند.

نظریه تغییر می‌تواند درک کاملی از نحوه مداخله در عمل (مثل مشکلی که مداخله می‌خواهد حل کند و تغییراتی که می‌خواهد پدید آورد، مسیری که به تغییر منجر می‌شود، افرادی که تحت تأثیر قرار می‌گیرند و شرایط موفقیت این مداخله)، فروض پشت هر مداخله (و قوت یا ضعف شواهدی که این فرضیات را پشتیبانی می‌کنند) و اثرهایی که بر سایر بخش‌ها دارد (تغییرات سیاستی، اقتصادی یا محیطی)، ارائه دهد. همچنین، به مجریان سیاست اجازه می‌دهد تا طراحی مداخله را بیازمایند و از کارایی آن اطمینان حاصل کنند.

در نظریه تغییر، استفاده از داده‌های موجود نقشی کلیدی دارد. با گردآوری شواهد موجود درباره مداخله، ارزیاب می‌تواند بخش‌هایی را که کیفیت داده‌ها ضعیف است بیابد. این کار کمک بزرگی به شناسایی پرسش‌هایی، که ارزیاب باید به آنها پاسخ دهد، می‌کند. در نهایت، نظریه تغییر باید با گسترش مشاهدات جدید توسعه یابد.

پس از این مرحله، نوبت به طراحی پرسش‌های ارزیابی می‌رسد. ارزیابی می‌تواند برای پاسخ به پرسش‌های متنوعی طراحی شود. بنابراین، لازم است از ابتدا مشخص شود که پرسش‌های هر ارزیابی کدام‌اند و انتظار می‌رود که پاسخ آنها چگونه، توسط چه کسانی و در چه زمانی مورد استفاده قرار گیرند. این پاسخ‌ها می‌توانند به انتخاب رویکرد ارزیابی مناسب نیز کمک کنند. قاعدتاً هر چه تعداد پرسش‌ها بیشتر باشد، چارچوب ارزیابی دقیق‌تر خواهد بود؛ اما باید در نظر داشت که زمان، منابع و روش ارزیابی هر سه محدودند و نمی‌توانند پرسش‌های زیادی را پوشش دهند.

به صورت کلی، ارزیابی به سه دسته تقسیم می‌شود: فرایند، اثربخشی و ارزش در برابر پول پرداختی. برای درک جزئیات هر ارزیابی و تفکیک پرسش‌هایی که به آنها پاسخ می‌دهند، به مطالعه این سه بخش می‌پردازیم.

### ۱-۱-۴. ارزیابی فرایند<sup>۱</sup>

ارزیابی فرایند، مقایسه بین ورودی، فعالیت‌ها و نتیجه برنامه‌ریزی شده با واقعیت است [۳]. به عبارت دیگر، ارزیابی فرایند به بررسی فعالیت‌هایی، که در اجرای یک سیاست دخالت داشته‌اند و مسیرهایی که این سیاست به افراد رسیده است، می‌پردازد. بنابراین تلاش بر آن است که به این پرسش‌ها پاسخ داده شود:

- آیا مداخله به نحوی که در نظر گرفته شده بود، اجرا شده است؟
- آیا طراحی انجام شده کار می‌کند؟
- کدام بخش کمتر یا بیشتر از پیش‌بینی کار می‌کند و علت آن چیست؟

### ۲-۱-۴. ارزیابی اثربخشی<sup>۲</sup>

ارزیابی اثربخشی نوع خاصی از ارزیابی‌هایی است که در پی پاسخ به پرسش رابطه علی است و آن اینکه اثر یک برنامه بر پیامد مورد نظر چیست؟ به عبارت دیگر، هدف از ارزیابی اثربخشی، تشخیص تغییرات بر پیامدی مشخص است که به علت اجرای یک سیاست صورت گرفته است. همچنین، تلاش می‌شود تا اندازه اثرگذاری سیاست نسبت به حالتی که مداخله صورت نمی‌گرفت، مقایسه شود. محاسبه این مقدار کاری پیچیده است؛ زیرا تغییرات صورت گرفته می‌تواند تحت تأثیر عوامل دیگری چون اتفاقات تصادفی، سایر مداخلات و سوگیری در انتخاب داده باشد.

### ۳-۱-۴. ارزیابی ارزش در برابر پول پرداختی<sup>۳</sup>

همان‌طور که گفته شد، اثر مداخله بر نتایج توسط روش ارزیابی اثربخشی محاسبه می‌شود. با این حال، این روش به تنهایی نمی‌تواند درباره

1. Process Evaluation  
2. Impact Evaluation  
3. Value-for-money Evaluation

بهبودن اجرای سیاست نظر بدهد. در این شرایط، ارزیابی ارزش در برابر پول پرداختی می‌تواند تحلیل دقیق‌تری ارائه دهد. این روش به پرسش‌هایی از این قبیل پاسخ می‌دهد: آیا منافع یک سیاست از هزینه‌های آن پیشی گرفته است؟ آیا این روش مداخله همچنان مؤثرترین روش برای استفاده از منابع است؟

به عبارت ساده‌تر، این روش هزینه‌ها و منافع نهایی کسب شده از طریق مداخله را با پیش‌بینی اولیه مقایسه می‌کند. همچنین می‌تواند به بررسی هزینه‌ها و منافع سایر روش‌هایی بپردازد که به همین هدف می‌رسید. بنابراین، این روش می‌تواند درباره معقول بودن ادامه اجرای یک مداخله از طریق مقایسه سود و هزینه‌های هر سیاست و سایر سیاست‌های پیشنهادی نظر بدهد.

بنابراین، هر نوع ارزیابی می‌تواند به پرسش‌هایی بپردازد که سایر روش‌ها از پاسخ به آن عاجزند. در ادامه، مهم‌ترین پرسش‌ها هر ارزیابی به‌طور خلاصه در شکل ۳ نمایش داده شده است.

شکل ۳. پرسش اصلی انواع ارزیابی [۳]



در میان انواع روش‌های ارزیابی، ارزیابی اثربخشی کمتر شناخته شده است. بسیاری از پژوهش‌های سیاستی در سال‌های اخیر به این نوع ارزیابی توجه ویژه‌ای نشان داده‌اند؛ زیرا ارزیابی اثربخشی به دغدغه اصلی سیاستگذار، یعنی اثرگذاری واقعی بر شاخص‌های مورد نظر، پاسخ می‌دهد. همچنین پیچیدگی‌های این نوع ارزیابی از سایر انواع آن بیشتر است. در ادامه به‌طور مشخص روش‌های ارزیابی اثربخشی مورد بررسی قرار خواهد گرفت.

## ۲-۴. روش‌های ارزیابی اثربخشی

هر یک از انواع ارزیابی روش‌های متعددی برای محاسبه دارد که انتخاب بین آنها براساس عوامل متعددی است. نظریه تغییر، نااطمینانی‌ها و فروض چند نمونه از این معیارها هستند.

برای انتخاب روش مناسب، باید چند نکته مهم مد نظر قرار داده شود:

الف) سیاستگذار ممکن است نیازهای متفاوتی داشته باشد که به تبع آن نتایج احتمالی هر سیاست را اولویت‌بندی کند؛ لذا سیاستگذار باید در مرحله انتخاب روش ارزیابی به‌طور فعال مشارکت داده شود تا نیازها و معیارهای وی برای تمایز قائل شدن میان سیاست‌های مختلف شناسایی شود. بدین ترتیب وی می‌تواند اولویت‌ها را مشخص کند، انتظارات را مدیریت و نتایج را تأیید کند.



ب) هیچ روشی نمی‌تواند به تمام پرسش‌های ارزیابی پاسخ دهد. از آنجایی که زمان و منابع محدودند، پرسش‌ها و روش‌ها باید اولویت‌بندی شوند. ج) هنگام انتخاب روش، باید دو عامل امکان‌پذیر بودن<sup>۱</sup> و متناسب بودن<sup>۲</sup> آن روش را در نظر گرفت. هزینه، زمان‌بندی و نرخ احتمالی پاسخ دادن چند نمونه از عواملی است که هنگام بررسی‌های منتهی به انتخاب روش باید در نظر گرفت. درباره این عوامل در بخش جمع‌آوری داده بیشتر صحبت خواهد شد.

با توجه به شرح مختصری که درباره انتخاب روش مناسب گفته شد، به بررسی انواع روش‌های ارزیابی اثربخشی می‌پردازیم:

### ۱-۲-۴. روش‌های ارزیابی اثربخشی نظریه-محور<sup>۳</sup>

با این روش می‌توان به این پرسش پاسخ داد که آیا علت تغییر اجرای سیاست بوده است یا عوامل دیگری بر نتیجه مشاهده شده تأثیر داشته‌اند. مثلاً برای جلوگیری از سوءتغذیه فرزندان والدین معتاد، پخش بسته‌هایی شامل اقلام غذایی ضروری پیشنهاد می‌شود. با این حال، نتایج نشان می‌دهد وزن فرزندان تغییر معناداری نداشته است. در نگاه اول، می‌توان ادعا کرد تهیه بسته غذایی نمی‌تواند مشکل سوءتغذیه را حل کند. با این حال، بررسی دقیق‌تر فرایند اجرای سیاست ممکن است نشان دهد که فرضاً مسئولان محلی توزیع‌کننده اقلام غذایی با ایجاد بازار سیاه، اقلام غذایی را با قیمتی بالاتر به دیگران فروخته‌اند. بنابراین، نبود رابطه معنادار بین توزیع بسته و وزن فرزندان به علت نبود اجرای درست سیاست بوده است، نه اثرگذار نبودن آن. ذکر این نکته ضروری است که این روش تخمین دقیقی از میزان اثر ارائه نمی‌دهد. در مقابل، به تعداد کسانی که مشمول آن شده‌اند و علت ایجاد تغییر می‌پردازد.

روش‌های نظریه-محور به طور خاص برای ارزیابی مداخلات پیچیده یا مداخله ساده در محیط‌های پیچیده مناسب‌اند. در این شرایط، که تعیین اندازه اثر اغلب دشوار است، این روش می‌تواند تأیید کند که آیا مداخله در جهت مورد نظر تأثیر گذاشته است یا خیر. در بسیاری از این روش‌ها، هدف محاسبه دقیق میزان تأثیر مداخله نیست. در مقابل، این روش‌ها در تلاش‌اند تا نشان دهند آیا مداخله به طور قطعی در تغییر ایجاد شده دخیل بوده است یا خیر. همچنین، توضیح دهند که چرا مداخله‌ای مؤثر بوده (یا نبوده) و بتوانند آن را به سایر جمعیت‌ها، مکان‌ها یا دوره‌های زمانی تعمیم دهند.

روش‌های ارزیابی اثربخشی نظریه-محور برای شرایطی کاربرد دارد که یک یا چند مورد زیر برقرار است:

۱. سیاست اعمال شده، پیچیده با ترکیبی از مداخلات است.

۲. مداخله برای ایجاد تغییر در سیستمی پیچیده طراحی شده است.

۳. نتایج در حال تولیدند و بنابراین در ابتدا قابل پیش‌بینی نیستند.

۴. امکان ایجاد پادواقع<sup>۴</sup> مناسب وجود ندارد.

۵. آیا همین نتایج در مکان یا شرایط متفاوتی حاصل می‌شد یا خیر.

### ۲-۲-۴. ارزیابی اثربخشی با استفاده از روش‌های آزمایشی<sup>۵</sup> و شبه‌آزمایشی<sup>۶</sup>

به کمک این دو رویکرد، تلاش می‌شود تا میزان اثربخشی سیاست از طریق تقسیم‌بندی نمونه به دو دسته مداخله و کنترل اندازه‌گیری شود. برای درک بهتر، ابتدا این روش را با مثالی بررسی می‌کنیم. فرض کنید دولت برای تشویق ادامه تحصیل فرزندان خانوارهای فقیر، ایده پخش یک وعده غذای گرم در مدرسه را اجرا می‌کند. برای بررسی اثرگذاری این روش، باید چند مدرسه تصادفی از دو منطقه فقیر شهر، که هیچ ارتباطی با یکدیگر ندارند، انتخاب شود (مسئله تفکیک افراد اهمیت زیادی در نتایج دارد و باید با دقت انجام شود). سپس در برخی

1. Feasibility  
2. Appropriateness  
3. Theory Based  
4. Counterfactual  
5. Experimental  
6. Quasi-experimental

از این مدارس (گروه آزمایش) یک وعده غذای گرم توزیع شود و در مدارس دیگر مداخله‌ای صورت نمی‌گیرد (گروه شاهد). سپس در بازه حداقل یک‌ساله بررسی شود که آیا میزان مراجعه والدین برای ثبت نام فرزندان‌شان و حضور کودک در مدرسه افزایش یافته است یا خیر. از این طریق، می‌توان تخمینی حدودی برای اثر اجرای این سیاست در ابعاد بزرگ‌تر داشت. بنابراین، مهم‌ترین مؤلفه این روش، پیدا کردن دو نمونه است که در تمامی شرایط با یکدیگر مشابه‌اند و تنها تفاوت آنها در دریافت یا دریافت نکردن وعده غذای روزانه است.

به صورت تخصصی‌تر در این روش، مقایسه (آماری) بین گروه مشارکت‌کننده در مداخله و گروهی، که تحت تأثیر این مداخله نبوده‌اند، انجام می‌شود. ابتدا باید جامعه‌ای مشابه جامعه مورد بررسی پیدا شود که تنها تفاوت بین آنها اجرای سیاست در آنها باشد. در این صورت، گروهی که تحت تأثیر سیاست نبوده‌اند، مثل تقریباً برای حالتی محسوب می‌شوند که گروه درگیر مداخله، اگر تحت تأثیر سیاست نبودند، برایشان رخ می‌داد. به این شرایط، رویکرد پادواقع گفته می‌شود. علت این کار آن است که داشتن گروه کنترل به ارزیاب اطمینان می‌دهد که هر تغییر صورت گرفته فقط ناشی از مداخله بوده است.

ذکر این نکته ضروری است که برای استفاده از این روش لازم است داده‌ها از کیفیت و کمیت کافی برای پشتیبانی از تحلیل برخوردار باشند، گروه شاهد با گروه آزمایش واقعاً قابل مقایسه باشند و در نهایت، تأثیر مداخله به اندازه کافی بزرگ باشد که بتواند از خطای تصادفی انتظاری داده‌ها قابل تشخیص باشد.

انتخاب بهترین روش معمولاً به این موارد بستگی دارد:

۱. ماهیت تخصیص گروه شاهد و آزمایش (برای مثال، تحت کنترل ارزیاب بوده یا ناشی از یک شوک خارجی؟)؛
۲. ماهیت گروه شاهد (می‌تواند همزمان با گروه مداخله یا مربوط به گذشته باشد)؛
۳. قالب داده‌های موجود (مثلاً گسسته‌اند یا روند زمانی را پوشش می‌دهند؟)؛
۴. حجم داده‌های در دسترس.

بر این اساس، روش‌های مورد استفاده در این گروه به چند دسته تقسیم می‌شوند:<sup>۲</sup>

**الف) روش‌های آزمایش تصادفی کنترل شده:**<sup>۳</sup> روشی برای مقایسه اثر مداخله با حالتی که اگر مداخله نبود، اتفاق می‌افتاد. این روش شرکت‌کنندگان را به صورت تصادفی در گروه کنترل و مداخله قرار می‌دهد (به همین دلیل می‌تواند تمامی عوامل شناخته شده و شناخته نشده را برای هر دو گروه مدیریت کند) و میزان تفاوت در نتیجه (تغییر مورد نظر) را به عنوان اثر مداخله لحاظ می‌کند. این روش زمانی کاربرد دارد که کسانی که تحت تأثیر مداخله هستند و کسانی که تحت تأثیر نیستند، هر دو به صورت تصادفی انتخاب شده باشند و همان‌طور که از نام این روش پیداست، در محیطی آزمایشگاهی این دو گروه تعیین شده‌اند که حضور افراد در هر یک از دو دسته به معنای واقعی تصادفی باشد و دو گروه در هیچ معیاری، غیر از تحت تأثیر مداخله بودن، تفاوتی نداشته باشند.

**ب) سری زمانی منقطع و تفاضل در تفاضل:**<sup>۴</sup> این روش‌ها برای محاسبه تأثیر یک سیاست، از داده‌های سری زمانی استفاده می‌کنند. در روش سری زمانی مختل شده (به کمک روش‌های تحلیل آماری) داده‌های بلندمدت قبل و بعد از مداخله مورد بررسی قرار می‌گیرد تا اثر سیاست (به عنوان عامل کنترل شده خارجی) محاسبه شود. برای مثال، در فرایند تصمیم برای تهیه یک وعده غذای رایگان، دو مدرسه را، که هیچ ارتباطی با هم ندارند، به صورت تصادفی انتخاب می‌کنند. سپس این وعده غذایی را در یک بازه طولانی به یک مدرسه می‌دهند و مدرسه دیگر را از این مداخله محروم می‌کنند. سپس نرخ بازگشت دانش‌آموزان به این دو مدرسه را با هم مقایسه می‌کنند. در روش تفاضل در تفاضل نیز از گروه شاهد و آزمایش استفاده می‌شود. در این حالت با مقایسه اختلاف میان میانگین تغییر در طول زمان در گروه آزمایش با

1. Proxy

۲. ذکر این نکته ضروری است که تمامی این روش‌ها تخمینی از اثر خالص مداخله را ارائه می‌دهند؛ اما در اینکه همه افراد یا زیرگروهی از آنان را مورد بررسی قرار می‌دهند، متفاوت‌اند.

3. Randomized Controlled Trials (RCT)

4. Interrupted Time Series Analysis and Difference-in-difference



میانگین تغییر در طول زمان برای گروه شاهد، تأثیر یک سیاست بررسی می‌شود. در روش تفاضل در تفاضل برخلاف روش آزمایش تصادفی کنترل شده، گروه تحت تأثیر مداخله و گروهی که تحت تأثیر مداخله نبودند، به صورت تصادفی انتخاب نشده‌اند و تفاوت‌های دیگری نیز با یکدیگر خواهند داشت که این کار ارزیابی اثر مداخله را سخت‌تر می‌کند.

### ۳-۴. روش‌های ارزیابی ارزش در برابر پول پرداختی

ارزیابی ارزش در برابر پول پرداختی، هزینه‌ها و منافع مداخله‌ها را با یکدیگر مقایسه می‌کند. دو روش متداول این ارزیابی، تحلیل هزینه فرصت (اجتماعی)<sup>۱</sup> و تحلیل مقرون به صرفه بودن<sup>۲</sup> است که هر دو برای مقایسه دو یا چند سیاست کاربرد دارند. در تحلیل مقرون به صرفه بودن، هزینه‌های سایر روش‌هایی که همین نتایج یا نتایج مشابه را می‌دهند، با یکدیگر مقایسه می‌شوند. این در حالی است که تحلیل هزینه-فرصت مرحله‌ای جلوتر می‌رود و اثر مداخلات مختلف بر رفاه اجتماعی را با در نظر گرفتن تمام هزینه‌ها و منافع مربوطه بر حسب پول (در صورت متناسب و امکان پذیر بودن) بررسی می‌کند. با این مقدمه، تعاریف دو روش مختلف ارزش در برابر پول پرداختی به این شرح خواهد بود:

#### ۱-۳-۴. تحلیل مقرون به صرفه بودن

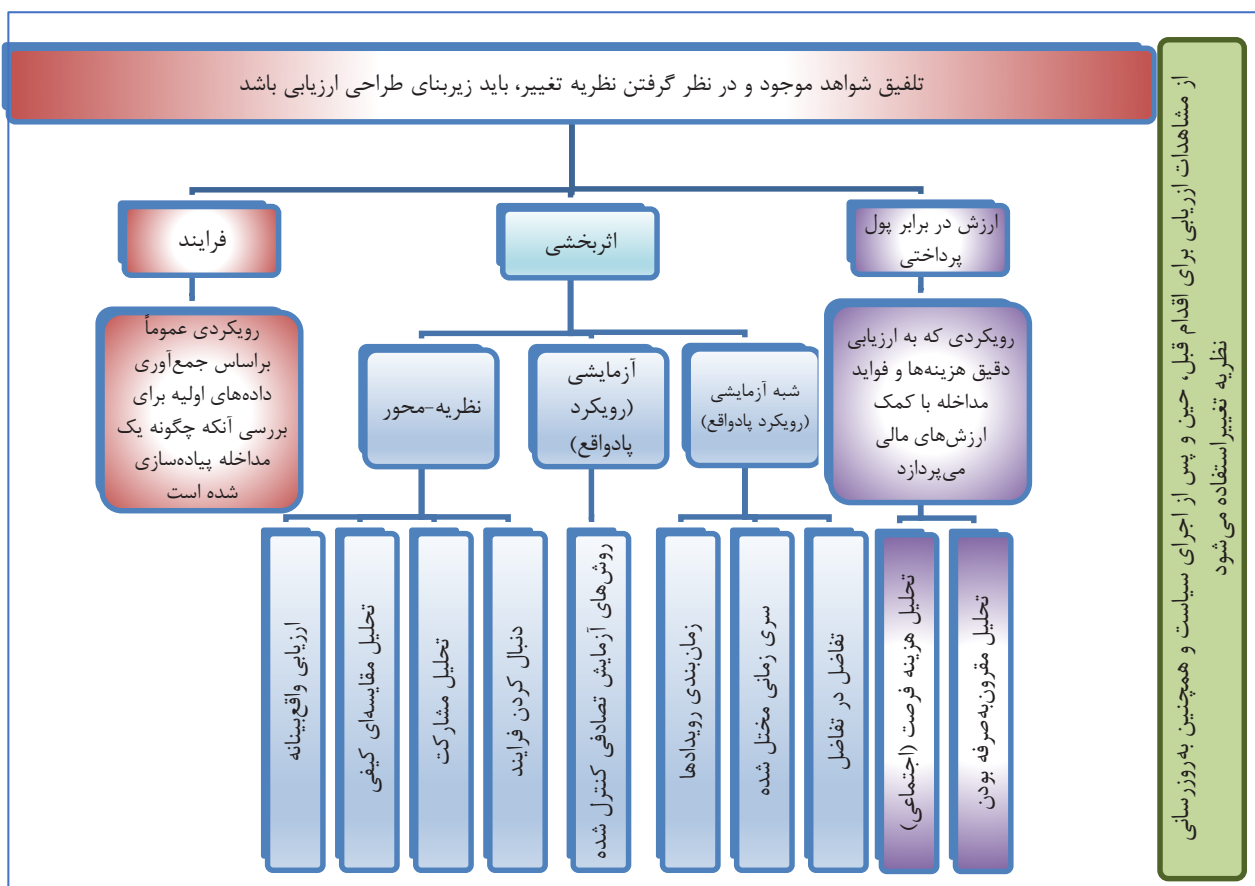
در این روش، به انتخاب بهینه‌ترین روش اجرای سیاست برای دستیابی به نتیجه‌ای مشخص پرداخته می‌شود. برای این کار، هزینه‌های اجرای روش‌های مختلف با یکدیگر مقایسه می‌شود. نقطه ضعف این روش آن است که گاهی منافع یا هزینه‌های یک سیاست قابلیت تبدیل به مقادیر کمی را ندارند و ارزش گذاری آنها دشوار است. مثلاً در تصمیم برای احداث خانه بهداشت، که آمار مرگ و میر و بیماری‌های یک منطقه را کاهش می‌دهد، نمی‌توان ارزش حفظ جان انسان‌ها را به راحتی به مقادیر کمی تبدیل کرد.

#### ۲-۳-۴. تحلیل هزینه - فرصت (اجتماعی)

برخلاف روش بالا، روش تحلیل هزینه - فرصت تمامی هزینه‌ها و منافع مداخله را کمی می‌کند. در این روش کلیه منافع و هزینه‌های انواع مداخلات ممکن یا حتی عدم مداخله مورد بررسی قرار می‌گیرد. به همین دلیل، امکان مقایسه بین گزینه‌های مختلف مداخله (حتی در حالتی که نتایج متفاوتی با یکدیگر داشته باشند) فراهم می‌شود. این روش دید کلی از شرایط مداخله (از جمله اثرهای زیست‌محیطی و اجتماعی) را فراهم می‌کند.

1. Social Cost-benefit Analysis  
2. Social Cost Effectiveness Analysis

شکل ۴. مراحل انجام ارزیابی موفق [۳]



## ۵. انواع داده‌ها و شیوه بهینه انتخاب آن

مرحله جمع‌آوری داده بخشی ضروری در ارزیابی هر سیاست حمایتی است که به برنامه‌ریزی دقیق و اصولی نیاز دارد؛ زیرا اگر طراحی آن به درستی انجام نشود، مداخله را به شدت محدود، گران‌قیمت و حتی ناممکن می‌کند. بر این اساس، انتخاب داده مناسب نقشی بسیار مهم در موفقیت ارزیابی خواهد داشت. در فرایند بهینه انتخاب داده، به چند سؤال کلیدی پاسخ داده می‌شود:

- چه نوع داده‌ای (عددی، کیفی و...) برای پاسخ به هر پرسش ارزیابی مورد نیاز است؟
- چه کسانی می‌توانند این داده‌ها را فراهم کنند؟
- آیا در دسترسی یا جمع‌آوری داده‌ها مشکلی وجود دارد؟
- داده‌ها از کدام بخش از جامعه هدف باید جمع‌آوری شوند؟
- روش تحلیل پیشنهادی چگونه بر داده‌های مورد نیاز تأثیر می‌گذارد؟

پاسخ به این پرسش‌ها می‌تواند راهگشای مناسبی برای انتخاب بهینه داده‌ها در ارزیابی یک سیاست حمایتی باشد. همچنین باید به کیفیت داده‌ها، رعایت حریم خصوصی افراد و عدم افشای اطلاعات شخصی نیز توجه شود. بر اساس آنچه گفته شد، داده‌های موجود در چهار دسته کلی گردآوری می‌شوند:



## ۱-۵. داده‌های ثبتي و نظارتی<sup>۱</sup>

جمع‌آوری این داده‌ها با هدف اجرای یک مداخله است. با وجود مشابهت‌ها، هریک از داده‌های ثبتي و نظارتی با اهداف متفاوتی جمع‌آوری می‌شوند که در ادامه تعریف شده و مورد بررسی قرار می‌گیرند.

### ۱-۱-۵. داده‌های ثبتي

این داده‌ها، برای اهداف مدیریتی (و نه پژوهشی) جمع‌آوری و عموماً به‌عنوان پایه نگارش آمار رسمی استفاده می‌شوند. داده‌های ثبتي، تراکنش‌های انجام شده و سوابق نگهداری شده که به‌طور معمول در حین ارائه یک خدمت ذخیره شده است، بخش قابل توجهی از این گروه داده را تشکیل می‌دهند.

از مزایای استفاده از این داده‌ها، می‌توان به کاهش هزینه‌های مالی ارزیابی و جلوگیری از ایجاد سختی برای پاسخ‌دهندگان اشاره کرد. با این حال، نباید از این نکته غافل شد که هدف اصلی از جمع‌آوری این داده‌ها، پاسخ به پرسش مورد نظر ما نبوده است. به همین دلیل، ممکن است که پاسخ داده به تمامی پرسش‌ها به اندازه کافی دقیق نباشد یا آن قدر منظم جمع‌آوری نشده باشد که بتواند شواهد مفیدی برای پرسش‌های ارزیابی ارائه کند.

### ۲-۱-۵. داده‌های نظارتی

این گروه داده، که به داده‌های مدیریت عملکرد<sup>۲</sup> نیز شناخته می‌شوند، در حین یک مداخله جمع‌آوری می‌شوند تا برای پرسش‌های سیاستی، تحقیقی و اجرایی پاسخ مناسبی ارائه دهند. این داده‌ها بیشتر اوقات کمی‌اند و تمامی جنبه‌های مداخله را پوشش می‌دهند. معمولاً از این داده‌ها برای پیگیری روند پیشرفت در دریافت سیاست (مداخله)، تشخیص قسمت‌هایی که مداخله طبق انتظار اجرا نشده است و اتخاذ تصمیماتی مناسب برای تحقق هدف سیاست استفاده می‌شود.

## ۲-۵. داده‌های پیمایشی<sup>۳</sup> در مقیاس بزرگ

این دسته‌بندی شامل داده‌های بلندمدت و در مقیاس بزرگ است که عموماً توسط دولت یا مراکز آمار محاسبه می‌شود و اهداف پژوهشی و آماری دارد. برای مثال، سیاستگذار با استفاده از این داده‌ها می‌تواند میزان شیوع فعالیت یا نگرشی را در جامعه بررسی کند و به تبع آن، مشکلات حاصل از این مسئله را دقیق‌تر تحلیل کند. از نمونه‌های این دسته داده‌ها در ایران، می‌توان به طرح آمارگیری نیروی کار اشاره کرد که با هدف شناخت ساختار و وضعیت جاری نیروی کار و تغییرات آن اجرا می‌شود.

با وجود مزایای قابل توجه، استفاده از این داده‌ها برای ارزیابی سیاست‌حمایتی با مشکلاتی نیز همراه است. هدف از جمع‌آوری داده‌های این بخش لزوماً با هدف ارزیابی هم‌راستا نیست. بنابراین ممکن است اطلاعات موجود در ارائه برخی جزئیات مورد نیاز ناتوان باشد (برای مثال، با وجود بزرگ بودن مقیاس داده‌ها، تعداد افرادی از نمونه که شرایط لازم برای ارزیابی مورد نظر را دارند محدود باشد). همچنین، ممکن است زمان‌بندی مشخص و ثابت داده‌ها نیز با هدف ارزیابی هم‌راستا نباشد. بنابراین، اگر در ارزیابی سیاست، فقط به جمع‌آوری این دسته از داده‌ها اکتفا شود، باید تمامی محدودیت‌های آن را در نظر گرفته باشیم.

از طرف دیگر - با توجه به مقیاس بزرگ این داده‌ها - نمونه‌گیری در ابعاد بزرگ باید انجام شود. به‌طور کلی، نمونه‌گیری از طریق دو روش کلی نمونه‌گیری احتمالی<sup>۴</sup> و نمونه‌گیری غیراحتمالی<sup>۵</sup> انجام می‌شود. در دسته اول، همان‌طور که از نامش مشخص است، افراد در نمونه‌گیری

1. Administrative and Monitoring Data  
2. Performance-management  
3. Survey Data  
4. Probability Sampling  
5. Non- probability Sampling

به‌طور تصادفی از جامعه مورد نظر (تمامی پاسخ‌دهندگان در جامعه هدف) انتخاب می‌شوند. اصلی‌ترین روش‌های استخراج نمونه انتخابی در این گروه شامل نمونه‌گیری تصادفی ساده<sup>۱</sup> (تمامی افراد احتمال انتخاب یکسان دارند)، نمونه‌گیری طبقه‌ای<sup>۲</sup> (گروه‌بندی جامعه و نمونه‌گیری احتمالی از هر گروه)، نمونه‌گیری خوشه‌ای<sup>۳</sup> (انتخاب خوشه‌ها به کمک احتمال و انتخاب افراد تنها از داخل خوشه‌های انتخابی) و نمونه‌گیری چندمرحله‌ای<sup>۴</sup> (ترکیب چند روش نمونه‌گیری) است. هدف تمامی این روش‌ها آن است که تخمین‌های نهایی بتوانند به‌خوبی جامعه را توصیف کنند.

دسته دوم (نمونه‌گیری غیراحتمالی) روشی راحت و کم‌هزینه است؛ اما باید با احتیاط استفاده شود. علت آن است که از نتایج این روش نمونه‌گیری نمی‌توان درباره جامعه هیچ استنتاجی کرد. یکی از معروف‌ترین رویکردها در این دسته‌بندی نمونه‌گیری سهمیه‌ای<sup>۵</sup> است. در این روش برای گروه‌های مختلف پاسخ‌دهنده، سهمیه در نظر گرفته می‌شود. در این حالت، کنترل روی ترکیب نمونه بیشتر خواهد بود. اگر این روش به‌درستی استفاده شود و سهمیه‌ها منطبق بر جامعه هدف وزن‌دهی شوند، می‌تواند جایگزین معتبری برای نمونه‌گیری تصادفی ساده باشد.

### ۳-۵. داده‌های جدید که به‌طور اختصاصی برای ارزیابی ایجاد شده‌اند

داده‌های این بخش به‌طور مشخص برای ارزیابی سیاست مورد نظر ایجاد می‌شوند. به همین دلیل، هم‌پوشانی کامل با ارزیابی مورد نظر دارند. برای ایجاد آن، می‌توان از روش‌هایی نظیر پرسش‌نامه، مصاحبه، مشاهدات جدید و... یاری گرفت. با وجود مزیت قابل توجهی که ذکر شد، نمی‌توان برای تمامی ارزیابی‌ها از داده‌های جدید استفاده کرد؛ زیرا ایجاد آن، نیازمند صرف زمان و هزینه قابل توجهی است. نمونه این نوع داده‌ها در ابعاد بسیار کوچک، پرسش‌نامه‌هایی است که به بررسی میزان اثرگذاری یک سیاست برای گروهی با هدف مشخص پیردازند.

### ۴-۵. داده‌های رسانه‌های اجتماعی

این داده‌ها اطلاعات جمع‌آوری شده از شبکه‌های اجتماعی است که دید افراد را به هر محتوا نشان می‌دهد. گسترش روش‌های اتوماتیک مثل استخراج از وبگاه<sup>۶</sup> به جمع‌آوری، پاک‌سازی و نگهداری و تحلیل سریع داده‌ها در ابعاد بزرگ کمک می‌کند. بنابراین، این داده‌ها به‌عنوان یکی از منابع نسبتاً جدید، بزرگ و دارای ظرفیت بالا برای اندازه‌گیری اثرهای یک مداخله قابل استفاده است.

## ۶. جمع‌بندی و نتیجه‌گیری

پایش و ارزیابی مستمر سیاست‌های حمایتی یکی از مؤلفه‌های اصلی نظام حمایتی کارآمد به حساب می‌آید و یکی از ارکان اصلی اهداف پیوست عدالت طرح‌ها و برنامه‌هاست. در یک دسته‌بندی کلی، ارزیابی سیاست شامل ارزیابی فرایند، ارزیابی اثربخشی و ارزیابی ارزش در برابر پول پرداختی می‌شود. ارزیابی فرایند تلاش می‌کند کلیه مراحل اجرای سیاست (از ورودی‌ها گرفته تا آخرین مرحله اجرا) را عارضه‌یابی کرده و گلوگاه‌ها و موانع هدف اولیه سیاست (به‌عنوان مثال ارائه خدمات به متقاضیان یا پرداخت نقدی به آنها) را بررسی و معرفی کند. ارزیابی اثربخشی نوع خاصی از ارزیابی‌هایی است که در پی پاسخ به پرسش رابطه علی است و آن اینکه اثر یک برنامه بر پیامد مورد نظر چیست؟ این

1. Simple Random Sampling  
2. Stratified Sampling  
3. Cluster Sampling  
4. Multi-stage Sampling  
5. Quota Sampling

۶. فرایند استفاده از ربات برای استخراج محتوا و داده از یک سایت را گویند.



نوع ارزیابی از ارزیابی فرایند گامی فراتر رفته و در پی بررسی این پرسش است که آیا با اجرای سیاست مورد نظر، هدف اصلی سیاست (به عنوان مثال کاهش فقر) محقق شده است یا خیر؟ ارزیابی ارزش در برابر پول پرداختی، هزینه‌ها و منافع مداخله‌ها را با یکدیگر مقایسه می‌کند. به بیان دیگر، این نوع ارزیابی در پی پاسخ به این پرسش است که عواید به دست آمده از اجرای سیاست مورد نظر چه تناسبی با هزینه‌های انجام شده دارد و آیا بهترین مداخله برای تحقق هدف مورد نظر بوده است؟ هر سه انواع ارزیابی کاربردها و مزایایی دارند؛ اما به طور مشخص ارزیابی اثربخشی در چند دهه گذشته بیشتر مورد توجه قرار گرفته و پژوهش‌های بسیاری را در اقتصاد و سایر رشته‌ها به خود اختصاص داده است. روش‌های ارزیابی اثربخشی نظریه-محور (ارزیابی واقع‌بینانه، تحلیل مقایسه‌ای کیفی، تحلیل مشارکت و دنبال کردن فرایند)، ارزیابی اثربخشی آزمایشی (روش‌های آزمایش تصادفی کنترل شده) و ارزیابی اثربخشی شبه‌آزمایشی (زمان‌بندی رویدادها، سری زمانی مختل شده و تفاضل در تفاضل) مهم‌ترین روش‌های ارزیابی اثربخشی است که مورد استفاده قرار می‌گیرد. همچنین نوع داده‌های موجود یا قابل دسترس در انتخاب نوع و روش ارزیابی نیز تأثیرگذار است.

جدول ۱. پیشنهاد توصیه سیاستی در رابطه با ارزیابی اثربخشی سیاست‌های حمایتی

ملاحظات	زمان‌بندی اجرا (کوتاه‌مدت، میان‌مدت، بلندمدت)	دستگاه معین	دستگاه متولی	الزامات و قیود اجرایی	توصیه سیاستی	نوع توصیه		ردیف
						اصلاح**	تداوم*	
	کوتاه‌مدت		وزارت تعاون، کار و رفاه اجتماعی		مشروط‌سازی تداوم اختصاص اعتبارات لازم برای اجرای سیاست‌های حمایتی به انجام ارزیابی اثربخشی سیاست‌ها	*		۱

\* تداوم یا تقویت آیتم‌ها یا اقدامات.  
\*\* اصلاح رویه‌ها یا ایجاد سازوکارها.  
مأخذ: یافته‌های پژوهش.

## منابع و مأخذ

- [۱] داده‌های طرح آمارگیری هزینه و درآمد خانوار، مرکز آمار ایران، ۱۴۰۲
- [2] World Bank (2022). Poverty and shared prosperity 2022: Correcting course. The World Bank.
- [3] Hurst, M. (2022). The Green Book: Central Government Guidance on Appraisal and Evaluation: HM Treasury, London, UK: OGL Press, 2022 (also available online on [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1063330/Green\\_Book\\_2022.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063330/Green_Book_2022.pdf)), ISBN 978-1-912225-57-6.
- [4] Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (Eds.) (2015). *Handbook of practical program evaluation* (p. 492). USA: John Wiley & Sons.
- [5] Britain, G. (2020). Magenta book: Central government guidance on evaluation. HM Treasury.
- [6] Mahler, D. G., Yonzan, N., & Lakner, C. (2022). The impact of COVID-19 on global inequality and poverty.



#### گزیده سیاستی

هرساله اعتبارات بسیاری برای پیشبرد طرح‌های حمایتی و محرومیت‌زدایی هزینه می‌شود؛ در حالی‌که معمولاً ارزیابی دقیقی از میزان اثربخشی آنها در دسترس نیست. در این گزارش سعی شده است کلان‌تصویری از انواع ارزیابی و روش‌های متداول هر یک از آنها ارائه شود.



مرکز پژوهش‌های مجلس شورای اسلامی

تهران، خیابان پاسداران، روبروی پارک نیاوران (ضلع جنوبی، پلاک ۸۰۲)

تلفن: ۷۵۱۸۲۰۰۰ صندوق پستی: ۱۵۸۷۵-۵۸۵۵ پست الکترونیک: [mrc@majles.ir](mailto:mrc@majles.ir)

وبسایت: [rc.majles.ir](http://rc.majles.ir)