Valid Inference on Functions of Causal Effects in the Absence of Microdata*

Vedant Vohra

January 22, 2025

Abstract

Economists are often interested in functions of multiple causal effects, a leading example of which is evaluating a policy's cost-effectiveness. The benefits and costs might be captured by multiple causal effects and aggregated into a scalar measure of cost-effectiveness. Oftentimes, the microdata underlying these estimates is inaccessible; only published estimates and their corresponding standard errors are available. We provide a method to conduct inference on non-linear functions of causal effects when the only information available is the point estimates and their standard errors. We apply our method to inference for the Marginal Value of Public Funds (MVPF) of government policies. (*JEL* C12, C21, H00)

*Vohra (vohra@ucsd.edu): UC San Diego. I am grateful to Julie Cullen, Graham Elliott, Itzik Fadlon, Stefan Faridani, Jacob Goldin, Xinwei Ma, Craig McIntosh, Harrison Mitchell, and Kaspar Wuthrich for helpful comments.

1 Introduction

There has been much innovation in econometrics to credibly estimate causal effects in a variety of settings and quantify the uncertainty in those estimates. Oftentimes, however, the researcher's eventual target of interest may not be a single causal effect, but rather a function of multiple causal effects. A leading example of this is determining a policy's cost-effectiveness: does a dollar of expenditure on the policy yield more than a dollar of benefits? For instance, a policy evaluation of a cash transfer program might consistently estimate its causal effect on a range of different outcomes and seek to determine whether a government should fund it. The benefits might include a treatment effect on health, education, and financial well-being; the costs might include potential labor market disincentives. These causal effects would then need to be aggregated into a scalar measure to obtain an estimate for the program's cost-effectiveness.

When the target parameter of interest is a scalar-valued, possibly non-linear, function of multiple causal effects, how can we conduct inference on this parameter? If we had access to the microdata underlying each causal effect, we could easily estimate the correlation across each causal effect and therefore estimate the variance of the function (e.g., Zellner, 1962). However, in many settings of interest, the microdata underlying the causal effects might be unavailable for ex-post analysis, making it infeasible to estimate the correlation structure on which the variance of the function depends. In this paper, we study the problem of inference on functions of multiple causal effects, without knowledge about the correlation structure across these causal effects.

To help further motivate the question, we focus on the problem of conducting inference on the Marginal Value of Public Funds (MVPF) for a given policy (Hendren and Sprung-Keyser, 2020). The MVPF is a popular metric for evaluating the welfare consequences of government policies. It is calculated as a non-linear function of multiple causal effects: the benefits that a policy provides to its recipients are divided by the policy's net cost to the government. Hendren and Sprung-Keyser (2020) use causal estimates available in existing studies to construct the MVPF for over a hundred policies. The only information available to estimate the MVPF and its variance are the causal effects and their corresponding standard errors reported in the original study. The microdata underlying these causal estimates is inaccessible to Hendren and Sprung-Keyser (2020) for ex-post analysis. The fundamental challenge for inference is that the variance of the MVPF depends not only on the standard errors of the causal effects but also on the correlation across the causal effects.

Estimates of the correlation structure across causal effects might be difficult to obtain for several reasons. First, one might be interested in a function of causal effects available in an existing publication, but the underlying microdata might be inaccessible. This could either be because the causal effects are estimated using privately held administrative microdata or the underlying replication data were not made publicly available. For instance, replication data are not publicly available for nearly half of all empirical papers published in the *American Economic Review* in recent years (Christensen and Miguel, 2018). Second, even if the microdata were available, it might be prohibitively costly to compute the correlation across causal effects when the causal effects rely on different data sources that have common units but are challenging to merge together. This might happen in cases where a unique identifier is missing to exactly merge the two data sources (e.g., in historical decennial Census data (Ruggles, Fitch, and Roberts, 2018)) or when the two data sources are housed at different federal agencies (e.g., administrative tax data and administrative crime records for the full population (Rose, 2018)).

We provide a simple inference procedure that yields valid confidence intervals on functions of causal effects without knowledge about the correlation across these causal effects. First, we ask what is the largest possible variance of the function given the information we observe, and what is the correlation structure under which this upper bound is attained? The implied confidence intervals using the upper bound of the variance are conservative but have close to exact coverage when the true correlation structure is close to the correlation structure under which this upper bound is attained. We show that using this upper bound alone, meaningful inference is possible in settings of interest to applied researchers where estimating the correlation across causal effects is not possible.

Second, we show how inference can be sharpened further when the causal effects being considered correspond to the effects of a randomized treatment on a range of different outcomes. We show that, in this setting, the correlation across the causal effects takes a particularly interpretable form, the sign of which might be known from prior studies, economic theory, or other data sources. Incorporating this information allows us to provide meaningfully tighter bounds on the variance. We cast the problem of finding the upper bound of the variance as an optimization problem, allowing us to flexibly incorporate additional setting-specific information to sharpen the upper bound on the variance, such as the bounded support of an outcome or known independence of two causal effects.

Finally, we re-cast the inference problem as a "breakdown" problem: instead of asking what is the largest possible variance of the function given the available information, we ask how large the variance can be before a policy-relevant conclusion *breaks down*. In the case of MVPF, one policy-relevant null hypothesis is whether a dollar spent on the policy provides the beneficiaries with less than one dollar of benefits, i.e., $H_0 : MVPF < 1$. Our proposed method asks, how plausible is the largest variance under which one can reject this null hypothesis. We provide an easily interpretable breakdown metric that takes a value between 0 and 1, where 0 implies that one can reject the null hypothesis under any correlation structure and 1 implies that one can't reject the null hypothesis under any correlation structure. The breakdown approach has the advantage of being comparable across policies: since the true correlation structure is unknown, it is unclear how close the upper bound of the variance might be to the true variance across different policies.

We illustrate our inference procedure by conducting inference on the MVPF for 8 different policies. We show that meaningful inference is possible in the absence of *any* microdata, using the upper bound of the variance alone. Hendren and Sprung-Keyser (2020) note that since the MVPF reflects the shadow price of redistribution, a welfare-maximizing government should have a positive willingness-to-pay to reduce the statistical uncertainty in the cost of redistribution. We illustrate how statistical uncertainty can be reduced by using setting-specific information about the sign of the correlation across outcomes. In fact, our novel characterization of the covariance matrix in randomized trials allows us to reduce the width of the MVPF confidence intervals beyond the worst case by up to 30% in the policies we consider. Finally, we compute the breakdown metric for the MPVF of multiple policies and illustrate how this metric can be useful to a policymaker choosing from a menu of policies.

Our work contributes to the literature on welfare analyses of government expenditure (e.g., Chetty, 2009; Heckman, Moon, Pinto, Savelyev, and Yavitz, 2010; Hendren and Sprung-Keyser, 2020). While the existing tools provide a unified framework to evaluate the welfare consequences of government policies, the statistical tools to conduct inference on welfare metrics under frequently encountered data limitations have been absent from the literature. Our inference tools strengthen the MVPF framework by allowing us to evaluate the uncertainty in welfare metrics. Hendren and Sprung-Keyser (2020) show that increasing spending on Policy A is welfare-improving by reducing spending on Policy B if and only if the MVPF of Policy A is greater than the MVPF of Policy B. We provide a valid test for the policy-relevant null hypothesis, $H_0 : MVPF_A < MVPF_B$.

Our work also relates to the growing literature of inference on causal effects under limited data availability (e.g., d'Haultfoeuille, Gaillac, and Maurel, 2022; Fan, Shi, and Tao, 2023). In contrast to the existing literature that operates in a setting where *some* microdata is available, our setting requires that we only have access to the causal effects and their standard errors but *none* of the underlying microdata.

Finally, Cocci and Plagborg-Møller (2024) adopt a similar approach to ours in a different setting: bounding the variance in the setting of calibrating parameters of a structural model to match empirical moments. In contrast to our setting where no additional information is available, Cocci and Plagborg-Møller (2024) focus on using moment selection tools to tighten the upper bound on the variance whereas we rely on an explicit characterization of the covariance structure to sharpen the upper bound.

2 Setting

Our starting point is a vector of estimated causal effects denoted by $\widehat{\beta} \in \mathbb{R}^d$. In our leading application, $\widehat{\beta}$ represents a vector of causal effects that are then aggregated to evaluate the cost-effectiveness of a policy. We assume that the estimated causal effects follow a joint Normal distribution with an asymptotic variance-covariance matrix $\mathbf{V}$. Since it is convention to report standard errors corresponding to the estimated causal effects, we assume we have access to consistent estimates for the diagonal entries of $\mathbf{V}$. In contrast, the off-diagonal entries of $\mathbf{V}$ are rarely reported and are not estimable in our setting since the microdata underlying the causal effects is unavailable. We summarize this in Assumption 1.

Assumption 1. *In an existing study, we observe a vector of estimates* $\widehat{\beta} = \left(\widehat{\beta}_1, \ldots, \widehat{\beta}_d\right)$ *for causal effects* $\beta = \left(\beta_1, \ldots, \beta_d\right)$ *along with their corresponding standard errors* $\widehat{\sigma} = \left(\widehat{\sigma}_1, \ldots, \widehat{\sigma}_d\right)$ *such that*

(i) $\widehat{\beta} \xrightarrow{p} \beta$

(ii) $\widehat{\beta}$ *follows a jointly normal distribution asymptotically,* $\sqrt{n}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} \mathcal{N}\left(0, \mathbf{V}\right)$*, where* $\mathbf{V}$ *is a* $d \times d$ *positive semi-definite variance-covariance matrix*

(iii) $\widehat{\sigma}$ *is a consistent estimator for the diagonal entries of* $\mathbf{V}$, $\left(\widehat{\sigma}_1^2, \ldots, \widehat{\sigma}_d^2\right) \xrightarrow{p} \left(\sigma_1^2, \ldots, \sigma_d^2\right)$.

Next, we assume that the causal effects are aggregated into a scalar estimate using a possibly non-linear function $f : \mathbb{R}^d \to \mathbb{R}$. This function is assumed to be continuously differentiable at β and $f'(\beta)$ is assumed to not be zero-valued. In our application, $f(\widehat{\beta})$ corresponds to an estimate of a policy's cost-effectiveness. We place no additional restrictions on the function $f(.)$, since in practice, this function would depend on the economic mapping between the estimated causal effects and the cost-effectiveness of the specific policy being analyzed. This is summarized in Assumption 2. We maintain Assumptions 1 and 2 throughout the paper.

Assumption 2. *The function* $f : \mathbb{R}^d \to \mathbb{R}$ *is continuously differentiable at* β *and* $f'(\beta)$ *is not zero-valued.*

Under Assumptions 1 and 2, we may use the delta-method to find the asymptotic distribution for $f(\widehat{\beta})$,

$$\sqrt{n}\Big(f(\widehat{\beta}) - f(\beta)\Big) \xrightarrow{d} \mathcal{N}(0, \tau^2)$$

where

$$\begin{aligned}
\tau^2 &= \sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_{ij} \frac{\partial f(\beta)}{\partial \beta_i}\frac{\partial f(\beta)}{\partial \beta_j} \\
&= \sum_{i=1}^{d}\left(\sigma_i \frac{\partial f(\beta)}{\partial \beta_i}\right)^2 + \sum_{\substack{i=1 \\ \{i,j:i\neq j\}}}^{d}\sum_{j=1}^{d} \sigma_{ij}\frac{\partial f(\beta)}{\partial \beta_i}\frac{\partial f(\beta)}{\partial \beta_j} \qquad (2.1)
\end{aligned}$$

and σ_{ij} is the covariance between β_i and β_j. Define $\rho_{ij} \equiv \frac{\sigma_{ij}}{\sigma_i \sigma_j}$ as the correlation coefficient between β_i and β_j.

The objective of this paper is to learn about τ^2. In our leading application where $f(\widehat{\beta})$ is the estimated cost-effectiveness of a policy, a policymaker may also be interested in the variance of the estimated cost-effectiveness: a risk-averse policymaker choosing to implement one of two equally cost-effective policies would choose to implement the policy that has lower statistical uncertainty in its estimated cost-effectiveness. The central challenge to estimating τ^2 is that it depends on σ_{ij} and in our setting, σ_{ij} is not estimable due to the data limitations discussed in Section 1. This raises the question, what can be learned about τ^2 when it is not possible to estimate σ_{ij} for $i \neq j$?

3 Inference Procedure

Since the asymptotic variance of $f(\widehat{\beta})$ depends on the correlation structure across causal effects and this correlation structure is not estimable, we consider an alternative approach to conducting inferences about $f(\beta)$. Specifically, we ask how large can the variance for $f(\widehat{\beta})$ be given the observed information and use this variance upper bound to test hypotheses about

$f(\beta)$.

The motivation for using an upper bound for the variance of $f(\widehat{\beta})$ is that any test relying on the upper bound guarantees size control, but trades it off against a loss of power. To formalize this rationale, consider the following hypothesis test:

$$H_0 : f(\beta) \leq k \quad \text{against} \quad H_1 : f(\beta) > k. \tag{3.1}$$

The optimality properties of testing H_0 with a test statistic using a consistent estimator of τ^2 are well-understood: in a limit experiment, the t-test controls size and is the Uniformly Most Powerful (UMP) test. When an estimate for τ^2 is not available, we may instead think of the above as a hypothesis test in the presence of nuisance parameters ρ_{ij} for all $i \neq j$. The problem of finding the UMP test in the presence of nuisance parameters corresponds to finding the least favorable distribution on the support of the nuisance parameters (Theorem 3.8.1 in Romano and Lehmann, 2005). A distribution is said to be least favorable if the power of the test under this distribution on the nuisance parameters is lower than the power of the test under any other distribution. While it is often prohibitively challenging to find the least favorable distribution (e.g., see Elliott, Müller, and Watson (2015)), this problem is greatly simplified in our setting with the following observation: the nuisance parameters only enter into the distribution of $f(\beta)$ through its variance, and power is minimized when variance is maximized. This provides a statistical rationale for inspecting the upper bound on the variance.

We can re-write Equation 2.1 as follows:

$$\tau^2 = \sum_{i=1}^{d} \left(\sigma_i \frac{\partial f(\beta)}{\partial \beta_i} \right)^2 + \sum_{\substack{i=1 \\ \{i,j:i \neq j\}}}^{d} \sum_{j=1}^{d} \rho_{ij} \sigma_i \sigma_j \frac{\partial f(\beta)}{\partial \beta_i} \frac{\partial f(\beta)}{\partial \beta_j} \tag{3.2}$$

Since β_i and σ_i can be consistently estimated given the observed data, obtaining an upper bound for τ^2 amounts to maximizing Equation 3.2 with respect to ρ_{ij} for $i \neq j$. We analyze the problem of obtaining an upper bound for the variance τ^2 separately for two cases. In

Section 3.1, we focus on the general case where $\widehat{\beta}$ is not necessarily an estimate for a causal parameter. In Section 3.2, we specialize to the case where $\widehat{\beta}$ corresponds to an estimate for a specific causal parameter and illustrate how we can sharpen inference in this setting.

3.1 Aggregating Regression Estimates

In the setting where no additional information is known about ρ_{ij}, it is easy to see from Equation 3.2 how one might maximize τ^2. If $\frac{\partial f(\beta)}{\partial \beta_i}\frac{\partial f(\beta)}{\partial \beta_j} > 0$ for some $i, j \in \{1, ..., d\}$, τ^2 is maximized when $\rho_{ij} = 1$. Similarly, if $\frac{\partial f(\beta)}{\partial \beta_i}\frac{\partial f(\beta)}{\partial \beta_j} < 0$ for some $i, j \in \{1, ..., d\}$, τ^2 is maximized when $\rho_{ij} = -1$. We can equivalently cast the problem of obtaining an upper bound on the variance τ^2 as the following optimization problem:

$$\begin{aligned}
&\underset{\{\rho_{ij}\}_{i,j=1}^{d}}{\text{Maximize}} && \tau^2 && && \textbf{SDP.1} \\
&\text{subject to} && \mathbf{V} \succeq 0 && && \text{(C.1)} \\
& && \rho_{ij} = \rho_{ji} && \forall\ i, j = 1, ..., d && \text{(C.2)} \\
& && \rho_{ij} \in [-1, 1] && \forall\ i, j = 1, ..., d && \text{(C.3)} \\
& && \rho_{ii} = 1 && \forall\ i = 1, ..., d && \text{(C.4)}
\end{aligned}$$

C.1 is a matrix inequality that states that the variance-covariance matrix must be positive semi-definite; C.2 requires the resulting variance-covariance matrix to be symmetric; and C.3 ensures that the correlation is bounded between -1 and 1. **SDP.1** is a well-defined semi-definite program (SDP) that can be solved using existing semi-definite programming tools (Grant and Boyd, 2008, 2014). The advantage of this approach is that we can flexibly add any known information about the correlation across estimates as constraints to the optimization problem, such as independence or the sign of correlation across two estimates.

We denote the maximum variance resulting from **SDP.1** as $\tau^2_{\max}$ and the correlation matrix that attains this upper bound as $\boldsymbol{\rho}_{\max}$. Then,

$$\begin{aligned}\tau_{\max}^2 &= \sum_{i=1}^{d}\left(\sigma_i \frac{\partial f(\beta)}{\partial \beta_i}\right)^2 + \sum_{\substack{i=1 \\ \{i,j:i\neq j\}}}^{d}\sum_{j=1}^{d} \sigma_i\sigma_j \left|\frac{\partial f(\beta)}{\partial \beta_i}\frac{\partial f(\beta)}{\partial \beta_j}\right| \\ &= \left(\sum_{i=1}^{d} \sigma_i \left|\frac{\partial f(\beta)}{\partial \beta_i}\frac{\partial f(\beta)}{\partial \beta_j}\right|\right)^2 \end{aligned} \tag{3.2}$$

Some implications arise from inspecting the worst-case variance. First, the confidence intervals constructed using $\tau_{\max}$ will have a higher coverage probability than when using τ by construction. While this will result in a loss of power, it guarantees size control; the coverage will only be exact when $\tau_{\max} = \tau$. Second, in settings where it is feasible but costly to estimate the covariance across estimates, we recommend that the researcher begin by testing their hypothesis of interest using the upper bound of the variance we provide. Rejecting a null hypothesis using the variance upper bound implies that the hypothesis would also be rejected using the true variance. This allows the researcher to test the hypothesis of interest while sidestepping the costs of computing the covariances we highlighted previously. Finally, the expression for $\tau_{\max}$ in Equation 3.2 aligns with Lemma 1 in Cocci and Plagborg-Møller (2024) who focus on minimizing the worst-case variance using moment selection tools when matching structural parameters to empirical moments in the over-identified case.

3.2 Aggregating Causal Effects from a Randomized Trial

In this sub-section, we specialize to the case where the causal effects being considered correspond to the effect of a randomized treatment on a range of different outcomes. As an example, consider the setting where a policymaker studies the effect of a pilot cash transfer program on a range of different outcomes through a randomized trial and then seeks to evaluate the program's cost-effectiveness. We show that in this setting, the off-diagonal entries of $\mathbf{V}$ take a particularly interpretable form.

Define $Y_{ij}(1)$ as the treated potential outcome j for unit i and $Y_{ij}(0)$ as the control potential outcome j for unit i, where $j \in \{1, ..., d\}$. Let $Z_i \in \{0, 1\}$ denote whether unit i is assigned to the treatment or control group. We assume that treatment assignment is independent of the vector of potential outcomes, i.e. $\Big(Y_{ij}(1), Y_{ij}(0)\Big) \perp Z_i$ for all $j = 1, \ldots, d$. The effect of the treatment Z_i is evaluated using d linear regressions of the form, $Y_{ij} = \alpha_j + \beta_j Z_i + \varepsilon_{ij}$. Let the vector of Average Treatment Effects (ATEs) be given by $\beta \in \mathbb{R}^d$ and our estimates be given by $\widehat{\beta} \in \mathbb{R}^d$. In this setting, the following proposition provides a simple characterization of the off-diagonal entries of $\mathbf{V}$.

Proposition 1. *The covariance between the ATE of a randomized binary treatment Z_i on any two outcomes Y_p and Y_q can be expressed as follows:*

$$\mathbb{C}\text{ov}\Big(\beta_p, \beta_q\Big) = \frac{\mathbb{C}\text{ov}(Y_{ip}, Y_{iq} \mid Z_i = 1)}{\mathbb{P}[Z_i = 1]} + \frac{\mathbb{C}\text{ov}(Y_{ip}, Y_{iq} \mid Z_i = 0)}{\mathbb{P}[Z_i = 0]}$$

The proof for Proposition 1 is provided in Appendix Section A. Proposition 1 establishes that covariance of the two ATEs β_p and β_q takes the following interpretable form: if the covariance of the two outcomes is positive in the treatment group *and* the control group, the covariance of the causal effects will also be positive. Intuitively, if the outcomes are positively correlated in treatment and control group, the effect of the treatment must also move in the same direction. Since the direction (and in some cases, the magnitude) of the covariance in the outcomes might be known from economic theory, alternative data sources, or prior evidence, this information can easily be added as constraints to **SDP.1** to find a (weakly) smaller upper bound on the variance. For instance, if *all* off-diagonal entries are known to be non-negative, we can solve the following optimization problem:

$$\underset{\{\rho_{ij}\}_{i,j=1}^{d}}{\text{Maximize}} \quad \tau^2 \qquad \textbf{SDP.2}$$

$$\text{subject to} \quad \mathbf{V} \succeq 0 \tag{C.1}$$

$$\rho_{ij} = \rho_{ji} \qquad \forall\, i, j = 1, ..., d \tag{C.2}$$

$$\rho_{ij} \in [-1, 1] \qquad \forall\, i, j = 1, ..., d \tag{C.3}$$

$$\rho_{ii} = 1 \qquad \forall\, i = 1, ..., d \tag{C.4}$$

$$\rho_{ij} \geq 0 \qquad \forall\, i, j = 1, ..., d \tag{C.5}$$

In Section 5, we show how incorporating sign constraints in real-world applications of interest can yield meaningfully tighter bounds on the variance in some cases.

In some settings, β_i and β_j will identify the ATE of a treatment on outcomes Y_i and Y_j with bounded support. When the outcome has bounded support, the support of the ATE will also bounded. In this case, one might be able to tighten the bounds on the variance further by incorporating this information into the optimization problem **SDP.2**. Hössjer and Sjölander (2022) derive the following bounds for the covariance of two bounded random variables.

Remark 1. *Suppose that outcome* $Y_i \in \left[\underline{Y_i}, \overline{Y_i}\right]$ *and* $Y_j \in \left[\underline{Y_j}, \overline{Y_j}\right]$. *Define* $\underline{\beta_i} = \underline{Y_i} - \overline{Y_i}$ *and* $\overline{\beta_i} = \overline{Y_i} - \underline{Y_i}$. *Define* $\underline{\beta_j}$ *and* $\overline{\beta_j}$ *analogously. Then, the ATEs* β_i *and* β_j *will be bounded:* $\beta_i \in \left[\underline{\beta_i}, \overline{\beta_i}\right]$ *and* $\beta_j \in \left[\underline{\beta_j}, \overline{\beta_j}\right]$. *Then, the covariance of* β_i *and* β_j, σ_{ij} *satisfies:*

$$\begin{aligned}
&-\min\left[\left(\beta_i - \underline{\beta_i}\right)\left(\beta_j - \underline{\beta_j}\right), \left(\overline{\beta_i} - \beta_i\right)\left(\overline{\beta_j} - \beta_j\right)\right] \\
&\leq \sigma_{ij} \\
&\leq \min\left[\left(\beta_i - \underline{\beta_i}\right)\left(\overline{\beta_j} - \beta_j\right), \left(\overline{\beta_i} - \beta_i\right)\left(\beta_j - \underline{\beta_j}\right)\right].
\end{aligned}$$

Finally, in some cases, it might be known that two estimates are uncorrelated if, for instance, the estimates are constructed using independent, non-overlapping samples. This information can incorporated in the optimization problem by fixing that correlation to be 0.

4 Breakdown Analysis

In Section 3, we provided a toolkit to quantify uncertainty in the estimated cost-effectiveness of a policy. In this section, we offer a method to quantify how "robustly" a policymaker can conclude whether a policy being considered is cost-effective. Specifically, instead of asking what is the largest possible variance of $f(\widehat{\beta})$ given the available information, we assess the plausibility of the correlation structure that yields the largest variance for $f(\widehat{\beta})$ without a conclusion of interest breaking down. This approach of assessing the plausibility of the point at which a conclusion of interest "breaks down" is referred to as breakdown analysis.[1] In the case of the MVPF, a policy-maker might be interested in testing whether a dollar spent on the policy provides the beneficiaries with more than one dollar of benefits, i.e., $H_0 : MVPF < 1$. We provide a Breakdown Statistic to assess how plausible is the worst-case correlation structure under which the policymaker can still reject this null hypothesis.

The Breakdown Statistic we provide is bounded between 0 and 1 and is easily interpretable. Intuitively, it captures how far we need to move from the worst-case correlation structure to be able to reject a hypothesis of interest: a lower Breakdown Statistic suggests that the rejection of a null hypothesis is more robust to different correlation structures. A Breakdown Statistic of 0 implies that we can reject the null hypothesis of interest under any correlation structure and a Breakdown Statistic of 1 implies that there is no feasible correlation structure under which we can reject the null hypothesis of interest. A Breakdown Statistic of 0.5 implies that the policy conclusion is valid under independence of all estimates.

Let z_α denote the $(1-\alpha)$ quantile of Normal distribution. $\|.\|_F$ denotes the Frobenius norm of a matrix. Suppose that we are interested in testing the hypothesis, $H_0 : f(\beta) < k$ vs. $H_1 : f(\beta) \geq k$. We compute the Breakdown Statistic as follows.

[1]See Manski and Pepper (2018); Masten and Poirier (2020); Diegert, Masten, and Poirier (2022); Rambachan and Roth (2023); Spini (2024) for recent examples.

Step 1: If $f(\widehat{\beta}) - z_\alpha \times \tau_{\max} \geq k$ where $\tau_{\max}$ is defined in Equation 3.2, we can reject the null hypothesis $H_0 : f(\beta) < k$ under even the worst case correlation structure. Since this implies that we can reject H_0 under any feasible correlation structure, we set the Breakdown Statistic equal to 0.

Step 2: If $f(\widehat{\beta}) - z_\alpha \times \tau_{\max} < k$, find the correlation structure that maximizes τ^2 *and* under which the $(1-\alpha)\%$ CI excludes all values less than k. We refer to this correlation matrix as $\boldsymbol{\rho}_B$, and find it by solving the following optimization problem.

$$
\begin{aligned}
&\underset{\{\rho_{ij}\}_{i,j=1}^{d}}{\text{Minimize}} && \|\boldsymbol{\rho}\|_F && && \textbf{SDP.3} \\
&\text{subject to} && \mathbf{V} \succeq 0 && && \text{(C.1)} \\
& && \rho_{ij} = \rho_{ji} && \forall\, i,j = 1,...,d && \text{(C.2)} \\
& && \rho_{ij} \in [-1,1] && \forall\, i,j = 1,...,d && \text{(C.3)} \\
& && \rho_{ii} = 1 && \forall\, i = 1,...,d && \text{(C.4)} \\
& && f(\widehat{\beta}) - z_\alpha \times \tau = k && && \text{(C.5)}
\end{aligned}
$$

SDP.3 finds the correlation structure with the lowest Frobenius norm that allows us to just reject H_0. Since the objective function is convex in the entries of the correlation matrix, there exists a unique $\boldsymbol{\rho}_B$ that minimizes **SDP.3**.

Step 3: If there is no feasible solution to **SDP.3**, set the Breakdown Statistic equal to 1. If there is a feasible solution, we calculate the following statistic:

$$\text{Breakdown Statistic} = \frac{\|\boldsymbol{\rho}_B - \boldsymbol{\rho}_{\max}\|_F}{2\sqrt{d(d-1)}}$$

In Section 5, we describe how the Breakdown Statistic can help a policy-maker choose from a menu of policies to implement.

5 Application

We illustrate our method by conducting inference on a metric that assesses the welfare implications of increasing expenditure on a range of government policies. Specifically, we apply the tools developed in Sections 3 and 4 to construct valid confidence intervals for the Marginal Value of Public Funds (MVPF) associated with a menu of government policies. We begin by outlining the MVPF framework for welfare analysis and highlight why our approach is particularly well-suited for valid inference on MVPFs. Then, we use our toolkit to quantify the uncertainty in the estimated MVPF for eight public policies.

Hendren and Sprung-Keyser (2020) popularized the MVPF as a unified metric for evaluating the welfare consequences of government expenditure. The MVPF quantifies the "bang-for-the-buck" of public expenditure: for instance, an MVPF of 1 indicates that a policy generates \$1 in benefits for every dollar of net government expenditure. Formally, the MVPF for a policy is defined as the ratio of the benefits provided to recipients of that policy to the net cost borne by the government:

$$MVPF = \frac{Benefits}{Net\ Government\ Costs} = \frac{\Delta W}{\Delta E - \Delta C},$$

where ΔW represents the benefits delivered to individuals, ΔE is the government's initial expenditure on the policy, and ΔC captures the reduction in government costs attributable to the policy's causal effects.

Four key features of the MVPF framework make our proposed method particularly well-suited for valid inference. First, the MVPF for a given policy is defined as a non-linear function of multiple causal effects. For example, consider the MVPF of an expanded Earning Income Tax Credit (EITC) program, *Paycheck Plus*. Miller, Katz, Azurdia, Isen, and Schultz (2017) estimate the causal effect of the tax credit on earnings, employment, and after-tax income. These causal effects, summarized in Table 1, are used to construct the MVPF:

$$\widehat{\beta} = \begin{bmatrix} \widehat{\beta}_1 & \widehat{\beta}_2 & \widehat{\beta}_3 & \widehat{\beta}_4 & \widehat{\beta}_5 & \widehat{\beta}_6 \end{bmatrix}' \\ = \begin{bmatrix} 0.009 & 0.025 & 654 & 33 & 645 & 192 \end{bmatrix}'$$

and

$$MVPF_{\text{Paycheck Plus}} \equiv f(\widehat{\beta}) = \frac{1399 \times (45 - \widehat{\beta}_1) + 1364 \times (34.8 - \widehat{\beta}_2)}{(\widehat{\beta}_3 - \widehat{\beta}_4) + (\widehat{\beta}_5 - \widehat{\beta}_6)} \\ = 0.996.$$

Second, in most cases, the only information available to construct the MVPF are the causal effects and their corresponding standard errors. For instance, the causal effects of Paycheck Plus are estimated using administrative tax data which is not publicly accessible and the correlation structure across causal effects is not reported in the original study. The only information available to estimate the MVPF of Paycheck Plus and its variance is provided in Table 1. To conduct inference on MVPF estimates, Hendren and Sprung-Keyser (2020) crucially assume a correlation structure across estimates. However, this might lead to confidence intervals that don't guarantee the appropriate coverage probability, as we discuss in Appendix Section B. Our method delivers valid inference using only the reported standard errors and importantly does not require the full variance-covariance matrix to be assumed or consistently estimated.

Third, Hendren and Sprung-Keyser (2020) show that increasing spending on Policy A is welfare-improving by reducing spending on Policy B if and only if the MVPF of Policy A is greater than the MVPF of Policy B. This presents a natural null hypothesis central to welfare analysis and policy choice, $H_0 : MVPF_A < MVPF_B$.[2] Our method provides a test for this null hypothesis that controls size under any correlation structure.

Fourth, since the MVPF reflects the shadow price of redistribution, a welfare-maximizing

[2]Here, we assume that the beneficiaries of both policies receive equal welfare weights.

government should be willing to pay to reduce the statistical uncertainty in the estimated cost of redistribution (Hendren and Sprung-Keyser, 2020). In Section 3.2, we show how mild setting-specific assumptions can be used to sharpen inference on the MVPF in many settings, establishing a systematic approach to reducing statistical uncertainty in the MVPF estimates.

Next, we apply our inference method to the MVPF of 8 government policies from different domains of public expenditure: three job-training programs (Job Start, Work Advance, Year Up), two cash transfers (Paycheck Plus, Alaska Universal Basic Income), a health insurance expansion (Medicare Part D), childcare expenditure (foster care provision), and an Unemployment Insurance (UI) expansion. The MVPF and the 95% confidence intervals constructed using **SDP.1** are shown in Figure 1. We defer the details of each policy and their corresponding MVPF calculations to Appendix Section C. Some key lessons emerge from examining Figure 1. First, absent any assumptions on the off-diagonal entries of variance-covariance matrix, we are able to reject the null hypothesis that MVPF of two job training programs (Job Start and Year Up) is greater than 1 under any correlation structure across causal effects. Specifically, we can conclude that $1 spent on a job training program provides its recipients with less than $1 of benefits. Second, we can use the estimated variance upper bound to test the null hypothesis, $H_0 : MVPF_{\text{Alaska UBI}} < MVPF_{\text{Job Start}}$.[3] We are able to reject this null hypothesis and conclude that increasing expenditure on this UBI policy by decreasing expenditure on this job-training program would increase welfare. Finally, our estimates suggest that there might meaningful statistical uncertainty in the relative ranking of policies by their MVPF. Absent estimates of the variance upper bound that we provide, a policymaker might erroneously conclude that increasing funding for UI extensions by decreasing funding for job training programs will be welfare improving. However, our results suggest that the statistical uncertainty in the estimates preclude this conclusion from the available information.

[3]Since the MVPFs of the two policies are constructed using independent non-overlapping samples, we assume that the two MVPFs are uncorrelated.

The only policy for which we have access to microdata underlying the causal effects is Medicaid Part D. Using this microdata, we are able to recover the full variance-covariance matrix and compare the exact confidence intervals, which are infeasible to compute for all other policies we consider. We compare three confidence intervals for the estimated MVPF for Medicaid Part D in Table 2: exact confidence intervals using the estimated correlated structure, confidence intervals implied by an assumption that all causal effects are uncorrelated, and the worst case confidence intervals implied by **SDP.1**. The exact confidence intervals rule out values of the MVPF smaller than 0.80 and larger than 1.95, while the worst-case confidence intervals allow us to rule out values less than 0.17 and larger than 2.57.[4] This suggests that while one can conduct meaningful inference on the MVPF for a given policy using the confidence intervals implied by **SDP.1**, a key takeaway for practitioners is that it may be valuable to report the estimated covariance matrix across causal effects whenever possible.

A policy-maker choosing from a menu of policies might care about whether we can robustly conclude that a policy "pays for itself" rather than focusing on the statistical uncertainty in the estimated returns from increasing expenditure on each policy. In order to answer the question of how robustly we can conclude that a policy pays for itself, we ask how far one would need to move away from the worst case correlation structure to be able to reject the null hypothesis that $H_0 : MVPF < 1$. The Breakdown Statistic, introduced in Section 4, serves as a concise summary measure for this purpose: the higher the Breakdown Statistic, the further one would need to move from the worst case correlation structure. For this reason, a risk-averse policymaker might prefer increasing expenditure on a policy with a lower Breakdown Statistic. We construct the Breakdown Statistic for a menu of policies in Table 3, fixing $k = 1$ in Section 4. Comparing Medicare Part D (0.59) and Foster Care Provision (0.67), we find that rejecting the null hypothesis that $H_0 : MVPF < 1$ for Foster Care Provision requires the largest deviation from the worst-case correlation structure.

[4]We also note that while in the case of Medicare Part-D, the confidence intervals implied by assuming independence are conservative, such an assumption can potentially lead to invalid confidence intervals in other cases.

Finally, we turn to the policies evaluated using randomized trials, the setting of interest in Section 3.2. In each case, we add additional sign restrictions to the optimization problem, as motivated by Proposition 1. In the case of Paycheck Plus, for instance, it seems plausible to assume that individuals with higher after-tax income are also more likely to have higher earnings and are more likely to participate in the labor force. Leveraging the additional sign restrictions, we compute the confidence intervals for the Paycheck Plus MVPF by solving **SDP.2**. Figure 2 shows the confidence intervals for the MVPF estimates of Job Start, Paycheck Plus, Work Advance, and Year Up, by including additional sign constraints. A key takeaway is that including the additional sign constraints motivated by Proposition 1 can meaningfully sharpen inference: in the case of Paycheck Plus, the data allow us to rule out values for the MVPF smaller than -0.38 and larger than 2.37, reducing the width of the confidence intervals beyond the worst case by nearly 30%.

Table 1: MVPF Calculation for Paycheck Plus

	(1)	(2)	(3)
	Year	Estimate	SE
Average Bonus Paid	2014	1399	
Average Bonus Paid	2015	1364	
Take-Up	2014	45.90%	
Take-Up	2015	34.80%	
Extensive Margin Labor Market ($\widehat{\beta}_1$)	2014	0.90%	0.65%
Extensive Margin Labor Market ($\widehat{\beta}_2$)	2015	2.5%	0.91%
Impact on After Tax Income ($\widehat{\beta}_3$)	2014	654	187.79
Impact on Earnings ($\widehat{\beta}_4$)	2014	33	43.35
Impact on After Tax Income ($\widehat{\beta}_5$)	2015	645	241.15
Impact on Earnings ($\widehat{\beta}_6$)	2015	192	177.71
WTP		1071	
Net Government Costs		1074	
MVPF		0.996	

Notes: The table reports the inputs to compute the MVPF for the Paycheck Plus program. The causal effects and their corresponding standard errors are reported in Miller, Katz, Azurdia, Isen, and Schultz (2017). The MVPF for Paycheck Plus is computed using these estimates in Hendren and Sprung-Keyser (2020).

Table 2: Inference for Medicare Part-D MVPF

(1)	(2)	(3)	(4)
MVPF	Exact CI	Independence CI	Worst-Case CI
1.37	[0.80, 1.95]	[0.52, 2.22]	[0.18, 2.57]

Notes: The table reports 95% confidence intervals for the MVPF of the introduction of Medicare Part-D, using causal effects reported in Wettstein (2020). Column 1 reports the point estimate for the MVPF computed by Wettstein (2020). Column 2 reports the exact confidence intervals for the estimated MVPF. The exact confidence intervals are computed with the Seemingly Unrelated Regression (SUR) approach of Zellner (1962) using the (publicly available) microdata underlying the causal effects in Wettstein (2020). Column 3 reports the confidence intervals under the assumption that all the causal effects are uncorrelated with each other, i.e., the off-diagonal entries of the variance-covariance matrix are equal to 0. Column 4 reports the confidence intervals computed using the method described in Section 3.1 by solving **SDP.1**.

Table 3: Breakdown Statistics for MVPF

	(1)
	Breakdown Statistic
Medicare Part D	0.59
Foster Care Provision	0.67
UI Extension	0.59

Notes: The table reports Breakdown Statistic for the MVPF of Medicare Part-D (Wettstein, 2020), Foster Care Provistion (Baron and Gross, 2022), and extension of Unemployment Insurance (Huang and Yang, 2021), using the method described in Section 4. The reported Breakdown Statistic is calculated with respect to the null hypothesis, $H_0 : MVPF < 1$.

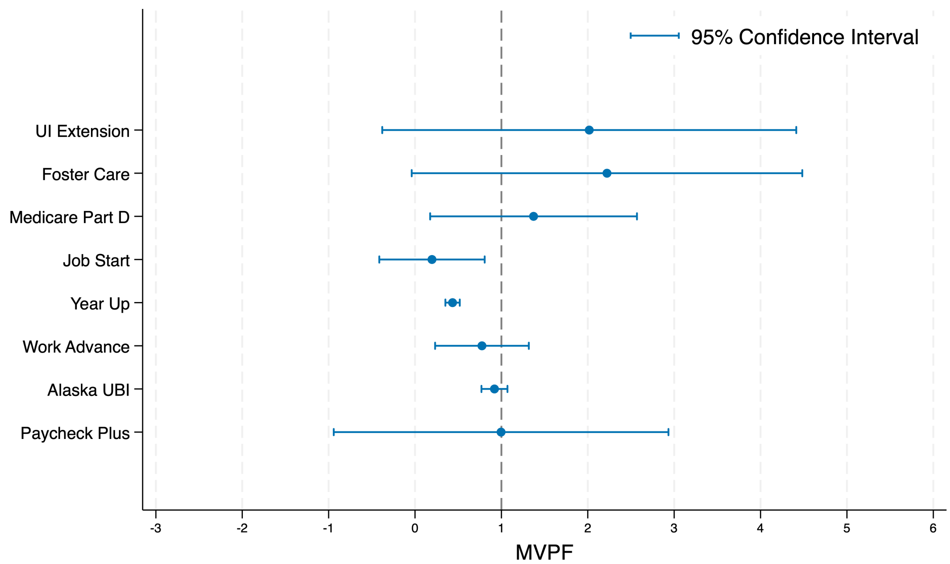

Figure 1: MVPF Confidence Intervals

Notes. The figure reports the 95% confidence intervals for the MVPF of eight different policies. The construction for the MVPF of each policy is detailed in Appendix Section C. The confidence intervals are computed using the method described in Section 3.1, by solving **SDP.1**.

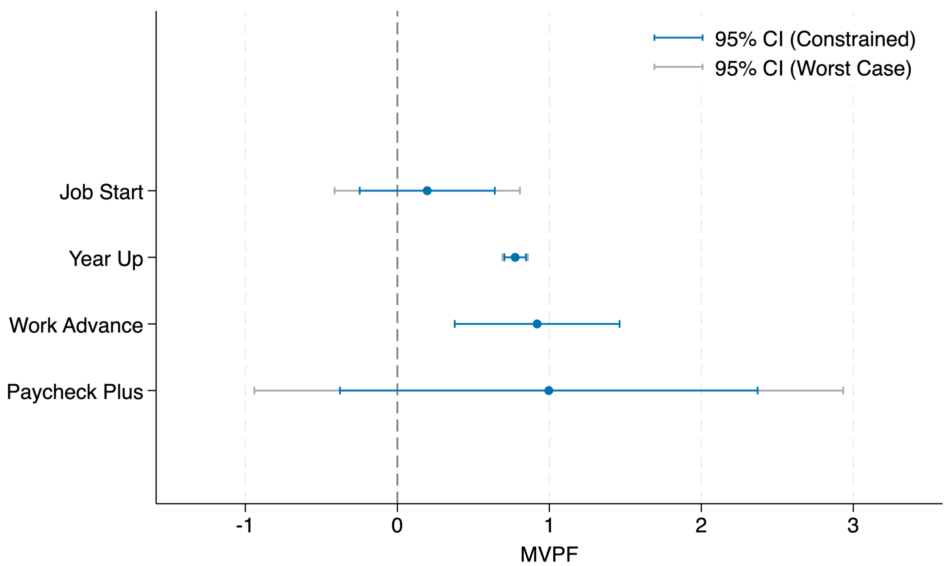

Figure 2: MVPF Confidence Intervals with Sign Constraints

Notes. The figure reports the 95% confidence intervals for the MVPF of four different policies that are evaluated using randomized trials. The construction for the MVPF of each policy is detailed in Appendix Section C. The confidence intervals are computed using the method described in Section 3.2.

Appendix

Table of Contents

A Proof of Proposition 1

Define $Y_{ij}(1)$ as the treated potential outcome j of unit i and $Y_{ij}(0)$ as the control potential outcome j of unit i, where $j = 1, ..., d$. Let $Z_i \in \{0, 1\}$ denote whether unit i is assigned the treatment or control group. We assume that treatment assignment is independent of the vector of potential outcomes, i.e. $\Big(Y_{ij}(1), Y_{ij}(0)\Big) \perp Z_i$ for all $j = 1, \ldots, d$. We decompose the potential outcomes for each j into an expectation component and an error term:

$$Y_{ij}(1) = \mathbb{E}[Y_{ij}(1)] + \varepsilon_{ij}(1)$$
$$Y_{ij}(0) = \mathbb{E}[Y_{ij}(0)] + \varepsilon_{ij}(0)$$

where $\varepsilon_{ij}(1) \equiv Y_{ij}(1) - \mathbb{E}[Y_{ij}(1)]$ and $\varepsilon_{ij}(0) \equiv Y_{ij}(0) - \mathbb{E}[Y_{ij}(0)]$. We rewriting the observed outcome variable, Y_{ij}, as follows.

$$\begin{aligned}
Y_{ij} &= Z_i Y_{ij}(1) + (1 - Z_i)\, Y_{ij}(0) \\
&= Z_i\Big(\mathbb{E}\left[Y_{ij}(1)\right] + \varepsilon_{ij}(1)\Big) + \Big(1 - Z_i\Big)\Big(\mathbb{E}\left[Y_{ij}(0)\right] + \varepsilon_{ij}(0)\Big) \\
&= \mathbb{E}\left[Y_{ij}(0)\right] + Z_i\Big(\mathbb{E}\left[Y_{ij}(1)\right] - \mathbb{E}\left[Y_{ij}(0)\right]\Big) + Z_i\varepsilon_{ij}(1) + (1 - Z_i)\,\varepsilon_{ij}(0) \\
&= \alpha_j + \beta_j Z_i + \varepsilon_{ij}
\end{aligned}$$

where $\alpha_j \equiv \mathbb{E}\left[Y_{ij}(0)\right], \beta_j \equiv \mathbb{E}\left[Y_{ij}(1) - Y_{ij}(0)\right]$, and $\varepsilon_{ij} \equiv Z_i\varepsilon_{ij}(1) + (1 - Z_i)\,\varepsilon_{ij}(0)$. Suppose that effect of the treatment Z_i is evaluated using d linear regressions:

$$\begin{aligned}
Y_{i1} &= \alpha_1 + \beta_1 Z_i + \varepsilon_{i1} \\
&\vdots \\
Y_{id} &= \alpha_d + \beta_d Z_i + \varepsilon_{id}
\end{aligned}$$

Let the vector of Average Treatment Effects (ATEs) be given by $\beta \in \mathbb{R}^d$ and our estimates be given by $\widehat{\beta} \in \mathbb{R}^d$. Under our maintained assumptions, we have that

$$\sqrt{n}(\widehat{\beta} - \beta) \to \mathcal{N}(0, \mathbf{V})$$

For any $p \neq q$, the (p, q)-th entry of the variance-covariance matrix $\mathbf{V}$ can be expressed as follows (Hansen, 2022):

$$\mathbb{C}\text{ov}[\beta_p, \beta_q] = \frac{\mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}]\mathbb{E}[Z_i]^2 - \mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}Z_i]\mathbb{E}[Z_i] - \mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}Z_i]\mathbb{E}[Z_i] + \mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}Z_i]}{\mathbb{E}[Z_i]^2(1 - \mathbb{E}[Z_i])^2} \tag{A.1}$$

Note that the term $\mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}Z_i]$ can be simplified as follows:

$$\begin{aligned}
\mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}Z_i] &= \mathbb{E}[\varepsilon_{ip}\varepsilon_{iq} \mid Z_i = 1]\mathbb{E}[Z_i] \\
&= \mathbb{E}\Big[Y_{ip}(1)Y_{iq}(1) - Y_{ip}(1)\mathbb{E}[Y_{iq}(1)] - \mathbb{E}[Y_{ip}(1)]Y_{iq}(1) + \mathbb{E}[Y_{ip}(1)]\mathbb{E}[Y_{iq}(1)] \mid Z_i = 1\Big]\mathbb{E}[Z_i] \\
&= \Big(\mathbb{E}[Y_{ip}(1)Y_{iq}(1) \mid Z_i = 1] - \mathbb{E}[Y_{ip}(1) \mid Z_i = 1]\mathbb{E}[Y_{iq}(1) \mid Z_i = 1]\Big)\mathbb{E}[Z_i] \\
&= \mathbb{E}[Z_i]\mathbb{C}\text{ov}\Big(Y_{ip}(1), Y_{iq}(1) \mid Z_i = 1\Big) = \mathbb{E}[Z_i]\mathbb{C}\text{ov}\Big(Y_{ip}, Y_{iq} \mid Z_i = 1\Big)
\end{aligned}$$

where the first equality follows from the Law of Total Probability and the third equality follows from random assignment of the treatment Z_i. Similarly, $\mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}]$ be simplified as follows:

$$
\begin{aligned}
\mathbb{E}[\varepsilon_{ip}\varepsilon_{iq}] &= \mathbb{E}\left[\Big(Z_i\varepsilon_{ip}(1) + (1 - Z_i)\,\varepsilon_{ip}(0)\Big)\Big(Z_i\varepsilon_{iq}(1) + (1 - Z_i)\,\varepsilon_{iq}(0)\Big)\right] \\
&= \mathbb{E}\Big[Z_i\varepsilon_{ip}(1)\varepsilon_{iq}(1)\Big] + \mathbb{E}\Big[\,(1 - Z_i)\,\varepsilon_{ip}(0)\varepsilon_{iq}(0)\Big] \\
&= \mathbb{E}\Big[Z_i\Big(Y_{ip}(1) - \mathbb{E}[Y_{ip}(1)]\Big)\Big(Y_{iq}(1) - \mathbb{E}[Y_{iq}(1)]\Big)\Big] \\
&\qquad + \mathbb{E}\Big[\,(1 - Z_i)\,\Big(Y_{ip}(0) - \mathbb{E}[Y_{ip}(0)]\Big)\Big(Y_{iq}(0) - \mathbb{E}[Y_{iq}(0)]\Big)\Big] \\
&= \mathbb{E}\Big[\Big(Y_{ip}(1) - \mathbb{E}[Y_{ip}(1)]\Big)\Big(Y_{iq}(1) - \mathbb{E}[Y_{iq}(1)]\Big) \mid Z_i = 1\Big]\mathbb{E}[Z_i] \\
&\qquad + \mathbb{E}\Big[\Big(Y_{ip}(0) - \mathbb{E}[Y_{ip}(0)]\Big)\Big(Y_{iq}(0) - \mathbb{E}[Y_{iq}(0)]\Big) \mid Z_i = 0\Big]\mathbb{E}[1 - Z_i] \\
&= \mathbb{E}[Z_i]\mathbb{C}\text{ov}\Big(Y_{ip}(1), Y_{iq}(1) \mid Z_i = 1\Big) + \Big(1 - \mathbb{E}[Z_i]\Big)\mathbb{C}\text{ov}\Big(Y_{ip}(0), Y_{iq}(0) \mid Z_i = 0\Big) \\
&= \mathbb{E}[Z_i]\mathbb{C}\text{ov}\Big(Y_{ip}, Y_{iq} \mid Z_i = 1\Big) + \Big(1 - \mathbb{E}[Z_i]\Big)\mathbb{C}\text{ov}\Big(Y_{ip}, Y_{iq} \mid Z_i = 0\Big)
\end{aligned}
$$

Plugging the above into A.1 and simplifying we obtain,

$$
\mathbb{C}\text{ov}[\beta_p, \beta_q] = \frac{\mathbb{C}\text{ov}(Y_{ip}, Y_{iq} \mid Z_i = 1)}{\mathbb{E}[Z_i]} + \frac{\mathbb{C}\text{ov}(Y_{ip}, Y_{iq} \mid Z_i = 0)}{1 - \mathbb{E}[Z_i]}
$$

B Inference Procedure in Hendren and Sprung-Keyser (2020)

To construct confidence intervals for Marginal Value of Public Funds estimates, Hendren and Sprung-Keyser (2020) adopt a "parametric bootstrap" approach. The inference procedure begins by specifying a correlation structure across the estimates. The correlation structure is user-specified to "maximise the width of our confidence intervals where estimates are from the same sample." Then, Hendren and Sprung-Keyser (2020) draw from a joint normal distribution centered at the estimates with the specified correlation structure. For each draw, they compute the MVPF obtaining a numerical distribution of the MVPF. The 2.5-th and 97.5-th quantiles of this numerical distribution are then used to construct the confidence intervals.

If the correlation structure in the first step of the inference procedure is correctly specified to maximize the width of the confidence intervals, this approach should yield valid inference on the MVPF and the resulting confidence intervals should align with our approach. However, the key departure in our approach from Hendren and Sprung-Keyser (2020) is finding the correlation structure that maximizes the width of the confidence intervals by solving **SDP.1** rather than assuming it. Indeed, the correlation structure that maximizes the confidence intervals may not be obvious to the analyst at the outset, and simply assuming a correlation structure does not provide the statistical guarantee of size control. Moreover, casting the problem of finding the appropriate correlation structure as an optimization problem allows us to flexibly incorporate additional setting-specific information such as known independence or the direction of correlation implied by Proposition 1.

C Details on MVPF Construction

In this Section, we detail the construction of the MVPF of all policies discussion in Section 5. In each case, we defer further discussion of the MVPF to the paper providing the MVPF construction for a given policy.

C.1 Paycheck Plus

The estimates used to construct the MVPF for Paycheck Plus are drawn from Miller, Katz, Azurdia, Isen, and Schultz (2017). The estimates are summarized in the following Table:

Table C.1: MVPF Calculation for Paycheck Plus

	(1)	(2)	(3)
	Year	Estimate	SE
Average Bonus Paid	2014	1399	
Average Bonus Paid	2015	1364	
Take-Up	2014	45.90%	
Take-Up	2015	34.80%	
Extensive Margin Labor Market ($\widehat{\beta}_1$)	2014	0.90%	0.65%
Extensive Margin Labor Market ($\widehat{\beta}_2$)	2015	2.5%	0.91%
Impact on After Tax Income ($\widehat{\beta}_3$)	2014	654	187.79
Impact on Earnings ($\widehat{\beta}_4$)	2014	33	43.35
Impact on After Tax Income ($\widehat{\beta}_5$)	2015	645	241.15
Impact on Earnings ($\widehat{\beta}_6$)	2015	192	177.71

We replicate the construction of the MVPF for Paycheck Plus from Hendren and Sprung-Keyser (2020), as follows:

$$
\begin{aligned}
MVPF_{\text{Paycheck Plus}} = f(\widehat{\beta}) &= \frac{1399 \times (45 - \widehat{\beta}_1) + 1364 \times (34.8 - \widehat{\beta}_2)}{(\widehat{\beta}_3 - \widehat{\beta}_4) + (\widehat{\beta}_5 - \widehat{\beta}_6)} \\
&= 0.996.
\end{aligned}
$$

C.2 Alaska UBI

The estimates used to construct the MVPF for Alaska UBI are drawn from Jones and Marinescu (2022).

Table C.2: MVPF Calculation for Alaska UBI

	(1)	(2)
	Estimate	SE
Full-Time Employment Effect (β_1)	0.001	0.016
Part-Time Employment Effect (β_2)	0.018	0.007

We replicate the construction of the MVPF for Alaska UBI from Hendren and Sprung-Keyser (2020), as follows:

$$MVPF_{\text{Alaska UBI}} = f(\widehat{\beta}) = \frac{1000}{1000 - \left(\beta_1 \times 5567.88 \times \frac{1000}{1602}\right) + \left(0.2 \times 0.5 \times \beta_2 \times \frac{1000}{1602} \times 80830.57\right)}$$
$$= 0.92.$$

C.3 Work Advance

The estimates used to construct the MVPF for Work Advance are drawn from Hendra, Greenberg, Hamilton, Oppenheim, Pennington, Schaberg, and Tessler (2016) and Schaberg (2017).

Table C.3: MVPF Calculation for Work Advance

	(1)	(2)
	Estimate	SE
Year 2 Earnings Effect (β_1)	1945	692.90
Year 3 Earnings Effect (β_2)	1865	664.40

We replicate the construction of the MVPF for Work Advance from Hendren and Sprung-

Keyser (2020), as follows:

$$MVPF_{\text{Work Advance}} = f(\widehat{\beta})\frac{\frac{\beta_1\times(1-0.003)}{1.03}+\frac{\beta_2\times(1-0.003)}{1.03^2}}{5641-940-\beta_1\times 0.003-\beta_2\times 0.003}$$
$$= 0.78$$

C.4 Year Up

The estimates used to construct the MVPF for Year Up are drawn from Fein and Hamadyk (2018).

Table C.4: Year Up

	(1) Estimate	(2) SE
Year 0 Earnings (β_1)	-5338	238
Year 1 Earnings (β_2)	5181	474
Year 2 Earnings (β_3)	7011	619
Discount Rate	3%	
Tax Rate	18.6%	
Per-Participant Cost	\$28,290	
Student Stipend	\$6,614	

We replicate the construction of the MVPF for Year Up from Hendren and Sprung-Keyser (2020), as follows:

$$MVPF_{\text{Year Up}} = f(\widehat{\beta}) = \frac{(1-0.186)\times(\beta_1+\beta_2/0.03+\beta_3/1.03^2)+6614}{28290-0.186\times(\beta_1+\beta_2+\beta_3)}$$
$$= 0.43$$

C.5 Job Start

The estimates used to construct the MVPF for Job Start are drawn from Cave et al. (1993).

We replicate the construction of the MVPF for Year Up from Hendren and Sprung-Keyser

Table C.5: MVPF Calculation for Job Start

	(1) Estimate	(2) SE
Year 1 Earnings Effect (β_1)	-499	151.65
Year 2 Earnings Effect (β_2)	-121	209.20
Year 3 Earnings Effect (β_3)	423	258.67
Year 4 Earnings Effect (β_4)	410	267.25
Year 1 AFDC Effect (β_5)	63	53.96
Year 2 AFDC Effect (β_6)	24	62.94
Year 3 AFDC Effect (β_7)	-3	85.47
Year 4 AFDC Effect (β_8)	-11	84.97
Year 1 Food Stamps Effect (β_9)	-45	35.66
Year 2 Food Stamps Effect (β_{10})	-42	34.83
Year 3 Food Stamps Effect (β_{11})	31	40.94
Year 4 Food Stamps Effect (β_{12})	31	45.21
Year 1 General Assistance Effect (β_{13})	24	23.54
Year 2 General Assistance Effect (β_{14})	7	15.14
Year 3 General Assistance Effect (β_{15})	-6	24.82
Year 4 General Assistance Effect (β_{16})	3	26.53

(2020), as follows:

$$MVPF_{\text{Job Start}} = f(\widehat{\beta}) = \frac{\sum_{i=1}^{4} \beta_i \times 0.993 + \sum_{i=5}^{16} \beta_i + 606.13}{4548}$$
$$= 0.20$$

C.6 Medicare Part D

The estimates used to construct the MVPF for Medicare Part D are drawn from Wettstein (2020).

Table C.6: MVPF Calculation for Introduction of Medicare Part D

	(1) Estimate	(2) SE
Effect on Labor Force Participation (β_1)	-0.10	0.03
Effect on Income (β_2)	-6665.40	1986.92
Semi-Elasticity of Demand for Insurance (β_3)	0.14	0.03

We replicate the construction of the MVPF for Introduction of Medicare Part D from Wettstein (2020), as follows:

$$MVPF_{\text{Medicare Part D}} = f(\widehat{\beta}) = \frac{0.65 \times \frac{\beta_1 \times -100}{25000} \times \frac{6126}{0.4}}{(0.65 + 0.65 \times \frac{\beta_3}{0.887} - \beta_1 - 0.28 \times \frac{\beta_2}{1588})/0.65}$$
$$= 1.37$$

C.7 Foster Care

The estimates used to construct the MVPF for Foster Care are drawn from Baron and Gross (2022). We replicate the construction of the MVPF for Foster Care from Baron and Gross

Table C.7: MVPF Calculation for Foster Care

	(1)	(2)
	Estimate	SE
Society's Willingness to Pay (β_1)	83854	29715
Cost Savings to the Government (β_2)	12188	6212

(2022), as follows:

$$MVPF_{\text{Foster Care}} = f(\widehat{\beta}) = \frac{\beta_1}{49920 - \beta_2}$$
$$= 2.22$$

C.8 UI Extension

The estimates used to construct the MVPF for UI Extension are drawn from Huang and Yang (2021).

We replicate the construction of the MVPF for extension of Unemployment Insurance

Table C.8: MVPF Calculation for UI Extension

	(1)	(2)
	Estimate	SE
Effect on Transfers from UI (β_1)	0.038	0.009
Effect on Transfers from Re-employment bonus(β_2)	0.019	0.011
Effect on Benefit Duration (β_3)	56.91	1.96
Effect on Unemployment Duration (β_4)	36.90	6.90

from Huang and Yang (2021), as follows:

$$MVPF_{\text{UI Extension}} = f(\widehat{\beta}) = \frac{0.77 \times \frac{\beta_1 + (\beta_2/2)}{\beta_2} + 0.23}{1 + (1/72.9) \times (\beta_3 - 55.8 - 0.5 \times (\beta_3 - 55.8) + 0.12 \times \beta_4)}$$
$$= 2.02$$

References

Baron, E. J., and M. Gross (2022): "Is there a foster care-to-prison pipeline? Evidence from quasi-randomly assigned investigators," Discussion paper, National Bureau of Economic Research. 3, C.7, C.7

Cave, G., et al. (1993): "JOBSTART. Final Report on a Program for School Dropouts.," *New York: MDRC.* C.5

Chetty, R. (2009): "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 1(1), 451–488. 1

Christensen, G., and E. Miguel (2018): "Transparency, reproducibility, and the credibility of economics research," *Journal of Economic Literature*, 56(3), 920–980. 1

Cocci, M. D., and M. Plagborg-Møller (2024): "Standard errors for calibrated parameters," *Review of Economic Studies.* 1, 3.1

d'Haultfoeuille, X., C. Gaillac, and A. Maurel (2022): "Partially linear models under data combination," Discussion paper, National Bureau of Economic Research. 1

Diegert, P., M. A. Masten, and A. Poirier (2022): "Assessing omitted variable bias when the controls are endogenous," *arXiv preprint arXiv:2206.02303.* 1

Elliott, G., U. K. Müller, and M. W. Watson (2015): "Nearly optimal tests when a nuisance parameter is present under the null hypothesis," *Econometrica*, 83(2), 771–811. 3

Fan, Y., X. Shi, and J. Tao (2023): "Partial identification and inference in moment models with incomplete data," *Journal of Econometrics*, 235(2), 418–443. 1

Fein, D., and J. Hamadyk (2018): "Bridging the opportunity divide for low-income youth: Implementation and early impacts of the year up program," *OPRE Report*, 65, 2018. C.4

GRANT, M., AND S. BOYD (2014): "CVX: Matlab software for disciplined convex programming, version 2.1," . 3.1

GRANT, M. C., AND S. P. BOYD (2008): "Graph implementations for nonsmooth convex programs," in *Recent advances in learning and control*, pp. 95–110. Springer. 3.1

HANSEN, B. (2022): *Econometrics*. Princeton University Press. A

HECKMAN, J. J., S. H. MOON, R. PINTO, P. A. SAVELYEV, AND A. YAVITZ (2010): "The rate of return to the HighScope Perry Preschool Program," *Journal of public Economics*, 94(1-2), 114–128. 1

HENDRA, R., D. H. GREENBERG, G. HAMILTON, A. OPPENHEIM, A. PENNINGTON, K. SCHABERG, AND B. L. TESSLER (2016): "Encouraging evidence on a sector-focused advancement strategy: Two-year impacts from the WorkAdvance demonstration," *New York: MDRC*. C.3

HENDREN, N., AND B. SPRUNG-KEYSER (2020): "A unified welfare analysis of government policies," *The Quarterly Journal of Economics*, 135(3), 1209–1318. 1, 5, 1, (document), B, C.1, C.2, C.3, C.4, C.5

HÖSSJER, O., AND A. SJÖLANDER (2022): "Sharp lower and upper bounds for the covariance of bounded random variables," *Statistics & Probability Letters*, 182, 109323. 3.2

HUANG, P.-C., AND T.-T. YANG (2021): "The welfare effects of extending unemployment benefits: Evidence from re-employment and unemployment transfers," *Journal of Public Economics*, 202, 104500. 3, C.8, C.8

JONES, D., AND I. MARINESCU (2022): "The labor market impacts of universal and permanent cash transfers: Evidence from the Alaska Permanent Fund," *American Economic Journal: Economic Policy*, 14(2), 315–340. C.2

Manski, C. F., and J. V. Pepper (2018): "How do right-to-carry laws affect crime rates? Coping with ambiguity using bounded-variation assumptions," *Review of Economics and Statistics*, 100(2), 232–244. 1

Masten, M. A., and A. Poirier (2020): "Inference on breakdown frontiers," *Quantitative Economics*, 11(1), 41–111. 1

Miller, C., L. F. Katz, G. Azurdia, A. Isen, and C. B. Schultz (2017): "Expanding the earned income tax credit for workers without dependent children: Interim findings from the paycheck plus demonstration in new york city," *New York: MDRC, September*. 5, 1, C.1

Rambachan, A., and J. Roth (2023): "A more credible approach to parallel trends," *Review of Economic Studies*, 90(5), 2555–2591. 1

Romano, J. P., and E. Lehmann (2005): "Testing statistical hypotheses," . 3

Rose, E. K. (2018): "The effects of job loss on crime: evidence from administrative data," *Available at SSRN 2991317*. 1

Ruggles, S., C. A. Fitch, and E. Roberts (2018): "Historical census record linkage," *Annual review of sociology*, 44(1), 19–37. 1

Schaberg, K. (2017): "Can Sector Strategies Promote Longer-Term Effects? Three-Year Impacts from the WorkAdvance Demonstration.," *MDRC*. C.3

Spini, P. E. (2024): "Robustness, heterogeneous treatment effects and covariate shifts," *arXiv preprint arXiv:2112.09259*. 1

Wettstein, G. (2020): "Retirement lock and prescription drug insurance: Evidence from medicare part d," *American Economic Journal: Economic Policy*, 12(1), 389–417. 2, 3, C.6, C.6

Zellner, A. (1962): "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias," *Journal of the American statistical Association*, 57(298), 348–368. 1, 2