Identification at the Stability Frontier: Trading Relevance for Exogeneity*

Nicolás Ronderos†

Abstract

This paper studies identification and estimation with imperfect instruments in a class of parameter-dependent moment models. We propose an approach in which instrument validity is characterized through a stability condition on parameter-dependent moment mappings. The key result establishes that local stability is equivalent to local identification: the model is well defined if and only if the corrected instrument retains sufficient effective relevance. The framework endogenizes the trade-off between relevance and exogeneity. Increasing the correction reduces contamination but weakens instrument strength, and the admissible degree of correction is disciplined by stability. The normalized correction parameter admits a regularization interpretation, with the stability region defining an identification-constrained regularization set whose boundary corresponds to the largest correction consistent with local identification. Branch selection can be guided by economic restrictions; absent such information, the approach provides a structured sensitivity analysis. We establish existence, uniqueness, convergence, consistency, and asymptotic normality of the resulting estimator. Monte Carlo simulations and an empirical application illustrate its performance.

Keywords: Identification; Stability; Imperfect Instruments; GMM

JEL Classification: C13,C26,C51

* Replication files containing the simulation programs, estimation routines, and scripts necessary to reproduce the results reported in the paper will be made available to ensure transparency and reproducibility. Replication materials will be deposited in a public repository upon acceptance.

† Ph.D in economics, Pontificia Universidad Javeriana. The author received no financial support for the research, authorship, or publication of this article. The author declares that there are no conflicts of interest related to this research.

1. Introduction

This paper studies identification and estimation with imperfect instruments. Its central claim is that, in a class of parameter-dependent moment models, local identification can be characterized as a stability property of the estimating equation. We develop this idea in an instrumental-variables setting where instruments may be relevant but not fully exogenous. In that environment, the key econometric problem is not only whether instrument validity fails, but how much of the instrument's relevance can still be used for identification once limited exclusion violations are admitted. This paper shows that the answer is disciplined by stability: identification is preserved only to the extent that the corrected instrument retains sufficient effective relevance. In this sense, the estimator trades relevance for exogeneity, but does so endogenously through the stability of the moment mapping rather than through externally imposed bounds.

Instrumental variables (IV) methods are a central tool for causal identification in empirical economics (Angrist and Krueger, 2001; Angrist and Imbens, 1995). Their validity relies on the exclusion restriction, which requires instruments to affect the outcome only through the endogenous regressor. In many empirical applications, however, the plausibility of this assumption is contested. When exclusion fails, conventional IV estimators become biased and inconsistent. These concerns have generated a large literature on imperfect instruments, partial identification, and sensitivity analysis (Conley, Hansen, and Rossi, 2012; Nevo and Rosen, 2012; Armstrong and Kolesár, 2021; Kolesár et al., 2015; Manski, 2003; Tamer, 2010). A common feature of this literature is that the admissible degree of invalidity is restricted externally, through researcher-imposed bounds or ordering assumptions. By contrast, the approach developed here determines the admissible correction internally: the relevant restriction is not imposed on the data-generating process from outside the estimator, but emerges from the requirement that the parameter-dependent moment mapping remain locally stable.

The paper develops this idea using a residual-based correction of the instrument. After projection onto the fitted first-stage index, the model is just-identified in the structural parameter, conditional on a correction parameter that is not itself point-identified from the data. The resulting estimator is defined by a parameter-dependent moment condition and admits a fixed-point representation. The key theoretical result is that local stability of this mapping is equivalent to local identification of the structural parameter. The estimator is therefore well defined if and only if the corrected instrument retains enough effective relevance after the residual-based adjustment. As the correction increases, contamination from exclusion violations is reduced, but the instrument is simultaneously weakened. At the boundary of the admissible region, effective relevance collapses, the mapping approaches instability, and local identification breaks down. In this framework, local stability and local identification therefore coincide.

This equivalence has several implications. First, it gives the relevance–exogeneity trade-off an internal econometric structure. Rather than treating relevance and exogeneity as competing desiderata balanced through ad hoc sensitivity parameters, the framework organizes them through a single derivative object that jointly characterizes admissibility, convergence, sensitivity, and identification. Second, it implies that the correction parameter admits a regularization interpretation. Normalizing the correction by first-stage strength yields a scale-free parameter that measures how much instrument relevance is being sacrificed to improve exogeneity. The stability region then defines an identification-constrained regularization set, and its boundary corresponds to the largest correction consistent with local identification. Third, this perspective clarifies why the method is particularly useful in settings with excess relevance but imperfect exogeneity: conventional IV estimators exploit all available first-stage variation as if it were valid, whereas the present framework discounts that variation endogenously, but only up to the point at which identification can still be sustained.

Under mild conditions, the correction path is monotone within each connected component of the admissible region. This monotonicity implies that admissible estimators are naturally ordered by the strength of the correction. The relevant branch, however, still depends on the direction of the exclusion violation. Branch selection may therefore be disciplined by economic restrictions, institutional knowledge, or auxiliary diagnostics. When such information is available, the boundary of the corresponding admissible branch provides a disciplined point-estimation rule. When it is not, the framework is more naturally interpreted as a structured sensitivity analysis over admissible correction paths. In either case, the contribution is not merely computational. The fixed-point representation is useful because it reveals that the same stability condition governs the existence of the estimator, the convergence of the iteration, the admissible degree of correction, and the local identification of the structural parameter.

The paper also studies the statistical consequences of this framework. We show that, for a given regularization parameter, the estimator satisfies existence, uniqueness, convergence, consistency, and asymptotic normality under standard regularity conditions. Moreover, when instruments are imperfect and bias declines monotonically along the admissible path, the relevant stability boundary has a risk interpretation: it corresponds to the largest bias-reducing correction consistent with local identification, and, under these conditions, to the mean-squared-error-optimal choice within the admissible region. This also yields a comparison with conventional k-class estimators. Because k-class procedures continue to use the full relevance of the original instrument, they retain non-vanishing asymptotic bias under exclusion violations. By contrast, the boundary-selected estimator can weakly dominate them whenever the reduction in squared bias exceeds the associated variance increase. The gains therefore arise precisely in environments where relevance is strong enough to survive correction, but not exogenous enough to be taken at face value.

The proposed framework connects to several strands of the econometric literature. First, it relates to work on estimation under misspecified moment conditions (Hansen, 1982; Hall and Inoue, 2003; Guggenberger and Smith, 2005). Second, it complements the literature on imperfect instruments and sensitivity analysis (Conley et al., 2012; Nevo and Rosen, 2012; Armstrong and Kolesár, 2021), but differs from it by replacing exogenous bounds on invalidity with stability-implied admissible corrections. Third, it is conceptually related to control-function approaches (Heckman, 1978; Wooldridge, 2015), although the correction operates on the instrument rather than on the structural equation. More broadly, the paper contributes a different perspective on identification itself: in the present setting, identification is not a primitive binary property of the model, but the outcome of a stability condition governing how much relevance can be traded away in order to improve exogeneity. The paper does not aim to provide a fully general theory for all parameter-dependent moment models, but rather to isolate a class of economically meaningful problems in which the stability–identification equivalence can be made explicit and operational.

What distinguishes the present framework is that the restriction governing identification is not imposed on the data-generating process, but on the behavior of the estimator itself. Conventional approaches either maintain exact exogeneity, impose bounds on violations, or modify the projection of the endogenous regressor while leaving the instrument space unchanged. In contrast, the proposed method alters the instrument directly through a residual-based correction and disciplines this adjustment through the stability of a parameter-dependent mapping. As a result, admissible degrees of instrument imperfection are not specified ex ante, but emerge endogenously from the requirement that the estimator remain well defined, locally stable, and identified. This shifts the role of identification from an assumption about the environment to a property of the estimation procedure, linking admissibility, convergence, and statistical performance within a single internal object.

Monte Carlo simulations and an empirical application illustrate the performance of the estimator. The simulations deliberately stress the method in settings known to challenge instrumental-variables procedures, including weak identification, heterogeneous invalidity, nonlinear exclusion failures, and mixtures of valid and invalid instruments (Staiger and Stock, 1997; Stock and Wright, 2000; Andrews, Stock and Sun, 2019). The empirical application revisits the institutions-and-growth design of Acemoglu, Johnson, and Robinson (2001), where the validity of settler mortality as an instrument has been widely debated (Albouy, 2012; Acemoglu et al., 2012). This setting is especially useful for the present framework because the instrument is clearly relevant, but its exclusion restriction remains controversial. It therefore provides a natural illustration of how relevance may be discounted, but not discarded, once exact exogeneity is relaxed.

The paper makes three main contributions. First, it introduces a class of parameter-dependent moment estimators in which local identification can be characterized as stability. Second, it develops a stability-based framework for estimation with imperfect instruments, in which the admissible degree of correction is determined endogenously rather than imposed externally. Third, it shows that stability is not merely a computational property, but an econometric object that jointly characterizes local identification, admissible regularization, and statistical performance. The remainder of the paper is organized as follows. Section 2 introduces the estimator and its moment representation. Section 3 presents the econometric framework. Section 4 studies identification, stability, existence, and uniqueness. Sections 5 and 6 establish consistency and asymptotic normality. Section 7 develops the stability-based calibration of the correction path. Section 8 relates the approach to existing estimators. Section 9 reports Monte Carlo evidence, Section 10 presents the empirical application, and Section 11 concludes.

2. Parameter-Dependent Moments and Stability

2.1 Parameter-Dependent Moment Conditions

We consider a class of moment conditions in which the instrument depends on the structural residual. Let $u(\beta) = y - x\beta$, and suppose that estimation is based on moment conditions of the form

$$\mathbb{E}\big[q\big(z, \lambda, u(\beta)\big)u(\beta)\big] = 0 \qquad (1)$$

where z denotes observable instruments, λ is a scalar parameter governing the correction, and $q(\cdot)$ is a measurable function that may depend on the residual $u(\beta)$. This formulation encompasses standard IV as a special case when $q(z, \lambda, u) = z$, and allows for more general corrections that adjust the instrument as a function of the structural residual.

The key feature of this class is that the moment condition depends on the parameter β both directly and through the residual. As a result, the solution to the moment condition cannot, in general, be obtained from a linear orthogonality condition, but instead requires solving a nonlinear consistency problem.

2.2 Fixed-Point Representation

The following result shows that this class of moment conditions admits a natural fixed-point representation.

Theorem 2.1 (Induced fixed-point representation)

Consider the parameter-dependent moment condition (1) where $u(\beta) = y - x\beta$. Suppose that:

1. $\mathbb{E}\big[q\big(z, \lambda, u(\beta)\big)x\big] \neq 0$ in a neighborhood of the solution,
2. the moment function is continuously differentiable in β,
3. the expectations are well defined and finite.

Then the moment condition can be equivalently written as

$$\beta = g(\beta; \lambda) \qquad (2)$$

Where $g(\beta; \lambda) = \mathbb{E}[q(z, \lambda, u(\beta))y]/\mathbb{E}[q(z, \lambda, u(\beta))x]$. In particular, any solution to the moment condition corresponds to a fixed point of the mapping $g(\beta; \lambda)$. See Appendix A.

Theorem 2.1 shows that parameter-dependent moment conditions naturally induce a self-mapping in the structural parameter. The dependence of the moment condition on $u(\beta)$ implies that the parameter determines the moment through the residual, while the moment determines the parameter through the ratio representation.

As a result, estimation in this class of models is inherently a fixed-point problem: the structural parameter must be consistent with the moment condition it generates. The mapping $g(\beta;\lambda)$ provides a natural representation of this consistency requirement. This representation is not imposed by the econometrician, but induced by the structure of the moment condition itself. It therefore provides a convenient framework for analyzing the existence, uniqueness, and stability of solutions, as well as their dependence on the correction parameter λ.

This class includes residual-based instrument corrections as the main case studied here, but also more broadly covers settings in which the estimating moment depends on the parameter through a residual transformation or other parameter-indexed adjustment.

The remainder of the paper exploits this fixed-point structure to study identification and admissibility in models with imperfect instruments. This section introduces a class of parameter-dependent moment models that underlies the analysis. The key feature is that the moment conditions depend on the parameter of interest through a residual-based transformation, inducing a mapping whose stability properties govern both existence and identification.

3. Econometric Framework

The parameter-dependent moment framework developed above applies to settings in which the moment condition depends on the structural parameter through a correction term. The IV model studied below provides the main environment in which the stability–identification equivalence can be developed in economically interpretable form. We now specialize that framework to a linear IV model with potentially invalid instruments, which provides the basis for the theoretical and empirical analysis. Consider the structural system

$$y_i = \alpha + \beta x_i + w_i'\mu + u_i, \qquad x_i = \alpha_x + w_i'\rho + z_i'\gamma + \nu_i \tag{3}$$

where x_i is potentially endogenous, w_i collects included exogenous controls, and $z_i \in \mathbb{R}^L$ denotes the excluded instruments. Let $W = [\mathbf{1}, \boldsymbol{w}]$, by the Frisch-Waugh-Lovell (FWL) theorem, estimation of β in the full model is equivalent to estimation in the residualized system $\tilde{y}_i = \beta\tilde{x}_i + \tilde{u}_i$ with $\tilde{y} = M_W y, \tilde{x} = M_W x$ and $\tilde{z} = M_W z$, where $M_W = I - W(W'W)^{-1}W'$.

If $L > 1$, the excluded instruments may be summarized through the first-stage fitted index $\hat{x} = P_{\tilde{z}}\tilde{x}$. This projected component will serve as the baseline instrument to which the residual-based correction is applied.

The use of the FWL theorem implies that this reduction does not sacrifice generality for the parameter of interest: intercepts, additional exogenous regressors, and multiple excluded instruments are absorbed before the correction is defined, so the fixed-point logic applies to the residualized problem without altering the identification argument for β. For notational simplicity, however, the analysis that follows is written in the residualized scalar case, and tildes are omitted from this point onward.

The same construction extends to settings with multiple endogenous regressors, where the corrected instrument becomes matrix-valued and the fixed-point mapping is vector-valued. For transparency, the main text focuses on the scalar case, while Appendix A develops the corresponding multivariate extension. Thus, we write

$$y_i = \beta x_i + u_i \tag{4}$$

where the variables are understood as residualized with respect to the included exogenous regressors.

Assumption 3.1 (data): $\{y_i, x_i, z_i\}_{i=1}^{n}$ is i.i.d with $\mathbb{E}[z_i^2] < \infty, \mathbb{E}[x_i^2] < \infty, \mathbb{E}[u_i^2] < \infty$.

Assumption 3.2 (Imperfect instrument): The exclusion restriction may fail, so that $\mathbb{E}[z_i u_i] = \delta$ where δ may differ from zero.

Remark: When $\delta \neq 0$, the classical IV moment $\mathbb{E}[z_i(y_i - \beta x_i)] = 0$ fails and TSLS becomes inconsistent.

Let $x_i = \gamma z_i + v_i$ denote the linear projection of x_i on z_i. Where γ denotes the first-stage coefficient. Let $\hat{\gamma}$ be the sample analogue, and the fitted values $\hat{x}_i = \hat{\gamma} z_i$

Assumption 3.3 (relevance): $\mathbb{E}[z_i x_i] = \gamma \mathbb{E}[z_i^2] \neq 0$ and $\gamma \neq 0$.

Assumption 3.4 (*Regularity of the first-stage estimation*): The first-stage estimator $\hat{\gamma}$ is consistent and asymptotically linear $\hat{\gamma} \to_p \gamma, \quad \sqrt{n}(\hat{\gamma} - \gamma) = 1/\sqrt{n} \sum_{i=1}^{n} \varphi_{\gamma,i} + o_p(1)$ with $\mathbb{E}[\varphi_{\gamma,i}] = 0$ and $\mathbb{E}[\varphi_{\gamma,i}^2] < \infty$.

3.1 Residual-Based Correction and Corrected Instrument

Let $\pi \in \Pi \subset \mathbb{R}$ denote the correction parameter. For any candidate value of β, define the structural residual:

$$u_i(\beta) = y_i - \beta x_i \tag{5}$$

Definition 3.1 (Fixed-Point instrument - $\hat{x}$ based)

For a correction parameter π, define the FP:

$$z_i^{FP}(\beta; \pi; \hat{\gamma}) := \hat{x}_i - \hat{\gamma}\lambda u_i(\beta) = \hat{\gamma}\big(z_i - \lambda u_i(\beta)\big), \qquad \lambda \equiv \frac{\pi}{\hat{\gamma}} \tag{6}$$

With its population counterpart $z_i^{FP}(\beta; \pi; \gamma)$. This specification is a particular case of the class defined in Section 2. If the instrument is correlated with the structural error, the residual contains information about this correlation. Subtracting a component proportional to the residual therefore adjusts the instrument to reduce this correlation. Larger values of π generate stronger corrections of the instrument.

Equation (6) is well suited to the present problem because it corrects the instrument along the direction in which exclusion failure enters the model, namely through the structural residual. When $Cov(z, u) \neq 0$, the residual $u(\beta)$ contains the misspecification signal that invalidates the classical IV moment. The correction in (6) therefore has a clear identification logic: it purges the projected instrument by subtracting a residual-based component that is informative about contamination, while retaining the fitted first-stage index as the underlying source of identifying variation. This makes the correction both economically interpretable and econometrically disciplined.

Because the corrected instrument is built from the fitted first-stage component, the raw correction parameter λ is confounded with first-stage strength. Writing $\lambda = \pi/\gamma$ isolates the correction from this nuisance scale effect. We therefore interpret λ as a *normalized regularization parameter*: it measures the size of the residual-based adjustment relative to instrument relevance. The FWL reduction also implies that this representation does not sacrifice generality for the parameter of interest. The scalar formulation is therefore a notational simplification of the identification problem, not a restriction on the underlying linear IV environment.

3.2 Moment Condition

Using the corrected instrument, we define the population moment condition:

$$\psi(\beta;\pi;\gamma) = \mathbb{E}\left[z_i^{FP}(\beta,\pi;\gamma)(y_i - x_i\beta)\right] = 0 \tag{7}$$

In the sample, the corresponding moment condition is:

$$\psi_n(\beta;\pi;\hat{\gamma}) = \frac{1}{n}\sum_{i=1}^{n} z_i^{FP}(\beta;\pi;\hat{\gamma})(y_i - x_i\beta) \tag{8}$$

3.3 Fixed-Point Representation

Because the corrected instrument depends on the structural residual, the moment condition depends on the parameter itself.

Definition 3.2 (Fixed-Point Mapping)

Define the population moments $A(\beta;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)y]$ and $B(\beta;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)x]$. The population mapping is defined as:

$$g(\beta;\pi;\gamma) = \frac{A(\beta;\pi;\gamma)}{B(\beta;\pi;\gamma)} \tag{9}$$

For a given correction parameter π the Fixed-Point GMM estimator is any solution $\hat{\beta}$ satisfying:

$$\hat{\beta}\,(\pi) = g_n\big(\hat{\beta}(\pi);\pi;\hat{\gamma}\big) \tag{10}$$

When the mapping is locally contractive, the estimator can be computed by fixed-point iteration. The sample analogue of (9) is defined as $g_n(\beta(\pi);\pi) = \frac{1}{n}\sum_{i=1}^{n} z_i^{FP}(\beta;\pi;\hat{\gamma})y_i \,/\, \frac{1}{n}\sum_{i=1}^{n} z_i^{FP}(\beta;\pi;\hat{\gamma})x_i$.

3.4 Z-Estimator Interpretation

The fixed-point estimator admits a natural interpretation within the generalized method of moments framework (Hansen, 1982; Chernozhukov and Hong, 2003). The estimator solves the sample moment condition

$$\frac{1}{n}\sum_{i=1}^{n}\psi(W_i,\beta,\pi,\hat{\gamma})=0 \tag{11}$$

Where $W_i=(y_i,x_i,z_i)$. The estimator can also be interpreted as a just-identified Z-estimator (Newey and McFadden, 1994; van der Vaart, 1998). The distinguishing feature relative to conventional GMM is that the moment function depends on the structural parameter through the residual-based correction of the instrument. The estimator is just-identified after projection onto the fitted first-stage index, even when the underlying first stage is built from multiple excluded instruments.

The statistical properties of the estimator therefore depend on the behavior of the mapping g, which is studied in the next section.

4. Stability, Identification, Existence and Uniqueness

Given the fixed-point representation in Section 2, estimation reduces to studying the mapping $\beta=g(\beta;\pi;\gamma)$. The properties of the estimator are governed by the local behavior of this mapping. In particular, its derivative determines existence, local uniqueness, and convergence.

Because the moment condition depends on the residual, effective instrument relevance is itself parameter-dependent. As a result, the same condition that ensures local stability of the mapping also governs local identification. The analysis therefore focuses on characterizing the set of correction parameters for which the mapping admits a locally stable fixed point.

4.1 Population Mapping

Let $g(\beta;\pi;\gamma)$ be the population mapping defined in Section 3. The population parameter is defined as any solution satisfying

$$\beta^*=g(\beta^*;\pi;\gamma) \tag{12}$$

Equivalently, β^* solves the population moment condition (7). The fixed-point representation provides an induced framework for analyzing identification and stability.

4.2 Local Stability and Contraction

The behavior of the fixed point depends on the derivative of the mapping with respect to β. Let

$$g'(\beta;\pi;\gamma)=\frac{\partial g(\beta;\pi;\gamma)}{\partial\beta} \tag{13}$$

The mapping is locally stable if the derivative satisfies $|g'(\beta^*;\pi;\gamma)| < 1$. Under this condition, small deviations from the fixed point are attenuated by the mapping, and the fixed-point iteration converges in a neighborhood of the solution.

This condition has a natural interpretation in the present framework. The denominator of the mapping $B(\beta;\pi;\gamma)$ depends on the effective relevance of the corrected instrument. When the identifying variation becomes small, the derivative approaches unity in magnitude and the mapping becomes nearly unstable.

This behavior parallels the weak-instrument problem in IV estimation, where estimators become highly sensitive to sampling noise when the identifying variation approaches zero.

4.3 Local Existence and Uniqueness

We now establish that the estimator is well defined under a local contraction condition. Let $D \subset \mathbb{R}$ be the parameter space for β, and $\Pi \subset \mathbb{R}$ be the admissible correction set.

Assumption 4.1 (Regularity of the mapping): The mapping $g(\beta,\pi;\gamma)$ is continuously differentiable in a neighborhood of the true parameter β^*.

Assumption 4.2 (Well-Defined Mapping): there exists $c_h > 0$ such that $\inf_{\beta\in D,\pi\in\Pi} |B(\beta;\pi;\gamma)| \geq c_h$ and similarly for the sample analogue with probability approaching one.

Theorem 4.1 (Local Existence and Uniqueness of the Fixed Point)

Suppose the mapping $g(\beta;\pi;\gamma)$ is continuously differentiable and locally contractive in a neighborhood of β^*, i.e. $|g'(\ \beta^*;\pi;\gamma)| < 1$. Then the fixed point exists and is unique in that neighborhood.

Proof (sketch).

The result follows from the contraction mapping theorem. When the derivative of the mapping is bounded by a constant strictly smaller than one in a neighborhood of the fixed point, the mapping is locally contractive and therefore admits a unique fixed point. Full details are provided in Appendix A.

4.4 Stability, Identification and Admissibility

The fixed-point estimator can be interpreted as trading instrument relevance for improved exogeneity through a residual-based correction. The magnitude of this correction is governed by a scalar parameter λ, which determines how much of the structural residual is subtracted from the instrument.

The stability of the mapping determines how much relevance can be sacrificed before identification breaks down. This subsection provides primitive conditions under which the mapping is well-defined and shows that stability induces an endogenous bound on admissible instrument invalidity.

Primitive conditions

Assumption 4.3 (baseline relevance, signal): $\mathbb{E}[zx] \neq 0$

Assumption 4.4 (Finite moments): $\mathbb{E}[z^2], \mathbb{E}[x^2], \mathbb{E}[u^2] < \infty$,

Assumption 4.5 (Residual informativeness): $Var(u) > 0$

Assumption 4.6 (bounded invalidity, noise): $|\mathbb{E}[zu]| < \infty$

Assumption 4.7 (interior admissible region): there exists a neighborhood $\mathcal{N}(\beta_0)$ such that $|\mathbb{E}[z^{FP}(\beta, \lambda)x]| > 0\ \forall \beta \in \mathcal{N}(\beta_0)$

Assumption 4.8 (smoothness): $g(\beta; \pi; \gamma)$ is continuously differentiable near β_0.

Proposition 4.1 (Stability, Identification and Admissibility)

Suppose Assumptions 4.3–4.8 hold. Then local stability of the FP-GMM mapping imposes an endogenous admissibility restriction on instrument invalidity: only those correction levels that preserve sufficient effective relevance are compatible with a locally well-defined and identified estimator. Increasing the correction improves exogeneity but weakens relevance, so admissibility is governed by a relevance–exogeneity trade-off internal to the mapping. As the stability margin shrinks, effective relevance vanishes and local identification breaks down at the boundary of the admissible region.

The correction (6) reduces contamination (improves exogeneity) but also attenuates signal (relevance). Stability requires that signal remains sufficiently large relative to noise after correction. Hence, the admissible set of invalidity is endogenously determined by the requirement that the corrected instrument retains a minimum signal-to-noise ratio. This yields an implicit identification region: $\mathcal{U} = \{\mathbb{E}[zu]: |g'(\beta)| < 1\}$ which replaces externally imposed bounds with a stability-determined constraint. See Appendix A.

Definition 4.1 (Admissible Stability Region)

For a given correction parameter π, define the stability region as:

$$\mathcal{A} \equiv \{\pi \in \Pi: |g'(\beta(\pi); \pi; \gamma)| < 1\} \tag{14}$$

Within this region the mapping is locally contractive and the fixed point is well defined.

The boundary of the stability region is characterized by $\partial\mathcal{A} \equiv \{\pi \in \Pi: |g'(\beta(\pi); \pi; \gamma)| = 1\}$. At the boundary $|g'(\beta(\pi); \pi; \gamma)| = 1$ the mapping loses local stability and the fixed-point iteration becomes highly sensitive to perturbations.

Theorem 4.2 (Local Identification Through Stability)

Let $\beta^*(\lambda)$ denote the population fixed point of the FP-GMM mapping $g(\beta; \lambda)$, and let the admissible region be defined by (14). Suppose Assumptions 4.3–4.8 hold, and suppose the effective relevance term $B(\beta, \lambda)$ is continuously differentiable and nondegenerate in a neighborhood of $(\beta^*(\lambda), \lambda)$.

Then, for any $\lambda \in \mathcal{A}$, the following are equivalent in a neighborhood of the fixed point:

1. The mapping is locally stable, i.e. $|g'(\beta^*(\lambda); \lambda)| < 1 - \varepsilon$

2. The population fixed point is locally unique and the structural parameter is locally identified
3. The corrected instrument retains nondegenerate effective relevance, so that $|B(\beta^*(\lambda), \lambda)| > 0$

Moreover, as λ approaches the boundary of $\mathcal{A}$, the effective relevance term approaches zero, the derivative approaches unity in magnitude, and local identification breaks down. Thus, the stability frontier coincides with the identification frontier.

The theorem formalizes the central role of stability in the parameter-dependent moment framework. Local stability is not only a computational property of the fixed-point iteration; it is equivalent to local identification of the structural parameter because both require the corrected instrument to retain enough effective relevance after the residual-based adjustment. In this sense, the admissible region is simultaneously a stability region and a local identification region. At the boundary, the correction becomes too strong relative to the remaining signal, the mapping loses contraction, and the parameter becomes weakly or non-locally identified. See Appendix A.

Corollary 4.1 (Branchwise Global Identification)

Let $\mathcal{A}_j \subset \mathcal{A}$ be a connected component of the admissible stability region, and let $\beta(\lambda)$ denote the population fixed-point solution of the FP-GMM mapping for $\lambda \in \mathcal{A}_j$. Suppose that:

(i) for every $\lambda \in \mathcal{A}_j$, the mapping admits a locally unique and stable fixed point;

(ii) the solution path $\beta(\lambda)$ is continuous on $\mathcal{A}_j$;

(iii) the derivative $\partial\beta(\lambda)/\partial\lambda$ preserves its sign on $\mathcal{A}_j$, so that the path is monotone (see Appendix D).

Then, conditional on the branch $\mathcal{A}_j$, the mapping $\lambda \mapsto \beta(\lambda)$ is globally ordered and one-to-one on $\mathcal{A}_j$. In particular, each admissible correction $\lambda \in \mathcal{A}_j$ corresponds to a unique value of the structural parameter, so identification holds globally along the branch.

Proof (sketch)

Local uniqueness follows from Theorem 4.2. Continuity and the sign-preserving derivative imply that $\beta(\lambda)$ is monotone on $\mathcal{A}_j$ (Appendix D). A continuous monotone function on a connected set is one-to-one, which establishes global identification along the branch.

4.5 Primitive Characterization of the Admissible Region

Proposition 4.2 (Primitive Characterization of the Admissible Region)

Let $\beta^*(\lambda)$ denote the population fixed-point associated with the FP-GMM mapping, and let the exact admissible region be defined by (14). Suppose there exists constants $\underline{r} > 0$, $\bar{e} < \infty$ and $\bar{d} < \infty$ such that, for all β in a neighborhood of $\beta^*(\lambda)$ we have $B(\beta, \lambda) \geq \underline{r} - |\lambda|\bar{d}$, $|\mathbb{E}[xu(\beta)]| \leq \bar{e}$.

Where $B(\beta, \lambda)$ denotes the effective relevance, $\underline{r}$ baseline relevance, $\bar{e}$ residual endogeneity and $\bar{d}$ correction-induced attenuation. Then a sufficient condition for local stability is

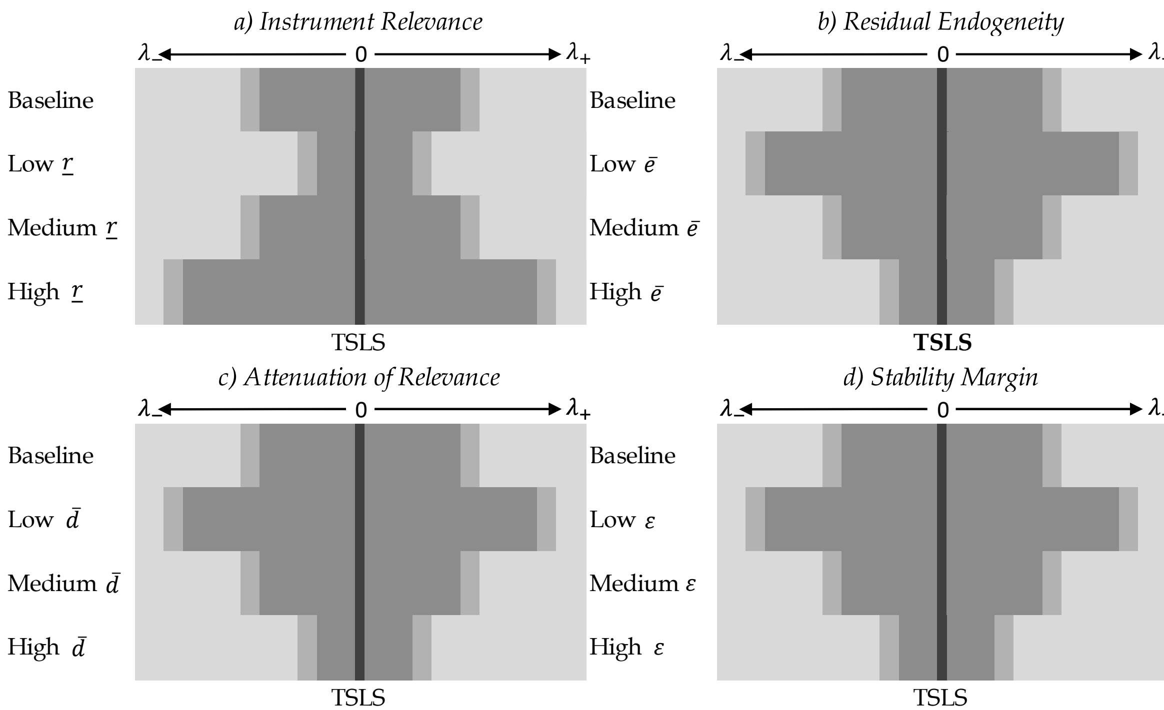
$$|\lambda| < \bar{\lambda}_{prim}(\varepsilon) := \frac{(1-\varepsilon)\underline{r}}{\bar{e}+(1-\varepsilon)\bar{d}} \tag{15}$$

For $\varepsilon \in (0,1)$. The interval $\mathcal{A}_{prim}(\varepsilon) = \left[-\bar{\lambda}_{prim}(\varepsilon), \bar{\lambda}_{prim}(\varepsilon)\right]$ satisfies $\mathcal{A}_{prim}(\varepsilon) \subseteq \mathcal{A}$.

Figure 1 provides a primitive comparative-statics representation of admissibility. While Definition 4.1 characterizes the admissible region exactly through the derivative of the fixed-point mapping, Proposition 4.2 yields a sufficient bound expressed in terms of economically interpretable primitives.

In the scalar case, admissibility takes the form of an interval on the λ-axis. This interval expands with baseline relevance $\underline{r}$, and contracts with residual endogeneity $\bar{e}$, correction-induced attenuation $\bar{d}$, and tighter stability margins ε. The figure illustrates these comparative statics, making clear that admissibility is determined by the signal structure of the model rather than by arbitrary tuning choices. See Appendix A.

Figure 1 The shaded interval represents a stylized primitive bound for the admissible region of λ. Stronger baseline relevance expands the interval, whereas larger residual endogeneity and faster attenuation of relevance contract it.



Baseline relevance. Higher $\underline{r}$ widens the admissible interval because the corrected instrument can sustain larger reductions in relevance before stability is lost. (b) Residual endogeneity. Higher $\bar{e}$ narrows the admissible interval because the correction amplifies instability more rapidly when residual endogeneity is larger. Panel (c) Attenuation of relevance. Higher $\bar{d}$ narrows the admissible interval because each unit of correction erodes effective instrument strength more quickly. Panel (d) Stability margin. Higher ε narrows the admissible interval by definition.

Thus, stability delivers local identification of the structural parameter only within the admissible region, while the correction parameter remains disciplined by stability rather than point-identified from the data.

Theorem 4.2 also clarifies how the present framework relates to Hansen (1982). In standard GMM, identification is studied through fixed moment restrictions and their Jacobian. Here, that perspective is extended to parameter-dependent moments generated by a residual-based correction, where local stability of the induced mapping becomes equivalent to local identification.

Proposition 4.3 (Stability-Constrained Correction Under Misspecified Moments)

Consider the parameter-dependent moment condition defined in (7), where the baseline instrument may violate the exclusion restriction. Suppose that the fixed-point mapping admits a locally stable solution for some $\lambda \in \mathcal{A}$.

Then, the residual-based correction induces an endogenous adjustment of the moment condition such that the structural parameter is locally identified within the admissible region. In this sense, misspecification arising from imperfect exogeneity is not eliminated, but disciplined through a stability-constrained correction.

The proposition highlights that the role of stability is not merely to ensure convergence, but to discipline the extent to which misspecification can be corrected. Rather than requiring valid instruments, the framework allows for imperfect exogeneity and determines the admissible correction endogenously through the fixed-point mapping.

Theorem 4.3 (Stability and the Relevance–Exogeneity Trade-off).

Let $\beta^*(\lambda)$ denote the population fixed point of the FP-GMM mapping $g(\beta;\lambda)$, and define the admissible region (14). Suppose:

(i) the correction is residual-based, so that $z^{FP}(\beta;\lambda)$ depends on $u(\beta)$;

(ii) baseline relevance is nondegenerate, so that $\mathbb{E}[\hat{x}x] > 0$

(iii) the mapping $g(\beta;\lambda)$ is continuously differentiable. Then:

1. Increasing λ generates a trade-off between reduced contamination and reduced effective relevance of the corrected instrument;
2. The admissible set of corrections is determined endogenously by the stability condition $|g'(\beta^*(\lambda);\lambda)| < 1'$
3. The same object $g'(\beta^*(\lambda);\lambda)$;$\lambda$) jointly governs:
    i. existence and local uniqueness of the estimator,
    ii. convergence of the fixed-point iteration,
    iii. sensitivity of the estimator, and
    iv. the admissible magnitude of the correction.

Hence, the fixed-point mapping provides a unified internal representation of the relevance–exogeneity trade-off, with stability determining how much relevance can be sacrificed to improve exogeneity. The parameter-dependent moment framework organizes this trade-off internally. The admissible magnitude of the correction is determined by the stability of a parameter-dependent mapping rather than by exogenous restrictions.

4.6 Convergence Rate

The contraction property also determines the rate at which the fixed-point iteration converges.

Corollary 4.2 (Convergence rate).

Under the conditions of Theorem 4.1, the fixed-point iteration satisfies for some $q < 1$

$$\left|\beta_j - \beta^*\right| \leq q^j |\beta_0 - \beta^*| \tag{16}$$

Under the contraction condition, the fixed-point iteration converges geometrically at rate $q = |g'(\beta^*; \pi; \gamma)|$. When the derivative approaches unity, convergence becomes slower, and the estimator becomes increasingly sensitive to sampling variation. Proofs are provided in Appendix A.

In the multivariate case, the scalar derivative condition is replaced by a Jacobian stability condition, so admissibility is governed by the spectral radius of the fixed-point mapping rather than by a one-dimensional contraction coefficient. Appendix A provides this extension and shows that the same relevance-exogeneity logic carries over, although the admissible region is no longer characterized by a monotone scalar path.

5. Consistency and Dominance of the Stability-Based Estimator

In this section we establish consistency of the fixed-point estimator $\hat{\beta}(\pi)$, defined as the solution to $\hat{\beta}(\pi) = g_n(\hat{\beta}(\pi); \pi; \hat{\gamma})$.

5.1 Population Fixed-Point

Let $\psi(\beta; \pi; \gamma)$ denote the population moment defined (7). The population parameter $\beta^*(\pi)$ is defined as the solution satisfying $\psi(\beta^*(\pi); \pi; \gamma) = 0$. Equivalently, $\beta^*(\pi)$ is the fixed point of the mapping $g(\beta; \pi; \gamma)$.

Assumption 5.1 (Compact parameter space): $D \subset \mathbb{R}$ is compact and contains $\beta(\pi)$.

Assumption 5.2 (Uniform Law of Large Numbers): $\sup_{\beta \in D} |\psi_n(\beta; \pi; \hat{\gamma}) - \psi(\beta; \pi; \gamma)| \rightarrow_p 0$

Assumption 5.3 (*Local Contraction*): there exists $\beta^*(\pi) \in D$ such that $|g'(\beta^*(\pi); \pi; \gamma)| < 1$.

Assumption 5.4 (*Non-degeneracy of the Denominator*): there exists $c_h > 0$ such that $|B(\beta; \pi; \gamma)| \geq c_h$ in a neighborhood of $\beta^*(\pi)$.

Theorem 5.1 (Consistency)

Under assumptions 5.1-5.4

$$\hat{\beta}(\pi) \to_p \beta^*(\pi) \quad (17)$$

As $n \to \infty$.

Proof (sketch)

Uniform convergence of the sample mapping to the population mapping, together with the existence of a locally unique fixed point and the contraction property of the mapping, imply convergence of the sample fixed point to the population fixed point. The full proof is given in Appendix B.

5.2 Asymptotic MSE Dominance over k-class Estimators

Conventional k-class estimator remains asymptotically biased then the exclusion restriction fails. More generally, let $\hat{\beta}_{k_n}$denote a k-class estimator generated by a sequence $k_n \to k_0$, under fixed-K, strong-instruments asymptotics. In the residualized scalar model considered here, its probability limit can be written as $plim\ \hat{\beta}_{k_n} = \beta_0 + B(k_0)$ where $B(k_0)$ captures the asymptotic distortion induced by instrument invalidity and, when $k_0 \neq 1$, an additional residual endogeneity component.

By contrast, under the contraction conditions ensuring existence and consistency of the fixed-point estimator, FP-GMM is consistent when the correction parameter equals its population value. Hence,

$$plim\ \hat{\beta}_{FP} = \beta_0 \quad (18)$$

The implication is immediate: whenever $B(k_0) \neq 0$, k-class estimators retain a non-vanishing asymptotic bias, while FP-GMM does not. Therefore, although FP-GMM may exhibit larger variance in small samples, its asymptotic mean squared error eventually becomes smaller than that of the corresponding k-class estimator.

This result applies not only to TSLS, but also to conventional k-class sequences such as LIML and Fuller under the standard fixed-K, strong-instrument asymptotic framework. Proofs are provided in Appendix B.

Theorem 5.2 (Asymptotic MSE dominance of FP-GMM over k-class estimators)

Suppose the instrument is imperfect, and let $\hat{\beta}_{k_n}$ be a k-class estimator such that $k_n \to k_0$ under fixed-K, strong-instrument asymptotics. Assume $plim\ \hat{\beta}_{k_n} = \beta_0 + B(k_0), \sqrt{n}\left(\hat{\beta}_{k_n} - \beta_0 - B(k_0)\right) \to_d N\big(0, V_k(k_0)\big)$ and $\sqrt{n}\big(\hat{\beta}_{FP} - \beta_0\big) \to_d N(0, V_{FP})$.

Then $MSE\big(\hat{\beta}_{k_n}\big) = B(k_0)^2 + V_k(k_0)/n + o(n^{-1})$, whereas $MSE\big(\hat{\beta}_{FP}\big) = V_{FP}/n + o(n^{-1})$. Consequently, if $B(k_0) \neq 0$, there exists a finite threshold $n^*(k_0)$ such that for all sufficiently large $n > n^*(k_0)$, we have

$$MSE\big(\hat{\beta}_{FP}\big) < MSE\big(\hat{\beta}_{k_n}\big) \quad (19)$$

A particularly relevant special case is given by conventional k-class sequences satisfying $k \to 1$, including TSLS, LIML, and Fuller. In that case, all such estimators share the same first-order asymptotic bias under instrument invalidity, while FP-GMM remains asymptotically unbiased.

The dominance result therefore reflects a bias–variance trade-off: while the stability-based estimator may exhibit higher variance in small samples, its asymptotic unbiasedness dominates the persistent bias of k-class estimators when instruments are invalid.

Remark 5.1

If the instrument is exogenous ($\mathbb{E}[zu] = 0$), the optimal correction parameter satisfies $\pi = 0$. The corrected instrument then reduces to the standard first-stage projection, and the fixed-point estimator coincides with k-class estimators. In this case both estimators share identical asymptotic properties.

Remark 5.2

Under instrument invalidity, k-class estimators remain misspecified, as they treat all first-stage relevance as valid and therefore retain non-vanishing bias. By contrast, FP-GMM introduces a correction parameter λ that discounts contaminated relevance, so its bias properties are inherently conditional on the chosen correction.

The dominance result should therefore be interpreted along the admissible correction path. As shown in Section 7, when instrument relevance is sufficiently strong, there exist admissible corrections that substantially reduce bias while preserving identification. In such cases, the boundary-selected stability-based estimator can achieve lower bias—and potentially lower mean squared error—than k-class estimators. The gains arise precisely in settings with strong but imperfectly exogenous instruments.

6. Asymptotic Normality

This section derives the asymptotic distribution of the fixed-point GMM estimator for a fixed regular correction parameter π using standard Z-estimator arguments (Newey and McFadden, 1994; van der Vaart, 1998). Appendix C further shows that the same limit theory extends to correction parameters selected at an operational boundary of the admissible region, provided that the selected endpoint remains bounded away from the exact instability frontier

6.1 Assumptions

We impose standard regularity conditions ensuring asymptotic normality of Z-estimators

Assumption 6.1 (Central Limit Theorem at the true fixed point): $\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\psi_i(\beta(\pi);\pi;\gamma) \Rightarrow N\big(0,\Omega(\pi)\big)$ for some finite $\Omega(\pi) > 0$.

Assumption 6.2 (Smoothness and non-degeneracy):

i. $\psi_i(\beta;\pi;\gamma)$ is continuously differentiable in β in a neighborhood of $\beta(\pi)$.

ii. The Jacobian $\mathcal{K}(\pi) \equiv \frac{\partial}{\partial\beta}\psi(\beta;\pi;\gamma)|_{\beta=\beta(\pi)}$ satisfies $\mathcal{K}(\pi) \neq 0$.

Assumption 6.3 (Consistency): $\hat{\beta}(\pi) \to_p \beta(\pi)$. This follows theorem 5.1.

6.2 Asymptotic Normality

Theorem 6.1 (Asymptotic normality for fixed π): Under assumptions 6.1-6.3

$$\sqrt{n}\left(\hat{\beta}(\pi) - \beta(\pi)\right) \Rightarrow N\left(0, \frac{\Omega(\pi)}{\mathcal{K}(\pi)^2}\right) \tag{20}$$

Moreover, the asymptotic variance admits the representation:

$$Avar\left(\hat{\beta}(\pi)\right) = \frac{\Omega(\pi)}{B(\beta(\pi);\pi;\gamma)^2\left(1-g'(\beta(\pi);\pi;\gamma)\right)^2} \tag{21}$$

Where $B(\beta;\pi) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)x]$, $g'(\beta;\pi;\gamma) = \partial g(\beta;\pi;\gamma)/\partial\beta|_{\beta=\beta(\pi)}$, $\Omega(\pi) = \mathbb{E}[(\hat{x}_i u_i^*)^2] - 2\pi\mathbb{E}[\hat{x}_i(u_i^*)^3] + \pi^2\mathbb{E}[(u_i^*)^4]$ and $u_i^* = u_i(\beta(\pi))$. A full proof and explicit expressions for $\Omega(\pi)$ are provided in Appendix C.

6.3 Feasible Variance Estimation

A consistent plug-in estimator of $\Omega(\pi)$ is

$$\hat{\Omega}(\pi) = \frac{1}{n}\sum_{i=1}^{n}\psi_i\left(\hat{\beta};\pi;\hat{\gamma}\right)^2 \tag{22}$$

And feasible standard error is obtained by replacing $\Omega(\pi)$ and $\mathcal{K}(\pi)$ by their sample's analogues (Appendix C).

Proposition 6.1 (Variance and stability of the fixed point)

The asymptotic variance of the fixed-point estimator is inversely related to the stability margin of the mapping. As $|g'(\beta(\pi);\pi;\gamma)| \to 1$, the denominator of the asymptotic variance shrinks and the estimator becomes increasingly variable.

The implicit function theorem further shows that sensitivity of the estimator with respect to the correction parameter increases as the mapping approaches instability (Appendix D).

Estimators operating near the stability boundary therefore exhibit both higher variance and greater sensitivity to the correction parameter. This reflects the fundamental bias–variance trade-off underlying the parameter-dependent moment framework. For this reason, subsequent inference is stated for operational boundaries satisfying a strict stability margin, rather than for the exact frontier at which the mapping loses local regularity.

7. Stability-Based Calibration of the Correction Path

For each value of the correction parameter π, let $\beta(\pi)$ denote the population fixed point satisfying

$$\beta(\pi) = g(\beta(\pi); \pi; \gamma), \quad equivalently \quad \psi(\beta(\pi), \pi; \gamma) = 0 \tag{23}$$

After projection onto the fitted first-stage index, the stability-based estimator is effectively just-identified in β, even when the first stage is built from multiple excluded instruments. The moment condition determines β conditional on π, but does not separately point-identify π. Accordingly, π is interpreted as a regularization parameter indexing a family of corrected estimators.

7.1 A Disciplined Calibration Problem

Section 4 shows that the admissible region is determined by local stability of the mapping. Section 7 takes that result as given and asks how the correction path should be interpreted and calibrated within that region. The relevant object is the family $\{\beta(\pi): \pi \in \mathcal{A}\}$, which traces the evolution of the estimator as the degree of correction varies over admissible values. The slope of this path is governed by the same derivative object that determines stability

$$g'(\beta; \pi; \gamma) = \frac{\pi\mathbb{E}[xu(\beta(\pi))]}{B(\beta(\pi); \pi; \gamma)} \tag{24}$$

which summarizes how the estimator responds to changes in the correction parameter. Rather than re-establishing admissibility, the role of this derivative here is to organize the correction path within the stable region and to motivate the calibration rules developed below.

7.2 Normalized Correction

The correction path becomes more transparent under the normalization $\lambda = \pi/\gamma$. Using this parametrization, $z_i^{FP}(\beta; \pi; \gamma) = \gamma(z_i - \lambda u_i(\beta))$, so that λ measures the correction relative to first-stage strength.

This separates two distinct objects that are confounded in π: instrument invalidity and instrument relevance. The parameter λ can therefore be interpreted as a normalized correction parameter, capturing how much instrument relevance is sacrificed to reduce endogeneity.

Because the correction parameter is not separately identified from the data, λ and should be interpreted as a regularization parameter indexing a family of admissible estimators rather than as a structural object of interest. The results in Section 4 imply that this regularization cannot be chosen freely: admissible values must lie within the stability region, which coincides with the set of locally identified solutions.

This interpretation is especially useful near the boundary. Values of λ beyond the admissible region correspond to over-regularization: the correction removes so much identifying variation that effective relevance collapses and the structural parameter is no longer locally identified. Boundary selection can therefore be understood as choosing the largest admissible degree of regularization consistent with identification.

The normalized parameter $\lambda = \pi/\gamma$ admits a natural interpretation in terms of instrument invalidity relative to relevance. Since π reflects the covariance between the instrument and the structural error, and γ captures first-stage relevance, λ can be interpreted as a normalized measure of exclusion violation per unit of identifying strength.

The sign of λ depends on the relative signs of π and γ: when both share the same sign, the correction is positive, whereas opposite signs imply a negative correction. This reflects whether the exclusion violation reinforces or offsets the first-stage relationship. While λ is related to the endogeneity of the endogenous regressor, it does not directly correspond to $Cov(x,\text{u})$, but rather to a rescaled measure of instrument-induced endogeneity.

7.3 Monotonicity and Direction of Correction

The stability region not only restricts admissible values of π, but also organizes the correction path. Let $\mathcal{A}_0 \subseteq \mathcal{A}$ be a connected component. Under mild conditions (see Appendix D), the mapping

$$\pi \mapsto \beta(\pi) \tag{25}$$

is continuous and monotone on $\mathcal{A}_0$. Monotonicity is a defining property of the correction path: it ensures that the estimator moves in a single direction as the correction increases, ruling out arbitrary fluctuations and yielding an ordered sequence of progressively stronger corrections. The sign of the slope $\beta'(\pi)$ determines the direction of this movement and provides a transparent characterization of how the estimator responds to changes in the correction parameter.

Importantly, monotonicity does not determine which point along the path should be selected. Instead, it provides a structured, one-dimensional path along which different degrees of correction can be evaluated.

The boundary represents a limiting configuration, as π approaches this boundary, the path becomes increasingly steep: small changes in π induce large changes in $\beta(\pi)$, reflecting the rapid loss of effective identifying strength.

Theorem 7.1 (MSE-optimal boundary under stability constrains)

Let $\beta^*(\lambda)$ denote the population fixed point of the FP-GMM mapping, and let $\hat{\beta}(\lambda)$ be its sample analogue. Let the admissible region be defined as $\mathcal{A} = [\lambda_-, \lambda_+] = \{\lambda: |g'(\beta^*(\lambda); \lambda)| < 1 - \varepsilon\}$, $\varepsilon \in (0,1)$. Suppose:

(i) Endogenous instruments: $\mathbb{E}[zu] \neq 0$, so that the baseline estimator at $\lambda = 0$ is biased.

(ii) Monotone bias reduction: There exists a direction $s \in \{-1, +1\}$ such that along the admissible branch $\mathcal{A}_s \subseteq \mathcal{A}$ for which $d(\beta^*(\lambda) - \beta_0)^2/d\lambda < 0$

(iii) Local variance growth near instability: The asymptotic variance $V(\lambda)$ is weakly increasing as λ approaches the boundary of $\mathcal{A}_s$

Then, for sufficiently large n, the mean squared error $MSE(\lambda) = (\beta^*(\lambda) - \beta_0)^2 + n^{-1}V(\lambda)$ is minimized at the boundary of the admissible region in the relevant direction: $\lambda^{MSE} \in \{\lambda_-, \lambda_+\}$. In particular, when bias reduction dominates variance growth along the admissible path, the optimal correction corresponds to the largest admissible correction consistent with stability.

The theorem provides a formal justification for the boundary selection rule. Under endogenous instruments, increasing the correction parameter reduces bias but weakens effective instrument strength. Stability imposes a constraint on how far this correction can be taken. When bias decreases monotonically along the admissible path, the boundary represents the maximal bias-reducing correction that remains consistent with identification. The selection rule therefore admits a risk-based interpretation: it corresponds to minimizing mean squared error subject to stability of the fixed-point mapping.

Proposition 7.1 (Boundary selection and k-class dominance)

Under the conditions of Theorem 7.1 and assuming instrument invalidity, the boundary-selected stability-based estimator minimizes mean squared error within the admissible region and weakly dominates k-class estimators whenever the reduction in squared bias along the admissible path exceeds the variance differential.

In particular, when k-class estimators exhibit non-vanishing bias and the admissible correction sufficiently reduces bias, the boundary stability-based estimator achieves strictly lower mean squared error.

The boundary rule therefore does not only solve a constrained optimization problem within FP-GMM, but also selects the point at which the estimator's advantage over conventional IV methods is most pronounced.

7.4 Boundary Calibration and Researcher Choice

Because the correction parameter is not identified from the data, its selection requires an external criterion. The stability condition restricts admissible values of π, but does not determine a unique choice within that region.

The monotone structure of the correction path implies that admissible estimators are naturally ordered. The relevant choice is therefore between boundary points of the admissible region, corresponding to maximal corrections in each direction.

The direction of the correction—i.e., the sign of λ—depends on economic reasoning, institutional knowledge, or prior information about the nature of the violation of the exclusion restriction. For example, if the instrument is believed to be positively correlated with the structural error, the appropriate correction corresponds to one side of the admissible region; if the correlation is negative, the opposite boundary is relevant.

While the sign of the exclusion violation is not identified from the structural moment itself, auxiliary data-based diagnostics may help discipline branch selection in applications. Such diagnostics do not eliminate the need for substantive judgment, but they can reduce reliance on purely untestable sign restrictions and make the choice of correction direction more empirically informed. Appendix D discusses several such diagnostics, including placebo outcomes, covariate imbalance, reduced-form decompositions, and subsample comparisons.

Thus, the estimator should be interpreted as selecting a boundary of the admissible region consistent with a given direction of correction, rather than as estimating a correction parameter from the data. Without credible sign information, the method should be interpreted as disciplined sensitivity analysis rather than a uniquely justified point estimator.

If that direction is misspecified, the selected boundary may no longer correspond to a bias-reducing correction. In that case, the procedure remains informative as a structured sensitivity analysis over admissible corrections, but its interpretation as a disciplined point-estimation rule is weakened.

7.5 Sample Implementation

This subsection describes a practical algorithm for selecting the correction parameter π. The procedure exploits the fact that admissible values of π must satisfy two conditions: (i) the fixed-point iteration must converge numerically, and (ii) the associated mapping must remain locally stable. The algorithm therefore searches over a grid of candidate values and retains those for which the fixed-point estimator exists and satisfies the stability condition.

Algorithm 7.1 (Stable-path calibration)

Input: grid $\Lambda = \{\pi_1/\hat{\gamma}, \dots, \pi_M/\hat{\gamma}\}$, tolerance levels $\varepsilon_\beta > 0$ and $\varepsilon_{g'} > 0$, maximum iterations J_{max}.

1. **Initialization:** set the starting values equation to the conventional IV estimator $\beta_0 = \hat{\beta}_{TSLS}$. For this iteration $\pi_0 = 0$.
2. **Fixed-point iteration:** for each $\pi \in \Pi$ generate the sequence $\beta_{j+1} = g_n(\beta_j, \pi; \hat{\gamma}),\ j = 0, \dots, J_{max} - 1$ by using (9), (6) and the residuals (5) in the previous iteration. Declare convergence if for some $j \le J_{max}$ $|\beta_j(\pi) - \beta_{j-1}(\pi)| \le \varepsilon_\beta$. Denote the resulting estimate by $\hat{\beta}(\pi)$.
3. **Local stability evaluation:** for each convergent sequence compute the sample contraction measure $\hat{q}(\pi) = |g_n'(\hat{\beta}(\pi), \pi; \hat{\gamma})|$ where $\hat{q}(\pi)$ denotes the estimates contraction coefficient (see Appendix A6).
4. **Admissible parameter set:** construct the set $\hat{\mathcal{A}} = \{\pi \in \Pi\text{: iteration converges and } \hat{q}(\pi) \le \varepsilon_{g'}\}$
5. **Boundary selection**: If theory, institutional knowledge, or sign restrictions imply a direction of bias correction, choose the corresponding endpoint of $\hat{\mathcal{A}}$. Let $\hat{\pi}_-/\hat{\gamma} = \min \hat{\mathcal{A}}$, $\hat{\pi}_+/\hat{\gamma} = \max \hat{\mathcal{A}}$.
6. **Final estimator**: The stability-based estimator is

$$\hat{\beta} = \hat{\beta}(\hat{\pi}_\pm/\hat{\gamma})$$

Remark (Sample implementation with multiple instruments)

When the first stage is generated by multiple excluded instruments $Z = (z_1, \dots, z_L)$, an unrestricted grid search over $\pi = (\pi_1, \dots, \pi_L)'$ is generally infeasible and destroys the ordered one-dimensional correction path used in Algorithm 7.1. A practical scalar implementation is obtained by imposing a proportionality restriction $\pi = \lambda\gamma$ where γ is the first-stage coefficient vector. Under this restriction, each instrument is corrected in proportion to its contribution to first-stage relevance, and the induced corrected fitted index becomes $x^M(\beta; \lambda) = \hat{x} - u(\beta)\lambda\,\gamma'\gamma$.

Hence the relevant grid search may be conducted over the scalar normalized parameter $\boldsymbol{\lambda} = \pi'\gamma/\gamma'\gamma$ or, equivalently, over the aggregate correction $\tilde{\lambda} = \pi'\gamma$. This preserves the one-dimensional admissible path and keeps the implementation consistent with the stability-based calibration developed in the scalar case.

Remark (FWL implementation).

When additional exogenous controls are included, the stability-based estimator can be implemented by residualizing all variables with respect to the control set using the Frisch–Waugh–Lovell theorem. The fixed-point procedure is then applied to the residualized variables. This ensures that identification and the stability-based admissible region are determined by variation orthogonal to the controls, without altering the structure of the estimator.

Economically, this means that the relevance–exogeneity trade-off is evaluated conditional on observed covariates, so that the correction operates only on the residual variation that remains after controlling for observable confounders.

8. Relation to Existing Estimators

This section situates FP-GMM relative to existing instrumental-variables estimators. The comparison is useful because the proposed estimator shares some features with standard IV procedures, k-class methods, control-function approaches, and imperfect-instrument frameworks, yet differs from each of them in the way the correction is constructed and disciplined. The purpose of the section is therefore not to claim equivalence, but to clarify where the estimator nests familiar procedures, where it departs from them, and what is gained by the stability-based formulation.

8.1 Relation to Conventional IV Estimators

When the correction parameter satisfies $\pi = 0$, the corrected instrument reduces to the original instrument and the moment condition becomes $\mathbb{E}[z_i u_i] = 0$. In this case, the stability-based estimator coincides with the conventional IV estimator. In particular, the sample solution reduces to standard two-stage least squares (TSLS). FP-GMM can therefore be interpreted as a generalization of IV estimation in which the instrument is adjusted using information from the structural residual.

This nesting is important for the economic interpretation of the method. At $\pi = 0$, all first-stage relevance is treated as valid identifying variation. Moving away from zero introduces a correction that discounts part of that relevance in order to improve exogeneity. FP-GMM therefore preserves the classical IV estimator as the zero-correction benchmark, while extending it to settings in which exact exclusion is considered implausible but instrument relevance remains informative.

8.2 Relation to k-Class Estimators

The k-class family modifies the projection of the endogenous regressor onto the instrument space through a scalar weighting parameter. Estimators such as TSLS, LIML, and Fuller all belong to this framework. FP-GMM differs in the object being adjusted: rather than altering the projection step, it corrects the instrument itself through a residual-based term that depends on the structural parameter. This produces parameter-dependent moment conditions and a fixed-point representation.

The distinction is substantive rather than merely algebraic. In FP-GMM, the correction traces an admissible path along which contaminated relevance is discounted, with stability determining how far that adjustment can be taken. By contrast, k-class estimators operate within a fixed instrument space and therefore do not provide an internal mechanism for disciplining admissible instrument invalidity. Appendix E shows that FP-GMM can be written in a form closely related to the k-class family, but the relationship is one of proximity rather than equivalence: the correction is data-dependent and determined by a fixed-point condition, not by a fixed weighting parameter.

8.3 Relation to Control-Function IV, Partial Identification and Imperfect-Instruments Frameworks

Control-function methods address endogeneity by augmenting the structural equation with an additional term that captures the correlation between the endogenous regressor and the structural error. FP-GMM is conceptually related, but differs in the object being adjusted: rather than modifying the structural equation, it corrects the instrument through the structural residual, thereby generating parameter-dependent moment conditions and a fixed-point representation. The relevance–exogeneity trade-off is therefore organized through the instrument itself rather than through an additional regressor.

FP-GMM is also related to the literature on imperfect instruments and partial identification, including Conley et al. (2012) and Nevo and Rosen (2012). Those approaches allow exclusion to fail, but typically rely on exogenous bounds or ordering restrictions on the extent of invalidity. By contrast, FP-GMM determines the admissible magnitude of the correction internally through the stability of the fixed-point mapping. The resulting trade-off is therefore represented not by an externally imposed sensitivity range, but by an admissible correction path whose boundary is disciplined by the estimator's own dynamics. Appendix E discusses these connections in more detail.

9. Monte Carlo Evidence

This section evaluates the finite-sample performance of the stability-based estimator using Monte Carlo simulations. The simulations are designed to stress the estimator in environments known to challenge instrumental variables methods.

We compare FP-GMM with conventional estimators including OLS, TSLS, LIML, and Fuller. All results are based on 500 replications and 10 fixed-point iterations.

The baseline simulations focus on a simple linear environment that highlights the main mechanism of the estimator. Additional designs incorporating heterogeneous exclusion violations, nonlinear invalidity, and multiple instruments are reported in Appendix F. The qualitative patterns are robust across these alternative specifications.

In all simulation exercises, the correction parameter is selected at the boundary of the admissible region in the direction of bias reduction. That is, we assume that the researcher correctly assigns the sign of the correction, so that the selected boundary corresponds to the appropriate direction of adjustment. This isolates the performance of the estimator along the stable correction path and abstracts from uncertainty regarding the direction of the violation.

The maintained sign restriction is therefore not merely a convenience. Because the correction path is monotone, selecting the wrong direction moves the estimator away from the bias-reducing branch and can generate substantial distortion. Absent credible prior information on the sign of the violation, FP-GMM should be interpreted primarily as a structured sensitivity device rather than as a uniquely disciplined point estimator

9.1 Linear Gaussian DGP

We begin consider a standard linear Gaussian design with a potentially invalid instrument.

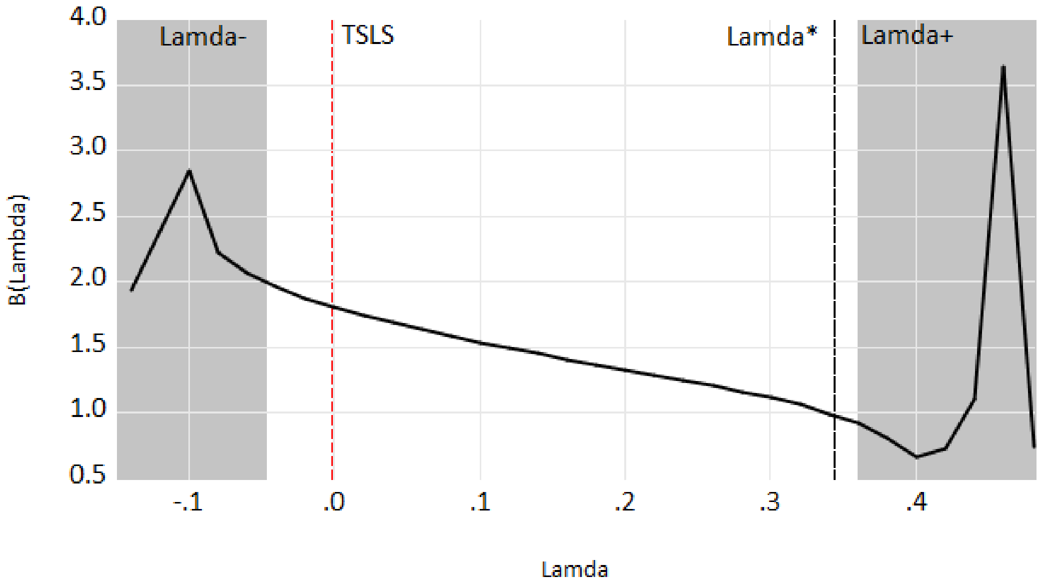
$$y_i = \beta x + u_i, \quad x_i = \gamma z_i + v_i, \quad u_i = \delta z_i + \varepsilon_i, \quad z_i, v_i, \varepsilon_i \sim \mathcal{N}(0,1) \tag{26}$$

The terms z_i, v_i, ε_i are mutually independent. The parameter δ governing the violation of the exclusion restriction determines the degree of instrument invalidity. When the violation is zero the FP estimator coincides with TSLS, while for nonzero values conventional IV estimators become inconsistent.

Figure 2 displays the estimated path $\beta(\lambda)$ based on 500 Monte Carlo replications over values of the correction parameter. The admissible region is defined by convergence of the fixed-point iteration and the local stability condition of the mapping. Over this region, $\beta(\lambda)$ follows a clearly monotone path. The selected estimate is obtained at the positive stability boundary, and this boundary lies very close to the theoretical optimal correction λ^*. In turn, this corresponds to the region where $\beta(\lambda)$ reaches the true structural value $\beta^* = 1$, indicating that the boundary rule delivers an accurate approximation to the optimal correction in this design.

The simulation uses $n = 5000, \beta^* = 1, \gamma = 0.5, \delta = 0.4$, 10 fixed-point iterations, $\varepsilon_{g'} = 0.25$, and $\varepsilon_\beta = 10^{-6}$. The admissible stability region is $\mathcal{A} = [\lambda_-, \lambda_+] = [-0.06, 0.36]$. The black dashed line marks the analytically derived optimal correction λ^*, and the red dashed line denotes the TSLS benchmark at $\lambda = 0$. Appendix F shows that the same qualitative pattern arises across alternative data-generating processes.

Figure 2 - Estimated $\widehat{\beta}(\lambda)$, across correction parameters, illustrating the monotone bias reduction within the stability region and the bias–stability trade-off that characterizes the FP estimator.



The sequence $\beta(\lambda)$ was estimated using 500 and the following parameters: $n = 5000, \beta^* = 1, \gamma = 0.5, \delta = 0.4, FP\ iterations = 10, \varepsilon_{g'} = 0.25, \varepsilon_\beta = 10^{-6}$. The admissible stability region is defined as $\mathcal{A} = [\lambda_-, \lambda_+] = [-0.06, 0.36]$. The black dashed vertical line indicates the theoretical optimal correction, derived analytically for this DGP. The location corresponding to the TSLS estimator (π=0) is also shown.

9.2 Linear Gaussian DGP and Instrument Invalidity

Table 1 reports the resulting bias across estimators. The first experiment considers a linear Gaussian data-generating process with a single instrument. The parameter δ controls the degree of instrument invalidity by introducing correlation between the instrument and the structural error.

Two patterns emerge clearly. First, as the degree of instrument invalidity increases, the bias of all IV estimators rises, reflecting the violation of the exclusion restriction. Second, the stability-based estimator substantially reduces bias relative to conventional IV estimators across all sample sizes. The magnitude of the improvement is relatively stable across n, indicating that the gain arises primarily from the correction mechanism rather than from sample size.

FP-GMM dramatically reduces bias relative to conventional IV estimators. Across the considered values of instrument invalidity, the bias of FP-GMM is between 60% and 90% smaller than TSLS, LIML, and Fuller, while exhibiting similar dispersion. The magnitude of the improvement remains stable across sample sizes, indicating that the gains arise primarily from the residual-based correction rather than from sample size.

Overall, the results illustrate two key properties of the estimator. First, the residual-based correction is effective at reducing the bias generated by moderate violations of the exclusion restriction. Second, when instrument endogeneity becomes severe, no estimator based on the same instrument can fully recover the true parameter, but FP still provides a substantial improvement relative to standard IV methods.

Table 1 – Monte Carlo results: Linear Gaussian DGP with instrument invalidity ($\boldsymbol{\beta = 1, \gamma = 1}$)

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} = 0.25$, $\varepsilon_b = 10^{-6}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$\delta = 0.4$			$\delta = 0.8$			$\delta = 1.2$			$\delta = 2$		
$n = 1000$	*FP-GMM*	-0.03	0.04	0.02	0.21	0.21	0.03	0.42	0.42	0.03	0.81	0.81	0.04
	OLS	0.20	0.20	0.02	0.40	0.40	0.03	0.60	0.60	0.03	1.00	1.00	0.04
	TSLS	0.40	0.40	0.03	0.80	0.80	0.04	1.20	1.20	0.05	2.00	2.00	0.07
	LIML	0.40	0.40	0.03	0.80	0.80	0.04	1.20	1.20	0.05	2.00	2.00	0.07
	Fuller	0.39	0.39	0.03	0.78	0.78	0.04	1.18	1.18	0.05	1.96	1.97	0.07
$n = 5000$	*FP-GMM*	-0.03	0.03	0.01	0.21	0.21	0.01	0.42	0.42	0.01	0.81	0.81	0.02
	OLS	0.20	0.20	0.01	0.40	0.40	0.01	0.60	0.60	0.01	1.00	1.00	0.02
	TSLS	0.40	0.40	0.02	0.80	0.80	0.02	1.20	1.20	0.02	2.00	2.00	0.03
	LIML	0.40	0.40	0.02	0.80	0.80	0.02	1.20	1.20	0.02	2.00	2.00	0.03
	Fuller	0.39	0.39	0.02	0.78	0.78	0.02	1.18	1.18	0.02	1.96	1.96	0.03
$n = 10000$	*FP-GMM*	-0.03	0.03	0.01	0.21	0.21	0.01	0.42	0.42	0.01	0.81	0.81	0.02
	OLS	0.20	0.20	0.01	0.40	0.40	0.01	0.60	0.60	0.01	1.00	1.00	0.01
	TSLS	0.40	0.40	0.01	0.80	0.80	0.01	1.20	1.20	0.02	2.00	2.00	0.02
	LIML	0.40	0.40	0.01	0.80	0.80	0.01	1.20	1.20	0.02	2.00	2.00	0.02
	Fuller	0.39	0.39	0.01	0.78	0.78	0.01	1.18	1.18	0.02	1.96	1.96	0.02

This table reports Monte Carlo results for a linear Gaussian data-generating process with a potentially invalid instrument. The parameter δ governs the degree of instrument invalidity. The table reports bias, RMSE, and standard deviation for the stability-based estimator and conventional estimators including OLS, TSLS, LIML, and Fuller. Results are based on 500 simulations.

9.3 Weak Instrument Design

The second experiment explores the interaction between instrument invalidity and instrument strength. The parameter γ governs the first-stage relationship between the instrument and the endogenous regressor, allowing the design to vary from strong to weak identification.

Table 2 reports the finite-sample bias of the estimators for different values of γ. As the instrument becomes weaker, the performance of all IV estimators deteriorates, reflecting the loss of identifying variation in the first stage.

The stability-based estimator follows the same pattern but generally exhibits lower bias than TSLS and LIML across most configurations. In the weakest-instrument cases Fuller performs slightly better, consistent with its known robustness in weak-IV settings.

When instruments are extremely weak ($\gamma = 0.05$), Fuller performs best, consistent with its well-known robustness in weak-IV settings. However, once moderate first-stage strength is present ($\gamma \geq 0.2$), FP-GMM substantially reduces bias relative to conventional IV estimators, with reductions of approximately 70–100% relative to TSLS and LIML.

When instruments are both weak and invalid, the estimation problem becomes severe for all IV estimators. As the sample size increases, however, the stability-based estimator converges toward the true parameter, reflecting its consistency under instrument invalidity.

Table 2 –Monte Carlo results: weak-IV design ($\boldsymbol{\beta = 1, \delta = 0.4}$)

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} =$	0.25, $\varepsilon_b = 10^{-6}$	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$\gamma = 0.05$			$\gamma = 0.1$			$\gamma = 0.2$			$\gamma = 0.5$		
$n = 1000$	*FP-GMM*	15.63	262.14	261.94	4.21	5.34	3.29	0.60	1.15	0.99	0.00	0.04	0.04
	OLS	0.02	0.04	0.03	0.04	0.05	0.04	0.08	0.08	0.03	0.16	0.16	0.03
	TSLS	15.74	262.17	261.96	4.61	5.48	2.97	2.04	2.07	0.38	0.80	0.80	0.08
	LIML	15.74	262.17	261.96	4.61	5.48	2.97	2.04	2.07	0.38	0.80	0.80	0.08
	FULLER	0.77	0.90	0.46	1.26	1.28	0.23	1.32	1.33	0.13	0.74	0.74	0.07
$n = 5000$	*FP- GMM*	8.96	9.69	3.68	3.68	3.94	1.41	0.39	0.81	0.71	0.00	0.02	0.02
	OLS	0.02	0.03	0.02	0.04	0.04	0.01	0.08	0.08	0.01	0.16	0.16	0.01
	TSLS	9.00	9.70	3.63	4.11	4.15	0.61	2.02	2.03	0.16	0.80	0.80	0.04
	LIML	9.00	9.70	3.63	4.11	4.15	0.61	2.02	2.03	0.16	0.80	0.80	0.04
	FULLER	0.86	0.89	0.20	1.31	1.32	0.09	1.33	1.33	0.06	0.74	0.74	0.03
$n = 10000$	*FP- GMM*	8.19	8.50	2.29	3.69	3.90	1.25	0.38	0.77	0.67	0.00	0.02	0.02
	OLS	0.02	0.02	0.01	0.04	0.04	0.01	0.08	0.08	0.01	0.16	0.16	0.01
	TSLS	8.33	8.56	1.98	4.08	4.10	0.43	2.01	2.01	0.11	0.80	0.80	0.03
	LIML	8.33	8.56	1.98	4.08	4.10	0.43	2.01	2.01	0.11	0.80	0.80	0.03
	FULLER	0.88	0.90	0.14	1.33	1.33	0.06	1.33	1.33	0.04	0.74	0.74	0.02

This table reports Monte Carlo results under a weak-instrument design where the first-stage relevance parameter is reduced. The table reports bias, RMSE, and standard deviation across estimators. Results are based on 500 simulations.

9.4 Bernoulli Switching Invalidity

To introduce non-Gaussian and asymmetric invalidity, we consider a switching design in which the direction of instrument invalidity varies across observations. The structural equations are

$$y_i = \beta x + u_i, \;\; x_i = \gamma z_i + v_i, \;\; u_i = \delta s_i z_i + \varepsilon_i, \quad z_i, v_i, \varepsilon_i \sim \mathcal{N}(0,1), \;\; s_i \sim Bernoulli(p) \qquad (27)$$

The terms $z_i, v_i, \varepsilon_i, s_i$ are mutually independent. The parameter δ determines the magnitude of the violation, while p determines the share of observations for which the instrument in invalid. When $p = 0.5$, the average violation is zero and the bias function becomes symmetric around $\pi = 0$.

Table 3 reports the results. When the instrument is valid for most observations, FP-GMM substantially reduces bias relative to conventional IV estimators. As p approaches 0.5, however, the mixture structure masks the direction of invalidity and the correction becomes less effective.

When invalidity is asymmetric ($p = 0.3$ or $p = 0.7$), FP-GMM reduces bias by approximately 75–80% relative to conventional IV estimators. When the mixture becomes symmetric ($p = 0.5$), the average violation cancels out and all estimators become approximately unbiased, as predicted by the design.

These results highlight that FP-GMM is not designed as a pretest for instrument validity. Rather, it provides a bias correction when instrument invalidity is considered plausible, selecting the largest adjustment compatible with stability of the fixed-point mapping.

Table 3 – Monte Carlo results: Bernoulli switching DGP with instrument invalidity. ($\boldsymbol{\beta = 1, \delta = 0.8, \gamma = 1}$)

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} = 0.9$, $\varepsilon_b = 10^{-12}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$p = 0.3$			$p = 0.45$			$p = 0.5$			$p = 0.55$			$p = 0.7$		
$n = 1000$	*FP-GMM*	-0.06	0.07	0.04	0.11	0.12	0.04	-0.15	0.16	0.04	-0.09	0.10	0.04	0.08	0.08	0.03
	OLS	-0.16	0.16	0.03	-0.04	0.05	0.03	0.00	0.03	0.03	0.04	0.05	0.03	0.16	0.16	0.03
	TSLS	-0.32	0.32	0.05	-0.08	0.10	0.06	0.00	0.06	0.06	0.08	0.10	0.05	0.32	0.32	0.05
	LIML	-0.32	0.32	0.05	-0.08	0.10	0.06	0.00	0.06	0.06	0.08	0.10	0.05	0.32	0.32	0.05
	FULLER	-0.31	0.32	0.05	-0.08	0.10	0.06	0.00	0.05	0.05	0.08	0.09	0.05	0.31	0.32	0.05
$n = 5000$	*FP- GMM*	-0.06	0.06	0.02	0.11	0.11	0.02	-0.15	0.15	0.02	-0.09	0.09	0.02	0.08	0.08	0.02
	OLS	-0.16	0.16	0.01	-0.04	0.04	0.02	0.00	0.02	0.02	0.04	0.04	0.02	0.16	0.16	0.02
	TSLS	-0.32	0.32	0.02	-0.08	0.08	0.02	0.00	0.02	0.02	0.08	0.08	0.02	0.32	0.32	0.02
	LIML	-0.32	0.32	0.02	-0.08	0.08	0.02	0.00	0.02	0.02	0.08	0.08	0.02	0.32	0.32	0.02
	FULLER	-0.32	0.32	0.02	-0.08	0.08	0.02	0.00	0.02	0.02	0.08	0.08	0.02	0.31	0.32	0.02
$n = 10000$	*FP- GMM*	-0.06	0.06	0.01	0.11	0.11	0.01	-0.15	0.15	0.01	-0.09	0.09	0.01	0.08	0.08	0.01
	OLS	-0.16	0.16	0.01	-0.04	0.04	0.01	0.00	0.01	0.01	0.04	0.04	0.01	0.16	0.16	0.01
	TSLS	-0.32	0.32	0.02	-0.08	0.08	0.02	0.00	0.02	0.02	0.08	0.08	0.02	0.32	0.32	0.02
	LIML	-0.32	0.32	0.02	-0.08	0.08	0.02	0.00	0.02	0.02	0.08	0.08	0.02	0.32	0.32	0.02
	FULLER	-0.31	0.31	0.02	-0.08	0.08	0.02	0.00	0.02	0.02	0.08	0.08	0.02	0.31	0.31	0.02

This table reports Monte Carlo results for a nonlinear Bernoulli switching data-generating process in which the instrument may violate the exclusion restriction. The table reports bias, RMSE, and standard deviation for the stability-based estimator and conventional estimators including OLS, TSLS, LIML, and Fuller. Results are based on 500 simulations.

10. Empirical Application: Institutions and Economic Development

We apply the estimator to the instrumental-variables design of Acemoglu, Johnson, and Robinson (2001), where historical settler mortality is used as an instrument for institutional quality in a regression of income on institutions. This application is a natural test case because the instrument is clearly relevant, but its exclusion restriction has been the subject of sustained debate. That combination makes the AJR setting especially useful for a framework that treats relevance as informative while allowing for limited departures from exact exogeneity.

The economic force of the AJR design lies not in first-stage relevance alone, but in the fact that this relevance is given causal meaning by a historically grounded exogeneity argument. Settler mortality matters because AJR interpret its effect on contemporary income as operating primarily through institutional development rather than through an independent direct channel. Once that exclusion claim is questioned, however, the issue is no longer whether the instrument is relevant, but how much identifying weight that relevance should retain. This is precisely the environment for which FP-GMM is designed.

Table 4 reports the estimates. The OLS coefficient is 0.52. Conventional IV estimators are substantially larger: TSLS and LIML are both about 0.93, while Fuller yields 0.86. The FP-GMM estimate is 0.43, well below the conventional IV estimates. Thus, once the identifying content of settler mortality is discounted to allow for limited exclusion violations, the estimated effect of institutions on income becomes materially smaller.

This pattern is economically meaningful. Conventional IV procedures treat the full relevance of settler mortality as valid identifying variation. FP-GMM instead allows part of that relevance to be absorbed by a correction for imperfect exogeneity. The resulting estimate remains positive, but it is substantially attenuated relative to TSLS, LIML, and Fuller. In that sense, the application suggests not that the instrument is uninformative, but that its relevance may overstate the causal effect if exact exclusion fails.

Table 4 – Institutional quality and economic development: AJR (2001) dataset

	$FP-GMM$	OLS	$TSLS$	$LIML$	$Fuller(4)$
$\hat{\beta}$	0.4257	0.5164	0.9294	0.9294	0.8627
$Robust\ Sd(\hat{\beta})$	(0.0632)	(0.0511)	(0.1728)	(0.1563)	(0.1571)

This table reports estimates of the effect of institutional quality on log income per capita using the dataset of Acemoglu, Johnson, and Robinson (2001). The dependent variable is the logarithm of GDP per capita, and institutional quality is measured by the protection against expropriation risk index. Columns report estimates obtained using the proposed fixed-point GMM estimator (FP-GMM), ordinary least squares (OLS), two-stage least squares (TSLS), limited information maximum likelihood (LIML), and the Fuller estimator with parameter 4. Historical settler mortality is used as the instrument for institutional quality, $n = 64$. Robust standard errors are reported in parentheses. For the LIML and Fuller estimators, standard errors are computed using the asymptotic results of Hansen, Hausman, and Newey (2008). For the OLS and TSLS we use Newey-West standard errors, and the FP-GMM uses equation (21) and (22) for the estimation of the standard errors.

The result also helps position the method within the AJR–Albouy debate. Albouy (2012) reports smaller effects under revised mortality data and alternative specifications, whereas Acemoglu et al. (2012) defend the original conclusions. The FP-GMM estimate lies closer to the smaller side of that debate. The point, however, is not to adjudicate the historical controversy directly, but to show that once exact exogeneity is relaxed, the relevance of the instrument can still be retained in a disciplined way rather than either accepted at face value or discarded entirely.

Figure 3 reinforces this interpretation. Over the admissible stability region, the empirical bias criterion declines monotonically toward the relevant boundary, consistent with the theory developed in Section 7. Under the maintained sign restriction that exclusion violations bias conventional IV estimates upward, the admissible correction lies on the positive branch, moving the estimate monotonically away from the TSLS benchmark toward smaller coefficients. The selected endpoint therefore corresponds to the strongest bias-reducing correction that remains compatible with convergence and local stability.

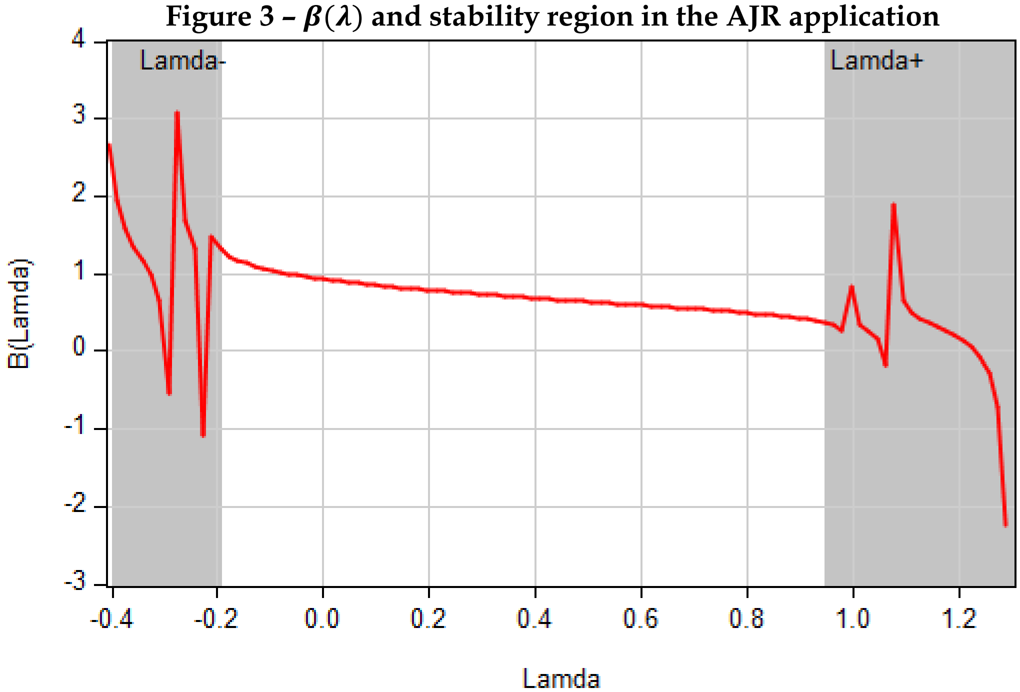
The AJR application illustrates the central role of FP-GMM in settings where relevance is credible but exact exogeneity is contestable. The method preserves the informational content of the instrument, but only to the extent that such relevance remains compatible with a stability-disciplined correction. In this case, that discipline implies a smaller estimated effect of institutions on long-run income than conventional IV procedures suggest.

A natural concern in this application is the relatively small sample size ($n = 64$), which raises the possibility that finite-sample variability could affect the stability region and the location of the selected boundary. Two considerations mitigate this concern. First, the correction is restricted to the admissible stability region, which explicitly avoids regions where the mapping becomes locally unstable and highly sensitive to sampling variation. Second, the empirical $\beta(\lambda)$ displayed in Figure 3 exhibits a clear monotone pattern over the admissible region, with the selected boundary located well within the range where convergence is robust. Taken together, these features suggest that the boundary selection is not driven by local irregularities in small samples, but reflects the underlying bias–stability trade-off emphasized by the theory.

Appendix F.5 examines comparable sample sizes under the same baseline design and varying degrees of instrument invalidity. The results indicate that, in such settings, conventional IV estimators may retain non-negligible bias, while the stability-based estimator exhibits lower mean squared error along the admissible correction path. This provides a useful reference for interpreting the empirical estimates.

If the negative boundary were chosen, the resulting estimate would be approximately 1.3 with a standard error of 0.066, indicating a large and statistically significant effect. Nonetheless, there is no empirical or historical basis in the literature to support a correction in that direction. As a result, the negative bound lacks economic plausibility and is not considered a credible candidate in the AJR application. This asymmetry further reinforces the interpretation that branch selection in this setting is primarily disciplined by economic plausibility rather than by sampling variation.

Figure 3 – $\beta(\lambda)$ and stability region in the AJR application



The figure reports the sample $\beta(\lambda)$ evaluated across values of the correction parameter λ using the AJR data (n = 64). The shaded regions indicate values of π for which the fixed-point fails to satisfy the stability conditions of the mapping. The admissible stability region corresponds to the interval (λ_-, λ_+). The fixed-point algorithm was implemented with tolerance levels $\boldsymbol{\varepsilon_\beta = 10^{-6}}$ and $\boldsymbol{\varepsilon_{g'} = 0.995}$. The sequence $\beta(\lambda)$ declines monotonically along the relevant branch, consistent with the theory in Section 7, and the selected endpoint corresponds to the largest admissible bias-reducing correction. The smooth behavior of the path within the admissible region suggests that the result is not driven by small-sample irregularities.

11. Conclusion

This paper studies identification and estimation with imperfect instruments through a class of parameter-dependent moment models in which stability governs identification. The central result establishes that local identification is equivalent to local stability of the induced mapping: the model is well defined if and only if the corrected instrument retains sufficient effective relevance.

This perspective yields a unified characterization of estimation with imperfect instruments. The admissible region of the correction parameter is simultaneously a stability region and a local identification region. Within this region, the estimator is well defined, locally unique, and convergent. At the boundary, effective relevance collapses, the mapping approaches instability, and identification breaks down. Stability therefore provides an operational criterion linking identification, existence, uniqueness, convergence, and sensitivity within a single object.

The framework also organizes the relevance–exogeneity trade-off internally. Increasing the correction reduces contamination from exclusion violations but weakens effective instrument strength, and the admissible degree of correction is disciplined by stability. The correction parameter admits a natural regularization interpretation, with the stability region defining an identification-constrained regularization set.

Within this structure, boundary selection plays a central role. The stability frontier corresponds to the largest admissible correction consistent with local identification. Under monotone bias reduction, this boundary also has a statistical interpretation: it minimizes mean squared error within the admissible region. Moreover, because conventional k-class estimators retain non-vanishing asymptotic bias under instrument invalidity, the boundary-selected estimator can weakly dominate them when the reduction in bias outweighs the associated variance increase. This establishes a direct link between stability-based identification and risk performance.

Economically, the approach is particularly relevant in environments with excess relevance and limited exogeneity, where instruments provide strong variation but may violate the exclusion restriction. In such settings, conventional IV estimators exploit all available relevance as if it were valid, whereas the present framework discounts relevance endogenously, but only up to the point at which identification can still be sustained.

The paper also establishes existence, uniqueness, convergence, consistency, and asymptotic normality of the estimator, providing a complete characterization of the approach. Monte Carlo simulations and an empirical application illustrate its practical implications, showing that substantial bias reductions can be achieved relative to conventional IV estimators in empirically relevant settings.

Several extensions remain open. First, extending the framework to multivariate settings raises nontrivial challenges, as admissible regions become multidimensional and stability must be characterized through the Jacobian. Second, developing data-driven criteria for selecting the direction of correction would reduce reliance on sign restrictions. Third, further work could explore connections with regularization methods and extend the framework to broader classes of moment conditions.

Overall, the paper shows that stability is not merely a computational property, but an econometric object that jointly characterize local identification, admissible correction, and statistical performance. By linking stability to identification and regularization, the framework provides a new way to organize estimation with imperfect instruments.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. American Economic Review, 91(5), 1369–1401.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2012). The colonial origins of comparative development: An empirical investigation: Reply. American Economic Review, 102(6), 3077–3110.

Aguirregabiria, V., & Mira, P. (2002). Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models. Econometrica, 70(4), 1519-1543.

Albouy, D. Y. (2012). The colonial origins of comparative development: An investigation of the settler mortality data. American Economic Review, 102(6), 3059–3076.

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics, 20, 46–63.

Andrews, D. W. K., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression. Annual Review of Economics, 11, 727–753.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. Journal of the American Statistical Association, 90(430), 431–442.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. Journal of Economic Perspectives, 15(4), 69–85.

Armstrong, T. B., & Kolesár, M. (2021). Sensitivity analysis using approximate moment condition models. Quantitative Economics, 12, 77–108.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. Journal of the American Statistical Association, 90, 443–450.

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly exogenous. Review of Economics and Statistics, 94(1), 260–272.

Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. Journal of econometrics, 34(3), 305-334.

Chernozhukov, V., & Hong, H. (2003). An MCMC approach to classical estimation. Journal of econometrics, 115(2), 293-346.

Chernozhukov, V., Hansen, C., & Jansson, M. (2015). Inference with many weak instruments. Review of Economic Studies, 82(2), 575–600.

Donald, S. G., & Newey, W. K. (2001). Choosing the number of instruments. Econometrica, 69(5), 1161–1191.

Fuller, W. A. (1977). Some properties of a modification of the limited information estimator. Econometrica, 45(4), 939–953.

Guggenberger, P., & Smith, R. J. (2005). Generalized empirical likelihood estimators and tests under partial, weak, and strong identification. Econometric Theory, 21, 667–709.

Hall, A. R., & Inoue, A. (2003). The large sample behaviour of the generalized method of moments estimator in misspecified models. Journal of Econometrics, 114(2), 361–394.

Hansen, C., Hausman, J., & Newey, W. (2008). Estimation with many instrumental variables. Journal of Business and Economic Statistics, 26(4), 398–422.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. Econometrica, 50(4), 1029–1054.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. Econometrica, 46, 931–959.

Kolesár, M. (2013). Estimation in instrumental variables models with many instruments. Review of Economics and Statistics, 95(2), 524–538.

Kolesár, M., Chetty, R., Friedman, J., Glaeser, E. L., & Imbens, G. W. (2015). Identification and inference with many invalid instruments. Journal of Business and Economic Statistics, 33(4), 474–484.

Manski, C. F. (2003). Partial Identification of Probability Distributions. New York: Springer.

Murphy, K. M., & Topel, R. H. (1985). Estimation and inference in two-step econometric models. Journal of Business and Economic Statistics, 3(4), 370–379.

Nagar, A. L. (1959). The bias and moment matrix of the general k-class estimators of the parameters in simultaneous equations. Econometrica: Journal of the Econometric Society, 575-595.

Nevo, A., & Rosen, A. M. (2012). Identification with imperfect instruments. Review of Economics and Statistics, 94(3), 659–671.

Newey, W. K., & McFadden, D. (1994). Large sample estimation and hypothesis testing. In R. F. Engle & D. McFadden (Eds.), Handbook of Econometrics (Vol. 4, pp. 2111–2245). Amsterdam: Elsevier.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. Econometrica: Journal of the Econometric Society, 999-1033.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. Econometrica, 65(3), 557–586.

Stock, J. H., & Wright, J. H. (2000). GMM with weak identification. Econometrica, 68(5), 1055–1096.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), Identification and Inference for Econometric Models. Cambridge: Cambridge University Press.

Tamer, E. (2010). Partial identification in econometrics. Annual Review of Economics, 2, 167–195.

Theil, H. (1958). Economic Forecasts and Policy. Amsterdam: North-Holland.

van der Vaart, A. W. (2000). Asymptotic Statistics. Cambridge: Cambridge University Press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. Journal of Human Resources, 50(2), 420–445

Appendix A - Fixed-Point Theory and Stability of the GMM Estimator

A1. Proof of Theorem 2.1

$g(\beta;\lambda) = \mathbb{E}[q(z,\lambda,u(\beta))y]/\mathbb{E}[q(z,\lambda,u(\beta))x]$.

Start from the moment condition:

$\mathbb{E}[q(z,\lambda,u(\beta))u(\beta)] = 0$.

Substituting $u(\beta) = y - x\beta$, we obtain:

$\mathbb{E}[q(z,\lambda,u(\beta))(y - x\beta)] = 0$.

Expanding:

$\mathbb{E}[q(z,\lambda,u(\beta))y] - \beta\mathbb{E}[q(z,\lambda,u(\beta))x] = 0$

Rearranging:

$\beta = \frac{\mathbb{E}[q(z,\lambda,u(\beta))y]}{\mathbb{E}[q(z,\lambda,u(\beta))x]}$.

Define the mapping

$g(\beta;\lambda) = \frac{\mathbb{E}[q(z,\lambda,u(\beta))y]}{\mathbb{E}[q(z,\lambda,u(\beta))x]}$.

Then any solution to the moment condition satisfies

$\beta = g(\beta;\lambda)$

which establishes the fixed-point representation.

A2. Population Mapping

Recall the modified instrument defined in the main text:

$z^{FP}(\beta;\pi;\gamma) = \hat{x} - \pi u(\beta)$

Define the population moments:

$A(\beta;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)y_i]$

$B(\beta;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)x_i]$

The population mapping is:

$$g(\beta;\pi;\gamma) = \frac{A(\beta;\pi;\gamma)}{B(\beta;\pi;\gamma)}$$

The fixed point satisfies:

$\beta(\pi) = g\big(\beta(\pi); \pi; \gamma\big)$

A3. Explicit Derivative of the Mapping

To study stability, we compute:

$$g'(\beta; \pi; \gamma) = \frac{\partial}{\partial\beta} g(\beta; \pi; \gamma)$$

Using quotient rule:

$$g'(\beta; \pi; \gamma) = \frac{A'(\beta;\pi;\gamma)B(\beta;\pi;\gamma) - A(\beta;\pi;\gamma)B'(\beta;\pi;\gamma)}{\big(B(\beta;\pi;\gamma)\big)^2}$$

Since:

$$z^{FP}(\beta; \pi; \gamma) = \hat{x} - \pi u(\beta) = \hat{x} - \pi(y - \beta x)$$

We have:

$$\frac{\partial z^{FP}}{\partial\beta} = \pi x$$

Therefore:

$$A'(\beta; \pi) = \pi\mathbb{E}[xy]$$

$$B'(\beta; \pi) = \pi\mathbb{E}[x^2]$$

$$g'(\beta; \pi) = \frac{\pi\big(\mathbb{E}[xy]B(\beta;\pi) - A(\beta;\pi)\mathbb{E}[x^2]\big)}{\big(B(\beta;\pi)\big)^2}$$

A4. Derivative at the Fixed Point

At the fixed point $\beta = \beta(\pi)$, we have:

$$A(\beta; \pi; \gamma) = \beta(\pi)B(\beta; \pi; \gamma)$$

Substituting:

$$g'(\beta; \pi; \gamma) = \frac{\pi\big(\mathbb{E}[xy]B(\beta;\pi) - \beta(\pi)B(\beta;\pi)\ \mathbb{E}[x^2]\big)}{\big(B(\beta;\pi)\big)^2} = \frac{\pi B(\beta;\pi)(\mathbb{E}[xy] - \beta(\pi)\mathbb{E}[x^2])}{\big(B(\beta;\pi)\big)^2}$$

Simplifying:

$$g'(\beta; \pi; \gamma) = \frac{\pi(\mathbb{E}[xy] - \beta(\pi)\mathbb{E}[x^2])}{B(\beta;\pi;\gamma)}$$

This is the key stability object.

A5. Interpretation

Note that

$$\mathbb{E}[xy] - \beta(\pi)\mathbb{E}[x^2] = \mathbb{E}[x(y - \beta(\pi)x)] = \mathbb{E}[xu(\beta)]$$

Thus:

$g'(\beta;\pi;\gamma) = \frac{\pi\mathbb{E}[xu(\beta)]}{B(\beta;\pi;\gamma)}$

This shows:

- Stability depends on residual endogeneity.
- Stability depends on relevance
- Instability arises when denominator shrinks (weak IV region).

A6. Stability condition

Define the stability region:

$\mathcal{A} \equiv \{\pi \in \Pi: |g'(\beta(\pi);\pi;\gamma)| < 1\}$

Explicitly:

$\left|\frac{\pi\mathbb{E}[xu(\beta)]}{B(\beta;\pi;\gamma)}\right| < 1$

A7. Closed form of g in observable moments for a given π

Recall that

$g'(\beta;\pi;\gamma) = \frac{\pi\left(\mathbb{E}[xy]B(\beta;\pi;\gamma) - A(\beta;\pi;\gamma)\mathbb{E}[x^2]\right)}{\left(B(\beta;\pi;\gamma)\right)^2}$

$A(\beta;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)y_i]$

$B(\beta;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta;\pi;\gamma)x_i]$

Consider:

$A(\beta;\pi;\gamma) = \mathbb{E}[(\hat{x} - (y - x\beta)\pi)y] = \mathbb{E}[\hat{x}y] - \pi\mathbb{E}[y^2] + \beta\pi\mathbb{E}[xy]$

$B(\beta;\pi;\gamma) = \mathbb{E}[(\hat{x} - (y - x\beta)\pi)x] = \mathbb{E}[\hat{x}x] - \pi\mathbb{E}[xy] + \beta\pi\mathbb{E}[x^2] = \mathbb{E}[\hat{x}^2] - \pi\mathbb{E}[xy] + \beta\pi\mathbb{E}[x^2]$

Therefore:

$g'(\beta;\pi;\gamma) = \frac{\pi\left(\mathbb{E}[xy]\left(\mathbb{E}[\hat{x}^2] - \pi\mathbb{E}[xy] + \beta\pi\mathbb{E}[x^2]\right) - \left(\mathbb{E}[\hat{x}y] - \pi\mathbb{E}[y^2] + \beta\pi\mathbb{E}[xy]\right)\mathbb{E}[x^2]\right)}{(\mathbb{E}[\hat{x}^2] - \pi\mathbb{E}[xy] + \beta\pi\mathbb{E}[x^2])}$

$g'(\beta;\pi;\gamma) = \frac{\pi\left(\mathbb{E}[xy]\mathbb{E}[\hat{x}^2] - \mathbb{E}[x^2]\mathbb{E}[\hat{x}y] + \pi\left(\mathbb{E}[x^2]\mathbb{E}[y^2] - \mathbb{E}[xy]^2\right)\right)}{(\mathbb{E}[\hat{x}^2] - \pi\mathbb{E}[xy] + \beta\pi\mathbb{E}[x^2])^2}$

A8. Theorem 3.1 (Local Existence and Uniqueness)

Suppose:

1. $B(\beta;\pi;\gamma)$ is continuous and bounded away from zero in a neighborhood of the fixed point
2. $g'(\beta;\pi;\gamma)$ exists and $|g'(\beta;\pi)| < 1$

Then:

- A unique fixed point exists locally.
- Fixed-point iteration converges locally.

Since the mapping is continuously differentiable and locally Lipschitz with constant strictly smaller than one, the scalar Banach fixed-point theorem applies.

A9. Spectral Radius Interpretation

In the scalar case

$\rho(g') = |g'|$

Thus, the condition

$\rho(g') < 1$

Is exactly the same condition

A10. Connection to Iterative Error Propagation

From fixed-point iteration:

$\beta_{j+1} - \beta(\pi) \approx g'(\beta(\pi); \pi)\left(\beta_j - \beta(\pi)\right)$

Thus

$|\beta_j - \beta(\pi)| \approx |g'|^j |\beta_0 - \beta(\pi)|$

If we have:

- $|g'| \ll 1$: fast decay
- $|g'| \approx 1$: slow decay and the error accumulate.
- $|g'| < 1$: divergence

A11. Boundary and Weak-GMM Region

Instability occurs when:

$|g'(\beta(\pi); \pi; \gamma)| = 1$

Or

$\left|\frac{\pi\mathbb{E}[xu(\beta)]}{B(\beta;\pi;\gamma)}\right| = 1$

Notice

$B(\beta; \pi; \gamma) \to 0$

Then stability necessarily fails. Thus, the boundary of $\mathcal{A}$ is tightly linked to loss of relevance.

A12. Stability and the λ −Dependent Identification Region

Since $\lambda = \pi/\gamma$, then

$\frac{d}{d\beta} z^{FP}(\beta_0) = \lambda x$

There exists $C > 0$ such that

$|g'(\beta_0)| < \frac{C\lambda}{|B(\beta_0)|}$

$B(\beta_0) = \mathbb{E}[\hat{x}x] - \lambda\mathbb{E}[ux]$

Thus

$B(\beta_0) = \mathbb{E}[\hat{x}x] - \lambda\mathbb{E}[ux]$

$|B(\beta_0)| = |\mathbb{E}[\hat{x}x]| - \lambda|\mathbb{E}[ux]|$

From the stability condition

$|g'(\beta_0)| < 1 \Longrightarrow |B(\beta_0)| > C\lambda$

Thus

$|\mathbb{E}[\hat{x}x]| - \lambda|\mathbb{E}[ux]| > C\lambda$

Rearrange

$|\mathbb{E}[ux]| < \frac{|\mathbb{E}[\hat{x}x]|}{\lambda} - C$

Using the proportionality between $\mathbb{E}[ux]$ and $\mathbb{E}[zu]$, we obtain

$|\mathbb{E}[zu]| \leq \varsigma(\lambda)|\mathbb{E}[zx]|$

With

$\varsigma(\lambda) = \frac{|\mathbb{E}[\hat{x}x]| - C\lambda}{\lambda\,|\mathbb{E}[zx]|}$

Therefore as λ increases i) correction becomes stronger, ii) relevance declines faster and iii) admissible invalidity shrinks.

A13. Global geometric convergence

We begin with a global bound based on the contraction property of the mapping. Let $C_r(\beta^*)$ denote a neighborhood of radius r around the fixed point.

Assumption A11 (Local contraction).

There exists a constant $q(r) < 1$ such that

$|g(\beta_1) - g(\beta_2)| \leq q(r)|\beta_1 - \beta_2|$

For $\beta_1, \beta_2 \in C_r(\beta^*)$. This assumption implies that the mapping is a contraction in the neighborhood $C_r(\beta^*)$.

Proposition A11 (geometric convergence)

Under assumption A11, for any initial value $\beta_0 \in C_r(\beta^*)$ the fixed-point satisfies

$$\left|\beta_j - \beta^*\right| \leq q^j |\beta_0 - \beta^*|, \quad j = 0,1,2,\ldots$$

Thus, the sequence converges to the fixed point at a geometric rate with ratio $q(r)$.

Proof

Using the contraction property

$$\left|\beta_{j+1} - \beta^*\right| = \left|g\left(\beta_j - g(\beta^*)\right)\right| \text{ since } \beta^* = g(\beta^*).$$

Applying the contraction inequality yields

$$\left|\beta_{j+1} - \beta^*\right| \leq q(r)\left|\beta_j - \beta^*\right|$$

Applying the inequality recursively gives

$$\left|\beta_{j+1} - \beta^*\right| \leq q(r)^j\left|\beta_j - \beta^*\right|$$

This establishes geometric convergence

A14. Local convergence rate

We now derive the local rate of convergence using a Taylor expansion of the mapping around the fixed point. Suppose the mapping $g(\beta)$ is differentiable in a neighborhood of β^*. Using a first-order Taylor expansion around β^*

$$g(\beta_j) = g(\beta^*) + g'(\beta^*)(\beta_j - \beta^*) + o(|\beta_j - \beta^*|)$$

Since $\beta^* = g(\beta^*)$

$$\beta_{j+1} - \beta^* = g'(\beta^*)(\beta_j - \beta^*) + o(|\beta_j - \beta^*|)$$

For β_j sufficiently close to β^*, the higher-order term becomes negligible and the iteration approximately satisfies

$$\beta_{j+1} - \beta^* \approx g'(\beta^*)(\beta_j - \beta^*)$$

Taking absolute values gives

$$\left|\beta_{j+1} - \beta^*\right| \approx |g'(\beta^*)|\left|(\beta_j - \beta^*)\right|$$

Applying the recursion yields

$$\left|\beta_j - \beta^*\right| \approx |g'(\beta^*)|^j |(\beta_0 - \beta^*)|$$

A15. Linear convergence

The result above implies the following property.

Proposition A2 (Local linear convergence).

If $|g'(\beta^*)| < 1$ then the fixed-point iteration converges locally at a linear rate with convergence factor $|g'(\beta^*)|$

The quantity $|g'(\beta^*)|$ therefore determines the speed of convergence of the algorithm. When the derivative is small, the iteration converges rapidly. Conversely, as the derivative approaches unity, the convergence rate slows down and the iteration becomes increasingly sensitive to perturbations in the parameter.

A16. Interpretation

The derivative of the mapping at the fixed point plays a central role in the behavior of the estimator. First, the condition $|g'(\beta^*)| < 1$ ensures local stability of the fixed point and guarantees convergence of the iterative algorithm.

Second, the magnitude of $|g'(\beta^*)|$ determines the speed of convergence. Smaller values imply faster convergence, while values close to unity imply slow convergence and increased sensitivity to sampling variation.

This observation links the statistical behavior of the estimator to the dynamic properties of the fixed-point mapping. In particular, when the derivative approaches unity the estimator becomes highly sensitive to perturbations, reflecting the loss of stability of the fixed-point iteration.

A17. Multivariate extension with multiple endogenous regressors

We now extend the fixed-point formulation to the case of multiple endogenous regressors. By the Frisch-Waugh-Lovell theorem, intercepts and included exogenous regressors can be partialled out without altering the identification argument for the parameter of interest. Let the residualized structural equation be

$y_i = x_i\beta_0 + u$

where

$x_i \in \mathbb{R}^p, \beta_0 \in \mathbb{R}^p$, and $u_i \in \mathbb{R}$. Let $z_i \in \mathbb{R}^L$ denote the excluded instruments, and let $\hat{x}_i = \mathbb{E}[x_i|z_i]$denote the fitted first-stage index. In sample notation, let X be the $n \times p$ matrix of endogenous regressors and let $\hat{X} = P_z X$ denote the matrix of first-stage fitted values. For any candidate value β, define the residual vector

$u(\beta) = y - X\beta \in \mathbb{R}^n$

Let $\lambda \in \mathbb{R}^p$ be a vector-valued regularization parameter. We define the corrected instrument matrix as

$$Z^M(\beta;\lambda) = \hat{X} - u(\beta)\lambda'$$

where $u(\beta)\lambda'$ is an $n \times p$ rank-one matrix. Thus, each column of the fitted first-stage matrix is corrected by a residual-based term whose magnitude is governed by the corresponding component of λ. The associated sample moment condition is

$$m_n(\beta;\lambda) = \frac{1}{n} Z^M(\beta;\lambda)' u(\beta) = \frac{1}{n}\hat{X}u(\beta) - \lambda\frac{1}{n}u(\beta)'u(\beta)$$

The estimator is defined as a solution to

$$m_n(\beta;\lambda) = 0$$

This formulation preserves the logic of the scalar case. The fitted first-stage matrix $\hat{X}$ remains the source of relevance, while the residual-based adjustment trades off that relevance against improved exogeneity. The role of λ is now vector-valued: it governs the direction and magnitude of the correction across the different endogenous regressors.

A18. Fixed-point mapping in vector form

Assume that the matrix

$$Q_{\hat{X}X,n} = \frac{1}{n}\hat{X}X$$

Is nonsingular. Rearranging the moment condition yields

$$\frac{1}{n}\hat{X}y - \frac{1}{n}\hat{X}X\beta - \lambda\frac{1}{n}u(\beta)'u(\beta) = 0$$

so that the estimator can be written as the fixed point of the vector-valued mapping

$$\beta = G_n(\beta;\lambda) := Q_{\hat{X}X,n}^{-1}\left[\frac{1}{n}\hat{X}y - \lambda\frac{1}{n}u(\beta)'u(\beta)\right]$$

The corresponding population mapping is

$$G_n(\beta;\lambda) := Q_{\hat{X}X}^{-1}\left[\frac{1}{n}Q_{\hat{X}y} - \lambda\frac{1}{n}\mathbb{E}[u_i(\beta)^2]\right]$$

where

$$Q_{\hat{X}X} = \mathbb{E}[\hat{x}_i x_i'] \quad Q_{\hat{X}y} = \mathbb{E}[\hat{x}_i y_i]$$

A population fixed point $\beta^*(\lambda)$ satisfies

$$\beta^*(\lambda) = G(\beta^*(\lambda);\lambda)$$

This representation makes clear that the multivariate stability-based estimator remains a parameter-dependent moment estimator whose solution is characterized by a fixed point. The main difference from the scalar case is that the mapping is now vector-valued and the admissible correction is no longer summarized by a scalar interval, but by a subset of $\mathbb{R}^p$.

Appendix A19. Jacobian and local stability

To study local stability, differentiate the population mapping with respect to β. Since

$\mathbb{E}[u_i(\beta)^2] = \mathbb{E}[(y_i - x_i'\beta)^2]$

its gradient is

$\nabla_\beta \mathbb{E}[u_i(\beta)^2] = -2\mathbb{E}[x_i u_i(\beta)]$

Hence, the Jacobian of the mapping is

$J_G(\beta;\lambda) = \frac{\partial G(\beta;\lambda)}{\partial \beta'} = 2Q_{\hat{X}X}^{-1}\lambda\mathbb{E}[x_i u_i(\beta)]'$

Thus, the Jacobian is the outer product of the vector $2Q_{\hat{X}X}^{-1}\lambda$ and the vector $\mathbb{E}[x_i u_i(\beta)]$. In particular, it is rank one. At the fixed point $\beta^*(\lambda)$

$J_G(\beta^*(\lambda);\lambda) = 2Q_{\hat{X}X}^{-1}\lambda\mathbb{E}[x_i u_i^*]'$

where $u_i^* = y_i - x_i\beta^*(\lambda)$

A sufficient condition for local stability is

$\rho\big(J_G(\beta^*(\lambda);\lambda)\big) < 1$

$\rho(.)$ denotes the spectral radius. More generally, if there exists a neighborhood $\mathfrak{B}$ of $\beta^*(\lambda)$ and a matrix norm $\|\cdot\|$ such that

$\sup_{\beta\in\mathfrak{B}} \| J_G(\beta;\lambda) \| \leq q < 1$

Then $G(\beta;\lambda)$ is a contraction on $\mathfrak{B}$. By the multivariate Banach fixed-point theorem, the fixed point is locally unique and the iteration

$\beta^{t+1} = G(\beta^t;\lambda)$

converges geometrically for initial values sufficiently close to $\beta^*(\lambda)$. Because $J_G(\beta;\lambda)$ is rank one, its only nonzero eigenvalue is

$\mu(\beta;\lambda) = 2\mathbb{E}[x_i u_i(\beta)]'Q_{\hat{X}X}^{-1}\lambda$

Therefore, local stability may be expressed equivalently as

$\left|2\mathbb{E}\left[x_i u_i\big(\beta^*(\lambda)\big)\right]'Q_{\hat{X}X}^{-1}\lambda\right| < 1$

This provides the multivariate analogue of the scalar derivative condition. The difference is that admissibility now depends on the joint direction of λ relative to the residual endogeneity vector $\mathbb{E}[x_i u_i(\beta)]$ and to the geometry of the first-stage matrix $Q_{\hat{X}X}$.

Define the multivariate admissible region as

$\mathcal{A} = \{\lambda \in \mathbb{R}^p : \rho\big(J_G(\beta^*(\lambda);\lambda)\big) < 1\}$

Unlike the scalar case, $\mathcal{A}$ is generally a multidimensional set rather than an interval, and there is no globally monotone correction path indexed by a single scalar. This is why the scalar case is especially useful for characterizing the trade-off between relevance and exogeneity in a transparent way.

Appendix A20. Interpretation and matrix-valued extension

The multivariate extension preserves the main intuition of the scalar framework. Larger values of λ impose stronger residual-based corrections, which can reduce contamination from exclusion violations but simultaneously attenuate the identifying content of the fitted first-stage matrix. Stability therefore continues to determine the admissible correction, but now in a multidimensional geometry. In particular, admissibility depends not only on the magnitude of λ, but also on its direction in parameter space.

The present extension considers a single-equation model with multiple endogenous regressors, so that $u_i(\beta)$ is scalar at the observation level and $u(\beta)$ is an $n \times 1$ residual vector in sample notation. A further extension would allow genuinely vector-valued structural errors, as in systems of equations or multivariate outcome models. In that case one may write

$$U(B) = Y - XB,$$

with $U(B) \in \mathbb{R}^{nxq}, B \in \mathbb{R}^{pxq}$, and define a matrix-valued correction parameter $\Lambda \in \mathbb{R}^{qxp}$. A natural corrected instrument takes the form

$$Z^M(B; \Lambda) = \hat{X} - U(B)\Lambda$$

which yields a matrix-valued moment condition and a vectorized fixed-point mapping for $vec(B)$. Local stability is then governed by the spectral radius of the Jacobian of that vectorized mapping. The admissible correction region becomes a subset of $\mathbb{R}^{pq}$, and the simple scalar ordering of the correction path is no longer available.

For this reason, the main text focuses on the scalar case, where the admissible region can be characterized directly and the trade-off between relevance and exogeneity is fully transparent. The multivariate extension developed here shows, however, that the fixed-point logic is not tied to the one-dimensional setting and can in principle be generalized to richer linear IV environments.

Appendix A21. Primitive inner bound for the admissible region

In the scalar case, the exact admissible stability region is defined implicitly by the derivative condition

$$\mathcal{A} = \{\lambda: |g'(\beta^*(\lambda); \lambda)| < 1\}$$

Where $\beta^*(\lambda)$ denotes the population fixed point. This characterization is exact, but implicit. We now derive a primitive inner bound that makes explicit how admissibility depends on baseline relevance, residual endogeneity, and correction-induced attenuation of instrument strength.

Recall that the derivative of the fixed-point mapping can be written as

$$g'(\beta;\pi;\gamma) = \frac{\pi\mathbb{E}[xu(\beta^*(\lambda))]}{B(\beta^*(\lambda);\lambda)}$$

Where $B(\beta,\lambda)$ denotes the effective relevance term in the denominator of the mapping. The key trade-off is immediate: larger values of λ strengthen the residual-based correction, but also reduce the effective identifying variation of the corrected instrument.

Suppose there exist constants $\underline{r} > 0, \bar{e} < \infty$, and $\bar{d} < \infty$ such that, for all β in a neighborhood of the relevant fixed point and all λ in a candidate interval,

$$B(\beta,\lambda) \geq \underline{r} - |\lambda|\bar{d}, \qquad |\mathbb{E}[xu(\beta)]| < \bar{e}$$

Here:

$\underline{r}$ measures baseline relevance,

$\bar{e}$ bounds residual endogeneity,

$\bar{d}$ bounds the rate at which correction attenuates relevance.

Then

$$g'(\beta^*(\lambda);\lambda) = \frac{|\lambda|\bar{e}}{\underline{r}-|\lambda|\bar{d}}$$

provided the denominator remains positive. To impose a strict stability margin, fix $\varepsilon \in (0,1)$ and require

$$g'(\beta^*(\lambda);\lambda) < 1-\varepsilon$$

A sufficient condition is therefore

$$\frac{|\lambda|\bar{e}}{\underline{r}-|\lambda|\bar{d}} < 1-\varepsilon$$

Rearranging,

$$|\lambda|\bar{e} < (1-\varepsilon)\underline{r} - (1-\varepsilon)|\lambda|\bar{d}$$

so that

$$|\lambda|\left[\bar{e} + (1-\varepsilon)\bar{d}\right] < (1-\varepsilon)\underline{r}$$

Hence a primitive sufficient condition for stability is

$$|\lambda| < \bar{\lambda}_{prim}(\varepsilon) := \frac{(1-\varepsilon)\underline{r}}{\bar{e}+(1-\varepsilon)\bar{d}}$$

This yields the primitive inner bound

$$\mathcal{A}_{prim}(\varepsilon) = \{\lambda: |\lambda| < \bar{\lambda}_{prim}(\varepsilon)\} = [-\bar{\lambda}_{prim}(\varepsilon), \bar{\lambda}_{prim}(\varepsilon)]$$

Thus, $\mathcal{A}_{prim}(\varepsilon) \subseteq \mathcal{A}$. Although this bound is generally conservative, it has the advantage of making the economics of admissibility transparent. The admissible region expands with baseline relevance $\underline{r}$, and contracts with residual endogeneity $\bar{e}$, correction-induced attenuation $\bar{d}$, and a stricter stability margin ε. The comparative statics are immediate:

$$\frac{\partial \bar{\lambda}_{prim}(\varepsilon)}{\partial \underline{r}} > 0, \quad \frac{\partial \bar{\lambda}_{prim}(\varepsilon)}{\partial \bar{e}} < 0 \qquad \frac{\partial \bar{\lambda}_{prim}(\varepsilon)}{\partial \bar{d}} < 0, \quad \frac{\partial \bar{\lambda}_{prim}(\varepsilon)}{\partial \varepsilon} < 0$$

In words, stronger baseline relevance allows larger admissible corrections, whereas larger residual endogeneity, faster attenuation of relevance, or a more conservative stability requirement all reduce the admissible interval.

This primitive characterization complements the exact derivative-based definition used in the main text. The exact admissible region remains the relevant theoretical object, but the bound above shows that admissibility is ultimately governed by primitive features of the signal structure of the model rather than by the numerical iteration alone.

Proposition A19 (Primitive inner bound for the admissible region).

Under the stated bounds on baseline relevance, residual endogeneity, and correction-induced attenuation, the interval

$$\mathcal{A}_{prim}(\varepsilon) = \left[-\frac{(1-\varepsilon)\underline{r}}{\bar{e}+(1-\varepsilon)\bar{d}}, \frac{(1-\varepsilon)\underline{r}}{\bar{e}+(1-\varepsilon)\bar{d}}\right]$$

is contained in the exact admissible stability region.

Appendix A22. Proof of theorem 4.2

Theorem 4.2 (Stability–Identification Equivalence)

Let $\beta^*(\lambda)$ denote the population fixed point of the FP-GMM mapping $g(\beta;\lambda)$, and let the admissible region be defined by (14). Suppose Assumptions 4.3–4.8 hold, and suppose the effective relevance term $B(\beta,\lambda)$ is continuously differentiable and nondegenerate in a neighborhood of $(\beta^*(\lambda),\lambda)$.

Then, for any $\lambda \in \mathcal{A}$, the following are equivalent in a neighborhood of the fixed point:

1. The mapping is locally stable, i.e. $g(\beta^*(\lambda);\lambda) < 1-\varepsilon$
2. The population fixed point is locally unique and the structural parameter is locally identified
3. The corrected instrument retains nondegenerate effective relevance, so that $B(\beta^*(\lambda),\lambda) > 0$

Moreover, as λ approaches the boundary of $\mathcal{A}$, the effective relevance term approaches zero, the derivative approaches unity in magnitude, and local identification breaks down. Thus, the stability frontier coincides with the identification frontier.

The theorem formalizes the central role of stability in the parameter-dependent moment framework. Local stability is not only a computational property of the fixed-point iteration; it is equivalent to local identification of the structural parameter because both require the corrected instrument to retain enough effective relevance after the residual-based adjustment. In this sense, the admissible region is simultaneously a stability region and a local identification region. At the boundary, the correction becomes too strong relative to the remaining signal, the mapping loses contraction, and the parameter becomes weakly or non-locally identified.

The proof proceeds in three steps.

Step 1. Stability implies local uniqueness and local identification

By definition, for $\lambda \in \mathcal{A}$

$g'(\beta^*(\lambda); \lambda) < 1 - \varepsilon$

for some $\varepsilon \in (0,1)$. Since $g(\beta; \lambda)$ is continuously differentiable in a neighborhood of $\beta^*(\lambda)$, the contraction mapping theorem implies that g is locally contractive around the fixed point. Hence the fixed point is locally unique.

Local uniqueness of the fixed point implies local identification of the structural parameter in the residualized just-identified problem, because any nearby parameter value must violate the population moment condition. Therefore, stability implies local identification.

Step 2. Local identification requires nondegenerate effective relevance

The derivative of the mapping depends on the effective relevance term $B(\beta, \lambda)$, which appears in the denominator of the fixed-point representation. When

$B(\beta^*(\lambda), \lambda) > 0$

the corrected instrument retains nondegenerate identifying variation, so the mapping is well defined and locally responsive to perturbations in β.

By contrast, if

$B(\beta^*(\lambda), \lambda) = 0$

the corrected instrument loses effective relevance. In that case the fixed-point derivative becomes singular or approaches unity in magnitude, the mapping ceases to be locally contractive, and the moment condition no longer pins down β uniquely in a neighborhood of the fixed point. Hence local identification fails.

Thus, local identification requires nondegenerate effective relevance.

Step 3. Equivalence and the boundary

Combining Steps 1 and 2, local stability, local uniqueness, and local identification are all equivalent to the requirement that the corrected instrument retain strictly positive effective relevance in a neighborhood of the fixed point.

Now consider a sequence $\lambda_n \to \partial\mathcal{A}$. By definition of the admissible region,

$|g'(\beta^*(\lambda_n); \lambda_n)| \to 1$

Since the derivative is governed by the ratio between residual endogeneity and effective relevance, this can occur only when the effective relevance term $B(\beta^*(\lambda_n), \lambda_n)$ becomes arbitrarily small. As a result, the corrected instrument approaches a weak-identification regime, the fixed point becomes increasingly sensitive to perturbations, and local identification breaks down at the frontier.

Therefore, the stability frontier coincides with the identification frontier.

Appendix A23. Proof of theorem 4.3

We proceed in three steps.

Step 1. Trade-off induced by residual-based correction

The corrected instrument takes the form

$z^{FP}(\beta, \lambda) = \hat{x} - \lambda u(\beta)$

The moment condition is

$\mathbb{E}[z^{FP}(\beta, \lambda) u(\beta)] = \mathbb{E}[\hat{x} u(\beta) - \lambda u(\beta)^2] = \mathbb{E}[\hat{x} u(\beta)] - \mathbb{E}[\lambda u(\beta)^2]$

- The first term reflects contamination through imperfect exclusion.
- The second term increases in magnitude with λ and reduces that contamination.

At the same time, the effective relevance term entering the mapping can be written a

$B(\beta, \lambda) = \mathbb{E}[\hat{x} x] - \lambda D(\beta)$

where

$D(\beta)$ captures attenuation of relevance induced by the correction. Thus, increasing λ

- reduces contamination (improves exogeneity),
- reduces effective relevance.

This establishes the trade-off.

Step 2. Stability determines admissibility

Rearranging the moment condition yields a self-mapping

$\beta = g(\beta; \lambda)$.

By standard arguments, local stability requires

$|g'(\beta^*(\lambda); \lambda))| < 1$

The derivative depends on:

- residual endogeneity $\mathbb{E}[\hat{x} u(\beta)]$,
- effective relevance $B(\beta, \lambda)$

Hence, admissibility is determined by the balance between these two forces.

Step 3. Single-object discipline

The same derivative g' governs:

- existence and uniqueness via contraction,
- convergence of the iteration,
- sensitivity of the estimator,
- and the admissible correction via stability.

No additional object is required to characterize these properties.

Thus, the trade-off is organized through a single internal object.

The residual-based correction induces a trade-off between exogeneity and relevance, and the stability of the fixed-point mapping determines the admissible region of that trade-off. This establishes the result.

Appendix B – Uniform Convergence and Proof of Theorem

B1. Uniform convergence of the Mapping

Define

$g_n(\beta;\pi;\hat{\gamma}) = \frac{\hat{A}(\beta;\pi;\hat{\gamma})}{\hat{B}(\beta;\pi;\hat{\gamma})}$; $g(\beta;\pi) = \frac{A(\beta;\pi;\gamma)}{B(\beta;\pi;\gamma)}$

Since both numerator and denominator are linear combinations of sample averages of square-integrable variables the uniform LLN implies

$\sup_{\beta \in D} |\hat{A}(\beta;\pi;\hat{\gamma}) - A(\beta;\pi;\gamma)| \to_p 0$

$\sup_{\beta \in D} |\hat{B}(\beta;\pi;\hat{\gamma}) - B(\beta;\pi;\gamma)| \to_p 0$

Because the denominator is bounded away from zero:

$\sup_{\beta \in D} |g_n(\beta;\pi;\hat{\gamma}) - g(\beta;\pi;\gamma)| \to_p 0$

B2. Consistency via Fixed-Point Argument

Define the operator

$T_\pi(\beta) = g(\beta;\pi;\gamma)$

By appendix A:

- $T_\pi(\beta)$ is continuous differentiable
- $|T'_\pi(\beta(\pi))| < 1$

Hence $\beta(\pi)$ is locally unique. Let $\hat{\beta}(\pi)$ satisfy

$\hat{\beta}(\pi) = g_n(\hat{\beta}(\pi);\pi;\hat{\gamma})$

Using uniform convergence:

$\hat{\beta}(\pi) - \beta(\pi) = g_n\big(\hat{\beta}(\pi);\pi;\hat{\gamma}\big) - g(\beta(\pi);\pi;\gamma)$

Add and subtract $g_n\big(\hat{\beta}(\pi);\pi;\hat{\gamma}\big)$

$\hat{\beta}(\pi) - \beta(\pi) = \Big(g_n\big(\hat{\beta}(\pi);\pi;\hat{\gamma}\big) - \ g_n\big(\hat{\beta}(\pi);\pi;\hat{\gamma}\big)\Big) + \Big(\ g_n\big(\hat{\beta}(\pi);\pi;\hat{\gamma}\big) - g(\beta(\pi);\pi;\gamma)\Big)$

The first term is $o_p(1)$. The second term is approximately:

$g'(\beta(\pi);\pi;\gamma)\left(\hat{\beta}(\pi) - \beta(\pi)\right)$

Thus:

$(1 - g')\left(\hat{\beta}(\pi) - \beta(\pi)\right)$

Is $o_p(1)$

Since $1 - g' \neq 0$

$\hat{\beta}(\pi) \to_p \beta(\pi)$

B3. Probability limit of a general k-class estimator

Consider the residualized scalar structural model

$y = x\beta_0 + u$

and let the k-class estimator be

$\hat{\beta}_{k_n} = (x'W_n x)^{-1} x'W_n y, \qquad W_n := I - k_n M_z = P_z + (1 - k_n) M_z$

where $P_z = Z(Z'Z)^{-1}Z'$ and $M_z = I - P_z$

Let

$Q_{zz} = plim\, n^{-1}Z'Z,\ \ Q_{Zx} = plim\ \ n^{-1}Z'x,\ \ Q_{xx} = plim\, n^{-1}x'x$

$q_{zu} = plim\, n^{-1}Z'u,\ Q_{xu} = plim\, n^{-1}x'u$

with Q_{zz} positive definite. Define

$A := Q_{xz}Q_{zz}^{-1}Q_{zx}, \quad C := Q_{xx} - A$

$\delta := Q_{xz}Q_{zz}^{-1}q_{zu}, \quad \eta := Q_{xu} - \delta$

Here A is the asymptotic variation in x explained by the instrument space, C is the residual variation orthogonal to that space, δ captures the asymptotic distortion transmitted through the invalid instrument, and η is the remaining endogeneity component outside the instrument space.

Proposition B3.1

Suppose

$k_n \to k_0 \in \mathbb{R}$ and

$D(k_0) := A + (1 - k_0)C \neq 0$

Then

$plim\, \hat{\beta}_{k_n} = \beta_0 + B(k_0)$

where $B(k_0) = \frac{\delta + (1-k_0)\eta}{A + (1-k_0)C}$

Proof

Using $y = x\beta_0 + u$

$\hat{\beta}_{k_n} - \beta_0 = \frac{x'W_n u}{x'W_n x}$

Divide numerator and denominator by n. First,

$n^{-1}x'W_n x = n^{-1}x'P_z x + (1 - k_n)n^{-1}x'M_z x$

By standard LLN arguments,

$n^{-1}x'P_z x \to_p Q_{xz}Q_{zz}^{-1}Q_{zx,} = A$

And

$n^{-1}x'M_z x = n^{-1}x'x - n^{-1}x'P_z x \to_p Q_{xx} - A = C$

Hence

$n^{-1}x'W_n x \to_p A + (1 - k_0)C$

Similarly,

$n^{-1}x'W_n u = n^{-1}x'P_z u + (1 - k_0)n^{-1}x'M_z u$

Now,

$n^{-1}x'P_z u \to_p Q_{xz}Q_{zz}^{-1}q_{zu} = \delta$

And

$n^{-1}x'M_z u = n^{-1}x'u - n^{-1}x'P_z u \to_p Q_{xu} - \delta = \eta$

Therefore,

$n^{-1}x'W_n u \to_p \delta + (1 - k_0)\eta$

Applying Slutsky's theorem,

$plim(\hat{\beta}_{k_n} - \beta_0) = \frac{\delta + (1-k_0)\eta}{A + (1-k_0)C}$

This proves the result.

Corollary B3.1

If $k_n \to 1$, then

$plim\ \hat{\beta}_{k_n} = \beta_0 + \frac{\delta}{A}$

This includes TSLS directly and also LIML and Fuller under the standard fixed-K, strong-instrument asymptotic framework. Hence all such estimators share the same first-order asymptotic bias under instrument invalidity.

Remark

The formula above separates two sources of distortion. The term δ arises from instrument invalidity. The additional term $(1 - k_0)\eta$ appears whenever the estimator does not converge to TSLS. Thus, fixed-$k_0 \neq 1$ estimators may inherit both invalid-instrument bias and residual endogeneity bias.

Appendix B4. MSE dominance of FP-GMM over general k-class estimators

Let $\hat{\beta}_{FP}$ denote the fixed-point GMM estimator evaluated at the population correction parameter π^*, so that $\hat{\beta}_{FP}$ is consistent for β_0. Suppose

$$\sqrt{n}\big(\hat{\beta}_{FP} - \beta_0\big) \to_d N(0, V_{FP})$$

And

$$\sqrt{n}\left(\hat{\beta}_{k_n} - \beta_0 - B(k_0)\right) \to_d N\big(0, V_k(k_0)\big)$$

Proposition B4.1

Under the preceding conditions,

$$MSE\big(\hat{\beta}_{FP}\big) = \frac{V_{FP}}{n} + o(n^{-1})$$

Whereas

$$MSE\big(\hat{\beta}_{k_n}\big) = B(k_0)^2 + \frac{V_k(k_0)}{n} + o(n^{-1}),$$

Proof

For FP-GMM, consistency implies zero asymptotic bias, so

$$MSE\big(\hat{\beta}_{FP}\big) = \frac{V_{FP}}{n} + o(n^{-1})$$

For the k-class estimator,

$$\hat{\beta}_{k_n} - \beta_0 = B(k_0) + \left(\hat{\beta}_{k_n} - \beta_0 - B(k_0)\right)$$

Hence

$$MSE\big(\hat{\beta}_{k_n}\big) = B(k_0)^2 + \frac{V_k(k_0)}{n} + o(n^{-1})$$

Theorem B4.1

If $B(k_0) \neq 0$, then FP-GMM has lower asymptotic MSE than the k-class estimator for all sufficiently large n. A first-order dominance threshold is

$$n^*(k_0) = \max\left\{0, \frac{V_{FP} - V_k(k_0)}{B(k_0)^2}\right\}$$

Hence, for all sufficiently large $n > n^*(k_0)$

$$MSE\big(\hat{\beta}_{k_n}\big) > MSE\big(\hat{\beta}_{FP}\big)$$

Corollary B4.1

$k_n \to 1$, then

$B(1) = \frac{\delta}{A}$

So, FP-GMM eventually dominates any conventional k-class sequence whose asymptotic behavior is governed by $k_n \to 1$, provided $\delta \neq 0$

Appendix B5. Relation to the correction parameter

The population correction parameter π^* is defined by the orthogonality condition

$\mathbb{E}[(\hat{x} - \pi^* u)u] = 0$

Therefore,

$\pi^* = \frac{\mathbb{E}[\hat{x}u]}{\mathbb{E}[u^2]} = \frac{\delta}{\sigma_u^2}$

where $\sigma_u^2 = \mathbb{E}[u^2]$

Substituting $\delta = \pi^* \sigma_u^2$ into the general bias formula yields

$B(k_0) = \frac{\pi^* \sigma_u^2 + (1-k_0)\eta}{A + (1-k_0)C}$

In particular, for conventional k-class sequences with $k_n \to 1$

$B(1) = \frac{\pi^* \sigma_u^2}{A}$

so the dominance threshold becomes

$n^*(1) = \max\left\{0, \frac{(V_{FP} - V_k(1))A^2}{(\pi^*)^2 \sigma_u^4}\right\}$

This expression gives a direct interpretation of the dominance region. Larger admissible corrections, summarized by larger $|\pi^*|$, imply larger asymptotic k-class bias under exclusion violations and therefore reduce the sample size required for FP-GMM to dominate in mean squared error.

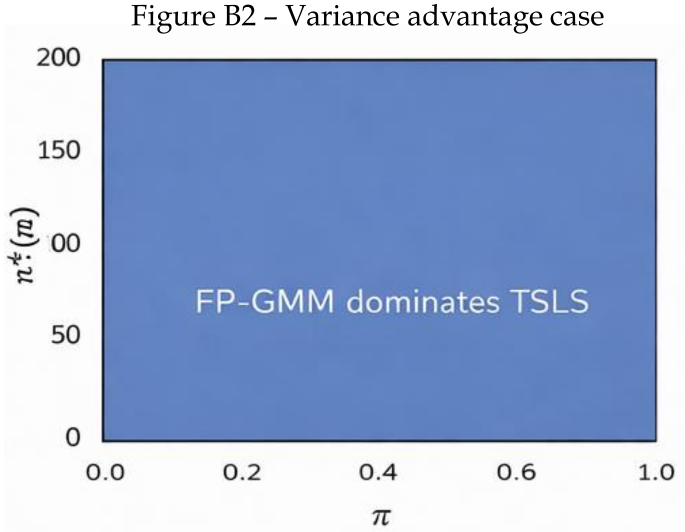
Implications

1. If $V_{FP} \leq V_{TSLS}$ dominates k-class for all sample sizes.
2. When the instruments are weak (ρ_{xz}^4 small) the dominance threshold decreases, implying that FP-GMM dominates earlier.
3. Larger violations of the exclusion restriction (larger π^*) increase the bias of k-class and strengthen the advantage of FP-GMM.

Figure B1 – Dominance threshold for FP-GMM

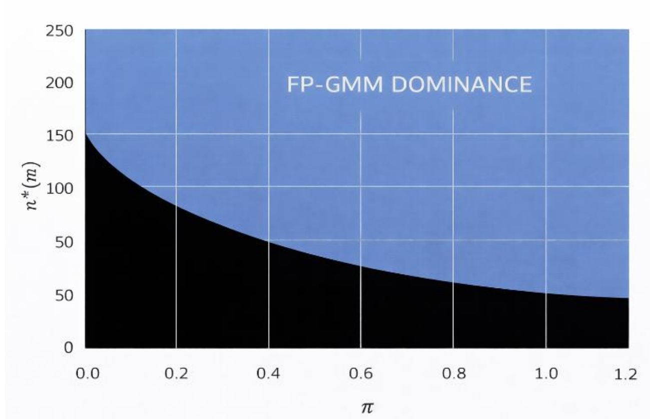
The figure plots the threshold function $n^*(\pi)$ defined by the equality $MSE\left(\hat{\beta}(k_n)\right) = MSE\left(\hat{\beta}_{FP}\right)$. The red curve gives the sample size threshold as a function of the correction parameter π. The blue region indicates the set (n, π) pairs for which FP-GMM dominates k-class in mean squared error.

Figure B2 – Variance advantage case



The figure illustrates the case $V_{FP} \leq V_{k-class}$. In this case the dominance threshold is non-positive for all π, so FP-GMM dominates k-class in mean squared error throughout the admissible parameter space. The blue region therefore covers the entire nonnegative sample-size domain.

Figure B3 – Instrument strength and the dominance threshold



The figure compares the threshold $n^*(\pi)$ under different levels of instrument relevant, indexed by the correlation ρ_{xy}. The blue area highlights the dominance region for the FP-GMM in the weak-instrument case.

Appendix C – Uniform Convergence and Asymptotic Normality of $\widehat{\boldsymbol{\beta}}$

C1. Preliminaries

Define

$$m_i(\beta;\pi;\gamma) \equiv z_i^{FP}(\beta;\pi;\gamma)(y_i - \beta x_i); \quad u_i(\beta) = y_i - \beta x_i; \quad z_i^{FP}(\beta;\pi;\gamma) = \hat{x}_i - \pi u_i(\beta)$$

Hence an explicit observable form is:

$$m_i(\beta;\pi;\gamma) = \big(\hat{x}_i - \pi u_i(\beta)\big)u_i(\beta) = \hat{x}_i u_i(\beta) - \pi u_i(\beta)^2$$

Let $\beta^* \equiv \beta(\pi)$ denote the population fixed point for the given π, and define $u_i^* \equiv u_i(\beta^*) = y_i - \beta^* x_i$. By definition of the fixed point.

$$M(\beta^*;\pi;\gamma) = \mathbb{E}[m_i(\beta^*;\pi)] = 0$$

C2. Derivative $\boldsymbol{\mathcal{K}(\pi)}$

Differentiate $m_i(\beta;\pi;\gamma)$ w.r.t β. First note:

$$\frac{\partial u_i(\beta)}{\partial \beta} = -x_i, \qquad \frac{\partial z_i^{FP}(\beta;\pi;\gamma)}{\partial \beta} = -\pi \frac{\partial u_i(\beta)}{\partial \beta} = \pi x_i$$

Product rule of $m_i(\beta;\pi;\gamma) \equiv z_i^{FP}(\beta;\pi;\gamma)u_i(\beta)$

$$\frac{\partial}{\partial \beta} m_i(\beta;\pi;\gamma) = \frac{\partial z_i^{FP}(\beta;\pi;\gamma)}{\partial \beta} u_i(\beta) + z_i^{FP}(\beta;\pi;\gamma)\frac{\partial u_i(\beta)}{\partial \beta}$$

Thus

$$\frac{\partial}{\partial \beta} m_i(\beta;\pi;\gamma) = \pi x_i u_i(\beta) - z_i^{FP}(\beta;\pi;\gamma)x_i$$

Taking expectations at $\beta = \beta^*$

$$\mathcal{K}(\pi) = \mathbb{E}\left[\frac{\partial}{\partial \beta} m_i(\beta^*;\pi;\gamma)\right] = \pi\mathbb{E}[xu^*] - \mathbb{E}[z^{FP}(\beta^*;\pi;\gamma)x]$$

Define $B(\beta^*;\pi;\gamma) \equiv \mathbb{E}[z^{FP}(\beta^*;\pi;\gamma)x]$

Then

$$\mathcal{K}(\pi) = \pi\mathbb{E}[xu^*] - B(\beta^*;\pi;\gamma)$$

Using Appendix A's fixed-point derivative identity

$$g'(\pi) = \frac{\pi\mathbb{E}[xu^*]}{B(\pi)}$$

We obtain a convenient factorization

$$\mathcal{K}(\pi) = B(\beta^*;\pi;\gamma)(g'(\pi) - 1) = -B(\beta^*;\pi;\gamma)\big(1 - g'(\pi)\big)$$

So

$\mathcal{K}(\pi)^2 = B(\beta^*;\pi;\gamma)^2\big(1-g'(\pi)\big)^2$

C3. Exact expression for $\boldsymbol{\Omega(\pi)}$

Under i.i.d sampling, the asymptotic variance of $\sqrt{n}\widehat{M_n}(\beta;\pi;\gamma)$ is:

$\Omega(\pi) = Var\big(m_i(\beta^*;\pi;\gamma)\big)$

Because $\mathbb{E}[m_i(\beta^*;\pi;\gamma)] = 0$ by (C2).

$\Omega(\pi) = \mathbb{E}[m_i(\beta^*;\pi;\gamma)^2]$

Using the explicit form (C1) at $\beta = \beta^*$:

$m_i(\beta^*;\pi;\gamma) = \hat{x}_i u_i^* - \pi(u_i^*)^2$

Therefore,

$\Omega(\pi) = \mathbb{E}[(\hat{x}_i u_i^* - \pi(u_i^*)^2)^2] = \mathbb{E}[(\hat{x}_i u_i^*)^2] - 2\pi\mathbb{E}[\hat{x}_i(u_i^*)^3] + \pi^2\mathbb{E}[(u_i^*)^4]$

This is the exact scalar expression in observable moments of $(\hat{x}_i u_i^*)$

C4. Proof of theorem 5.1

We use the standard Z-estimator expansion. The estimator satisfies

$0 = \widehat{M_n}\big(\hat{\beta};\pi;\hat{\gamma}\big)$

By a mean-value expansion around β^*

$0 = \widehat{M_n}(\beta^*;\pi;\hat{\gamma}) + \left[\frac{\partial}{\partial\beta}\widehat{M_n}\big(\tilde{\beta};\pi;\hat{\gamma}\big)\right]\big(\hat{\beta}-\beta^*\big)$

For some random $\tilde{\beta}$ between $\hat{\beta}$ and β^*. Rearrange:

$\sqrt{n}\big(\hat{\beta}-\beta^*\big) = -\left[\frac{\partial}{\partial\beta}\widehat{M_n}\big(\tilde{\beta};\pi;\hat{\gamma}\big)\right]^{-1}\sqrt{n}\widehat{M_n}(\beta^*;\pi;\hat{\gamma})$

By assumption 5.1

$\sqrt{n}\widehat{M_n}(\beta^*;\pi;\hat{\gamma}) \Rightarrow N\big(0,\Omega(\pi)\big)$

By consistency and a LLN for derivatives

$\frac{\partial}{\partial\beta}\widehat{M_n}\big(\tilde{\beta};\pi;\hat{\gamma}\big) \to_p \mathcal{K}(\pi)$

With $\mathcal{K}(\pi) \neq 0$. Hence, by Slutsky's theorem

$\sqrt{n}\big(\hat{\beta}-\beta^*\big) \Rightarrow N\left(0,\frac{\Omega(\pi)}{\mathcal{K}(\pi)}\right)$

Therefore

$Avar\left(\hat{\beta}(\pi)\right) = \frac{\Omega(\pi)}{B(\pi)^2\left(1-g'(\pi)\right)^2}$

This completes the proof.

C5. Feasible estimation of $\boldsymbol{\Omega(\pi)}$ and $\boldsymbol{\mathcal{K}(\pi)}$

A consistent estimation of $\Omega(\pi)$ is

$\widehat{\Omega}(\pi) = \frac{1}{n}\sum_{i=1}^{n}\widehat{m}_i(\pi)^2, \quad \widehat{m}_i(\pi) = m_i(\hat{\beta}(\pi);\pi)$

A consistent estimator of $\mathcal{K}(\pi)$ is the sample analogue of (C4)

$\widehat{\mathcal{K}}(\pi) = \frac{1}{n}\sum_{i=1}^{n}[\pi x_i\hat{u}_i(\pi) - z_i^{FP}(\pi)\, x_i]$

Where $\hat{u}_i(\pi) = y_i - \hat{\beta}(\pi)x_i, \quad z_i^{FP}(\pi) = \hat{x}_i - \pi\hat{u}_i$

Thus a feasible asymptotic variance estimator is

$Avar\left(\hat{\beta}(\pi)\right) = \frac{\widehat{\Omega}(\pi)}{\widehat{\mathcal{K}}(\pi)^2}$

Remark (Dependence / Cluster / HAC)

If observations are dependent or clustered, replace $\widehat{\Omega}(\pi)$ by the appropriate long-run variance estimator for $\Omega(\pi)$, keeping the same sandwich form $\widehat{\mathcal{K}}(\pi)^{-2}\widehat{\Omega}(\pi)$.

C6. Asymptotic normality at an operational boundary of the admissible region

The asymptotic normality result in Section 6 is stated for a fixed correction parameter π. The sample implementation developed in Section 7, however, selects π as an endpoint of the admissible correction set. Because the exact boundary of the stability region corresponds to loss of local contraction, inference is more naturally formulated for an operational boundary, defined as an endpoint of a strictly stable region bounded away from the instability frontier. This formulation matches Algorithm 7.1, which retains only values of π for which the iteration converges numerically and the sample contraction measure remains below a fixed tolerance.

Let $q(\pi)$ denote the population contraction coefficient evaluated at the fixed point $\beta(\pi)$, and define the trimmed admissible region

$\mathcal{A}_\varepsilon = \{\pi \in \Pi: q(\pi) \le 1 - \varepsilon\}$

For a given direction of correction, define the corresponding operational boundary by

$\pi_\varepsilon^- := inf\,\mathcal{A}_\varepsilon, \qquad \pi_\varepsilon^+ := sup\mathcal{A}_\varepsilon$

The sample implementation delivers an estimated admissible set $\hat{\mathcal{A}}_{n,\varepsilon}$ and the corresponding operational endpoint $\hat{\pi}_\varepsilon^{op}$. The final estimator is

$\hat{\beta}_{\varepsilon}^{op} \coloneqq \hat{\beta}(\hat{\pi}_{\varepsilon}^{op})$

Assumption C6.1 (Regular operational boundary)

The endpoint $\hat{\pi}_{\varepsilon}^{op}$ is unique, belongs to the interior of the parameter space Π, and satisfies

$q(\hat{\pi}_{\varepsilon}^{op}) \leq 1 - \varepsilon$

for some fixed $\varepsilon > 0$

Assumption C6.2 (Consistency of endpoint selection)

The estimated operational endpoint satisfies

$\hat{\pi}_{\varepsilon}^{op} \to_p \pi_{\varepsilon}^{op}$

Assumption C6.3 (Negligible first-step boundary error)

The endpoint selection error is asymptotically negligible at the $\sqrt{n}$ scale:

$\hat{\pi}_{\varepsilon}^{op} - \pi_{\varepsilon}^{op} = o_p(n^{-1/2})$

This condition is satisfied, for example, if the admissible set is searched over a deterministic grid whose mesh shrinks faster than $n^{-1/2}$, and the sample stability measure converges uniformly to its population counterpart.

Theorem C6.1 (Asymptotic normality at the operational boundary)

Suppose Assumptions 6.1–6.3 and C6.1–C6.3 hold. Then

$\sqrt{n}(\hat{\beta}_{\varepsilon}^{op} - \beta_{\varepsilon}^{op}) \to_d N\left(0, V(\pi_{\varepsilon}^{op})\right)$

where $V(\pi_{\varepsilon}^{op})$ is the asymptotic variance in Theorem 6.1 evaluated at the operational boundary π_{ε}^{op}

Proof

Write

$\hat{\beta}_{\varepsilon}^{op}(\hat{\pi}_{\varepsilon}^{op}) - \beta(\pi_{\varepsilon}^{op}) = [\hat{\beta}_{\varepsilon}^{op}(\hat{\pi}_{\varepsilon}^{op}) - \hat{\beta}_{\varepsilon}^{op}(\pi_{\varepsilon}^{op})] + [\hat{\beta}_{\varepsilon}^{op}(\pi_{\varepsilon}^{op}) - \beta(\pi_{\varepsilon}^{op})]$

By smoothness of the sample moment condition with respect to π, together with Assumption C6.3,

$\hat{\beta}_{\varepsilon}^{op}(\hat{\pi}_{\varepsilon}^{op}) - \hat{\beta}_{\varepsilon}^{op}(\pi_{\varepsilon}^{op}) = o_p(n^{-1/2})$

Therefore,

$\sqrt{n}\left(\hat{\beta}_{\varepsilon}^{op}(\hat{\pi}_{\varepsilon}^{op}) - \beta(\pi_{\varepsilon}^{op})\right) = \sqrt{n}\left(\hat{\beta}_{\varepsilon}^{op}(\hat{\pi}_{\varepsilon}^{op}) - \beta(\pi_{\varepsilon}^{op})\right) + o_p(1)$

Since π_{ε}^{op} is fixed and lies strictly inside the regular stability region, Theorem 6.1 applies at $\pi = \pi_{\varepsilon}^{op}$yielding

$$\sqrt{n}\left(\hat{\beta}_{\varepsilon}^{op}\left(\hat{\pi}_{\varepsilon}^{op}\right) - \beta\left(\pi_{\varepsilon}^{op}\right)\right) \to_d N\left(0, V\left(\pi_{\varepsilon}^{op}\right)\right)$$

The result follows by Slutsky's theorem.

Remark

The theorem applies to an operational boundary, not to the exact instability frontier $q(\pi) = 1$. This distinction is important. As emphasized in Section 6 and Section 7, variance and sensitivity increase sharply as the mapping approaches instability. At the exact boundary, the regular linearization underlying Theorem 6.1 may fail, whereas a trimmed boundary $q(\pi) \leq 1 - \varepsilon$ preserves a non-degenerate stability margin.

Appendix D – Regularization of $\boldsymbol{\pi}$

This appendix provides the formal arguments underlying the selection rule for the correction parameter π. In particular, it derives the derivative of the population fixed-point estimator with respect to π, establishes conditions under which $\beta(\pi)$ is monotone over the stability region.

Throughout the appendix, let $\beta(\pi)$ denote the population fixed-point estimator defined by $\beta(\pi) = g(\beta(\pi); \pi; \gamma)$. Which is equivalent to the moment condition $\mathbb{E}\left[z_i^{FP}(\beta, \pi; \gamma)(y_i - x_i\beta)\right] = 0$.

D1. Implicit characterization of $\boldsymbol{\beta(\pi)}$

Under the regularity conditions stated in Section 4, the moment function is continuously differentiable in a neighborhood of the solution. By the implicit function theorem

$$\frac{\partial\beta(\pi)}{\partial\pi} = -\frac{\frac{\partial\mathbb{E}\left[z_i^{FP}(\beta,\pi;\gamma)(y_i - x_i\beta)\right]}{\partial\pi}}{\frac{\partial\mathbb{E}\left[z_i^{FP}(\beta,\pi;\gamma)(y_i - x_i\beta)\right]}{\partial\beta}}|_{\beta=\beta(\pi)}$$

The partial derivatives of the moment function are

$$\frac{\partial\mathbb{E}\left[z_i^{FP}(\beta,\pi;\gamma)(y_i - x_i\beta)\right]}{\partial\pi} = -\mathbb{E}[u_i(\beta)^2]$$

$$\frac{\partial\mathbb{E}\left[z_i^{FP}(\beta,\pi;\gamma)(y_i - x_i\beta)\right]}{\partial\beta} = \frac{\partial\mathbb{E}\left[\hat{x}_i u_i(\beta) - \pi u_i(\beta)^2\right]}{\partial\beta}$$

$$\frac{\partial\mathbb{E}[u_i(\beta)]}{\partial\beta} = -x_i$$

Thus

$$\frac{\partial\mathbb{E}\left[\hat{x}_i u_i(\beta) - \pi u_i(\beta)^2\right]}{\partial\beta} = -\mathbb{E}[\hat{x}_i x_i] + 2\pi\mathbb{E}[u_i(\beta)x_i]$$

Therefore

$$\frac{\partial\beta(\pi)}{\partial\pi} = \frac{\mathbb{E}\left[u_i(\beta)^2\right]}{-\mathbb{E}[\hat{x}_i x_i] + 2\pi\mathbb{E}[u_i(\beta)x_i]}$$

D2. Sign of the derivative

Since $\mathbb{E}[u_i(\beta)^2] > 0$. The sign of $\frac{\partial\beta(\pi)}{\partial\pi}$ is fully determined by the denominator $D(\pi) = -\mathbb{E}[\hat{x}_i x_i] + 2\pi\mathbb{E}[u_i(\beta)x_i]$. Thus,

$$sign\left(\frac{\partial\beta(\pi)}{\partial\pi}\right) = sign\big(D(\pi)\big)$$

Accordingly, the derivative of $\beta(\pi)$ has a constant sign on any interval over which $D(\pi)$ does not change sign, this motivates the following condition.

Assumption E1. (Sign-preserving slope)

Over the population stability region $\mathcal{A}$, the quantity

$D(\pi) = -\mathbb{E}[\hat{x}_i x_i] + 2\pi\mathbb{E}[u_i(\beta)x_i]$

Does not change sign. Under assumption E2, the mapping $\pi \to \beta(\pi)$ is monotone over $\mathcal{A}$.

D3. Economic Interpretation

The denominator changes sign when the residual-based correction becomes large enough to offset the original identifying variation of the instrument. This corresponds to a regime in which the correction not only removes contamination but also begins to eliminate the signal required for identification. As a result, the correction path may lose monotonicity and become locally unstable. Restricting attention to regions where the denominator preserves its sign ensures that the estimator moves in a well-defined direction along the stable correction path.

The denominator admits a natural interpretation

$\mathbb{E}[\hat{x}_i x_i]$: baseline instrument relevance

$\mathbb{E}[u_i(\beta)x_i]$: residual endogeneity

Thus,

$D(\pi) = -relevance + correction \times endogeneity$

Monotonicity holds as long as the correction term does nor overturn the sign of the relevance component.

D4. Sufficient conditions for constant sign

We now provide sufficient conditions under which $D(\pi)$ does not change sign on a connected admissible region $\mathcal{A}_0$.

Assumption E.4.1 (Stable endogeneity sign)

$\mathbb{E}[u_i(\beta)x_i]$

does not change sign for $\pi \in \mathcal{A}_0$

Assumption E.4.2 (Bounded correction relative to relevance)

For all $\pi \in \mathcal{A}_0$

$\mathbb{E}[\hat{x}_i x_i] > 2|\pi||\mathbb{E}[u_i(\beta)x_i]|$

Proposition E.4.1 (Sign preservation)

Under Assumptions E.4.1–E.4.2,

$D(\pi)$

does not change sign on $\mathcal{A}_0$, and therefore $\beta(\pi)$ is monotone.

Proof

From (E.2),

$D(\pi) = -\mathbb{E}[\hat{x}_i x_i] + 2\pi\mathbb{E}[u_i(\beta)x_i]$

By Assumption E.4.2, the second term is strictly smaller in magnitude than the first, so the sign of $D(\pi)$ is determined by $\mathbb{E}[\hat{x}_i x_i]$, which is constant.

D5. Link with stability

The condition above is closely related to the stability restriction. Note that

$g'(\beta;\pi;\gamma) = \frac{\pi\mathbb{E}[xu(\beta)]}{B(\beta;\pi;\gamma)}$

As $|g'| \to 1$, the effective instrument strength $B(\beta;\pi;\gamma) \to 0$, and the denominator $D(\pi)$ approaches zero. Thus, violations of the sign condition correspond to regions where the mapping becomes nearly unstable.

D6. Implication

The sign condition therefore holds naturally within the interior of the admissible region and may fail only near its boundary. This implies that monotonicity is not an additional restriction, but rather a property that emerges from the same forces that govern stability.

D7. Auxiliary diagnostics for data-informed branch selection

The branch-selection problem in Section 7 arises because the structural moment condition determines the parameter of interest conditional on a correction parameter, but does not identify the direction of the correction. Stability restricts the admissible magnitude of the correction, yet the sign of the relevant branch must still be disciplined by external information. This appendix outlines how auxiliary empirical evidence may be used to inform that choice in applications.

The key point is that the sign of the exclusion violation is not identified from the main structural moment itself. Accordingly, the procedures discussed below should not be interpreted as recovering the sign from the data in a point-identified sense. Their role is more modest: they provide empirical diagnostics that may help discipline branch selection and reduce reliance on purely untestable sign restrictions.

A first source of information comes from placebo or negative-control outcomes. If the instrument displays systematic association with variables that should not be affected through the endogenous regressor, the sign of that association may reveal the likely direction of a direct channel. In settings where credible placebo variables are available, this provides a natural diagnostic for the sign of exclusion failure.

A second source of information comes from covariate imbalance patterns. Suppose the instrument is systematically correlated with observed covariates that are themselves known, on theoretical or institutional grounds, to affect the outcome in a particular direction. Then the sign of the instrument–covariate association may provide indirect evidence on the likely sign of the exclusion violation. This does not identify the structural error directly, but it can be informative about the direction in which the instrument is contaminated.

A third source of information comes from reduced-form decompositions. Since the reduced-form covariance satisfies

$$Cov(z, y) = \beta Cov(z, x) + Cov(z, u),$$

any benchmark or plausible range for β implies a corresponding sign for the residual component $Cov(z, u)$. In applications, this benchmark may come from auxiliary estimates, external evidence, or economically plausible bounds. Such calculations do not identify the sign uniquely, but they can indicate whether one branch is more consistent with the observed reduced-form relationship than the other.

A fourth source of information comes from subsample comparisons or heterogeneous environments. If the exclusion violation is believed to be stronger in some subsamples than in others, the direction of auxiliary sign diagnostics may be compared across groups. A branch that is supported consistently across empirically relevant subsamples is more credible than one that appears only under isolated specifications.

These diagnostics may be used jointly rather than in isolation. In practice, the researcher may construct a sign assessment based on the cumulative evidence from placebo outcomes, covariate imbalance, reduced-form decompositions, and subsample comparisons. When that evidence points clearly in one direction, the corresponding boundary of the admissible region becomes the natural branch for calibration. When the evidence is weak, mixed, or contradictory, the method should be interpreted more cautiously as a structured sensitivity analysis over admissible branches rather than as a uniquely disciplined point-estimation rule.

This perspective is fully consistent with the logic of the parameter-dependent moment framework developed in the main text. Stability determines how far the correction can be taken, while auxiliary empirical evidence may help discipline the direction in which that correction should operate. The procedure therefore remains partially theory-guided, but it need not rely exclusively on unstructured prior beliefs about the sign of exclusion failure.

D8. Proof of Theorem 7.1

The mean squared error of the estimator can be decomposed as

$$MSE(\lambda) = (\beta^*(\lambda) - \beta_0)^2 + \frac{1}{n}V(\lambda)$$

where the first term corresponds to squared asymptotic bias and the second term to variance. By assumption (i), the baseline estimator at $\lambda = 0$ is biased due to instrument endogeneity. By assumption (ii), along the admissible branch $\mathcal{A}_s$, the bias is weakly decreasing in λ, so that

$$\frac{d}{d\lambda}(\beta^*(\lambda) - \beta_0)^2 < 0$$

The admissible region $\mathcal{A}_s$ is compact, since it is defined by the stability condition

$$|g'(\beta^*(\lambda); \lambda)| < 1 - \varepsilon$$

By assumption (iii), the variance term $V(\lambda)$ is weakly increasing as λ approaches the boundary of $\mathcal{A}_s$. However, the variance term enters the MSE scaled by $1/n$, whereas the bias term enters at order one.

Therefore, for sufficiently large n, the reduction in squared bias dominates the increase in variance along the admissible path. It follows that the MSE is minimized at the largest admissible value of λ in the relevant direction, i.e., at the boundary of $\mathcal{A}_s$.

Hence,

$$\lambda^{MSE} \in \{\lambda_-, \lambda_+\}.$$

This establishes the result.

Appendix E - Relation with k-class estimators and imperfect instruments estimation

This appendix shows that the modified instrumental variable estimator can be written in a form closely related to the k-class family.

E1. Moment condition

The estimator is defined by

$$\mathbb{E}[z^{FP}(\beta;\pi;\gamma)(y-\beta x)] = 0$$

With

$$z^{FP}(\beta;\pi;\gamma) = \hat{x} - \pi(y-\beta x)$$

In sample form

$$\left(\hat{x} - \pi(y-\beta x)\right)'(y-\beta x) = 0$$

E2. Expansion

Expanding the expression

$$\hat{x}'(y-\beta x) - \pi(y-\beta x)'(y-\beta x) = 0$$

Rearranging,

$$\hat{x}'y - \beta\hat{x}'x - \pi(y-\beta x)'y + \pi\beta(y-\beta x)'x = 0$$

E3. Solving for the estimator

Collecting terms in β yields

$$\beta[\hat{x}'x - \pi(y-\beta x)'x] = \hat{x}'y - \pi(y-\beta x)'y$$

Therefore,

$$\hat{\beta}(\pi) = \frac{\hat{x}'y - \pi(y-\beta x)'y}{x'x - \pi(y-\beta x)'x} = \frac{x'\hat{x}'y - \pi x'(y-\beta x)'y}{x'\hat{x}'x - \pi x'(y-\beta x)'x}$$

E4. Classical k-class estimators

The k-class family (Theil 1958; Nagar 1959) provides a general class of instrumental variable estimators defined as:

$$\hat{\beta}_{K-CLASS}(k) = \frac{x'(I-kM_Z)y}{x'(I-kM_Z)x}$$

Where $M_Z = I - P_z$ and $P_z = z'(z'z)^{-1}z'$. Special cases include i) two-stage least squares $k = 1$ (Theil, 1958) and ii) limited-information maximum likelihood (LIML) $k = \hat{k}_{LIML}$ (Anderson and Rubin, 1949).

The LIML estimator is known to have smaller finite-sample bias that TSLS in many settings, particularly when instruments are weak.

Fuller (1977) proposed a modification of the LIML estimator that further reduces finite-sample bias. The Fuller estimator replaces the LIML parameter with $k_{Fuller} = \hat{k}_{LIML} - \alpha/(n-K)$ where K denotes the number of instruments and α is typically chosen as 1 or 4.

All estimators in this class operate by modifying the projection of the regressors onto the instrument space through the scalar parameter k, while leaving the instrument set unchanged.

E5. Connection with k-class estimators

The k-class estimator can be written as:

$$\hat{\beta}_{K-CLASS}(k) = \frac{x'P_z y - kx'M_z y}{x'P_z x - kx'M_z x}$$

Comparing both expressions $\hat{\beta}_{K-CLASS}(k)$ and $\hat{\beta}(k)$ shows that the proposed estimator introduces a correction term proportional to the structural residual rather than to the projection residual. Thus, the estimator behaves like a k-class estimator with **data-dependent correction determined by the fixed-point condition.**

E6. Stability-implied admissible range for instrument invalidity

This appendix clarifies how the parameter-dependent moment framework relates to imperfect-instrument approaches that impose exogenous bounds on the degree of instrument invalidity. In particular, we show that the stability region for the correction parameter π induces an implicit admissible range for the covariance $Cov(z,u)$. This provides a direct link between the parameter-dependent moment framework and approaches such as Conley, Hansen, and Rossi (2012) and Nevo and Rosen (2012), while also highlighting an important conceptual difference: in FP-GMM the admissible magnitude of instrument invalidity is not imposed by the researcher, but instead arises from the stability of the estimator.

Consider the linear model

$$y_i = \beta x_i + u_i$$

With candidate instrument z, and recall from the main test that the FP-GMM corrections is based on the instrument

$$z^{FP}(\beta) = \hat{x} - \pi u(\beta)$$

$$u(\beta) = y - \beta x$$

Where $\hat{x} = P_z x$.

For a given $\hat{\pi}$, the estimator is defined as the fixed point of the moment mapping, and admissible values of π are restricted to the stability region

$$\mathcal{A} \equiv \{\pi \in \Pi: |g'(\beta(\pi); \pi; \gamma)| < 1\}$$

E7. Mapping from π to instrument invalidity

In the just-identified scalar case,

$$\hat{x} = \delta z$$

$$\delta = \frac{\mathbb{E}[xz]}{\mathbb{E}[z^2]}$$

The population parameter satisfies

$$\pi^* = \frac{\mathbb{E}[\hat{x}u]}{\mathbb{E}[u^2]}$$

Substituting $\hat{x} = \delta z$

$$\pi^* = \frac{\delta\mathbb{E}[zu]}{\mathbb{E}[u^2]} = \frac{\mathbb{E}[xz]\mathbb{E}[zu]}{\mathbb{E}[z^2]\mathbb{E}[u^2]}$$

Hence

$$\mathbb{E}[zu] = \frac{\pi^*\mathbb{E}[z^2]\mathbb{E}[u^2]}{\mathbb{E}[xz]}$$

This establishes a one-to-one linear relation between the correction parameter π^* and the covariance $cov(z, u) = \mathbb{E}[zu]$, provided $\mathbb{E}[xz] \neq 0$.

Proposition F1 (stability-implied admissible range for $\boldsymbol{Cov(z, u)}$)

Suppose $\mathbb{E}[xz] \neq 0$, $\mathbb{E}[u^2] > 0$ and let $\mathcal{A} \subset \mathbb{R}$ denote the stability region of the FP-GMM mapping. Define the induced set

$$\Delta_{stable} = \left\{\delta_u \in \mathbb{R}: \delta_u = \frac{\pi\mathbb{E}[z^2]\mathbb{E}[u^2]}{\mathbb{E}[xz]} \; for\ some\ \pi \in \mathcal{A}\right\}$$

Then Δ_{stable} is the set of instrument-error covariances compatible with a stable stability-based estimator. In particular, if

$$\mathcal{A}_{stable} = [\pi_-, \pi_+]$$

Then

$$\Delta_{stable} = \left[\frac{\pi_-\mathbb{E}[z^2]\mathbb{E}[u^2]}{\mathbb{E}[xz]}, \frac{\pi_+\mathbb{E}[z^2]\mathbb{E}[u^2]}{\mathbb{E}[xz]}\right]$$

Up to the ordering induced by the sign of $\mathbb{E}[xz]$.

Proof

The relation

$\mathbb{E}[zu] = \frac{\pi^* \mathbb{E}[z^2]\mathbb{E}[u^2]}{\mathbb{E}[xz]}$

Is linear in π. Since $\mathcal{A}_{stable}$ is the set of admissible correction parameters, the image of $\mathcal{A}_{stable}$, under this linear map is precisely the set of admissible values of $\mathbb{E}[zu]$. If $\mathcal{A}_{stable}$ is an interval, then its image is also an interval.

Interpretation

Proposition F1 shows that the parameter-dependent moment framework implicitly restricts the degree of instrument invalidity. However, unlike Conley, Hansen, and Rossi (2012), who posit an exogenous bound on $\mathbb{E}[zu]$, or Nevo and Rosen (2012) , who impose restrictions on the relative endogeneity of z and x, the FP-GMM approach generates the admissible range for $\mathbb{E}[zu]$ through the dynamic stability of the estimator.

This distinction is important. In Conley-type approaches, the researcher specifies a set such as

$\mathbb{E}[zu] \in [\gamma_-, \gamma_+]$

In the parameter-dependent moment framework, no such range is imposed a priori. Instead, the admissible set is

$\mathbb{E}[zu] \in \Delta_{stable}$

Where Δ_{stable} is determined by the fixed-point stability condition. Thus, the stability-based estimator can be interpreted as replacing exogenous invalidity bounds with stability-implied invalidity bounds.

E8. Relation to Nevo and Rosen (2012)

Nevo and Rosen (2012) derive an identified set for β under the assumption that the instrument is less endogenous than the regressor, typically expressed as:

$|Corr(z, u)| \leq |Corr(x, u)|$

This assumption yields an interval for the structural parameter, often bounded by OLS and IV. By contrast, FP-GMM does not impose any restriction on the relative correlation between z and u. Instead, it restricts the admissible correction through the stability region $\mathcal{A}_{stable}$ which in turn induces the admissible covariance set Δ_{stable}.

In this sense, the two approaches differ along two dimensions:

1. Source of restriction.

Nevo–Rosen restrict the data-generating process through a correlation inequality. FP-GMM restricts the estimator through the contraction property of the fixed-point mapping.

2. Object delivered.

Nevo–Rosen characterize a set of admissible parameter values. FP-GMM selects a point estimate by solving the corrected moment condition inside the stable region.

Accordingly, the stability-based estimator should not be interpreted as a special case of Nevo–Rosen or Conley. Rather, it provides a distinct point-estimation framework in which the admissible magnitude of instrument invalidity is determined by estimator stability.

The previous result provides an alternative interpretation of the stability-based estimator. The estimator can be viewed as selecting a point from a stability-admissible class of imperfect-instrument corrections. The admissible class is not determined by researcher-imposed bounds on $Cov(z, u)$, but by whether the resulting moment mapping remains locally stable.

This strengthens the connection between fixed-point stability and identification developed in the main text. Stability does not merely guarantee existence and convergence of the estimator; it also implicitly determines which degrees of instrument invalidity are compatible with economically and statistically meaningful estimation.

Appendix F -Monte Carlo Evidence

Across all data generating processes considered, the estimated sequence $\beta(\lambda)$ exhibits a monotone pattern within the stability region. This empirical regularity is consistent with the theoretical monotonicity property of the correction path established in Section 7 and confirms that the estimator evolves in an ordered manner as the correction increases.

In these simulations, the sign of the normalized correction parameter λ is assumed to be correctly specified by the researcher, so that the selected boundary corresponds to the appropriate direction of adjustment. This assumption allows us to isolate the behavior of the estimator along the admissible path, abstracting from uncertainty regarding the direction of the exclusion violation.

The results also illustrate the relevance–exogeneity trade-off that underlies the parameter-dependent moment framework. As the correction increases, the estimator moves along a structured path while the effective strength of the instrument declines. This behavior becomes particularly pronounced near the stability boundary, where the estimator becomes increasingly sensitive as identification weakens.

F1. Opposing sources of instrument invalidity‡

This design introduces two channels of exclusion violation with opposite signs.

$y_i = \beta x_i + u_i, \quad x_i = \gamma z_i + 0.7 w_i + v_i$

$u_i = \delta_1 z_i + \delta_2 w_i + \varepsilon_i$

where

$v_i, \varepsilon_i, z_i \sim \mathcal{N}(0,1)$

$Corr(z, w) = \rho$

and mutually independent. The parameters δ_1 and δ_2 generate two opposite sources of instrument invalidity. Only z is used as instrument.

This design introduces two distinct channels through which the instrument violates the exclusion restriction, each operating with opposite sign. This configuration generates a more complex form of instrument invalidity than the baseline design, since the instrument simultaneously contains components that affect the estimator in different directions.

Despite this additional complexity, the estimated sequence $\beta(\lambda)$ exhibits a clear monotone pattern within the stability region across all specifications considered. As the correction parameter increases, the estimator evolves along an ordered path, reflecting the progressive adjustment induced by the residual-based correction.

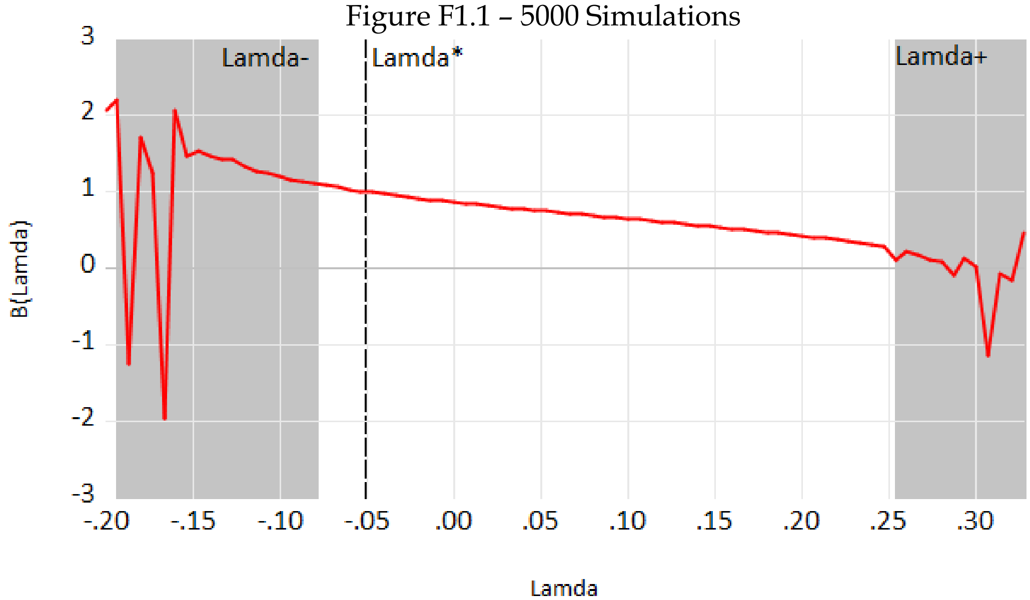
‡ G1 code.

This result is consistent with the theoretical analysis developed in Section 7. In particular, monotonicity ensures that the correction path is well-behaved and free of reversals, allowing the effect of increasing the correction to be interpreted in a transparent and systematic way.

A notable feature of this design is that the presence of opposing channels of invalidity increases the curvature of the correction path. The adjustment occurs more gradually because the residual-based correction must simultaneously offset two conflicting sources of correlation between the instrument and the structural error.

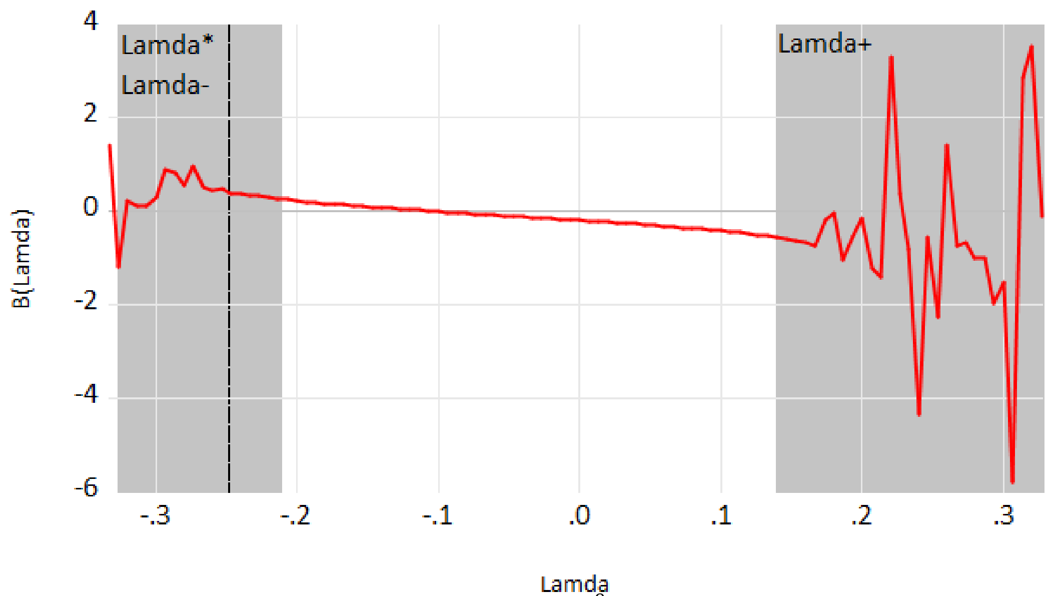
Nevertheless, the simulations show that the ordered structure predicted by the theory remains robust even under this more complex form of misspecification, and the relevance–exogeneity trade-off continues to govern the behavior of the estimator as the stability boundary is approached.

Figure F1.1 – 5000 Simulations



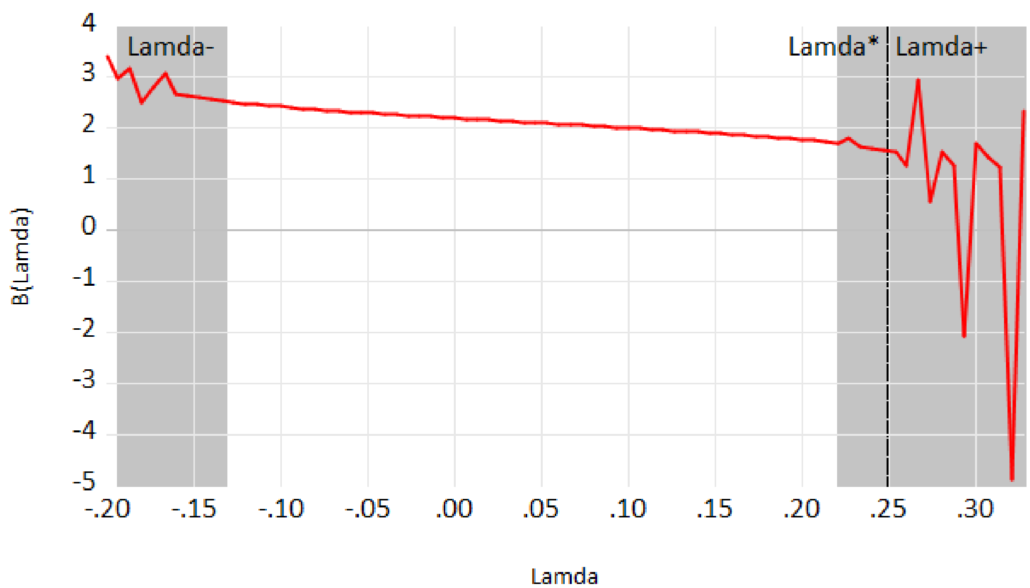
$n = 5000, \gamma = 1, \beta^* = 1, \delta_1 = 0.8, \delta_2 = -2, \varepsilon_\beta = 10^{-9}, \varepsilon_{g'} = 0.2, FP - Iterations = 10, \rho = 0.5$

Figure F1.2 – 5000 Simulations



$n = 5000, \gamma = 1, \beta^* = 1, \delta_1 = -0.8, \delta_2 = -2, \varepsilon_\beta = 10^{-9}, \varepsilon_{g'} = 0.2, FP - Iterations = 10, \rho = 0.5$

Figure F1.3 – 5000 Simulations



$n = 5000, \gamma = 1, \beta^* = 1, \delta_1 = 0.8, \delta_2 = 2, \varepsilon_\beta = 10^{-9}, \varepsilon_{g'} = 0.2, FP - Iterations = 10, \rho = 0.5$

Table F1 reports the finite-sample performance of the estimators under the opposing-invalidity design. Several patterns emerge. First, conventional IV estimators exhibit large biases when the instrument contains strong exclusion violations. Because both channels of invalidity contaminate the instrument, estimators such as TSLS, LIML, and Fuller inherit substantial asymptotic bias.

Second, the FP estimator systematically reduces the magnitude of this bias across all configurations. Although the correction cannot fully eliminate bias when the instrument is severely contaminated, the residual-based adjustment significantly mitigates the distortion relative to conventional IV estimators.

Third, the gains achieved by the FP estimator remain stable across sample sizes. This pattern reflects the structural nature of the correction mechanism: the improvement arises from modifying the instrument rather than from improved sampling precision.

Overall, the results confirm that the FP estimator remains effective even when instrument invalidity arises through multiple and potentially conflicting channels.

Table F1 – Monte Carlo results: Opposing sources of instrument invalidity

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} = 0.2$, $\varepsilon_b = 10^{-9}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$\delta_1 = 0.8 \;\; \delta_2 = -2$			$\delta_1 = -0.8 \;\; \delta_2 = -2$			$\delta_1 = 0.8 \;\; \delta_2 = 2$		
$n = 1000$	*FP-GMM*	0.03	0.05	0.04	-0.89	0.89	0.03	0.91	0.91	0.03
	OLS	-0.45	10.09	0.03	-1.05	23.47	0.03	1.05	23.48	0.03
	TSLS	-0.13	3.13	0.04	-1.20	26.85	0.04	1.20	26.88	0.04
	LIML	-0.13	3.13	0.04	-1.20	26.85	0.04	1.20	26.88	0.04
	Fuller	-0.14	3.36	0.04	-1.19	26.73	0.04	1.20	26.76	0.04
$n = 5000$	*FP-GMM*	0.03	0.03	0.02	-0.89	0.89	0.01	0.90	0.90	0.01
	OLS	-0.45	10.06	0.01	-1.05	23.47	0.01	1.05	10.49	0.01
	TSLS	-0.13	2.99	0.02	-1.20	26.81	0.02	1.20	11.99	0.01
	LIML	-0.13	2.99	0.02	-1.20	26.81	0.02	1.20	11.99	0.01
	Fuller	-0.14	3.23	0.02	-1.19	26.69	0.02	1.19	11.94	0.01
$n = 10000$	*FP-GMM*	0.03	0.03	0.01	-0.89	0.89	0.01	0.91	0.91	0.01
	OLS	-0.45	10.06	0.01	-1.05	23.48	0.01	1.05	10.51	0.01
	TSLS	-0.13	3.00	0.01	-1.20	26.83	0.01	1.20	12.01	0.01
	LIML	-0.13	3.00	0.01	-1.20	26.83	0.01	1.20	12.01	0.01
	Fuller	-0.14	3.25	0.01	-1.19	26.72	0.01	1.20	11.96	0.01

Monte Carlo results based on 500 replications under a data generating process with two sources of instrument invalidity with opposite signs. Results report bias, root mean squared error (RMSE), and standard deviation for FP-GMM, OLS, TSLS, LIML, and Fuller estimators. The experiment illustrates how the residual-based correction mitigates bias even when instrument invalidity arises from multiple channels.

F2. Group Heterogeneity [§]

We consider a discrete instrument with heterogeneous invalidity across groups.

$y_i = \beta x_i + u_i, \quad x_i = \gamma z_i + v_i$

With:

$u_i = \delta z + \eta(z_i x_i) + \varepsilon_i$

[§] G2 code.

With

$z_i, \nu_i, \varepsilon_i \sim \mathcal{N}(0,1)$

Mutually independent.

This design introduces a discrete instrument with heterogeneous invalidity across groups. The purpose of the design is to make the exclusion violation less transparent by allowing the relationship between the instrument and the structural error to vary across latent subpopulations. The interaction structure therefore generates a setting in which the direction and magnitude of instrument invalidity differ across groups.

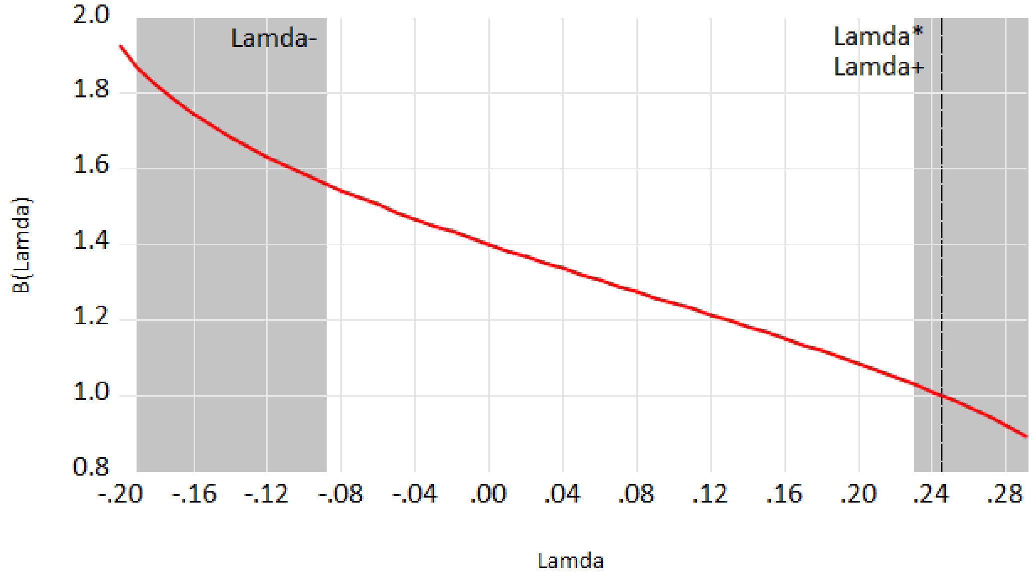
Despite this more complex source of heterogeneity, the estimated sequence $\beta(\lambda)$ continues to exhibit a stable and ordered pattern within the admissible region. As the correction parameter increases, the estimator evolves smoothly along the correction path, reflecting the progressive adjustment induced by the residual-based correction. Compared to the Gaussian benchmark, the path is typically flatter and may display weaker curvature, reflecting the fact that group-specific invalidity partially masks the aggregate direction of distortion.

This behavior is economically and statistically intuitive. When exclusion violations differ across groups, the residual-based correction has less direct information about a common direction of endogeneity. As a result, the adjustment of the estimator becomes more gradual. Nevertheless, the simulations show that the correction path remains well-behaved under this heterogeneity, and the relevance–exogeneity trade-off continues to govern the evolution of the estimator as the stability boundary is approached

Figure F2.1 – 5000 Simulations

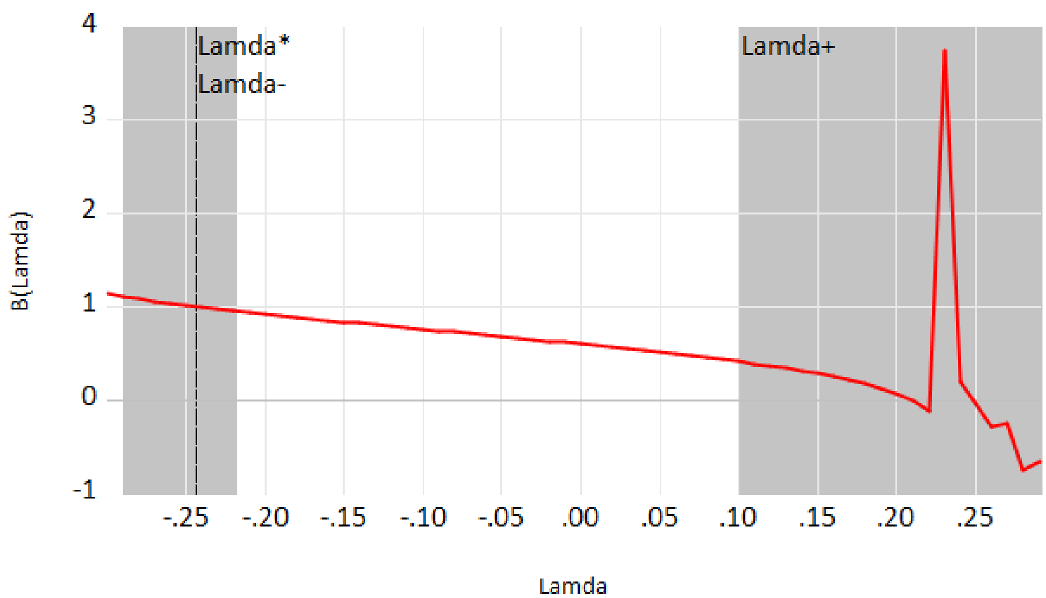
$n = 5000, \gamma = 1, \beta^* = 1, \delta = 0.4, \eta = 0.4, \varepsilon_\beta = 10^{-10}, \varepsilon_{g'} = 0.05, FP - Iterations = 10$

Figure F2.2 – 5000 Simulations



$n = 5000, \gamma = 1, \beta^* = 1, \delta = 0.4, \eta = -0.4, \varepsilon_\beta = 10^{-10}, \varepsilon_{g'} = 0.05, FP - Iterations = 10$

Figure F2.3 – 5000 Simulations



$n = 5000, \gamma = 1, \beta^* = 1, \delta = -0.4, \eta = -0.4, \varepsilon_\beta = 10^{-10}, \varepsilon_{g'} = 0.05, FP - Iterations = 10$

Table F2 reports the finite-sample performance of the estimators under group-specific instrument invalidity.

Three features stand out. First, conventional IV estimators continue to inherit substantial bias because they aggregate valid and invalid components of the instrument without accounting for the heterogeneity in the exclusion violation. Second, the FP estimator systematically lowers the bias across the reported configurations, although the gains are smaller than in the baseline Gaussian design. This is expected, since heterogeneous invalidity weakens the ability of the residual-based correction to identify a single dominant direction of adjustment.

Third, the overall performance of the FP estimator remains stable across sample sizes, suggesting that the main improvement comes from the correction mechanism itself rather than from purely finite-sample effects. In this sense, the design illustrates both the strength and the limitation of the method: the FP estimator continues to reduce bias when invalidity is heterogeneous, but the magnitude of the gain depends on how clearly the average direction of instrument contamination can be recovered from the data.

Table F2 – Monte Carlo results: Group heterogeneity in instrument invalidity

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} = 0.05$, $\varepsilon_b = 10^{-10}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$\delta = 0.4$ $\eta = 0.4$			$\delta = 0.4$ $\eta = -0.4$			$\delta = -0.4$ $\eta = -0.4$		
$n = 1000$	*FP-GMM*	0.10	0.13	0.08	0.11	0.11	0.04	-0.10	0.10	0.04
	OLS	0.20	4.50	0.04	0.20	4.47	0.04	-0.20	2.07	0.03
	TSLS	0.40	8.95	0.06	0.40	8.98	0.06	-0.40	4.03	0.04
	LIML	0.40	8.95	0.06	0.40	8.98	0.06	-0.40	4.03	0.04
	Fuller	0.39	8.78	0.06	0.39	8.80	0.06	-0.39	3.96	0.04
$n = 5000$	*FP-GMM*	0.10	0.11	0.05	0.10	0.11	0.05	-0.09	0.10	0.02
	OLS	0.20	4.48	0.02	0.20	2.02	0.02	-0.20	2.01	0.02
	TSLS	0.40	8.94	0.03	0.40	3.99	0.03	-0.40	4.02	0.03
	LIML	0.40	8.94	0.03	0.40	3.99	0.03	-0.40	4.02	0.03
	Fuller	0.39	8.76	0.03	0.39	3.91	0.03	-0.39	3.94	0.03
$n = 10000$	*FP-GMM*	0.10	0.11	0.04	0.11	0.11	0.03	-0.10	0.10	0.01
	OLS	0.20	4.50	0.01	0.20	2.01	0.01	-0.20	2.03	0.01
	TSLS	0.40	8.97	0.02	0.40	4.00	0.02	-0.40	4.03	0.02
	LIML	0.40	8.97	0.02	0.40	4.00	0.02	-0.40	4.03	0.02
	Fuller	0.39	8.80	0.02	0.39	3.92	0.02	-0.39	3.95	0.02

Monte Carlo results based on 500 replications for a design with a discrete instrument and group-specific exclusion violations. The table reports bias, RMSE, and standard deviation for FP-GMM, OLS, TSLS, LIML, and Fuller estimators. The design is intended to challenge the monotonicity of the bias criterion by introducing heterogeneous invalidity across groups. The results show that the FP estimator continues to reduce bias relative to conventional IV estimators, although the gains are smaller than in the baseline design.

F3. Binary instrument invalidity**

To evaluate robustness to nonlinear exclusion violations, we consider

** G3 Code

$z^* = \xi \quad \xi \sim \mathcal{N}(0,1) \quad z = 1\{z^* > 0\}$

$y_i = \beta x_i + u_i, \quad x_i = \gamma(z_i) z_i + v_i$

$u_i = \delta(z_i) z + \varepsilon_i$

Where:

$$\gamma(z) = \begin{cases} \gamma_0 & z = 0 \\ \gamma_1 & z = 1 \end{cases} \quad \delta(z) = \begin{cases} \delta_0 & z = 0 \\ \delta_1 & z = 1 \end{cases}$$

$\varepsilon_i, z_i \sim \mathcal{N}(0,1)$

Mutually independent.

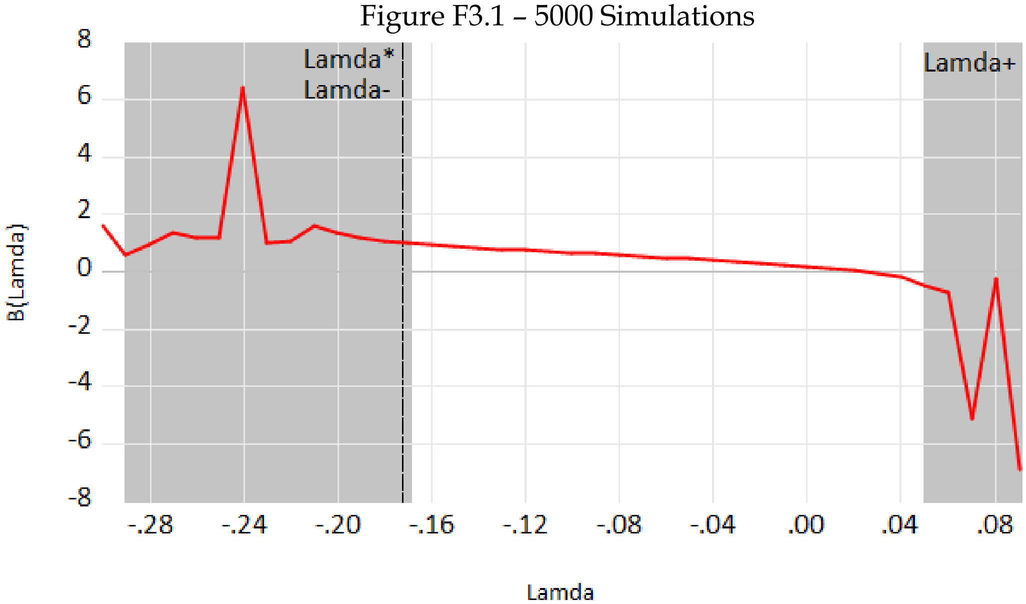
The third design evaluates robustness to nonlinear exclusion violations by allowing the instrument to affect the structural equation through a binary channel. This generates a form of invalidity that is discrete and inherently nonlinear, departing more sharply from the linear structure considered in the main theoretical development.

Even in this setting, the estimated sequence $\beta(\lambda)$ retains the same qualitative ordered behavior within the admissible region. As the correction parameter increases, the estimator evolves along a stable path, although the trajectory may become less smooth and exhibit sharper local changes, reflecting the discontinuous nature of the underlying source of invalidity.

This behavior is informative for the interpretation of the theory. While the analytical results in Section 7 are developed under a linear framework, the simulation evidence shows that the correction path remains well-behaved even when the exclusion violation arises through a nonlinear mechanism. This suggests that the structured evolution of the estimator is not restricted to linear environments, but extends to more general forms of misspecification.

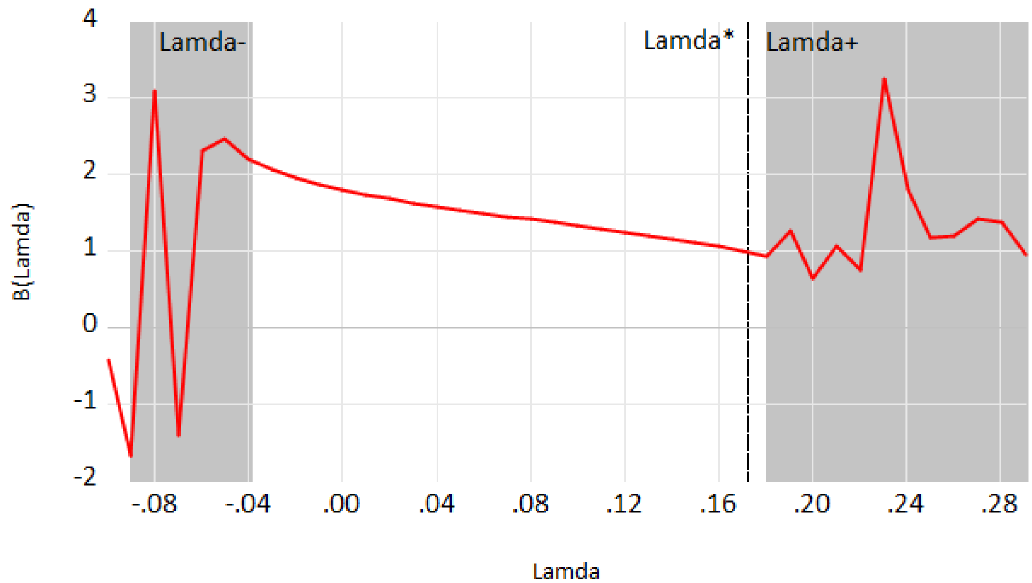
Overall, the results confirm that the relevance–exogeneity trade-off continues to govern the behavior of the estimator, even in the presence of nonlinear and discrete sources of instrument invalidity.

Figure F3.1 – 5000 Simulations



$n = 5000, \gamma_1 = 1, \gamma_0 = 0, \beta^* = 1, \delta_1 = -0.8, \delta_0 = 0, \varepsilon_\beta = 10^{-10}, \varepsilon_{g'} = 0.025, FP - Iterations = 20$

Figure F3.2 – 5000 Simulations



$n = 5000, \gamma_1 = 1, \gamma_0 = 0, \beta^* = 1, \delta_1 = 0.8, \delta_0 = 0, \varepsilon_\beta = 10^{-10}, \varepsilon_{g'} = 0.025, FP - Iterations = 20$

Table F3 presents the corresponding Monte Carlo evidence for the binary-invalidity design. The results show that the FP estimator continues to outperform conventional IV estimators in terms of bias across the reported specifications. The improvement is especially notable because the source of invalidity is nonlinear and therefore less directly aligned with the linear correction used by the estimator. In contrast, TSLS, LIML, and Fuller remain strongly affected by the invalid component of the instrument, as they continue to treat the contaminated instrument as if it satisfied the exclusion restriction.

At the same time, the gains from the correction are not uniform across all cases. Because the binary channel creates abrupt variation in the exclusion violation, the correction becomes less precise than in the Gaussian benchmark. Nevertheless, the FP estimator still achieves a clear reduction in bias, which indicates that the residual-based correction remains informative even when invalidity is discrete and nonlinear.

Overall, this design provides additional support for the central message of the paper: the monotone bias-reduction mechanism of the FP estimator is not confined to linear-Gaussian environments, but extends to settings where instrument invalidity is highly nonlinear.

Table F3 – Monte Carlo results: Binary instrument invalidity

500 simulations, 20 fixed-point iterations, $\varepsilon_{g'} = 0.025$, $\varepsilon_b = 10^{-10}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$\gamma_1 = 1, \gamma_0 = 0, \delta_1 = -0.8, \delta_0 = 0$			$\gamma_1 = 1, \gamma_0 = 0, \delta_1 = 0.8, \delta_0 = 0$		
$n = 1000$	*FP-GMM*	-0.67	0.67	0.07	0.72	0.73	0.08
	OLS	-0.16	1.65	0.03	0.16	1.61	0.03

	TSLS	-0.80	8.02	0.08	0.80	8.07	0.08
	LIML	-0.80	8.02	0.08	0.80	8.07	0.08
	Fuller	-0.74	7.41	0.07	0.74	7.45	0.07
$n = 5000$	*FP-GMM*	-0.67	0.68	0.03	0.74	0.74	0.04
	OLS	-0.16	1.61	0.01	0.16	1.60	0.01
	TSLS	-0.80	7.99	0.04	0.80	8.03	0.04
	LIML	-0.80	7.99	0.04	0.80	8.03	0.04
	Fuller	-0.74	7.40	0.03	0.74	7.43	0.04
$n = 10000$	*FP-GMM*	-0.69	0.69	0.02	0.74	0.74	0.03
	OLS	-0.16	1.61	0.01	0.16	1.61	0.01
	TSLS	-0.81	8.06	0.03	0.80	8.01	0.03
	LIML	-0.81	8.06	0.03	0.80	8.01	0.03
	Fuller	-0.75	7.45	0.02	0.74	7.41	0.02

Monte Carlo results based on 500 replications for a design in which the exclusion restriction fails through a binary and nonlinear channel. The table reports bias, RMSE, and standard deviation for FP-GMM, OLS, TSLS, LIML, and Fuller estimators. Despite the nonlinear structure of the exclusion violation, the FP estimator continues to reduce bias relative to conventional IV estimators, supporting the robustness of the monotone bias-correction mechanism.

F4. One Imperfect Instrument and Multiple Valid Instruments[††]

$y_i = \beta x + u_i, \quad x_i = \gamma_1 z_{1i} + \sum_{j=1}^{K} \gamma_2 z_{ji} + v_i, \quad u_i = \delta z_{1i} + \varepsilon_i, \quad z_i, v_i, \varepsilon_i \sim \mathcal{N}(0,1)$

This design introduces multiple instruments, only one of which satisfies the exclusion restriction. This configuration mimics empirical environments in which some instruments may be contaminated while others remain valid.

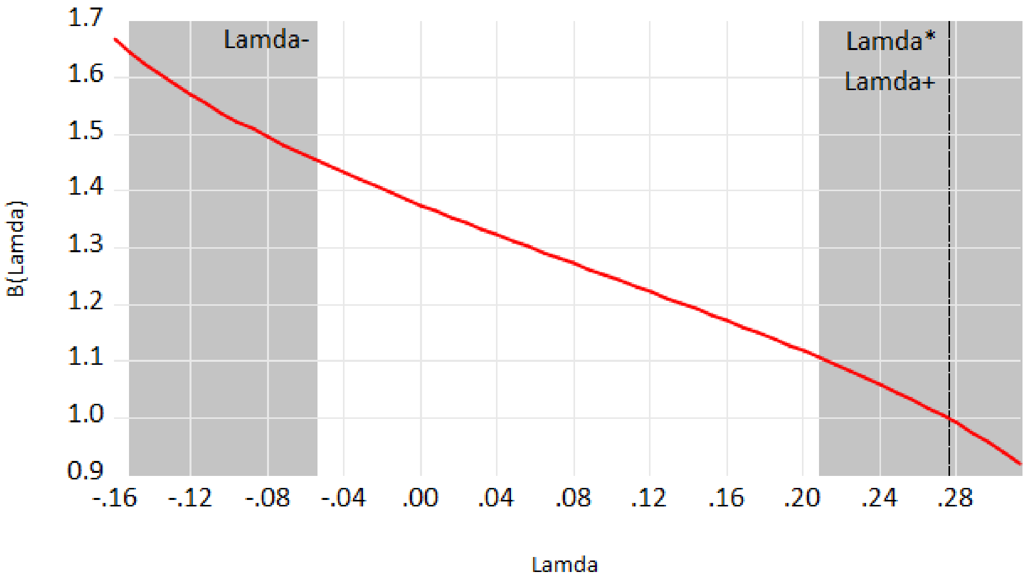
The estimated sequence $\beta(\lambda)$ again exhibits a stable and ordered pattern within the admissible region. As the correction parameter increases, the residual-based adjustment progressively attenuates the component of the instrument that is correlated with the structural error, while preserving the identifying variation provided by the valid instrument.

The evolution of the estimator along the correction path reflects the interaction between these two forces. The presence of a valid instrument provides a strong source of identifying variation, which allows the estimator to adjust more rapidly as the correction increases. As a result, the path typically becomes steeper relative to the baseline case, indicating that small changes in the correction parameter induce larger changes in the estimate.

These findings show that the structured behavior of the correction path extends naturally to settings with multiple instruments and heterogeneous validity. The relevance–exogeneity trade-off continues to govern the evolution of the estimator, while the stability boundary delineates the limit beyond which further correction would compromise identification.

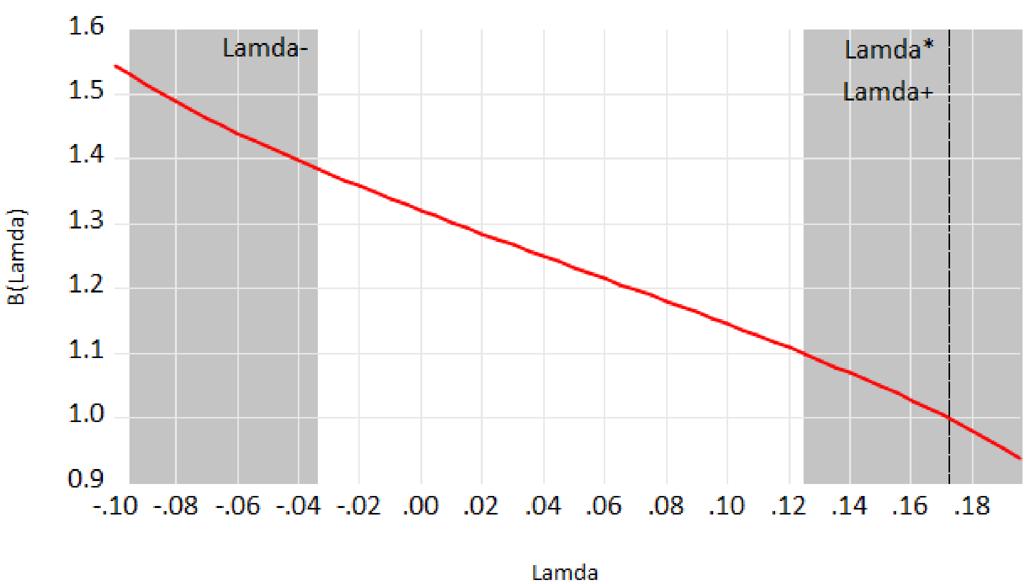
[††] G4 Code.

Figure F4.1 – 5000 Simulations



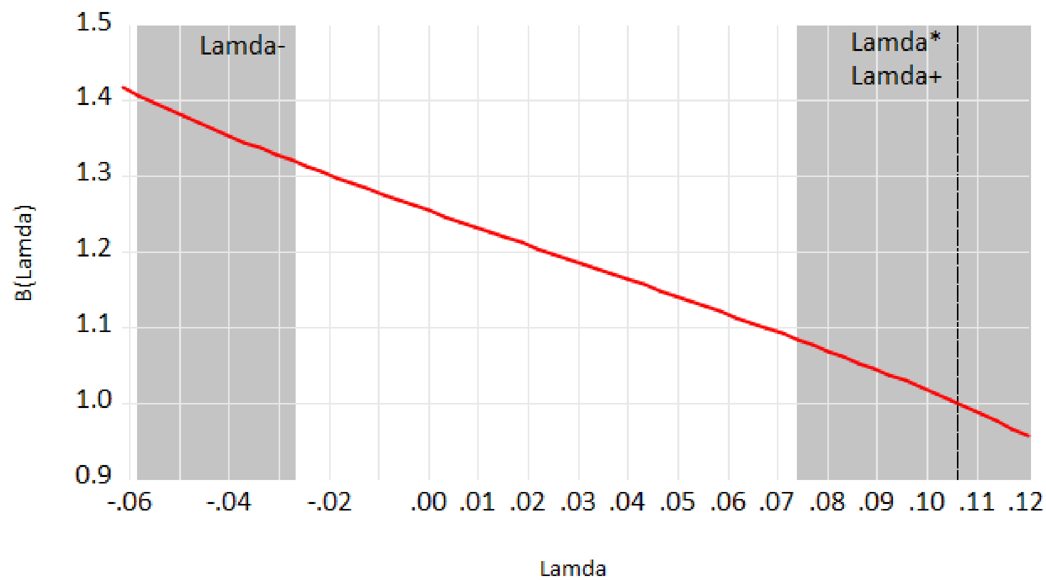
$n = 5000, K = 2, \delta = 0.4, \gamma_1 = 1, \gamma_2 = 0.25, \varepsilon_\beta = 10^{-12}, \varepsilon_{g'} = 0.01, FP - Iterations = 10$

Figure F4.2 – 5000 Simulations



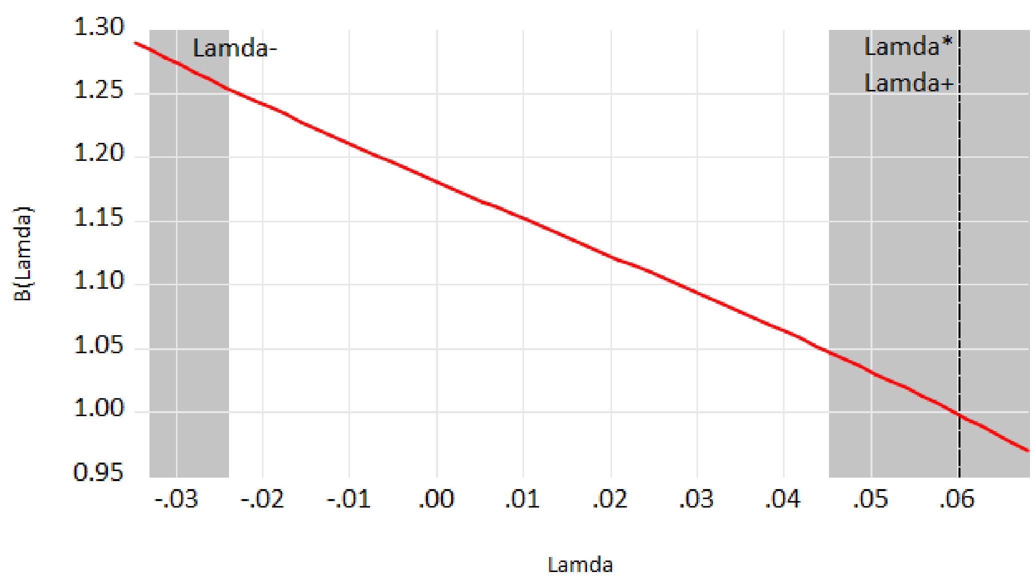
$n = 5000, K = 5, \delta = 0.4, \gamma_1 = 1, \gamma_2 = 0.25, \varepsilon_\beta = 10^{-12}, \varepsilon_{g'} = 0.01, FP - Iterations = 10$

Figure F4.3 – 5000 Simulations



$n = 5000, K = 10, \delta = 0.4, \gamma_1 = 1, \gamma_2 = 0.25, \varepsilon_\beta = 10^{-12}, \varepsilon_{g'} = 0.01, FP - Iterations = 10$

Figure F4.4 – 5000 Simulations



$n = 5000, K = 20, \delta = 0.4, \gamma_1 = 1, \gamma_2 = 0.25, \varepsilon_\beta = 10^{-12}, \varepsilon_{g'} = 0.01, FP - Iterations = 10$

Table F4 summarizes the finite-sample properties of the estimators. The results show that when one instrument is invalid, conventional IV estimators inherit bias from the contaminated instrument because the projection step combines information from both instruments.

The FP estimator reduces this bias by partially removing the component of the instrument correlated with the structural error. The bias reduction is particularly pronounced when the invalid instrument contributes strongly to the first-stage variation.

At the same time, the variance of the estimator remains comparable to that of standard IV estimators, reflecting the fact that the correction operates directly on the instrument rather than altering the structure of the regression.

These results highlight an important advantage of the FP approach: the estimator can partially correct instrument contamination even when valid and invalid instruments coexist.

Table F4 – Monte Carlo results: Multiple instruments with partial invalidity

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} = 0.01$, $\varepsilon_b = 10^{-12}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$K = 2$			$K = 5$			$K = 10$			$K = 20$		
$n = 1000$	*FP-GMM*	0.18	0.18	0.02	0.15	0.16	0.02	0.13	0.13	0.02	0.09	0.09	0.02
	OLS	0.20	0.20	0.02	0.18	0.18	0.02	0.15	0.16	0.02	0.13	0.13	0.02
	TSLS	0.38	0.38	0.03	0.32	0.32	0.03	0.25	0.25	0.03	0.18	0.18	0.02
	LIML	0.38	0.38	0.03	0.33	0.33	0.03	0.26	0.27	0.03	0.19	0.19	0.02
	FULLER	0.38	0.38	0.03	0.32	0.32	0.03	0.25	0.25	0.03	0.18	0.18	0.02
$n = 5000$	*FP- GMM*	0.18	0.18	0.01	0.15	0.15	0.01	0.13	0.13	0.01	0.09	0.09	0.01
	OLS	0.20	0.20	0.01	0.18	0.18	0.01	0.16	0.16	0.01	0.13	0.13	0.01
	TSLS	0.38	0.38	0.02	0.32	0.32	0.01	0.26	0.26	0.01	0.18	0.18	0.01
	LIML	0.38	0.38	0.02	0.33	0.33	0.01	0.27	0.27	0.01	0.19	0.19	0.01
	FULLER	0.38	0.38	0.02	0.32	0.32	0.01	0.26	0.26	0.01	0.18	0.18	0.01
$n = 10000$	*FP- GMM*	0.17	0.17	0.01	0.15	0.15	0.01	0.13	0.13	0.01	0.09	0.09	0.01
	OLS	0.19	0.19	0.01	0.18	0.18	0.01	0.16	0.16	0.01	0.13	0.13	0.01
	TSLS	0.38	0.38	0.01	0.32	0.32	0.01	0.26	0.26	0.01	0.18	0.18	0.01
	LIML	0.38	0.38	0.01	0.33	0.33	0.01	0.27	0.27	0.01	0.19	0.19	0.01
	FULLER	0.38	0.38	0.01	0.32	0.32	0.01	0.26	0.26	0.01	0.18	0.18	0.01

Monte Carlo results for a design with multiple instruments where only a subset satisfies the exclusion restriction. The table reports bias, RMSE, and standard deviation for the competing estimators across sample sizes. The FP estimator reduces bias by correcting the contaminated instrument while preserving the identifying variation provided by the valid instrument.

Taken together, the additional designs indicate that the monotonicity of the bias criterion is not a fragile feature of the Gaussian benchmark, but a robust empirical regularity of the FP-GMM mapping across a broad class of nonlinear and heterogeneous exclusion violations.

F5. Small-sample performance under varying instrument invalidity

This appendix evaluates the finite-sample performance of the FP-GMM estimator under small sample sizes and varying degrees of instrument invalidity. The objective is to assess how the stability-based correction behaves when sampling variability is non-negligible and the bias–variance trade-off becomes more pronounced.

We consider the same linear Gaussian data-generating process used in Table 1 of the main text. The structural model, first-stage relationship, and error distributions are identical, ensuring direct comparability with the baseline Monte Carlo design. The only modifications are the sample size and the magnitude of the exclusion violation.

Specifically, we examine sample sizes $n \in \{50,100,200\}$, which represent empirically relevant small-sample environments. For each sample size, we vary the degree of instrument invalidity by setting $\delta \in \{0.4,0.8,1.2,2.0\}$, where δ governs the strength of the correlation between the instrument and the structural error. This parametrization allows us to trace the performance of the estimator across progressively more severe violations of the exclusion restriction.

For each configuration, we report Monte Carlo averages of bias, variance, and mean squared error for the FP-GMM estimator and benchmark estimators, including OLS, TSLS, LIML, and Fuller. The FP-GMM estimator is implemented using the same stability-based calibration procedure described in Section 7, with the admissible region determined by convergence and local stability of the fixed-point mapping.

The results in Table F.5 show that the stability-based estimator systematically improves mean squared error relative to conventional IV estimators across all sample sizes and degrees of instrument invalidity. While TSLS, LIML, and Fuller estimators exhibit substantial bias that persists even as sample size increases, the FP-GMM estimator reduces bias by trading off instrument relevance for improved exogeneity along the admissible correction path.

This bias reduction comes at the cost of moderately higher variance, particularly in small samples. However, the simulations indicate that the variance increase is dominated by the reduction in squared bias, resulting in uniformly lower RMSE for the FP-GMM estimator. This pattern is consistent with the theoretical characterization in Section 7, where the stability boundary corresponds to a bias–variance trade-off governed by the admissible correction.

Importantly, the performance gains persist across sample sizes $n = 50,100,200$, suggesting that the advantage of the stability-based correction is not driven by sampling noise, but reflects a structural improvement in the estimator when instruments are imperfect. Conventional IV estimators remain centered around a biased pseudo-parameter, whereas the FP-GMM estimator moves along a disciplined correction path that reduces endogeneity while preserving sufficient identifying variation.

Table F5 - Monte Carlo results: Finite-Sample Linear Gaussian DGP with instrument invalidity ($\beta = 1, \gamma = 1$)

500 simulations, 10 fixed-point iterations, $\varepsilon_{g'} = 0.25$, $\varepsilon_b = 10^{-6}$		*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*	*BIAS*	*RMSE*	*SD*
		$\delta = 0.4$			$\delta = 0.8$			$\delta = 1.2$			$\delta = 2$		
$n = 50$	*FP-GMM*	0.04	0.13	0.12	0.26	0.29	0.13	0.49	0.53	0.18	0.86	0.88	0.18
	OLS	0.22	0.25	0.11	0.41	0.42	0.12	0.62	0.64	0.13	0.99	1.01	0.18
	TSLS	0.46	0.49	0.17	0.83	0.84	0.18	1.28	1.30	0.24	2.03	2.06	0.35
	LIML	0.46	0.49	0.17	0.83	0.84	0.18	1.28	1.30	0.24	2.03	2.06	0.35
	Fuller	0.45	0.48	0.17	0.81	0.82	0.17	1.25	1.27	0.22	1.99	2.02	0.33
$n = 100$	*FP-GMM*	0.00	0.08	0.08	0.24	0.25	0.07	0.48	0.49	0.09	0.85	0.86	0.10
	OLS	0.19	0.20	0.07	0.39	0.40	0.07	0.62	0.63	0.08	1.00	1.00	0.11
	TSLS	0.40	0.41	0.11	0.81	0.82	0.14	1.24	1.25	0.17	2.01	2.02	0.23
	LIML	0.40	0.41	0.11	0.81	0.82	0.14	1.24	1.25	0.17	2.01	2.02	0.23
	Fuller	0.39	0.40	0.10	0.79	0.80	0.13	1.21	1.22	0.17	1.97	1.98	0.22
$n = 200$	*FP-GMM*	0.00	0.06	0.06	0.25	0.26	0.06	0.46	0.46	0.06	0.85	0.85	0.07
	OLS	0.19	0.20	0.06	0.40	0.40	0.06	0.60	0.61	0.06	1.00	1.00	0.08
	TSLS	0.40	0.40	0.07	0.81	0.82	0.09	1.21	1.22	0.10	2.01	2.02	0.17
	LIML	0.40	0.40	0.07	0.81	0.82	0.09	1.21	1.22	0.10	2.01	2.02	0.17
	Fuller	0.39	0.39	0.07	0.80	0.80	0.09	1.19	1.19	0.09	1.97	1.98	0.16

This table reports Monte Carlo results for a linear Gaussian data-generating process with a potentially invalid instrument. The parameter δ governs the degree of instrument invalidity. The table reports bias, RMSE, and standard deviation for the stability-based estimator and conventional estimators including OLS, TSLS, LIML, and Fuller. Results are based on 500